федеральное государственное бюджетное образовательное учреждение
высшего профессионального образования
«Московский педагогический государственный университет»
Математический факультет
Кафедра элементарной математики и методики обучения математике

Заплетина Ольга Максимовна

«Организация заочных интеллектуальных соревнований по теории вероятностей и статистике для школьников»

Выпускная квалификационная работа
по направлению 050100.68 – «Педагогическое образование»
Магистерская программа «Математика и информационные технологии»

Научный руководитель
зав. кафедрой математики МИОО
к.ф.-м.н., директор ГБОУ ЦПМ
Ященко И. В.

Москва
2014

ОГЛАВЛЕНИЕ

Введение

Раздел, посвященный теории вероятностей, присутствует в некоторых школьных учебниках (13-15), однако зачастую, учителя не включали ее в учебный план. Во многом это связано с несистематичностью изложения вероятно-статистического материала. Поэтому в 2003 году было принято решение о включении элементов теории вероятностей и статистики в школьный курс математики общеобразовательной школы (инструктивное письмо № 03–93ин/13–03 от 23.09.2003 Министерства образования РФ «О введении элементов комбинаторики, статистики и теории вероятностей в содержание математического образования основной школы», «Математика в школе», № 9 за 2003 г.). Принятый Министерством образования в том же году документ предусматривал постепенное, поэтапное включение этих разделов в школьные курсы, давая возможность преподавательскому сообществу подготовиться к соответствующим изменениям.

С этого момента начинает выходить ряд учебных пособий, дополняющих существующие учебники алгебры (16-22), а также методические пособия в помощь учителям. С 2004 года апробацию этих учебных пособий провели в ряде московских школе для 7—9 классов, после чего в школьных программах страны заняли свое место разделы статистики и теории вероятностей. С 2012 года задания на проверку знаний в этой области вошли в Единый Государственный Экзамен и в Государственную Итоговую Аттестацию.

Сразу встала проблема согласованности основных определений и обозначений, используемых в описанных выше пособиях, главной причиной которой является столь же несогласованное изложение материалов в университетской литературе, на которую по большей части опирается построения школьной программы.

В вузовских учебниках по теории вероятностей и математической статистике встречается существенный терминологический разнобой и различия в обозначениях и записи формул. Если для взрослого исследователя эта ситуация не представляет затруднений, то для студента она может составлять сложность, по крайней мере в начале знакомства с этой дисциплиной. Тем более неприемлема она для школьника. Необходимость сопоставления различных обозначений для одних и тех же объектов может поставить в тупик неподготовленного ученика.

Попытки внедрения в школу строгой базы преподавания теории вероятностей и статистики показали, что необходимо объединение усилий авторов школьных пособий для выработки согласованных позиций по унификации основных определений и используемых обозначений. В 2004 году выходит учебник, написанный группой специалистов по теории вероятностей (Ю.Н. Тюрин, А.А. Макаров и др., Теория вероятностей и статистика), ставший основным пособием на последующие годы.

Появление учебника и разработанной системы контрольных работ осуществило внедрение в массовую школу базового уровня преподавания предмета. Что является вполне достаточным для значительной части учащихся. Однако только школьного учебника не достаточно для удовлетворения когнитивной потребности заинтересованного школьника. Безусловно, следует дорожить таким проявлением интереса. Можно предложить различные формы расширения кругозора: дополнительная литература, кружковая деятельность. Олимпиады по школьным и внешкольным предметам стали привычным явлением в наше время. Логичным и естественным представляется проведение олимпиады и по нашему предмету.

С 2008 года развернут проект по проведению заочной интернет-олимпиады по теории вероятности и статистике для школьников. В отличие от ограниченных по времени очных математических соревнований, наша олимпиада, к проведению которой автор присоседился в 2011 году, дает школьникам целый месяц на поиски информации и знаний, необходимых для решения задач.

Немаловажную роль играют задания типа эссе — это некоторые сочинения школьников на заданную тему по теории вероятностей или статистике. Именно эти задания и будут пристально рассмотрены в данной работе.

Глава I

Концепция развития математического образования и комплекс мер по ее реализации

§ 1. О необходимости преподавания теории вероятностей и статистики в школе

Попытки введения в школьный курс теории вероятностей и статистики были и раньше. Успешными они не оказались. Изложение материала начиналось с комбинаторики, которая была пугающе непроста для неподготовленного слушателя. Дальнейшее строгое и последовательное изложение начал теории вероятностей хорошо укладывалось бы в голове студента-математика, но не рядового школьника (и, зачастую, его учителя). Выделение статистики как способа представления информации не происходило вовсе.

За последние 15-20 лет информационная насыщенность жизни сильно увеличилась. Изучение статистики дает инструменты обработки и осмысления больших объемов количественной информации и перевода ее в качественную. Умения составлять и читать различного рода диаграммы, оперировать со средними, различными индексами и т.д., является полезным умением в наше время. Современный человек все чаще оказывается в ситуации, когда необходимо сделать выбор, тщательно взвесив и оценив все условия.

Школьные учебники (физики, химии) описывают окружающий мир строго детерминированным, живущим по формулам, законам и схемам, исключающим вероятностную составляющую. В то же время эти законы и схемы служат лишь инструментами в сложных стохастических явлениях с высокой долей случайностей, и без понимания этого старшеклассник и студент оказываются дезориентированными. Современная жизнь все чаще ставит индивида в ситуацию, когда необходимо сделать выбор.

Эффективность этого решения зависит, в том числе, от имеющегося опыта, вероятностно-статистической культуры, которые могут быть приобретены и воспитаны.

Страхование плотно вошло в нашу жизнь. Каждый человек, имеющий автомобиль (а таких семей сейчас большинство), решает для себя вопросы необходимости и выбора опций добровольного страхования автомобиля, ищет баланс между затраченными средствами и возможными выплатами.

Пенсионное, медицинское страхование заставляют обычного человека оценивать вероятности тех или иных неблагоприятных событий.

Вопросы сохранения и увеличения собственных денежных средств требуют анализа рисков (хотя бы на элементарном «бытовом» уровне) вложений в те или иные активы, что связано со сбором и оцениванием статистических данных.

Замечено, что до 20% современных выпускников слабо разбираются в понятиях «дробь», «доля» и «процент». При этом почти все будут пользоваться различными банковскими продуктами, в том числе кредитами. Финансовая грамотность невозможна при математической безграмотности. Курс теории вероятностей и статистики дает школьнику дополнительную возможность попрактиковаться в работе с этими не простыми для него объектами.

Почти каждый будущий студент, даже гуманитарного профиля, столкнется с изучением теории вероятностей в высшем учебном заведении. А даже у студентов-математиков возникают проблемы в понимании этого предмета и, особенно, математической статистики. Здесь большой подмогой могут стать первоначальные знания, полученные в школе.

§ 2. Концепция преподавания теории вероятностей и статистики в школе

Для того чтобы говорить о преподавании в школе теории вероятностей и статистики, посмотрим, с чего начинается изучение геометрии. Хорошо знакомые любому ребенку фигуры: круг, прямая линия, прямоугольник, треугольник и т.д., которые до этого играли в его мире лишь предметную роль — солнце, его лучи, дом, его крыша, — приобретают новые формальные свойства, теряя при том качественную конкретику. Свойства окружности не зависят от того, что она изображает: желтое солнце или синюю тарелку. Усвоение формальных правил происходит легко, потому что они естественным образом обобщают имеющийся опыт.

Объекты теории вероятностей: события, шансы, независимость, — не обладают такой же наглядностью и интуитивной понятностью, как геометрические образы и их количественные характеристики (длина, площадь).

Изучение теории вероятностей в высших учебных заведениях предполагает вначале введение определений и обобщенных формулировок. Такой подход только испугает и вряд ли окажется продуктивным для обычного школьника. В лучшем случае он будет способен применить вызубренную формулу в стандартной задаче, а в худшем — будет считать, что теория вероятностей ограничивается лишь комбинаторикой.

Задача же преподавателя состоит в том, чтобы дать ребенку «общие знания и представления, позволяющие ему ориентироваться в статистической информации», и выработать у него «верное представление о законе больших чисел и его роли в статистическом понимании изменчивых явлений» (22).

Таким образом, школьная теория вероятностей должна быть максимально приближена к жизненным сюжетам, с которыми регулярно сталкиваются школьники и их родители, решая вопросы страхования, кредитования, пенсионного обеспечения и т.д. В ее изложении там, где это возможно, стоит избегать математического формализма. Классические сюжеты некоторых задач, пришедших из азартных игр, стоит так же исключить из школьного курса, заменив знакомыми сегодняшнему школьнику ситуациями.

Акцент должен быть сделан не на вычислениях и доказательствах, а на осознании школьниками понятий об изменчивости, средних, случайных величинах, выборочных исследованиях. Задания без однозначного ответа и ситуации, требующие домысливания и обсуждения, способствуют формированию системы базовых понятий и интуиции.

И наконец, кто и как донесет до школьников вышеизложенную концепцию. Значительная часть сегодняшних школьных учителей боится предмета не меньше, чем их подопечные. По-видимому, это связано с тем, что их самих в свое время учили сухой формальной теории без применения к практическим задачам. Есть надежда, что ситуация может измениться, когда в школу придут молодые учителя, сами прошедшие школьный курс теории вероятностей.

В распоряжении учителей должны быть в том числе внеклассные методы работы, которые могут заинтересовать школьника. Обработка в Exel рядов данных, полученных в результате метеонаблюдений, индивидуальных или коллективных вероятностных экспериментов (например, с подбрасыванием монетки или кубика), дает наглядное представление о методах изучаемой дисциплины. Проведение небольшого социологического опроса с последующей визуализацией и статистической обработкой может оказаться несложным и увлекательным занятием.

На сегодняшний день требования к знаниям школьника зафиксированы в образовательном стандарте и выражаются через контрольные работы, проводимые через интернет-систему СтатГрад. Такой метод проверки знаний зарекомендовал себя как удобный и эффективный. И используется не только в московских школах, но и ряде школ из регионов.

В заключение отметим, что проблемы в реализации внедрения в школы имеются не только в России. Во многих странах (США, Великобритания, Швейцария, Южная Корея, Япония и др.) статистика и теория вероятностей давно вошли в школьные образовательные стандарты. Но несмотря на это там нередко встречается непонимание сути методов и интерпретаций вероятностных фактов. Как правило имеется перекос в сторону техники применения вероятностных методов без осмысления сути задач (22).

Глава II

Заочные интернет-олимпиады по теории вероятностей и статистике

§ 1. Концепция олимпиады

В содержание курса школьной математики в соответствии с образовательным стандартом с 2004 года входят элементы теории вероятностей и статистики. С 2012 года задачи по вероятности и представлению данных входят в единый государственный экзамен по математике.

Вместе с тем известно, что одним из важнейших аспектов образования является популяризация знаний. Если по алгебре, геометрии и другим традиционным разделам школьной математики нет недостатка в дополнительной, научно-популярной и популярной литературе для взрослых и детей разных возрастов, то анализ состояния дел в области теории вероятностей показывает явный недостаток как популярной литературы, так и иных форм популяризационной работы [1, 4].

Незначительное количество популярных изданий для школьников было выпущено в Советском Союзе преимущественно в 50-е – 80-е годы прошлого века (например, [8, 9]). Число новых материалов можно считать исчезающее малым, даже с учетом переводных зарубежных изданий на тему вероятности и статистики для школьников (например, [12]). В то же время количество зарубежной популярной литературы по статистике и вероятности увеличивается. Частично это связано с ростом значимости вероятностно-статистической линии в школе, частично – с ростом роли стохастических методов в разных отраслях мирового хозяйства [1].

Популяризационная деятельность должна осуществляться в различных формах. Помимо литературы большую ценность имеют кружки, олимпиады и другие интеллектуальные соревнования для детей разных возрастов и уровней. В последние годы мы наблюдаем стихийный рост количества и качества интернет-сайтов, посвященных популярной математике, в частности, в области теории вероятностей и статистики.

На сайте http://www.mccme.ru/ с 2008 года проходит ставшая уже почти традиционной ежегодная заочная интернет-олимпиада по теории вероятностей для школьников. Статьи, посвященные этой олимпиаде, публиковались в журнале «Математика» [3, 5].

Впервые олимпиада была проведена по инициативе Юрия Николаевича Тюрина, Алексея Алексеевича Макарова и Ивана Валериевича Ященко. В отборе и составлении задач в разные годы участвовали многие математики и педагоги: Ю.Н. Тюрин, Е.А. Бунимович, В.А. Булычев, П.В. Семенов, А.А. Макаров, И.В. Ященко и др. Последние несколько лет олимпиада проводилась в основном силами сотрудников научно-методической лаборатории теории вероятностей МЦНМО.

Правила самой олимпиады крайне просты; невысоки и требования к участникам. Олимпиада открыта для всех желающих и проходит в течение календарного месяца, обычно с 20 марта по 20 апреля. Каждый желающий должен заявить письменно (адрес teorver.olympiads@gmail.com) или через форму регистрации на сайте http://olimpiada.ru/ о своем желании участвовать. Это можно сделать заранее или перед окончанием олимпиады, прислав заявку вместе с решениями.

Задачи олимпиады рассчитаны на школьников 6–11 классов. Устоялась традиция, когда оргкомитет не ограничивает «возраст задачи» сверху. Любую задачу мы лишь рекомендуем учащимся, начиная с некоторого класса. Например, задача, которая может быть решена с использованием только интуитивных представлений или конечного перебора, классического определения вероятностей и т.п., рекомендуется учащимся, начиная с 6 класса. Если при решении имеется необходимость использовать несложные преобразования в рамках алгебры событий, задача рекомендуется учащимся 7 класса и старше. Если в задаче участвуют характеристики случайных величин, то такую задачу мы предлагаем школьникам с 8 класса и т.п. При этом, если одиннадцатиклассник успешно решил задачу для 6 или 7 класса, он получает за нее полный балл, так же как и шестиклассник.

Возрастная дифференциация возникает на этапе проверки и награждения победителей. Отдельно выделяются победители (3 первых места) среди учащихся 6–7 классов, отдельно среди 8–9 классов и отдельно среди старшеклассников 10–11 классов.

За последние несколько лет мы с удовольствием наблюдаем значительный прогресс многих наших постоянных участников. Особенной активностью отличаются школьники 7–9 классов. Удивительно, но за годы проведения олимпиады в ней участвовало несколько десятков учащихся старших классов, но никто из них не показал интересных результатов.

Вероятно, это связано с тем, что олимпиада не имеет официального статуса и не предоставляет победителям и призерам льгот при поступлении в вузы.

Задания олимпиады выкладываются в свободном доступе на сайте (для того, чтобы войти на сайт и увидеть задания, регистрация не нужна) и объявляется дата, до которой оргкомитет принимает решения от участников. Согласно правилам олимпиады участники могут пользоваться любой помощью, справочной литературой и т.п. Неспортивным и некрасивым считается лишь откровенное «использование взрослого труда» и выдавание его за свою работу. До сих пор при проверке работ у нас ни разу не возникало сомнений в самостоятельности работы школьника.

Свои решения участник может присылать в любом формате от текстового файла до фотографий тетрадных страниц. Можно пользоваться обычной почтой. Единственное требование – работа должна быть читаемой. Разумеется, такая свобода выбора предоставляется участникам потому, что число их сравнительно невелико. Каждый год в олимпиаде регистрируется около 100 участников, а работ оргкомитет получает от 20 до 40. Оргкомитет работает над популяризацией олимпиады, стараясь привлечь больше участников, но при этом мы понимаем, что бóльшее число участников потребует значительно более строгий регламент и привлечения значительных сил к проверке работ.

Материалы прошлых лет публикуются на сайте в разделе «Архив», Кроме того, задачи олимпиад 2008–2011 годов и их решения опубликованы отдельной книгой [10].

§ 2. Задания-эссе, классификация

Отличительная особенность данной олимпиады состоит в том, что помимо собственно олимпиадных задач, подразумевающих решение, с 2010 года олимпиада включает 3 задания-эссе. От участников требуется проанализировать предложенную ситуацию и написать короткое «сочинение на заданную статистическую тему».

Участник олимпиады погружается в неопределенную ситуацию, где от него требуется фантазия, оценочная деятельность, учитывающая реальные ограничения и природу данных. Действия в неопределенной ситуации играют крайне важную роль в формировании общей статистической и математической культуры учащихся, поскольку вместо выполнения шагов известного или изученного алгоритма участник олимпиады оказывается в роли исследователя, самостоятельно планирующего эксперимент. Он самостоятельно определяет существенные и малосущественные особенности случайного эксперимента, самостоятельно интерпретирует полученные результаты, изобретает метод описания полученных данных и формирования гипотезы.

Сразу скажем, что ни в одном из предложенных эссе мы не предлагаем школьникам проверять сформулированные гипотезы, поскольку для этого математический аппарат, доступный школьнику, явно недостаточен.

В некоторых ситуациях выдвижение гипотезы само по себе является крайне непростым действием, требующим от учащегося высокой культуры регулятивной деятельности. Кроме того, возникают ситуации, когда количество возможных правдоподобных гипотез, связанных с описанной в задании ситуации, велико. В этих случаях авторы составляют задания таким образом, чтобы либо явно сформулировать гипотезу, либо косвенным образом направить действия учащегося (см. примеры заданий-эссе ниже).

Важнейшей частью выполнения заданий эссе является поиск информации и данных, необходимых либо для поиска закономерности, либо для проверки данных на соответствие определенному предположению (гипотезе). Некоторые задания требуют от учащихся самостоятельного сбора данных с помощью опросов и коротких социологических исследований Таковы, например, задания о проблеме профессора Хаги, о выборе числа из числового отрезка. Другие задания предполагают самостоятельный поиск данных в интернете. Таковы задания о связи влажности воздуха и интенсивности снегопада, о законе Бенфорда (тексты заданий и комментарии к ним см. ниже).

Некоторые задания-эссе предназначены для выработки критического научного мышления учащихся. Учащийся погружается в ситуацию, когда известно, что в приведенных рассуждениях имеется какого-либо рода неточность, ошибка или необоснованный вывод. Учащимся предлагается понять, в чем недостатки проведенного исследования и сделать самостоятельные исследовательские шаги. Надо сказать, что задания такого типа оказались наиболее привлекательными для участников олимпиады.

Эта часть олимпиады традиционно вызывает повышенный интерес среди участников и пользуется популярностью. Среди присланных эссе есть очень яркие, оригинальные работы, которые мы размещаем на сайте в разделе «Решения». Материалы четырёх первых олимпиад в издательстве МЦНМО изданы отдельной книгой, в которую вошли наиболее удачные работы школьников.

Все эссе можно условно разделить на несколько типов.

1. Проверка какого-либо утверждения либо с помощью самостоятельно собранных данных, либо с помощью необработанных данных, предоставленных «в комплекте» с заданием. Как правило, в этом случае участнику олимпиады частично предлагается алгоритм действий (см. ниже эссе 1 и 2).

2. Самостоятельное исследование, связанное с опросом одноклассников, друзей, родителей и последующая обработка данных с выдвижением правдоподобной гипотезы или опровержением неправдопдобной (см. эссе 3 и 5). Здесь важно умение систематизировать полученные данные, представлять их наиболее подходящим способом и выдвигать гипотезы.

3. Поиск статистического метода для решения задачи (см. эссе 4). Выработка конструктивного мышления и навыков. Ничем не ограниченное пространство поиска оказывается очень привлекательным для школьников. Эссе про оценку численности толпы оказалось самым часто выбираемым среди участников всех возрастных групп. Более того, никто из учащихся не предложил метод, казавшийся авторам задачи наиболее естественным, зато было предложено множество оригинальных методов. Одна участница (7 класс), чувствуя, что придуманный ею способ, как и всякий другой, нуждается в оценке точности и, не имея возможности провести такую оценку, нашла замечательный выход: проведя оценку численности людей на фотографии тремя различными методами, она убедилась в том, что все три оценки не очень далеки друг от друга и тем самым, вероятно, дают правдоподобные результаты. После этого участница усреднила полученные результаты, интуитивно чувствуя, что усредненная оценка, скорее всего, лучше, чем каждая отдельная. Здесь статистическая интуиция проявляется в бóльшей степени, чем в умении применять готовые рецепты из учебника.

Задания первых трех типов, как правило, нацелены на выработку регулятивных умений учащихся по самостоятельному планированию и проведению экспериментов, а также выработке методов интерпретации полученных данных.

4. Поиск ошибки в сложном и распространенном рассуждении, представленном в описании задания (см., например, эссе 6). При выполнении заданий этого типа вырабатывается критически-деструктивное мышление, являющейся составной частью мыслительной культуры. Говоря конкретно о выполнении эссе про связь влажности воздуха и количество выпавшего снега, отметим, что школьникам удалось заметить многие недостатки в рассуждениях автора утверждений, но главную проблему – непригодность линейной регрессии для описания модели – не отметил никто из участников. Это лишний раз говорит о том, базовые идеи статистики находятся на границе интуитивного и осознанного.

Заметим, что эссе оценивались отдельно от прочих задач, независимо от возраста участника.

§ 2.1. Проверка какого-либо утверждения

1. (2010) Закон Ньюкомба–Бенфорда.

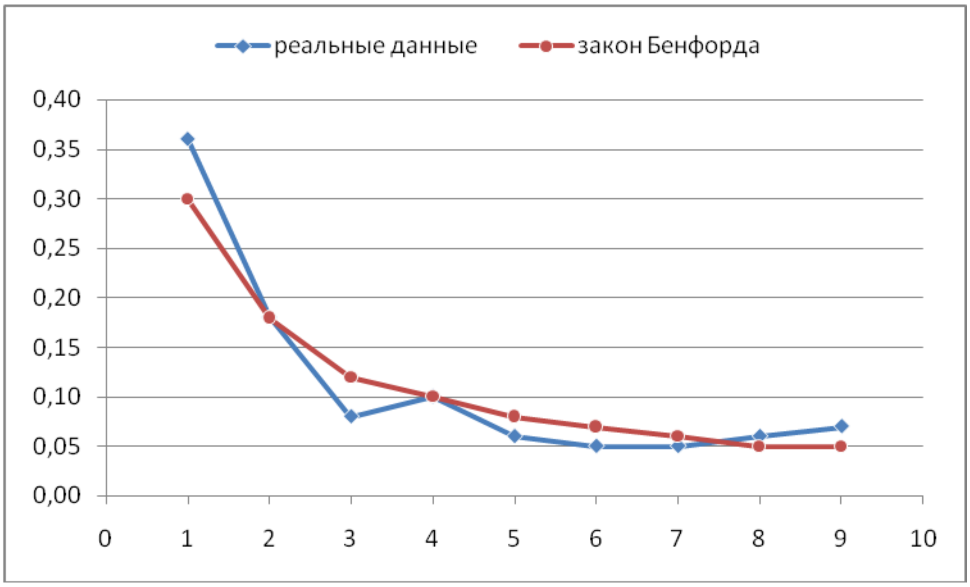
В 1881 году Саймон Ньюкомб заметил интересную закономерность, получившую впоследствии название “закон Бенфорда” (в честь переоткрывателя Фрэнка Бенфорда). Закон гласит, что в любом большом массиве однородных числовых данных (площади стран, высоты гор, длины рек, курсы валют, значения физических величин и т.п.) цифра 1 встречается на первом месте примерно в шесть раз чаще, чем цифра 9. Или, в общем случае, вероятность встретить на первом месте цифру k равна $\lg\frac{k+1}{k}$, где $k=1,\ldots,9$. Обсуждение закона можно найти в статье академика В.И.Арнольда «Антинаучная революция и математика» или «Статистика первых цифр степеней двойки и передел мира» (В.И.Арнольд. «Жесткие» и «мягкие» математические модели. М., МЦНМО, 2008). Напишите небольшое эссе на тему закона Бенфорда, в котором

а) проверьте, выполняется ли этот закон для населения субъектов Российской Федерации (см. файл data1.xls);

б) опишите любой другой набор однотипных величин, у которых на первом месте может стоять любая цифра, и проверьте согласованность этих данных с законом Бенфорда.

Вариант эссе а). Построим таблицу по данным из файла data1.xls. В первом столбце укажем цифры от 1 до 9. Во втором — частоту, с которой каждая из цифр стояла на первом месте в наших данных о населении (с точностью до двух цифр после запятой). В последнем столбце поместим теоретические значения вероятностей встретить каждую из цифр на первом месте, вычисленные по закону Бенфорда. Например, для цифры 3 в последнем столбце будет стоять $\lg\left(\frac{3+1}{3}\right)=\lg\frac{4}{3}\approx 0{,}125$.

k	1	2	3	4	5	6	7	8	9
	0,357	0,179	0,083	0,095	0,06	0,048	0,048	0,06	0,071
$\lg\frac{k+1}{k}$	0,301	0,176	0,125	0,097	0,079	0,067	0,058	0,051	0,046



Построим график, где по оси OX цифры от 1 до 9, а по оси OY процент их появления на первом месте в теории и на практике. Судя по полученному графику, можно сделать вывод, что данные довольно хорошо согласуются с законом Бенфорда. Кроме цифры 3: в реальных данных она встречается намного реже, чем в теории. И цифры 1: она, наоборот, встречается чаще в реальных данных. Однако, возможно, это связано с тем, что мы рассмотрели не очень большой объем данных: всего 84 числа. В целом же закон Бенфорда, проверенный на основе данных о населении субъектов Российской Федерации, вполне правдоподобен.

2. (2010) Прогноз погоды.

Считается, что лучший прогноз погоды на завтра «Завтра погода будет такая же, как сегодня». Метеорологи часто говорят, что точность этого прогноза около 80%, то есть вероятность того, что погода завтра будет похожа на сегодняшнюю равна примерно 0,8. Зайдите на сайт www.rp5.ru или используйте любой другой источник с архивом погоды за несколько месяцев. Выберите свой город и проверьте, насколько правы метеорологи, которые так говорят. Выполняя это задание, составьте таблицы погоды, учитывая температуру воздуха, облачность и направление ветра за месяц.

3. (2014) Жребий с одной монетой.

Вот задача, с которой однажды столкнулись пятеро: нужно было бросить справедливый жребий, чтобы выбрать одного из них. Из подручных средств была только монетка. Возникло два предложения, как можно организовать жребий.

Первое предложение. *Все пятеро бросают монетку по очереди. Те, у кого выпала решка, выбывают из розыгрыша, а те, у кого выпал орёл, – остаются. Оставшиеся еще раз бросают по монетке – делают новый круг. И так далее. Если на каком-то круге все выбросят решки, то круг не считается и разыгрывается еще раз. Бросание продолжается до тех пор, пока на каком-то из кругов не останется только один участник, у которого выпал орёл. Он и считается победителем жребия.*

Второе предложение. *Каждому участнику присвоим номер от* 0 *до* 4. *Затем бросаем монету* 3 *раза. Если выпал орёл, то записываем единицу. Если решка, – записываем ноль. Получается двоичная последовательность длины три. Интерпретируем ее как число от* 0 *до* 7. *Например, если бросания дали «орёл-решка-орёл», получается*

$$101_2 = 1 \cdot 2^2 + 0 \cdot 2 + 1 = 5.$$

Если полученное число от 0 *до* 4*, то выигрывает тот, чей номер получился. Если число от* 5 *до* 7*, то бросания не засчитываются и всё повторяется еще раз.*

Приведут ли рано или поздно оба способа к цели, то есть не может ли получиться бесконечная серия бросаний? Какой способ в среднем быстрее выявит победителя? Можно ли как-нибудь подсчитать или хотя бы оценить математическое ожидание числа бросков (то есть времени), которое придётся сделать, если пользоваться первым способом? А если пользоваться вторым способом?

Попробуйте придумать какой-нибудь свой метод. Может быть, вам удастся сделать его более экономичным, чем наши.

В общем, мы предлагаем вам провести вероятностный анализ задачи, причем настолько полный и подробный, насколько у вас получится.

Возможная идея исследования. Оба предложенных варианта результативны, в том смысле, что бесконечная последовательность бросков получиться не может. Докажем это для первого способа. Бесконечная последовательность бросков может случиться, только если, начиная с какого-то момента, все оставшиеся участники выбрасывают то же самое, что и первый из них. То есть, если первый из оставшихся выбрасывает орла, то и все остальные – тоже. Если он выбрасывает решку, то и все остальные тоже выбрасывают решку. И это продолжается до бесконечности. Предположим, оставшихся участников k. Вероятность того, что все они, выбросят ту же сторону монеты, что и первый из них, равна $0,5^{k-1}$. Вероятность того, что они сделают это бесконечное число раз, равна бесконечному произведению

$$0,5^{k-1}\cdot 0,5^{k-1}\cdot 0,5^{k-1}\cdot\ldots\cdot 0,5^{k-1}\cdot\ldots.$$

Это произведение не может быть отрицательным. С другой стороны, оно меньше любого положительного числа. Значит, оно равно нулю. Таким образом, вероятность бесконечной последовательности бросков нулевая. Примерно такое же рассуждение показывает, что и второй способ рано или поздно приведет к успеху.

Оценить математическое ожидание времени, затраченного на жребий можно разными способами. Грубую оценку можно сделать, исходя из того что на каждом «кругу жребия» выбывает в среднем половина участников.

Второй способ – более экономичный. Жребий даст результат, если при трехкратном бросании монеты выпало число от 0 до 4 (из возможного набора значений от 0 до 7); вероятность этого равна $\frac{5}{8}$. Легко догадаться (а можно доказать), что по этой причине математическое ожидание числа серий по три броска равно $\frac{8}{5}=1,6$. Значит, математическое ожидание числа бросков одной монеты равно $3\cdot 1,6=4,8$. Это число меньше, чем 5 – наименьшее число бросков, которое придется сделать пяти участникам, если они используют первый способ.

§ 2.2. Самостоятельное исследование

1. (2011) Проблема профессора Кадзуо Хаги.

Профессор Кадзуо Хага из университета Цукубы изобретатель и энтузиаст оригамики — геометрического оригами. Однажды он поставил интересный вопрос.

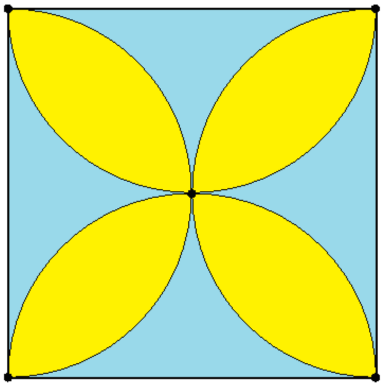


Рис. 1

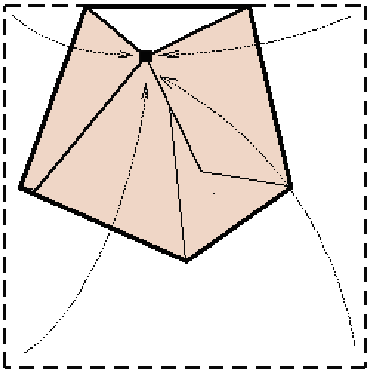


Рис. 2

На рис. 1 бумажный квадрат разделен на желтые и голубые области четырьмя полуокружностями. Получается изящный орнамент, напоминающий цветок.

Если выбрать точку в какой-нибудь голубой области и затем последовательно сделать сгибы так, что каждая вершина квадрата наложится на эту точку, то в результате получится пятиугольник (рис. 2). Поэтому объединение голубых фигур профессор Хага назвал областью пятиугольников или просто 5-областью. Желтым цветом показана 6-область. Четыре вершины и центр квадрата — точки, приводящие к четырехугольникам.

Профессор Хага неоднократно рассказывал школьникам и учителям математики о распределении точек по числу вершин получающегося многоугольника. Он пишет следующее:

«*Я заметил, что, если попросить выбрать случайную точку, то намного больше людей отмечает точку, приводящую к шестиугольнику, чем к пятиугольнику. Очень редко встречаются те, кто выбирает точки, дающие четырехугольник. Возникает вопрос: пропорционально ли число людей, о которых идет речь, площадям областей?*

Чтобы ответить на этот вопрос, найдем площади областей. Пусть длина стороны квадрата равна 1. *Тогда радиус каждого полукруга равен* $0{,}5$*, а его площадь равна* $\frac{\pi}{8}$. *Значит, общая площадь четырех полукругов равна* $\frac{\pi}{2}$. *Таким образом, площадь лепестков, то есть площадь области шестиугольников равна* $\frac{\pi}{2}-1\approx 0{,}57$. *А площадь области пятиугольников равна* $1-\left(\frac{\pi}{2}-1\right)=2-\frac{\pi}{2}\approx 0{,}43$.

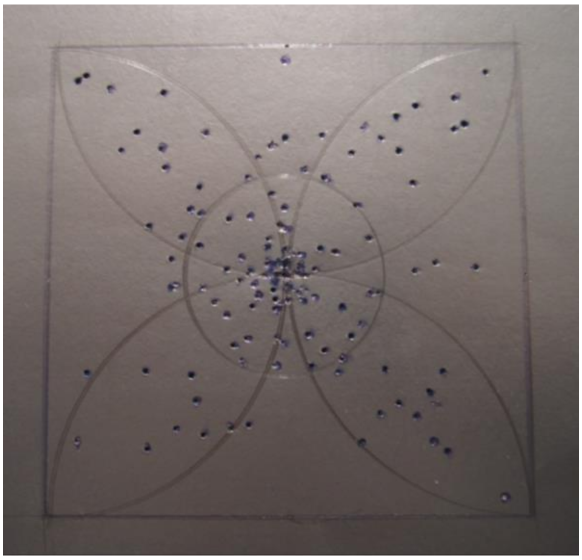
Оказывается, разница площадей не очень велика. Отношение числа людей, выбирающих точки, лежащие в соответствующих областях несоразмерно отношению площадей. Было бы интересно узнать причину, по которой точки вне лепестков обладают меньшей привлекательностью, чем точки внутри них».

Проведите эксперимент. Для этого Вам потребуется несколько десятков бумажных квадратов. Можно использовать белые чистые салфетки, разрезанные на квадраты без сгибов. Попросите как можно больше людей (своих друзей, а также взрослых) отметить произвольную точку на чистом бумажном квадрате. Если собрать все отмеченные точки на одном рисунке, то получится распределение точек. Может быть, свойства этого распределения помогут объяснить, почему любители шестиугольников встречаются гораздо чаще, чем мог бы ожидать профессор Хага, сравнивая площади областей?

Эссе одного из участников олимпиады (редактировано).

Для ответа на вопрос профессора Кадзуо Хаги: «Почему точки вне лепестков на его диаграмме обладают меньшей привлекательностью, чем точки внутри них?», нужно было проверить, так ли это.

Мы взяли много бумажных квадратных листочков размером 9×9 см^2 (на таких листочках делают записи в офисах) и провели опрос. Мы предлагали поставить точку в произвольном месте на квадратном листочке. Чтобы отличить ответы респондентов (мужчина, женщина, школьник) листочки были разные. Например, школьникам давались листочки с рисунком на обратной стороне. Опрос проводился в офисе коммерческой организации, в организации госструктуры, в общеобразовательной и музыкальной школах. Для проведения опроса привлекались помощники — родители. Всего в опросе участвовало 147 респондентов.



Обработка результатов. Получив 147 листочков с точками в разных местах, нужно было собрать их на одном рисунке. Для этого на плотном листе бумаги был нарисован квадрат-трафарет. Далее последовательно брался квадрат с точкой, накладывался на трафарет, и точка прокалывалась иголкой. Так удалось получить на одном рисунке все точки в виде дырочек — распределение множества точек (см. фотографию). Нарисовав 4 полуокружности радиусом 4,5 см, можно было видеть, что точки распределились в основном так, как и говорил профессор — внутри лепестков.

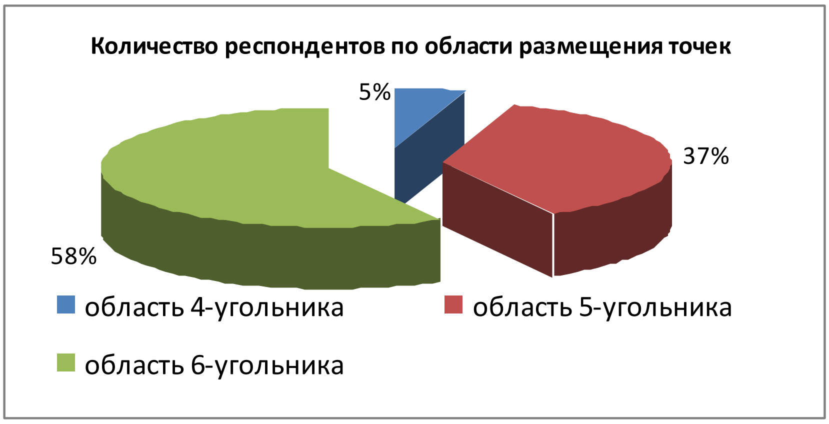
В результате было получено:

«6-угольников» — 84 чел. или 58%

«5-угольников» — 55 чел. или 37%

«4-угольников» — 8 чел. или 5%

Полученные данные показаны на диаграмме.



Анализ результатов. Большинство респондентов (58%) поставили точку в области шестиугольников. Было замечено, что некоторые люди рассматривают листок как чистый лист бумаги, писать на котором нужно слева вверху с красной строки — это также область шестиугольников.

Соразмерность отношения площадей и отношения количества людей. При проведении опроса было замечено, что многие люди стараются поставить точку в середину. Так как попасть в самый центр довольно трудно, точки попадают в область «6-угольников», которая расположена близко к центру. Поэтому я решила посмотреть на площади внутри «центральной» области.

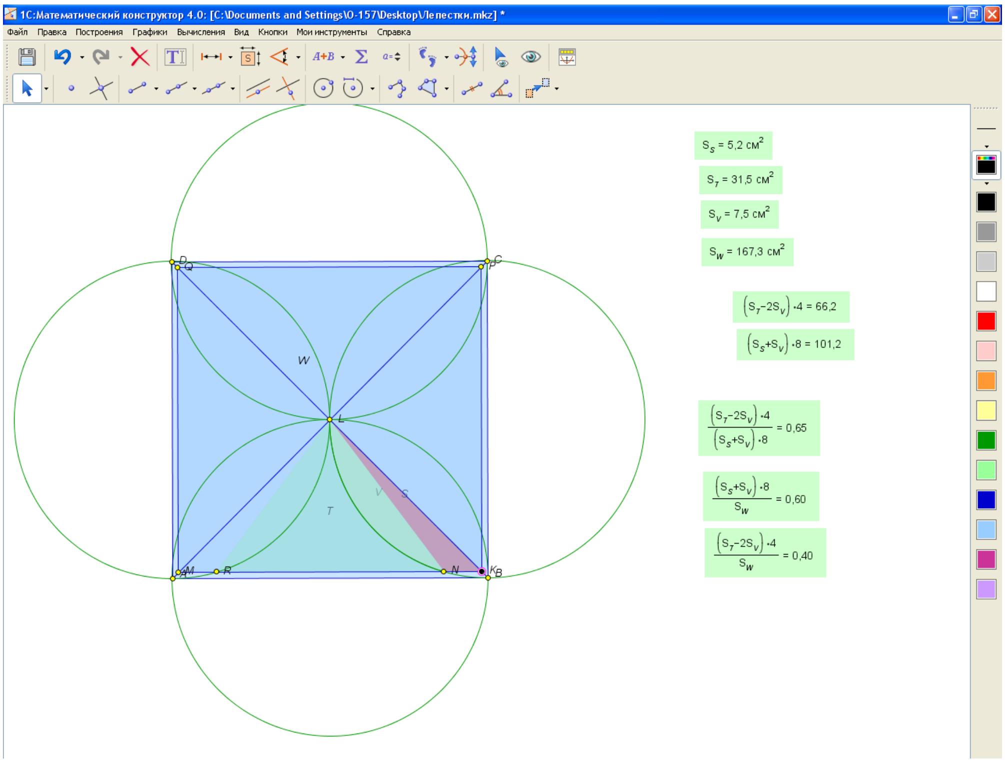
На фото видно, что довольно густо заполнена область вблизи центра (менее половины радиуса, с помощью которого строились лепестки). Если принять за единицу сторону квадрата, то радиус круга, в который попало большинство точек, равен 0,2. Площадь круга с радиусом приблизительно равна 0,13, что составляет 13% от общей площади квадрата.

Число точек, попавших в круг, радиусом 0,2 равно 87 или около 60% от общего числа точек. Откуда следует вывод, что на площади 13% разместилось 60% точек. Площадь «области пятиугольников» внутри маленького круга составила 0,031 (я ее считала приблизительно, как число несколько меньшее площади четырех треугольников площадью 0,009 каждый; основание — 0,09, высота — 0,2) то есть 24% площади круга.

Число точек, попавших в область пятиугольников маленького круга, равно 7 (8%).

Также я решила, воспользовавшись программой «Математический конструктор», посмотреть на соотношения площадей в квадрате, который будет несколько меньше данного, т.е. откинув из рассмотрения несколько редких точек, которые ставят близко к краям.

Расчёты в «Математическом конструкторе» показали, что при небольшом отступе от краев основного квадрата, площадь «области пятиугольников» занимает 40%, а площадь «области шестиугольников» — 60%. А это уже сопоставимо с отношениями количества людей (если еще выбросить из рассмотрения людей, выбирающих «область четырехугольника»).



<u>Заключение.</u> Я считаю, что если учесть «психологию» людей, которые неуютно чувствуют себя на краю, и соотносить площади, откинув края квадрата, то отношения количества людей и отношения площадей могут оказаться вполне соразмерными. А точки внутри лепестков обладают большей привлекательностью в связи с тем, что они занимают

больше центральной части и в них проходят оси симметрии, приятные глазу.

2. (2011) Страховая компания.

Страховая компания АВС страхует автомобили. Страховая стоимость автомобиля зависит от его возраста. Агенты компании рассчитывают уценку очень просто — машины старше двух лет каждый год теряют в цене 10%.

По данным сайта www.auto.ru или других источников проведите исследование на тему — соответствует ли оценочная политика страховой компании АВС практике, сложившейся на рынке подержанных автомобилей. Свое исследование напишите в виде небольшого эссе (2-3 страницы). При выполнении работы учтите, что среди автомобилей бывают тюнингованные (особым образом оборудованные и отделанные) и разбитые. Цены на те и другие значительно отличаются от средней цены.

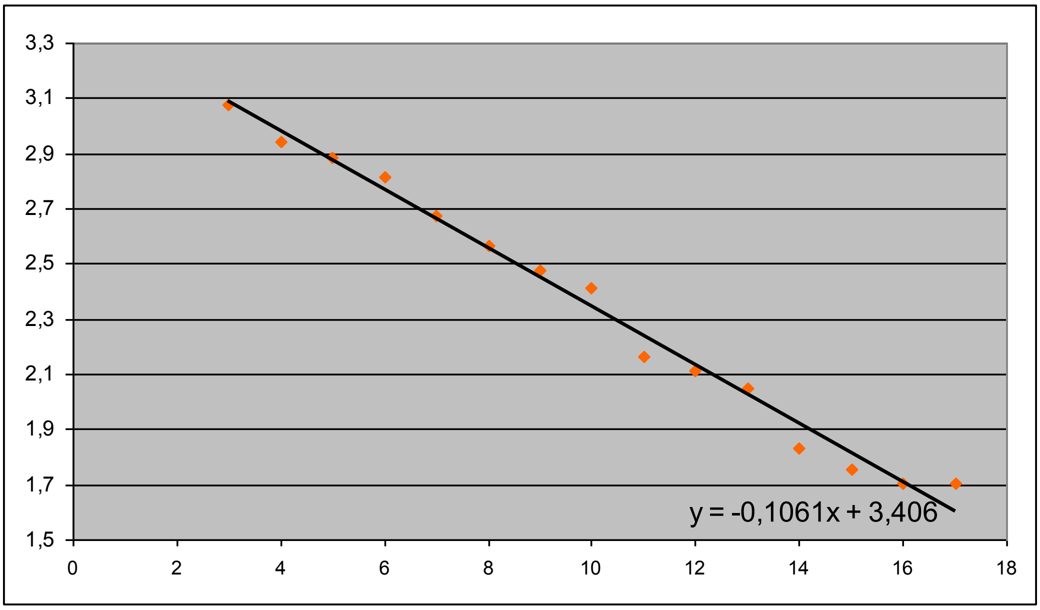
Возможная идея исследования. Исследуем динамику цен на распространенные в Москве подержанные автомобили «Фольксваген-Пассат». Составим таблицу «Год выпуска» — «Цена» для Пассатов, выставленных на продажу через Интернет-сайт Auto.ru. Будем учитывать только машины одной модели 1995 – 2009 годов выпуска (более старых машин в продаже почти нет). 11 февраля 2011 года на продажу было выставлено 1565 машин.

Среди автомобилей встречаются битые (попавшие в аварию), не прошедшие таможню, машины с правым рулем, а также эксклюзивные машины с дорогостоящими нестандартными устройствами, полноприводные, с художественной окраской, дорогой отделкой салона и т.п. Все эти обстоятельства приводят к резкому отличию их цен от цен обычных авто. Чтобы убрать влияния этих выбросов, вместо среднего арифметического вычислим медиану цен Пассатов за каждый год (вместо медианы можно взять урезанное среднее, отбросив перед усреднением, скажем, 25 % крайних значений).

Получим таблицу из 15 строк. Предполагая, что цена автомобиля убывает в геометрической прогрессии (уменьшается на одно и то же число процентов), получаем, что логарифм цены должен убывать линейно год от года. Добавим в таблицу логарифм цены (переменную y).

Возраст машины x_i (лет)	Медиана цен m_i (тыс. $)	$y_i = \ln m_i$
17	5,492	1,703292
16	5,5	1,704748
15	5,8	1,757858
14	6,25	1,832581
13	7,784	2,05207
12	8,25	2,110213
11	8,7	2,163323
10	11,15	2,411439
9	11,9	2,476538
8	13	2,564949
7	14,5	2,674149
6	16,7	2,815409
5	17,9	2,884801
4	19	2,944439
3	21,7	3,077312

Для пар $(x_i; y_i)$, где $i = 1...15$ построим диаграмму рассеивания. На построенной диаграмме видно, что точки близки некоторой прямой. Это означает, что предположение о примерно одинаковом процентном удешевлении подержанных машин оправдывается (по крайней мере, для Пассатов).



Однако никакую количественную оценку удешевления мы еще не сделали. Чтобы сделать ее, воспользуемся методом наименьших квадратов. Построим линейную регрессию $y = ax + b$. (Вычисления производились с помощью Excel). Получается линейная зависимость.

$$y = -0{,}1061x + 3{,}406.$$

Эта прямая показана на диаграмме рассеивания. Осталось узнать, на сколько же дешевеют машины ежегодно? Натуральный логарифм цены (в тыс. $) ежегодно понижается на 0,1061, следовательно, цена умножается на $e^{-0{,}1061} \approx 0{,}8993$, то есть составляет $89{,}9\,\%$ от цены предыдущего года. Таким образом, по нашим расчетам средняя машина ежегодно дешевеет на $10{,}1\,\%$.

Наши выводы для Пассатов весьма хорошо согласуются с ценовой политикой упомянутой страховой компании, которая упоминалась в начале.

3. (2012) Выбор числа.

Психологи и другие специалисты утверждают, что человек не может сделать чисто случайный выбор из предложенных альтернатив. Например, известно, что если человеку предложить выбрать совершенно случайную цифру, то он, скорее всего, выберет цифру 7. Причины этого до конца неизвестны, хотя есть различные гипотезы.

Мы предлагаем Вам экспериментально проверить, правдоподобна ли гипотеза: «*значительное число людей чаще всего подсознательно выбирают числа, далекие от краев и от середины предложенного числового ряда, то есть делящее ряд чисел близко к отношению 1:3 или 3:1*».

Мы приготовили файл esse1.xls. Его можно скачать здесь или с помощью отдельной ссылки на странице «Задачи».

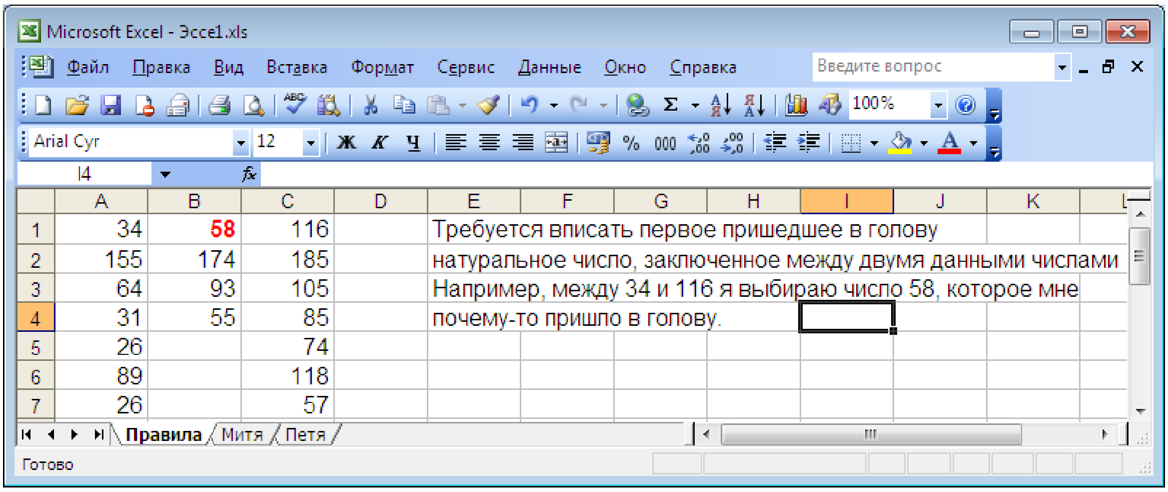
A	B	C
34		116
155		185
64		105
31		85
26		74
89		118
26		57
100		129
188		229
148		236
178		268
97		161
115		165
111		183
156		208

A	B	C
113		187
120		162
91		181
15		84
30		81
57		86
120		215
47		124
17		78
105		129
24		101
11		85
120		152
159		235
31		104

75		140
52		131
153		245
15		69
84		180
18		78
117		198
128		193
126		208
22		46

142		183
37		107
189		270
110		210
180		260
28		84
88		150
61		87
164		228
75		137

В каждой строке таблицы — начало и конец натурального ряда: в ячейке A — начальное число, в ячейке C — конечное число. Респонденту (испытуемому) предлагается быстро вписать в ячейку B любое число, заключенное между этими числами. Это нужно сделать в каждой строке. Всего одному респонденту предлагается 50 строк.



На рисунке показано начало эксперимента. Можно все опыты с разными людьми фиксировать в одном и том же файле на разных страницах.

Затем нужно определить, в каком отношении каждое выбранное число разбивает отрезок между первым и последним числом. В показанном примере число 58 отсекает от отрезка от 34 до 116 примерно 29%, поскольку

$$\frac{58-34}{116-34} \approx 0,29.$$

Все вычисления удобно проводить здесь же — в Excel. Когда все отношения получены, их следует проанализировать для каждого человека и для всех вместе. Нужно сгруппировать полученные доли (проценты) в интервалы длиной, например, 5% (длину интервала нужно подобрать так, чтобы результаты были наглядны). Затем нужно найти, какая доля результатов попадает в каждый интервал. Результаты удобно представить с помощью диаграмм.

Гипотезу придется отвергнуть, как неправдоподобную, если вблизи точек 25% и 75% не наблюдается заметного преобладания частот. Если же оно наблюдается, то следует признать, что гипотеза правдоподобна. Чем больше выбранных чисел группируется около точек 25% или 75%, тем более правдоподобна гипотеза (к сожалению, этого недостаточно, чтобы наверняка утверждать, что она верна).

Может быть, вам удастся найти другие правдоподобные гипотезы или закономерности, которые незаметны, если число респондентов мало.

Чем респондентов больше, тем лучше. Если ограничиться 50 числами для каждого респондента, то для проявления закономерностей нужно опросить хотя бы $20-30$ человек — одноклассников, друзей, родственников. Не обязательно, чтобы респондент сам вводил числа в компьютер. Если ему это непривычно, он может это делать на распечатке или просто продиктовать вам. Нужно только помнить, что все опросы должны проводиться независимо друг от друга.

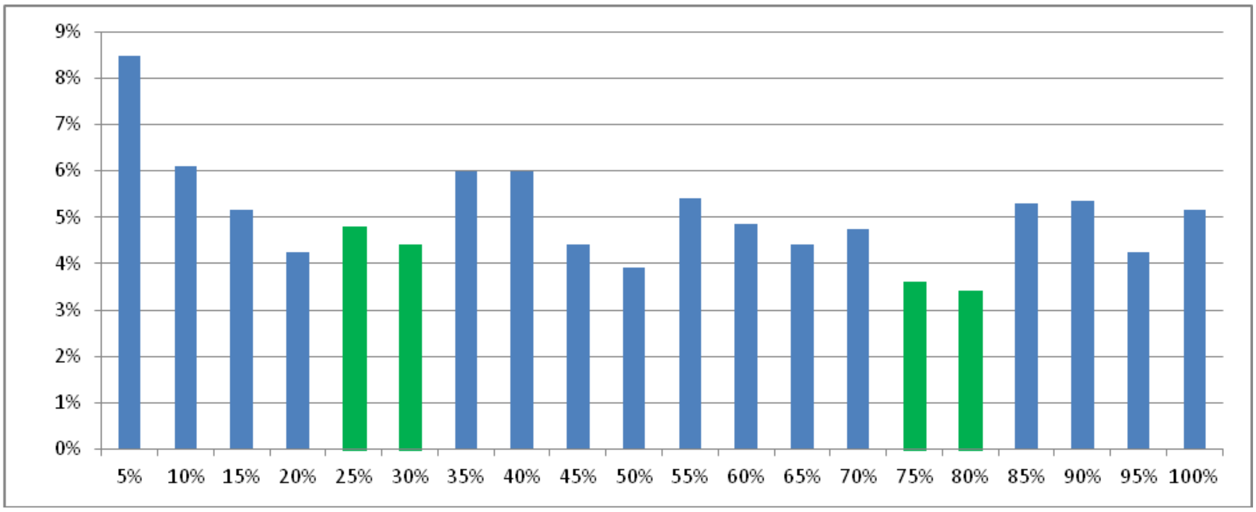
Эссе одного из участников олимпиады (редактировано).

Всего опрошено 33 респондента, из них 19 мужчин и 14 женщин. Было получено 1650 результатов, по 50 от каждого респондента. 44 результата пришлось исключить из обработки, потому что они оказались ошибочными (внесенные числа не попадали в указанные пределы). В итоге обрабатывалось 1606 результатов.

Опрашиваемые – люди от 27 до 58 лет. К сожалению, мне не удалось собрать данные с одноклассников или ровесников (начались каникулы), но интересно было бы это сделать и сравнить полученные результаты.

Далее, сгруппировав полученные доли в интервалы по 5% и определив, какая доля результатов попадает в каждый интервал, я получил диаграмму1.

Диаграмма 1



На диаграмме не наблюдается преобладания частот попаданий результатов вблизи точек 25% и точек 75 %, даже можно сказать наоборот – идет понижение. Считаю гипотезу необоснованной. Возможно, для подтверждения гипотезы требуется гораздо большее число опрашиваемых респондентов или форма опроса должна быть другой.

Когда я собирал результаты, то отмечал, мужчина или женщина заполняли таблицу. Отдельно по мужчинам результаты в диаграмме №2, а по женщинам в диаграмме 3. За 100% процентов принято общее количество результатов, т.е. 1606

Диаграмма 2

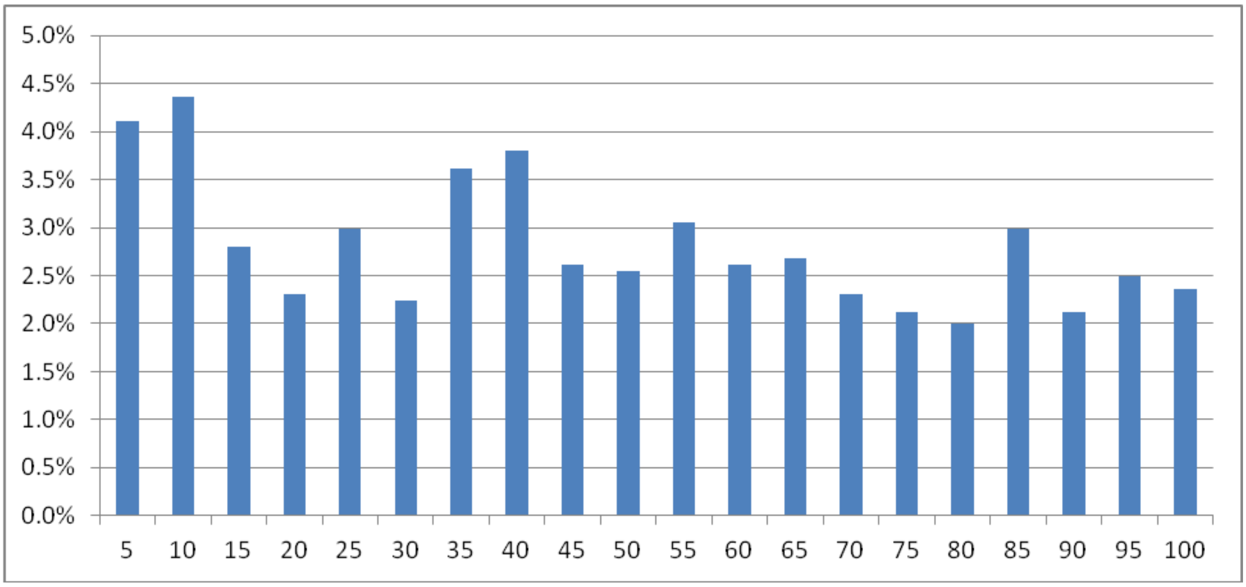
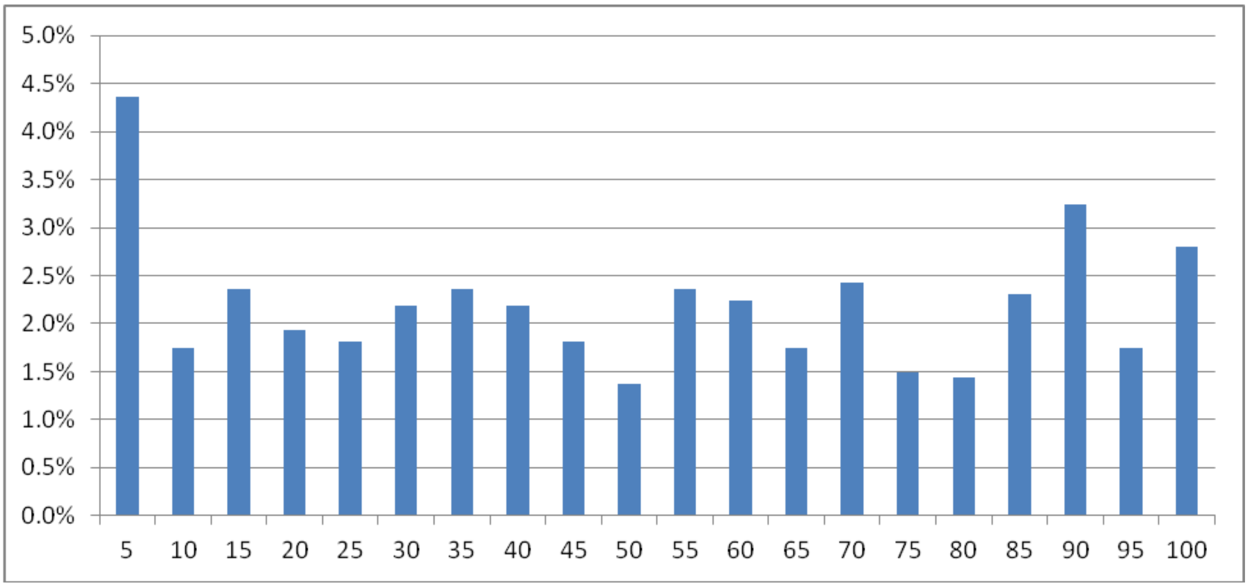


Диаграмма 3



По диаграммам 2 и 3результат такой же – гипотеза не подтверждается. Но некоторые закономерности для моей совокупности респондентов всё же могу отметить. Частота результатов от 0 % до 10%, т.е. в начале предлагаемых числовых отрезков, преобладает и составляет 14,6% от всех результатов.

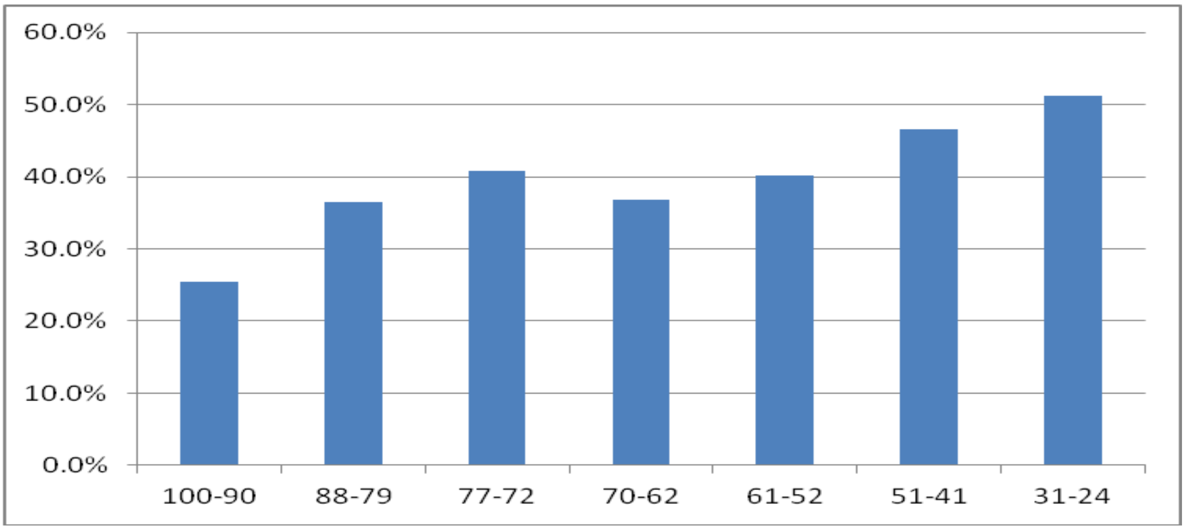
В пределах от 30% до 40% частота попаданий составляет 12%, от 50% до 60% – 10,3%, а в пределах от 80% до 90% частота попаданий составляет 10,6%.

Я проделал ещё некоторые попытки выявить закономерности в полученных результатах. Результаты расположились так: от 0 % до 50% доля частоты попадания результатов составила 51,6%, в 50% попали 1,6%, а более 50% доля составила 46,5%. Полученные результаты закономерностью не назовешь.

Частота попадания результатов в центральную зону увеличивается с уменьшением длины задаваемого числового интервала. И при минимальной длине отрезка (от 24 до 31) частота результатов попадающих в центральную зону (от 30% до 70%) составила 51,2% от всех результатов для этих отрезков.

Диаграмма 4 показывает, как зависит частота попадания в центральную зону от длины отрезка, предложенного респонденту.

Диаграмма 4



При большой длине отрезка частота попадания в центральную зону была минимальной (при длине 90 – 100 она составила 25,4% от всех результатов для этих отрезков). При уменьшении длины отрезка частота попадания в центральную зону растет. Например, для отрезков длиной 24 –

31 частота попадания в центральную зону в 2 раза выше, чем для отрезков длиной 90 –100.

Это уже больше похоже на закономерность или гипотезу.

Комментарий. Исходный текст работы Кирилла труден для понимания, поэтому пришлось значительно его редактировать, сохраняя структуру и мысли автора, насколько мы их поняли. Удалена диаграмма, которая была явно неудачной и лишней. Значительной переделке подверглась заключительная часть эссе, где автор пытается сформулировать наблюдение – с ростом длины отрезка все меньшая доля респондентов «попадает в его центр». Надеюсь, автор нас простит. А если мы неверно его поняли – поправит.

Почти все авторы эссе «Выбор числа» показали, что выдвинутая гипотеза несостоятельна (собственно, у нас не было никакого определённого мнения относительно ее справедливости). Несколько авторов обнаружили преобладание частот вблизи концов отрезков.

И все же мы отдали предпочтение работе Кирилла Стефанова из 24 лицея г.Нерюнгри, поскольку Кирилл, несмотря на неловко выраженные мысли (недостаток литературных навыков Кирилл скоро преодолеет), сделал, явно сформулировал и наглядно представил на диаграмме занятное и неожиданное наблюдение – **чем длиннее отрезок, тем меньшая часть людей попадает в его центральную часть.** Как пишет автор, «это уже больше похоже на закономерность или гипотезу». Не исключено, что именно эту гипотезу – гипотезу Стефанова – мы предложим проверить в следующей олимпиаде.

4. (2012) Секрет успеха.

В таблице представлены некоторые данные по 38 странам мира. В первом столбце — расходы государства на дошкольное образование. Второй столбец — расходы на начальное и среднее общее образование (в процентах от ВВП на душу населения). Третий столбец — средний балл, полученный школьниками этой страны в тестировании PISA 2006 года — чем выше балл, тем лучше школьники справились с тестом. Мы приготовили файл esse2.xls. Его можно скачать на странице «Задачи».

Интересно, наблюдается ли связь между финансированием образования и результатами теста PISA? Попробуйте исследовать представленные данные, сделать выводы и обосновать их. В качестве инструмента можно использовать диаграммы рассеивания и другие подходящие статистические средства. Не забудьте, что если вы высказываете гипотезы, то лучше эти гипотезы обосновать. Напишите на эту тему эссе — 2 – 3 страницы.

Страны	Расходы на дошкольное образование	Расходы на начальное и среднее образование	Средний международный балл PISA
Норвегия	11	40	490
США	20	47	474
Ирландия	16	52	501
Швейцария	11	83	530
Нидерланды	16	71	531
Австрия	19	54	505
Исландия	23	51	506
Дания	15	53	513
Швеция	16	61	502
Великобритания	21	48	495
Бельгия	15	47	520
Германия	17	65	504

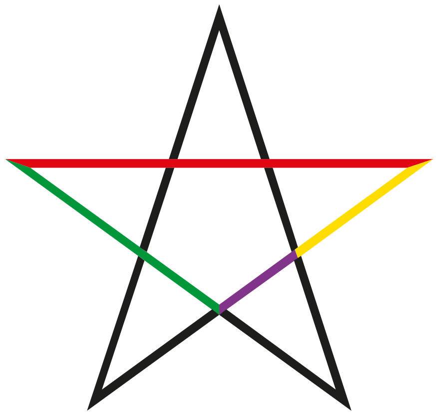
Финляндия	14	41	548
Япония	14	48	523
Франция	16	48	496
Испания	18	47	480
Италия	24	55	462
Новая Зеландия	19	62	522
Словения	29	31	504
Израиль	15	63	442
Корея	15	53	547
Чешская Республика	16	48	510
Португалия	23	55	466
Эстония	10	73	515
Венгрия	25	74	491
Словакия	18	34	492
Литва	23	57	486
Латвия	25	66	486
Польша	31	73	495
Мексика	15	31	406
Россия	22	22	476
Чили	20	30	411
Аргентина	12	32	381
Румыния	11	36	415
Болгария	29	63	413
Уругвай	11	22	427
Бразилия	14	33	370
Иордания	17	37	384

Среди присланных нам эссе на тему **«Секрет успеха»** мы не сумели выбрать достойное публикации. Практически все авторы заметили отсутствие видимой связи между финансированием дошкольного образования и результатами теста PISA, а также наличие слабой связи между финансированием среднего образования и результатами теста. Однако только один автор (участник 91224) догадался проверить достоверность обнаруженной связи. Этот автор пишет: «ни один из парных коэффициентов корреляции не показывает наличие сильной связи. …Можно убедиться, что в нашей модели есть неучтенные факторы, которые могли бы сыграть большую роль. … Поэтому строить регрессионную модель, опираясь только на эти данные, нет смысла».

К сожалению, на этом исследование заканчивается – обнаружив отсутствие значимой связи, автор, вероятно, сделал верный вывод, но не стал фантазировать на тему неучтённых факторов. А жаль – это и было бы самым интересным: сотни учёных ломают голову над вопросом, почему в одних странах школьники пишут тест PISA намного лучше, чем в других.

5. (2013) Золотое сечение.

Число φ (фи) называется «золотым числом» или числом золотого сечения. Золотое сечение разбивает отрезок на две части таким образом, что большая часть относится к меньшей части, как весь отрезок относится к большей части.



Из данного определения следует: $\frac{\varphi}{1}=\frac{(1+\varphi)}{\varphi}$, откуда $\varphi^2-\varphi-1=0$. Решив уравнение, получаем: $\varphi=\frac{\sqrt{5}+1}{2}\approx 1{,}618$.

Последовательность Фибоначчи

1, 1, 2, 3, 5, 8, 13, …

связана с золотым сечением — в ней отношение каждого члена к предыдущему приближается к φ. Золотое сечение легко увидеть в правильной пятиконечной звезде — точка пересечения двух отрезков разбивает каждый из них в золотом отношении.

Считается, что золотое сечение играет важную роль в природе и искусстве. Многие полагают, что древние мастера создавали свои гармоничные произведения, используя число φ и его степени. Например, пропорции пирамиды Хеопса, храмов, барельефов, предметов быта и украшений из гробницы Тутанхамона свидетельствуют о том, что египетские мастера пользовались золотым сечением. В фасаде древнегреческого храма Парфенона также присутствуют золотые пропорции. Лука Пачоли, современник и друг Леонардо да Винчи, называл это отношение «божественной пропорцией». Сам Леонардо да Винчи нашел в человеческом теле множество отношений φ, φ^2, φ^3 и т.д. Начиная с Леонардо да Винчи, многие художники сознательно использовали пропорции «золотого сечения». Московский архитектор Жолтовский также использовал золотое сечение в своих проектах.

Известно, что Сергей Эйзенштейн искусственно построил фильм «Броненосец Потёмкин» по правилам золотого сечения, разбив ленту на пять частей (в первых трёх действие развивается на корабле, в двух последних — в Одессе), причем граница между третьей и четвертой частью в точке золотого сечения. Утверждается даже, что композиция некоторых произведений А.С.Пушкина и Л.Н.Толстого использует золотые точки.

Термин «золотое сечение» был введён Мартином Омом в 1835 году.

В связи с чудесными свойствами золотого сечения возникло множество утверждений о том, что человеческому глазу золотые пропорции симпатичнее других, что человек подсознательно выбирает среди всех пропорций золотые как наиболее привлекательные и т.п.

Однако есть также основания считать, что значимость золотого сечения в искусстве преувеличена и базируется на ошибках и неподтвержденных мифах. Например, при обсуждении оптимальных соотношений сторон прямоугольников (форма листов бумаги, фотографий, телевизионных экранов) были испытаны разные варианты. Оказалось, что большинство людей считает прямоугольники, у которых длина больше ширины в φ раз «слишком вытянутыми».

Мы предлагаем Вам провести самостоятельное исследование, цель которого — дать аргументы за или против мнения о том, что «золотое сечение — сечение гармонии, приятное глазу». Мы не будем ограничиваться одним экспериментом, а поставим три эксперимента.

1. Выбор формы для блокнота.

2. Разбиение отрезка на две части.

3. Выбор разбиения прямоугольника (дверцы для шкафа).

В первом тесте золотое сечение — отношение двух сторон прямоугольника.

Во втором и третьем тестах золотое сечение используется иначе — как разбиение отрезка или прямоугольника по высоте.

Для проведения исследования Вам потребуется скачать файл тестов, распечатать его, разрезать по пунктирным линиям на карточки и предложить родителям, друзьям и соседям пройти эти тесты. Запомните:

1. Каждый человек (респондент) проходит тест **только один раз**.

2. Каждый человек проходит тест **самостоятельно**, опираясь только на свое мнение.

3. Всего нужно опросить **не менее 30** (а лучше — больше) респондентов.

4. Для каждого **нового респондента придется снова распечатывать** и разрезать тест.

После сбора данных их нужно обработать. Каждый тест обрабатывается своим способом и по каждому тесту делаются свои выводы.

Файл с тестами можно скачать на сайте олимпиады.

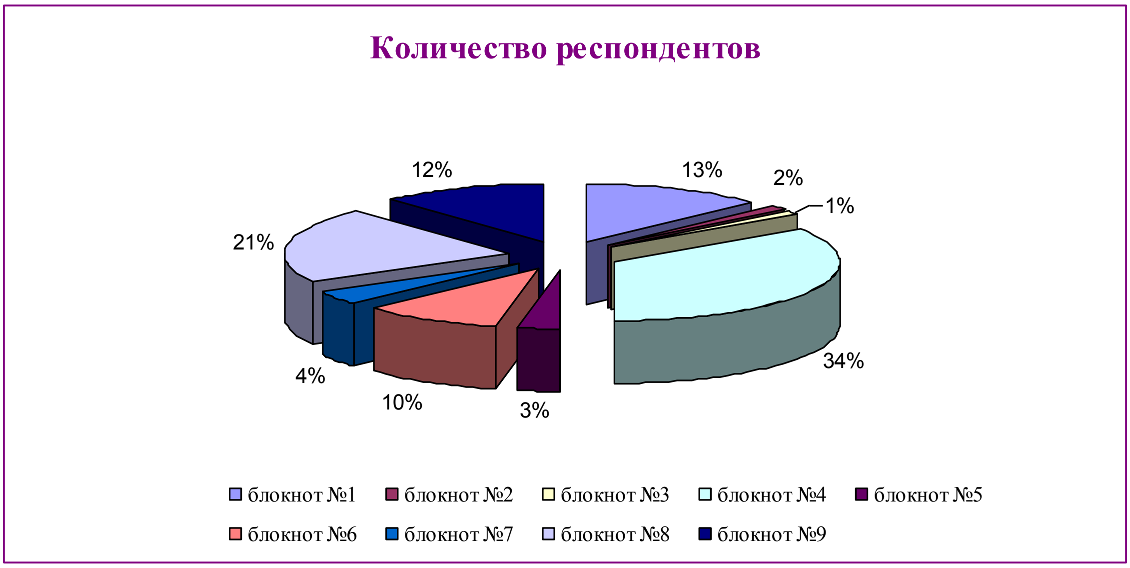
Файл с правилами и примерами обработки можно скачать на сайте.

Как и какие выводы можно сделать — пусть Вам подскажут Ваши знания, фантазия и статистическая интуиция.

Эссе одного из участников олимпиады (редактировано).

В исследовании, считают ли люди «золотое сечение» наиболее гармоничным и приятным для глаза, принимали участие 100 респондентов. Опрос проводился в общеобразовательной школе.

В первом тесте респондентам было предложено выбрать две наиболее удобные для них формы блокнота из 9 предложенных. На диаграмме показано распределение предпочтений.

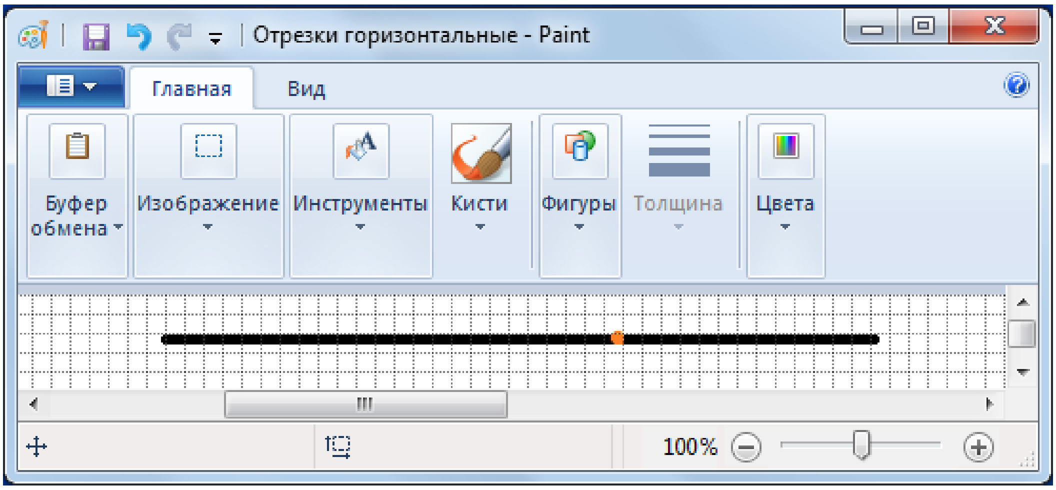


Как показывает диаграмма, самым популярным вариантом оказался вариант 4, сильно уступает ему вариант 8, примерно одинаково популярны варианты 1, 6 и 9, варианты же 2, 3, 5 и 7 оказались непредпочтительны для абсолютного большинства респондентов.

Из всех вариантов наиболее приближенным к $\frac{1}{\varphi}$ отношением длины и ширины блокнота обладают варианты 5 и 6. По результатам опроса вариант 5 оказался не самым популярным (3%), а вариант 6 набрал 10% от общего количества голосов.

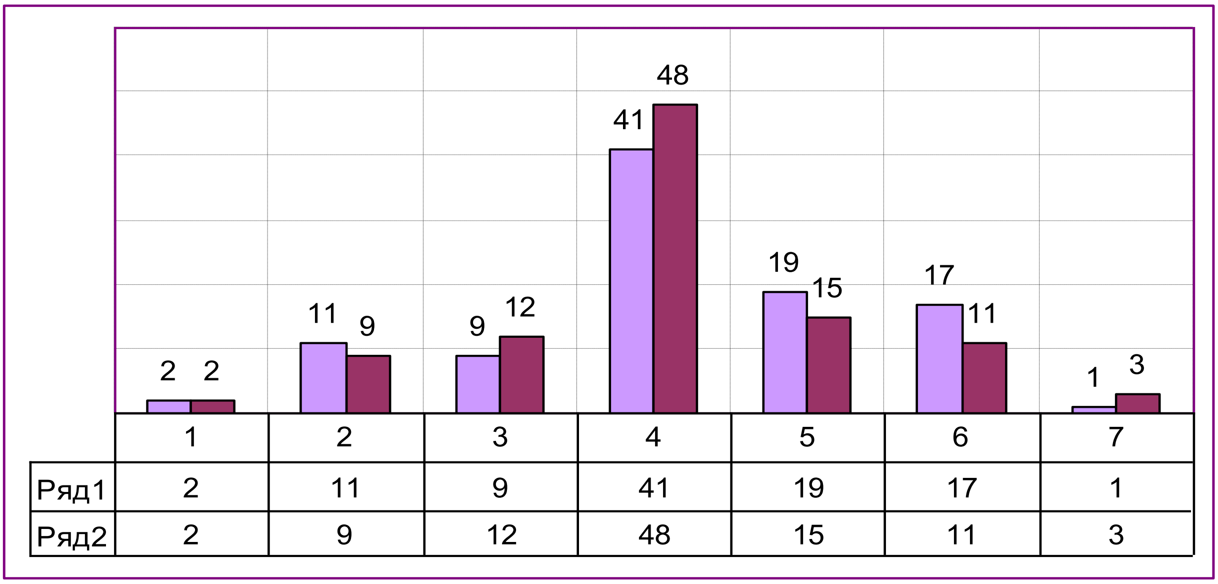
Можно сделать вывод: большинству не нравятся блокноты с «золотым» отношением длины и ширины листа.

Во втором тесте респондентам предлагалось выбрать произвольную точку на двух данных отрезках. Для удобства опроса мы не стали распечатывать тест, а нарисовали в программе Paint нужные отрезки, а опрашиваемые ставили на своём отрезке точку. Чтобы они не видели точки, поставленные предыдущими респондентами, мы сузили окно и подобрали масштаб так, чтобы респондент видел только свой отрезок.



После того, как все точки были отмечены, мы заметили, что приличное количество точек попало в область, близкую к середине отрезка. Для обработки полученных результатов мы разбили каждый отрезок на 7 равных частей и проанализировали, в какую из них попало наибольшее количество точек. Число «7» было выбрано нами неспроста — ведь в таком случае точки, делящие отрезок в отношении, близком к 1:φ попадают в 3 и 5 области, а точки, делящие отрезок в отношениях, близких к 1:1, попадают в 4 область.

Мы вычислили количество точек, попавших в каждую из частей отрезка, и построили диаграмму.



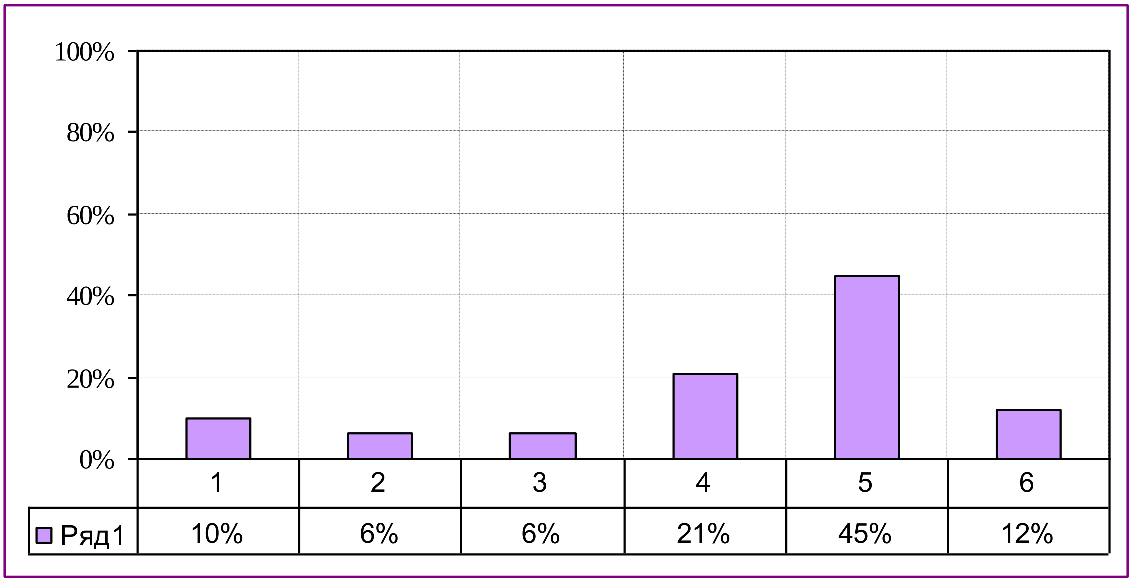
	1	2	3	4	5	6	7
Ряд1	2	11	9	41	19	17	1
Ряд2	2	9	12	48	15	11	3

Видно, что наибольшее количество (41% для первого отрезка, 48% для второго отрезка) точек попало в 4 область, т.е. ближе к середине отрезка. Также видно, что частота точек медленно повышается с приближением к середине. Точек же, близких к значению «золотого сечения», т.е. попавших в области 3 и 5, не так много — 28% для первого отрезка, 27% для второго отрезка, а в области 1 и 7 и того меньше — 3% и 5%.

По результатам второго теста можно сделать такие выводы: большая часть людей выбирает на отрезке точку, которая делит отрезок в отношении не $1:\varphi$ или $\varphi:1$, а примерно пополам.

В последнем, третьем эксперименте респонденты выбирали наиболее привлекательный на их взгляд дизайн декора дверцы шкафа из 6 вариантов.

Результаты опроса:



Из диаграммы видно, что предпочтение отдано 4 и 5 вариантам. При этом отношением длин декоративных панелей, близким к «золотому», обладают именно варианты 4 и 5.

Из результатов этого опыта можно сделать следующий вывод — что наиболее гармоничными люди считают отношения, близкие к «золотому сечению».

Итак, по результатам всех трёх опросов оказалось, что люди считают более красивыми и гармоничными отношения, близкие к «золотому сечению», однако чаще используют в обычной жизни и подсознательно выбирают отношения, далёкие от «золотого» (1:1, 3:5 и т.д.)

§ 2.3. Поиск статистического метода

1. (2012) Численность толпы.

На фотографии — толпа людей. Как оценить (подсчитать приблизительно) число людей в этой толпе? Может быть, можно что-то придумать?

Попробуйте придумать подходящий метод, воспользоваться им и сделать оценку численности толпы на фотографии. Напишите подробно, что за метод вы придумали, почему он должен правильно работать, как им пользоваться и что у вас получается. С интересом ждем ваших результатов — сколько же людей на фото?

Эссе одного из участников олимпиады (редактировано).

Количество людей на фотографии можно приблизительно вычислить многими способами. Для их сравнения введем понятия погрешность и трудоемкость (сложность).

Погрешность – это разность между приближенным и действительным значением. Самый точный метод (наименьшая погрешность) – посчитать количество людей на фото, но он является самым трудоемким. Трудоемкость (сложность) – время, потраченное на подсчеты (прямо пропорционально количеству подсчитанных людей).

Впрочем, помимо метода прямого подсчета, есть другие, менее точные и менее трудоемкие.

Возьмем, например, метод «Регионы». Он заключается в том, что мы разбиваем всё фото на равные регионы – прямоугольники, и подсчитываем количество людей в левом верхнем и в правом нижнем регионах (т.к. плотность людей на фото увеличивается справа налево и снизу вверх), это количество умножаем на количество регионов. Точность и трудоемкость этого метода зависит от количества регионов: чем больше регионов, тем больше точность и трудоемкость.

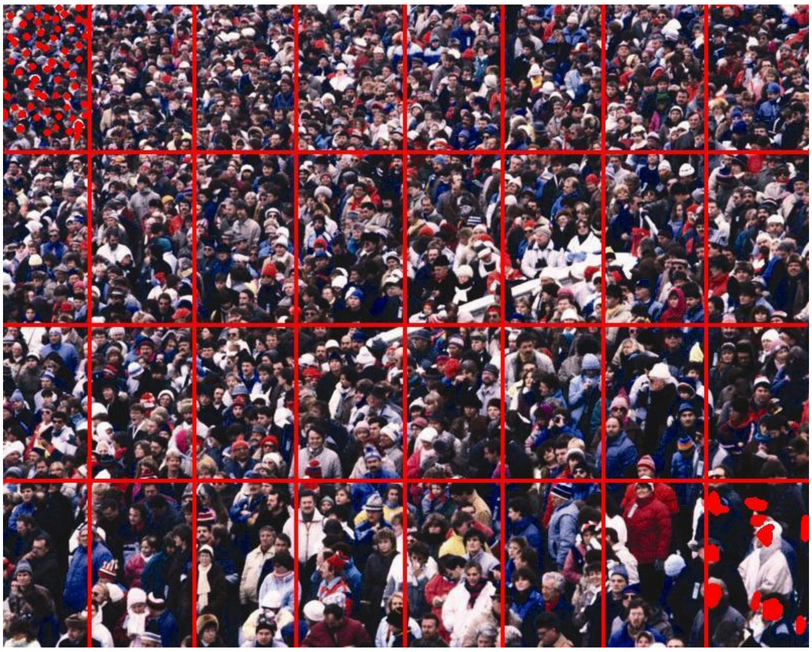


Рис. 1. Метод «Регионы» (красным отмечены посчитанные люди)

В левом верхнем регионе 60 человек. В правом нижнем регионе – 10 человек. В среднем в регионах по 35 человек. Количество регионов – 32. Человек на фото $32 \cdot 35 = 1120$.

Гораздо более простым является метод «Прямоугольник». Для этого вычисляем среднюю «длину» и среднюю «ширину» в людях прямоугольника – фотографии.

Рис. 2. Метод «Прямоугольник» (красным отмечены посчитанные люди)

Ширина слева равна 40 человек, справа – 20 человек. Средняя ширина - 30 человек.

Длина снизу равна 20 человек, сверху – 50 человек. Средняя длина – 35 человек.

Площадь (количество людей) равна 30*35=1050.

Как вариант этого метода – метод «Прямоугольники»: разбиваем фотографию на несколько равных прямоугольников и вычисляем «площадь» каждого из них методом «Прямоугольник».

Метод «Прямоугольники», безусловно, более точен, чем «Прямоугольник», но и более трудоемок.

Трудоемкость этого метода пропорциональна количеству прямоугольников.

Прямоугольник	Средняя ширина	Средняя длина	Площадь
1	(16+30)/2=23	(20+34)/2=27	621
2	(12+14)/2=13	(10+22)/2=16	208
3	(14+10)/2=12	(16+12)/2=14	168
4	(24+20)/2=22	(20+24)/2=22	484

Общее количество людей оценивается числом 621+208+168+484=1481.

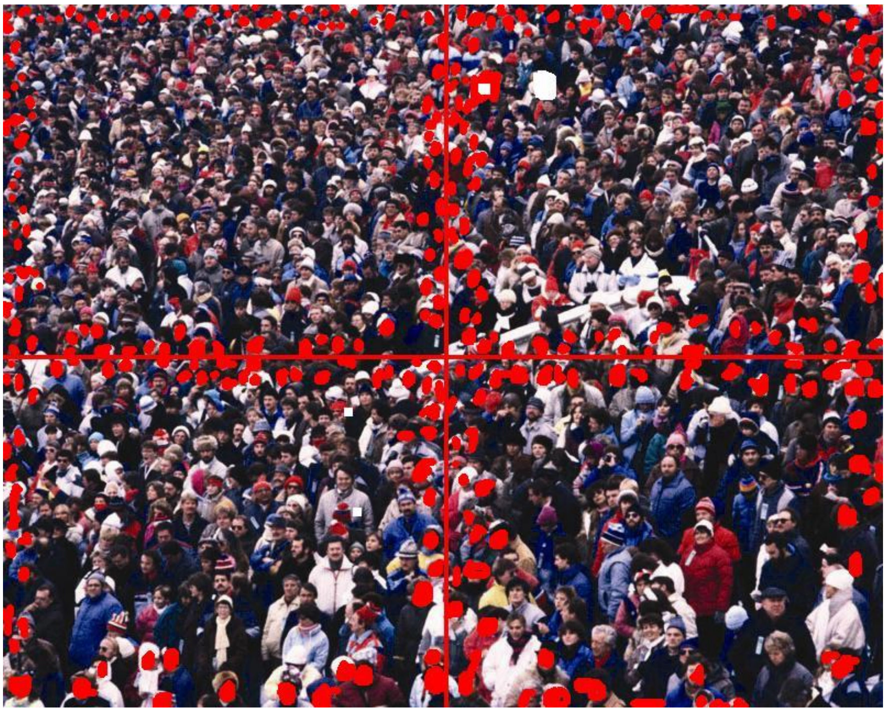


Рис. 3. Метод «Прямоугольники» (красным отмечены посчитанные люди)

На мой взгляд, самым приемлемым с точки зрения практических подсчетов является метод «Прямоугольник». Однако из всех полученных мной результатов наиболее близким к истине является результат, полученный методом «Прямоугольники».

Ответ: Примерно 1300 человек (среднее арифметическое всех результатов).

Комментарий. Мы взяли на себя смелость немного отредактировать текст участника. Например, разность между точным значением и оценкой мы назвали *погрешностью* (а не точностью).

Разумеется, в работе отсутствует оценка погрешности (это весьма непростая задача), однако заслугой участника является то, что она задумалась о связи погрешности и трудоёмкости измерения – чем меньше трудоемкость, тем, как правило, больше погрешность (меньше точность).

Другое достоинство работы состоит в том, что Алия применила три разных подхода к подсчету количества людей и получила результаты одного порядка (1120, 1050 и 1481). Это, безусловно, дает автору основание считать, что ее методы работают более-менее в согласии друг с другом и погрешность не очень велика (вспомним, что по разным публикациям в СМИ оценки численности толпы во время митингов различаются в два, а то и в десять раз).

Почему-то автор считает, что лучший метод «Прямоугольники». Возможно, это подсказывает интуиция, хотя трудно сказать, ошибается она в этом случае или нет. Но вместо того, чтобы выбрать «самый лучший результат» Алия усредняет результаты, полученные тремя разными способами. Видимо, именно здесь проявляется статистическая интуиция автора. Усреднение трёх разных оценок (четвёртая оценка) является третьим достоинством работы.

Подводя итоги, скажем, что автор работы трижды проявил высокую статистическую культуру и интуитивную грамотность:

1. Заметила связь между точностью и трудоёмкостью и поставила вопрос о компромиссе между ними.

2. Задумалась о надёжности своих выводов и, не умея оценить погрешность, пошла другим путем: убедилась в том, что три разных метода дают приемлемо близкие оценки.

3. Погадав, какая из оценок достовернее, участница сделала лучшее, что могла: придумала метод получения четвёртой оценки – усреднение имеющихся трёх.

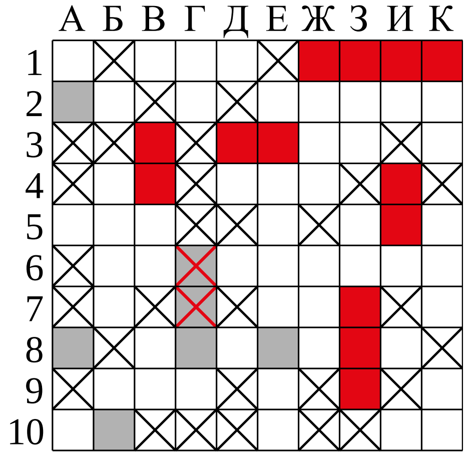
Эти три обстоятельства привели к тому, что мы признали эту работу лучшей среди всех эссе, представленных на эту тему.

2. (2014) Морской бой.

Надеемся, вы умеете играть в морской бой. Если нет, то правила игры найдутся в интернете. А можно спросить у знающих людей.

Морской бой – игра вероятностная. Почти каждый раз игрок делает выстрел наудачу. Вероятность попадания зависит от того, сколько на поле противника осталось свободного места, как расположены корабли, какой стратегии (схемы игры) придерживается сам игрок.

Мы остановимся только на одном аспекте – на расположении кораблей в начале игры. Зависит ли вероятность выигрыша от того, как именно противник расставил корабли? Каких правил расстановки нужно придерживаться, чтобы увеличить свои шансы на победу? А как корабли расставлять не следует? Попробуйте провести вероятностный анализ этой задачи настолько полный, насколько это возможно. Чем более глубокий и аргументированный анализ вы проведете, тем лучше и тем больше баллов вы заработаете на этой непростой задаче.



Возможная идея исследования. Очевидно, вероятность попасть в корабль тем меньше, чем больше осталось клеток, где корабль может прятаться. Если корабль противника стоит «в открытом море», то при попадании в него из дальнейшей игры исключается не только сам корабль, но и вокруг лежащие клетки. Значит, нужно минимизировать число клеток, расположенных вокруг кораблей. Это удастся сделать, если расставить корабли вдоль берегов или даже в углах поля. В особенности это касается больших кораблей. Именно на эту тему и предлагалось пофантазировать в данном задании. Многие участники предложили именно эту идею и даже сделали количественные попытки оценить вероятность поражения малых кораблей в начале боя.

3. (2014) Куда лететь быстрее?

Правда ли, что самолёт тратит на полёт разное время, в зависимости от того, на восток или на запад он летит? Всегда ли это так? Зайдите на сайт крупного аэропорта или крупной авиакомпании, а лучше – нескольких авиакомпаний или аэропортов. Выберите перелёты с востока на запад и обратно. Соберите и обработайте нужную информацию. Насколько отличаются продолжительности полётов в ту и в другую сторону? Зависит ли это от расстояния? Нужно придумать статистический показатель, описывающий разницу. Устойчивая ли она? Трудность в том, что просто разность времени туда и времени обратно нам мало что даст – ведь нужно учитывать и очень дальние перелёты и сравнительно короткие. Если разница есть, с чем она может быть связана? Можно ли оценить постоянство и силу этого удивительного фактора? Действительно ли виноваты восток и запад? Может быть, похожая картина наблюдается и с другими полётами, например, с севера на юг и обратно? Вопросов множество. Попробуйте разыскать, описать, проанализировать данные и пофантазировать.

1. Сбор и представление данных

Мы сделали случайную выборку, в которую вошли 70 регулярных рейсов (для более представительных результатов нужно больше, но для иллюстрации достаточно и этого).

На восток			На запад			$T_2 - T_1$	$T_1 + T_2$	δ
Рейс	Пункт А	T_1	T_2	Пункт В	Рейс			
R2 731	Москва	1:55	2:00	Оренбург	R2 732	0:05	3:55	0,0213
PS 9262	Париж	3:05	3:20	Киев	PS 9263	0:15	6:25	0,0390
MS-768	Барселона	4:00	4:30	Каир	MS-767	0:30	8:30	0,0588
VX90	С.-Франциско	5:05	6:00	Вашингтон	VX77	0:55	11:05	0,0827
AM680	Мехико	5:09	5:57	Монреаль	AM681	0:48	11:06	0,0721
V0 7201	Панама Сити	2:20	2:00	Каракас	V0 7208	–0:20	4:20	0,0769
EK0090	Женева	6:20	6:50	Дубай	EK0089	0:30	13:10	0,0380
KC 140	Тбилиси	3:35	4:10	Алматы	KC 139	0:35	7:45	0,0753
AB8486	Берлин	2:10	2:20	С.-Петербург	AB8487	0:10	4:30	0,0370
SU-2393	Цюрих	3:20	3:30	Москва	SU-2392	0:10	6:50	0,0244
TA-917	Лима	5:00	5:05	Сан-Паоло	TA-916	0:05	10:05	0,0083
WS-724	Ванкувер	4:26	5:03	Торонто	WS-719	0:37	9:29	0,0650
TK-636	Триполи	2:45	2:55	Стамбул	TK-635	0:10	5:40	0,0294
ET-910	Абуджа	4:30	4:40	Аддис-Абеба	ET-911	0:10	9:10	0,0182
UT345	Москва	3:20	3:50	Омск	UT346	0:30	7:10	0,0698
S7 3261	Новосибирск	4:10	4:20	Якутск	S7 3262	0:10	8:30	0,0196
W2 3121	Рейкявик	3:30	3:45	Берлин	W2 3124	0:15	7:15	0,0345
RO 416	Мадрид	3:35	4:00	Бухарест	RO 415	0:25	7:35	0,0549
IB 5885	Лиссабон	2:20	2:20	Любляна	IB 5884	0:00	4:40	0,0000
SU1410	Москва	2:20	2:30	Екатеринбург	SU1411	0:10	4:50	0,0345
AZ 7180	Милан	1:55	2:05	София	AZ 7179	0:10	4:00	0,0417
UA 122	Вашингтон	7:25	8:25	Лондон	UA 123	1:00	15:50	0,0632
FI 644	Вашингтон	5:45	6:20	Рейкьявик	FI 645	0:35	12:05	0,0483
DL 1610	Денвер	2:49	3:18	Детройт	DL 1645	0:29	6:07	0,0790
UA 1279	Денвер	2:38	2:48	Новый Орлеан	UA 295	0:10	5:26	0,0307
DL 1577	Лос-Анджелес	3:58	3:56	Новый Орлеан	DL 1325	–0:02	7:54	0,0042
AM 7901	Мехико	3:47	3:51	Панама	AM 7920	0:04	7:38	0,0087
AR 1365	Лима	4:35	5:25	Буэнос-Айрес	AR 1364	0:50	10:00	0,0833
SU 6276	Рим	3:40	3:45	С.-Петербург	SU 3276	0:05	7:25	0,0112
UT 439	Москва	4:05	4:20	Томск	UT 440	0:15	8:25	0,0297
SU 2193	Амстердам	3:15	3:30	Москва	SU 2192	0:15	6:45	0,0370
KL 1395	Амстердам	2:50	3:00	С.-Петербург	KL 1396	0:10	5:50	0,0286
KL 1363	Амстердам	1:55	2:05	Варшава	KL 1364	0:10	4:00	0,0417
B2 868	Амстердам	2:30	2:40	Минск	B2 867	0:10	5:10	0,0323
OU 451	Амстердам	1:55	2:05	Загреб	OU 450	0:10	4:00	0,0417

TP 802	Лиссабон	2:35	2:40	Милан	TP 809	0:05	5:15	0,0159
S7 4302	Лиссабон	5:25	5:35	Москва	S7 4301	0:10	11:00	0,0152
TP 536	Лиссабон	3:20	3:40	Берлин	TP 537	0:20	7:00	0,0476
TP 692	Лиссабон	3:20	3:35	Вена	TP 691	0:15	6:55	0,0361
AB 5032	Лондон	1:35	1:40	Гамбург	AB 5033	0:05	3:15	0,0256
BA 798	Лондон	2:50	3:05	Хельсинки	BA 799	0:15	5:55	0,0423
UN 7401	Лондон	3:55	4:10	Москва	UN 7402	0:15	8:05	0,0309
BA 840	Лондон	3:10	3:30	Киев	BA 841	0:20	6:40	0,0500
BA 884	Лондон	3:10	3:25	Бухарест	BA 887	0:15	6:35	0,0380
IB 5300	Барселона	4:10	4:35	Москва	IB 5757	0:25	8:45	0,0476
VY 7892	Барселона	4:00	4:15	С.-Петербург	VY 7893	0:15	8:15	0,0303
SU 6741	С.-Петербург	2:50	2:40	Уфа	SU 6742	–0:10	5:30	0,0303
SU 6711	С.-Петербург	2:25	2:35	Пермь	SU 6712	0:10	5:00	0,0333
SU 6765	С.-Петербург	2:15	2:15	Казань	SU 6766	0:00	4:30	0,0000
SU 6753	С.-Петербург	2:25	2:30	Самара	SU 6754	0:05	4:55	0,0169
SU 6541	С.-Петербург	4:20	4:50	Новосибирск	SU 6542	0:30	9:10	0,0545
SU1478	Москва	4:30	4:50	Абакан	SU1479	0:20	9:20	0,0357
SU 1654	Москва	6:20	7:25	Чита	SU 1655	1:05	13:45	0,0788
SU 1742	Москва	8:20	8:50	Ю.-Сахалинск	SU 1743	0:30	17:10	0,0291
S7 121	Москва	5:15	5:45	Братск	S7 122	0:30	11:00	0,0455
UT 465	Москва	5:25	5:35	Иркутск	UT 466	0:10	11:00	0,0152
UN 2101	Москва	4:10	5:10	Норильск	UN 102	1:00	9:20	0,1071
SU 330	Москва	6:10	6:40	Улан-Батор	SU 331	0:30	12:50	0,0390
UN 273	Москва	3:15	3:35	Астана	UN 204	0:20	6:50	0,0488
S7 4461	Москва	7:55	8:20	Пекин	S7 4462	0:25	16:15	0,0256
KE 924	Москва	8:20	9:20	Сеул	KE 923	1:00	17:40	0,0566
SU 270	Москва	9:15	10:00	Бангкок	SU 271	0:45	19:15	0,0390
9W 64	Дели	4:25	4:20	Бангкок	9W 63	–0:05	8:45	0,0095
JL 740	Дели	8:20	8:35	Токио	JL 749	0:15	16:55	0,0148
OZ 6378	Дели	9:45	10:25	Сеул	OZ 6377	0:40	20:10	0,0331
KE 868	Улан-Батор	3:00	3:35	Сеул	KE 867	0:35	6:35	0,0886
SV 760	Эр-Рияд	4:00	4:30	Дели	SV759	0:30	8:30	0,0588
WY 686	Эр-Рияд	2:35	2:20	Маскат	WY 685	–0:15	4:55	0,0508
LX 1310	Цюрих	3:00	3:15	С.-Петербург	LX 1311	0:15	6:15	0,0400
AF 1497	Касабланка	2:00	2:00	Париж	AF 1197	0:00	4:00	0,0000

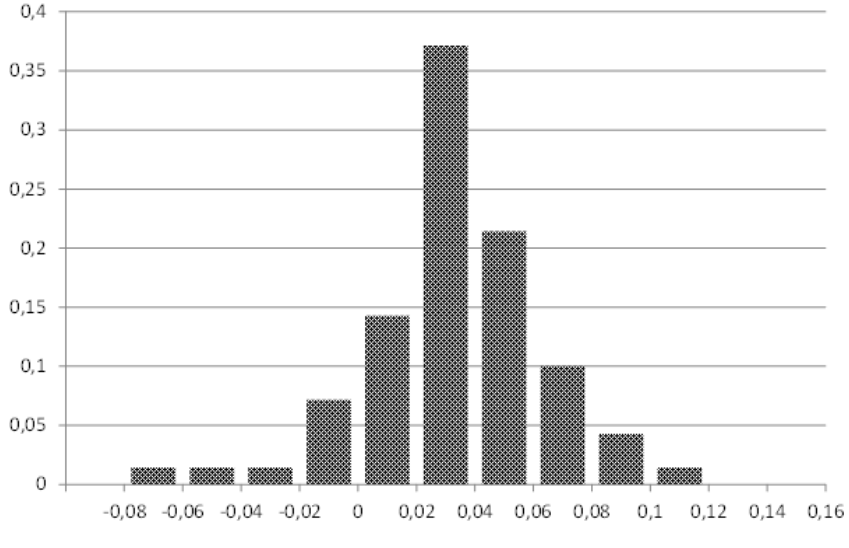
Для убедительности мы включили в таблицу номера рейсов, хотя можно было обойтись без них. Мы считаем, что авиакомпании выполняют рейсы туда и обратно, как правило, одним и тем же самолетом, поэтому тип самолета на разницу во времени полетов значительно не влияет. Время полета на восток мы обозначили T_1, а время обратного полета – T_2. Сразу приходит в голову разность $T_2 - T_1$. В таблице видно, что эта разность почти везде положительна. Пять строк, где разность отрицательна, выделены. Похоже, что *обычно на восток самолеты летят действительно быстрее*.

Описание данных

Можно ли сравнивать, например, рейс С.-Петербург–Пермь–С.-Петербург (разница 10 минут) с рейсом Москва–Сеул–Москва (разница час)? Конечно, нет – ведь от Сеула до Москвы в четыре с лишним раза дальше, чем от Петербурга до Перми. Разумный показатель задержки должен учитывать общую длительность полетов. Разделим разность на $T_1 + T_2$. Получается показатель

$$\delta = \frac{T_2 - T_1}{T_1 + T_2},$$

не зависящий напрямую от дальности перелетов, который характеризует *относительную задержку западного рейса.* Разумеется, показатель δ не идеален – он не учитывает, что маршруты редко проходят строго с востока на запад и обратно. Чаще самолеты летят «наискосок» – на северо-восток или юго-восток. Кроме того, не учтены широты, в которых проходит полет. Но это уже что-то. Посмотрим на распределение δ. Для этого сгруппируем значения подходящим образом, например, шагом 0,02 и построим гистограмму.



На рисунке без вычислений видно, что величина δ концентрируется вблизи значения 0,03. Вычисления подтверждают это – среднее арифметическое относительной задержки равно $0,0347$.

Не случайность ли?

Могла ли такая средняя задержка получиться случайно просто потому, что мы не очень удачно выбрали рейсы? Не будет ли картина иной, если мы возьмем другие авиакомпании и другие города? Можно попытаться разобраться в этом подробно с помощью *стандартного отклонения,* но это требует весьма глубоких знаний.

Применим более простое и интуитивное рассуждение. Сформулируем гипотезу*: на самом деле никакой разницы во времени нет*, и в среднем длительность полетов на восток и на запад одинакова. Тогда *положительные и отрицательные значения* δ *должны быть равновозможны*. Поэтому, если исключить из рассмотрения нулевые значения, вероятности положительных и отрицательных значений δ должны равняться 0,5:

$$\mathrm{P}(\delta<0) = \mathrm{P}(\delta>0) = 0{,}5.$$

Всего у нас 70 значений, из них 3 нулевых (рейсы Лиссабон–Любляна, Санкт-Петербург–Казань и Касабланка–Париж). Остается 67 ненулевых значений. Считая, что наша гипотеза верна, спросим себя, а сколько разумно ждать положительных и отрицательных значений? Ясно, что положительными должна оказаться примерно половина. *А у нас* 62 *из* 67 *значений оказались положительными!* Вероятность того, что при указанных условиях положительны будут 62 значения или больше, равна

$$C_{67}^{62}0{,}5^{67}+C_{67}^{63}0{,}5^{67}+\ldots+C_{67}^{67}0{,}5^{67}\approx 0{,}000000000000071,$$

что *ничтожно мало* по сравнению с вероятностью противоположного события «меньше чем 62 значения положительны». Скорее всего, исходная гипотеза ошибочна, и на самом деле вероятность положительного значения δ намного выше, чем вероятность отрицательного.

Мы, если не доказали, то *почти научно обосновали*, что полеты на запад *в среднем дольше полетов на восток и что это не случайность.*

Почему так?

Конечно, странно думать, что разница во времени получается оттого, что Земля «вращается под самолетом». Разумеется, она вращается, но самолет находится в земной атмосфере и поэтому «вращается вместе с ней», точно так же, как идущие люди или едущие автомобили. Истинная причина – устойчивое воздушное течение, попросту говоря – ветер, который постоянно дует на восток в верхних слоях атмосферы. Это течение называется *западным переносом воздушных масс*. Подробнее о западном переносе и причинах, вызывающих его, можно прочесть, например, в Википедии.

Мы можем даже оценить (найти приблизительно) скорость западного переноса на определенных широтах: ведь нам известны расстояния между городами и время в пути. Попробуйте сделать эти приближенные расчеты самостоятельно – они не сложнее задачки на движение для 7 – 8 класса.

§ 2.4. Поиск ошибки

1. (2010) Мобильные телефоны и женское образование.

В таблице из файла data2.xls во втором столбце указано число грамотных (умеющих читать и писать) женщин $X = \{x_1, \dots, x_{40}\}$, а в третьем количество мобильных телефонов $Y = \{y_1, ..., y_{40}\}$ по странам Европы. Постройте диаграмму рассеивания для пар $\{X, Y\} = \{(x_1, y_1), \dots, (x_{40}, y_{40})\}$. Пользуясь диаграммой, напишите небольшое эссе, посвященное следующим вопросам.

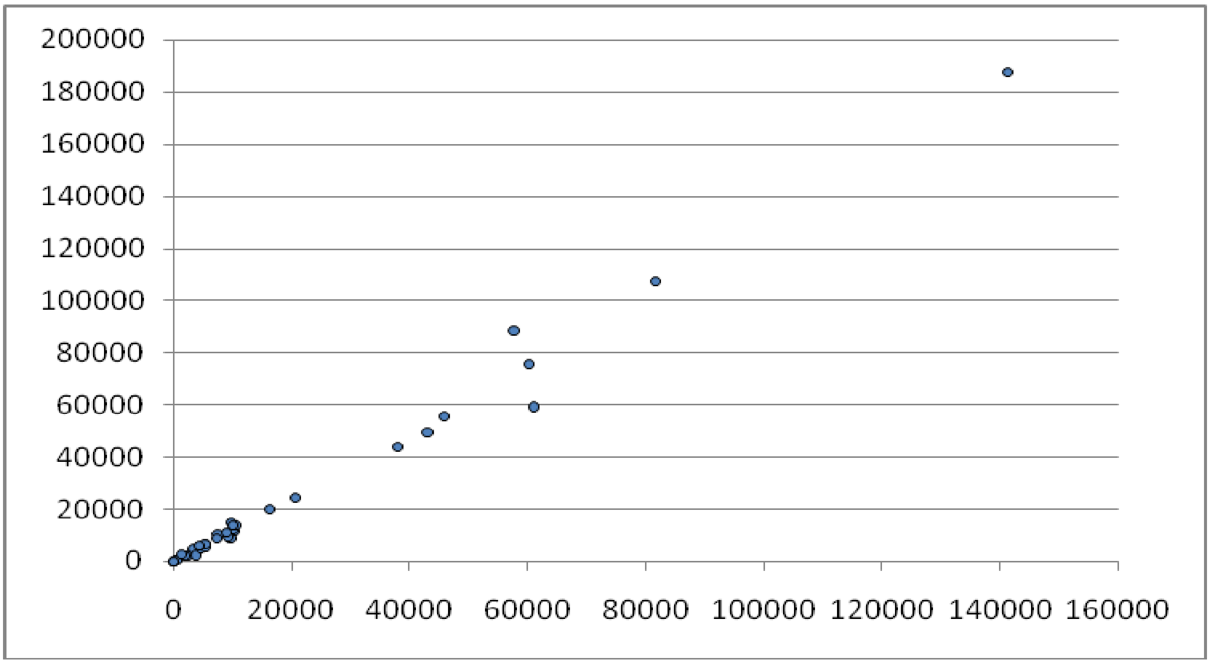
а) Можно ли считать, что чем больше в стране грамотных женщин, тем больше сотовых телефонов или наоборот?

б) Предположим, что, изучив европейский опыт, правительство некоторой развивающейся страны начало интенсивную программу по ликвидации безграмотности среди женщин, и через три года доля грамотных женщин в стране заметно выросла. Следует ли ожидать, что число мобильных телефонов за это время тоже пропорционально изменится?

в) Предположим, что правительство развивающейся страны закупило и раздало населению большое число мобильных телефонов. Следует ли рассчитывать на то, что число грамотных женщин в этой стране быстро увеличится?

г) Если на диаграмме прослеживается связь между числом грамотных женщин и числом мобильных телефонов, то можно ли объяснить связь?

Возможная идея эссе. Построим диаграмму рассеивания.



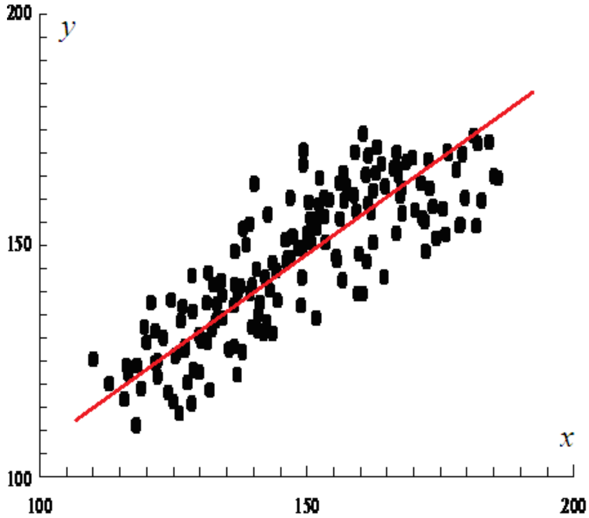
Почти все точки укладываются в одну линию. Такой вид диаграммы рассеивания должен говорить о наличии положительной взаимосвязи между количеством грамотных женщин и количеством мобильных телефонов в стране. Однако это не говорит о том, что чем больше телефонов, тем больше образованных женщин или наоборот. Рассмотрим еще одну величину N – численность населения в стране. При примерно равных условиях, характерных для большинства стран Европы, можно говорить о наличии почти линейной зависимости между X и N, поскольку чем больше население страны, тем больше найдется в этой стране грамотных женщин: $X = a_1 N + b_1$. Похожая зависимость имеет место для количества сотовых телефонов: $Y = a_2 N + b_2$. Значит, имеется почти линейная связь между X и Y:

$$Y = \frac{a_2}{a_1} X + \left(b_2 - b_1 \frac{a_2}{a_1} \right)$$

Получаем, что величины X и Y, скорее всего, связаны через третью величину, но не сами по себе.

2. (2011) Как связан рост мальчиков и рост девочек?

Однажды учительница математики решила показать своим ученикам, что рост мальчиков и рост девочек — независимые случайные величины. Для этого учительница провела исследование. В каждом классе она выбрала случайным образом 10 мальчиков и 10 девочек, случайно разбила их на пары мальчик-девочка и записала рост мальчика x_k и девочки y_k в каждой паре. Получились пары $(x_k\,;\,y_k)$.

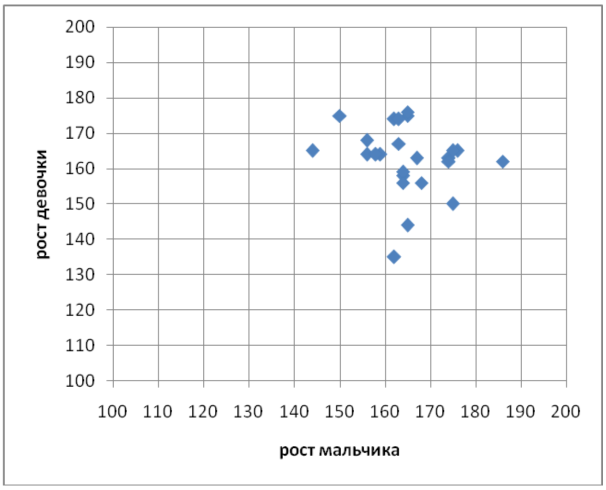


Когда учительница отметила все результаты на диаграмме рассеивания (см. рис.) она к своему ужасу обнаружила, что точки сгруппированы возле наклонной прямой (см. рис). Значит, между ростом мальчиков и ростом девочек есть очевидная связь! Как же так?

Напишите короткое эссе (1 – 2 страницы), в котором попробуйте объяснить, ошиблась ли учительница в своих выводах, и если да, то в чем состоит ее вероятная ошибка.

Эссе одного из участников олимпиады (редактировано).

Перед нами диаграмма рассеивания. Она показывает примерный характер взаимосвязи между двумя числовыми характеристиками. Если почти все точки образуют наклонное облако, то связь между рассматриваемыми величинами есть. На представленной диаграмме между ростом мальчиков и ростом девочек есть очевидная связь. Но я считаю, что рост мальчиков и рост девочек — две независимые случайные величины. На диаграмме рассеивания, которую составила учительница, видно, что почти в каждой паре мальчик и девочка были приблизительно одного роста, именно поэтому на диаграмме рассеивания прослеживается связь. Так как учительница выбирала детей из одного класса, скорее всего, эти дети приблизительно одного возраста, а значит, и близкого роста (например, пары первоклассников, пары второклассников, пары десятиклассников и т.д.).



Для того чтобы проследить возможное наличие связи или её отсутствие, надо выбрать другие пары: провести исследование среди детей одного возраста. Я думаю, что если в одной паре девочка будет выше мальчика, то в другой паре — наоборот, и тогда точки не будут группироваться возле наклонной прямой.

Я провела опрос среди учеников 7-х классов ГОУ СОШ №1000 об их росте и случайным образом распределила их по парам (мальчик, девочка) конечно, это очень небольшое количество (48 опрошенных — 24 пары), чтобы делать достоверные выводы, но у меня получилась совершенно другая диаграмма рассевания.

Глядя на эту диаграмму можно предположить, что рост мальчиков и рост девочек — две независимые случайные величины.

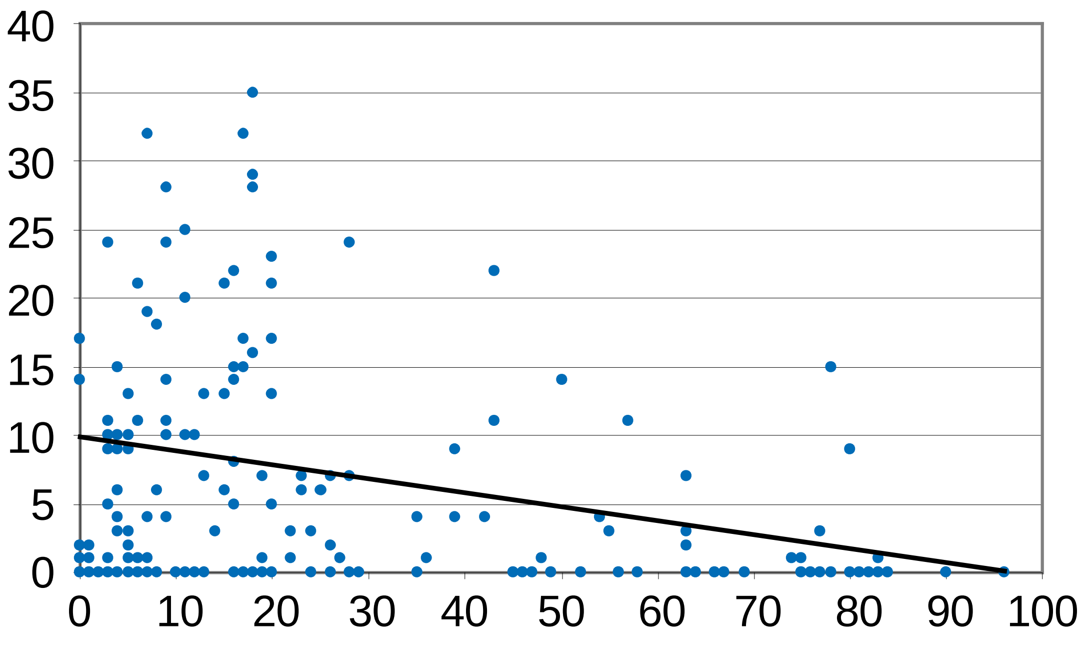
3. (2013) Влажность воздуха и снег.

Одним февральским утром, отворачивая лицо от колючей метели, господин Ноибуро Кавасаки задался вопросом — а как связаны влажность воздуха и количество снега? Чтобы ответить на вопрос, он зашел на сайт метеостанции префектуры Ниигата и нашел данные о влажности нижних слоев воздуха (ниже 2 000 м) и верхних слоев воздуха (10 000 — 12 000 м) за много дней наблюдений. Кроме того, Кавасаки-сенсэй присовокупил к этим данным информацию о толщине снежного покрова, выпавшего в этот день в месте наблюдения.

Влажность верхних слоев атмосферы (φ_B) и нижних слоев атмосферы (φ_H) измеряется в процентах, а толщина снежного покрова (H) в миллиметрах выпавшего снега.

Вот что получилось, когда г-н Кавасаки составил диаграммы рассеивания.

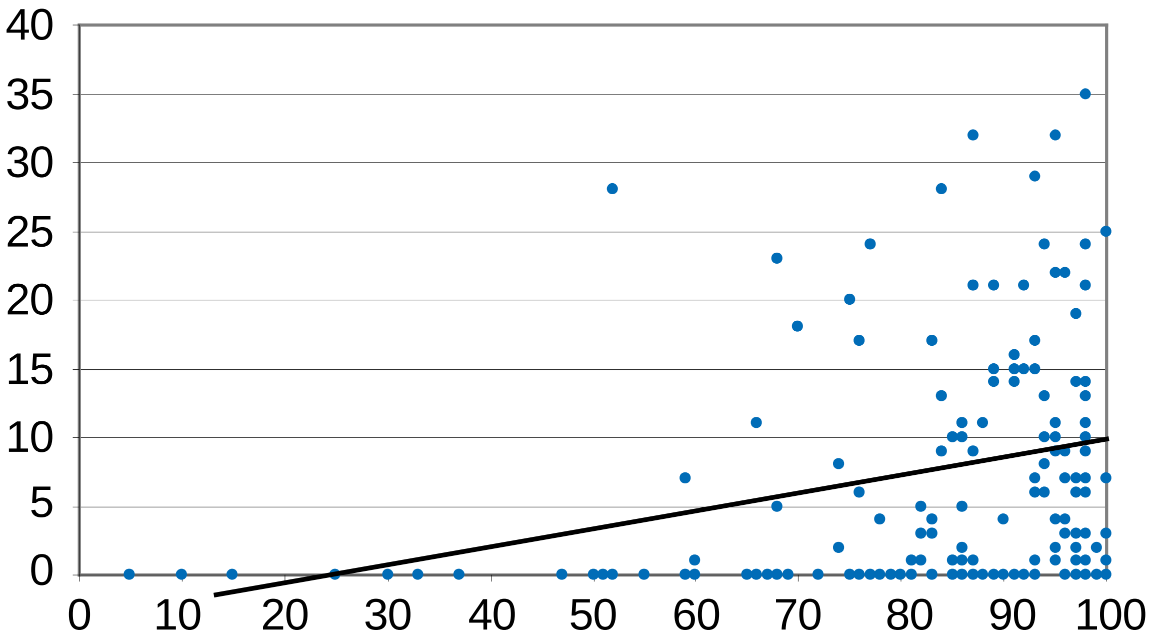
Диаграмма φ_B — H



Кавасаки-сенсэй рассчитал коэффициент линейной корреляции Пирсона. Получилось $r_{\varphi_B,H} = -0,24$. После некоторых раздумий г-н Кавасаки написал: «Имеется отрицательная корреляция. Значение коэффициента корреляции приблизительно $-0,24$. Абсолютное значение не так велико, я думаю. Однако если верхние слои воздуха сухие, то выпадает много снега».

Затем точно так же, г-н Кавасаки и его ученики построили диаграмму рассеивания для влажности нижних слоев воздуха и количества снега.

Диаграмма φ_H — H



Коэффициент линейной корреляции Пирсона оказался в этом случае равен $r_{\varphi_H,H} = 0,28$. Г-н Кавасаки так прокомментировал результат: «Аналогично, здесь имеется слабая положительная корреляция. Значение коэффициента корреляции приблизительно 0,28. Полагаю, и это абсолютное значение не слишком велико. Однако если верхние слои воздуха влажные, то выпадает много снега».

Рассмотрите полученные диаграммы. Попробуйте найти соответствующие данные в архивах сайтов метеонаблюдений (accuweather.com, rp5.ru и т.д.). Постройте свои диаграммы по наблюдениям данных величин. Разберитесь в том, что представляет собой коэффициент корреляции Пирсона и что он показывает. Подумайте, нет ли ошибок в действиях и рассуждениях господина Ноибуро Кавасаки. Напишите в коротком эссе, что Вы думаете по этому поводу. Вдруг у Вас появятся собственные мысли о том, как лучше интерпретировать полученные данные о влажности воздуха и снегообразовании.

Возможная идея исследования

Здесь нужно заметить, что прямая наименьших квадратов, скорее всего, не является подходящим средством изучения данных — форма облака точек наводит на мысль, что линейной зависимости вовсе нет. Коэффициент корреляции Пирсона также служит мерой линейной связи. Если линейной связи нет, то вычисления превращаются в малоосмысленную арифметическую игру. Эта мысль ни в одном из присланных эссе не фигурирует. Поэтому мы не сочли возможным помещать в разборе задач присланные нам тексты.

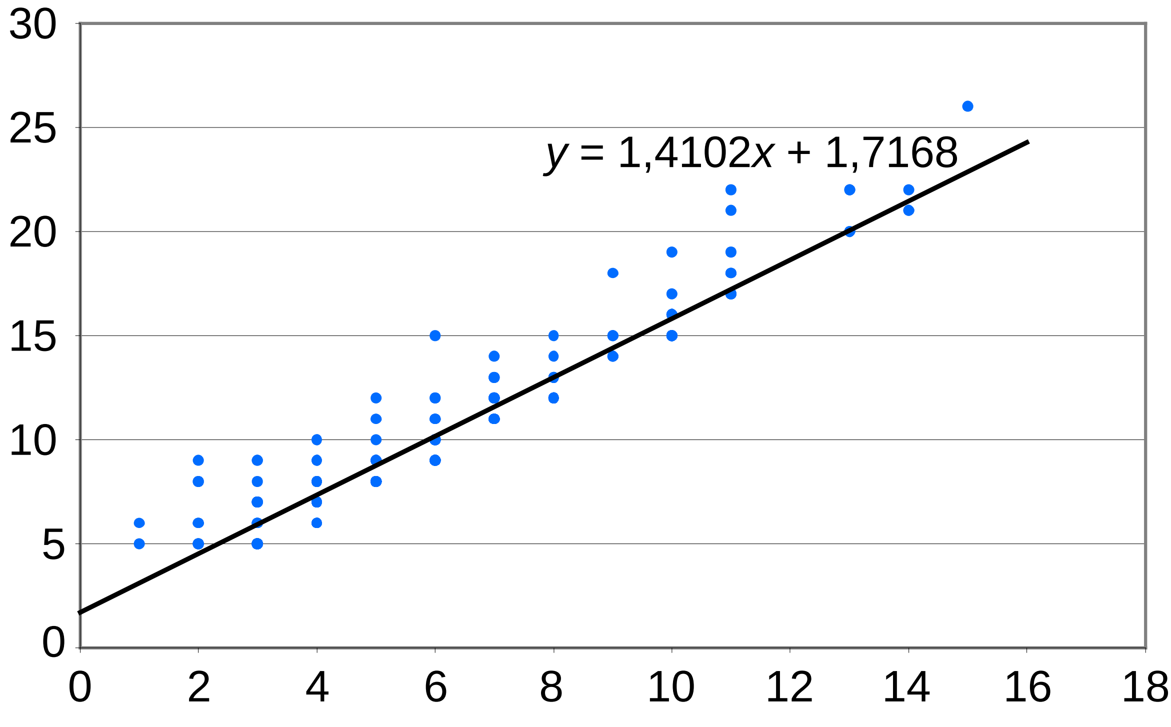
Другое соображение, связанное с этой ситуацией, присутствует у двух авторов присланных эссе. Соображение такое: даже если связь есть, нельзя утверждать, что она причинно-следственная, то есть, что влажность явно определяет количество снега. Может быть, существует третий параметр, который мы не видим, и который влияет и на то, и на другое? А может быть, таких незамеченных параметров много?

4. (2013) Скорость ветра.

Учитель математики Ноибуро Кавасаки из средней школы при университете Цукубы (Токио) увлекается метеорологической статистикой. Со своими школьниками он проводит обработку данных, пользуясь информацией, полученной из архивов метеорологических сайтов, и пытается делать выводы и даже прогнозы.

Однажды Кавасаки-сенсэй решил посмотреть, наблюдается ли линейная зависимость между средней скоростью ветра и скоростью максимальных порывов ветра в одном и том же месте. Для этого он брал данные различных метеостанций, строил диаграммы рассеивания и прямые, наилучшим образом описывающие эти диаграммы. Для построения прямых г-н Кавасаки использовал метод наименьших квадратов.

На рисунке показана одна из таких диаграмм, построенная по данным за несколько недель из архива погоды г.Понта-Делгада (Португалия).



На оси абсцисс откладывается средняя скорость ветра (м/с) за 10-минутный промежуток времени в конце каждого трехчасового периода, на оси ординат — максимальная скорость порыва ветра (м/с) за этот же промежуток.

Затем г-н Кавасаки построил прямую наименьших квадратов, которая служит осью облака рассеивания. Посмотрев на полученное облако и на уравнение прямой, г-н Кавасаки утверждает, что связь, безусловно, есть, и что максимальная скорость ветра в среднем в 1,41 раза превосходит среднюю скорость ветра.

Нет ли ошибки в рассуждениях г-на Ноибуро Кавасаки? Можете ли Вы предложить свои соображения: как подсчитать, во сколько раз максимальная скорость ветра в среднем превосходит среднюю скорость?

Возможная идея исследования

Анализируя данные и действия господина Ноибуро Кавасаки, можно было обратить внимание на то, что коэффициент прямой наименьших квадратов не является средним отношением скоростей. Отношение это равно отношению ординаты и абсциссы точки на диаграмме. Даже наименьшее из этих отношений существенно больше, чем угловой коэффициент прямой. Здесь мы имеем дело с неверным толкованием описательных параметров.

Один из авторов эссе заметил, что количество точек над прямой значительно больше, чем под прямой. Из этого автор сделал вывод, что прямая проведена неверно. Здесь мы должны огорчить автора — дело в том, что число точек над и под прямой оценить по диаграмме нельзя, поскольку случайно совпавшие точки изображаются одной и той же точкой. Прямая проведена верно.

§ 2.5. Анализ результатов решения заданий-эссе

Посмотрим на статистику присланных решений с 2011 года. В таблице приведены количества присланных непустых решений. Подчеркнутые числа отвечают задачам первого типа нашей классификации (проверка какого-либо утверждения), выделенные жирным — самостоятельное статистические исследование, выделенные и подчеркнутые — на поиск статистического метода, остальные задания-эссе на поиск ошибки в рассуждениях.

	1	2	3	Число участников
2011	**5**	5	**0**	24
2012	**7**	**5**	**9**	18
2013	**4**	2	1	7
2014	5	**5**	**3**	16

Как показала практика, задания типа эссе привлекают участников, при этом доля присланных решений, в которых не было замечено ни одной разумной мысли, мала. К сожалению, олимпиада проводится не достаточно долго, чтобы судить о предпочтениях школьников. Хотя задания, основанные на жизненном опыте обычных людей, привлекают куда больший интерес юных исследователей, чем строго сформулированные задачи с диаграммами зависимости влажности воздуха и выпавших осадках.

Постараемся оценить вклад заданий каждого типа в общий массив самостоятельной интеллектуально деятельности участников. Для этого введем некоторый нормированный параметр, назовем его — индекс решаемости. Выясним долю решавших к числу участников отдельно по каждому году и каждому типу заданий, затем, усредним полученные значения по годам. Поскольку, мы хотим получить нормированный показатель, то каждое из четырех полученных чисел разделим на 4.

Заметим, что максимальное значение нашего параметра равно 0,25. Используя статистику решения задач за последние 4 года, получаем:

Тип задания-эссе	Индекс решаемости
Проверка какого-либо утверждения	0,08
Самостоятельное статистическое исследование	0,09
Поиск статистического метода	0,08
Поиск ошибки	0,05

С помощью нашего параметра, можно утверждать, что самостоятельные статистические исследования нравятся участникам больше, хоть и не намного, других, а задачи на поиск ошибки востребованы чуть ли не в 2 раза меньше.

За все годы проведения олимпиады мы ни разу не усомнились в самостоятельности присланных работ. В этом нет ничего удивительного, ведь олимпиада направлена на заинтересованных детей, которым мало школьных знаний, и не является промежуточным фактором для каких-либо дальнейших соревнований.

Мы старались формулировать задания таким образом, чтобы из условия не возникало однозначного направления решения. Чтобы дети могли подумать, пофантазировать, порассуждать самостоятельно, не опираясь на оставленные в тексте зацепки. При этом мы хотим привить и развить умение и способность участников критиковать себя, чужие рассуждения, строить правдоподобные рассуждения и использовать их для решения новых задач. Так же важным мы видим развитие логики, развитие умения анализировать информацию, заинтересованность окружающим миром и самостоятельный поиск закономерностей и случайностей.

§ 3. Задания, предполагающие решения, классификация

Помимо заданий-эссе олимпиада традиционно включает 16-19 традиционных задач по комбинаторике, статистике и теории вероятностей.

Задачи этого раздела олимпиады можно классифицировать следующим образом:

1. **Комбинаторика**. К этому типу задач отнесем задачи, для решения которых достаточно знаний комбинаторных формул, перечисления равновозможных вариантов и выбора из них, удовлетворяющих какому-либо условию.

2. **Простые задачи на вероятность**. Для решения этого типа достаточно иметь интуитивное представление теории вероятностей. Понимать, что такое случайное событие и уметь применять формулы оперирования вероятностью.

3. **Случайная величина**, ее свойства, математическое ожидание, дисперсия и пр.

4. **Сложные задачи на вероятность**. Требуются нетривиальные знания, довольно серьезная техника оперирования с вероятностью, теория меры и интеграла.

Если первые четыре типа заданий можно решить, почерпнув знания преимущественно из учебника по теории вероятностей и статистике Тюрина, то для решения сложных задач этих знаний недостаточно. Такие задания под силу только тем, кто уделит раздумьям и попыткам достаточно много времени, будет изучать литературу. На сайте нашей олимпиады (http://terver.mccme.ru/) мы приводим список рекомендуемой литературы а так же предлагаем вниманию олимпиады прошлых лет полностью с решениями.

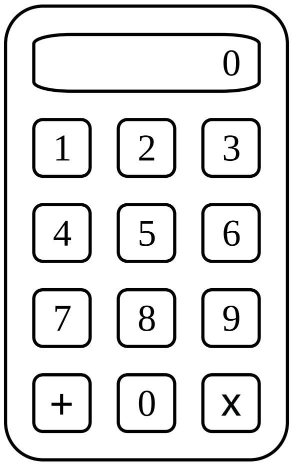
Авторы сознательно включают в олимпиаду задачи, решение которых требует сложных преобразований. Таких задач немного, но они должны быть в заочной олимпиаде, само участие в которой предполагает исследовательскую работу участника.

Со всеми заданиями олимпиады можно ознакомиться на сайте олимпиады http://terver.mccme.ru/. Кроме того, в 2011 году в издательстве МЦНМО вышла книжка, содержащая все задания олимпиад 2008 —2011 годов. Ожидается следующее, более полное издание.

В заключение выражаем убеждение, что пришло время включать вероятностные задачи в классические олимпиады и математические турниры, такие как Турнир Ломоносова, региональный тур всероссийской олимпиады школьников и др.

§ 3.1. Примеры задач с решениями

1. (2010) Калькулятор.



а) На клавиатуре калькулятора есть цифры от 0 до 9 и знаки двух действий (см. рисунок). Вначале на дисплее написано число 0. Можно нажимать любые клавиши. Калькулятор выполняет действия в последовательности нажатий. Если знак действия нажать подряд несколько раз, то калькулятор запомнит только последнее нажатие. Кнопка со знаком умножения <×> сломалась и не работает. Рассеянный Учёный нажал несколько кнопок в случайной последовательности. Какой результат получившейся цепочки действий более вероятен — четное число или нечетное?

б) Решите предыдущую задачу, если кнопку со знаком умножения починили.

Решение. а) Заметим, что хотя бы одно действие сложения выполнено, даже если Учёный набрал только одно число – тем самым он прибавил это число к нулю.

Обозначим A событие «результат нечетный». Четность результата определяется последним слагаемым. Поясним это подробнее.

Пусть с вероятностью p предпоследнее число нечетное. Тогда результат останется нечетным, только если прибавить четное число. Значит, вероятность нечетного результата равна $p\cdot\frac{1}{2}$.

Предпоследнее число было четным с вероятностью $1-p$. Результат будет нечетным, только если последнее слагаемое нечетно. Значит, в этом случае событие A имеет вероятность $(1-p)\cdot\frac{1}{2}$.

Складывая вероятности этих несовместных событий, получим:

$$\mathrm{P}(A) = p \cdot \frac{1}{2} + (1-p) \cdot \frac{1}{2} = \frac{1}{2}.$$

Таким образом, если разрешено только сложение, четный и нечетный результаты имеют равные шансы.

б) Рассмотрим последнее сложение в цепочке действий. После него получился либо четный, либо нечетный результат, причем вероятности этих событий равны, как мы видели при решении задачи а).

Если цепочка оканчивается ровно одним умножением, то оно дает нечетный результат, только если оба последних числа нечетные. Вероятность $\frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$.

Если цепочка оканчивается двумя умножениями, то нечетный результат получится, только если все три сомножителя нечетные. Вероятность $\frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8}$ и так далее. Если цепочка оканчивается k действиями умножения, то при этом условии вероятность нечетного результата равна $\frac{1}{2^{k+1}}$.

Если обозначить p_k вероятность того, что цепочка действий оканчивается ровно k действиями умножения, то

$$\mathrm{P}(A) = p_0 \cdot \frac{1}{2} + p_1 \cdot \frac{1}{4} + p_2 \cdot \frac{1}{8} + \ldots + p_k \cdot \frac{1}{2^{k+1}} + \ldots + p_n \cdot \frac{1}{2^{n+1}},$$

где n — общее число действий в цепочке.

Вынесем $\frac{1}{2}$ за скобки:

$$\mathrm{P}(A) = \frac{1}{2}\left(p_0 + p_1 \cdot \frac{1}{2} + p_2 \cdot \frac{1}{4} + \ldots + p_n \cdot \frac{1}{2^n}\right) < \frac{1}{2}(p_0 + p_1 + \ldots + p_n) = \frac{1}{2} \cdot 1 = \frac{1}{2},$$

поскольку сумма всех p_k равна единице — ведь каким–то образом цепочка из n действий все же оканчивается.

Значит, если разрешено умножение, четный результат имеет больше шансов. Очевидно, что вероятность четного результата намного выше, чем нечетного, если умножений в конце много — оценка сверху, которую мы сделали при переходе к неравенству, довольно грубая.

2. (2012) Автобус.

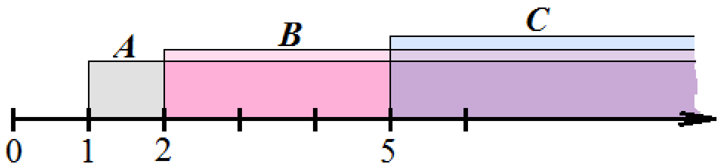
Аня ждёт автобус. Какое событие имеет наибольшую вероятность?

A = {Аня ждет автобуса не меньше минуты}

B = {Аня ждет автобуса не меньше двух минут.}

C = {Аня ждет автобуса не меньше пяти минут.}

Решение. Расположим события на временной оси.



На рисунке видно, что событие A включает в себя событие B, а событие B включает в себя событие C, то есть событие A наиболее обширное:

$$C \subset B \subset A.$$

Следовательно,

$$\mathrm{P}(C) \le \mathrm{P}(B) \le \mathrm{P}(A).$$

Ответ: Наиболее вероятно событие A.

3. (2013) Оплата электричества.

На рисунке показано платежное поручение на оплату электричества некоторой энергосбытовой компании.

Получатель платежа ОАО XXX

Код РР

Период: Март 2013 г.

Номер лицевого счета XXX

Ф.И.О.: XXX

Адрес: XXX

Код платежа	Тарифная зона	Показания счетчика		Расход факт.	Тариф (руб.)	Сумма к оплате (руб.)
		Текущее	Предыдущее			
13	***пик (Т1)***	1347	1270	77.0	**4,03**	310,31
2	***ночь (Т2)***	1402	1337	65.0	**1,01**	65,65
15	***п/пик (Т3)***	1298	1214	84.0	**3,39**	284,76

Подпись абонента: ***Итого к оплате:*** **660** *р.* **72** *к.*

Каждый месяц клиент передаёт компании показания трёхтарифного счётчика, установленного в квартире. Из показаний за текущий месяц вычитаются соответствующие показания за прошлый месяц, получается фактический расход за месяц по каждой из трёх тарифных зон (пик, ночь, полупик). Затем расход по каждой зоне умножается на цену одного киловатт-часа в этой зоне. Складывая полученные суммы, клиент получает общую сумму оплаты за месяц. В данном примере клиент заплатит 660 р.72 коп.

Компания ведет учёт расхода и оплаты электроэнергии, пользуясь данными, полученными от клиента. Проблема состоит в том, что компания иногда путает полученные шесть чисел, переставляя их произвольном порядке, правда, следит за тем, чтобы текущее показание оставалось больше, чем предыдущее. В результате расчёт компании может оказаться ошибочным. Если компания считает, что клиент должен больше, чем он заплатил, компания требует доплатить разность.

Пользуясь данными изображенной квитанции, найдите:

а) максимально возможную сумму доплаты за март 2013 года, которую компания потребует у клиента;

б) математическое ожидание разности между суммой, которую насчитает компания, и суммой, которую заплатил клиент.

Решение. а) Очевидно (несложно показать), что сумма, которую потребует компания, будет наибольшей, если расход по самому высокому тарифу будет наибольшим возможным, по среднему — наибольшим возможным из оставшихся. Наибольшая возможная сумма при этом 1058 р. 06 коп. Компания потребует доплатить

1058 р. 6 коп. – 660 р.72 коп. = 397 р. 34 коп.

б) Решим задачу в общем виде, считая, что клиент передал шесть различных чисел $a<b<c<d<e<f$. Рассмотрим одну из тарифных зон (например, пик, для определенности). Для этой зоны можно выбрать два любых числа из данных шести и поставить их в соответствующем порядке — по возрастанию. Это можно сделать $C_6^2=15$ способами. По условию все способы равновозможны. Поэтому математическое ожидание случайной величины X_1 «Расход по тарифу «Пик» есть среднее арифметическое пятнадцати чисел:

$$\begin{aligned}15\mathrm{E}\,X_1=(f-e)+(f-d)+(f-c)+(f-b)+(f-a)&+\\+(e-d)+(e-c)+(e-b)+(e-a)&+\\+(d-c)+(d-b)+(d-a)&+\\+(c-b)+(c-a)&+\\+(b-a)&.\end{aligned}$$

Приведём подобные:

$$\mathrm{E}\,X_1=\frac{5f+3e+d-c-3b-5a}{15}.$$

Очевидно, такое же ожидание имеют случайные величины X_2 и X_3 «Расход по тарифу «Ночь» и «Расход по тарифу «Полупик». Полная сумма оплаты S, рассчитанная компанией, равна

$$S = t_1 X_1 + t_2 X_2 + t_3 X_3,$$

где t_1, t_2 и t_3 — цены одного киловатт-часа по соответствующим тарифам. Следовательно,

$$\mathrm{E}\,S = t_1 \cdot \mathrm{E}\,X_1 + t_2 \cdot \mathrm{E}\,X_2 + t_3 \cdot \mathrm{E}\,X_3 = \frac{5f + 3e + d - c - 3b - 5a}{15} \cdot \left(t_1 + t_2 + t_3\right),$$

Определим значения переменных в соответствии с данными задачи:

$$f = 1402;\ e = 1347;\ d = 1337;\ c = 1298;\ b = 1270;\ a = 1214;$$

$$t_1 + t_2 + t_3 = 4,03 + 1,01 + 3,39 = 8,43.$$

Получаем:

$$\mathrm{E}\,S = \frac{5 \cdot 1402 + 3 \cdot 1347 + 1337 - 1298 - 3 \cdot 1270 - 5 \cdot 1214}{15} \cdot 8,43 =$$

$$= \frac{1210}{15} \cdot 8,43 = 680,02.$$

Математическое ожидание разности между суммой, которую потребует сбытовая компания и суммой, которую заплатит клиент, равно

$\mathrm{E}\,S - 660,72 = 680,02 - 660,72 = 19,03$ (руб.)

Ответ: а) 397 руб. 34 коп.; **б)** 19руб. 3 коп.

4. (2014) Чаепитие Федоры Егоровны.

У Федоры Егоровны есть 3 чайные чашки. Однажды она их вымыла и начала новую жизнь. Каждый раз, чтобы выпить чаю, Федора берёт наудачу какую-нибудь чашку, а потом ставит её к остальным, но уже не моет. Найдите вероятность того, что во время пятого чаепития с момента начала новой жизни Федора Егоровна использует последнюю чистую чашку.

Решение. Указанное событие наступит, только если четыре первых раза Федора использовала какие-то две чашки (причем обе хотя бы по разу), а в пятый раз – последнюю чистую чашку. Для определенности пронумеруем чашки. Тогда существует 16 способов четыре раза использовать чашки 1 и 2. При этом два из них не подходят – это способы 1111 и 2222. Значит, всего 14 способов использовать четыре раза ровно две чашки (каждую хотя бы раз). Точно так же, 14 способов четыре раза выпить чаю из чашек 1 и 3 и еще 14 способов использовать чашки 2 и 3. Значит, всего 42 способа использовать ровно две чашки за четыре чаепития и третью чашку – в пятый раз. Общее число способов использовать три чашки пять раз равно $3^5 = 243$. Искомая вероятность равна $\frac{42}{243} = \frac{14}{81}$.

Ответ: $\frac{14}{81}$.

5. (2012) Первая пара.

На сушке в случайном порядке (как достали из стиральной машины) висит n пар носков. Двух одинаковых пар нет. Носки висят за сохнущей простыней, поэтому Рассеянный Ученый достает по одному носку на ощупь и сравнивает каждый новый носок со всеми предыдущими. Найдите математическое ожидание числа носков, снятых к моменту, когда у Учёного окажется какая-нибудь пара.

Решение. Назовем ξ_n случайную величину, равную числу снятых носков при условии, что на сушке висит n пар. Очевидно, $\mathrm{E}\xi_1 = 2$. Пусть теперь $n > 1$. Попробуем установить рекуррентную связь между ξ_n и ξ_{n-1}. Пронумеруем носки в том порядке, в каком Рассеянный Ученый их снимает с сушки. Очевидно, способ нумерации не играет роли. В этой нумерации какой-то носок является последним (его номер $2n$, и он никогда не будет снят, поскольку носок с номером $n+1$ наверняка образует пару с каким-либо предыдущим). Отличим как-нибудь этот последний носок. Пусть он будет белым.

Обозначим $p_{j,n}$ вероятность события $\xi_n = j$. Очевидно, $p_{1,n} = 0$.

Найдем $p_{2,n}$. Все зависит от того, где висит другой белый носок. Если он висит на первом месте (вероятность этого $\frac{1}{2n-1}$), то второй носок не может образовать с ним пару. В этом случае $p_{2,n} = 0$.

Если другой белый носок висит на втором месте (вероятность этого $\frac{1}{2n-1}$), то же самое – $p_{2,n} = 0$.

Если он висит на третьем или последующих местах (вероятность этого равна $\frac{2n-3}{2n-1}$), то $p_{2,n} = p_{2,\,n-1}$ – это вероятность получить пару со второй попытки, выбрасывая из последовательности оба белых носка. По формуле полной вероятности получаем:

$$p_{2,\,n} = \frac{2n-3}{2n-1} p_{2,\,n-1} \;.$$

Будем рассуждать так же относительно $p_{3,\,n}$.

1. Если белый носок висит на одном из двух первых мест (с вероятностью $\frac{2}{2n-1}$), то вероятность того, что третий носок даст пару, равна вероятности того, что второй носок даст пару в последовательности без белых носков: $p_{3,\,n} = p_{2,\,n-1}$.

2. Если белый носок на третьем месте (вероятность $\frac{1}{2n-1}$), то вероятность того, что третий носок даст пару, равна 0, поскольку он белый, а парный к нему висит в конце и потому недоступен.

3. Если белый носок висит далее третьего места (вероятность $\frac{2n-4}{2n-1}$), то вероятность того, что третий носок даст пару, равна $p_{3,\,n-1}$, поскольку в этом случае наличие белых носков роли не играет. Формула полной вероятности дает:

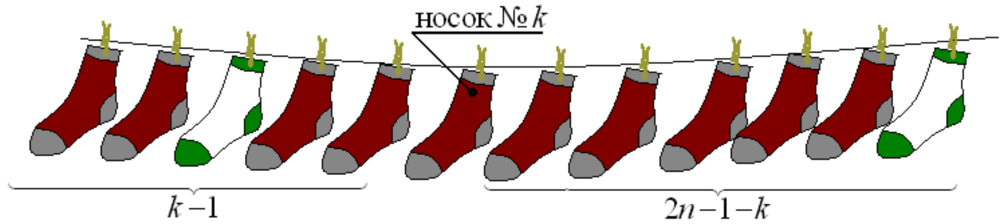
$$p_{3,\,n} = \frac{2}{2n-1} p_{2,\,n-1} + \frac{2n-4}{2n-1} p_{3,\,n-1}$$

Рассуждая аналогично и дальше, получаем:

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

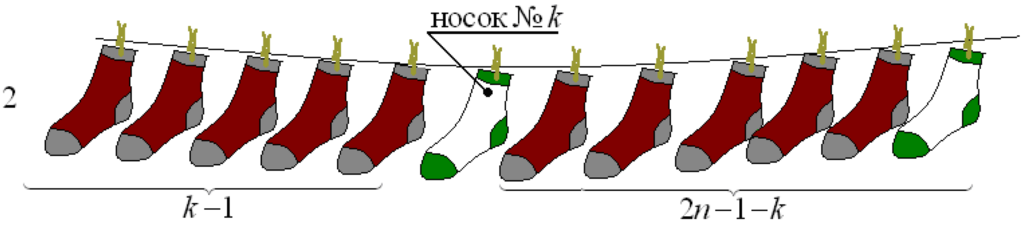
для всех k от 2 до n.

1



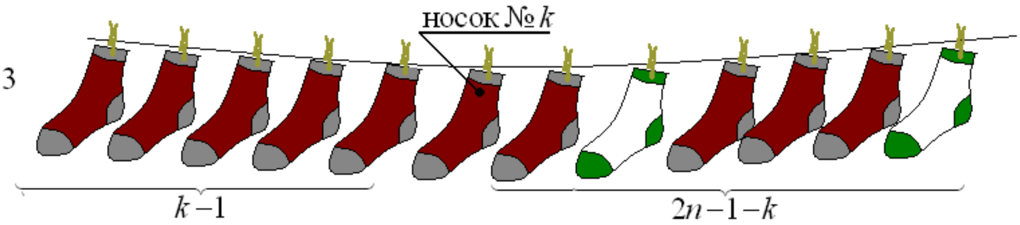
Здесь k-й носок имеет номер $k-1$ в последовательности, из которой удалены белые носки. Поэтому в этом случае $p_{k,n}=p_{k-1,n-1}$.

2



Здесь k-й носок белый,а поэтому он не может образовать пару ни с каким из предыдущих: $p_{k,n}=0$.

3



Здесь белый носок висит дальше, чем k-й, а поэтому белые можно не рассматривать вообще: $p_{k,n}=p_{k,n-1}$.

Если всего пар $n-1$, то вероятность того, что первая пара получится на $n+1$ носке, равна нулю: $p_{n+1,\,n-1}=0$. Поэтому для $k=n+1$ получаем:

$$p_{n+1,\,n}=\frac{n}{2n-1}\,p_{n,\,n-1},$$

Найдем теперь математическое ожидание:

$$\mathrm{E}\xi_n = 2p_{2,n}+3p_{3,n}+4p_{5,n}+\ldots+np_{n,n}+(n+1)\,p_{n+1,n}=$$

$$=2\cdot\frac{2n-3}{2n-1}p_{2,\,n-1}+3\cdot\frac{2}{2n-1}p_{2,\,n-1}+3\frac{2n-4}{2n-1}p_{3,\,n-1}+$$

$$+4\cdot\frac{3}{2n-1}p_{3,\,n-1}+4\frac{2n-5}{2n-1}p_{4,\,n-1}+\ldots$$

$$\ldots+n\cdot\frac{n-1}{2n-1}p_{n-1,\,n-1}+n\frac{n-1}{2n-1}p_{n,\,n-1}+(n+1)\frac{n}{2n-1}p_{n,\,n-1}.$$

Группируя слагаемые при $p_{k,\,n-1}$, получаем:

$$\mathrm{E}\xi_n = \frac{4n}{2n-1}p_{2,\,n-1} + \frac{6n}{2n-1}p_{3,\,n-1} + \frac{8n}{2n-1}p_{4,\,n-1} + \ldots + \frac{2n^2}{2n-1}p_{n,\,n-1} =$$
$$= \frac{2n}{2n-1}\left(2p_{2,n-1} + 3p_{3,n-1} + 4p_{4,n-1} + \ldots + np_{n,n-1}\right) = \frac{2n}{2n-1}\cdot \mathrm{E}\xi_{n-1}.$$

Полученное соотношение

$$\mathrm{E}\xi_n = \frac{2n}{2n-1}\cdot \mathrm{E}\xi_{n-1}$$

дает возможность вычислить $\mathrm{E}\xi_n$, зная, что $\mathrm{E}\xi_1 = 2$:

$$\mathrm{E}\xi_n = \frac{(2n)\cdot(2n-2)\cdot\ldots\cdot 4}{(2n-1)\cdot(2n-3)\cdot\ldots\cdot 3}\cdot \mathrm{E}\xi_1 = \frac{(2n)\cdot(2n-2)\cdot\ldots\cdot 4\cdot 2}{(2n-1)\cdot(2n-3)\cdot\ldots\cdot 3\cdot 1} = \frac{(2n)!!}{(2n-1)!!}.$$

Этот результат можно представить иначе:

$$\mathrm{E}\xi_n = \frac{\left((2n)!!\right)^2}{(2n-1)!!\cdot(2n)!!} = \frac{\left(2^n n!\right)^2}{(2n)!} = \frac{4^n (n!)^2}{(2n)!} = \frac{4^n}{C_{2n}^n}.$$

Для получения приближения при больших n воспользуемся формулой Стирлинга: $n! \approx \sqrt{2\pi n}\cdot n^n e^{-n}$. Получаем:

$$\mathrm{E}\xi_n = \frac{4^n (n!)^2}{(2n)!} \approx \frac{4^n 2\pi n\cdot n^{2n}\cdot e^{-2n}}{e^{-2n}\sqrt{2\pi\cdot 2n}\,(2n)^{2n}} = \sqrt{\pi n}.$$

Согласие можно продемонстрировать при большом n. Скажем, если на сушке висит 2012 пар носков, то точная формула дает

$$\mathrm{E}\xi_{2012} = 79{,}509,$$

а приближенная даст число

$$\mathrm{E}\xi_{2012} \approx \sqrt{2012\pi} = 79{,}504.$$

Погрешность менее $0{,}065\,\%$. Таким образом, если на веревке 4024 носка, то Рассеянному Учёному в среднем повезет уже на семьдесят девятой – восьмидесятой попытке.

§ 3.3. Анализ результатов решения задач

Посмотрим на статистику присланных решений с 2011 года. В таблице указано среднее по числу участников, приславших верное или частично верное решение, и по количеству задач данного типа. Сверху указан тип задач, согласно нашей классификации.

	1	2	3	4	Число участников
2011	7	5,83	2,17	0,67	24
2012	2,09	2,83	0,33	0	18
2013	4,4	2,38	0,29	2,86	7
2014	6,78	5	2,86	1,67	16

Как видно из таблицы, школьники чаще всего берутся за более простые задачи, для решения которых достаточно иметь интуитивное представление о теории вероятностей, владеть приемами комбинаторики, и уметь грамотно оперировать понятием о случайном событии и операциями над множеством случайных событий.

Найдем и тут вклад каждого типа задач в общий массив самостоятельной интеллектуальной деятельности участников.

Тип задач	Индекс решаемости
Комбинаторика	0,12
Простые задачи на вероятность	0,09
Случайная величина	0,03
Сложные задачи на вероятность	0,04

Вспомним, что вклады заданий-эссе попадают в диапазон от 0,05 до 0,09. Значит, они вызывают у участников олимпиады больший интерес, чем сложные вычислительные задачи. В то же время очень радуют ненулевые показатели действительно непростых задач. Мы получаем грамотные решения далеко не тривиальных задач от 6-ти и 7-миклассников, при этом не возникает сомнений, что ребенок справился сам.

Пусть наша заочная олимпиада проводится не так долго и количество участников несравнимо с олимпиадами более серьезных масштабов, но можно с уверенность сказать, что она нашла свою целевую аудиторию и показала очень хорошие результаты. Мы считает создание такого соревнования успешным и планируем в последующие годы развернуть масштабную рекламу и, возможно, добьемся присвоения более серьезного статуса.

Заключение

Как уже отмечалось выше, введение курса теории вероятностей и статистики в школьную программу крайне необходимо, однако не стоит забывать, что одного школьного учебника не достаточно для популяризации данной дисциплины. На помощь приходит заочная олимпиада, дающая детям возможность задуматься над вопросами, не вошедшими в школьную программу.

На данный момент наша олимпиада не имеет такого масштаба, как традиционные олимпиады по математике, информатике, физике и т.д. За последние 5 лет на олимпиаду регистрировалось от 50 до 100 школьников ежегодно, и только 10-15% из них присылали свои решения. По-видимому, это связано с тем, что отсутствие официального статуса олимпиады не дает льгот абитуриентам. В дальнейшем мы планируем привлекать к олимпиаде существенно большее количество участников.

Опыт проведения олимпиады показал, что даже при отсутствии информационной компании находятся школьники, которые проявляют к ней интерес. Что, в свою очередь, ведет к повышению квалификации их преподавателей и улучшению ситуации в целом.

Литература

1. Е.А.Бунимович, В.А.Булычев, И.Р.Высоцкий и др., О теории вероятностей и статистике в школьном курсе, Математика в школе, №7, Москва, Школьная пресса, 2009, с. 3-13.
2. Ю.Н.Тюрин, А.А.Макаров, И.Р.Высоцкий, И.В.Ященко, Преподавание теории вероятностей и статистики в школе по учебному пособию Ю.Н. Тюрина и др. «Теория вероятностей и статистика», Математика в школе, №7, Москва, Школьная пресса, 2009, с. 14-31.
3. И.Р. Высоцкий, Олимпиадные задачи по ТВ, Первое сентября, Январь 2012, с. 61-62.
4. Е.А.Бунимович, И.Р.Высоцкий и др., Терминология, обозначения и соглашения в школьном курсе теории вероятностей и статистики, Издательский дом Первое Сентября, Математика, №17, 2009, с. 13-27.
5. И.Р.Высоцкий, В.В.Бородкина и др., Заочная олимпиада по теории вероятностей и математической статистике, Издательский дом Первое Сентября, Математика, №15, 2009, с. 17-25
6. Ю.Н.Тюрин, И.Р.Высоцкий, Методические подходы к преподаванию теории вероятностей и статистики, Издательский дом Первое Сентября, Математика, №10, 2010, с. 8-19.
7. Ю.Н.Тюрин, А.А.Макаров, И.Р.Высоцкий, И.В.Ященко, Теория вероятностей и статистика, Издательство МЦНМО, 2-е издание, 2008.
8. А.Н. Колмогоров, И.Г. Журбенко, А.В. Прохоров, Введение в теорию вероятностей, Москва, Наука, 1982.
9. А. Мостеллер, Пятьдесят занимательных вероятностных задач с решениями, Москва, Наука, Физматлит, 3-е издание, 1985.
10. И.Р.Высоцкий, П.И.Захаров, В.В.Нестерова, И.В.Ященко, Теория вероятностей и статистика, Задачи заочных Интернет-олимпиад, Москва, Издательство МЦНМО, 2011.

11. Ю.Н.Благовещенский, Тайны корреляционных связей в статистике, Москва, «Научная книга», 2009.
12. К.Л. Чжун, Ф. Аит-Сахлиа, Элементарный курс теории вероятностей. Стохастические процессы и финансовая математика, Москва, БИНОМ. Лаборатория знаний, 2007.
13. Математика: Учебники для 5–6 классов общеобразовательных учреждений. Под редакцией Г.В.Дорофеева, И.Ф. Шарыгина, – М.: Просвещение, Дрофа, 2000–2009. Алгебра: Учебники для 7–9 классов общеобразовательных учреждений. Под редакцией Г.В.Дорофеева, – М.: Просвещение, 2000–2009. Алгебра и начала анализа: Учебники для 10–11 классов общеобразовательных учреждений./ Г.В.Дорофеев, Л.В.Кузнецова, Е.А.Седова, – М.: Просвещение, Дрофа, 2000–2009.
14. Виленкин Н.Я. Алгебра и математический анализ. 11 класс – М.: Мнемозина 2003–2009
15. Зубарева И.И., Мордкович А.Г. Математика: Учебники для 5–6 классов общеобразовательных учреждений. – М.: Мнемозина, 2002–2009.
16. Бунимович Е.А., Булычев В.А. Основы статистики и вероятность. 5–11 кл.: учебное пособие – М.: Дрофа, 2008.
17. Бунимович Е.А., Булычев В.А. Вероятность и статистика. 5–9 кл. Электронное учебное пособие. – М.: Дрофа, 2007.
18. Дорофеев Г.В. Математика 5–9. учебное пособие / Е.А.Бунимович, В.А.Булычев. Просвещение. – 2009.
19. Макарычев Ю.Н., Миндюк Н.Г. Алгебра: элементы статистики и теории вероятностей: учеб. Пособие для учащихся 7–9 кл. общеобразоват. учреждений / Ю.Н.Макарычев, Н.Г.Миндюк; под ред. С.А.Теляковского. – 3-е изд. – М.: Просвещение, 2005.

20. Ткачева М.В., Федорова Н.Е. Элементы статистики и вероятность: Учеб. Пособие для 7–9 кл. общеобразоват. Учреждений / М.В.Ткачева, Н.Е.Федорова. – М. Просвещение, 2004.
21. Бродский И.Л., Литвиненко Р.А. Вероятность и статистика 7–9 классы. Решение задач из учебников под ред. Г.В.Дорофеева. – М.: АРКТИ, 2006.
22. И.Р.Высоцкий, И.В.Ященко, Типичные ошибки в преподавании теории вероятностей и статистики, Математика в школе, №5, Москва, Школьная пресса, 2014, с. 30-41.

Выпускная квалификационная работа «Организация заочных интеллектуальных соревнований по теории вероятностей и статистике для школьников».

Выпускная квалификационная работа выполнена мной совершенно самостоятельно. Все использованные в работе материалы и концепции из опубликованной научной литературы и других источников имеют ссылки на них.

11.06.2014 Заплетина О.М.