Типичные ошибки в преподавании теории вероятностей и статистики

Цель статьи – помочь учителю сориентироваться в содержании и преподавании теории вероятностей и статистики основной и старшей школы на основе анализа типичных методических ошибок. Использованы материалы уроков, методических статей и руководств, а также опыт работы курсов повышения квалификации учителей, проводимых в МЦНМО и МИОО с 2005 года.

ВВЕДЕНИЕ

«Что мы хотим от детей?»

Недавно один из авторов этой статьи имел любопытную беседу с уважаемым коллегой, учителем и популяризатором математики, автором хороших книжек. Коллега жаловался на составителей московской контрольной работы по теории вероятностей в восьмом классе, которые дали в работе вопрос, можно ли считать некоторое (описанное) событие маловероятным?. Как ответить? Нет определенного ответа! Что составители хотят от детей? Разве можно спрашивать такое в математике? Математическая задача должна иметь один-единственный определенный верный ответ!

Уважаемый коллега относится к теории вероятностей как к набору фактов и формул, в которые дети должны подставлять известные числа, чтобы получить результат. Такой подход мы видим в преподавании анализа: зная алгоритм, школьник успешно решает задачи, не понимая смысла своих действий. Это проявляется в результатах диагностических работ, когда экстремум функции находит больше школьников, чем отвечает на вопрос о значении производной по рисунку на клетчатой бумаге.

Часто от ученых-математиков мы слышим, что вероятность и статистика сложны для школьников, и поэтому вообще не нужны в школе. Такое мнение складывается из-за трудностей, которыми сопровождается изучение теории вероятностей в вузах, традиционно дедуктивное, опирающееся на комбинаторику и обширные сведения из математического анализа. Механистический подход к теории вероятностей характерен для многих поколений нашей высшей школы. Будет обидно, если школьная теория вероятностей пойдет по этому пути. Заставляя школьников считать по формулам, не задумываясь над смыслом явления и результата, мы непременно снова придем в тупик, как приходили уже дважды в начале и в середине XX века.

На самом деле ни комбинаторика, ни анализ напрямую не связаны с базовыми идеями статистики и теории вероятностей. Они служат лишь средством перечисления элементов обширных вероятностных пространств и доказательства теорем. Но ни то, ни другое не является первостепенным.

Теория вероятностей вторична, как и всякая строгая теория. Первичны опыт и интуиция, без которых не нужно заниматься теорией. Важно умение работать с базовыми понятиями без формул, опираясь на здравый смысл. Требуется осмысленное отношение к шансам событий (особенно к маловероятным событиям, играющим большую роль в жизни).

Чтобы пояснить сказанное, проведем аналогию между вероятностью и другими разделами математической дидактики. Для этого придется заглянуть в начальную школу и в дошкольный период.

Маленький ребенок знакомится с количеством и числом и понимает, что с ними можно делать что-то полезное (например, посчитать сдачу в магазине или время в пути). Позже число формализуется в курсе арифметики и алгебры. Если же интуитивного понятия о числе нет, то и алгебры не будет.

Свойства рисунка (длины, углы, соединения линий) формализуются геометрией. Если школьник в детстве не рисовал человечков, солнышко и дым из трубы, то геометрия для него превращается в абстрактный набор фактов, которые плохо подкреплены опытом. **Любая математическая теория формализует привычные наглядные или мыслимые образы. Нет образов – нет теории.**

Здесь следует сказать о взаимосвязи статистики и теории вероятностей. Развивая проведенную аналогию, заметим, что статистика относится к теории вероятностей в некотором смысле так же, как черчение относится к геометрии. А именно: **статистика дает методы численного описания массовых явлений, а теория вероятностей предоставляет для этого теоретическую основу**, подобно тому, как геометрия предоставляет теоретическую базу любому профессиональному изображению от наброска карандашом до компьютерной графики.

Из этого вытекает важный методический **принцип первичности статистики**. Только вдумавшись в природу случайности и научившись описывать простейшие случайные явления минимальным набором числовых средств, учащийся с помощью учителя может ставить перед собой вопрос о теоретических инструментах изучения случайности. Для массового обучения характерен индуктивный подход: переход от конкретности и наглядности к абстракции и обобщению. Индуктивность проявляется в том, что сначала школьника следует познакомить с частотой явления, представлением и описанием числовых данных, и только позже – с их теоретическими аналогами: вероятностью и случайной величиной.

Главная проблема в том, что события менее наглядны, чем фигуры, числа или выражения, а шансы и изменчивость не так интуитивны, как длина, площадь или объем. Событие и его шансы – особые типы мыслимых объектов, формализация которых в математические происходит значительно сложнее, чем формализация рисунка (переход к геометрии) или количества (к арифметике или алгебре).

Вторая проблема состоит в том, что большинству детей до определенного возраста чуждо представление об изменчивости и неустойчивости явлений. Вопрос о том, в каком возрасте ребенок готов к восприятию изменчивых моделей, и какими эти модели должны быть, предстоит изучать. В советский период наука о законах познавательной деятельности, была уничтожена [5]. Ее место заняла парадигма чистого листа бумаги.

Но если в раннем детстве отторжение изменчивости, вероятно, служит защитной функцией, которая упрощает адаптацию к социальным и природным условиям, минимизируя число необходимых правил поведения, то позже отсутствие общих представлений о случайности и изменчивости мешает ориентированию человека в мире, полном неопределенностей. Формальное изучение статистики в зрелом возрасте не улучшает положения [6].

Поэтому концептуальные искажения восприятия стохастических явлений характерны не только для школьников, но и для взрослых. Пытаясь сохранить жесткую алгоритмичность и простоту преподавания, учителя порождают ошибки и целые ошибочные концепции в преподавании статистики и вероятности (для краткости, говоря о теории вероятностей в школе, часто будем опускать слово «теория»). Накопленный опыт позволяет выделить наиболее типичные заблуждения и ошибки. Ниже мы анализируем их и предлагаем приемлемые методические решения возникающих проблем.

ЦЕЛИ ИЗУЧЕНИЯ ТЕОРИИ ВЕРОЯТНОСТЕЙ И СТАТИСТИКИ В ШКОЛЕ

Цели изучения математики для разных целевых групп школьников разные. Это же касается вероятности и статистики.

На минимальном и базовом уровне мы видим необходимость вооружить школьника общими знаниями и представлениями, позволяющими ему ориентироваться в статистической информации, предоставляемой СМИ. Вторая важная цель – выработка верного представления о законе больших чисел и его роли в статистическом понимании изменчивых явлений.

На профильном уровне эти цели дополняются необходимостью познакомить школьников с основами вероятностного моделирования в упрощенных случайных экспериментах, которые нам представляет практика массовых общественных и экономических явлений. Помимо этого на профильном и частично на базовом уровне возникает возможность использовать школьный аппарат алгебры, геометрии и анализа для поддержки и теоретического обоснования вероятностных фактов. То есть вероятность мотивирует изучение ряда традиционных разделов школьной математики, таких как элементы математического анализа, прогрессии и др.

РОЛЬ И МЕСТО ВЕРОЯТНОСТИ И СТАТИСТИКИ В СИСТЕМЕ МАТЕМАТИЧЕСКИХ ЗНАНИЙ ШКОЛЬНИКОВ

Теория вероятностей и статистика вводятся в школьную математику не только в России, но и в других странах, что говорит о росте значимости базовых стохастических знаний в современном обществе. Говоря о значимости, заметим лишь, что с 2003 года в международное исследование PISA включаются вопросы на применение вероятностных и статистических знаний. Согласно данным центра оценки качества образования (www.centeroko.ru), российские школьники при выполнении заданий этого раздела показывают низкие результаты [4].

В [1] отмечается, что школьное образование в России традиционно направлено на выработку представлений о жестких связях между явлениями окружающего мира (законах). Школьная математика при этом остается консервативным предметом, ориентированным на дальнейшее математическое образование, нужное далеко не всем. Негативные последствия оторванности школьной математики от запросов общества ощущаются из года в год все сильнее. Математика не востребована. Многие старшеклассники и их родители не видят в изучении математики пользы.

С другой стороны, статистика и вероятность, безусловно, доставляют ясные приложения математики к повседневным изменчивым явлениям, чего не может делать алгебра или геометрия. Элементарная статистическая культура является составной частью общественной культуры, значительно более важной, чем умение строить и исследовать жесткие модели процессов, решать уравнения или неравенства.

Следовательно, задача ученых, методистов и учителей – не отказываться от статистики и вероятности в школе, а освободить их от лишних технических наслоений и создать ясный школьный предмет, который отличается от современной статистики и вероятности так же, как школьная физика отличается от современной теоретической физики. Школьное содержание должно быть вплотную приближено к вопросам страхования, торговли, банковских услуг, медицинского обслуживания, управления и принятия решений, к погрешностям измерений любого рода – к явлениям, известным школьникам или их родителям из повседневного опыта, явлениям, значимость которых не вызывает сомнений.

Как отмечалось, следует делать акцент не на доказательствах и вычислениях, а на ясном понимании учащимися концепций изменчивости, средних, рассеивания, выборочных исследований, случайных величин, закона больших чисел. Многие понятия и факты становятся доступны школьникам через эксперимент, а не с помощью дедуктивных выводов и формул. Ради формирования системы базовых понятий и интуиции нужно, чтобы школьники на уроках статистики и вероятности размышляли, обсуждали и спорили. Поэтому **мы считаем принципиально важными задачи без однозначного ответа задания и ситуации, которые провоцируют домысливание и субъективизацию. Другой принцип – как можно больше значимых для школьника сюжетов.** С 2005 года в Москве и в ряде других регионов внедряется практико-ориентированный подход с акцентом на понимании вероятностных ситуаций и описании изменчивости, а не на количественных отношениях между вероятностями. Такой же подход в настоящее время характерен для составителей заданий по вероятности и статистике для ГИА и ЕГЭ.

В первую очередь нужно позаботиться о том, чтобы базовые концепции статистики и вероятности были понятны и привычны для учителей математики, которым трудно перестроиться с преподавания абстрактных фактов на применение математики при обсуждении практических задач. Мы видим это по реакции учителей и школьников на появление в ЕГЭ практико-ориентированных задач на расчет сдачи, выбор альтернативы, расчет времени поездки, которые оказались субъективно сложнее, чем преобразования логарифмических выражений.

Трудности в преподавании вероятности и статистики вызваны не столько недостатком знаний у учителей, сколько формальным и часто бесцельным преподаванием вероятности в вузах, а также скудностью и низким качеством методических пособий. Все это является прямым следствием отсутствия школьных традиций преподавания вероятности и статистики. По мере развития методики и накопления опыта эти проблемы постепенно уйдут, уступив место следующим. Важнейшим периодом становления нового школьного предмета следует считать 8 – 9 лет. По истечении этого периода в школу приходят молодые учителя, начавшие изучать данный предмет в средних классах и поэтому не рассматривающие данный предмет как новый или слишком трудный.

Проблемы с преподаванием статистики и теории вероятностей характерны не только для России. Во многих странах (США, Великобритания, Швейцария, Южная Корея, Япония и др.), где статистика и вероятность давно вошли в школьные программы, нередко встречается непонимание методов и интерпретаций вероятностных фактов (см, например, [9]). Если в России мы видим болезни роста, связанные с недостаточным опытом, то в упомянутых странах, как правило, опыт достаточен. Проблема в некритическом использовании статистических и вероятностных вычислений «по аналогии», в отсутствии навыков поиска подходящего решения задачи и интерпретации результата. В конце статьи мы приводим два типичных примера.

АНАЛИЗ ТИПИЧНЫХ ОШИБОК И ЗАБЛУЖДЕНИЙ

1. Невозможные, достоверные и случайные события. Большинство учебных пособий трактует случайное событие, как событие, которое в условиях некоторого эксперимента может произойти, а может не произойти. Учебники тоже грешат этим. При этом создаётся ложное противопоставление случайных событий невозможным и достоверным. Характерный пример из учебника.

Укажите, какие из следующих событий являются достоверными, какие – невозможными, какие – случайными.

А. В будущем году в Москве выпадет снег.

Б. Футбольный матч «Спартак» - «Динамо» закончится вничью.

В. На день рождения вам подарят говорящего крокодила.

Г. Завтра будет контрольная по математике.

Д. 30 февраля будет дождь.

Во-первых, здесь имеется терминологическая ошибка: случайные события противопоставлены невозможным или достоверным. Если принять это противопоставление, мы придём к серьезным трудностям, когда объединение или пересечение двух случайных событий может не быть случайным событием.

Пример. При бросании монеты события A = {Орел} и B = {Решка}, безусловно, случайные, но их объединение является достоверным, а их пересечение – невозможным.

Во-вторых, события А – Д не относятся к случайному эксперименту, а являются «событиями вообще». Такой подход разрушителен для дальнейшей формализации случайных явлений[1]. Итак, вторая ошибка – **смешение событий в бытовом смысле слова и случайных событий случайного эксперимента**.

Как правильно построить изложение материала и примеры

События окружающей жизни, которые рассматриваются сами по себе, в отрыве от других, плохо поддаются математическому описанию. Им невозможно назначить вероятности. Сначала следует описать случайный эксперимент. Когда это сделано, можно говорить о случайных событиях, возникающих в этом эксперименте.

Теория вероятностей изучает не всякие события, а только те, которые можно формализовать в рамках случайного эксперимента.

Не всякое жизненное событие удается так формализовать. Но и напротив, – несложно придумать хорошо формализованное событие, которое в житейском смысле событием не является.

Пример. Событие «Новая лампочка никогда не перегорит». В обычном смысле это вообще не событие, поскольку его невозможно зафиксировать – сколько нужно сидеть около этой лампочки и ждать, чтобы убедиться в том, что она НИКОГДА не перегорит?

В рамках формального эксперимента – это случайное событие, которое является пересечением бесконечного числа событий «Лампочка не перегорит через год», «Лампочка не перегорит через два года» и так далее.

[1] Наиболее неудачным в данном списке представляется событие «30 февраля будет дождь», поскольку, если все же выделить в нем эксперимент, уточнив регион и время суток, то мы получим эксперимент «наблюдение над погодой в N-ске в NN часов 30 февраля». В этом эксперименте невозможно оценить вероятность событий, поскольку эксперимент настолько уникален, что никогда не происходит.

В случайном эксперименте определяется множество элементарных событий[2]. Все подмножества этого множества – случайные события. Среди них есть **достоверное** и **невозможное**.

Достоверное событие состоит из всех элементарных событий эксперимента, поэтому наступит обязательно и имеет вероятность 1. Но от этого оно не перестает быть случайным.

Невозможное событие не содержит элементарных событий вовсе. Это – пустое множество. Оно обязательно не наступит и имеет вероятность 0. Но оно все равно случайное.

Еще бывают события с неопределённой вероятностью. Их можно называть неопределёнными, а можно не давать им специального названия.

Как переформулировать некоторые пункты в рассмотренной задаче

А. Случайный эксперимент – наблюдение над погодой в Москве во все дни 2014 года. Оцените степень достоверности события
«В какой-то день выпадет снег».
Придумайте несколько других случайных событий этого эксперимента. Оцените, если возможно, степень их достоверности.
Б. В случайном эксперименте наблюдают результат футбольного матча «Спартак» - «Зенит». Является ли событие «Ничья» достоверным или невозможным? Придумайте несколько других случайных событий в этом эксперименте.

О событиях В, Г и Д лучше не говорить, поскольку В и Г требуют уточнения слишком многих условий эксперимента. Вероятность события Д определить нельзя, поскольку Д относится к принципиально невозможному эксперименту.

2. Что, если монета встала на ребро? В ответ на бородатую шутку школьников «*А что делать, если монета встала на ребро?*» учителя часто ввязываются в длительную и бесплодную дискуссию по поводу этого события. Здесь тоже имеется непонимание. Дело в том, что **монета не встанет на ребро**. Не потому что мы «не учитываем» такую возможность – здесь просто нечего учитывать или не учитывать. Просто такого исхода «ребро» нет в эксперименте.

Правильный выход из ситуации. Как ответить школьнику

В эксперименте «Бросание монеты» явление «Ребро» не является событием (ни возможным, ни невозможным – просто никаким). По определению

[2] Определение элементарного события дать нельзя – это основное неопределяемое понятие, такое же, как точка в геометрии. Нужно просто сказать, что элементарные события или исходы – это простейшие события, из которых в данном эксперименте наступает один и только один.

эксперимент «Бросание монеты» состоит ровно из двух элементарных исходов «Орел» и «Решка», и никаких других элементарных исходов нет. Объединяя их, получить событие «Ребро», невозможно. Поэтому монета не встанет на ребро, не повиснет в воздухе и не закатится в щель.

Когда же мы физически бросаем монету, то мы моделируем случайный эксперимент с помощью подручных средств. Физическая монета может закатиться под плинтус, и если это случилось, нам приходится признать, что модель эксперимента оказалась неудачной и придется монетку перебросить.

Фантазировать о непредусмотренных событиях полезно, но **случайный эксперимент состоит из тех и только тех элементарных событий, которые предусмотрены в эксперименте.**

3. Зачем нужна теория вероятностей? Тем более – школьникам. Обычный ответ на этот вопрос:

Вероятностные и статистические методы в настоящее время глубоко проникли в приложения. Они используются в физике, технике, экономике, биологии и медицине. Особенно возросла их роль в связи с развитием вычислительной техники.

Это замечательное и верное наблюдение. Если Вы вашим школьникам расскажете что-нибудь такое, то они обязательно воспылают желанием изучать вероятность вкупе с прочей математикой, которая тоже проникла.

Как не очень сложно сказать правду

Помните задачу про перчатки?

В ящике 20 левых и 20 правых перчаток. Сколько нужно вытащить на ощупь перчаток, чтобы наверняка получилась одна пара?

Детерминистская математика в лице принципа Дирихле говорит – наверняка хватит 21 перчатки. Ответ точный, но… что скажет теория вероятностей?

После того, как извлечена одна перчатка, вероятность случайно вытащить еще четыре перчатки на ту же руку равна

$$\frac{19}{39}\cdot\frac{18}{38}\cdot\frac{17}{37}\cdot\frac{16}{36}\approx 0{,}047.$$

Значит, вероятность получить пару, взяв наудачу только пять перчаток, равна 0,95. Вы готовы поверить в такое событие? Ведь оно практически достоверное. Если готовы, то не нужно отсчитывать в темноте 21 перчатку. Возьмите пять штук. А если не готовы – возьмите… шесть. Теория вероятностей **практически гарантирует** вам нужную пару. Это смешная игровая задача. А вот серьезные задачи, которые также решает теория вероятностей.

1. Каким должен быть минимальный запас медикаментов в городе, чтобы практически наверняка хватило в случае стихийного бедствия, характерного для данной местности?

2. Сколько нужно запасти порций курицы и рыбы в самолет, чтобы практически наверняка не было недовольных пассажиров?
3. Сколько нужно иметь операторов в банке, чтобы скопление клиентов в очереди было очень редким явлением?
4. Сколько денег нужно заложить в банкомат, чтобы практически наверняка хватило на день?
5. Каким должен быть минимальный страховой взнос, чтобы страховая компания практически наверняка получила установленную прибыль?

Иногда практически достоверно удается решить задачу, которая вовсе не имеет определенного решения. Например:

Сколько купить саженцев, чтобы из них прижилось хотя бы 10?

Вместо достоверных, но бессмысленных ответов **теория вероятностей часто даёт разумные и практически достоверные ответы**.

4. Переоценка комбинаторики. Отметим одну общую вредную тенденцию: многим учителям математики свойственно переоценивать роль комбинаторики в преподавании теории вероятностей. Часто учитель формально излагает комбинаторные факты и формулы, а затем предлагает задачи со словом «вероятность» в качестве примера применения.

***Пример**: у Буратино в правом кармане 4 серебряных и 2 золотые монеты. Буратино, не глядя, перекладывает три каких-то монеты в левый карман. Какова вероятность того, что обе золотые монеты оказались в одном кармане?*

Эта задача, появляясь в разных формулировках в разных сборниках и на экзаменах, вызвала массу обсуждений. Как выяснилось, большинство специалистов видят в ней комбинаторную задачу со следующим решением.

Общее число комбинаций из 6 монет по 3 равно C_6^3. Выбрать две золотые монеты из двух и одну серебряную из четырех и положить их в левый карман можно $C_2^2 \cdot C_4^1$ способами. Столько же есть способов положить эти выбранные монеты в правый карман. Получаем:

$$p = \frac{2C_2^2 C_4^1}{C_6^3} = \frac{2 \cdot 1 \cdot 4}{20} = 0{,}4.$$

Вот типичное проявление проблемы: вероятность в вузах часто преподается как приложение комбинаторики, и этот подход проецируется на школу. Мы не против комбинаторики. Но важнее воспитание вероятностного мышления. Давайте учить в первую очередь этому. Недолгое раздумье позволяет резко упростить множество возможных исходов. Нам неважно, где оказалась первая золотая монета. Давайте рассмотрим возможные размещения второй

Решение. Мысленно присвоим золотым монетам номера 1 и 2. Первая золотая монета окажется в каком-то кармане. В этот карман, кроме нее, попадут еще две монеты из пяти оставшихся. Значит, вероятность того, что вторая золотая монета случайно окажется в том же кармане, равна $\frac{2}{5}=0,4$.

Важно то, что отсутствие комбинаторики не сужает класс возможных задач. В изучении и преподавании теории вероятностей комбинаторика должна играть вспомогательную роль, и нужна она там, где вероятностные пространства обширные и без нее нельзя обойтись. Нужно использовать хорошие и важные задачи с простыми вероятностными множествами, а не наполнять голову школьника комбинаторными отношениями, выдавая их за суть науки.

5. Прогнозы. В одном из методических изданий встретилась задача.

Две фирмы собирают компьютеры из комплектующих деталей. Первая фирма собирает «левые» компьютеры, вероятность их поломки равна 2,4%. Вторая фирма собирает компьютеры из качественных деталей. Вероятность их поломки равна 0,6%. Первая фирма продала 6000 компьютеров. Вторая фирма – 22500 компьютеров. Какая фирма получит больше жалоб на качество?

Решение*.* x_1 *- количество неисправных компьютеров первой фирмы,* x_2 *- второй фирмы. Тогда*

$$\frac{x_1}{6000}=0,024;\ \frac{x_2}{22500}=0,006.\ \textit{Значит, }\ x_1=144;\ x_2=135,\ \ x_1>x_2.$$

Вероятно, следует уточнить, что речь идет о поломке в течение какого-то времени, скажем, в течение гарантийного срока. Неявно предполагается также, что жаловаться будут те и только те, у кого сломается компьютер. Но, даже абстрагируясь от нечеткости условия[3], и оставив только суть, следует признать, что задача не корректна ни в постановке, ни, тем более, в решении.

Верный ответ на вопрос задачи в том виде, в каком он поставлен, может быть только один – неизвестно, какая фирма получит больше рекламаций (жалоб).

[3] Неудачно использованное слово «левые» несет неоправданный негативный оттенок. Можно подумать, что фирма подпольная. В дальнейшем выясняется, что слово «левые» противопоставлено обороту «из качественных деталей». То есть, первая фирма просто закупает дешевые комплектующие. И, вероятно, продает свои компьютеры дешевле, чем вторая. Таким образом, клиенты второй фирмы просто доплачивают за надежность.

В чем ошибка? Автор задачи механически подходит к вероятностному вопросу, выдавая математические ожидания величин за их достоверные значения. На самом деле компьютеров сломается столько, сколько сломается, и клиентов с рекламациями будет столько, сколько будет. 144 и 135 – это вовсе не x_1 и x_2, а математические ожидания этих величин, то есть теоретические средние значения. Всего лишь центры, вокруг которых возникает случайное рассеивание. Более вероятно, что недовольных клиентов у первой фирмы будет больше. Это нам подсказывает интуиция и это можно доказать. Но интуиция говорит и другое: средние 144 и 135 довольно близки, а рассеивание обеих величин значительно, поэтому недовольных клиентов у второй фирмы может оказаться больше. Расчет[4] показывает, что вероятность события $x_1 > x_2$ равна 0,696, вероятность события $x_1 < x_2$, равна 0,283 и есть еще событие $x_1 = x_2$, вероятность которого равна 0,021.

Если событие $x_1 = x_2$ (жалобщиков поровну), можно сбросить со счетов как маловероятное, то событие $x_1 < x_2$ (жалоб больше у второй фирмы) вполне вероятно (0,283). Нельзя уверенно утверждать, что у первой фирмы жалоб будет больше.

Как удачно использовать задачу в 8 или 9 классе, и какие выводы следует сделать

Сначала обсудим более простую задачу – сколько первой фирме ждать жалоб? Математическое ожидание числа гарантийных случаев у первой фирмы $n_1 p_1 = 6000 \cdot 0,024 = 144$.

Не вдаваясь в тонкости, можно попытаться дать прогноз: предположим, что число обращений в первую фирму будет 144 ± 20, то есть от 124 до 164 (отклонение 20 взято на глазок). Это не очень определенный прогноз, и степень его достоверности нам неизвестна.

Попробуем дать прогноз 144 ± 10, то есть от 134 до 154 обращений. Этот прогноз более определенный? Да. Он более достоверный? Конечно, нет, потому что вероятность более узкого интервала меньше:

$$\mathrm{P}\left(134 \le x_1 \le 154\right) < \mathrm{P}\left(124 \le x_1 \le 164\right).$$

Не вычисляя вероятности, мы сделали важное качественное наблюдение: **чем прогноз определеннее[5], тем он менее достоверен** (вероятность его

[4] Чтобы точно выяснить, насколько более вероятно преобладание жалоб у первой фирмы, нужно провести обширные вычисления. При желании и возможности учитель с учениками в несколько приемов могут вывести формулу:

$$\mathrm{P}\left(x_1 > x_2\right) = \sum_{k=0}^{6000} \left(C_{6000}^{k} 0,024^{k} 0,976^{6000-k} \cdot \sum_{j=0}^{k-1} C_{22500}^{j} 0,006^{j} 0,976^{22500-j} \right).$$

Вычисления на доске наверно будут несколько утомительными, зато компьютер (например, Excel) позволяет выполнить их за несколько минут. Результат, как уже сказано, равен 0,696.

[5] Определенность здесь используется в бытовом смысле. Чем шире прогнозируемый интервал значений, построенный около некоторого центра, тем прогноз менее определен в том смысле, что он дает меньше информации.

осуществления меньше). Можно дать совершенно определенный прогноз $x_1 = 144$, как это сделал автор задачи, но его достоверность крайне низка.

Сообщим учащимся без вычислений, что

$$\mathrm{P}\left(124 \le x_1 \le 164\right) \approx 0{,}92, \quad \mathrm{P}\left(134 \le x_1 \le 154\right) \approx 0{,}62, \text{ а } \mathrm{P}\left(x_1 = 144\right) \approx 0{,}034.$$

Таким образом, сделанный наугад прогноз 144 ± 20 оказался вполне вероятным, прогноз 144 ± 10 не очень хороший, а автор задачи взял за достоверный факт маловероятное событие.

Вернемся к исходной задаче. Как мы уже знаем, математическое ожидание числа рекламаций у первой фирмы $\mathrm{E}x_1 = 144$. У второй фирмы математическое ожидание числа рекламаций вычисляется так же:

$$\mathrm{E}x_2 = n_2 p_2 = 22500 \cdot 0{,}006 = 135.$$

Будем прогнозировать количество жалоб 124 – 164 для первой фирмы и 115 – 155 (то есть 135 ± 20) для второй фирмы, Опираясь только на это, мы уже понимаем, что жалоб у первой фирмы, скорее всего будет больше, но может случиться и иначе: ведь 155 заметно больше, чем 124

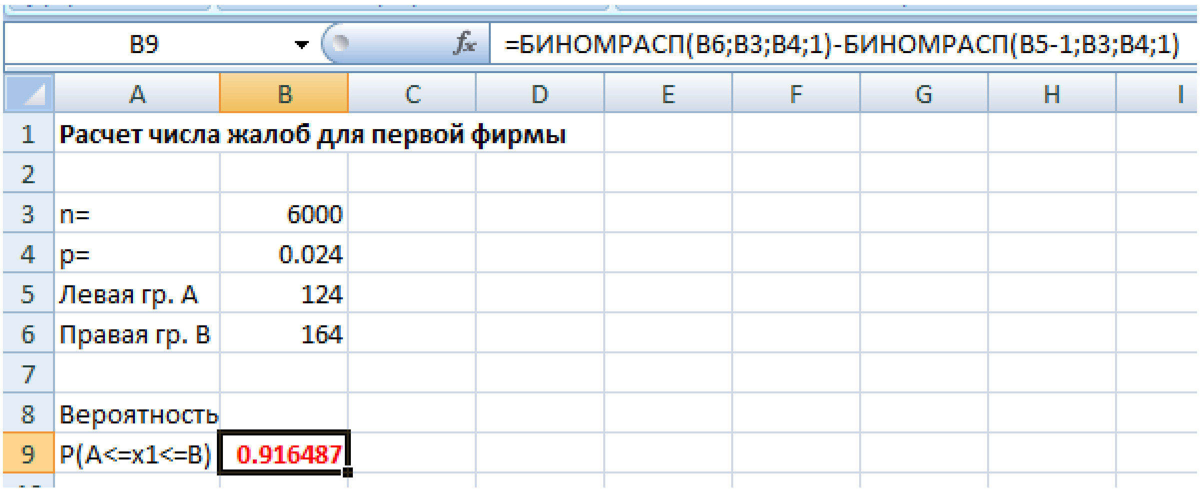
Мы решили задачу на качественном, но интуитивном уровне. Точное количественное решение описано в ссылке на предыдущей странице, но оно и не требуется. Важно подвести учащихся к мысли, что прогнозирование – приложение теории вероятностей и что, как правило, любой прогноз – палка о двух концах: чем он более определенный, тем он менее достоверный. И наоборот. В каждом конкретном случае приходится искать промежуточное решение – компромисс между определенностью и достоверностью.

1. При работе на профильном уровне учащимся можно предложить провести вычисления вероятностей сделанных прогнозов, то есть доверительных интервалов. Требуется знание формулы биномиальной вероятности и умение работать с Excel или аналогичным средством. Выражение для расчета выглядит следующим образом:

$$\mathrm{P}\left(a \le x_1 \le b\right) = \sum_{k=a}^{b} C_{6000}^{k} 0{,}024^{k} 0{,}976^{6000-k}$$

и преобразуется в разность $\sum_{k=0}^{b} C_{6000}^{k} 0{,}024^{k} 0{,}976^{6000-k} - \sum_{k=0}^{a-1} C_{6000}^{k} 0{,}024^{k} 0{,}976^{6000-k}$.

При расчете в Excel формула записывается с помощью функции БИНОМРАСП() и размещается в одной ячейке (на рисунке в ячейке B9).



B9 f_x =БИНОМРАСП(B6;B3;B4;1)-БИНОМРАСП(B5-1;B3;B4;1)

	A	B	C	D	E	F	G	H	I
1	**Расчет числа жалоб для первой фирмы**								
2									
3	n=	6000							
4	p=	0.024							
5	Левая гр. A	124							
6	Правая гр. B	164							
7									
8	Вероятность								
9	P(A<=x1<=B)	0.916487							

Аналогично можно вычислить вероятности подобных прогнозов для второй фирмы.

6. Объем случайной выборки. В другом методическом пособии, выпущенном институтом развития образования Якутии, обнаружилась интересная статья учителя математики, в которой описан урок о выборочных обследованиях. Процитируем текст, опуская несущественное.

Возникла проблема определения уровня знаний девятиклассников в некотором регионе России, например, в Республике Саха (Якутия). Понятно, что для получения точных результатов исследования необходимо, чтобы в исследовании участвовали все девятиклассники всего региона. Допустим, что мы провели контрольную работу из шести заданий и посмотрим, сколько учеников правильно решает то или иное количество задач. …Для получения вполне достоверной информации по интересующему нас вопросу достаточно провести выборочное исследование, ограничившись проверкой знаний сравнительно небольшой части школьников.

В ситуациях, подобных нашей, обычно ограничиваются обследованием 5 – 10 % всей изучаемой совокупности. Кроме того, осуществляют случайный отбор учащихся, обеспечивая одинаковую вероятность попадания в выборку любого объекта генеральной совокупности.

Пусть в некотором селении девятиклассников всего 716 учащихся. Выберем случайным образом из генеральной совокупности всего 50 учащихся (6,9 %).

В чем ошибка? На самом деле для достижения определенной уверенности в обследовании большой генеральной совокупности специалисты берут выборку вполне определенного объема, **который не зависит от численности совокупности.** Если изучаемый признак имеет всего лишь два значения (как в описанном случае: либо учащийся решил нужное число задач, либо нет), то можно рассуждать следующим образом.

Пусть интересующая нас случайная величина p – доля учащихся всей совокупности, верно решающих не менее пяти задач. Мы хотим оценить эту долю с точностью, например, $0,025$. В качестве оценки мы используем соответствующую долю учащихся выборки, то есть частоту события «Учащийся в выборке верно решил не менее пяти задач». Обозначим эту частоту $\hat{p}$.

Сделанная по выборке частотная оценка $\hat{p}$ должна отличаться от истинного значения p не больше, чем на $0,025$. Допустимая вероятность ошибки пусть будет $0,05$ (вероятность того, что отклонение оценки от истинной доли больше, чем $0,025$). Иными словами, мы хотим получить такую выборочную оценку $\hat{p}$, что

$$\mathrm{P}\left(\left|\hat{p}-p\right|>0,025\right)\le 0,05.$$

Какого объема n выборку следует взять, чтобы это неравенство выполнялось? При больших значениях числа испытаний (учащихся в выборке) истинное распределение $\hat{p}$ заменяют нормальным распределением[6], из свойств которого следует, что удвоенное стандартное отклонение величины $\hat{p}$ должно быть не больше допустимой ошибки[7], то есть

$$2\sqrt{\frac{p(1-p)}{n}}\le 0,025, \quad \text{откуда } n\ge 6400p(1-p).$$

Поскольку $p(1-p)\le 0,25$, можно положить $n\ge 6400\cdot 0,25=1600$. Тогда нужное неравенство будет заведомо выполняться. Как видим, для оценки n не нужно знать p, а если все же про p что-то известно заранее, то оценку n можно несколько уменьшить[8].

Обратим внимание – в оценке объема выборки n **не используется размер совокупности**. Поэтому странно измерять объем выборки долями генеральной совокупности. Если в селении 716 девятиклассников, то для получения почти достоверных результатов нужно обследовать намного больше, чем 50 человек. А вот если мы имеем, скажем, 70000 девятиклассников, то достаточно обследовать 1600 (при указанной допустимой ошибке и достоверности), не думая о том, сколько это процентов генеральной совокупности.

Что из этого и как следует сказать школьникам 7 – 9 класса

Удивительным и важным свойством выборочного обследования является то, что объем выборки не зависит от размера совокупности. В несложных обследованиях достаточная точность (погрешность 3 – 5 %) с большой вероятностью (обычно берут 0,9 – 0,95) достигается при выборке объемом

---

[6] Близость нормального распределения и распределения оценки частоты является предметом центральной предельной теоремы теории вероятностей.

[7] Множитель перед стандартным отклонением зависит от допустимой вероятности ошибки. При допустимой ошибке 0,05 множитель следует взять 2. На самом деле это даже с небольшим запасом. При множителе 2 вероятность ошибочного прогноза не превышает 0,0456. Если же мы хотим точно следовать допустимой вероятности ошибки 0,05, то можно взять не 2, а 1,96.

[8] Например, если заранее известно, что пять или больше задач решает не менее 80 % учащихся, то $p\ge 0,8$, поэтому можно взять $n\ge 6400\cdot 0,8\cdot 0,2=1024$.

1600 – 2000. Таким образом, можно обследовать совокупность большого размера на распространенность какого-либо признака. Имеется и неудобная сторона: если совокупность невелика, то приходится обследовать ее всю или почти всю. В таких случаях применяют специальные выборочные методы.

Другой важный аспект – **репрезентативность**, то есть представительность выборки. Выборка должна хорошо отражать изучаемое свойство всей совокупности. Чаще всего наилучший способ создать репрезентативную выборку – сделать ее с помощью случайного выбора.

К сказанному можно добавить, что точность обследования растет медленно – пропорционально $\sqrt{n}$. Пусть, например, в каком-то обследовании для достижения точности 10% пришлось взять выборку $n = 400$ человек. Предположим, заказчик исследования считает достигнутую точность недостаточной и требует, чтобы допустимая ошибка была понижена до 2%, то есть в пять раз. При той же достоверности исследования выборку придется увеличить в $5^2 = 25$ раз. Теперь она должна быть $400 \cdot 25 = 10000$ человек. В результате расходы на составление выборки, само обследование и обработку данных вырастают даже не в 5, а в 25 раз. Не исключено, что, узнав об этом, заказчик ограничится более скромными требованиями.

7. Роль монеты и кубика. Учителя математики в преподавании теории вероятностей, к сожалению, мало используют эксперимент. Да и методик экспериментального обучения школьников пока еще очень мало, они бессистемны и разрозенны. При изучении статистики и вероятности эксперимент необходим так же, как лабораторные работы на уроках физики. К счастью, для проведения вероятностного эксперимента часто не нужны дорогостоящие установки или реактивы[9].

На семинаре в Елецком университете в 2010 году учительница рассказала, что один ее ученик на уроке заявил: «Вы говорите, что теория вероятностей глубоко проникла во все сферы жизни, а мы на уроках все время решаем задачки про кубики и монеты. Зачем?». Учительница ответила так.

– Видишь ли, чтобы научиться вычислять вероятности событий окружающей жизни, сперва нужно потренироваться на простых событиях, вероятности которых вычислять легче.

Ответ, разумеется, верный. Действительно, в случайных опытах с монетами, кубиками, картами, жребиями элементарные события равновозможны, и поэтому эти опыты легко описывать, к тому же они позволяют иллюстрировать и проверять утверждения и рассуждения. Но главное не это.

Как хорошо ответить на такой вопрос

Монета и кубик позволяют не только иллюстрировать вероятностные эксперименты, но и моделировать их.

[9] Как правило, стоимость оборудования для проведения вероятностного эксперимента в классе, составляет 2 рубля на одного ученика – пластиковый стаканчик за рубль и десять гривенников.

Предположим, нам нужно оценить, сколько мальчиков и девочек разумно ожидать среди 30 первоклассников на будущий год.

Можно ли провести расчет? Да, можно, но школьники этого, как правило, не умеют[10]. Можно ли провести такой эксперимент вживую? Нужно дождаться будущего года и посчитать пришедших детей, но это исследование будет уже неактуально. Зато уже сейчас можно бросить 30 монет и посмотреть, сколько выпадет орлов и решек[11]. Если этот эксперимент провести много раз (например, 100), то мы сможем по результатам наблюдений сделать разумные верхние и нижние оценки количества девочек и мальчиков, даже не зная многих фактов теории вероятностей. Это называется статистическим моделированием или методом Монте-Карло.

Чтобы провести такой эксперимент в классе, в котором 25 человек, достаточно, чтобы каждый ученик четыре раза бросил 30 монет, подсчитывая количество выпавших при каждом бросании орлов. Все 100 бросаний по 30 монет в каждом будут проведены в течение 5 – 10 минут. Мы провели такой эксперимент. Результаты в таблице:

N	орлов	N	орлов	N	орлов	N	орлов	N	орлов
1	**17**	21	**14**	41	**18**	61	**8**	81	**13**
2	**15**	22	**16**	42	**13**	62	**14**	82	**13**
3	**13**	23	**19**	43	**17**	63	**19**	83	**18**
4	**13**	24	**19**	44	**17**	64	**13**	84	**11**
5	**19**	25	**14**	45	**18**	65	**14**	85	**13**
6	**15**	26	**11**	46	**17**	66	**19**	86	**14**
7	**20**	27	**13**	47	**16**	67	**13**	87	**16**
8	**16**	28	**10**	48	**19**	68	**14**	88	**13**
9	**12**	29	**17**	49	**12**	69	**14**	89	**13**
10	**11**	30	**12**	50	**12**	70	**14**	90	**9**
11	**14**	31	**14**	51	**17**	71	**10**	91	**17**
12	**15**	32	**15**	52	**12**	72	**13**	92	**11**
13	**13**	33	**19**	53	**13**	73	**16**	93	**15**
14	**15**	34	**19**	54	**14**	74	**15**	94	**17**
15	**20**	35	**20**	55	**17**	75	**14**	95	**14**
16	**9**	36	**15**	56	**11**	76	**16**	96	**9**
17	**19**	37	**15**	57	**14**	77	**14**	97	**16**
18	**14**	38	**16**	58	**17**	78	**18**	98	**10**

[10] Ситуация снова сводится к расчетам по биномиальной формуле. Нужно уменьшать число k при общем числе испытаний $n=30$ до тех пор, пока вероятность $\mathrm{P}(s \le k)$ того, что число успехов s (например, мальчиков) не больше k, не станет меньше, чем мы готовы допустить. Мы получим разумную нижнюю оценку числа мальчиков и, тем самым, верхнюю оценку числа девочек. В силу симметрии те же оценки будут верны, если поменять мальчиков и девочек местами.

[11] Будем считать, что появления мальчика и девочки равновозможны. На самом деле эти вероятности зависят от разных обстоятельств, но, как правило, мало отличаются от 0,5.

19	**15**	39	**17**	59	**14**	79	**12**	99	**19**
20	**16**	40	**12**	60	**13**	80	**18**	100	**16**

Как видим, при бросании 30 монет число орлов может сильно отличаться от 15. В этой серии опытов получилось от 8 до 20 орлов из 30 монет.

Поставим перед школьниками вопрос – следует ли директору школы рассчитывать на то, что число мальчиков в будущем первом классе будет от 8 до 20? Составим таблицу частот:

Число орлов	8	9	10	11	12	13	14	15	16	17	18	19	20
Сколько раз выпало	1	3	3	5	7	15	17	10	10	11	5	10	3
Частота	0.01	0.03	0.03	**0.05**	**0.07**	**0.15**	**0.17**	**0.1**	**0.1**	**0.11**	**0.05**	**0.1**	0.03

Дальше можно поступать по-разному. Например, договоримся считать маловероятными события с вероятностью не больше 0,1, и не будем на них рассчитывать. Возьмем 15 мальчиков (ожидание) за середину и будем отступать от этой точки влево и вправо на один шаг до тех пор, пока суммарная частота (оценка вероятности) внутри выбранного интервала не станет 0,9 или больше. В таблице нужный отрезок частот выделен. Сумма выделенных частот (в столбцах от 11 до 19 орлов) как раз 0,9. Таким образом, **в наших допущениях о достоверности прогноза нужно ожидать от 11 до 19 мальчиков, и, видимо, остальные будут девочками[12].**

Непростая и важная вероятностная задача прогнозирования решена вообще без формул с помощью моделирования. Заодно мы показали, что обычная монета может быть инструментом изучения случайных явлений в природе и обществе.

Можно придумать много экспериментов, где статистическое моделирование с помощью монет и кубиков оказывается результативным. Но легко перестараться.

1. Для статистического моделирования лучше подходят компьютерные программы, чем монеты и кубики.
2. Моделирование действительно сложных экспериментов, которые представляют интерес, само по себе непростая задача.

Главное – убедить школьников, что **монеты и кубики (и другие простые генераторы равновозможных событий) – это средства моделирования изменчивых процессов и явлений.**

[12] Справедливости ради скажем, что если бы на этом же уровне значимости 0,1 мы решали задачу теоретически, то получили бы тот же интервал – от 11 до 19. Так в данном случае проявился закон больших чисел. Мы здесь не обсуждаем, насколько хорошо статистическое моделирование приближает результаты теоретических расчетов. Ясно, что это зависит от числа экспериментов – чем больше, тем, скорее всего, лучше. Наши 100 экспериментов дали нулевую разницу между теорией и практикой. Может быть, нам просто повезло.

А как обстоят дела за рубежом?

В заключение хотим отметить, что проблемы имеются не только в России. Во введении уже отмечалось, что во многих странах статистика и теория вероятностей давно вошли в школьные образовательные стандарты. Казалось бы, дело должно обстоять лучше, чем у нас. Однако и там нередко встречается непонимание сути методов и интерпретаций вероятностных фактов. Как правило, и это видно из содержания учебников, имеется перекос в сторону техники применения вероятностных методов без осмысления сути задач. Приведем лишь два примера из большого числа имеющихся.

8. Что вероятнее? В одном из проектов АТЭС рабочая группа создавала тест для проверки готовности старшекурсников образовательных специальностей к профессии учителя математики. В разделе «Вероятность и статистика» была предложена следующая задача.

Баскетболист утверждает, что в среднем попадает в корзину 70 раз из 100 попыток. Тренер решил проверить это и обнаружил, что из 100 попыток баскетболист попал в корзину только 55 раз. Что более вероятно – что баскетболист преувеличил свои достижения, а 55 – его обычный результат или что он говорит правду, но у него просто неудачный день?

Предлагаемое авторами решение. Предположим, что вероятность попадания 0,7. Тогда стандартное отклонение числа попаданий при ста бросках равно $\sqrt{100 \cdot 0,7 \cdot 0,3} = 4,58$.

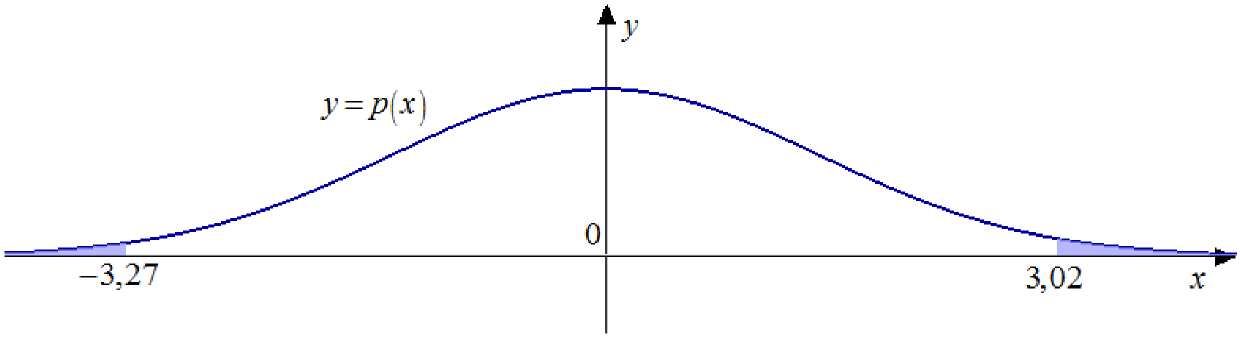
Имеет смысл рассмотреть наиболее вероятное событие, включающее 55 попаданий, но не включающее 56 попаданий, то есть событие «при среднем числе попаданий 70 случилось 55 попаданий или меньше». В силу центральной предельной теоремы эта вероятность приближенно равна вероятности того, что случайная величина ξ, имеющая стандартное нормальное распределение, не превосходит $\dfrac{55-70}{4,58} = -3,27$.

Теперь предположим, что истинное среднее 55, то есть вероятность попадания в действительности близка к 0,55. Тогда стандартное отклонение числа попаданий при ста бросках равно

$$\sqrt{100 \cdot 0,55 \cdot 0,45} = 4,97.$$

Проверим, насколько вероятно событие «при среднем 55 случилось 70 попаданий или больше». В этом случае вероятность не менее 70 попаданий близка к вероятности того, что случайная величина ξ, распределенная по стандартному нормальному закону, окажется не меньше, чем $\dfrac{70-55}{4,97} = 3,02$.

Вероятности событий $\xi \le -3,27$ и $\xi \ge 3,02$ показаны площадями закрашенных на рисунке фигур.



Очевидно, $\mathrm{P}(\xi \le -3{,}27) < \mathrm{P}(\xi \ge 3{,}02)$. Значит, событие «Случайно не меньше 70 при среднем 55» более вероятно, чем «Случайно не больше 55 при среднем 70». Наверно, баскетболист привирает.

Здесь мы видим применение готового рецепта к искусно скомпонованной и более-менее естественной задаче. Казалось бы, все хорошо.

В чем ошибка? Нет ошибки. Есть глубокое непонимание полученных результатов. Почему? Давайте найдем эти вероятности, пользуясь, например, таблицей стандартного нормального распределения или Excel:

$$\mathrm{P}(\xi \le -3{,}27) = 0{,}00054, \qquad \mathrm{P}(\xi \ge 3{,}02) = 0{,}00126.$$

Если вы не сразу заметили фигуры на рисунке, так это потому, что это нелегко. Обе вероятности пренебрежимо малы. Бессмысленно сравнивать вероятности двух практически невероятных событий. Это все равно, что всерьез рассуждать, что более вероятно при прогулке по Москве: случайно встретить дикого пингвина или бродячего слона.

Как правильно распорядиться сюжетом

Баскетболист утверждает, что в среднем попадает в корзину 70 раз из 100 попыток. Тренер решил проверить это и обнаружил, что из 100 попыток баскетболист попал в корзину только 55 раз. Насколько вероятно, что баскетболист говорит правду?

Получилась осмысленная задача на биномиальную вероятность. Можно воспользоваться нормальным приближением, как поступили авторы задачи.

Решение с помощью нормального приближения: вероятность попасть не более 55 раз в корзину при ожидаемом числе попаданий 70, приближенно равна вероятности того, что стандартная нормальная величина ξ не превосходит

$$\frac{55-70}{\sqrt{100\cdot 0{,}3\cdot 0{,}7}} = -3{,}27. \quad \mathrm{P}(\xi \le -3{,}27) = 0{,}00054.$$

Дополнительная вероятность $0{,}99946$ за то, что баскетболист несколько преувеличил свою меткость. Такое событие многие люди склонны считать практически достоверным.

9. Эми готовится к экзамену. Еще одна задача из того же источника.

Школьники Вашего класса получили задачу: *Эми готовится к экзамену по математике и еженедельно выполняет подготовительные тесты. Ее результаты (в процентах от максимального балла) показаны в таблице.*

Неделя (x)	1	2	3	4	5	6
y%	38	63	67	75	71	82

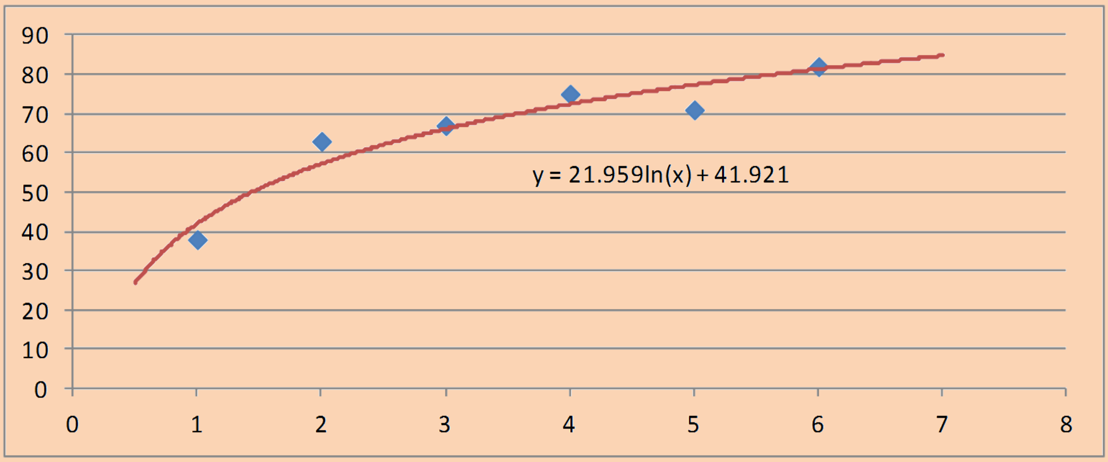
а) Постройте диаграмму рассеивания по данным таблицы;
б) Определите наиболее подходящую модель, которая описывает оценки Эми с течением времени.

Один школьник предложил линейную модель, другой школьник предложил квадратичную модель, а третий – логарифмическую.

А. Какой из ответов верный?

Б. Как убедить других школьников в том, что их модели не подходят?

Предлагаемое решение. А. Если мы рассмотрим диаграмму рассеивания, то увидим, что логарифмическая модель – правильная, потому что в целом повторяет форму диаграммы. Похоже, что оценки Эми со временем возрастают все медленнее.



Б. Линейная модель не подходит, потому что, используя ее для прогнозирования, мы увидим, что оценки Эми скоро превысят 100%, чего не может быть (логарифмическая модель также дает превышение 100%, но это случится много позже; можно считать, что к этому времени Эми уже закончит подготовку).

Если использовать квадратичную модель, то подходящий квадратный трехчлен имеет отрицательный старший коэффициент. В этом случае оценки со временем начнут уменьшаться. Кроме того, они достигнут отрицательных значений, что также неправдоподобно.

В чем здесь странность? Дело в том, что авторы задачи, говоря о модели, не указали никаких желаемых свойств этой модели – меру близости теоретических и наблюдаемых значений, непрерывность, возможность строить прогнозы и т.п. Если никакие свойства не указаны, то самой лучшей моделью, безусловно, будет точная модель. Например, сама таблица. Или можно построить многочлен Лагранжа, который пройдет точно через все точки. Можно придумать множество хороших и разных моделей[13].

Если линейная модель не устраивает авторов задачи, поскольку линейная функция преодолеет 100% рубеж, то почему устраивает логарифмическая? Рассуждение про «много позже» неубедительно. А если хороша логарифмическая, которая не скоро станет слишком большой, чем плоха квадратичная, которая тоже весьма не скоро станет слишком маленькой?

Как построить обсуждение со школьниками?

Скорее всего, никак. Если в российской школе когда-нибудь будут рассматриваться регрессионные модели[14], то, вероятно, тогда мы и будем о них говорить. Но в любом случае мы должны понимать, что **перед построением какой-либо модели следует иметь ясное представление, зачем она нужна и каковы ее свойства**. Например, будет она использоваться для экстраполяции или для интерполяции? Должна ли она быть непрерывной? Должна ли отражать вероятностные характеристики процесса?

Учитель может найти множество примеров, связанных с неаккуратным толкованием корреляций и регрессионных моделей, в книге [7]. Заблуждения в этой области поистине неисчерпаемы в своем разнообразии. Мир полон ложных корреляций, а литература – неудачных моделей.

Ссылки

1. 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.
2. 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.
3. Е.А.Бунимович, И.Р.Высоцкий, А.А.Макаров и др., Терминология, обозначения и соглашения в школьном курсе теории вероятностей и статистики, Издательский дом Первое Сентября, Математика, №17, 2009, с. 13-27.
4. Ковалева Г.С. и др. Основные результаты международного исследования образовательных достижений учащихся PISA-2006, М.: ЦОКО, 2007.
5. А.В.Петровский. Запрет на комплексное исследование детства // Репрессированная наука, Л.: Наука, 1991, с.126–135.

[13] Возможно, лучшую модель может предложить сама Эми, как наиболее информированный персонаж, но у нас нет ее телефона. Справедливости ради скажем, что обе задачи – про баскетболиста и про Эми – были исключены из теста.

[14] Некоторая попытка в качестве эксперимента поговорить о регрессиях, корреляциях и методе наименьших квадратов в учебном пособии для старшеклассников, предпринята в [10]. Но и там соответствующий раздел не содержит задач, поскольку отсутствие опыта преподавания не позволяет на данный момент понять, какие задачи и упражнения на эту тему важны, востребованы и возможны в настоящее время.

6. K.Kakihana, S.Watanabe. Statistic Education for Lifelong Learning. // The 6th East Asia Regional Conference on Mathematical Education. Vol.3, p. 318–322. 2013, ICMI, Phuket, Thailand.
7. Ю.Н.Благовещенский. Тайны корреляционных связей. М.: ИНФРА-М, 2009.
8. David S. Salsburg. The Lady Tasting Tea: How Statistics Revolutionized Science in the Twentieth Century. NY.: 2001.
9. M.Isoda, O.Gonzales. Survey on Elementary, Junior and High School Teachers' Statistical Literacy – The Need for Teacher training in Variability. //Journal of Scince Education Research, 2012, #37, Japan Society for Science Education, pp.61-76.
10. Ю.Н.Тюрин, А.А.Макаров, И.Р.Высоцкий, И.В.Ященко. Теория вероятностей и статистика. Экспериментальный учебник для 10-11 классов. М., МЦНМО, 2014.