Случайность
под контролем
вероятность и статистика
9-й класс

Сергей Яснов

Случайность под контролем: вероятность и статистика, 9-й класс

https://litres.ru/74656339
SelfPub; 2026

Аннотация

Семьдесят процентов дождя — это почти наверняка промокнуть? А если другое приложение обещает всего двадцать? С такого спора начинается знакомство с вероятностью и статистикой — через понятные ситуации, а не пугающие формулы.

Научитесь собирать данные и оценивать выборки, понимать среднее, медиану и моду, читать диаграммы без ловушек. Разберётесь, как задавать пространство исходов, считать вероятность, находить варианты с помощью таблиц и схем, работать с событиями «и» и «или», зависимыми и независимыми испытаниями. И увидите, где числа помогают сделать вывод, а где не дают гарантий.

Книга адресована девятиклассникам, которым нужны ясные объяснения и уверенность в задачах. От условия до проверки

ответа — понятный маршрут, который поможет применять знания в учёбе и повседневной жизни.

Книга создана с помощью ИИ. Обложка: GPT Image - Лицензия

Содержание

Случайность и данные: где встречаются вероятность и статистика 6
Собираем и упорядочиваем данные: таблицы, частоты и выборки 17
Находим главное в данных: среднее, медиана, мода и размах 30
Конец ознакомительного фрагмента. 37

Сергей Яснов Случайность под контролем: вероятность и статистика, 9-й класс

Случайность под контролем: вероятность и статистика, 9-й класс

Сергей Яснов

Случайность и данные: где встречаются вероятность и статистика

Во вторник утром 9-й «Б» собрался на экскурсию в ботанический сад. До поездки оставалось три дня, а за окном ветер гонял по двору жёлтые листья. Лиза открыла приложение на телефоне:

— Здесь написано: вероятность дождя — семьдесят процентов.

— А у меня двадцать, — возразил Артём. — Значит, дождя почти не будет.

Класс загудел. Брать зонты или нет? Одни уже представляли промокшие куртки, другие не хотели весь день таскать лишние вещи. И главное: как могут два приложения показывать такие разные числа?

Учительница не стала выбирать за класс. Она спросила:

— А что именно обещает прогноз? И на каких данных он основан?

Вопрос оказался интереснее спора. Чтобы ответить, ребятам предстояло разобраться, как случайность уживается с закономерностями и почему цифры помогают принимать решения, но не дают волшебной гарантии.

Когда будущее не расписано по минутам

Слово «случайность» часто понимают как «всё происходит без причины». Но это не совсем так. Дождь возникает не из пустоты: на погоду влияют температура, влажность, движение воздушных масс и множество других факторов. Просто человек не всегда располагает всей нужной информацией и не может с полной точностью предсказать, что произойдёт в конкретном месте в конкретный час.

Случайным мы называем событие, исход которого заранее нельзя уверенно определить. При этом случайность не обязательно означает полный хаос. Можно не знать, выпадет ли дождь в пятницу, но оценить, насколько он вероятен.

Представь обычный кубик. Если он правильный, у него шесть равновозможных граней. Вероятность выпадения каждой грани — 1 / 6. Это не обещание, что после шести бросков обязательно выпадет каждая грань ровно один раз. Может три раза подряд выпасть четвёрка. Но если бросать кубик очень много раз, доли разных результатов обычно становятся ближе к 1 / 6.

Вероятность помогает рассуждать о шансах. Она не сообщает: «Вот что точно случится». Она говорит: «При таких условиях этот исход возможен с такой-то степенью вероятности».

А что же делает статистика? Она помогает собирать сведения о том, что уже наблюдали, приводить их в порядок и находить в них смысл. Если вероятность обращена к возможным исходам, то статистика чаще всего начинает с фак-

тов и данных.

Данные — это сведения, собранные для ответа на вопрос. Например: температура воздуха каждый час, число опоздавших автобусов за неделю или результаты контрольной работы.

Наблюдение — отдельный зафиксированный факт. Например: «В среду автобус № 17 пришёл на остановку в 8:12» — одно наблюдение. А список времени его прибытия за двадцать учебных дней — уже набор данных.

Так же устроена разница между отдельным броском кубика и серией бросков. Один результат — наблюдение. Записанные результаты всех бросков — данные, которые можно анализировать.

От прогноза к исследованию

Учительница предложила классу не гадать, какое приложение «правильнее», а сначала уточнить вопрос. Ведь «Будет ли дождь?» — вопрос слишком широкий.

Когда именно? Во время всей поездки или только в пути к саду? В какой местности? Считается ли дождём короткий моросящий дождик? Чем точнее вопрос, тем полезнее данные для ответа.

Класс сформулировал его так: **«Насколько вероятен дождь в ботаническом саду во время экскурсии в пятницу?»**

Затем ребята стали разбираться, почему прогнозы могли отличаться. Приложения могли обновить данные в раз-

ное время, использовать разные модели расчёта или по-разному учитывать особенности местности. Кроме того, прогноз осадков относится к определённому месту и промежутку времени. Если одно приложение показывает шанс дождя в течение дня, а другое — только во время экскурсии, сравнивать их числа напрямую нельзя.

Значение вроде «70%» обычно означает оценку вероятности осадков в определённом месте и в указанный период. Это не означает, что дождь займёт ровно семьдесят процентов времени или что на семидесяти процентах территории обязательно прольётся дождь. Точная трактовка зависит от сервиса, поэтому полезно смотреть, к какому времени и месту относится прогноз.

— Значит, семьдесят процентов — это не «семьдесят процентов дождя»? — уточнил Артём.

— Именно, — ответила учительница. — И уж точно не гарантия, что каждый человек из класса промокнет.

Чтобы посмотреть, что происходило раньше, ребята нашли школьный дневник погоды и выбрали двенадцать дней с похожей погодой и временем года. Для каждого дня записали, был ли дождь в районе сада. Допустим, в пяти случаях из двенадцати дождь был.

Долю таких дней можно записать так:

$5 / 12 \approx 0{,}42 \approx 42\%$.

Это не точный прогноз на пятницу. Это результат, который получился в небольшой группе прошлых наблюдений:

примерно в четырёх случаях из десяти дождь шёл. Такие данные полезны для размышления, но их нельзя принимать за обещание будущего.

Вот здесь встречаются статистика и вероятность. Статистика сообщает: «В наших записях дождь был в пяти днях из двенадцати». На основе этих данных можно осторожно оценить вероятность дождя в похожей ситуации. Но оценка будет лучше или хуже в зависимости от того, насколько подходящими были дни для сравнения и насколько аккуратно ребята записывали погоду.

Один факт — ещё не закономерность

Представь, что автобус сегодня опоздал. Можно ли сделать вывод, что он всегда ходит с опозданием? Нет. Сегодня могла быть пробка или дорожные работы. Одно наблюдение рассказывает о конкретном случае, но почти ничего не говорит о типичной ситуации.

А если записывать время прибытия автобуса каждый учебный день в течение месяца, появится более содержательная картина. Можно выяснить, сколько раз он опаздывал, насколько сильно задерживался и бывают ли опоздания чаще по определённым дням. Чем больше подходящих и аккуратно собранных наблюдений, тем легче заметить повторяющиеся особенности.

Такую повторяющуюся особенность называют **закономерностью**. Например, при длительном наблюдении выяснилось, что автобус по понедельникам обычно задерживает-

ся чаще, чем по пятницам. Это закономерность в собранных данных, а не гарантия, что каждый понедельник он опоздает.

С закономерностями связана важная ловушка: хочется сделать быстрый вывод по маленькой выборке. Допустим, три последних раза автобус приходил вовремя. Это хорошая новость для этих трёх поездок, но ещё не доказательство, что он всегда приходит вовремя. Возможно, именно в эти дни на дороге было свободно. Возможно, наблюдений слишком мало.

С погодой похожая история. Если в июне три раза шёл дождь, это не позволяет уверенно заключить, что июль будет дождливым. Для сравнения нужны данные, подходящие по месту, сезону и условиям. А если записи собирались только в дни, когда класс специально проверял прогноз после сильного дождя, выборка могла получиться односторонней.

Полезный вопрос к любому набору данных: **«Как именно его собрали?»** Если наблюдения выбраны неудачно, даже аккуратные подсчёты могут привести к ошибочному выводу.

Что считать, а что — предсказывать

Чтобы не путать статистику с вероятностью, посмотрим на несколько знакомых ситуаций.

Результаты контрольной. Одна оценка показывает, как ученик справился с одной работой. Несколько оценок по разным темам дают более полное представление о том, что получается хорошо, а что стоит повторить. Статистика может

помочь увидеть, например, что задачи на проценты решаются уверенно, а уравнения с дробями вызывают трудности. Но средний балл не объясняет автоматически, почему возникли ошибки: для этого нужно рассмотреть сами работы и условия их выполнения.

Спортивный турнир. Если команда выиграла семь матчей из десяти, это важные данные о прошлых играх. Они могут помочь оценить её шансы в следующем матче. Но победа не становится гарантированной: состав соперника, травмы, стратегия и обстоятельства игры изменятся. Данные помогают рассуждать, а не отменяют неопределённость.

Число пассажиров в автобусе. Если ежедневно записывать, сколько человек садится на остановке в 8:00, можно заметить, что в будни пассажиров обычно больше, чем в выходные. Статистика помогает описать поток и планировать поездку. Вероятность может пригодиться, если нужно оценить шанс найти свободное место в переполненном автобусе — но для такой оценки потребуются наблюдения за разными днями.

Бросок кубика. Здесь можно рассуждать о вероятности заранее: у правильного кубика шанс выпадения шестёрки равен 1 / 6. Если много раз бросать кубик и записывать результаты, мы получим данные. По ним можно проверить, похож ли кубик на честный. Получается, один и тот же пример показывает, как вероятность и статистика работают вместе: первая описывает возможные шансы, вторая изучает реаль-

ные результаты.

Карта повседневных ситуаций

Когда встречаешь число или прогноз, попробуй задать три вопроса: что именно наблюдаем, какие данные нужны и какой вывод можно сделать? Эта простая карта помогает не перепрыгивать от одного факта к смелому предположению.

Ситуация
Что наблюдаем?
Какие данные нужны?
Какой вывод можно сделать?

Дождь в день экскурсии
Осадки в нужном месте и в нужное время
Прогнозы с временем обновления, записи о погоде в похожие дни
Оценить шанс дождя и решить, нужна ли защита от него

Автобус опаздывает
Время его прибытия
Записи за много учебных дней в одинаковое время
Оценить, насколько часто бывают опоздания

Команда играет в турнире
Победы, поражения и состав соперников
Результаты прошлых матчей и сведения о нынешней игре

Оценить шансы команды, не объявляя победу заранее

В классе пишут контрольную
Оценки и ошибки по темам
Работы учеников и результаты нескольких заданий
Найти темы, которые стоит повторить

В очереди в столовой
Время ожидания
Замеры в разные дни и в разное время
Понять, когда очередь обычно короче

Обрати внимание: вывод всегда должен подходить к данным. Если мы записывали время автобуса только утром, нельзя уверенно говорить, что вечером он опаздывает так же часто. Если наблюдали одну команду, нельзя по её результатам судить обо всех командах турнира.

Как класс принимал решение

После небольшой проверки ребята не стали выбирать приложение с самым приятным числом. Они посмотрели, когда обновлялся каждый прогноз, относился ли он к ботаническому саду и совпадал ли указанный промежуток с временем экскурсии. Затем учительница предложила подумать не только о вероятности, но и о последствиях.

Если вероятность дождя невелика, но промокнуть и замёрзнуть очень неприятно, лёгкий дождевик может оказать-

ся разумной страховкой. Если в прогнозе небольшой шанс кратковременного дождя, а у всех есть куртки с капюшоном, тяжёлые зонты могут быть лишними. Решение зависит не только от цифры, но и от того, насколько неудобен возможный исход.

Это практическая сторона рассуждений о вероятности: мы не ждём абсолютной уверенности, а выбираем разумный вариант с учётом информации и возможных последствий. Поэтому важны не только вопрос «Каков шанс?», но и вопрос «Что случится, если прогноз не оправдается?»

В итоге класс договорился за день до поездки снова проверить прогноз, а каждому — взять лёгкую куртку с капюшоном. Никто не объявил, что дождь обязательно будет или обязательно обойдёт сад стороной. Ребята просто использовали доступные сведения, оценили неопределённость и выбрали удобный план.

Мини-задание: найди статистику вокруг себя

Сегодня попробуй заметить три ситуации, в которых могут пригодиться вероятность или статистика. Например: успеешь ли ты на автобус, сколько времени занимает дорога до школы, как часто в твоей команде забивают голы или какой десерт чаще выбирают дома.

Для каждой ситуации запиши:

1. Что я хочу узнать? Сформулируй вопрос как можно точнее.

2. Какие данные мне нужны? Подумай, что именно

нужно наблюдать и сколько раз.

3. Какой вывод я смогу сделать? Не обещай больше, чем позволяют собранные сведения.

Например: «Как часто я опаздываю на первый урок?» Чтобы ответить, нужно записывать время прихода не один день, а, скажем, две учебные недели. Один случай опоздания — наблюдение. Записи за много дней — данные. По ним можно оценить частоту опозданий, но нельзя без дополнительных сведений точно определить их причину.

А теперь вопрос, на который невозможно честно ответить без данных: **какой маршрут до школы обычно занимает меньше времени — мой обычный или запасной?** Можно выбрать тот, который кажется короче, но это будет догадка. Чтобы сравнить маршруты, придётся несколько раз засечь время поездки по каждому из них, желательно в похожие дни и часы.

В начале этой главы класс спорил из-за двух прогнозов. К концу он научился задавать более точный вопрос, отличать отдельное наблюдение от набора данных и осторожно оценивать возможный исход. Вероятность помогает рассуждать о шансах; статистика помогает находить смысл в собранных сведениях. Вместе они не устраняют неопределённость, но делают наши решения разумнее. А следующий шаг — научиться собирать данные так, чтобы их было удобно сравнивать и анализировать.

Собираем и упорядочиваем данные: таблицы, частоты и выборки

Случайность под контролем: вероятность и статистика, 9-й класс — максимально простым и понятным языком

В пятницу после уроков девятиклассники собрались выбрать фильм для общего просмотра. На доске уже появились три предложения, в чате — ещё пять, а из-под тетрадей выглядывали записки с коротким «за комедию». Кто-то отправил сообщение дважды. Кто-то написал: «Можно что-нибудь интересное». И вот простой вопрос «Что будем смотреть?» внезапно превратился в маленькое расследование: сколько человек за каждый вариант, чьё мнение учтено и можно ли вообще доверять подсчёту?

Статистика начинается не с калькулятора и не с красивой диаграммы. Она начинается с порядка. Сначала нужно точно понять, что мы хотим узнать, затем собрать ответы по понятным правилам, а после — разложить их так, чтобы никто не потерялся между строчками.

Шаг 1. Задаём понятный вопрос

Первое подозрение у нашего класса возникло к самому вопросу. «Что посмотреть?» — звучит естественно, но каж-

дый может понять его по-своему. Один назовёт конкретный фильм, другой — жанр, третий предложит сериал, а четвёртый перечислит сразу пять вариантов. Подсчитать такие ответы трудно: мы будто пытаемся сравнить яблоки, названия деревьев и рецепт пирога.

Вопрос должен точно показывать, что именно мы собираем. Например:

Какой один жанр ты предпочёл бы для общего просмотра фильма с классом? Выбери один вариант: комедия, приключения, фантастика, детектив, ужасы, другое (укажи жанр), затрудняюсь ответить.

Это лучше, потому что участник понимает и тему, и правило выбора. Варианты ответа не должны пересекаться: если разрешён только один жанр, один голос нельзя одновременно записать и в «комедию», и в «приключения». При этом список должен оставлять место для ответа, который мы не предусмотрели: для этого есть «другое». А «затрудняюсь ответить» честнее, чем вынуждать человека выбирать наугад.

Неудачная формулировка: «Какой фильм лучше и что-нибудь предложите?»

Улучшенная формулировка: «Какой один жанр ты предпочёл бы для общего просмотра фильма с классом? Выбери один вариант из списка».

Первая формулировка смешивает два вопроса: о фильме и о жанре. Вторая задаёт один вопрос и объясняет, как отвечать. Если хочется узнать конкретные названия фильмов,

это лучше спросить отдельно.

Проверка шага: сможет ли любой участник ответить на вопрос одинаковым способом? Если ответы можно понять по-разному, формулировку стоит уточнить до начала опроса, а не после подсчёта.

Шаг 2. Решаем, кого спрашиваем

В 9Б — 24 ученика. Если мы хотим выбрать фильм именно для 9Б, целевая группа — все 24 ученика. В статистике такую группу называют **совокупностью**: это все, о ком мы хотим сделать вывод.

Но в первом сборе ответили только 12 человек. Эти 12 ответов — **выборка**, то есть часть совокупности, по которой мы пытаемся судить о большей группе. Выборка может быть полезной, но важно не путать её со всей группой.

Если ответили преимущественно те, кто первым увидел сообщение в чате или оказался рядом с доской, это не случайный срез класса. Возможно, активные участники чаще отвечают, а занятые или отсутствующие — реже. Поэтому 12 ответов помогают понять мнение ответивших, но не дают права уверенно говорить: «Весь класс хочет фантастику».

Для решения внутри 9Б лучше постараться опросить каждого. Если же мы хотим узнать мнение всей школы, опросить один класс недостаточно: девятиклассники одного класса могут отличаться от учеников других классов и возрастов. Для более широкого вывода нужно продумать, как выбирать участников из разных классов, а не просто спрашивать тех,

кто оказался поблизости.

Сразу договоримся и о способе сбора. Например, каждый ученик отвечает один раз, до конца пятницы, в общей анкете или на одинаковом бумажном бюллетене. Имена для подсчёта не нужны, если можно надёжно обеспечить один ответ от каждого другим способом. Если ответы пришли и на бумаге, и в чате, повторы нужно проверить: одно сообщение — не всегда один человек.

Проверка шага: известно ли, кто входит в целевую группу, сколько человек в ней и сколько действительно ответило? Это три разных числа, и полезно не смешивать их.

Шаг 3. Сохраняем ответы, не превращая их в кашу

После первоначальной проверки у класса осталось 12 ответов от разных учеников:

1. фантастика;
2. комедия;
3. комедия;
4. приключения;
5. фантастика;
6. детектив;
7. комедия;
8. приключения;
9. ужасы;
10. фантастика;
11. фантастика;
12. приключения.

Это **сырые данные**: ответы пока записаны так, как их можно прочитать по одному. Они ещё не показывают общую картину — её нужно получить подсчётом.

На этом этапе важно соблюдать несколько простых правил.

• **Считать людей, а не сообщения.** Если один ученик написал в чат, а затем сдал записку с тем же выбором, его нельзя посчитать дважды.

• **Не считать повторяющиеся ответы ошибкой.** Три ответа «комедия» — это три голоса, а не один и тот же ответ, случайно попавший в список.

• **Не превращать пропуск в голос.** Пустой бюллетень не означает «затрудняюсь ответить», если человек этого не отметил.

• **Не решать за участника.** Если при правиле «выбери один жанр» кто-то отметил два, не надо молча выбрать за него один из вариантов. Нужно заранее решить, можно ли уточнить такой ответ или его следует считать недействительным для этого вопроса.

• **Объединять одинаковые ответы по ясному правилу.** Например, «фантастика» и «научная фантастика» могут относиться к одной категории, если это определено до подсчёта. Важно не объединять ответы выборочно только потому, что так получается удобнее.

Если выяснилось, что некоторые ученики вообще не ответили, это тоже можно записать — как число пропусков. Но

пропуск не надо добавлять к жанрам: человек, не заполнивший вопрос, не проголосовал за «другое».

Шаг 4. Группируем ответы в таблицу

Теперь складываем вместе одинаковые варианты. **Абсолютная частота** — это количество ответов данной категории. Например, «фантастика» встретилась 4 раза, значит, её абсолютная частота равна 4.

Вот таблица для нашего опроса:

Вариант ответа	**Абсолютная частота**	**Доля от 12 ответов**
Фантастика	4	$4/12 \approx 33{,}3\%$
Комедия	3	$3/12 = 25\%$
Приключения	3	$3/12 = 25\%$
Детектив		

1
1/12 ≈ 8,3%

Ужасы
1
1/12 ≈ 8,3%

Другое
0
0/12 = 0%

Затрудняюсь ответить
0
0/12 = 0%

Всего
12
12/12 = 100%

Доля показывает, какую часть всех действительных ответов составляет категория. Чтобы найти её, частоту делят на число действительных ответов. Для фантастики это 4/12, или примерно 33,3%. Иными словами, среди ответивших примерно каждый третий выбрал фантастику.

Проверка суммы — важный момент, почти как пересчёт сдачи у кассы:

4 + 3 + 3 + 1 + 1 + 0 + 0 = 12.

Сумма абсолютных частот совпала с числом действительных ответов. Значит, при условии, что категории не пересекаются и каждый ответ попал ровно в одну из них, мы никого не потеряли и никого не посчитали дважды.

С долями есть небольшая тонкость. В таблице проценты округлены, поэтому 33,3% + 25% + 25% + 8,3% + 8,3% дают 99,9%, а не ровно 100%. Это не ошибка в ответах, а следствие округления. Точные доли составляют 4/12, 3/12, 3/12, 1/12 и 1/12; вместе они дают 12/12, то есть 100%.

Если при проверке частоты не складываются в нужное число, не стоит сразу «подправлять» какую-нибудь строку. Лучше вернуться к исходному списку: возможно, один ответ пропущен, один человек посчитан дважды или для ответа не нашлось категории.

Шаг 5. Читаем таблицу осторожно

По нашим 12 ответам фантастика лидирует: за неё четыре голоса. Комедия и приключения идут следом — по три голоса. Но лидерство не означает, что весь класс выбрал фантастику. Её предпочла только треть ответивших, а две трети выбрали что-то другое.

Здесь важно различать **наибольшую частоту** и **большинство**. Наибольшая частота означает, что у категории голосов больше, чем у каждой другой категории по отдельности. Большинство — это больше половины всех голосов. Четыре из двенадцати — наибольшая частота, но не большин-

ство.

Поэтому аккуратный вывод звучит так:

Среди 12 ответивших фантастика оказалась самым популярным вариантом: её выбрали 4 человека из 12. Эти ответы не описывают автоматически мнение всех 24 учеников.

А вот фраза «Весь класс хочет смотреть фантастику» выходит за пределы данных. Она делает вывод о 24 учениках на основании ответов 12, причём эти 12 могли отвечать по собственному желанию, а не быть выбраны случайным образом.

Хорошая статистика не только считает, но и указывает, **чьи именно ответы** были посчитаны. Это не мелкий шрифт, а часть честного вывода.

Шаг 6. Сравниваем два способа опроса

Класс решил провести повторный сбор: теперь всем 24 ученикам отправили одну и ту же анкету с одинаковой формулировкой и правилом «один ответ на человека». Заполнить её можно было до конца дня. Ответили 22 человека; двое не заполнили вопрос.

Результаты двух сборов выглядят так:

Способ сбора
Действительных ответов
Фантастика
Комедия
Приключения
Детектив

Ужасы

Первые записки и сообщения
12
4
3
3
1
1

Единая анкета для всего класса
22
7
6
5
2
2

В обоих сборах лидирует фантастика. В первом её выбрали 4 из 12 человек — около 33,3%, во втором 7 из 22 — около 31,8%. Это похоже на один и тот же общий сигнал: фантастика часто встречалась среди ответов, хотя убедительного большинства не получила.

Второй сбор надёжнее для описания именно 9Б: ответило больше учеников, все получили одинаковый вопрос, а правило одного ответа было понятным. Но и он не идеален: два

человека не ответили. А если оба пропустивших вопрос особенно не любят фантастику, их отсутствие может немного изменить картину.

Есть и ещё одно ограничение: эти два сбора проходили не одновременно. Результаты могли отличаться не только из-за способа опроса, но и из-за того, что часть учеников не участвовала в первом сборе или передумала к следующему дню. Поэтому нельзя объявить: «Онлайн-анкета сама по себе заставляет людей выбирать иначе». Мы можем сказать осторожнее: единая анкета дала более полные и сопоставимые ответы для нашего класса.

Бумажный опрос тоже может быть хорошим. Например, каждому ученику выдают одинаковый анонимный бюллетень, каждый выбирает один вариант, а затем все бюллетени складывают в коробку. Электронная форма тоже удобна, если все могут ею воспользоваться и система не позволяет отправить несколько ответов. Важнее не то, бумага это или экран, а то, получают ли участники одинаковые условия, может ли каждый ответить и можно ли проверить число ответов.

Если цель — выбрать фильм для всего класса, самый честный следующий шаг — напомнить двум неответившим об анкете, а затем принять решение по ответам класса. Если цель — узнать мнение всей школы, даже 22 ответа из одного класса будут лишь небольшой выборкой, а не точным портретом всех учеников.

Мини-практикум: попробуй сам

Представь, что собраны такие восемь ответов: комедия, фантастика, комедия, приключения, комедия, фантастика, ужасы, приключения.

Составь для них таблицу с абсолютной частотой и долей. Проверь, что сумма частот равна числу ответов.

Проверка: комедия — 3 ответа, фантастика — 2, приключения — 2, ужасы — 1. Всего 3 + 2 + 2 + 1 = 8. Доли равны соответственно 3/8 = 37,5%, 2/8 = 25%, 2/8 = 25% и 1/8 = 12,5%. Вместе они составляют 100%.

А теперь добавь к списку ответ «комедия и приключения», хотя вопрос разрешал выбрать только один вариант. Как поступить? Не угадывать за человека. Если можно, уточнить его выбор; если нет — отметить ответ отдельно и применить заранее выбранное правило. Главное — не менять правила подсчёта после того, как стало ясно, какой вариант лидирует.

Сбор данных похож на сбор улик: каждая мелочь важна, но сама по себе ещё не раскрывает всю картину. Чёткий вопрос задаёт направление, аккуратная процедура защищает от повторов и пропусков, таблица помогает увидеть частоты, а описание выборки не позволяет сказать больше, чем показывают ответы. Сначала мы выясняем, что именно узнали и от кого, и только потом решаем, какой вывод выдерживает проверку.

Именно такой порядок помогает не спутать отдель-

ные ответы с общей картиной. Чтобы данные можно было проверить, важно заранее определить, что считать, собирать ответы по одним и тем же правилам, а затем записывать их так, чтобы были видны частоты. И всегда нужно помнить, кого именно мы опросили: выводы о выборке не обязательно верны для всех.

Таблица показывает, сколько раз встречается каждый вариант, но не всегда сразу подсказывает, что можно сказать о данных в целом. Чтобы найти их характерные значения и понять, насколько сильно они различаются, пригодятся среднее, медиана, мода и размах.

Находим главное в данных: среднее, медиана, мода и размах

Случайность под контролем: вероятность и статистика, 9-й класс — максимально простым и понятным языком

В тренерской комнате спорили два тренера. У каждого на столе лежал список очков, набранных ведущим игроком его команды за пять матчей.

Первый список: 18, 19, 19, 19, 20.

Второй: 12, 17, 19, 21, 26.

— В среднем оба игрока набирали по 19 очков за матч, — сказал один тренер. — Значит, результат одинаковый.

— Средний — одинаковый, а вот игра совсем разная, — возразил второй. — Мой игрок почти всегда приносит команде около 19 очков. А твой то проваливается, то выдаёт рекорд.

Оба тренера были правы — каждый со своей стороны. Среднее действительно совпало. Но одно число не рассказало всей истории. Чтобы разобраться, понадобятся четыре показателя: среднее, медиана, мода и размах. Они не спорят за единственное звание «главного». Каждый отвечает на свой вопрос.

Одна карточка — четыре показателя

Начнём с первого списка: 18, 19, 19, 19, 20. Данные уже расположены по порядку — от меньшего значения к большему. Это удобно: некоторые показатели ищутся именно в упорядоченном ряду.

Представим, что перед нами лежит карточка с четырьмя подписями.

Карточка	**Как найти**	**Что показывает**
Среднее	(18 + 19 + 19 + 19 + 20) / 5 = 95 / 5 = 19	Сколько очков приходилось на матч в среднем
Медиана	Среднее, третье значение в ряду: 19	Какое значение находится посередине
Мода	Чаще всего встречается 19	Какой результат повторяется чаще других
Размах	20 − 18 = 2	Насколько различаются крайние значения

Одна карточка — четыре разных взгляда на одни и те же данные. Теперь разберём, чем каждый полезен и где может подвести.

Среднее: если сложить всё и поделить поровну

Среднее арифметическое знакомо многим: складываем все значения и делим сумму на их количество.

Среднее = сумма всех значений / количество значений.

В нашем примере игрок набрал 18 + 19 + 19 + 19 + 20 = 95 очков за пять матчей.

95 / 5 = 19.

Получается, в среднем он набирал 19 очков за матч. Среднее удобно, когда нужно оценить общий уровень, сравнить результаты или, например, рассчитать, сколько товара в среднем продаётся за день.

Есть и наглядный образ: среднее похоже на точку равновесия. Если представить числа на числовой прямой, их отклонения от среднего в сумме уравновесят друг друга. Для нашего ряда среднее равно 19: значение 18 отстоит от него на 1 влево, а значение 20 — на 1 вправо. Остальные числа стоят прямо на отметке 19.

Но у среднего есть важная особенность: в его расчёт входят все значения. Поэтому необычно большое или маленькое число может заметно его сдвинуть.

Вернёмся ко второму списку: 12, 17, 19, 21, 26.

12 + 17 + 19 + 21 + 26 = 95.

95 / 5 = 19.

Среднее такое же, как в первом списке. Но второй игрок набирал то 12 очков, то 26. Первое значение гораздо ниже среднего, второе — заметно выше. Среднее сложило их влияние в один результат, но не показало, насколько неровно шла игра.

Это не ошибка среднего. Просто оно отвечает на конкретный вопрос: «Какова сумма результатов, если разделить её поровну между всеми наблюдениями?» Оно не обещает, что именно такое значение встретилось в данных или что каждый отдельный результат был на него похож.

Выброс: одно необычное значение меняет картину

Представь, что ты записываешь, сколько минут уходит на дорогу до школы. За пять обычных дней получилось: 12, 13, 13, 14 и 15 минут.

Среднее:

12 + 13 + 13 + 14 + 15 = 67.

67 / 5 = 13,4 минуты.

А теперь на шестой день автобус задержался из-за перекрытой дороги, и путь занял 53 минуты. Если сравнить пять дней, где последний результат заменён на этот необычный, получится ряд: 12, 13, 13, 14, 53.

12 + 13 + 13 + 14 + 53 = 105.

105 / 5 = 21 минута.

Среднее выросло с 13,4 до 21 минуты. Один очень долгий путь сильно изменил картину. Такое заметно отличающееся

значение часто называют выбросом.

Важно не перепутать «необычное» с «неправильным». Если 53 появилось из-за ошибки при записи, данные нужно проверить и исправить. Если автобус действительно задержался, значение настоящее. Его не стоит молча выкидывать: лучше указать, что произошло, и подумать, отвечает ли среднее на нужный вопрос. Например, для оценки обычного пути до школы может быть полезнее медиана, а для планирования поездки с учётом возможных задержек необычный случай тоже важен.

Медиана: середина упорядоченного ряда

Чтобы найти медиану, сначала расположи значения по порядку. Медиана — это значение, которое оказывается в середине этого ряда.

В примере с обычными поездками данные уже упорядочены:

12, 13, 13, 14, 15.

Здесь пять значений. Среднее, третье по счёту, — 13. Значит, медиана равна 13 минутам.

Если добавить очень долгую поездку, ряд станет таким:

12, 13, 13, 14, 53.

Посередине по-прежнему стоит 13. Медиана не изменилась, хотя среднее заметно выросло. Это полезное свойство: медиана обычно меньше зависит от крайних значений, потому что важнее всего её место в упорядоченном ряду, а не точное расстояние до каждого числа.

Если количество значений чётное, одного центрального числа нет: посередине стоят два. Тогда медиана равна среднему арифметическому этих двух значений.

Например, упорядоченный ряд: 6, 8, 8, 10, 13, 15. Два центральных значения — 8 и 10.

(8 + 10) / 2 = 9.

Медиана равна 9. Её не обязательно встретить в исходных данных: в этом ряду такого значения нет. Это нормально — медиана показывает центр упорядоченного ряда, а не обязательно одно из наблюдений.

Медиана особенно удобна, когда данные содержат редкие, но очень большие или маленькие значения. Например, если в группе доходы большинства людей похожи, но у одного человека доход во много раз выше, среднее может оказаться выше того, что получает типичный участник. Медиана в такой ситуации часто лучше передаёт центральное положение данных. Но и она не расскажет, насколько разбросаны значения вокруг центра.

Мода: самое частое значение

Мода — это значение, которое встречается чаще других.

В первом списке, 18, 19, 19, 19, 20, мода — 19: это число появилось три раза. Мода отвечает на практический вопрос: «Какое значение встречается чаще всего?» Например, какой размер обуви чаще покупают в магазине или какой маршрут чаще выбирают пассажиры.

У моды есть особенность: она может быть не единствен-

ной. В ряду 1, 2, 2, 3, 3, 4 чаще других встречаются и 2, и 3. Значит, мод две.

А иногда моды нет. Во втором списке с баскетбольными результатами каждое число встречается один раз: 12, 17, 19, 21, 26. Ни одно значение не встречается чаще других, поэтому единственной наиболее частой величины здесь нет. В таких данных говорят, что мода отсутствует.

Конец ознакомительного фрагмента.

Текст предоставлен ООО «Литрес».

Прочитайте эту книгу целиком, купив полную легальную версию на Литрес.

Безопасно оплатить книгу можно банковской картой Visa, MasterCard, Maestro, со счета мобильного телефона, с платежного терминала, в салоне МТС или Связной, через PayPal, WebMoney, Яндекс.Деньги, QIWI Кошелек, бонусными картами или другим удобным Вам способом.