ТЕХНОЛОГИЧЕСКАЯ КАРТА ЗАНЯТИЯ

Тема занятия: Основы теории вероятностей и математической статистики

Аннотация к занятию: на данном уроке обучающиеся продолжат знакомство с основами теории вероятностей и математической статистики. В первой части урока они поговорят о случайных величинах, а также обсудят основные характеристики случайных величин и выборки. Во второй части урока закрепят полученные знания решением задач по данной теме.

Цель занятия: знакомство обучающихся с основами теории вероятностей и математической статистики.

Задачи занятия:

- познакомить с понятиями «случайная величина», «выборка», «математическое ожидание»;
- научить определять, что случайные величины независимы;
- применить полученные знания на практике.

Ход занятия

Этап занятия	Время	Деятельность педагога	Комментарии, рекомендации для педагогов
Организационный этап	2 мин.	Здравствуйте! Мы продолжаем изучать теорию вероятностей.	Приветствие. Создание в классе атмосферы психологическ ого комфорта
Постановка цели и задач занятия. Мотивация учебной деятельности обучающихся	10 мин.	Давайте проверим, хорошо ли вы разобрались в понятиях. Обратимся к сказке «12 месяцев». **Вопрос для обсуждения** Назовите мне события, которые вы увидели в данном фрагменте. **Ответы обучающихся** После 31 наступит 32 декабря и т. д. В декабре распустятся подснежники. **Вопрос для обсуждения** Как бы вы могли назвать данные события? **Ответы обучающихся** Невозможные.	Просмотр фрагмента сказки «12 месяцев»: https://www.youtube.com/watch?v=y6UB3y-vgg8 Просмотр фрагмента новостей: https://youtu.be/gJnu45D4L2A

		Вопрос для обсуждения Может у кого-то есть иное мнение? Я предлагаю посмотреть другой сюжет. Это новости Первого канала о событии, которое произошло в 2012 году в Белграде. Какой вывод можно сделать из того, что мы увидели? **Ответы обучающихся** Мы ошиблись по поводу второго события. Так как в реальной жизни такое событие произошло, то его уже нельзя назвать невозможным. Это случайное событие.	Способствовать обсуждению мотивационных вопросов
Изучение нового материала	50 мин.	Начнём! На предыдущем уроке мы обсудили понятие события. Событие характеризует результат эксперимента, в котором возможно всего два исхода — успешный или неуспешный. Иногда одного только успеха/неуспеха не хватает для описания результата эксперимента. Часто в ходе эксперимента мы получаем в качестве результата некоторое число. Итак, численное описание результата эксперимента называется случайной величиной. Приведём примеры случайной величины из жизни. Пусть мы бросаем несколько кубиков и смотрим сумму чисел на этих кубиках. Сумма этих чисел — случайная величина. Или, например, мы выбираем из группы людей случайного человека и измеряем ему рост. Это тоже будет случайной величиной. Обратите внимание, что в первом случае результатом эксперимента является натуральное число. Во втором случае результат — это произвольное положительное число, которое необязательно является целым. Так вот, случайная величина, которая может принимать лишь конечное количество значений, называется	Для справки: Сайт: https://resh.edu.ru/subject/lesson/3751/conspect/326747/ Перед уроком рекомендуется ознакомиться с материалами, представленными на сайте.

		дискретной случайной величиной. Мы с вами для простоты будем рассматривать только дискретные случайные величины. Случайная величина принимает различные значения с некоторыми вероятностями. Пусть $x_1, x_2, \ldots, x_n$ — это все возможные значения некоторой случайной величины X. Тогда сопоставим каждому значению вероятность: $p_1, p_2, \ldots, p_n$, где p_i — это вероятность X_i. Такое соответствие называется распределением случайной величины. Составим таблицу, где в верхней строке будут лежать значения нашей случайной величины, а в нижней — соответствующие вероятности. Такая таблица полностью описывает наше распределение. Отметим, что сумма вероятностей должна быть равна единице. В качестве примера рассмотрим эксперимент с броском игрального кубика ещё раз. Результат эксперимента — случайная величина X, принимающая натуральные значения от 1 до 6 с вероятностью ⅙ каждое. Таблица распределения случайной величины X указана на слайде. Со случайными величинами можно производить арифметические преобразования. Давайте решим небольшую задачу на эту тему. Условие. Пусть X — случайная величина, равная результату броска кубика. Найдите распределение случайной величины (X − 2)2. Решите задачу самостоятельно, нарисовав таблицу распределения искомой случайной величины. Теперь посмотрим на задачу вместе. Нарисуем таблицу распределения случайной величины икс. Добавим в эту таблицу ещё одну строчку — значение случайной величины $(X-2)^2$. Если X равно	

		1, то $(X-2)^2$ равно 1. Если двойке, то $(X-2)^2$ равно нулю. При тройке получается единица, при четвёрке четвёрка, при пятёрке девять, при шестёрке 16. Мы видим, что значения в единице и тройке склеились. Поэтому значение 1 принимается с вероятностью 2/6, то есть одна третья. Остальные значения — 0, 4, 9, 16 принимаются с вероятностью ⅙. Получается вот такая таблица для распределения случайной величины $(X-2)^2$. Теперь рассмотрим примеры из жизни, где встречаются случайные величины. Представим некоторую страну, где регулярно измеряют рост своих граждан. Для них закон распределения роста может выглядеть так, как написано в таблице. Здесь мы видим, что наша случайная величина — это рост, а значения, которые она может принимать, — это ниже 150, от 150 до 160, от 160 до 170 и так далее, то есть рост с шагом 10 см. Мы сделали это для простоты, чтобы иметь дело с дискретной случайной величиной. В таблице записаны соответствующие вероятности, которые мы вычислили экспериментально. Обратите внимание, что на практике, как здесь, точных вероятностей различных значений мы не знаем. Это существенно отличается от наших модельных задач с кубиками, когда мы знали точные вероятности событий. Но даже если мы не знаем точного распределения, всё равно можно визуализировать распределение вероятностей различных исходов. Для этого по оси абсцисс откладывают значения случайной величины, а по оси ординат — соответствующие им вероятности. Соединение полученных точек образует ломаную линию, которая и показывает график распределения случайных величин.	

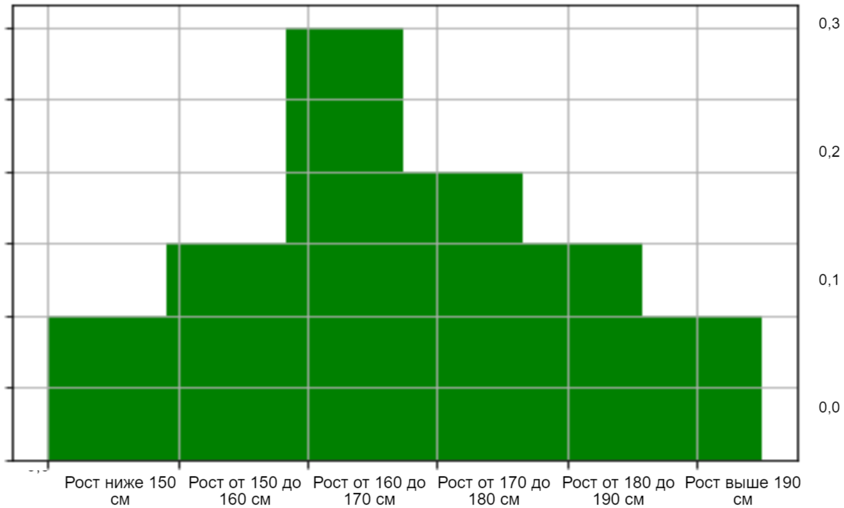
		Кроме того, можно визуализировать распределение в виде гистограммы. Столбцы здесь соответствуют различным значениям случайной величины, а высота столбца — вероятности. Законов распределения случайных величин существует очень много, некоторые из них даже имеют названия, потому что очень часто встречаются в жизни. С одним таким распределением давайте вместе познакомимся. Нормальное распределение описывает множество явлений в нашей жизни. Представьте себе распределение значений, которое опишет рост всех взрослых жителей страны. Бывают люди очень высокие, бывают низкие, но большинство людей среднего роста. Так вот, нормальное распределение — это закон, по которому большинство объектов близки к среднему значению по какому-либо признаку. При этом меньшая их часть находится либо сильно ниже, либо сильно выше этого значения. График нормального распределения вы видите на картинке. Оказывается, нормальное распределение постоянно встречается в природе. Оно появляется, если суммировать много-много различных факторов, каждый из которых вносит незначительную роль в общую сумму всех факторов. Например, наш рост определяется большим количеством генов, поэтому рост популяции нормально распределён. Давайте порассуждаем. Мы изучили понятие случайной величины и рассмотрели примеры распределений: один теоретический — распределение очков на кубике, и один практический — распределение роста людей в популяции. Как правило, в жизни у нас нет возможности узнать распределение теоретически. Поэтому единственный вариант, который мы можем себе позволить — это собрать данные в ходе эксперимента и по этим данным	

восстановить примерное распределение изучаемого признака. Так вот, набор данных, который соответствует одной и той же случайной величине, мы будем называть выборкой.
Представьте, что мы обзваниваем случайных людей и спрашиваем их рост, а затем записываем данные в таблицу. То, каких мы людей выбираем, полностью случайно, поэтому данные о первом человеке будут независимы от данных о втором человеке. Когда мы имеем дело с выборкой, мы обычно считаем, что это значения независимых друг от друга случайных величин.

Независимость случайных величин — очень важное понятие в теории вероятностей. При создании выборки обязательно нужно следить за независимостью отдельных элементов. Представьте, что мы хотим узнать распределение роста жителей России, а выбираем для этого только жителей из одного города. Понятно, что эти данные будут зависимы между собой и не будут представлять распределения роста по стране в целом. Тогда выборка получится, как говорят социологи, нерепрезентативной. Этого необходимо избегать по-настоящему случайным выбором объектов для измерения.

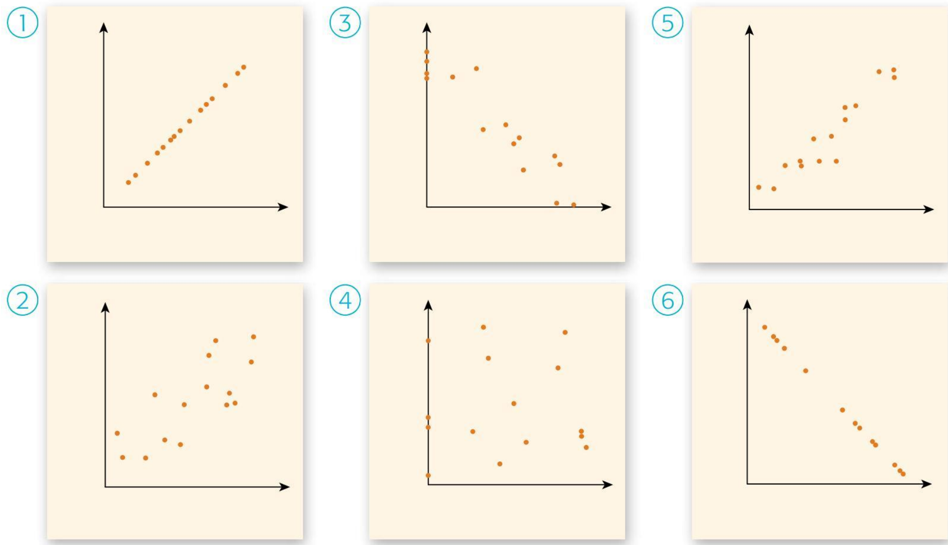
Чуть ранее мы говорили о случайных величинах. Мы сказали, что случайная величина — это численное описание результата эксперимента. Пусть наш эксперимент — бросок шестигранного кубика. Тогда случайная величина — это число, которое выпало на кубике.
Мы обсудили, что дискретная случайная величина X задаётся конечным множеством значений $x_1, x_2, \ldots, x_n$ и распределением вероятностей на этом множестве, то есть числами $p_1, p_2, \ldots, p_n$, где $p_i = P(X = x_i)$.

		Отметим, что дискретная случайная величина может принимать не только конечное, но и счётное число значений. Например, если эксперимент — это измерение количества конфет в пачке M&Ms, а случайная величина X — это, собственно, количество конфет, с математической точки зрения удобно считать, что X может принимать любые целые неотрицательные значения (0, 1, 2, 3, ...), хоть в реальности конфет в пачке вряд ли больше, например, тысячи (и точно меньше количества атомов во вселенной, то есть 10^{81}. Тем не менее, для описания некоторых экспериментов недостаточно одной случайной величины. Например, если наш эксперимент — это выбор случайного школьника из Казани, то этот эксперимент описывается сразу огромным количеством характеристик, каждая из которых будет случайной величиной: рост случайно выбранного человека (обозначим эту случайную величину через A), вес случайно выбранного человека (B), количество волос у него на голове (C), его оценка за последнюю контрольную работу (D), количество мороженых, съеденных за последний месяц (E) и так далее. Предположим, что у нас есть некоторая выборка объектов (в данном случае школьников), каждый из которых описывается этими признаками. Признаки являются случайными величинами. Если для одной случайной величины мы распределяем вероятности как одномерный график, то в многомерном случае мы уже не можем построить такой график для взаимосвязей между случайными величинами.	



Тем не менее, эти взаимосвязи присутствуют. Закон, по которому связаны случайные величины, описывающие один и тот же объект, называется многомерным распределением вероятностей.

Если признаков всего два, то мы можем визуализировать точки выборки на графике. Получится визуализация соответствующего двумерного распределения вероятностей. Вот некоторые примеры распределений. Чуть позже мы обсудим их более подробно.



Часто в задачах анализа данных нам бывает важно установить зависимость между разными характеристиками объектов. Например, мы можем задаться вопросом, влияет ли количество мороженых, съеденных за месяц, на оценку за последнюю контрольную? Если влияет, то можно было бы попытаться использовать одну переменную для предсказания другой.

Более прикладные примеры можно найти, если рассмотреть стандартные задачи, для которых часто применяют машинное обучение. Одной из таких задач является кредитный скоринг. Представьте, что вы — банк, который выдаёт кредиты клиентам. К вам приходит клиент и просит у вас кредит на определенную сумму под определённый процент. Вам необходимо принять решение:

выдавать кредит или не выдавать. Для принятия решения вам нужны данные — признаки клиента:

- месячная зарплата,
- средние месячные траты,
- количество текущих кредитов,
- количество погашенных кредитов,
- суммарное количество просрочек по предыдущим выплатам,
- наличие собственного жилья и так далее.

Ваша цель — оценить вероятность того, что клиент вернёт вам кредит. Если вероятность больше некоторого порога (например, 90%), то кредит стоит выдать, если нет — отказать в выдаче.
Для решения этой задачи в распоряжении банка есть данные о клиентах, которые уже брали кредит, для которых известно, вернули они кредит или нет. Эти данные — тоже случайная величина (назовем ее Y), которая принимает два значения:

- 0 (если кредит не вернули),
- 1 (если кредит вернули).

Наша цель при решении задачи машинного обучения — предсказать значение этой случайной величины Y. Более подробно о такой постановке задачи мы погоорим, когда будем изучать машинное обучение.

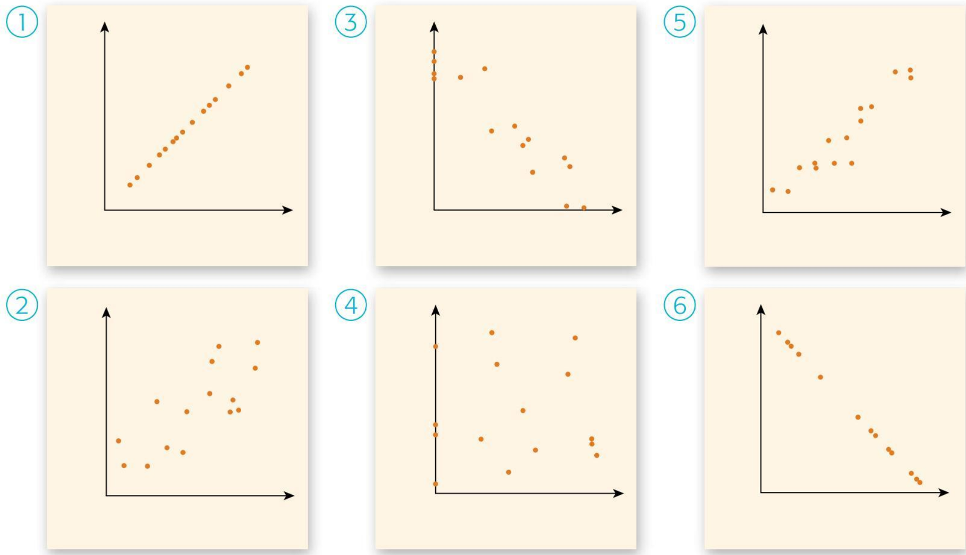
Перед решением данной задачи нужно на этапе первичного анализа данных отобрать только те признаки, которые могут помочь определить значение целевой переменной. Некоторые признаки к таковым не относятся. Например, признак «рост клиента» вряд ли поможет определить значение целевой переменной, потому что платёжеспособность клиента и его рост скорее всего никак не связаны.

		Независимость случайных величин Определение. Дискретные случайные величины X и Y называются независимыми, если для любых значений x и y, которые могут принимать X и Y, справедливо равенство $P(X = x, Y = y) = P(X = x) \cdot P(Y = y)$. Иными словами, случайные величины независимы, если для любых x и y события «$X = x$» и «$Y = y$» независимы. Интуитивный смысл определения независимости таков: если мы знаем значение случайной величины X для данного эксперимента, это не даёт нам никакой информации о значении случайной величины Y. Замечание. Независимых случайных величин в реальном мире практически не существует, потому что всё в мире взаимосвязано. Или, по крайней мере, доказать независимость случайных величин в реальном мире невозможно. Например, являются ли независимыми результаты двух бросков одного кубика? Нет, потому что некоторые внешние обстоятельства, такие как температура воздуха, высота подброса и т.д. могут повлиять на вероятность обоих бросков (напомним, что шестёрка на кубиках в среднем выпадает чаще, чем единица, потому что сторона с шестеркой легче). Тем не менее, при решении математических задач мы часто предполагаем, что некоторые случайные величины являются независимыми, поскольку с ними гораздо удобнее работать. Например, влиянием несимметричности игрального кубика пренебрегают при решении задач, потому что это влияние очень мало. Тем не менее, независимые случайные величины встречаются в идеализациях. Например, если мы независимо выберем два случайных числа m, n в промежутке от 1 до N, где N — общее количество людей на планете, а затем рассмотрим величины X и Y — рост соответствующих людей, то X и Y будут независимыми	

случайными величинами, так как людей мы выбирали независимо.

Зависимость в выборках

Как для реальных данных определить, что случайные величины независимы? Рассмотрим несколько примеров, показанных на картинке. Перед нами 6 различных выборок с данными, данные в каждой выборке характеризуются двумя признаками: X и Y, отложенными по оси абсцисс и ординат. Будем считать, что все признаки принимают только целые значения, чтобы можно было рассуждать о дискретных случайных величинах. Посмотрите на выборки и ответьте, в каком случае можно предположить, что случайные величины X и Y независимы.



		Рассмотрим данные примеры по очереди. ● На первой картинке мы имеем линейную зависимость Y от X. Более того, из картинки кажется, что X просто равен Y, но давайте считать, что $Y = aX + b$, где a, b — некоторые константы. Тогда, естественно, случайные величины X и Y не являются независимыми. В самом деле, рассмотрим некоторое число x, для которого $X = x$ с ненулевой вероятностью. Тогда события $\{X = x\}$ и $\{Y = ax + b\}$ совпадают: из одного следует другое. Поэтому, если $P(Y = ax + b) \neq 1$, то $P(X = x, Y = ax + b) = P(X = x) \neq P(X = x) \cdot P(Y = ax + b)$. Это противоречит независимости X и Y. Примером таких случайных величин может быть, например, температура воздуха, измеренная в градусах Цельсия и Фаренгейта. ● На шестой картинке мы также видим точную линейную зависимость одного признака от другого, поэтому случайные величины также не являются независимыми. В общем случае, если значение одной случайной величины полностью определяет значение другой, то они независимыми не являются. ● На второй картинке мы видим, что между признаками X и Y, по всей видимости, нет функциональной зависимости: по значению X нельзя точно определить значение Y. Тем не менее, можно в целом сказать, что чем больше значение X, тем в среднем больше и значение Y. Можно сказать, что они связаны линейной зависимостью с шумом: $Y = X + Z$, где Z — шум, то есть небольшая случайная добавка, которая может быть как положительной, так и отрицательной. X и Y также не являются независимыми: значение X даёт нам некоторую информацию о значении Y. Примером таких признаков могут быть, например, рост и вес человека: чем больше рост, тем в среднем больше вес. То же самое можно сказать про пятую и третью картинки.	

- На четвёртой картинке не наблюдается никаких зависимостей в данных. Это не означает, что эти случайные величины точно независимые, но данных недостаточно, чтобы утверждать зависимость.

Распределение суммы независимых случайных величин
Поговорим о применениях независимости случайных величин при решении задач. Пусть даны случайные величины X и Y. У каждой из них известно распределение вероятностей. Например, пусть X и Y имеет равномерное распределение на {1, 2, 3, 4, 5, 6}: $P(X = a) = P(Y = a) = ⅙$ для всех $a \in \{1, 2, 3, 4, 5, 6\}$.

Можно ли узнать распределение суммы этих двух случайных величин, используя только эту информацию? Ответ: сделать это невозможно. Чтобы продемонстрировать это, достаточно предъявить два примера, при которых получаются разные ответы.

- Пусть случайные величины X и Y — это результаты последовательных бросков одного и того же кубика. Тогда вероятность того, что сумма бросков равна 12, составляет $P(X + Y = 12) = P(X = 6, Y = 6) = P(X = 6) \cdot P(Y = 6) = ⅙ \cdot ⅙ = 1/36$.
- Пусть случайная величина X — это результат броска кубика, а случайная величина Y — результат того же броска того же кубика. Тогда X и Y всегда принимают равные значения (потому что это одна и та же случайная величина). Следовательно, вероятность того, что сумма бросков равна 12, равна $P(X + Y = 12) = P(X = 6) = ⅙$.

Значения получились различными, поэтому распределение суммы случайных величин не восстанавливается однозначно по самим распределениям случайных величин. Какая же разница между первой и второй ситуацией? Дело в том, что в первой ситуации случайные величины X и Y были независимы. Поэтому мы смогли сделать вывод о том, что P(X = 6, Y = 6) = P(X = 6) · P(Y = 6).

Итак, если случайные величины X и Y независимы, то по распределению X и Y можно вычислить распределение случайной величины Z, равной X + Y. Приведём пример, как это сделать, имея таблицы распределений X и Y.

Выпишем распределения X и Y.

X	1	2	3	4
	½	⅓	¼	1/12

Y	1	2	3	4
	⅕	1/10	½	⅕

Найдём распределение Z = X + Y. Для этого можно выписать таблицу, в клетки которой мы будем записывать вероятности P(X = x, Y = y). Поскольку случайные величины независимы, эта вероятность равна произведению вероятностей P(X = x) и P(Y = y).

	X	1	2	3	4
Y		½	⅓	¼	1/12

1	⅕	1/10	1/15	1/20	1/60
2	1/10	1/20	1/30	1/40	1/120
3	½	1/4	1/6	1/8	1/24
4	⅕	1/10	1/15	1/20	1/60

Для примера найдём вероятность того, что Z = 5. Для этого необходимо просуммировать вероятности всех возможных вариантов, как можно получить 5 из суммы двух чисел. Это выделенные клетки в таблице:

	X	1	2	3	4
Y		½	⅓	¼	1/12
1	⅕	1/10	1/15	1/20	1/60
2	1/10	1/20	1/30	1/40	1/120

3	½	1/4	1/6	1/8	1/24
4	⅕	1/10	1/15	1/20	1/60

Таким образом,
$P(Z = 5) = P(X = 1, Y = 4) + P(X = 2, Y = 3) + P(X = 3, Y = 2) + P(X = 4, Y = 1) =$
$= 1/10 + 1/6 + 1/40 + 1/60 = 37/120$.
Итоговая таблица для распределения суммы:

Z	2	3	4	5	6	7	8
	1/10	7/60	1/3	37/120	7/60	11/120	1/60

Аналогичным образом можно находить распределения любых функций от двух независимых случайных величин, например, разности, произведения, частного.

Ранее мы научились задавать случайную величину с помощью её распределения. Но распределение — это слишком подробная характеристика случайной величины или выборки. Стоит ввести некоторые дополнительные показатели распределения, которые будут описывать его свойства. Такими показателями являются среднее значение распределения и типичное отклонение точки от среднего значения. На картинке среднее значение изображено красной точкой посередине купола, а типичное отклонение —

		двусторонней стрелкой. Итак, начинаем с характеристик среднего. Если нам вдруг понадобилось описать распределение случайной величины одним-единственным числом, то, скорее всего, это число будет средним значением случайной величины. Введём определение. Дана случайная величина X, которая принимает значения x_1, ..., x_n с некоторыми вероятностями. Тогда математическим ожиданием случайной величины называется вот такая сумма: x_1 умножить на вероятность того, что X равно x_1 плюс и так далее плюс x_n умножить на вероятность того, что X равно x_n. Обозначается математическое ожидание буквой E, от английского слова expectation. В советских учебниках иногда матожидание обозначалось буквой M, но такое обозначение постепенно уходит из употребления. У нас получается взвешенная сумма всех возможных значений случайной величины: каждое значение входит в сумму с соответствующей вероятностью. На самом деле интуиция математического ожидания простая — это среднее значение нашей случайной величины. Представьте, например, что вероятности всех иксов равны между собой. Тогда математическое ожидание получается равным среднему арифметическому всех возможных значений нашей случайной величины. Теперь пусть у нас вместо случайной величины есть выборка с данными X_1, ..., X_N — например, мы измерили рост у N различных людей. Тогда, если мы посчитаем среднее арифметическое значение	

роста, то с ростом количества людей среднее арифметическое будет всё более близко подходить к математическому ожиданию случайной величины X.

Укажем важные свойства математического ожидания. Рассмотрим всего два свойства. Первое: из-под знака матожидания можно выносить константу. В самом деле, если сначала домножить на константу, а потом усреднить, получится то же самое, как если сначала усреднить, а потом домножить на константу.

Следующее важное свойство математического ожидания — линейность. Пусть у нас есть две произвольные случайные величины X и Y. Тогда оказывается, что математическое ожидание суммы X плюс Y равно сумме математических ожиданий.

Представьте, что X — это длина до пояса, а Y — длина тела выше пояса. Тогда как посчитать среднее значение роста человека? Можно сложить среднее значение длины до пояса и среднее значение длины выше пояса. По сути в этом и заключается правило линейности.

Третье и последнее свойство: математическое ожидание константы равно константе. Если случайности нет, то среднее значение равно самому числу.

Решим простую задачку на тему математического ожидания. Условие задачи: бросаются 5 шестигранных кубиков. Найдите математическое ожидание суммы выпавших очков. Поставьте видео на паузу и попробуйте решить задачу самостоятельно, воспользовавшись свойством линейности математического ожидания.

Решение. Пусть X_1, X_2,..., X_5 — это случайные величины, равные количеству очков, выпавших на первом, втором, третьем, четвёртом и пятом кубиках. Наша случайная величина X представляет собой сумму $X_1 + X_2 + \ldots + X_5$. Тогда, по свойству линейности, математическое ожидание X равно сумме математических ожиданий X_1, X_2, ..., X_5. Поскольку кубики ничем не отличаются, достаточно посчитать математическое ожидание X_1, а затем умножить его на 5. Это будет ответом на нашу задачу.

Считаем математическое ожидание X_1. X_1 принимает всего 6 значений: 1, 2, 3, 4, 5, 6, все с одной и той же вероятностью $\frac{1}{6}$. Тогда, по определению, математическое ожидание X_1 равно $\frac{1}{6}$ умножить на 1 + $\frac{1}{6}$ умножить на 2 плюс и так далее плюс $\frac{1}{6}$ умножить на 6. Ясно, что это просто равно 3 с половиной, потому что можно усреднить по парам: 1+ 6 пополам это 3,5, аналогично 2 + 5 пополам и 3+4 пополам.

Можно записать итоговый ответ. Он равен 5 умножить на 3,5 равно 17,5. Задача решена.

В решённой задаче всё было просто: из условия было известно распределение случайной величины, у которой мы хотели посчитать математическое ожидание. Но в жизни всё, конечно же, не так. На практике у дата сайентиста обычно есть некоторая выборка с данными.

Если x_1, ... X_n — это наша выборка, то по ней можно посчитать среднее значение икса, равное $X_1 + \ldots + X_n$ делить на n. Среднее значение некоторой величины мы будем обозначать чертой сверху над выражением. Например, среднее значение икса будем обозначать как X с чертой.
Среднее значение выборки обычно называют выборочным средним,

		чтобы подчеркнуть, что его посчитали по выборке, а не теоретически, по распределению случайной величины. Обратите внимание, что с ростом n выборочное среднее, конечно же, стремится к истинному математическому ожиданию. С математическим ожиданием и средним значением всё более-менее понятно. Впрочем, с ними есть одна важная проблема, которую мы тоже кратко обсудим. Дело в том, что в данных могут быть значения, которые сильно отличаются от остальных. Например, в стране может быть житель, который играет в баскетбол, и его рост будет 253 см, или житель с очень низким ростом в 54 см. Это может значительно сдвинуть среднее значение. Поэтому такие выбросы могут сделать среднее значение нерепрезентативным. Если вы знаете, что в данных есть выбросы, тогда вместо среднего арифметического лучше посчитать медиану. Медиана — это число, которое делит упорядоченный ряд значений пополам. Значения в одной половине меньше медианы, а в другой — больше. Медиана и среднее значение могут совпадать только если нам очень повезёт. В примере на картинке математическое ожидание сильно смещено именно из-за выбросов, в отличие от медианы, которая остаётся стабильной. Поговорим о характеристиках разброса выборки и разброса случайной величины. Сначала обсудим характеристику разброса выборки под названием выборочная дисперсия. Что такое разброс? Разброс — это некоторое число, которое определяет, насколько сильно элементы нашей выборки отклоняются от среднего значения в нашей	

выборке.
Итак, пусть дана выборка $X_1, X_2, \dots, X_n$. Давайте визуализируем её на картинке. Красной вертикальной чертой отметим среднее значение нашей выборки.

Рассмотрим какую-нибудь точку выборки X_k и посмотрим на величину $X_k - a$. Отметим её на картинке. Отклонение X_k от среднего равно $X_k - a$. В данном случае выражение будет отрицательным, поэтому, строго говоря, отклонение будет равно модулю $X_k - a$.

Но сейчас мы эту отрицательность исправим. Давайте рассмотрим среднее значение выражения $(X - a)$ в квадрате по всем элементам нашей выборки. Эта величина называется выборочной дисперсией и обозначается как D большое с крышечкой от X. Крышечка здесь нужна именно для того, чтобы подчеркнуть, что выборочная дисперсия является характеристикой выборки, а не случайной величины. Впрочем, на практике дисперсией часто называют выборочную дисперсию, но в этом видео мы постараемся так не делать.

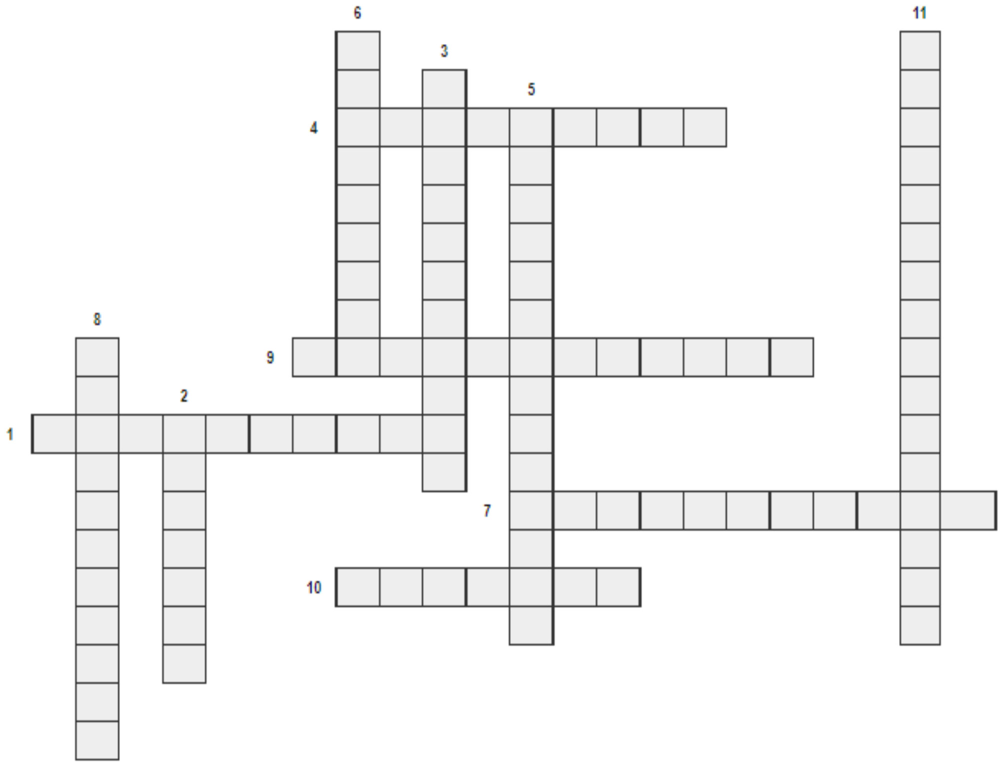
Выборочную дисперсию можно записать как среднее значение от квадрата икс минус среднее значение икс. Получается выражение с двумя вложенными чертами на слайде. Поначалу выглядит страшно, но к этому быстро привыкаешь. Вторая черта означает, что мы усредняем квадрат отклонения. Можно расписать это по всем элементам выборки, как это сделано внизу: выборочная дисперсия равна сумме по всем элементам выборки X_i − X среднее в квадрате, делить на n.

У вас наверняка возник вопрос, почему мы возводим именно в квадрат. Действительно, вместо квадрата мы могли бы брать модуль

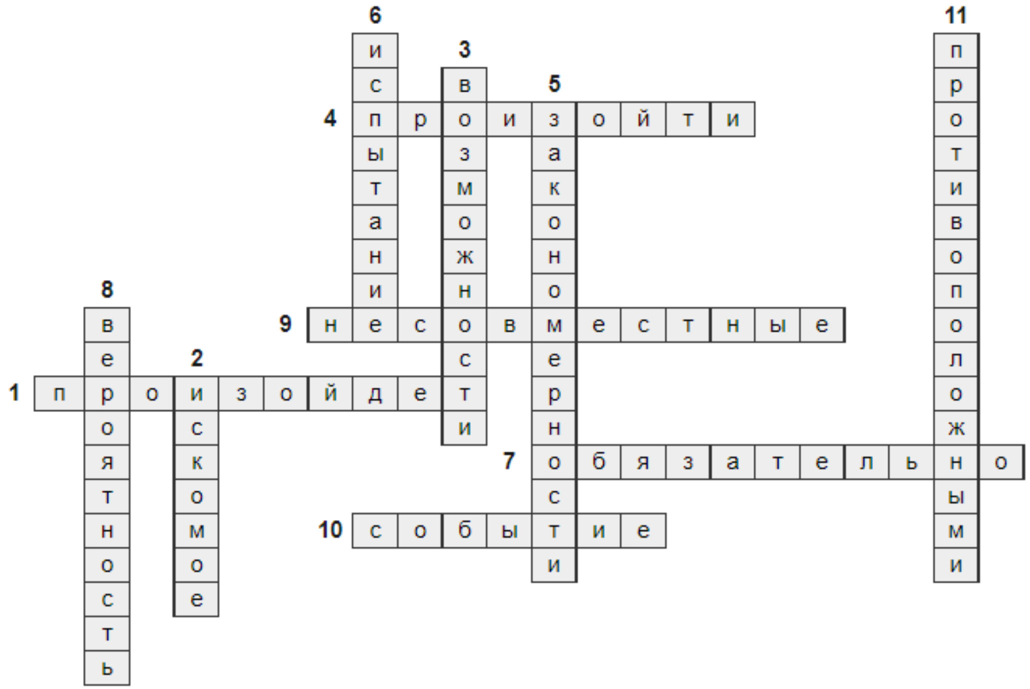
		отклонения, но тогда дисперсия бы не обладала некоторыми замечательными теоретическими свойствами, которые мы с вами ещё обсудим. Квадрат — функция гораздо круче, чем модуль, сравните хотя бы сложность решения квадратных уравнений и решения уравнения с модулями. Второе делать гораздо муторнее, нужно разбирать какие-то случаи, а с квадратными уравнением проще — написал дискриминант и всё. Вернёмся к нашим вычислениям. У нас есть проблема. Дело в том, что получившееся выражение для выборочной дисперсии имеет странную единицу измерения. Пусть изначальная величина X измеряется в сантиметрах. Выходит, дисперсия должна измеряться в сантиметрах квадратных. И как тогда рисовать её на нашей картинке, если она в других единицах измерения? Нет, так не пойдёт. Для удобства из выборочной дисперсии часто извлекают корень. Получившееся выражение называют стандартным отклонением и обозначают std или буквой сигма. Стандартное отклонение имеет ту же размерность, что и изначальные элементы выборки. Посчитаем стандартное отклонение на примере выборки значений роста некоторых людей. Всего 10 человек. Перед вами таблица этой выборки. Сначала посчитаем среднее значение роста. Выписывать его не будем: поверьте мне на слово, что оно равно 175,3 см. Выписываем таблицу с отклонениями от среднего. Для этого из значения роста вычитаем среднее значение, получается вот такой набор отклонений.	

		Наконец, усредняем квадраты отклонений из таблицы и получаем исходное значение выборочной дисперсии, равное 169,81. Стандартное отклонение, равное корню из дисперсии, оказывается примерно равным 13,3 см. Получается, в среднем рост человека отличается от среднего на 13 см. Давайте ещё раз изобразим это на картинке. Точки выборки с ростом отметим синим, красным проведём линию среднего значения. Теперь укажем стрелку в обе стороны от среднего, по длине равную сигма. Мы видим, что большая часть точек этой стрелкой покрывается. Вывод, который мы делаем: большая часть значений отклонения, как правило, попадает в диапазон одной сигмы от среднего значения. То есть я с довольно высокой уверенностью могу сказать, что ваш рост, уважаемый слушатель, находится в границе 175 плюс минус 13 см. К этому наблюдению мы ещё вернёмся. Давайте введём ту же самую дисперсию, только теперь не для выборки, а для случайной величины. Пусть X — наша случайная величина, a — математическое ожидание икса. Тогда X − a — это случайная величина, равная размеру отклонения случайной величины от центра. Наконец, дисперсией мы будем называть математическое ожидание квадрата этого отклонения, то есть, записывая это выражение полностью, выходит математическое ожидание от X − EX в квадрате. Дисперсия аналогична среднему квадратичному отклонению выборки.	

		Сравните эти два выражения: они полностью одинаковы, но верхняя черта заменяется знаком математического ожидания. Отличие в том, что дисперсию можно вычислить именно для случайной величины, а выборочную дисперсию — для выборки. Теперь давайте вернёмся к нормальному распределению, которое мы рассматривали, когда говорили о распределении вероятностей. У нормального распределения математическое ожидание будет в центре купола распределения, а стандартное отклонение будет определять ширину купола. Стандартное отклонение на картинке изображено стрелочкой. Давайте обозначим матожидание буквой a, а стандартное отклонение — буквой сигма. Интересный факт заключается в том, что, оказывается, если мы отложим вправо-влево от среднего значения величину, равную утроенному значению сигма, то в полученный диапазон, обозначенный на картинке длинной стрелочкой, наша случайная величина будет попадать с огромной вероятностью: больше 99,7 процентов. Таким образом, если у вас есть выборка из нормального распределения, то вы можете с уверенностью утверждать, что за пределы трёх сигм ваше значение не выйдет. Это оказывается очень полезно на практике. Например, при оценке всевозможных рисков. Если риски распределены нормально, то мы можем контролировать вероятность появления слишком больших рисков.	
Закрепление изученного материала	15 мин.	**Предлагаю самостоятельно решить задачи:** 1. Случайная величина X имеет константное распределение, то есть P(X = C) = 1 для некоторой константы C. Докажите, что X независима с любой случайной величиной Y.	Обучающиеся решают задачи

		2. Случайные величины X и Y независимы. Докажите, что случайные величины X^2 и Y^2 также независимы. Замечание. На самом деле, утверждение верно для случайных величин $f(X)$ и $g(Y)$ для произвольных функций f и g.  	

Этап подведения итогов занятия (рефлексия)	8 мин.	Давайте порешаем кроссворд по теме: «Основы теории вероятностей». По вертикали: 2. Единственное событие из всех возможных событий, которым вы интересуетесь, это ___________ событие. 3. Равновозможные события - это события, которые имеют равные ____________ . 5.Теория вероятностей — раздел математики, в котором изучаются случайные явления (события) и выявляются _________ при массовом их повторении. 6. Всякое действие с различными исходами, которые реализуются при данном комплексе условий ________ . 8. P(A) означает ____________ события. 11. Полная система событий может состоять из двух событий, называемых _____________ . По горизонтали: 1. Полная система событий — это совокупность несовместных событий, наступление хотя бы одного в данном опыте _____________ . 4. Невозможное событие, если оно в данном опыте не может ____________ . 7. Достоверное событие, если оно происходит в данном опыте __________________ . 9. Если никакие два события не могут произойти в данном опыте вместе, то они _____________ . 10.Результат испытания — это случайное __________ .	Педагог способствует размышлению обучающихся над заполнением кроссворда

			
Информация о домашнем задании, инструктаж по его применению	5 мин.	Дома повторите основные определения.	

Рекомендуемые ресурсы для дополнительного изучения:

1. Начала статистики. [Электронный ресурс] – Режим доступа: https://resh.edu.ru/subject/lesson/3751/conspect/326747/.
2. Теории вероятностей. [Электронный ресурс] – Режим доступа: https://habr.com/ru/company/JetBrains-education/blog/498188/.
3. Теории вероятностей. [Электронный ресурс] – Режим доступа: https://mipt.ru/online/hi-Math/teorver/.