Четвертая технологическая революция строится на вездесущем и мобильном Интернете, искусственном интеллекте и машинном обучении.

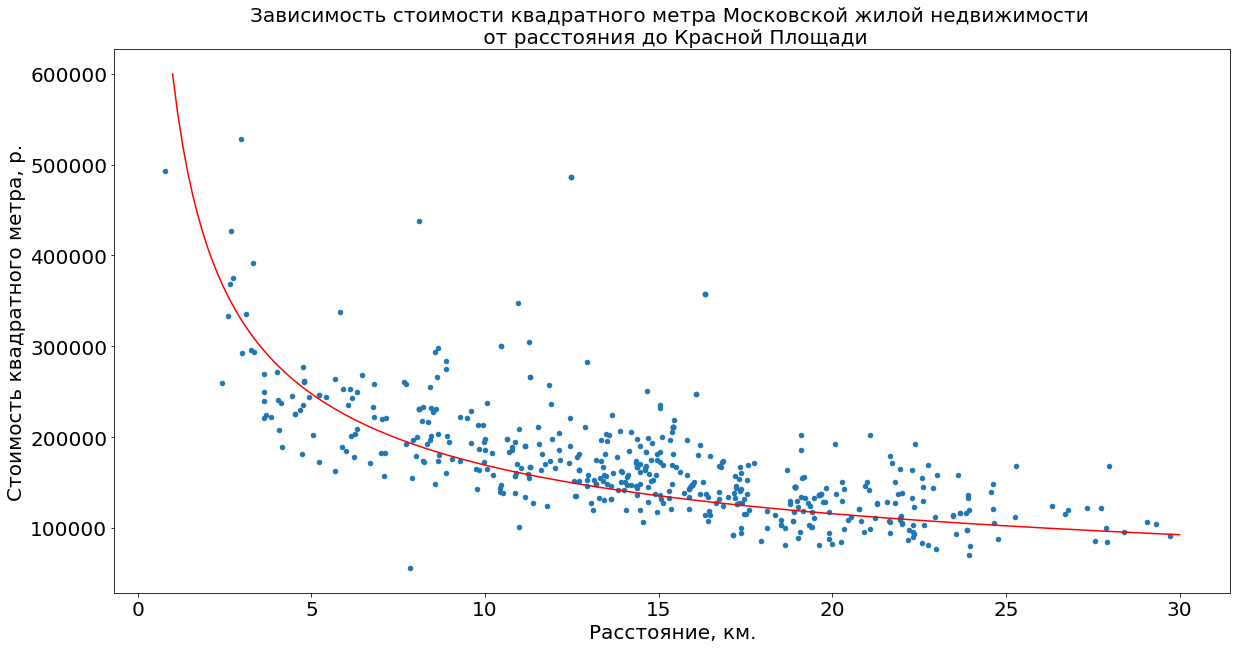
Президент Всемирного экономического форума

Клаус Мартин Шваб

Нации уделяющие значительное внимание научно-исследовательским и опытно-конструкторским работам в области машинного обучения займут лидирующие позиции в автоматизации будущего. Основными сферами применения достижений исследовательских разработок станут:

* Цифровая и распределенная экономика
* Автоматизация и сокращение издержек
* Автономный транспорт и роботизация
* Оптимизация логистики и цепей поставок
* Оптимизация энергетических сетей
* Автоматизация банковских услуг
* Автоматизация юридических услуг
* Мониторинг сельского хозяйства
* Персональная медицина
* Персональные образовательные траектории
* Автономные системы вооружений.

**“**[**Preparing** **for** **the** **future** **of** **artificial** **intelligence**](https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf)”, Доклад начальника штаба президента США по национальной науке и технике, Вашингтон, 2016.



## Введение

## Машинное обучение применяется всюду, где собираются данные. По сути машинное обучение, это наука о том, как восстановить неизвестную зависимость по экспериментальным данным, как провести функцию по точкам . На рисунке выше приведена экспериментальная зависимость стоимости квадратного метра Московской жилой недвижимости в зависимости от одного фактора – расстояния до Красной Площади. 500 синих точек на графике – 500 московских квартир. Красная кривая – восстановленная зависимость. Математические методы решающие подобные задачи известны не одну сотню лет.

## Машинное обучение – развитие этих методов для случая, когда точки принадлежат многомерным пространствам, а ответы принимают не только числовые значения. Задача определения стоимости квадратного метра квартиры в зависимости от целого набора её признаков (года постройки дома, района города, этажа, площади квартиры и отдельных комнат, качества ремонта и т.д.) – задача машинного обучения. Иными словами, методы машинного обучения позволяют по набору известных признаков некоторого объекта предсказать значение некоторого неизвестного признака этого объекта. Методы машинного обучения позволяют построить алгоритм оценки стоимости квадратного метра, «цифрового риелтора», который «осмотрев» квартиру выносит вердикт о её стоимости.

## Различные особенности множества признаков и множества допустимых значений восстанавливаемой функции привели к появлению массы методов машинного обучения.

## Основные понятия и определения.

Попытаемся формализовать постановку задачи машинного обучения.

Пусть нам задано множество *объектов* , множество *допустимых ответов* и существует *целевая функция* (target function) , значения которой известны только на конечном подмножестве объектов . Совокупность пар называется *обучающей выборкой* (training sample).

Задача машинного обучения заключается в том, чтобы по выборке *восстановить зависимость* , то есть построить *решающую функцию* (decision function) , которая приближала бы целевую функцию , причем не только на объектах обучающей выборки, но и на всём множестве .

Решающая функция должна допускать эффективную компьютерную реализацию; по этой причине её называют *алгоритмом.*

## Объекты и признаки

*Признак (feature)*  объекта - это результат измерения некоторой характеристики объекта. Формально признаком называется отображение где множество допустимых значений признака. В зависимости от природы множества признаки делятся на несколько типов:

Если , то *бинарный* признак, например наличие балкона у квартиры;

Если конечное множество, то категориальный признак, например район города (Центральный, Сокольники, Тушинский …);

Если , то вещественный признак, например жилая площадь квартиры, количество комнат, год постройки дома

Пусть имеется набор признаков , тогда вектор составленный из значений признаков ) называют признаковым описанием объекта . Совокупность признаковых описаний всех объектов выборки , записанную в виде таблицы размера , назваются *матрицей объектов-признаков:*

*Например, если предположить, что наличие балкона, площадь квартиры, количество комнат, расстояние до Красной Площади, то некоторые три квартиры могли бы быть описаны следующей матрицей объектов-признаков.*

## Ответы и типы задач.

## В зависимости от природы множества допустимых ответов Y задачи машинного обучения делятся на следующие типы. Если , то это задача классификации (classification) на M непересекающихся классов. В этом случае всё множество объектов X разбивается на классы, и алгоритм a(x) должен давать ответ на вопрос «какому классу принадлежит ?». В некоторых приложениях классы называют образами и говорят о задаче распознавания образов (pattern recognition). Если , то это задача классификации на M пересекающихся классов (к примеру классификация новостей по темам: новость «Путин провел двустороннюю встречу с представителями ЦК КНДР» может быть отнесена как к теме «Президент России», так и к теме «Китай»). В простейшем случае эта задача сводится к решению M независимых задач классификации с двумя непересекающимися классами. Если , то это задача восстановления регрессии (regression estimation). Задачи прогнозирования (forecasting), когда x ∈ X — описание прошлого поведения объекта x, y ∈ Y — описание некоторых характеристик его будущего поведения, являются частными случаями классификации или восстановления регрессии.

## Алгоритмы и их обучение.

Одним из простейших алгоритмов машинного обучения являются *линейные алгоритмы* с набором параметров :

Процесс подбора оптимальных параметров по обучающий выборке , позволяющих решить поставленную задачу с наилучшим качеством, называют *обучением* (training, learning) алгоритма .

В случае задачи предсказания стоимости квадратного метра жилья на основе его признакового описания линейный алгоритм будет выглядеть так:

Итак, в задачах машинного обучения чётко различаются два этапа:

* этапе *обучения* в процессе которого по выборке строит алгоритм
* этапе *применения (*prediction*)* для новых объектов алгоритм выдает ответы .

Этап обучения как правило сводится к поиску параметров модели, доставляющих оптимальное значение заданному функционалу качества.

## Функционал качества

*Функция потерь* (loss function) – это неотрицательная функция , характеризующая величину ошибки алгоритма на объекте . Если , то ответ называется корректным.

Функционал качества алгоритма на выборке :

Функционал называют также функционалом *средних потерь*. Пример функционала средних потерь с квадратичной функцией ошибки, который может применяться в задаче оценки стоимости квадратного метра:

Классический *метод обучения*, называемый *минимизацией средних потерь*, заключается в том, чтобы найти набор параметров алгоритм , доставляющий минимальное значение функционалу на заданной обучающей выборке :

## Вопросы по прочитанному.

1. Что такое объект?

То, для чего нужно делать предсказания.

То, с помощью чего измеряется качество предсказаний.

То, что необходимо предсказывать.

1. Что такое признаки?

То, с помощью чего измеряется качество предсказаний.

То, с помощью чего описываются ответы.

То, с помощью чего описываются объекты.

1. Что такое алгоритм?

Функция, которая принимает на вход объекты и выдаёт подсчитанные для них признаки.

Функция, которая принимает на вход предсказания на обучающей выборке и выдаёт оценку качества этих предсказаний.

Функция, которая принимает на вход объект и выдаёт предсказанный ответ.

1. Выберите верные утверждения про признаки.

Набор значений признаков на объекте представляет собой вектор определённой размерности.

**Правильно**

Полный набор признаков для объекта удобно задавать с помощью вектора, размерность которого совпадает с числом признаков.

Признаки задаются только вещественными числами.

**правильно, этот вариант не должен быть выбран**

Предсказания для объектов делаются на основе значений признаков.

**Правильно**

Алгоритмы принимают на вход именно признаковые описания объектов.

Признаки могут иметь только значения 0 или 1.

**Правильно, этот вариант не должен быть выбран**

1. Выберите вещественные признаки из списка

Количество детей

**Правильно**

Наличие у клиента банка военного билета

**правильно, этот вариант не должен быть выбран**

Город, в котором прописан клиент

**правильно, этот вариант не должен быть выбран**

Температура воздуха

**Правильно**

Год рождения

**Правильно**

1. Выберите признаки, которые могут рассматриваться только как категориальные (и не могут рассматриваться как бинарные или вещественные).

Тарифный план клиента мобильного оператора

**Правильно**

Тарифных планов много, и вряд ли все из них можно сравнивать.

Наличие у клиента банка военного билета

**правильно, этот вариант не должен быть выбран**

Цвет автомобиля

**Правильно**

Цветов бывает больше двух.

Тип дома — кирпичный, блочный, панельный и т.д.

**Правильно**

Тут конечное число вариантов и уже приведено больше двух.

Скорость интернет-соединения

**правильно, этот вариант не должен быть выбран**

Год рождения

**правильно, этот вариант не должен быть выбран**

1. Выберите из списка задачи классификации.

Определение возраста человека по его активности в сети.

**правильно, этот вариант не должен быть выбран**

Поиск групп схожих пользователей мобильного оператора.

**правильно, этот вариант не должен быть выбран**

Предсказание тарифного плана, который клиент захочет себе подключить.

**Правильно**

Возможных тарифных планов конечное число, поэтому задача относится к задачам многоклассовой классификации.

Определение стоимости одежды по фотографии.

**правильно, этот вариант не должен быть выбран**

1. Выберите верные утверждения про функционал качества

Функционал качества определяет уровень шума в признаках.

Функционал качества позволяет определить, насколько данный алгоритм подходит для решения задачи на конкретной выборке.

## Вопросы на общую математическую подготовку

1. Найдите производную при :









1. Сколько решений у следующей системы линейных уравнений?

2*x*=3

2*x*=4

Ни одного

**Правильно**

Нет такого x, чтобы он одновременно был равен 1.5 и 2.

Одно

Конечное число, больше единицы

Бесконечно много

1. Что такое ранг матрицы?

Максимальное число линейно независимых строк.

**Правильно**

Это утверждение верное.

Максимальное число линейно независимых столбцов.

**Правильно**

Это утверждение верное.

Максимальное число различных элементов.

**правильно, этот вариант не должен быть выбран**

Число решений системы линейных уравнений с данной матрицей коэффициентов и нулевой правой частью.

**правильно, этот вариант не должен быть выбран**

1. Предположим, что в некоторой популяции до 60 лет доживает 50%, а до 80 лет — 20%. Какова вероятность (от 0 до 1), что случайно выбранный шестидесятилетний представитель популяции доживёт до восьмидесяти? Запишите ответ с точностью до одного знака после десятичной точки.



**Правильный ответ** По формуле условной вероятности 0.2/0.5 ​= 0.4.

1. В супермаркете 60% яблок из Турции и 40% яблок из Индии. 10% турецких и 15% индийских яблок — червивые. Какова вероятность, что яблоко, купленное в этом магазине, окажется червивым? Запишите ответ от 0 до 1 с двумя знаками после десятичной точки.



**Правильный ответ**

Пусть событие A происходит, если яблоко из Турции, а событие B — если яблоко червивое. Воспользуемся формулой полной вероятности:

𝖯(B)=𝖯(B|A)𝖯(A)+𝖯(B|~A)𝖯(~A)=0.1⋅0.6+0.15⋅0.4=0.12

1. Журнал "COSMOPOLITAN" читает 6.4% целевой аудитории вашего продукта, а журнал "Караван историй" — 3.7%. Если вы разместите рекламу в обоих журналах, что можно сказать о доле *p* целевой аудитории, которая увидит рекламу хотя бы один раз?

p≤0.037

0.037≤*p*≤0.064

0.037≤*p*≤0.101

0.064≤*p*≤0.101

## Вопросы для аналитиков и специалистов по машинному обучению.

1. Какой из перечисленных ниже функционалов качества используются в задачах классификации и является дифференцируемым?









1. Почему в градиентном спуске на каждой итерации делается шаг в сторону антиградиента?

Антиградиент функционала ошибки зависит только от одного объекта.

Антиградиент легко найти, в отличие от других направлений.

Антиградиент совпадает с направлением наискорейшего убывания.

**Правильно**

Таким образом ошибка на выборке на каждой итерации минимизируется.

1. Градиент какой функции/функционала и по какому аргументу используется в градиентном спуске при обучении линейной регрессии?

Алгоритма — то есть скалярного произведение вектора признаков —по вектору весов.

Функционала ошибки — например, среднеквадратичной ошибки — по вектору весов.

**Правильно**

Такой градиент показывает в какую сторону нужно сдвинуть веса, чтобы уменьшить ошибку на выборке.

Функционала ошибки — например, среднеквадратичной ошибки — по прогнозам алгоритма.

1. В чём заключается отличие градиентного спуска от стохастического градиентного спуска?

В стохастическом градиентном спуске на каждой итерации делается шаг в случайном направлении.

В стохастическом градиентном спуске на каждой итерации используется лишь одно слагаемое в функционале ошибки.

В стохастическом градиентном спуске на каждой итерации к антиградиенту добавляется нормальный шум.

1. Для чего при обучении линейных классификаторов используются верхние оценки на пороговую функцию потерь?

Чтобы перейти к функции потерь, на которой градиентный спуск будет быстрее сходиться.

 Чтобы перейти к функции потерь, которая будет легко вычислима — без этого градиентная оптимизация будет занимать слишком много времени.

Чтобы заменить разрывную пороговую функцию потерь на гладкую функцию — без этого градиентная оптимизация будет невозможна.

1. Предположим, вы решаете задачу восстановления регрессии с использованием некоторого функционала ошибки . Как будет выглядеть новый функционал при добавлении ​-регуляризатора с коэффициентом регуляризации ?



**Правильно**

Добавочное слагаемое является квадратом *L*2​ нормы весов линейной модели.







1. Метод опорных векторов (SVM) — это линейный классификатор, использующий:

Кусочно-линейную функцию потерь (hinge loss) и *L*2​ регуляризатор.

**Правильно**

Логистическую функцию потерь и *L*1​ регуляризатор.

Экспоненциальную функцию потерь и *L*2​ регуляризатор.

Квадратичную функцию потерь и *L*1​ регуляризатор.

Логистическую функцию потерь и *L*2​ регуляризатор.

1. Выберите предположения, выполнение которых необходимо и достаточно для того, чтобы метод наименьших квадратов давал несмещённые и состоятельные оценки истинных коэффициентов регрессии.

Полнота ранга *X*: ни один из признаков не является линейной комбинацией других признаков.

**Правильно**

Линейность отклика:  y=*Xβ*+*ε*

**Правильно**

Гомоскедастичность ошибок

**правильно, этот вариант не должен быть выбран**



Нормальность ошибок

**правильно, этот вариант не должен быть выбран**



Случайность ошибок:

**Правильно**



Случайность выборки: наблюдения (*xi*​,*yi*​) независимы

**Правильно**

1. Предположим, вы оцениваете качество работы алгоритма при помощи кросс-валидации с разбиением на *k* блоков. Сколько раз будет проведено обучение модели?









1. К какому из указанных чисел будет ближе значение метрики AUC-ROC для алгоритма, возвращающего случайный ответ для любого объекта?

Невозможно указать точное значение без информации о выборке.

0

0.33

0.5

1

1. Зачем нужно масштабировать признаки перед обучением линейной модели?

Линейная модель не будет иметь смысла при обучении на признаках с разным масштабом.

**правильно, этот вариант не должен быть выбран**

Градиентный спуск может очень медленно сходиться при разном масштабе признаков.

**Правильно**

Благодаря масштабированию выборка будет занимать меньше места в оперативной памяти, что позволит ускорить процесс обучения.

1. Выберите верные утверждения про бинарное кодирование категориальных признаков.

Бинарное кодирование категориальных признаков является способом регуляризации линейных моделей.

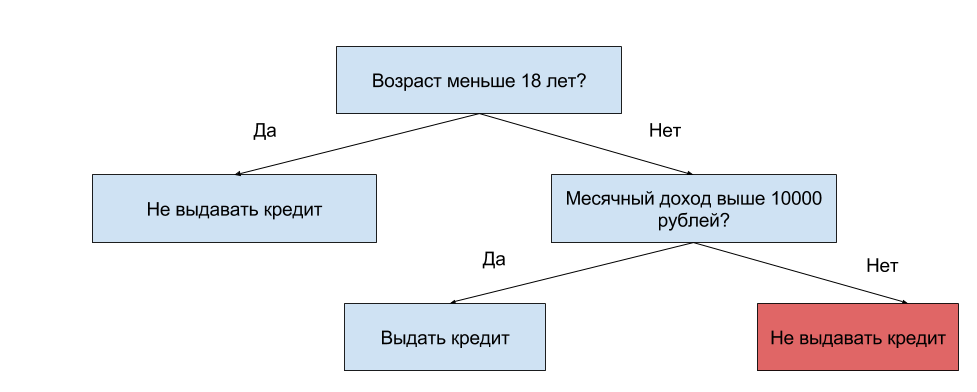
**правильно, этот вариант не должен быть выбран**

Если обучить над бинарным кодированием линейную модель, то получится, что каждому значению исходного категориального признака будет соответствовать свой вес.

**Правильно**

Для кодирования требуется столько бинарных признаков, сколько значений мог принимать исходный категориальный признак.

1. Рассмотрим решающее дерево для решения задачи кредитного скоринга, которое основано на двух признаках: возрасте и месячном доходе.



Пусть дана следующая выборка из пяти объектов (первый признак — возраст, второй — месячный доход):

[20, 8000], [15, 15000], [28, 9500], [24, 30000], [30, 20000]

Для скольки из них будет принято положительное решение о выдаче кредита?



1. Как в общем устроен процесс построения решающего дерева?

Жадно — начинаем с одной вершины, разбиваем её на две, после чего рекурсивно повторяем процедуру для новых дочерних вершин.

**Правильно**

Верно.

Жадно — начинаем с дерева, у которого в каждом листе находится по одному объекту, и удаляем из него вершины, пока улучшается качество.

Полным перебором — вычисляем качество каждого возможного дерева, выбираем лучшее.

Аналитически — можно в явном виде выписать формулы, задающие структуру оптимального дерева.

1. Мы пытаемся найти лучшее разбиение вершины *m*, и хотим оценить качество конкретного способа, который разобьёт вершину *m* на вершины *l* и *r*. Обозначим через ​∣*Xm*​∣,∣*Xl*​∣,∣*Xr*​∣ количество объектов в вершинах *m*, *l* и *r* соответственно, через H(X) *-*  значение критерия информативности на выборке *X*.

По какой из формул следует вычислять ошибку такого разбиения?









1. Можно ли решать задачу регрессии с помощью решающих деревьев?

Да, можно — достаточно лишь выбрать критерий информативности, оценивающий разброс вещественных ответов.

Нет, нельзя — деревья могут выдавать столько различных ответов, сколько в дереве листьев, то есть конечное число. А в задаче регрессии бесконечно много возможных ответов.

Нет, нельзя — критерии информативности зависят от распределения объектов по классам, такие распределения нельзя построить в задачах регрессии.

1. Выберите верные утверждения про обучение случайного леса.

Каждое дерево обучается по случайной подвыборке признаков.

Как правило, строятся деревья небольшой глубины, поскольку этого достаточно для восстановления сложных зависимостей.

Каждое дерево обучается по случайной подвыборке объектов.

Каждое дерево обучается независимо от остальных деревьев в композиции.

В каждой вершине оптимальный признак для разбиения выбирается из случайного подмножества признаков.

1. Какие величины предсказывает *N*-й базовый алгоритм в градиентном бустинге?

Производные (с минусом) функции потерь, вычисленные в точках, соответствующих ответам композиции  на обучающей выборке.

**Правильно**

Разницу между истинными ответами и ответами композиции  на обучающей выборке.

Производные (с минусом) функции потерь, вычисленные в точках, соответствующих ответам предыдущего базового алгоритма  на обучающей выборке.

1. Чем градиентный бустинг отличается от случайного леса?

Базовые алгоритмы, как правило, выбираются достаточно простыми — например, это могут быть неглубокие деревья.

Градиентный бустинг может строить алгоритмы только для задач регрессии.

Каждый следующий алгоритм в градиентном бустинге обучается так, чтобы исправить ошибки предыдущих базовых алгоритмов.

1. В задаче кластеризации мы пытаемся:

Восстановить отображение вектора признаков в метку класса по набору известных пар (признаки, метка).

Восстановить отображение вектора признаков в действительное число по набору известных пар (признаки, метка).

Сгруппировать похожие объекты (близкие в пространстве признаков) в кластеры, поместив при этом непохожие друг на друга объекты в разные кластеры.

1. Какие параметры есть у метода *k* ближайших соседей?

Число деревьев

**правильно, этот вариант не должен быть выбран**

Число соседей *k*

**Правильно**

Функция активации

**правильно, этот вариант не должен быть выбран**

Функция весов объектов

**Правильно**

1. Какие из перечисленных алгоритмов нуждаются в задании количества кластеров?

DBSCAN

k-Means

EM-алгоритм

1. С помощью какого алгоритма можно разделять смеси распределений?

Алгоритм K-средних

EM-алгоритм

**Правильно**

Верно.

Алгоритм Прима

1. Выберите верные утверждения про метод главных компонент

Метод главных компонент — это метод понижения размерности с потерями, если число новых признаков *d* меньше ранга исходной матрицы "объекты-признаки".

Метод главных компонент является линейным методом понижение размерности — новые признаки являются линейными комбинациями исходных признаков.

Метод главных компонент позволяет найти нелинейную поверхность, при проецировании выборки на которую дисперсия оказывается максимальной.

Матрица весов *W*, которая строится в методе главных компонент, является ортогональной.

1. Что неотрицательного в методе неотрицательных матричных разложений?

Элементы раскладываемой матрицы

**Правильно**

Потому что элементы приближающей матрицы неотрицательны и было бы странно приближать такой матрицей, матрицу с отрицательными элементами.

Определители всех угловых миноров раскладываемой матрицы

**правильно, этот вариант не должен быть выбран**

Определители матриц-множителей

**правильно, этот вариант не должен быть выбран**

Элементы матриц-множителей

**Правильно**

Это верно по определению неотрицательного матричного разложения.

Значения функции потерь

**Правильно**

Как и во всех остальных задачах, функция потерь отрицательных значений не принимает.

1. Какую функцию активации *σ* нужно взять в однослойной нейронной сети , чтобы получилась линейная регрессия?



**Правильно**





1. Какую функцию активации *σ* нужно взять в однослойной нейронной сети , чтобы получилась классификация (не логическая регрессия) ?





**Правильно**



1. Какие параметры являются мерами разброса распределений?

Математическое ожидание.

**правильно, этот вариант не должен быть выбран**

Среднеквадратическое отклонение.

**Правильно**

Среднеквадратическое отклонение — это корень из дисперсии; как и дисперсия, оно характеризует разброс распределения.

Интерквартильный размах.

**Правильно**

Интерквартильный размах — мера разброса.

Медиана.

**правильно, этот вариант не должен быть выбран**

Дисперсия.

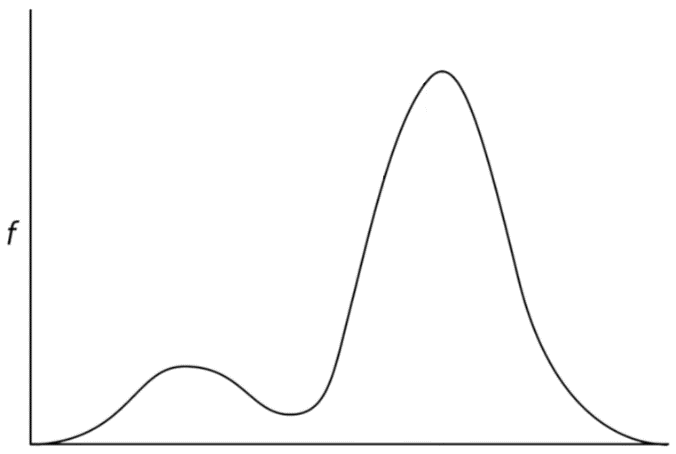
**Правильно**

Дисперсия — наиболее часто используемая мера разброса.

Мода.

**правильно, этот вариант не должен быть выбран**

1. Пусть график плотности распределения случайной величины *X* выглядит следующим образом:



Выберите верные утверждения о средних такой случайной величины.

Медиана *X* меньше её математического ожидания

**правильно, этот вариант не должен быть выбран**

Медиана менее чувствительна к выбросам, в свою очередь небольшое количество экстремальных значений сильно влияют на математическое ожидание.

Мода *X* больше её медианы.

**Правильно**

Мода совпадает с положением большего пика, а медиана смещается в сторону меньшего.

Математическое ожидание и медиана *X* совпадают, а мода не определена.

**правильно, этот вариант не должен быть выбран**

Мода *X* больше её математического ожидания.

**Правильно**

Мода совпадает с положением большего пика, а математическое ожидание смещается в сторону меньшего.

1. Пусть *X*∼*F*(*x*) — случайная величина с произвольным распределением, неизвестным математическим ожиданием EX и известной дисперсией DX. Как выглядит доверительный интервал для с приближённым уровнем доверия  объем выборки?



**Правильно**







1. Пусть *H*0​  - нулевая, основная гипотеза, *H*1 – альтернативная, конкурирующая гипотеза​, если достигаемый уровень значимости *p*≤*α*, то:

*H*0​ не отвергается

*H*1​ верна

*H*0​ отвергается в пользу *H*1​

**Правильно**

*H*1​ отвергается в пользу *H*0​

*H*0​ верна

1. При проверке некоторой гипотезы значение статистики составило 54123432,22. Достигаемый уровень значимости p — это:

вероятность справедливости нулевой гипотезы.

вероятность справедливости нулевой гипотезы при таком значении статистики.

вероятность получить такое значение статистики.

ничего из перечисленного

**Правильно**

Достигаемый уровень значимости — вероятность получить такое или ещё более экстремальное значение статистики при справедливости нулевой гипотезы.

1. Ошибка первого рода — это:

отвержение верной нулевой гипотезы

**Правильно**

принятие неверной нулевой гипотезы

1. Объём выборки, необходимый для построения доверительного интервала заданной ширины:

увеличивается с уменьшением требуемой ширины

**Правильно**

увеличивается с увеличением требуемой ширины

**правильно, этот вариант не должен быть выбран**

увеличивается с ростом дисперсии выборки

**Правильно**

увеличивается с ростом *α*

**правильно, этот вариант не должен быть выбран**

уменьшается с ростом *α*

**Правильно**

уменьшается с ростом дисперсии выборки

**правильно, этот вариант не должен быть выбран**

1. С помощью каких из этих величин по данным социологического опроса можно оценить силу взаимосвязи между возрастом (в годах) и средним годовым доходом (в рублях)?

Коэффициент V Крамера

**правильно, этот вариант не должен быть выбран**

Коэффициент корреляции Пирсона

**Правильно**

Так мы измерим силу линейной взаимосвязи между нашими непрерывными признаками

Коэффициент корреляции Спирмена

**Правильно**

Так мы измерим силу монотонной взаимосвязи между нашими непрерывными признаками

Разность между средними доходами молодых и старых

**правильно, этот вариант не должен быть выбран**

Коэффициент корреляции Мэтьюса

**правильно, этот вариант не должен быть выбран**

1. С помощью каких из этих величин по данным социологического опроса можно оценить силу взаимосвязи между полом и семейным положением (холост/в браке/в разводе)?

Разность между средними долями женатых среди мужчин и среди женщин

**правильно, этот вариант не должен быть выбран**

Коэффициент V Крамера

**Правильно**

Коэффициент корреляции Мэтьюса

**правильно, этот вариант не должен быть выбран**

Коэффициент корреляции Спирмена

**правильно, этот вариант не должен быть выбран**

Коэффициент корреляции Пирсона

**правильно, этот вариант не должен быть выбран**

1. Какой из статистических тестов используется для проверки на нормальность?

Манна-Уитни

Шапиро-Уилка

**Правильно**

Стьюдента

Вавилова-Черенкова

1. Какой из трёх методов поправки на множественную проверку позволяет отвергнуть больше всего гипотез?

Бенджамини-Хохберга

**Правильно**

Холма

Бонферрони

1. В каком случае два ненулевых вектора будут линейно зависимы?

 Если их можно просуммировать с коэффициентами, хотя бы один из которых ненулевой, и получить ноль.

**Правильно**

Это одно из определений линейной зависимости: существует взвешенная сумма, в которой хотя бы один вес ненулевой, и которая равна нулю.

Если один можно получить домножением другого на число.

**Правильно**

Это одно из определений линейной зависимости: один вектор выражается через другие.

Если поэлементное перемножение этих векторов даст нулевой вектор

**правильно, этот вариант не должен быть выбран**

Если существует непрерывная функция, устанавливающая зависимость одного вектора от другого.

**правильно, этот вариант не должен быть выбран**

1. Что такое размерность векторного пространства?

Число различных векторов в нем.

Максимальный размер набора линейно независимых векторов из него.

**Правильно**

Максимальное значение элемента вектора.