1. Мистер Смит отец двоих детей. Мы встретили его, прогуливающегося по улице с маленьким мальчиком, которого он с гордостью представил нам, как своего сына. Какова вероятность того, что другой ребёнок мистера Смита тоже мальчик?

1/3

2/3

1/2

4/3

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?

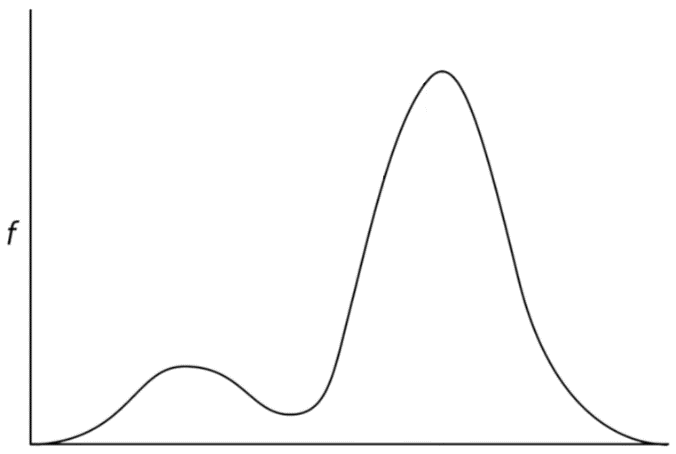
p≤0.037

0.037≤*p*≤0.064

0.037≤*p*≤0.101

0.064≤*p*≤0.101

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

Мода *X* больше её медианы.

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

Мода *X* больше её математического ожидания.

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

увеличивается с уменьшением требуемой ширины

увеличивается с увеличением требуемой ширины

увеличивается с ростом дисперсии выборки

увеличивается с ростом α уровня значимости (допустимая вероятность ошибки I рода)

уменьшается с ростом α уровня значимости

уменьшается с ростом дисперсии выборки

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

Манна-Уитни

Шапиро-Уилка

Стьюдента

Вавилова-Черенкова

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

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

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

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

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









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

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

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. 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. 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. Ошибка первого рода — это:

отвержение верной нулевой гипотезы

принятие неверной нулевой гипотезы