

Вопросы. Теория вероятностей и статистика.

1. (YA) Элементы теории вероятностей: пространство элементарных исходов, вероятностная мера. Свойства вероятностной меры. Условная вероятность. Классическая вероятностная схема. Формулы полной вероятности и Байеса.
2. (SE) Вероятностное пространство. Независимые события. Теорема сложения. Условная вероятность. Полная система событий. Формула полной вероятности. Формула Байеса.
3. (YA) Случайные величины и типы распределений: дискретное, абсолютно непрерывное, сингулярное, смесь. Примеры стандартных распределений: Бернулли, биномиальное, Пуассона, геометрическое, гипергеометрическое, равномерное, показательное, нормальное, гамма.
4. (SE) Случайная величина и её функция распределения. Совместное распределение случайных величин. Распределение суммы независимых случайных величин.
5. (YA) Числовые характеристики случайных величин: математическое ожидание, дисперсия, моменты старших порядков, свойства. Совместное распределение, ковариация и корреляция.
6. (YA) Предельные теоремы теории вероятностей: ЦПТ и ЗБЧ. Теорема Муавра-Лапласа.
7. (ВШЭ-НОД) Генеральная совокупность и выборка. Выборочное распределение и выборочные характеристики (среднее, дисперсия, ковариация, коэффициент корреляции).
8. (ВШЭ-НОД) Статистическое оценивание. Точечные оценки. Линейность, несмещенность, эффективность и состоятельность оценок. Интервальные оценки, доверительный интервал. Метод моментов и метод наибольшего правдоподобия для точечной оценки параметров распределения.

9. (ВШЭ-НОД) Статистические выводы и проверка статистических гипотез. Ошибки первого и второго рода. Уровень доверия и проверка значимости. Проверка гипотез о параметрах нормального распределения.
10. (УА) Описательная статистика. Эмпирическая СВ, теорема Гливенко-Кантелли. Выборочное среднее, выборочная дисперсия (смещенная и несмещенная), выборочные моменты. Гистограмма, как оценка плотности, выборочная функция распределения. Состоятельность оценок, асимптотическая нормальность оценок, несмещенность оценок.
11. (УА) Точечные оценки параметрических и непараметрических семейств. Методы моментов и максимального правдоподобия. Эффективность и асимптотическая эффективность оценок.
12. (УА) Проверка гипотез. Ошибки 1 и 2 рода, уровень значимости и доверительная область. Критерии и их примеры.

Задачи прошлых лет.

1. (ВШЭ. ML. 2020) Какова вероятность того, что сумма n независимых случайных величин, равномерно распределенных на $[0, 1]$ будет не более 1.
2. (ВШЭ. ML. 2020) Нужно написать программу, которая случайным образом генерирует точки в круге радиуса 1. Точки должны быть равномерно распределены по площади круга. Как бы Вы это организовали?
3. (ВШЭ. ML. 2021) Школьник Ваня приболел, и его мама решила вызвать врача домой. У врача есть статистика по району, где живет Ваня. У 90 % больных детей этого района — грипп, у остальных 10 % — ветрянка. Других болезней в этом районе не зафиксировано. Один из основных симптомов ветрянки — это сыпь, она появляется в 95 % случаях заболевания ветрянкой. Однако, во время гриппа она тоже возможна и появляется в 8 % случаях. Осмотрев Ваню, врач обнаружил сыпь. Какова вероятность того, что у Вани ветрянка?
4. (ВШЭ. ML. 2021) Каждый из девяти единичных квадратов 3×3 - квадрата случайным образом окрашен в красный или синий цвет с вероятностью $1/2$. Определите вероятность того, что ни один из четырех квадратов 2×2 не является полностью красным.
5. (ВШЭ. PROG. 2020) Найдите математическое ожидание числа бросков монетки до выпадения первого орла.
6. (ВШЭ. PROG. 2021) У Пети есть для игральных кубика. Один из них честный — имеет равные вероятности выпадения каждой из граней. Второй кубик бракованный — шестерка на нем выпадает с вероятностью $2/3$, а все остальные — с вероятностью $1/15$.

Петя случайным образом выбирает один из двух кубиков, после этого бросает его три раза. Какой будет вероятность выпадения шестерки при третьем броске при условии того, что первые два раза выпала шестерка?

7. (ВШЭ. СКН) Случайные величины X и Y независимы и экспоненциально распределены, X — с параметром 1, а Y — с параметром 2. Пусть $Z = \max(X, Y)$. Найдите математическое ожидание случайной величины Z .
8. (ВШЭ. СКН) Пусть $Y_i \sim N(0, 1)$ при всех $1 \leq i \leq n$ и все Y_i независимы, $\varepsilon_{ij} = Y_i - Y_j$. Покажите, что выполнено

$$E \left[\max_{1 \leq i, j \leq N} |\varepsilon_{ij}| \right] = 2E \left[\max_{1 \leq i \leq N} Y_i \right],$$

а также вычислите дисперсию распределения ε_{ij} при $i \neq j$.

9. (ВШЭ. СКН) На потоке 120 студентов. Каждый студент может выбрать одного из 10 семинаристов. Оказалось, что 20 человек выбрали Алексея. Верно ли, что студенты выбирают его со статистически значимо большей вероятностью, чем случайно, на уровне значимости 5 %. Придумайте статистический критерий для проверки гипотезы, воспользуйтесь им и сделайте выводы.
10. (ВШЭ. НОД-МСК. 2022) Случайная величина X задана функцией плотности вероятности

$$f(x) = \frac{1}{\pi} e^{-\frac{(x-1)^2}{\pi}}.$$

Найдите мат. ожидание случайной величины $Z = X^2 - X + 1$.

11. (ВШЭ. НОД-МСК. 2022) Предполагается, что результаты измерений приведенные в таблице распределены равномерно, с функцией плотности вероятности

$$f(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & x < a, x > b \end{cases}$$

Оцените параметры a, b , используя метод моментов.

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
1.8080	3.2900	3.4050	2.4490	3.1140	1.7600	2.3230	3.3160	3.0820	3.4050

12. (ВШЭ. НОД-МСК. 2021) Вася подбросил монетку 103 раза, и у него 51 раз выпала решка. Потом Миша подбросил ее 53 раза и получил 25 орлов. Оцените вероятность выпадения решки с помощью метода максимального правдоподобия.
13. (ВШЭ. НОД-МСК. 2019) Независимые одинаково распределенные случайные величины X_1, \dots, X_{20} принимают только значения 2 и 3, при этом значение 3 принимается с вероятностью 0,2. Найти вероятность того, что сумма данных величин будет равна 46.
14. (ВШЭ. НОД-МСК. 2019) Программист желает найти доверительный интервал для математического ожидания времени работы программы с надежностью $\gamma = 0,95$. Выборочное среднее времени работы 65 циклов составило 0,09 сек. Времена работы программы независимы и распределены нормально с постоянными параметрами, при этом генеральное среднееквадратичное отклонение равно 0,02. Найти искомый доверительный интервал.