1 Листок 1. Вперёд, в рукопашную!

Минитеория:

1. http://bdemeshev.github.io/pr201/ или http://pokrovka11.wordpress.com
2. Константы. Строчные английские буквы, a, x, z.
3. События. Заглавные английский буквы начала алфавита A, B, C, D. Вероятность $\mathbb{P}(A)$.
4. Случайные величины. Заглавные английский буквы конца алфавита X, Y, W, Z. Математическое ожидание $\mathbb{E}(X)$.

Задачи:

1. В вазе пять неотличимых с виду конфет. Две без ореха и три — с орехом. Маша ест конфеты выбирая их наугад до тех пор, пока не съест первую конфету с орехом. Обозначим X — число съеденных конфет. Найдите $\mathbb{P}(X = 2)$, $\mathbb{P}(X > 1)$, $\mathbb{E}(X)$
2. Саша и Маша по очереди подбрасывают кубик до первой шестёрки. Посуду будет мыть тот, кто первым выбросит шестерку. Маша бросает кубик первой. Какова вероятность того, что посуду будет мыть Маша? Сколько в среднем раз они будут бросать кубик?
3. Две команды равной силы играют в волейбол до трех побед одной из них, не обязательно подряд. Ничья невозможна. Из-за равенства сил можно считать, что вероятность победы каждой равна 0.5. Величина N – количество сыгранных партий. Составьте табличку возможных значений N с их вероятностями. Найдите $\mathbb{P}(N \text{ — четное})$, $\mathbb{E}(N)$
4. Какова вероятность того, что у 30 человек не будет ни одного совпадения дней рождений?
5. Неправильную монетку с вероятностью «орла» равной 0.7 подбрасывают до первого «орла». Чему равно среднее количество подбрасываний? Орлов? Решек? Какова вероятность чётного числа бросков? Как изменятся ответы, если вероятность орла будет равна p?
6. Вы играете в следующую игру. Кубик подкидывается неограниченное число раз. Если на кубике выпадает 1, 2 или 3, то соответствующее количество монет добавляется на кон. Если выпадает 4 или 5, то игра оканчивается и Вы получаете сумму, лежащую на кону. Если выпадает 6, то игра оканчивается, а Вы не получаете ничего. Изначально на кону лежит ноль рублей.
   (a) Какова вероятность того, что игра рано или поздно закончится выпадением 6-ки?
   (b) Какова ожидаемая продолжительность игры?
   (c) Чему равен ожидаемый выигрыш в эту игру?
   (d) Чему равен ожидаемый выигрыш в эту игру, если изначально на кону лежит 100 рублей?
   (e) Изменим условие: если выпадает 5, то сумма на кону сгорает, а игра продолжается. Как изменятся ответы на предыдущие вопросы?
7. Саша и Маша подкидывают монетку до тех пор, пока не выпадет последовательность РОО или ООР. Если игра закончится выпадением РОО, то выигрывает Саша, если ООР, то — Маша. Случайная величина X — общее количество подбрасываний, Y — количество выпавших решек.

(a) У кого какие шансы выиграть?

(b) $\mathbb{P}(X = 4)$, $\mathbb{P}(Y = 1)$, $\mathbb{E}(X)$, $\mathbb{E}(Y)$

(c) Решите аналогичную задачу для ОРО и ООР.

8. Вася подкидывает кубик до тех пор, пока на кубике не выпадет единица, или пока он сам не скажет «Стоп». Вася получает столько рублей, сколько выпало на кубике при последнем броске. Вася хочет максимизировать свой ожидаемый выигрыш.

   (a) Как выглядит оптимальная стратегия? Чему равен ожидаемый выигрыш при использовании оптимальной стратегии?

   (b) Какова средняя продолжительность игры при использовании оптимальной стратегии?

   (c) Как выглядит оптимальная стратегия и чему равен ожидаемый выигрыш, если за каждое подбрасывание Вася платит 35 копеек?

9. Саша и Маша решили, что будут заводить новых детей до тех пор, пока в их семье не будут дети обоих полов. Обозначим X — количество детей в их семье. Найдите $\mathbb{P}(X = 4)$, $\mathbb{E}(X)$

10. В каждой вершине треугольника по ёжику. Каждую минуту с вероятностью 0.7 каждый ежик независимо от других двигается по часовой стрелке, с вероятностью 0.3 — против часовой стрелки. Обозначим T — время до встречи всех ежей в одной вершине. Найдите $\mathbb{P}(T = 1)$, $\mathbb{P}(T = 2)$, $\mathbb{P}(T = 3)$, $\mathbb{E}(T)$.

2 Листок 2. Хочу ещё задач!

1. Наугад из четырех тузов разных мастей выбираются два. $\mathbb{P}$(они будут разного цвета)?

2. События A и B несовместны, то есть не могут произойти одновременно. Известны вероятности $\mathbb{P}(A) = 0.3$, $\mathbb{P}(B) = 0.4$. Найдите[1] $\mathbb{P}(A^c \cap B^c)$.

3. Вероятность $\mathbb{P}(A) = 0.3$, $\mathbb{P}(B) = 0.8$. В каких пределах может лежать $\mathbb{P}(A \cap B)$?

4. Множество исходов $\Omega = \{a, b, c\}$, $\mathbb{P}(\{a, b\}) = 0,8$, $\mathbb{P}(\{b, c\}) = 0,7$. Найдите $\mathbb{P}(\{a\})$, $\mathbb{P}(\{b\})$, $\mathbb{P}(\{c\})$

5. «Amoeba». A population starts with a single amoeba. For this one and for the generations thereafter, there is a probability of $3/4$ that an individual amoeba will split to create two amoebas, and a $1/4$ probability that it will die out without producing offspring. Let the random variable X be the number of generations before the death of all the amoebas. Find the probabilities $\mathbb{P}(X = 2)$, $\mathbb{P}(X = 3)$, $\mathbb{P}(X = \infty)$

6. Вася нажимает на пульте телевизора кнопку «On-Off» 100 раз подряд. Пульт старый, поэтому в первый раз кнопка срабатывает с вероятностью $\frac{1}{2}$, затем вероятность срабатывания падает. Какова вероятность того, что после всех нажатий телевизор будет включен, если сейчас он выключен?

7. Suppose the probability to get a head when throwing an unfair coin is p, what's the expected number of throwings in order to get two consecutive heads? The expected number of tails?

[1] Событие A^c — это событие противоположное событию A, иногда обозначается $\bar{A}$

8. Вам предложена следующая игра. Изначально на кону 0 рублей. Раз за разом подбрасывается правильная монетка. Если она выпадает орлом, то казино добавляет на кон 100 рублей. Если монетка выпадает решкой, то все деньги, лежащие на кону, казино забирает себе, а Вы получаете красную карточку. Игра прекращается либо когда Вы получаете третью красную карточку, либо в любой момент времени до этого по Вашему выбору. Если Вы решили остановить игру до получения трех красных карточек, то Ваш выигрыш равен сумме на кону. При получении третьей красной карточки игра заканчивается и Вы не получаете ничего.

   (a) Как выглядит оптимальная стратегия в этой игре, если Вы заинтересованы в максимальном среднем выигрыше?

   (b) Чему при этом будет равен средний выигрыш?

9. Есть три комнаты. В первой из них лежит сыр. Если мышка попадает в первую комнату, то она находит сыр через одну минуту. Если мышка попадает во вторую комнату, то она ищет сыр две минуты и покидает комнату. Если мышка попадает в третью комнату, то она ищет сыр три минуты и покидает комнату. Покинув комнату, мышка выходит в коридор и выбирает новую комнату наугад (т.е. может зайти в одну и ту же). Сейчас мышка в коридоре. Сколько времени ей в среднем потребуется, чтобы найти сыр?

10. Илье Муромцу предстоит дорога к камню. От камня начинаются ещё три дороги. Каждая из тех дорог снова оканчивается камнем. И от каждого камня начинаются ещё три дороги. И каждые те три дороги оканчиваются камнем... И так далее до бесконечности. На каждой дороге живёт трёхголовый Змей Горыныч. Каждый Змей Горыныч бодрствует независимо от других с вероятностью (хм, Вы не поверите!) одна третья. У Василисы Премудрой существует Чудо-Карта, на которой видно, какие Змеи Горынычи бодрствуют, а какие — нет. Какова вероятность того, что Василиса Премудрая сможет найти на карте бесконечный жизненный путь Ильи Муромца проходящий исключительно мимо спящих Змеев Горынычей?

11. У Пети — монетка, выпадающая орлом с вероятностью $p \in (0;1)$. У Васи — с вероятностью $1/2$. Они одновременно подбрасывают свои монетки до тех пор, пока у них не окажется набранным одинаковое количество орлов. В частности, они останавливаются после первого подбрасывания, если оно дало одинаковые результаты. Сколько в среднем раз им придётся подбросить монетку?

12. Треугольник с вершинами $(0;0)$, $(2;0)$ и $(1;1)$. Внутри него случайным образом выбирается точка, X – абсцисса точки. Найдите $\mathbb{P}(X > 1)$, $\mathbb{P}(X \in [0.5;1])$, $\mathbb{E}(X)$

13. Треугольник с вершинами $(0;0)$, $(2;0)$ и $(2;1)$. Внутри него случайным образом выбирается точка, X – абсцисса точки. Найдите $\mathbb{P}(X > 1)$, $\mathbb{P}(X \in [0.5;1])$. Что больше, $\mathbb{E}(X)$ или 1?

14. Исследовательница Мишель подкидывает игральный кубик неограниченное количество раз и складывает выпадающие количества очков.

   (a) Чему примерно равна вероятность того, что однажды сумма в точности будет равна 123456789?

   (b) Чему точно равна указанная вероятность?

15. Упрямая исследовательница Мишель подбрасывает монетку до тех пор, пока количество орлов не окажется в точности равным удвоенному количеству решек. Монетка выпадает

орлом с вероятностью p. Какова вероятность того, что Мишель будет подкидывать монетку вечно?

16. Исследовательница Мишель хочет встать утром с правой ноги с вероятностью $1/\sqrt{2}$, и с вероятностью $(\sqrt{2}-1)/\sqrt{2}$ — с левой. Однако для проведения случайных экспериментов у неё есть только одна правильная монетка. Как с помощью правильной монетки ей добиться цели?

3 Листок 3. К чёрту условности!

1. Имеется три монетки. Две «правильных» и одна — с «орлами» по обеим сторонам. Петя выбирает одну монетку наугад и подкидывает ее два раза. Оба раза выпадает «орел». Какова условная вероятность того, что монетка «неправильная»?

2. Два охотника одновременно выстрелили в одну утку. Первый попадает с вероятностью 0,4, второй — с вероятностью 0,7. В утку попала ровно одна пуля. Какова условная вероятность того, что утка была убита первым охотником?

3. Кубик подбрасывается два раза. Найдите вероятность получить сумму равную 8, если при первом броске выпало 3.

4. Игрок получает 13 карт из колоды в 52 карты. Какова вероятность, что у него как минимум два туза, если известно, что у него есть хотя бы один туз? Какова вероятность того, что у него как минимум два туза, если известно, что у него есть туз пик?

5. В урне 7 красных, 5 желтых и 11 белых шаров. Два шара выбирают наугад. Какова вероятность, что это красный и белый, если известно, что они разного цвета?

6. В урне 5 белых и 11 черных шаров. Два шара извлекаются по очереди. Какова вероятность того, что второй шар будет черным? Какова вероятность того, что первый шар — белый, если известно, что второй шар — черный?

7. Примерно 4% коров заражены «коровьим бешенством». Имеется тест, который дает ошибочный результат с вероятностью 0,1. Судя по тесту, новая партия мяса заражена. Какова вероятность того, что она действительно заражена?

8. В школе три девятых класса, «А», «Б» и «В», одинаковые по численности. В «А» классе 30% обожают учителя географии, в «Б» классе — 40% и в «В» классе — 70%. Девятиклассник Петя обожает учителя географии. Какова вероятность того, что он из «Б» класса?

9. Ген карих глаз доминирует ген синих. Т.е. у носителя пары bb глаза синие, а у носителя пар BB и Bb — карие. У диплоидных организмов (а мы такие :)) одна аллель наследуется от папы, а одна — от мамы. В семье у ых родителей два сына — кареглазый и синеглазый. Кареглазый женился на синеглазой девушке. Какова вероятность рождения у них синеглазого ребенка?

10. Из колоды в 52 карты извлекается одна карта наугад. Являются ли события «извлечен туз» и «извлечена пика» независимыми?

11. Из колоды в 52 карты извлекаются по очереди две карты наугад. Являются ли события «первая карта — туз» и «вторая карта — туз» независимыми?

12. Известно, что $\mathbb{P}(A) = 0,3$, $\mathbb{P}(B) = 0,4$, $\mathbb{P}(C) = 0,5$. События A и B несовместны, события A и C независимы и $\mathbb{P}(B|C) = 0,1$. Найдите $\mathbb{P}(A \cup B \cup C)$.

13. У тети Маши — двое детей, один старше другого. Предположим, что вероятности рождения мальчика и девочки равны и не зависят от дня недели, а пол первого и второго ребенка независимы. Для каждой из четырех ситуаций найдите условную вероятность того, что у тёти Маши есть дети обоих полов.

    (a) Известно, что хотя бы один ребенок — мальчик.

    (b) Тетя Маша наугад выбирает одного своего ребенка и посылает к тете Оле, вернуть учебник по теории вероятностей. Это оказывается мальчик.

    (c) Известно, что старший ребенок — мальчик.

    (d) На вопрос: «А правда ли тетя Маша, что у вас есть сын, родившийся в пятницу?» тётя Маша ответила: «Да».

14. В своём труде 1763 года «An Essay towards solving a Problem in the Doctrine of Chances» Томас Байес решает следующую задачу: «Given the number of times in which an unknown event has happened and failed: Required the chance that the probability of its happening in a single trial lies somewhere between any two degrees of probability that can be named». А вам слабо?

    Имеется монетка, возможно неправильная. Априорная вероятность выпадения орла, p, равномерно распределена на отрезке $[0; 1]$. Монетка выпала 5 раз орлом и 7 раз решкой. Какова апостериорная вероятность того, что $p < 0.4$?

    В байесовском подходе p — это констната или случайная величина?

    В каком году умер Томас Байес?

    http://rstl.royalsocietypublishing.org/content/53/370.full.pdf

4 Листок 4. Use R

5 Листок 5. Use R or die!

1. Самая простая. Случайная величина N имеет пуассоновское распределение с $\lambda = 2$. С помощью симуляций оцените $\mathbb{E}(N^3)$, $\mathbb{P}(N \geqslant 4)$, $\mathbb{P}(N \geqslant 10 \mid N \geqslant 5)$, $\mathbb{E}(N \mid N \geqslant 5)$. Функция `rpois` может помочь :)

2. Случайные величины $X_1, \ldots, X_5$ имеют равномерное распределение на отрезке $[0; 1]$ и независимы. С помощью симуляций оцените $\mathbb{P}(\min\{X_1, \ldots, X_5\} > 0.2)$, $\mathbb{P}(\min\{X_1, \ldots, X_5\} > 0.2 \mid X_1 + X_2 < 0.5)$, $\mathbb{E}(\min\{X_1, \ldots, X_5\})$, $\mathbb{E}(\min\{X_1, \ldots, X_5\} \mid X_1 + X_2 < 0.5)$

3. Случайные величины X_1, X_2 независимы и обе имеют биномиальное распределение с параметрами $n = 16$, $p = 0.7$. Величина Y задана формулой $Y = X_1/(1 + X_2)$. С помощью симуляций оцените $\mathbb{P}(Y > 0.5)$, $\mathbb{E}(Y)$, $\mathbb{P}(Y > 0.5 \mid X_1 > 10)$, $\mathbb{E}(Y \mid X_1 > 10)$. Функция `rbinom` в помощь!

4. В колоде 52 карты. Мы вытаскиваем карты из колоды до первого туза, пусть X — количество вытянутых карт. С помощью симуляций оцените $\mathbb{E}(X^2)$, $\mathbb{P}(X > 10)$, $\mathbb{P}(X > 5 \mid X < 15)$, $\mathbb{E}(X^2 \mid X < 15)$

5. Иван Федорович Крузенштерн случайным образом с возможностью повторов выбирает 10 натуральных чисел от 1 до 100. Пусть X — минимум этих чисел, а Y — максимум. С помощью симуляций оцените $\mathbb{P}(Y > 3X)$, $\mathbb{E}(XY)$, $\mathbb{P}(Y > 3X \mid Y < X^2)$, $\mathbb{E}(XY \mid Y < X^2)$

6 Листок f. Эф большое и эф малое

1. Функция плотности случайной величины X равна 5 при $x = 7$. Найдите примерно вероятность того, что X попадёт в отрезок $[7; 7.001]$.

2. Случайнай величина X имеет функцию плотности f. С помощью $o(\Delta)$ и $f(x)$ запишите вероятность вероятность $\mathbb{P}(X \in [x; x + \Delta])$.

3. Величина X распределена на отрезке $[0; 1]$ и на нём имеет функцию плотности $f(t) = 3t^2$. Найдите функции плотности и функции распределения величин $Y = \ln X$, $Z = X^2$, $W = (X - 0.5)^2$.

4. Может ли функция плотности принимать значение больше 2015? Может ли предел $\lim_{x\to\infty} f(x)$ не равняться нулю?

5. Функция плотности случайной величины X имеет вид:

$$f(t) = \begin{cases} \frac{3}{16}t^2, t \in [-2; 2] \\ 0, t \notin [-2; 2] \end{cases}$$

   Найдите:

   (a) $\mathbb{P}(X > 1)$, $\mathbb{E}(X)$, $\mathbb{E}(X^2)$, $\mathrm{Var}(X)$, σ_X

   (b) $\mathbb{E}(X|X > 1)$, $\mathbb{E}(X^2|X > 1)$, $\mathrm{Var}(X|X > 1)$

   (c) Функцию распределения случайной величины X

   (d) Решите предыдущие пункты для

$$f(t) = \begin{cases} t/8, t \in [0; 4] \\ 0, t \notin [0; 4] \end{cases}$$

6. Величина X распределена на отрезке $[0; 2]$ и имеет на нём функцию распределения $F(x) = x^2/4$. Найдите $\mathbb{P}(X \in [1; 1.5])$, $\mathbb{P}(X < 1)$, $F(-5)$, $F(10)$, функцию плотности величины X

7. Если возможно, найдите функцию распределения и функцию плотности величины X принимающей значения 1, 2, 3 и 4 с вероятностями 0.1, 0.2, 0.3 и 0.4 соответственно.

8. Величина X равномерна на отрезке $[-2; 1]$, а величина Y — это расстояние от числа X до числа (-1). Найдите фукнцию плотности Y, $\mathbb{E}(Y)$.

9. Глафира случайным образом равномерно выбирает случайную внутри треугольника с вершинами $(0; 0)$, $(0; 2)$ и $(3; 3)$. Пусть X и Y — абсцисса и ордината выбранной точки. Найдите функцию плотности X, функцию плотности Y.

10. Прямой убыток от пожара в миллионах рублей равномерно распределен на $[0; 1]$. Если убыток оказывается больше 0.7, то страховая компания выплачивает компенсацию 0.7.

    (a) Найдите функцию распределения потерь от пожара.

    (b) Чему равны средние потери?

11. Пусть X — неотрицательная случайная величина с функцией плотности $f(t)$ и $\mathbb{E}(X) < \infty$. При каком c функция $g(t) = c \cdot tp(t)$ также будет функцией плотности?

12. Завтрашняя цена акции — случайная величина с функцией плотности $f(x) = \frac{3}{4}\max\{x(2-x), 0\}$.

    (a) Постройте график функции плотности

    (b) Найдите функцию распределения Васиного дохода, средний доход и дисперсию дохода, если:

        i. У Васи есть 10 акций

        ii. У Васи есть опцион-пут, дающий ему право продать акции по цене 1.2 рубля

        iii. У Васи есть опцион-колл, дающий ему право купить акции по цене 1 рубль

13. В соревнованиях по прыжкам в длину участвовали n спортсменов. Результаты их прыжков, величины X_i, независимы и одинаково распределены с функцией плотности f и функцией распределения F.

    (a) Найдите функцию распределения и функцию плотности длины наилучшего прыжка

    (b) Найдите функцию распределения и функцию плотности длины наихудшего прыжка

    (c) Найдите вероятность того, что Сидоров и Петров прыгнули меньше чем на t метров, Иванов прыгнул от t до $t + \Delta$ метров, а остальные прыгнули больше, чем на t метров.

    (d) Найдите функцию распределения и функцию плотности длины прыжка бронзового призёра соревнований

14. Большой Адронный Коллайдер запускают ровно в полночь. Оставшееся время до Конца Света — случайная величина X распределенная равномерно от 0 до 16 часов. Когда произойдет Конец Света, механические часы остановятся и будут показывать время Y.

    (a) Найдите $\mathbb{P}(Y < 2)$

    (b) Постройте функцию плотности величины Y

    (c) Найдите $\mathbb{E}(Y)$, $\mathrm{Var}(Y)$

    (d) Найдите $\mathrm{Cov}(X, Y)$

7 Листок 7. Разлагай и властвуй!

1. Из грота ведут 10 штреков, с длинами 100м, 200м, ...1000м. Самый длинный штрек оканчивается выходом на поверхность. Остальные — тупиком. Вася выбирает штреки наугад, в тупиковый штрек два раза не ходит. Какова вероятность того, что Вася посетит самый короткий штрек? Какой в среднем путь он нагуляет прежде чем выберется на поверхность?

2. У Маши 30 разных пар туфель. И она говорит, что мало! Пёс Шарик утащил без разбору на левые и правые 17 туфель. Какова вероятность того, что у Маши останется ровно 13 полных пар? Величина X — количество полных целых оставшихся пар, Y — количество полных пар, доставшихся Шарику. Найдите $\mathbb{E}(X)$, $\mathbb{E}(Y)$, $\mathrm{Var}(X)$, $\mathrm{Var}(Y)$.

3. У меня в кармане 3 рубля мелочью. Среди монет всего одна монета достоинством 50 копеек. Я извлекаю монеты по одной наугад. Я останавливаюсь после того, как извлеку монету в 50 копеек. Какую сумму в среднем я извлеку?

4. «Модница». В шкатулке у Маши 100 пар серёжек. Каждый день утром она выбирает одну пару наугад, носит ее, а вечером возвращает в шкатулку. Проходит год.

   (a) Сколько в среднем пар окажутся ни разу не надетыми?

   (b) Сколько в среднем пар окажутся надетыми не менее двух раз?

   (c*) Как изменятся ответы, если каждый день Маша покупает себе новую пару серёжек и вечером добавляет ее в шкатулку?

5. Вовочка получает пятерку с вероятностью 0.1, четверку — с вероятностью 0.2, тройку — с вероятностью — 0.3 и двойку с вероятностью 0.4. В этом четверти он писал 20 контрольных. Какова вероятность того, что все оценки у Вовочки одинаковые? Сколько разных оценок он в среднем получит?

6. «Судьба Дон Жуана» У Васи n знакомых девушек (их всех зовут по-разному). Он пишет им n писем, но, по рассеянности, раскладывает их в конверты наугад. Величина X обозначает количество девушек, получивших письма, написанные лично для них. Найдите $\mathbb{E}(X)$, $\mathrm{Var}(X)$.

7. Над озером взлетело 20 уток. Каждый из 10 охотников один раз стреляет в случайно выбираемую им утку. Величина Y — количество убитых уток, X — количество попавших в цель охотников. Найдите $\mathbb{E}(X)$, $\mathrm{Var}(X)$, $\mathbb{E}(Y)$, $\mathrm{Var}(Y)$, если охотники стреляют без промаха. Как изменится ответ, если вероятность попадания равна 0.7?

8. Вокруг новогодней ёлки танцуют хороводом 27 детей. Мы считаем, что ребенок высокий, если он выше обоих своих соседей. Величина X — количество высоких детей в хороводе. Найдите $\mathbb{E}(X)$, $\mathrm{Var}(X)$. Вероятность совпадения роста будем считать равной нулю.

9. По 10 коробкам наугад раскладывают 7 карандашей. Каково среднее количество пустых коробок? Дисперсия?

10. Внутри каждой упаковки шоколадки находится наклейка с изображением одного из 30 животных. Предположим, что все наклейки равновероятны, величина X — это количество шоколадок, которые купить, чтобы собрать полную коллекцию наклеек. Чему равны $\mathbb{E}(X)$, $\mathrm{Var}(X)$? Как это объяснить ребёнку?

11. Из колоды в 52 карты извлекается 5 карт. Сколько в среднем извлекается мастей? Достоинств? Тузов? Дисперсии этих величин?

12. За круглым столом сидят в случайном порядке n супружеских пар, всего — $2n$ человек. Величина X — число пар, где супруги оказались напротив друг друга. Найдите $\mathbb{E}(X)$ и $\mathrm{Var}(X)$

13. В задачнике N задач. Из них a — Вася умеет решать, а остальные не умеет. На экзамене предлагается равновероятно выбираемые n задач. Величина X — число решенных Васей задач на экзамене. Найдите $\mathbb{E}(X)$ и $\mathrm{Var}(X)$

14. Кубик подбрасывается n раз. Величина X_1 — число выпадений 1, а X_6 — число выпадений 6. Найдите $\mathrm{Corr}(X_1, X_6)$

Пуассоновский поток событий. Обозначим: $X[a; a+\Delta]$ — количество происшествий на интервале $[a; a+\Delta]$, $X_t = X[0;t]$ — количество происшествий за период $[0;t]$.

Если:

1. На малом интервале времени вероятность одного происшествия примерно пропорциональна длине интервала, $\mathbb{P}(X[a; a+\Delta] = 1) = \lambda\Delta + o(\Delta)$.

2. На малом интервале времени несколько происшествий происходят существенно реже одного происшествия, $\mathbb{P}(X[a; a+\Delta] \geqslant 2) = o(\Delta)$.

3. Стационарность приращений. Распределение случайной величины $X[a; a+\Delta]$, количества происшествий на интервале $[a; a+\Delta t]$, зависит только от Δ, но не от a.

4. Независимость приращений. Количество происшествий на непересекающихся интервалах времени независимы.

То:

1. Время между $(i-1)$-ым и i-ым происшествием, Y_i, имеет экспоненциальное распределение $Y_i \sim exp(\lambda)$.
$$f(y) = \begin{cases} \lambda e^{-\lambda y},\ y \geqslant 0 \\ 0,\ y < 0 \end{cases}$$
$$\mathbb{E}(Y_i) = 1/\lambda$$
$$\mathrm{Var}(Y_i) = 1/\lambda^2$$

2. Количество происшествий за единицу времени, X, имеет пуассоновское распределение $X \sim Pois(\lambda)$
$$\mathbb{P}(X = k) = e^{-\lambda}\frac{\lambda^k}{k!}$$
$$\mathbb{E}(X) = \lambda$$
$$\mathrm{Var}(X) = \lambda$$

3. Отсюда смысл λ — среднее количество событий за единицу времени, дисперсия количества событий за единицу времени

4. Количество событий за период времени $[0;t]$, величина X_t, имеет пуассоновское распределение $X_t \sim Pois(\lambda t)$
$$\mathbb{P}(X_t = k) = e^{-\lambda t} \cdot \frac{(\lambda t)^k}{k!}$$
$$\mathbb{E}(X_t) = \lambda t$$
$$\mathrm{Var}(X_t) = \lambda t$$

5. Величины Y_i независимы

6. Сумма двух независимых пуассоновских процессов с интенсивностями λ_1 и λ_2 — пуассоновский процесс с интенсивностью $\lambda_1 + \lambda_2$

Замена $Bin(n,p)$ на $Pois(\lambda = np)$ дает погрешность не более $\min\{p, np^2\}$

8 Листок (n+1). За время моего дежурства происшествий не было!

1. Маша и Саша пошли в лес по грибы. Саша собирает все грибы, а Маша — только подберезовики. Саша в среднем находит один гриб за одну минуту, Маша — один гриб за десять минут. Какова вероятность того, за 8 минут они найдут ровно 13 грибов? Какова вероятность того, что следующий гриб им попадется позже, чем через минуту, если Маша только что нашла подберезовик?

2. Пост майора ГИБДД Иванова И.И. в среднем ловит одного нарушителя в час. Какова вероятность того, за первые полчаса дежурства будет не меньше двух нарушителей? Какова вероятность того, что следующего нарушителя ждать еще более 40 минут, если уже целых три часа никто не превышал скорость?

3. Оля и Юля пишут смс Маше. Оля отправляет Маше в среднем 5 смс в час. Юля отправляет Маше в среднем 2 смс в час. Какова вероятность того, что Маша получит ровно 6 смс за час? Сколько времени в среднем проходит между смс, получаемыми Машей от подруг?

4. Кузнечики на большой поляне распределены по пуассоновскому закону, в среднем 3 кузнечика на квадратный метр. Какой следует взять сторону квадрата, чтобы вероятность найти в нем хотя бы одного кузнечика была равна $0,8$?

5. В магазине две кассирши (ах, да! две хозяйки кассы). Допустим, что время обслуживания клиента распределено экспоненциально. Тетя Зина обслуживает в среднем 5 клиентов в час, а тетя Маша - 7. Два клиента подошли к кассам одновременно.

   (a) Какова вероятность того, что тетя Зина обслужит клиента быстрее?

   (b) Как распределено время обслуживания того клиента, который освободится быстрее?

   (c) Каково условное среднее время обслуживания клиента тетей Зиной, если известно, что она обслужила клиента быстрее тети Маши?

6. Время между приходами студентов в столовую распределено экспоненциально; в среднем за 10 минут приходит 5 студентов. Время обслуживания имеет экспоненциальное распределение; в среднем за 10 минут столовая может обслужить 6 студентов. Столовая находится в динамическом равновесии, то есть закон распределения длины очереди стабилен (это не означает, что длина очереди не меняется).

   (a) Какова вероятность того, что в очереди ровно n студентов?

   (b) Какова средняя длина очереди?

   Подсказка: если сейчас в очереди n человек, то через малый промежуток времени Δ…

7. The arrival of buses at a given bus stop follows Poisson law with rate 2. The arrival of taxis at the same bus stop is also Poisson, with rate 3. What is the probability that next time I'll go to the bus stop I'll see at least two taxis arriving before a bus? Exactly two taxis?

8. Время, которое хорошо обученная свинья тратит на поиск трюфеля — экспоненциальная случайная величина со средним в 10 минут. Какова вероятность того, что свинья за 20 минут не найдет ни одного трюфеля?

9. Величина X распределена экспоненциально с параметром λ, а константа $a > 0$. Как распределена величина $Y = aX$?

10. В гирлянде 25 лампочек. Вероятность брака для отдельной лампочки равна 0,01. Какова вероятность того, что гирлянда полностью исправна? Оцените точность ответа при использовании распределения Пуассона.

11. По некоему предмету незачет получило всего 2% студентов. Какова вероятность того, что в группе из 50 студентов будет ровно 1 человек с незачетом? Оцените точность ответа при использовании распределения Пуассона.

12. Вася испек 40 булочек. В каждую из них он кладет изюминку с $p = 0,02$. Какова вероятность того, что всего окажется 3 булочки с изюмом? Оцените точность ответа при использовании распределения Пуассона.

13. В офисе два телефона — зеленый и красный. Входящие звонки на красный — Пуассоновский поток событий с интенсивностью $\lambda_1 = 4$ звонка в час, входящие на зеленый – с интенсивностью $\lambda_2 = 5$ звонка в час. Секретарша Василиса Премудрая одна в офисе. Перед началом рабочего дня она подбрасывает монетку и отключает один из телефонов, зеленый — если выпала решка, красный — если орел. Обозначим Y_1 время от начала дня до первого звонка.

    (a) Найдите функцию плотности Y_1

    (b) Верно ли, что процесс количества звонков, которые услышит Василиса, имеет независимые приращения?

14. Владелец салуна «Огненная зебра» закрывает заведение, если в течение 5 минут никто не заказывает виски. Посетители заказывают в среднем один виски в минуту. Заказы представляют собой пуассоновский поток. Пусть X — время от открытия до закрытия таверны. Найдите $\mathbb{E}(X)$, $\mathrm{Var}(X)$.

15. Рассмотрим определение пуассоновского процесса, а именно: вероятность ровно одного события за интервал времени $[a; a + \Delta]$ есть $\lambda\delta + o(\Delta)$, вероятность не менее двух событий за тот же интервал времени есть $o(\Delta)$. Докажите, что время между событиями имеет экспоненциальное распределение.

16. Количество трапперов, заходящих в салун «Огненная зебра», — пуассоновский поток с единичной интенсивностью. Какова вероятность того, что через время t от момента открытия в салун зайдёт чётное количество трапперов?

9 Листок. Совместная плотность

1. На первом шаге значение X выбирается случайно равномерно на отрезке $[0; 1]$. На втором шаге значение Y выбирается случайно и равномерно от 0 до получившегося X.

    (a) Найдите функции плотности $f(y|x)$, $f(x)$, $f(x, y)$, $f(y|x)$, $f(y)$

    (b) Найдите $\mathbb{E}(X)$, $\mathbb{E}(Y)$, $\mathbb{E}(X^2)$, $\mathbb{E}(Y^2)$

    (c) Найдите $\mathrm{Var}(X)$, $\mathrm{Var}(Y)$, $\mathrm{Cov}(X, Y)$, $\mathrm{Corr}(X, Y)$

    (d) Найдите $\mathbb{E}(X|Y)$, $\mathbb{E}(Y|X)$, $\mathbb{E}(X^2|Y)$, $\mathbb{E}(Y^2|X)$

(e) Найдите $\mathrm{Var}(Y|X)$, $\mathrm{Var}(X|Y)$

(f) Найдите $\mathbb{P}(Y > 0.2|X = 0.5)$, $\mathbb{P}(Y > 0.2|X < 0.5)$

2. Величины Y_1 и Y_2 независимы и экспоненциально распределены с параметром λ, а $S = Y_1 + Y_2$.

   (a) Найдите плотности $f(y_1, s)$, $f(y_1|s)$

   (b) Прокомментируйте простыми словами вид функции $f(y_1|s)$

3. Величины X и Y имеют совместную функцию плотности

$$f(x,y) = \begin{cases} 4xy, \ 0 < x < 1, \, 0 < y < 1 \\ 0, \ \text{иначе} \end{cases}$$

   (a) Найдите $\mathbb{P}(X > 0.5)$, $\mathbb{P}(X + Y > 0.5)$, $\mathbb{P}(X = Y + 0.2)$, $\mathbb{P}(X \leqslant Y)$, $\mathbb{P}(Y > 0.5|X > 0.5)$

   (b) Найдите $\mathbb{E}(X)$, $\mathbb{E}(XY)$, $\mathrm{Cov}(X, Y)$, $\mathrm{Corr}(X, Y)$

   (c) Найдите $\mathbb{E}(Y|X)$, $\mathbb{E}(Y^2|X)$, $\mathrm{Var}(Y|X)$

   (d) Зависимы ли величины X и Y?

   (e) Найдите совместную функцию распределения $F(x, y)$

4. Величины X и Y имеют совместную функцию распределения

$$F(x,y) = \begin{cases} 1 - e^{-x} - e^{-y} + e^{-x-y}, \ x > 0, \, y > 0 \\ 0, \ \text{иначе} \end{cases}$$

   (a) Найдите $\mathbb{P}(X > 0.5)$, $\mathbb{P}(X + Y > 0.5)$, $\mathbb{P}(X = Y + 0.2)$, $\mathbb{P}(X \leqslant Y)$, $\mathbb{P}(Y > 0.5|X > 0.5)$

   (b) Найдите частные функции распределения $F(x)$, $F(y)$

   (c) Найдите совместную функцию плотности $f(x, y)$

   (d) Найдите $\mathbb{E}(X)$, $\mathbb{E}(XY)$, $\mathrm{Cov}(X, Y)$, $\mathrm{Corr}(X, Y)$

   (e) Зависимы ли величины X и Y?

5. Величины X_1 и X_2 независимы и равномерны на отрезке $[0; 1]$. Пусть $L = \min\{X_1, X_2\}$, а $R = \max\{X_1, X_2\}$. Найдите совместную функцию плотности пары (L, R)

6. Точка случайно равномерно выбирается внутри треугольника с вершинами в $(0; 0)$, $(3; 3)$ и $(1; 2)$. Пусть X и Y — абсцисса и ордината этой точки. Найдите совместную функцию плотности пары (X, Y)

7. Приведите пример пары X и Y у которой нет совместной функции плотности, однако X и Y равномерно распределены на отрезке $[0; 1]$

10 Листок. Всё нормально!

1. Величины X_1, ..., X_n распределены нормально $\mathcal{N}(4, 100)$ и независимы.

   (a) Найдите вероятности $\mathbb{P}(X_1 > 4)$, $\mathbb{P}(X_1 \in [2; 20]$, $\mathbb{P}(X_1 < -5)$

   (b) Найдите вероятности $\mathbb{P}(\bar{X}_{100} > 10)$, $\mathbb{P}(\bar{X}_{36} \in [0; 5])$

   (c) Найдите такое число a, что $\mathbb{P}(X_1 > a) = 0.3$

   (d) Найдите такое число b, что $\mathbb{P}(\bar{X}_{100} \in [4 - b; 4 + b]) = 0.5$

   Решите эту задачу двумя способами: с использованием R и с помощью таблиц. В R могут оказаться полезны функции `pnorm` и `qnorm`.

2. Величина W имеет функцию плотности $f(w) = c \cdot \exp(5w - 2w^2)$. Найдите $\mathbb{E}(W)$, $\mathrm{Var}(W)$, c.

3. Величина X нормальна $\mathcal{N}(\mu, \sigma^2)$.

   (a) Выпишите функцию плотности случайной величины X, $f(x)$

   (b) Найдите точку максимума и точки перегиба функции f

4. Величина X распределена нормально $\mathcal{N}(0, \sigma^2)$.

   (a) Найдите $\mathbb{E}(|X|)$

   (b) Найдите функцию плотности $|X|$

5. Известно, что $\ln Y \sim \mathcal{N}(\mu; \sigma^2)$. Найдите $\mathbb{E}(Y)$, $\mathrm{Var}(Y)$, медиану и моду величины Y.

6. Величина X имеет стандартное нормальное распределение.

   (a) Найдите $\mathbb{E}(Y)$ для $Y = \max\{X, 0\}$

   (b) Найдите $\mathbb{E}(X|X < 0)$ и $\mathrm{Var}(X|X < 0)$

11 Неравенство Чебышёва и Маркова

1. С помощью неравенства Чебышева оцените вероятности

   (a) $\mathbb{P}(-2\sigma < X - \mu < 2\sigma)$, если $\mathbb{E}(X) = \mu$, $\mathrm{Var}(X) = \sigma^2$

   (b) $\mathbb{P}(8 < Y < 12)$, если $\mathbb{E}(Y) = 10$, $\mathrm{Var}(Y) = 400/12$

   (c) $\mathbb{P}(-2 < Z - \mathbb{E}(Z) < 2)$, если $\mathbb{E}(Z) = 1$, $\mathrm{Var}(Z) = 1$

   (d) Найдите точные значения, если дополнительно известно, что $X \sim \mathcal{N}(\mu; \sigma^2)$, $Y \sim U[0; 20]$ и $Z \sim Exp(1)$.

2. Известно, что $\mathbb{E}(X) = 100$, какой должна быть дисперсия величины X, чтобы вне зависимости от закона распределения величины X можно было бы гарантировать, что $\mathbb{P}(X \in [90; 110]) \geqslant 0.95$? А как решить аналогичный вопрос для $\mathbb{P}(X \in [90; 100]) \geqslant 0.95$?

3. Известно, что X — неотрицательная случайная величина с $\mathbb{E}(X) = 10$. В каких пределах может лежать вероятность $\mathbb{P}(X < 20)$?

12 Листок. Полный беспредел

Предел по вероятности, $\operatorname{plim}_{n\to\infty} X_n = X$:

$$\mathbb{P}(|X_n - X| > \varepsilon) \to 0,\ \forall \varepsilon > 0$$

Закон Больших Чисел (формулировка Хинчина): Если X_i независимы, одинаково распределены и математическое ожидание $\mathbb{E}(X_i)$ существует, то $\operatorname{plim} \bar{X}_n = \mathbb{E}(X_1)$

Центральная Предельная Теорема (формулировка Линдеберга-Леви): Если X_i независимы, одинаково распределены с математическим ожиданием μ и дисперсией σ^2, то:

$$\frac{\bar{X} - \mu}{\sqrt{\frac{\sigma^2}{n}}} \to \mathcal{N}(0;1) \text{ (строго)}$$

$$\bar{X}_n \approx \mathcal{N}\left(\mu; \frac{\sigma^2}{n}\right) \text{ (практично)}$$

Дельта-метод: Если Z_n асимптотически нормальны, то есть

$$\frac{Z_n - \mu}{\sqrt{\frac{\sigma^2}{n}}} \to \mathcal{N}(0;1),\ \text{(строго)}$$

$$Z_n \approx \mathcal{N}\left(\mu; \frac{\sigma^2}{n}\right) \text{ (практично)}$$

и f — дифференциируемая функция, такая что $f'(\mu) =\neq 0$,
то:

$$\frac{f(Z_n) - f(\mu)}{\sqrt{\frac{(f'(\mu))^2\sigma^2}{n}}} \to \mathcal{N}(0;1) \text{ (строго)}$$

$$f(Z_n) \approx \mathcal{N}\left(f(\mu); \frac{(f'(\mu))^2\sigma^2}{n}\right) \text{ (практично)}$$

1. Величины X_1, …, X_n независимы и равномерно распределены на отрезке $[0;1]$. Найдите $\operatorname{plim} \bar{X}_n$, $\operatorname{plim} 1/(1+\bar{X}_n)$, $\operatorname{plim} \sum_{i=1}^n \ln X_i/n$, $\operatorname{plim} \sqrt[n]{X_1 \cdot \ldots \cdot X_n}$, $\operatorname{plim}(X_1 \cdot \ldots \cdot X_n)$, $\operatorname{plim} \max\{X_1, \ldots, X_n\}$, $\operatorname{plim} \min\{X_1, \ldots, X_n\}$, $\operatorname{plim} \frac{\sum (X_i - \bar{X})^2}{n}$, $\operatorname{plim} X_1/\bar{X}$.

2. Величины X_1, …, X_n независимы и равномерно распределены на отрезке $[0;1]$. Найдите примерный закон распределения величин $\bar{X}_{100}$, $S_{100} = X_1 + \ldots + X_{100}$. Найдите примерно $\mathbb{P}(\bar{X}_{100} > 0.55)$, $\mathbb{P}(S_n \in [50;60])$, $\mathbb{E}(\bar{X}_{100} | \bar{X}_{100} > 0.6)$. Найдите такое число a, для которого $\mathbb{P}(S_n < a) = 0.65$.

3. Количество смс за сутки, посылаемое каждым из 160 абонентов, имеет пуассоновское распределение со средним значением 5 смс в сутки. Какова вероятность того, что за двое суток абоненты пошлют в сумме более 1700 сообщений?

4. Каждый день цена акции равновероятно поднимается или опускается на один рубль. Сейчас акция стоит 1000 рублей. Найдите вероятность того, что через сто дней акция будет стоить больше 1030 рублей.

5. Вероятность выпадения монетки «орлом» равна 0.63.

(a) Какова вероятность, что в 100 испытаниях выборочная доля выпадения орлов будет отличаться от истинной вероятности менее, чем на 0.07?

(b) Каким должно быть минимальное количество испытаний, чтобы вероятность отличия выборочной доли и истинной вероятности менее чем на 0.02 была больше 0.95?

6. Величины X_1, …, X_n независимы и одинаково распределены с математическим ожиданием 10 и дисперсией 20. Найдите примерный закон распределения величин $\bar{X}^2$, $(1+\bar{X})/(\bar{X}^2+5)$ при большом n.

7. Величина X имеет биномиальное распределение $Bin(n,p)$ и n велико. Какое распределение примерно имеют величины $\ln(X/n)$? $X/(n-X)$?

8. (*) Случайные величины X и Y независимы и равновероятно принимают значения от 1 до n. Вероятность того, что сумма $X+Y$ является квадратом натурального числа, обозначим p_n. Найдите предел $\lim_{n\to\infty} \sqrt{n}p_n$.

9. (*) Величины X_1, …, X_n независимы и равномерны на отрезке $[0;1]$. Вероятность того, что сумма любых двух соседних иксов меньше единицы, обозначим p_n. Найдите предел $\lim_{n\to\infty} p_n^{1/n}$.

13 Листок. Многомерное нормальное

1. Ермолай Лопахин решил приступить к вырубке вишневого сада. Однако выяснилось, что растут в нём не только вишни, но и яблони. Причём, по словам Любови Андреевны Раневской, среднее количество деревьев (а они периодически погибают от холода или жары, либо из семян вырастают новые) в саду распределено в соответствии с нормальным законом (X — число яблонь, Y — число вишен) со следующими параметрами:

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 25 \\ 125 \end{pmatrix}; \begin{pmatrix} 5 & 4 \\ 4 & 10 \end{pmatrix}\right) \tag{1}$$

Найдите вероятность того, что Ермолаю Лопахину придется вырубить более 150 деревьев. Каково ожидаемое число подлежащих вырубке вишен, если известно, что предприимчивый и последовательный Лопахин, не затронув ни одного вишнёвого дерева, начал очистку сада с яблонь и все 35 яблонь уже вырубил?

Автор: Кирилл Фурманов, Ира Чернухина

2. В самолете пассажирам предлагают на выбор «мясо» или «курицу». В самолет 250 мест. Каждый пассажир с вероятностью 0.6 выбирает курицу, и с вероятностью 0.4 — мясо. Сколько порций курицы и мяса нужно взять, чтобы с вероятностью 99% каждый пассажир получил предпочитаемое блюдо, а стоимость «мяса» и «курицы» для компании одинаковая?

Как изменится ответ, если компания берет на борт одинаковое количество «мяса» и «курицы»?

3. Сэр Фрэнсис Гальтон — учёный XIX-XX веков, один из основоположников как генетики, так и статистики — изучал, среди всего прочего, связь между ростом детей и родителей. Он исследовал данные о росте 928 индивидов. Обозначим X_1 — рост случайного человека, а X_2 — среднее арифметическое роста его отца и матери. По результатам исследования Гальтона:

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim \mathcal{N}\left[\begin{pmatrix} 68.1 \\ 68.3 \end{pmatrix}; \begin{pmatrix} 6.3 & 2.1 \\ 2.1 & 3.2 \end{pmatrix}\right]$$

(a) Обратите внимание на то, что дисперсия роста детей выше дисперсии среднего роста родителей. С чем это может быть связано? Учтите, что рост детей измерялся уже по достижении зрелости, так что разброс не должен быть связан с возрастными различиями.

(b) Рассчитайте корреляцию между X_1 и X_2

(c) Один дюйм примерно равен 2.54 сантиметра. Пусть X_1' и X_2' — это те же X_1 и X_2, только измеренные в сантиметрах. Найдите вектор математических ожиданий и ковариационную матрицу вектора $X' = (X_1', X_2')$.

(d) Определите, каков ожидаемый рост и дисперсия роста человека, средний рост родителей которого составляет 72 дюйма?

(e) Найдите вероятность того, что рост человека превысит 68 дюймов, если средний рост его родителей равен 72 дюймам. Подсказка: используйте предыдущий пункт и нормальность распределения!

Автор: Кирилл Фурманов, Ира Чернухина

4. «Регрессия к среднему»

Каждый день независимо от других муж дарит Машке случайное количество роз. Логарифм количества роз (в тысячах цветов), подаренных мужем в день t, ε_t, имеет нормальное распределение $N(0, 1)$. Улыбчивость Машки в день t, обозначаемая Y_t, зависит от количества цветов, подаренных в этот и в предыдущей день, $Y_t = \varepsilon_t + 0.5\varepsilon_{t-1}$. Подружка Машки не наблюдает подарки Машкиного мужа, ε_t, однако видит улыбчивость Машки, Y_t. Машкина подружка хочет спрогнозировать завтрашнюю улыбчивость Машки, исходя из прошлой информации.

(a) Найдите $\mathrm{Corr}(Y_t, Y_{t-1})$

(b) Найдите прогноз Машкиной улыбчивости завтра при известной сегодняшней улыбчивости $\mathbb{E}(Y_t|Y_{t-1})$

(c) Проинтерпретируйте величину коэффициента при Y_{t-1}, становится ли Машка в среднем улыбчивей или грустней со временем?

(d) Исследователь Вениамин собрал данные по 1000 семей и изобразил диаграмму рассеяния в осях (рост мамы в 20 лет, рост дочки в 20 лет). Далее он провел наиболее похожую на эти точки прямую. Наклон этой прямой скорее всего будет около 1, меньше 1, больше 1?

(e) Если Вовочка плохо пишет контрольную, то его лишают мороженого. После этого успеваемость Вовочки как правило улучшается. В чём сходство этой ситуации с данной задачей?

(f) Найдите прогноз $\mathbb{E}(Y_t|Y_{t-1}, Y_{t-2})$

5. Пара случайных величин X и Y имеет совместное нормальное распределение:

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 1 \\ -2 \end{pmatrix}; \begin{pmatrix} 5 & 6 \\ 6 & 10 \end{pmatrix}\right)$$

(a) Найдите корреляцию X и Y

(b) Найдите собственные числа и собственные векторые ковариационной матрицы

(c) Постройте линии уровня совместной функции плотности

6. Случайная величина X нормально распределена, $N(0;4)$. При фиксированном X случайная величина Y нормально распределена, $N(2X-1,9)$.

   (a) Найдите безусловный закон распределения величины Y

   (b) Найдите $\mathrm{Cov}(X,Y)$

   (c) Найдите $\mathbb{P}(Y>2)$

   (d) Найдите $\mathbb{E}(X|Y)$, $\mathbb{E}(Y|X)$, $\mathrm{Var}(X|Y)$, $\mathrm{Var}(Y|X)$

7. Маша прячется от Медведей в случайной точке на числовой прямой. Место, где спряталась Маша — случайная величина X, имеющая нормальное распределение, $N(0;4)$. Каждый из n Медведей, обнюхав числовую прямую, имеет своё мнение о том, где спряталась Маша. Эти мнения — случайные величины Y_i. При фиксированном X случайные величины Y_1, …, Y_n условно независимы и нормально распределены, $N(X,9)$.

   (a) Михаилу Потапычу, Медведю номер 1, кажется, что Машей сильнее всего пахнет в точке Y_1. Где ему следует искать Машу, т.е. чему равно $\mathbb{E}(X|Y_1)$?

   (b) Теперь n Медведей объединились и зная Y_1, Y_2, …, Y_n хотят понять, где же разумнее всего искать Машу. Помогите им посчитать $\mathbb{E}(X|Y_1,Y_2,\ldots,Y_n)$:

      i. Найдите безусловный закон распределения вектора Y_1, …, Y_n

      ii. Маленькое техническое задание. Пусть I — единичная матрица, а S — матрица строевого леса, то есть матрица, в которой все элементы равны единицам. Найдите $(aI+bS)^{-1}$. Подсказка: ответ имеет вид $cI+dS$.

      iii. Найдите $\mathbb{E}(X|Y_1,\ldots,Y_n)$

8. Величины X_1 и X_2 имеют совместное нормальное распределение, причем каждая из них имеет стандартное нормальное распределение $\mathcal{N}(0;1)$, а корреляция между ними равна ρ. Найдите $\mathbb{P}(X_1>0,X_2>0)$.

9. Величина X имеет стандартное нормальное распределение $\mathcal{N}(0;1)$. Храбрая исследовательница Мишель подкидывает правильную монетку. Если монетка выпадает орлом, то Мишель домножает величину X на единицу, а иначе — на минус единицу, и получает величину Y.

   (a) Какое распределение имеет величина Y?

   (b) Верно ли, что пара X и Y имеет совместное нормальное распределение?

   (c) Чему равна корреляция X и Y?

   (d) Приведите пример таких случайных величин Z и W, что каждая из них имеет нормальное распределение, корреляция между ними равна 0.5, однако распределение пары (Z,W) не является совместным нормальным.

14 Листок. Статистика ноль

1. Имеется пять чисел: x, 9, 5, 4, 7. При каком значении x медиана будет равна среднему?

2. Измерен рост 25 человек. Средний рост оказался равным 160 см. Медиана оказалась равной 155 см. Машин рост в 163 см был ошибочно внесен как 173 см. Как изменятся медиана и среднее после исправления ошибки?

3. Возможно ли чисто теоретически, что риск катастрофы в расчете на 1 час пути больше для самолета, чем для автомобиля, а в расчете на 1 километр пути — наоборот?

4. Деканат утверждает, что если студента N перевести из группы A в группу B, то средний рейтинг каждой группы возрастет. Возможно ли это?

5. Есть три группы по 10 человек, две группы по 20 человек и одна группа по 40 человек. У каждой из групп свой преподаватель.

   (a) Каков средний размер группы, для которой читает лекции наугад выбранный профессор?

   (b) Каков средний размер группы, в которой учится наугад выбранный студент?

   (c) Творческий вопрос. Мы ловим студентов наугад и спрашиваем каждого размер группы, в которой он учится. Можно ли как-то восстановить средний размер группы с точки зрения преподавателя?

6. Приведите примеры случайных величин, для которых:

   (a) $\mathrm{Med}(X+Y)=\mathrm{Med}(X)+\mathrm{Med}(Y)$

   (b) $\mathrm{Med}(X+Y)\neq\mathrm{Med}(X)+\mathrm{Med}(Y)$

   (c) $\mathrm{Med}(X^k)=\mathrm{Med}(X)^k$ для всех k

   (d) $\mathrm{Med}(X^2)\neq\mathrm{Med}(X)^2$

7. Исследователь Вениамин измерил рост пяти случайно выбранных человек. Какова вероятность того, что истинная медиана роста лежит между минимумом и максимумом из этих пяти наблюдений? Предположим, что рост имеет непрерывное распределение.

8. Во время Второй Мировой войны американские военные собрали статистику попаданий пуль в фюзеляж самолёта. По самолётам, вернувшимся из полёта на базу, была составлена карта повреждений среднестатистического самолёта. С этими данными военные обратились к статистику Абрахаму Вальду с вопросом, в каких местах следует увеличить броню самолёта.

   Что посоветовал Абрахам Вальд и почему?

9. Два лекарства испытывали на мужчинах и женщинах. Каждый человек принимал только одно лекарство. Общий процент людей, почувствовавших улучшение, больше среди принимавших лекарство A. Процент мужчин, почувствовавших улучшение, больше среди мужчин, принимавших лекарство B. Процент женщин, почувствовавших улучшение, больше среди женщин, принимавших лекарство B.

   (a) Возможно ли это?

   (b) Какое лекарству нужно порекомендовать больному, не зная его пола?

10. Из набора чисел $\{2, 4, 10, 14\}$ случайным образом равновероятно по очереди выбираются три числа с возможностью повторения.

    (a) Найдите закон распеределения (табличку с вероятностями) величины X_1. Найдите закон распределения величины X_2.

    (b) Найдите совместный закон распределения пары X_1, X_2. Найдите совместный закон распределения пары X_1, X_3.

    (c) Являются величины X_1, X_2, X_3 независимыми? Одинаково распределеннными?

    (d) Верно ли, что $\mathbb{E}(X_1) = \mathbb{E}(X_2) = \mathbb{E}(X_3)$? Верно ли, что $\mathrm{Var}(X_1) = \mathrm{Var}(X_2) = \mathrm{Var}(X_3)$?

    (e) Верно ли, что $\mathrm{Cov}(X_1, X_2) = \mathrm{Cov}(X_1, X_3)$?

    (f) Как изменятся ответы на предыдущие вопросы, если числа выбираются без возможности повторения?

11. Из фиксированного множества N чисел случайным образом выбирают n чисел. Известно, что если бы выбирать наугад всего одно число, то тогда математическое ожидание и дисперсие этого одного случайного числа были бы равны $\mathbb{E}(X_1) = \mu$ и $\mathrm{Var}(X_1) = \sigma^2$. Чему равны $\mathbb{E}(\bar{X}_n)$ и $\mathrm{Var}(\bar{X}_n)$, если:

    (a) Мы выбираем n чисел из N с возвращениями

    (b) Мы выбираем n чисел из N без возвращений

    (c) Во что превращаются полученные формулы при $n = 1$? при $n = N$? при $N \to \infty$?

12. Исследовательница Мишель подбрасывает кубик 100 раз. Пусть X_1 — количество выпадений единицы, а X_6 — количество выпадений шестёрки.

    (a) Как распределена величина X_1? Величина X_6? Найдите $\mathbb{E}(X_1)$, $\mathrm{Var}(X_1)$.

    (b) Верно ли, что величины X_1 и X_6 независимы? Одинаково распереденны?

    (c) Найдите $\mathrm{Cov}(X_1, X_1 + X_2 + X_3 + X_4 + X_5 + X_6)$, $\mathrm{Cov}(X_1, X_6)$

    (d) Найдите $\mathrm{Corr}(X_1, X_6)$, проинтерпретируйте эту величину

13. В множестве A всего два чисел, $A = \{24, 42\}$. Случайным образом из множества A выбираются 3 числа с возможностью повторений. Явно найдите закон распределения выборочного среднего, выборочной медианы, выборочной моды, выборочного минимума и выборочного максимума.

15 Листок. Случайная выборка

1. Создайте случайную выборку объемом $n = 1000$ из равномерного на отрезке $[0; 1]$ распределения.

    (a) Найдите выборочные характеристики: среднее, медиану, минимум и максимум, стандартную ошибку, 10%-ый и 95%-ый квантили.

    (b) Постройте гистограмму распределения, выборочную функцию распределения для первых 20 чисел из случайной выборки

    (c) Повторите данный опыт для нормального $N(5, 1)$ распределения и для экспоненциального распределения с параметром $\lambda = 1$

(d) Насколько сильно выборочные характеристики отличаются от истинных?

2. Придумайте способ, как сгенерировать 100 одинаково распределенных случайных величин, таких что $\sum_{i=1}^{100} X_i = 50$. Будут ли эти величины X_i зависимы?Модифицируйте способ, так чтобы он давал одинаково распределенные величины, такие что $\sum_{i=1}^{100} Y_i^2 = 50$. Будут ли эти новые величины Y_i зависимы?

3. Придумайте детерминистическую функцию, такую, которая бы превращала одну равномерную на $[0;1]$ случайную величину X в

   (a) случайную величину Y, принимающую значения 1 и 0 с вероятностями 0.7 и 0.3 соответственно

   (b) случайную величину Z с функцией плотности $f(z) = 2z$ на отрезке $z \in [0;1]$

   (c) пару независимых одинаково распределенных случайных величин (Y_1, Y_2), принимающих значения 1 и 0 с вероятностями 0.7 и 0.3 соответственно

   (d) пару независимых равномерных на $[0;1]$ случайных величин

4. Постройте случайную выборку в $n = 200$ наблюдений из двумерного нормального распределения с параметрами:

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 25 \\ 125 \end{pmatrix}; \begin{pmatrix} 5 & 4 \\ 4 & 10 \end{pmatrix}\right) \tag{2}$$

   (a) Посчитайте выборочную ковариацию, выборочную корреляцию

   (b) Постройте диаграмму рассеяния, нанесите на диаграмму рассеяния линию $y(x) = E(Y|X = x)$

   (c) Насколько сильно выборочные характеристики отличаются от истинных?

5. Создайте 500 выборок объемом $n = 20$ каждая из равномерного на отрезке $[0;1]$ распределения и вычислите выборочное среднее для каждой из выборок.

   (a) Каково теоретическое математическое ожидание и дисперсия каждого из выборочных средних?

   (b) Постройте гистограмму выборочных средних

   (c) На фоне функции плотности стандартного нормального распределения изобразите в подходящем масштабе гистограмму стандартизированных выборочных средних

16 Максимально правдоподобно — 1!

1. Кот Матроскин каждый день приходит к пруду поймать «рыбку». В пруду встречаются караси, щуки и бегемоты. Кот Матроскин хочет оценить вероятность p поймать карася. От своей бабушки Кот Матроскин достоверно знает, что щуки встречаются в два раза чаще карасей.

   (a) Оцените $\hat{p}_{KM}$ методом максимального правдоподобия, если Кот Матроскин ловил «рыбку» четыре дня и имеются наблюдения: $X_1 =$ щука, $X_2 =$ карась, $X_3 =$ карась, $X_4 =$ бегемот.

(b) Постройте оценку $\hat{p}_{KM}$ методом максимального правдоподобия в общем виде. То есть, Кот Матроскин ходил на пруд n дней, поймал $Y_{\text{к}}$ карасей, $Y_{\text{щ}}$ щук и $Y_{\text{б}}$ бегемотов.

(c) Зависимы ли величины $Y_{\text{к}}$, $Y_{\text{щ}}$ и $Y_{\text{б}}$ бегемотов? Как распределена величина $Y_{\text{б}}$? Найдите $\mathbb{E}(Y_{\text{б}})$, $\mathrm{Var}(Y_{\text{б}})$.

(d) Найдите $\mathbb{E}(\hat{p}_{KM})$, $\mathrm{Var}(\hat{p}_{KM})$

(e) Является ли оценка $\hat{p}_{KM}$ несмещенной, состоятельной?

(f) Постройте аналогичную оценку $\hat{p}$ для Пса Шарика. В отличие от Кота Матроскина Пёс Шарик не знает, что щуки встречаются в два раза чаще карасей. Является ли оценка Пса Шарика несмещенной и состоятельной?

(g) Какая из двух оценок является более эффективной? Почему?

2. «Про зайцев». В темно-синем лесу, где трепещут осины, живут $n >> 0$ зайцев. Мы случайным образом отловили 100 зайцев. Каждому из них на левое ухо мы завязали бант из красной ленточки и потом всех отпустили. Через неделю будет снова отловлено 100 зайцев. Из них случайное количество S зайцев окажутся с бантами.

   (a) Постройте ML оценку для неизвестного параметра n

   (b) Рассмотрим случайные величины X_1, X_2, …, X_{100}, — индикаторы того, оказался ли соответствующий повторно отловленный заяц с бантом или без. С помощью этих величин постройте оценку для n методом моментов.

3. У Васи есть два одинаковых золотых слитка неизвестной массы m каждый и весы, которые взвешивают с некоторой погрешностью. Сначала Вася положил на весы один слиток и получил результат $Y_1 = m + u_1$, где u_1 — случайная величина, ошибка первого взвешивания. Затем Вася положил на весы сразу оба слитка и получил результат $Y_2 = 2m + u_2$, где u_2 — случайная величина, ошибка второго взвешивания. Оказалось, что $y_1 = 0.9$, а $y_2 = 2.3$.

   Используя ML оцените вес слитка m и параметр погрешности весов b, если

   (a) u_i — независимы и $N(0; b)$

   (b) u_i — независимы и $U[-b; b]$

4. Задача о немецких танках

   Предположим, что все выпущенные танки имеют порядковый номер. От самого первого выпущенного танка, имеющего номер 1, до самого последнего танка, имеющего номер n. В бою удалось подбить танки с номерами 15, 29 и 23.

   (a) Постройте оценку количества танков методом моментов

   (b) Постройте оценку количества танков методом максимального правдоподобия

   (c) Постройте несмещенную оценку количества танков с наименьшей дисперсией

17 Максимально правдоподобно — 2!

`Минитеория`

Метод моментов (MM, method of moments): найти θ из уравнения $\bar{X}_n = \mathbb{E}(X_i)$
Метод максимального правдоподобия (ML, maximum likelihood):
найти θ при котором вероятность получить имеющиеся наблюдения будет максимальной

Наблюдаемая информация Фишера: $\hat{I} = -\frac{\partial^2 l}{\partial^2 \theta}(\hat{\theta})$, $\widehat{\mathrm{Var}}(\hat{\theta}_{ML}) = \hat{I}^{-1}$.
Байесовский подход (bayesian approach):
1. Сделать изначальное предположение о распределении $\hat{\theta}$
2. Обновлять закон распределения $\hat{\theta}$ по формуле условной вероятности

Пусть $l(\theta)$ — логарифмическая функция правдоподобия ($l(\theta) = \ln(f(X_1, \ldots, X_n, \theta))$).
Ожидаемая информация Фишера $I(\theta) = \mathbb{E}\left[\left(\frac{\partial l}{\partial \theta}\right)^2\right] = -E\left(\frac{\partial^2 l}{\partial \theta^2}\right)$
Сколько информации о неизвестном θ содержится в выборке $X_1, \ldots, X_n$

Неравенство Крамера-Рао (Cramer-Rao) («слишком хорошей оценки не бывает»):
Если $\hat{\theta}$ - несмещенная оценка и ..., то $\mathrm{Var}(\hat{\theta}) \geqslant \frac{1}{I(\theta)}$

Оценки ML — самые лучшие (асимптотически несмещенные и с минимальной дисперсий):
Если X_i — iid, ..., и $n \to \infty$ то $\hat{\theta}_{ML} \sim N(\theta, \frac{1}{I(\theta)})$.

1. Допустим, что X_i — независимы и имеют закон распределения, заданный табличкой:

X	-1	0	2
$\mathbb{P}()$	θ	$2\theta - 0.2$	$1.2 - 3\theta$

Имеется выборка: $X_1 = 0$, $X_2 = 2$.

   (a) Найдите оценки $\hat{\theta}_{ML}$ и $\hat{\theta}_{MM}$

   (b) Первоначально ничего о θ не было известно и поэтому предполагалось, что θ распределена равномерно на $[0.1; 0.4]$. Как выглядит условное распределение θ, если известно что $X_1 = 0$, $X_2 = 2$?

   (c) Постройте ML и MM оценки для произвольной выборки $X_1, X_2, \ldots X_n$

2. Пусть Y_1 и Y_2 независимы и распределены по Пуассону. Известно также, что $\mathbb{E}(Y_1) = e^a$ и $\mathbb{E}(Y_2) = e^{a+b}$. Найдите ML оценки для a и b.

3. Пусть $X_1, \ldots, X_n$ распределены одинаково и независимо. Оцените значение θ с помощью ML (везде) и MM (в «а» и «б»), оцените дисперсию ML оценки, если функция плотности X_i, $p(t)$ имеет вид:

   (a) $\theta t^{\theta-1}$ при $t \in [0; 1]$;
   (b) $\frac{2t}{\theta^2}$ при $t \in [0; \theta]$

   (c) $\frac{\theta e^{-\frac{\theta^2}{2t}}}{\sqrt{2\pi t^3}}$ при $t \in [0; +\infty)$;
   (d) $\frac{\theta\left(\ln^{\theta-1} t\right)}{t}$ при $t \in [1; e]$;
   (e) $\frac{e^{-|t|}}{2\left(1-e^{-\theta}\right)}$ при $t \in [-\theta; \theta]$

4. Пусть $X_1, \ldots, X_n$ - независимы и экспоненциальны с параметром λ. Постройте MM и ML оценки параметра λ. Оцените дисперсию ML оценки.

5. Пусть $X_1, \ldots, X_n$ - независимы и $N(\mu; \sigma^2)$. Значение σ^2 известно. Постройте MM и ML оценки параметра μ.

6. Пусть X_i независимы и одинаково распределены $N(\alpha, 2\alpha)$
По выборке $X_1, \ldots, X_n$ постройте оценку для α с помощью ML и MM. Оцените дисперсию ML оценки.

7. Пусть $Y_1 \sim N(0; \frac{1}{1-\theta^2})$. Найдите ML оценку для θ. Оцените дисперсию ML оценки.

8. Пусть X_1, X_2,..., X_n независимы и их функции плотности имеет вид:
$f(x) = \begin{cases} (k+1)x^k, & x \in [0;1]; \\ 0, & x \notin [0;1]. \end{cases}$
Найдите оценки параметра k с помощью ML и MM. Оцените дисперсию ML оценки.

9. Пусть X_1, X_2,..., X_n независимы и равномерно распределены на отрезке $[0;\theta]$, $\theta > 1$
   (a) Постройте MM и ML оценки для неизвестного θ.
   (b) Как изменятся ответы на «а», если исследователь не знает значений самих X_i, а знает только количество X_i оказавшихся больше единицы?

10. В озере водятся караси, окуни, щуки и налимы. Вероятности их поймать занесены в табличку

Fish	Карась	Окунь	Щука	Налим
$\mathbb{P}()$	0.1	p	p	$0.9-2p$

Рыбак поймал 100 рыб и среди пойманных 100 рыб он посчитал количества карасей, окуней, щук и налимов.

   (a) Постройте $\hat{p}_{ML}$
   (b) Найдите ожидаемую и наблюдаемую информацию Фишера
   (c) Несмещенная оценка $\hat{\theta}$ получена по 100 наблюдениям: X_1, ..., X_{100}. В каких пределах может лежать $\mathrm{Var}(\hat{\theta})$?

11. Известно, что X_i — независимы и имеют закон распределения, заданный таблицей:

X_i	0	1
$\mathbb{P}()$	p	$1-p$

   (a) Постройте $\hat{p}_{ML}$
   (b) Найдите ожидаемую и наблюдаемую информацию Фишера. Постройте возможные графики $I(p)$.
   (c) Пусть $\hat{\theta}$ — несмещенная оценка, полученная по 100 наблюдениям: X_1, ..., X_{100}. В каких пределах может лежать $\mathrm{Var}(\hat{\theta})$?

12. Пусть X_i независимы и имеют экспоненциальное распределение с параметром λ, т.е. $p(t) = \lambda e^{-\lambda t}$.
   (a) Найдите $I(\lambda)$, если наблюдаются X_1, ..., X_n
   (b) Пусть $\lambda = 1/\theta$, т.е. $p(t) = \frac{1}{\theta}e^{-\frac{1}{\theta}t}$. Найдите $I(\theta)$, если наблюдается X_1, ..., X_n

13. Пусть X_i — независимы и одинаково распределены. Пусть $I_{X_i}(\theta)$ - информация Фишера о θ, получаемая при наблюдении X_i.
   (a) Верно ли, что $I_{X_1}(\theta) = I_{X_2}(\theta)$?
   (b) Как найти $I_{X_1,\ldots,X_n}(\theta)$ зная $I_{X_i}(\theta)$?

14. Пусть X — равномерна на участке $[0;2a]$. С какой вероятностью интервал $[0.9X;1.1X]$ накрывает неизвестное a? Постройте 95%-ый доверительный интервал для a вида $[0;kX]$.

15. Пусть X — экспоненциальна с параметром λ и $\mu = \mathbb{E}(X)$. С какой вероятностью интервал $[0.9X; 1.1X]$ накрывает μ? Постройте 90%-ый доверительный интервал для μ вида $[0; kX]$.

16. Пусть X_i — независимы и нормальны $N(\mu, 1)$. Какова вероятность того, что интервал $[\bar{X}_{10} - 1; \bar{X}_{10} + 1]$ накроет неизвестное μ? Постройте 90%-ый доверительный интервал для μ вида $[\bar{X}_{10} - k; \bar{X}_{10} + k]$.

17. Величины $X_1, \ldots, X_n$ — независимы и одинаково распределены с функцией плотности $f(t) = \frac{\theta(\ln t)^{\theta-1}}{t}$ при $t \in [1; e]$. По выборке из 100 наблюдений оказалось, что $\sum \ln(\ln(X_i)) = -30$

    (a) Найдите ML оценку параметра θ и ожидаемую и наблюдаемую информацию Фишера

    (b) Постройте 95% доверительный интервал для θ

18. Величины $X_1, \ldots, X_n$ — независимы и одинаково распределены с функцией плотности $\frac{\theta e^{-\frac{\theta^2}{2t}}}{\sqrt{2\pi t^3}}$ при $t \in [0; +\infty)$. По выборке из 100 наблюдений оказалось, что $\sum 1/X_i = 12$

    (a) Найдите ML оценку параметра θ и информацию Фишера $I(\theta)$

    (b) Пользуясь данными по выборке постройте оценку $\hat{I}$ для информации Фишера

    (c) Постройте 90% доверительный интервал для θ Hint: $\mathbb{E}(1/X_i) = 1/\theta^2$ (интеграл берется заменой $x = \theta^2 a^{-2}$)

19. Время, которое Вася тратит на задачу — равномерно распределенная случайная величина: на простую — от 1 до 15 минут, на сложную — от 10 до 20 минут. Известно, что на некую задачу Вася потратил 13 минут.

    (a) С помощью метода максимального правдоподобия определите, простая она или трудная.

    (b) С помощью байесовского подхода посчитайте вероятности того, что задача была простая, если на экзамене было 7 легких и 3 трудных задачи.

20. Известно, что X_i независимы, $\mathbb{E}(X_i) = 5$, $\mathrm{Var}(X_i) = 4$ и n велико. Как примерно распределены следующие величины:

    (a) $\bar{X}_n$,

    (b) $Y_n = (\bar{X}_n + 3)/(\bar{X}_n + 6)$,

    (c) $Z_n = \bar{X}_n^2$,

    (d) $W_n = 1/\bar{X}_n$

21. Известно, что X_i независимы и равномерны на $[0; 1]$.

    (a) Найдите $\mathbb{E}(\ln(X_i))$, $\mathrm{Var}(\ln(X_i))$, $\mathbb{E}(X_i^2)$, $\mathrm{Var}(X_i^2)$

    (b) Как примерно распределены величины $X_n = \frac{\sum \ln(X_i)}{n}$, $Y_n = (X_1 \cdot X_2 \cdots X_n)^{1/n}$, $Z_n = \left(\frac{\sum X_i^2}{n}\right)^3$ при больших n?

22. Величины X_i независимы и имеют функцию плотности $f(x) = a \cdot x^{a-1}$ на отрезке $[0; 1]$.

    (a) Постройте оценку $\hat{a}$ методом моментов, укажите ее примерный закон распределения

(b) По 100 наблюдениям оказалось, что $\sum X_i = 25$. Посчитайте численное значение $\hat{a}$ и оцените дисперсию случайной величины $\hat{a}$.

23. Начинающий каратист Вася тренируется бить кирпичи ударом ладони. Каждый день он бьёт ладонью по кирпичу до пор, пока тот не расколется от одного удара. Предположим, что вероятность разбить кирпич с одного удара равна p и неизменна во времени. Величины X_1, X_2, …, X_n — количества ударов которые потребовались Васе в соответствующий день.

    (a) Найдите оценку p методом максимального правдоподобия

    (b) Найдите достаточную статистику T

    (c) Выразите $\widehat{\mathrm{Var}}(\hat{p})$ через достаточную статистику T

    (d) Найдите $\mathbb{P}(X_1 = 1 \mid T = t)$.

24. Продавщица Глафира отдаёт псу Шарику в конце каждого дня нерасфасованные остатки мясного фарша. Фарш фасуется упаковками по a грамм, поэтому нерасфасованный остаток в i-ый день, X_i, случаен и равномерно распределен на отрезке $[0; a]$. Пёс Шарик хорошо помнит все X_1, …, X_n. Помогите псу Шарику:

    (a) Найдите оценку a методом максимального правдоподобия

    (b) Найдите достаточную статистику T

    (c) Выразите $\widehat{\mathrm{Var}}(\hat{a})$ через достаточную статистику T

    (d) Найдите $\mathbb{P}(X_1 < 10 \mid T = t)$.

25. Величины X_i равномерны на отрезке $[-a; 3a]$ и независимы. Есть несколько наблюдений, $X_1 = 0.5$, $X_2 = 0.7$, $X_3 = -0.1$.

    (a) Найдите $\mathbb{E}(X_i)$ и $\mathbb{E}(|X_i|)$

    (b) Постройте оценку метода моментов, используя $\mathbb{E}(X_i)$

    (c) Постройте оценку метода моментов, используя $\mathbb{E}(|X_i|)$

    (d) Постройте оценку обобщёного метода моментов используя моменты $\mathbb{E}(X_i)$, $\mathbb{E}(|X_i|)$ и взвешивающую матрицу
    $$W = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$$

    (e) Найдите оптимальную теоретическую взвешивающую матрицу для обобщённого метода моментов

    (f) Постройте двухшаговую оценку обобщённого метода моментов, начав со взвешивающей матрицы W

26. Величины X_i имеют Пуассоновское распределение с параметром λ и независимы. Есть несколько наблюдений, $X_1 = 5$, $X_2 = 7$, $X_3 = 1$.

    (a) Найдите $\mathbb{E}(X_i)$ и $\mathbb{E}((X_i - \bar{X})^2)$

    (b) Постройте оценку метода моментов, используя $\mathbb{E}(X_i)$

    (c) Постройте оценку метода моментов, используя $\mathbb{E}((X_i - \bar{X})^2)$

(d) Постройте оценку обобщёного метода моментов используя моменты $\mathbb{E}(X_i)$, $\mathbb{E}((X_i-\bar{X})^2)$ и взвешивающую матрицу

$$W = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}$$

(e) Найдите оптимальную теоретическую взвешивающую матрицу для обобщённого метода моментов

(f) Постройте двухшаговую оценку обобщённого метода моментов, начав со взвешивающей матрицы W

18 Листок. Проверка гипотез

1. Кальямпуди Радхакришна Рао и Карл Харальд Крамер строят доверительный интервал для μ по случайной выборке из n наблюдений. Наблюдения, величины X_i, распределены нормально и независимы друг от друга. Рао знает величину σ^2; а Крамер не знает, и поэтому вынужден использовать $\hat{\sigma}^2$ при проверке гипотезы.

   (a) Какова вероятность того, что Рао получит более короткий доверительный интервал при $n = 10$?

   (b) К чему стремится эта вероятность при $n \to \infty$?

2. Как распределено P-значение при верной H_0? Вовочка использует следующий статистический критерий: «Если P-значение больше 0.95, то H_0 отвергается». Чему де-факто равна вероятность ошибки первого рода в этом случае? Разумно ли использовать данный критерий?

19 Листок $g()$. И потребленье возрастает, а производство отстаёт!

Обобщенный бином Ньютона для произвольной степени $a \in \mathbb{R}$:

$$(1+x)^a = \sum_{k=0}^{\infty} C_a^k x^k, \quad \text{где } C_a^k = \frac{a(a-1)(a-2)\cdot \ldots \cdot (a-k+1)}{k!}$$

1. Найдите коэффициент при x^{17} в многочлене $(1+x^5+x^7)^{20}$.

2. Найдите и запишите максимально просто производящие функции для последовательностей

   (a) 1, 1, 1, 1, 0, 0, 0, ...

   (b) 1, 1, 1, 1, 1, 1, 1, ...

   (c) 1, 4, 6, 4, 1, 0, 0, ...

3. Найдите первые четыре коэффициента в многочленах $(1+2x)^{-3}$ и $(1+x+x^2)^{-4}$.

4. Кубик подбрасывают три раза. Обозначим результаты подбрасываний X_1, X_2 и X_3. Найдите производящие функции для величин X_1, X_1+X_2 и $X_1+X_2+X_3$.

5. Производящая функция случайной величины X линейна и проходит через точки $(1,1)$ и $(5,4)$. Какие значения принимает величина X и с какими вероятностями?

6. На месте убийства преступники оставили улику — маленький-маленький кусочек графика производящей функции. Сможет ли Шерлок Холмс по этому кусочку восстановить все вероятности?