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:

Задача 1.1. *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* $\mathbb{P}(X=2)$, $\mathbb{P}(X>1)$, $\mathbb{E}(X)$

Задача 1.2. *В коробке находится четыре внешне одинаковые лампочки, две из них исправны. Лампочки извлекают из коробки по одной до тех пор, пока не будут извлечены обе исправные.*

1. *Какова вероятность того, что опыт закончится извлечением трёх лампочек?*
2. *Каково ожидаемое количество извлеченных лампочек?*

Задача 1.3. *Маша подкидывает монетку. Если она выпадает орлом, то Маша подкидывает монетку ещё один раз, если решкой — то ещё два раза. После этого Маша идёт в кино! Пусть* X *— количество выпавших орлов. Найдите вероятности* $\mathbb{P}(X=0)$, $\mathbb{P}(X=1)$, … *и ожидание* $\mathbb{E}(X)$.

Задача 1.4. *Две команды равной силы играют в волейбол до трёх побед одной из них, не обязательно подряд. Ничья невозможна. Из-за равенства сил будем считать, что вероятность победы каждой равна* 0.5. *Величина* N *— количество сыгранных партий. Составьте табличку возможных значений* N *с их вероятностями. Найдите* $\mathbb{P}(N$ *— чётное*$)$, $\mathbb{E}(N)$

Задача 1.5. *Какова вероятность того, что у 30 человек не будет ни одного совпадения дней рождений? Сколько человек должно собраться, чтобы вероятность совпадения дней рождения превысила* $1/2$*? Сколько в среднем человек должно войти в комнату, чтобы впервые произошло совпадения дней рождения?*

Задача 1.6. *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?*

Задача 1.7. *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*?*

Задача 1.8. *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.*

1. *Какова вероятность того, что игра рано или поздно закончится выпадением 6-ки?*
2. *Какова ожидаемая продолжительность игры?*
3. *Чему равен ожидаемый выигрыш в эту игру?*
4. *Чему равен ожидаемый выигрыш в эту игру, если изначально на кону лежит 100 рублей?*

5. *Изменим условие: если выпадает 5, то сумма на кону сгорает, а игра продолжается. Как изменятся ответы на предыдущие вопросы?*

Задача 1.9. *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.*

1. *У кого какие шансы выиграть?*
2. $\mathbb{P}(X = 4)$, $\mathbb{P}(Y = 1)$, $\mathbb{E}(X)$, $\mathbb{E}(Y)$
3. *Решите аналогичную задачу для ОРО и ООР.*

Задача 1.10. *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.*

1. *Как выглядит оптимальная стратегия? Чему равен ожидаемый выигрыш при использовании оптимальной стратегии?*
2. *Какова средняя продолжительность игры при использовании оптимальной стратегии?*
3. *Как выглядит оптимальная стратегия и чему равен ожидаемый выигрыш, если за каждое подбрасывание Вася платит 35 копеек?*

Задача 1.11. *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* $\mathbb{P}(X = 4)$, $\mathbb{E}(X)$

Задача 1.12. *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 Хочу ещё задач!

Задача 2.1. *Наугад из четырех тузов разных мастей выбираются два.* $\mathbb{P}($*они будут разного цвета*$)$?

Задача 2.2. *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] $\mathbb{P}(A^c \cap B^c)$.

Задача 2.3. *Вероятность* $\mathbb{P}(A) = 0.3$, $\mathbb{P}(B) = 0.8$. *В каких пределах может лежать* $\mathbb{P}(A \cap B)$?

Задача 2.4. *Множество исходов* $\Omega = \{a, b, c\}$, $\mathbb{P}(\{a, b\}) = 0,8$, $\mathbb{P}(\{b, c\}) = 0,7$. *Найдите* $\mathbb{P}(\{a\})$, $\mathbb{P}(\{b\})$, $\mathbb{P}(\{c\})$

Задача 2.5. *«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* $\mathbb{P}(X = 2)$, $\mathbb{P}(X = 3)$, $\mathbb{P}(X = \infty)$

Задача 2.6. *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?*

---

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

Задача 2.7. *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?*

Задача 2.8. *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.*

1. *Как выглядит оптимальная стратегия в этой игре, если Вы заинтересованы в максимальном среднем выигрыше?*

2. *Чему при этом будет равен средний выигрыш?*

Задача 2.9. *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, например, может зайти в одну и ту же. Сейчас мышка в коридоре. Сколько времени ей в среднем потребуется, чтобы найти сыр?*

Задача 2.10. *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?*

Задача 2.11. *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?*

Задача 2.12. *Треугольник с вершинами $(0; 0)$, $(2; 0)$ и $(1; 1)$. Внутри него случайным образом выбирается точка, X – абсцисса точки. Найдите $\mathbb{P}(X > 1)$, $\mathbb{P}(X \in [0.5; 1])$, $\mathbb{E}(X)$*

Задача 2.13. *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?*

Задача 2.14. *Исследовательница Мишель подкидывает игральный кубик неограниченное количество раз и складывает выпадающие количества очков.*

1. *Чему примерно равна вероятность того, что однажды сумма в точности будет равна 123456789?*

2. *Чему точно равна указанная вероятность?*

Задача 2.15. *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?*

Задача 2.16. *Исследовательница Мишель хочет встать утром с правой ноги с вероятностью* $1/\sqrt{2}$, *и с левой с вероятностью* $1 - 1/\sqrt{2}$. *Однако для проведения случайных экспериментов у неё есть только одна правильная монетка. Как с помощью правильной монетки ей добиться цели?*

3 К чёрту условности!

Задача 3.1. *Имеется три монетки. Две «правильных» и одна — с «орлами» по обеим сторонам. Петя выбирает одну монетку наугад и подкидывает её два раза. Оба раза выпадает «орел». Какова условная вероятность того, что монетка «неправильная»?*
Задача 3.2. *Два охотника одновременно выстрелили в одну утку. Первый попадает с вероятностью 0.4, второй — с вероятностью 0.7.*

1. *Какова вероятность того, что в утку попала ровно одна пуля?*
2. *Какова условная вероятность того, что утка была убита первым охотником, если в утку попала ровно одна пуля?*

Задача 3.3. *Кубик подбрасывается два раза. Найдите вероятность получить сумму равную 8, если при первом броске выпало 3.*
Задача 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?*
Задача 3.5. *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?*
Задача 3.6. *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?*
Задача 3.7. *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?*
Задача 3.8. *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?*
Задача 3.9. *Ген карих глаз доминирует ген синих. Следовательно, 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 кареглазых родителей два сына — кареглазый и синеглазый. Кареглазый женился на синеглазой девушке. Какова вероятность рождения у них синеглазого ребенка?*
Задача 3.10. *Из колоды в 52 карты извлекается одна карта наугад. Являются ли события «извлечен туз» и «извлечена пика» независимыми?*
Задача 3.11. *Из колоды в 52 карты извлекаются по очереди две карты наугад. Являются ли события «первая карта — туз» и «вторая карта — туз» независимыми?*
Задача 3.12. *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* $\mathbb{P}(A \cup B \cup C)$.
Задача 3.13. *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.*

1. *Известно, что хотя бы один ребенок — мальчик.*

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

3. *Известно, что старший ребенок — мальчик.*

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

4 Use Julia/R/python/...or die!

Задача 4.1. *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 :)*

Задача 4.2. *Случайные величины* $X_1, \ldots, X_5$ *имеют равномерное распределение на отрезке* $[0;1]$ *и независимы. С помощью симуляций оцените* $\mathbb{P}(\min\{X_1,\ldots,X_5\} > 0.2)$, $\mathbb{P}(\min\{X_1,\ldots,X_5\} > 0.2 \mid X_1+X_2 < 0.5)$, $\mathbb{E}(\min\{X_1,\ldots,X_5\})$, $\mathbb{E}(\min\{X_1,\ldots,X_5\} \mid X_1+X_2 < 0.5)$

Задача 4.3. *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!*

Задача 4.4. *В колоде 52 карты. Мы вытаскиваем карты из колоды до первого туза, пусть* X *— количество вытянутых карт. С помощью симуляций оцените* $\mathbb{E}(X^2)$, $\mathbb{P}(X > 10)$, $\mathbb{P}(X > 5 \mid X < 15)$, $\mathbb{E}(X^2 \mid X < 15)$

Задача 4.5. *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* $\mathbb{P}(Y > 3X)$, $\mathbb{E}(XY)$, $\mathbb{P}(Y > 3X \mid Y < X^2)$, $\mathbb{E}(XY \mid Y < X^2)$

5 Эф большое и эф малое

Минитеория:

1. $F_X(t) = \mathbb{P}(X \leqslant t)$
2. $F(t) = \int_{-\infty}^{t} f(a)\,da$, $f(t) = F'(t)$.

Задачи:

Задача 5.1. *Функция плотности случайной величины* X *равна* 5 *при* $x = 7$. *Найдите примерно вероятность того, что* X *попадёт в отрезок* $[7; 7.001]$.

Задача 5.2. *Случайная величина* X *имеет функцию плотности* f. *С помощью* $o(\Delta)$ *и* $f(x)$ *запишите вероятность* $\mathbb{P}(X \in [x; x + \Delta])$.

Задача 5.3. *Величина* X *распределена на отрезке* $[0; 1]$ *и на нём имеет функцию плотности* $f(t) = 3t^2$. *Найдите функции плотности и функции распределения величин* $Y = \ln X$, $Z = X^2$, $W = (X - 0.5)^2$.

Задача 5.4. *Может ли функция плотности принимать значение больше 2015? Может ли предел* $\lim_{x\to\infty} f(x)$ *не равняться нулю?*

Задача 5.5. *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:

1. $\mathbb{P}(X > 1)$, $\mathbb{E}(X)$, $\mathbb{E}(X^2)$, $\mathrm{Var}(X)$, σ_X
2. $\mathbb{E}(X|X > 1)$, $\mathbb{E}(X^2|X > 1)$, $\mathrm{Var}(X|X > 1)$
3. *Функцию распределения случайной величины* X
4. *Медиану величины* X, *40%-ую квантиль величины* X

Задача 5.6. *Функция плотности случайной величины* X *имеет вид:*

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

Найдите:

1. $\mathbb{P}(X > 1)$, $\mathbb{E}(X)$, $\mathbb{E}(X^2)$, $\mathrm{Var}(X)$, σ_X
2. $\mathbb{E}(X|X > 1)$, $\mathbb{E}(X^2|X > 1)$, $\mathrm{Var}(X|X > 1)$
3. *Функцию распределения случайной величины* X
4. *Медиану величины* X, *40%-ую квантиль величины* X

Задача 5.7. *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

Задача 5.8. *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.*

Задача 5.9. *Величина* X *равномерна на отрезке* $[-2; 1]$, *а величина* Y *— это расстояние от числа* X *до числа* (-1). *Найдите фукнцию плотности* Y, $\mathbb{E}(Y)$.

Задача 5.10. *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.

Задача 5.11. *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.*

1. *Найдите функцию распределения потерь от пожара.*
2. *Чему равны средние потери?*

Задача 5.12. *Пусть* X *— неотрицательная случайная величина с функцией плотности* $f(t)$ *и* $\mathbb{E}(X) < \infty$. *При каком* c *функция* $g(t) = c \cdot t \cdot f(t)$ *также будет функцией плотности?*

Задача 5.13. *Завтрашняя цена акции — случайная величина с функцией плотности* $f(x) = \frac{3}{4}\max\{x(2-x), 0\}$.

1. *Постройте график функции плотности*
2. *Найдите функцию распределения Васиного дохода, средний доход и дисперсию дохода, если:*
   (a) *У Васи есть 10 акций*
   (b) *У Васи есть опцион-пут, дающий ему право продать акции по цене 1.2 рубля*
   (c) *У Васи есть опцион-колл, дающий ему право купить акции по цене 1 рубль*

Задача 5.14. *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.

1. *Найдите функцию распределения и функцию плотности длины наилучшего прыжка*
2. *Найдите функцию распределения и функцию плотности длины наихудшего прыжка*
3. *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.*
4. *Найдите функцию распределения и функцию плотности длины прыжка бронзового призёра соревнований*

Задача 5.15. *Светофор попеременно горит пешеходу то 40 секунд зелёным, то 60 секунд — красным. Законопослушный Вася Бубликов подходит к светофору в случайный момент времени. Пусть* X *— время ожидания до возможности перейти дорогу.*

1. *Найдите функцию распределения величины* X *и постройте её график*
2. *Найдите* $\mathbb{E}(X)$ *и* $\mathrm{Var}(X)$

Задача 5.16. *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.

1. *Найдите* $\mathbb{P}(Y < 2)$
2. *Постройте функцию плотности величины* Y
3. *Найдите* $\mathbb{E}(Y)$, $\mathrm{Var}(Y)$
4. *Найдите* $\mathrm{Cov}(X, Y)$

6 Разлагай и властвуй!

Задача 6.1. *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?*

Задача 6.2. *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* $\mathbb{E}(X)$, $\mathbb{E}(Y)$, $\mathrm{Var}(X)$, $\mathrm{Var}(Y)$.

Задача 6.3. *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?*

Задача 6.4. *«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.*

1. *Сколько в среднем пар окажутся ни разу не надетыми?*
2. *Сколько в среднем пар окажутся надетыми не менее двух раз?*

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

Задача 6.5. *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?*

Задача 6.6. *«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* $\mathbb{E}(X)$, $\mathrm{Var}(X)$.

Задача 6.7. *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?

Задача 6.8. *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.*

Задача 6.9. *По 10 коробкам наугад раскладывают 7 карандашей. Каково среднее количество пустых коробок? Дисперсия?*

Задача 6.10. *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?*

Задача 6.11. *Из колоды в 52 карты извлекается 5 карт. Сколько в среднем извлекается мастей? Достоинств? Тузов? Дисперсии этих величин?*

Задача 6.12. *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* $\mathrm{Var}(X)$

Задача 6.13. *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* $\mathrm{Var}(X)$

Задача 6.14. *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:

Задача 6.15. *На малом интервале времени вероятность одного происшествия примерно пропорциональна длине интервала,* $\mathbb{P}(X[a; a+\Delta] = 1) = \lambda\Delta + o(\Delta)$.

Задача 6.16. *На малом интервале времени несколько происшествий происходят существенно реже одного происшествия,* $\mathbb{P}(X[a; a+\Delta] \geqslant 2) = o(\Delta)$.

Задача 6.17. *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.

Задача 6.18. *Независимость приращений. Количество происшествий на непересекающихся интервалах времени независимы.*

То:

Задача 6.19. *Время между* $(i-1)$*-ым и* i*-ым происшествием,* Y_i*, имеет экспоненциальное распределение* $Y_i \sim exp(\lambda)$.

$$f(y) = \begin{cases} \lambda e^{-\lambda y}, \ y \geqslant 0 \\ 0, \ y < 0 \end{cases}$$
$$\mathbb{E}(Y_i) = 1/\lambda$$
$$\mathrm{Var}(Y_i) = 1/\lambda^2$$

Задача 6.20. *Количество происшествий за единицу времени,* X*, имеет пуассоновское распределение* $X \sim Pois(\lambda)$

$$\mathbb{P}(X = k) = e^{-\lambda}\frac{\lambda^k}{k!}$$
$$\mathbb{E}(X) = \lambda$$
$$\mathrm{Var}(X) = \lambda$$

Задача 6.21. *Отсюда смысл* λ *— среднее количество событий за единицу времени, дисперсия количества событий за единицу времени*

Задача 6.22. *Количество событий за период времени* $[0; t]$*, величина* X_t*, имеет пуассоновское распределение* $X_t \sim Pois(\lambda t)$

$$\mathbb{P}(X_t = k) = e^{-\lambda t} \cdot \frac{(\lambda t)^k}{k!}$$
$$\mathbb{E}(X_t) = \lambda t$$
$$\mathrm{Var}(X_t) = \lambda t$$

Задача 6.23. *Величины* Y_i *независимы*

Задача 6.24. *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 $\min\{p, np^2\}$

7 За время моего дежурства происшествий не было!

Задача 7.1. *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?*

Задача 7.2. *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?*

Задача 7.3. *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?*

Задача 7.4. *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* $0,8$*?*

Задача 7.5. *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.*

1. *Какова вероятность того, что тетя Зина обслужит клиента быстрее?*
2. *Как распределено время обслуживания того клиента, который освободится быстрее?*
3. *Каково условное среднее время обслуживания клиента тетей Зиной, если известно, что она обслужила клиента быстрее тети Маши?*

Задача 7.6. *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).*

1. *Какова вероятность того, что в очереди ровно* n *студентов?*
2. *Какова средняя длина очереди?*

Подсказка: если сейчас в очереди n *человек, то через малый промежуток времени* $\Delta\ldots$

Задача 7.7. *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?*

Задача 7.8. *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?*

Задача 7.9. *Величина* X *распределена экспоненциально с параметром* λ*, а константа* $a > 0$. *Как распределена величина* $Y = aX$*?*

Задача 7.10. *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.*

Задача 7.11. *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.*

Задача 7.12. *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.*

Задача 7.13. *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.*

1. *Найдите функцию плотности* Y_1
2. *Верно ли, что процесс количества звонков, которые услышит Василиса, имеет независимые приращения?*

Задача 7.14. *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* $\mathbb{E}(X)$, $\mathrm{Var}(X)$.

Задача 7.15. *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.*

Задача 7.16. *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?*

Задача 7.17. *На Краю Вселенной давным-давно работает парикмахерская. В ней счётное количество занумерованных по порядку парикмахеров. Каждый парикмахер независимо от других обслуживает клиента за экспоненциально распределенное время с параметром* λ. *Клиенты приходят в парикмахерскую пуассоновским потоком с интенсивностью* μ. *Клиент всегда выбирает свободного парикмахера с наименьшим номером.*

Какую долю времени будет в среднем занят парикмахер номер n*?*

8 Совместная плотность и вероятность

Минитеория:

1. $f(x|y) = \frac{f(x,y)}{f(y)}$
2. $f(x) = \int_{-\infty}^{+\infty} f(x,y)\,dy$
3. $\mathbb{E}(g(X,Y)) = \int\int g(x,y) \cdot f(x,y)\,dx\,dy$
4. Если $f(x,y) = f(x) \cdot f(y)$, то величины X и Y независимы

Задачи:

Задача 8.1. *Эксперимент может окончиться одним из шести исходов:*

	$X = -1$	$X = 0$	$X = 2$
$Y = 0$	*0.1*	*0.2*	*0.3*
$Y = 4$	*0.2*	*0.1*	*0.1*

Найдите: $\mathrm{Var}(X)$, $\mathrm{Var}(Y)$, $\mathrm{Cov}(X,Y)$, $\mathrm{Corr}(X,Y)$, $\mathrm{Cov}(X,Y|X \geqslant 0)$

Задача 8.2. *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.*

1. *Найдите функции плотности $f(y|x)$, $f(x)$, $f(x,y)$, $f(x|y)$, $f(y)$*
2. *Найдите $\mathbb{E}(X)$, $\mathbb{E}(Y)$, $\mathbb{E}(X^2)$, $\mathbb{E}(Y^2)$*
3. *Найдите $\mathrm{Var}(X)$, $\mathrm{Var}(Y)$, $\mathrm{Cov}(X,Y)$, $\mathrm{Corr}(X,Y)$*
4. *Найдите $\mathbb{E}(X|Y)$, $\mathbb{E}(Y|X)$, $\mathbb{E}(X^2|Y)$, $\mathbb{E}(Y^2|X)$*
5. *Найдите $\mathrm{Var}(Y|X)$, $\mathrm{Var}(X|Y)$*
6. *Найдите $\mathbb{P}(Y > 0.2|X = 0.5)$, $\mathbb{P}(Y > 0.2|X < 0.5)$*

Задача 8.3. *Величины Y_1 и Y_2 независимы и экспоненциально распределены с параметром $\lambda = 5$, а $S = Y_1 + Y_2$.*

1. *Выпишите и нарисуйте функцию плотности $f(y_1)$*
2. *Найдите совместную функцию плотности $f(y_1, y_2)$*
3. *Найдите плотности $f(y_1, s)$, $f(y_1|s)$*
4. *Прокомментируйте простыми словами вид функции $f(y_1|s)$*

Задача 8.4. *Величины X и Y имеют совместную функцию плотности*

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

1. *Найдите $\mathbb{P}(X > 0.5)$, $\mathbb{P}(X + Y > 0.5)$, $\mathbb{P}(X = Y + 0.2)$, $\mathbb{P}(X \leqslant Y)$, $\mathbb{P}(Y > 0.5|X > 0.5)$*
2. *Найдите $\mathbb{E}(X)$, $\mathbb{E}(XY)$, $\mathrm{Cov}(X,Y)$, $\mathrm{Corr}(X,Y)$*
3. *Найдите $\mathbb{E}(Y|X)$, $\mathbb{E}(Y^2|X)$, $\mathrm{Var}(Y|X)$*
4. *Зависимы ли величины X и Y?*
5. *Найдите совместную функцию распределения $F(x,y)$*

Задача 8.5. *Величины X и Y имеют совместную функцию распределения*

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

1. *Найдите $\mathbb{P}(X > 0.5)$, $\mathbb{P}(X + Y > 0.5)$, $\mathbb{P}(X = Y + 0.2)$, $\mathbb{P}(X \leqslant Y)$, $\mathbb{P}(Y > 0.5|X > 0.5)$*
2. *Найдите частные функции распределения $F(x)$, $F(y)$*
3. *Найдите совместную функцию плотности $f(x,y)$*
4. *Найдите $\mathbb{E}(X)$, $\mathbb{E}(XY)$, $\mathrm{Cov}(X,Y)$, $\mathrm{Corr}(X,Y)$*
5. *Зависимы ли величины X и Y?*

Задача 8.6. *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)*

Задача 8.7. *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)

Задача 8.8. *Приведите пример пары* X *и* Y *у которой нет совместной функции плотности, однако* X *и* Y *равномерно распределены на отрезке* $[0;1]$

Задача 8.9. *Известно, что* $Y=2X-3$, *а* $Z=6-3X$. *Найдите* $\mathrm{Corr}(X,Y)$, $\mathrm{Corr}(X,Z)$

Задача 8.10. *Пусть* X *и* Y *независимы.*

1. *Найдите* $\mathrm{Cov}(X,Y)$, $\mathrm{Cov}(X^3,Y^2-5Y)$, $\mathrm{Corr}(X,Y)$
2. *Выразите* $\mathrm{Var}(X+Y)$ *и* $\mathrm{Var}(X-Y)$ *через* $\mathrm{Var}(X)$ *и* $\mathrm{Var}(Y)$

Задача 8.11. *Кубик подбрасывается два раза,* X *— сумма очков,* Y *— разность очков, число при первом броске минус число при втором. Найдите* $\mathbb{E}(XY)$, $\mathrm{Cov}(X,Y)$, $\mathrm{Corr}(X,Y)$

Задача 8.12. *Пусть* X *равновероятно принимает значения -1, 0, +1, а* $Y=X^2$

1. *Найдите* $\mathrm{Cov}(X,Y)$*;*
2. *Верно ли, что* X *и* Y *независимы?*

Задача 8.13. *Пусть* X *равномерно на* $[-1;1]$, $Y=X^2$

1. *Найдите* $\mathrm{Cov}(X,Y)$*;*
2. *Верно ли, что* X *и* Y *независимы?*

Задача 8.14. *Паук сидит в начале координат. Равновероятно он может сместиться на единицу вверх, вниз, влево или вправо (по диагонали паук не ползает). Пусть* X *и* Y *— это абсцисса и ордината паука после первого шага.*

1. *Найдите* $\mathrm{Cov}(X,Y)$*?*
2. *Верно ли, что* X *и* Y *независимы?*

Задача 8.15. *Кубик подбрасывается* n *раз. Пусть* X_1 *— число выпадений 1, а* X_6 *— число выпадений 6.*

Найдите $\mathrm{Corr}(X_1,X_6)$.

Подсказка: $\mathrm{Cov}(X_1,X_1+\ldots+X_6)=\ldots$

Задача 8.16. *Вероятность дождя в субботу 0.5, вероятность дождя в воскресенье 0.3. Корреляция между наличием дождя в субботу и наличием дождя в воскресенье равна* r.

Какова вероятность того, что в выходные вообще не будет дождя?

Задача 8.17. *Пусть* $\mathbb{P}(A|B)>\mathbb{P}(A)$. *Можно ли определить знак* $\mathrm{Cov}(1_A,1_B)$*?*

Задача 8.18. *Вася наблюдает значение* X, *но не наблюдает значение* Y*; при этом он знает, что* $\mathrm{Var}(X)=3$, $\mathrm{Var}(Y)=8$, $\mathrm{Cov}(X,Y)=-3$, $\mathbb{E}(Y)=3$, $\mathbb{E}(X)=2$. *Задача Васи — спрогнозировать* Y *с помощью линейной функции от* X, *т.е. построить* $\hat{Y}=aX+b$.

Васю штрафуют за неправильный прогноз на сумму $(Y-\hat{Y})^2$. *Найдите* a *и* b

Задача 8.19. *Случайные величины* X *и* Y *зависимы, случайные величины* Y *и* Z *зависимы. Верно ли, что случайные величины* X *и* Z *зависимы?*

Задача 8.20. *Пусть* $\mathrm{Cov}(X,Y)>0$, $\mathrm{Cov}(Y,Z)>0$. *Верно ли, что* $\mathrm{Cov}(X,Z)>0$*?* $\mathrm{Cov}(X,Z)\geqslant 0$*?*

Задача 8.21. *Кольцо задается системой неравенств:* $x^2+y^2\geqslant 1$ *и* $x^2+y^2\leqslant 4$. *Случайным образом, равномерно на этом кольце, выбирается точка,* X *и* Y *— её координаты.*

Чему равна корреляция X *и* Y*? Зависимы ли* X *и* Y*?*

9 Всё нормально!

Задача 9.1. *Величины* $X_1, \ldots, X_n$ *распределены нормально* $\mathcal{N}(4, 100)$ *и независимы.*

1. *Найдите вероятности* $\mathbb{P}(X_1 > 4)$, $\mathbb{P}(X_1 \in [2; 20])$, $\mathbb{P}(X_1 < -5)$
2. *Найдите вероятности* $\mathbb{P}(\bar{X}_{100} > 10)$, $\mathbb{P}(\bar{X}_{36} \in [0; 5])$
3. *Найдите такое число* a*, что* $\mathbb{P}(X_1 > a) = 0.3$
4. *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`.

Задача 9.2. *Величина* W *имеет функцию плотности* $f(w) = c \cdot \exp(5w - 2w^2)$*. Найдите* $\mathbb{E}(W)$, $\mathrm{Var}(W)$, c.

Задача 9.3. *Величина* X *нормальна* $\mathcal{N}(\mu, \sigma^2)$.

1. *Выпишите функцию плотности случайной величины* X, $f(x)$
2. *Найдите точку максимума и точки перегиба функции* f

Задача 9.4. *Величина* X *распределена нормально* $\mathcal{N}(0, \sigma^2)$.

1. *Найдите* $\mathbb{E}(|X|)$
2. *Найдите функцию плотности* $|X|$

Задача 9.5. *Известно, что* $\ln Y \sim \mathcal{N}(\mu; \sigma^2)$*. Найдите* $\mathbb{E}(Y)$, $\mathrm{Var}(Y)$*, медиану и моду величины* Y.

Задача 9.6. *Величина* X *имеет стандартное нормальное распределение.*

1. *Найдите* $\mathbb{E}(Y)$ *для* $Y = \max\{X, 0\}$
2. *Найдите* $\mathbb{E}(X|X < 0)$ *и* $\mathrm{Var}(X|X < 0)$

Задача 9.7. *Величина* X *имеет нормальное распределение* $\mathcal{N}(\mu; \sigma^2)$*, а функция* $g(t)$ *дифференцируема и не слишком быстро растёт при* $t \to \infty$.

1. *Докажите, что* $\mathbb{E}((X - \mu)f(X)) = \sigma^2 \mathbb{E}(f'(X))$.
2. *Найдите* $\mathbb{E}(X^4)$ *и* $\mathbb{E}(X^6)$ *для* $X \sim \mathcal{N}(0; 1)$.

10 Долой неравенство Чебышёва и Маркова

Задача 10.1. *С помощью неравенства Чебышева оцените вероятности*

1. $\mathbb{P}(-2\sigma < X - \mu < 2\sigma)$*, если* $\mathbb{E}(X) = \mu$, $\mathrm{Var}(X) = \sigma^2$
2. $\mathbb{P}(8 < Y < 12)$*, если* $\mathbb{E}(Y) = 10$, $\mathrm{Var}(Y) = 400/12$
3. $\mathbb{P}(-2 < Z - \mathbb{E}(Z) < 2)$*, если* $\mathbb{E}(Z) = 1$, $\mathrm{Var}(Z) = 1$
4. *Найдите точные значения, если дополнительно известно, что* $X \sim \mathcal{N}(\mu; \sigma^2)$, $Y \sim U[0; 20]$ *и* $Z \sim Exp(1)$.

Задача 10.2. *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* $\mathbb{P}(X \in [90; 100]) \geqslant 0.95$*?*

Задача 10.3. *Известно, что* X *— неотрицательная случайная величина с* $\mathbb{E}(X) = 10$*. В каких пределах может лежать вероятность* $\mathbb{P}(X < 20)$*?*

11 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:

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

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

Задача 11.1. *Величины* X_1, …, X_n *независимы и равномерно распределены на отрезке* $[0;1]$. *Найдите* $\operatorname{plim} \bar{X}_n$, $\operatorname{plim} 1/(1+\bar{X}_n)$, $\operatorname{plim} \sum_{i=1}^n \ln X_i/n$, $\operatorname{plim} \sqrt[n]{X_1 \cdot \ldots \cdot X_n}$, $\operatorname{plim}(X_1 \cdot \ldots \cdot X_n)$, $\operatorname{plim} \max\{X_1, \ldots, X_n\}$, $\operatorname{plim} \min\{X_1, \ldots, X_n\}$, $\operatorname{plim} \frac{\sum (X_i - \bar{X})^2}{n}$, $\operatorname{plim} X_1/\bar{X}$.

Задача 11.2. *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* $\mathbb{P}(S_n < a) = 0.65$.

Задача 11.3. *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?*

Задача 11.4. *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.*

Задача 11.5. *Вероятность выпадения монетки «орлом» равна 0.63.*

1. *Какова вероятность, что в 100 испытаниях выборочная доля выпадения орлов будет отличаться от истинной вероятности менее, чем на 0.07?*

2. *Каким должно быть минимальное количество испытаний, чтобы вероятность отличия выборочной доли и истинной вероятности менее чем на 0.02 была больше 0.95?*

Задача 11.6. *Величины* $X_1, \ldots, X_n$ *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.

Задача 11.7. *Величина* X *имеет биномиальное распределение* $Bin(n,p)$ *и* n *велико. Какое распределение примерно имеют величины* $\ln(X/n)$*?* $X/(n-X)$*?*

Задача 11.8. *(*) 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* $\lim_{n\to\infty}\sqrt{n}p_n$.

Задача 11.9. *(*) Величины* $X_1, \ldots, X_n$ *независимы и равномерны на отрезке* $[0;1]$. *Вероятность того, что сумма любых двух соседних иксов меньше единицы, обозначим* p_n. *Найдите предел* $\lim_{n\to\infty} p_n^{1/n}$.

12 Условное математическое ожидание

Задача 12.1. *Пусть совместное распределение* X *и* Y *задано таблицей:*

	$X=-1$	$X=1$
$Y=-1$	*1/8*	*4/8*
$Y=2$	*2/8*	*1/8*

1. *Найдите* $\mathbb{E}(Y|X)$, $\operatorname{Var}(Y|X)$
2. *Убедитесь, что* $\mathbb{E}(\mathbb{E}(Y|X)) = \mathbb{E}(Y)$

Задача 12.2. *Монетка выпадает орлом с вероятностью* p. *Эксперимент состоит из двух этапов. На первом этапе монетку подкидывают 100 раз и записывают число орлов,* Z. *На втором этапе монетку подбрасывают до тех пор пока не выпадет столько орлов, сколько выпало на первом этапе. Обозначим число подбрасываний монетки на втором этапе буквой* X. *Найдите* $\mathbb{E}(X|Z)$, $\operatorname{Var}(X|Z)$, $\mathbb{E}(X)$, $\operatorname{Var}(X)$

Задача 12.3. *Автобусы приходят на остановку через случайные промежутки времени (Пуассоновский поток с параметром* λ*). В первый день Вася приходит на остановку и замеряет время до первого автобуса. Пусть это время* X. *На следующий день Вася приходит на остановку и считает, сколько автобусов придет в течении времени* X. *Он получает количество автобусов* N.

1. *Найдите* $\mathbb{E}(N)$ *и* $\operatorname{Var}(N)$
2. *Найдите* $\mathbb{P}(N>0)$

Задача 12.4. *Известно, что* $U \sim U[0;1]$, *а величина* X *имеет распределение Рэлея (Rayleigh density):*

$$f_X(x) = \begin{cases} \frac{x}{\sigma^2}e^{-x^2/2\sigma^2}, & \text{если } x \geqslant 0 \\ 0, & \text{иначе} \end{cases} .$$

1. *Найдите* $\mathbb{E}(X|Y)$
2. *Найдите* $\hat{X} = aY + b$ *так, чтобы* $E[(X-\hat{X})^2]$ *была минимальной.*

Задача 12.5. *Маша собрала* n *грибов в лесу наугад. В лесу есть рыжики, мухоморы и лисички. Рыжики попадаются с вероятностью* $p_R > 0$, *лисички — с вероятностью* $p_L > 0$, *мухоморы — с вероятностью* $p_M > 0$, $p_R + p_M + p_L = 1$. *Пусть* R *— количество собранных рыжиков,* L *— лисичек, а* M *— мухоморов. Найдите:*

1. $\mathbb{E}(R+L|M)$, $\mathbb{E}(M|R+L)$
2. $\mathbb{E}(R|L)$
3. $\mathrm{Var}(R|L)$
4. $\mathbb{E}(R+L|L+M)$
5. $\mathbb{E}(R|R-L)$ *(? похоже не решается в явном виде)*
6. $\mathrm{Var}(R|L)$
7. $\mathbb{P}(\mathbb{E}(R|L)=0)$
8. $\mathbb{P}(R=0|L)$
9. $\mathbb{E}\left(\left(\frac{p_M}{p_R+p_M}\right)^{100-L}\right)$

Задача 12.6. *Пусть* X *и* Y *— независимые пуассоновские случайные величины с параметрами* λ_1 *и* λ_2*. Найдите* $\mathbb{E}(X|X+Y)$

Задача 12.7. *Вася, Петя и Коля играют в карточного «дурака» втроём. Вася проигрывает с вероятностью* p_1*, Петя — с вероятностью* p_2*, Коля — с вероятностью* p_3*. Естественно,* $p_1+p_2+p_3=1$*. Всего они сыграли* n *партий. Обозначим количества проигранных ими партий* X_1*,* X_2 *и* X_3*, соответственно. Найдите* $\mathbb{E}(X_1|X_1+X_2)$*. Может получится и* $\mathrm{Var}(X_1|X_1+X_2)$*?*

Задача 12.8. *Пусть* X_1*, …,* X_{100} *независимы и равномерны на* $[0;1]$*. Пусть* $L=\max\{X_1,X_2,\ldots,X_{80}\}$ *а* $R=\max\{X_{81},X_{82},\ldots,X_{100}\}$ *и* $M=\max\{X_1,\ldots,X_{100}\}$

Найдите

1. $\mathbb{P}(L>R|L)$ *и* $\mathbb{P}(L>R|R)$ *и* $\mathbb{P}(L>R|M)$, $\mathbb{P}(L>R|L,M)$
2. $\mathbb{E}(X_1|L)$, $\mathbb{E}(X_1|\min\{X_1,\ldots,X_{100}\})$
3. $\mathbb{E}(\min\{X_1,\ldots,X_{100}\}|\max\{X_1,\ldots,X_{100}\})$
4. $\mathbb{E}(\min\{X_1,\ldots,X_{100}\}|X_1)$
5. *Нарисуйте условную функцию распределения* $\mathbb{P}(X_1\leqslant t|L)$

Задача 12.9. *Величины* X_1*,* X_2*, … независимы и одинаково распределены* $\mathcal{N}(2;9)$*. Мы складываем случайное количество* N *слагаемых. Величина* N *независима от* X_i *и распределена по Пуассону с параметром* $\lambda=10$*. Обозначим сумму буквой* $S=\sum_{i=1}^{N}X_i$*. Найдите* $\mathbb{E}(S)$*,* $\mathrm{Var}(S)$ *и* $Cov(S,N)$

Задача 12.10. *Неправильный кубик выпадает с вероятностью* 0,5 *шестеркой вверх. Остальные пять граней выпадают равновероятно. Случайная величина* X *– остаток от деления номера грани на два,* Y *– остаток от деления номера грани на три. Найдите*

1. *Закон распределения* $\mathbb{E}(X\mid Y)$*,* $\mathbb{E}(Y\mid X)$
2. *Выразите* $\mathbb{E}(Y\mid X)$ *через* X*, а* $\mathbb{E}(X\mid Y)$ *через* Y
3. *Найдите* $Cov(\mathbb{E}(Y\mid X),\mathbb{E}(X\mid Y))$*,* $Cov(\mathbb{E}(Y\mid X),X)$*,* $Cov(Y,X)$

Задача 12.11. *Цена литра молока,* X*, распределена равномерно на отрезке* $[1;2]$*. Количество молока, которое дает корова Мурка,* Y*, распределено экспоненциально с* $\lambda=1$*. Надои не зависят от цены. Величина* Z *— выручка кота Матроскина от продажи всего объема молока.*

Найдите

1. *$E(Z|X)$, $\mathrm{Var}(Z|X)$, корреляцию Z и X*

2. *Закон распределения $E(Z|X)$*

3. *Функцию плотности величины $\mathrm{Var}(Z|X)$*

Задача 12.12. *Кубик подбрасывают бесконечное количество раз. Величина X — номер подбрасывания, когда впервые выпала единица, а Y — номер подбрасывания, когда впервые выпала шестерка. Найдите $\mathbb{E}(Y|X)$.*

Задача 12.13. *The random variables X_1, X_2, … are independent uniformly distributed on $[0;1]$. I am summing them until the first X_i greater than 0.5 is added. After this term I stop. Let's denote by S the total sum and by N — the number of terms added. Find $\mathbb{E}(S|N)$, $\mathrm{Var}(S|N)$, $\mathbb{E}(S)$, $\mathrm{Var}(S)$*

Задача 12.14. *Известно, что $X = (Z_1 + Z_2)^2 + Z_3$ и $Y = (Z_1 + Z_2)^3 + Z_3$, величины Z_i независимы, $Z_1 \sim U[0;2]$, $Z_2 \sim N(1,4)$, $Z_3 \sim N(-2,9)$. Найдите*

1. *$\mathbb{E}(X \mid Z_1)$, $\mathbb{E}(X \mid Z_2)$, $\mathbb{E}(X \mid Z_3)$*

2. *$\mathbb{E}(Y \mid Z_1)$, $\mathbb{E}(Y \mid Z_2)$, $\mathbb{E}(Y \mid Z_3)$*

Задача 12.15. *Вася случайно выбирает между 0 и 1 число X_1, затем случайно выбирает между 0 и X_1 число X_2, затем X_3 между 0 и X_2, и так до бесконечности.*

1. *Найдите $\mathbb{E}(X_n)$, $\mathrm{Var}(X_n)$;*

2. *Найдите функцию плотности распределения X_n;*

3. *Найдите $\mathbb{E}(X_2|X_1, X_3)$;*

4. *Найдите $\mathrm{plim}\, X_n$*

Задача 12.16. *Приведите пример:*

1. *X и Y зависимы, но $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$*

2. *X и Y зависимы, но $\mathbb{E}(Y|X) = \mathbb{E}(Y)$ (as)*

3. *$\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$, но $\mathbb{E}(Y|X) \neq \mathbb{E}(Y)$ (as)*

Задача 12.17. *Пусть X — равномерна на отрезке $[0;1]$. В шляпе лежат две свернутые бумажки. На одной бумажке написано X, на другой X^2. Вы тяните одну бумажку наугад. Пусть Z — число, написанное на вытянутой Вами бумажке, а W — число на другой бумажке. Увидев число Вы решаете, оставить себе эту бумажку, или отказаться от этой и забрать оставшуюся. Ваш выигрыш — число на оставшейся у Вас бумажке.*

1. *Найдите $\mathbb{E}(W|Z)$*

2. *Максимально подробно (кубическое уравнение там будет суровое, не решайте его) опишите стратегию максимизирующую Ваш выигрыш*

3. *Как изменится результат, если на одной бумажке написано значение X, а на второй — значение случайной величины имеющей такое же распределение, как и X^2, но независимой от X?*

13 Многомерное нормальное

Задача 13.1. *Пара X и Y имеет двумерное нормальное распределение*

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim \mathcal{N}\left[\begin{pmatrix} 1 \\ -2 \end{pmatrix}; \begin{pmatrix} 9 & -1 \\ -1 & 4 \end{pmatrix}\right]$$

1. *Найдите* $\mathbb{E}(X)$, $\operatorname{Var}(X)$, $\mathbb{E}(X+3Y-7)$, $\operatorname{Var}(X+3Y-7)$, $\operatorname{Cov}(X-Y, 2X+3Y)$, $\operatorname{Corr}(X-9, X+3Y)$
2. *Найдите* $\mathbb{P}(X>5)$, $\mathbb{P}(X+Y>5)$
3. *Найдите* $\mathbb{E}(X|Y)$, $\operatorname{Var}(X|Y)$, $\mathbb{P}(X>1|Y=1)$
4. *Найдите* $\mathbb{E}(Y-3|X)$, $\operatorname{Var}(2X+7Y+2|Y)$, $\mathbb{P}(Y+X>1|X=2)$

Задача 13.2. *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

Задача 13.3. *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»?

Задача 13.4. *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:*

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

1. *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.*
2. *Рассчитайте корреляцию между X_1 и X_2*

3. *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 $X' = (X'_1, X'_2)$.*

4. *Определите, каков ожидаемый рост и дисперсия роста человека, средний рост родителей которого составляет 72 дюйма?*

5. *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

Задача 13.5. *«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.

1. *Найдите* $\operatorname{Corr}(Y_t, Y_{t-1})$

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

3. *Проинтерпретируйте величину коэффициента при Y_{t-1}, становится ли Машка в среднем улыбчивей или грустней со временем?*

4. *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?*

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

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

Задача 13.6. *Пара случайных величин X и Y имеет совместное нормальное распределение:*

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

1. *Найдите корреляцию X и Y*

2. *Найдите собственные числа и собственные векторые ковариационной матрицы*

3. *Постройте линии уровня совместной функции плотности*

Задача 13.7. *Случайная величина X нормально распределена, $N(0;4)$. При фиксированном X случайная величина Y нормально распределена, $N(2X-1,9)$.*

1. *Найдите безусловный закон распределения величины Y*

2. *Найдите* $\mathrm{Cov}(X, Y)$

3. *Найдите* $\mathbb{P}(Y > 2)$

4. *Найдите* $\mathbb{E}(X|Y)$, $\mathbb{E}(Y|X)$, $\mathrm{Var}(X|Y)$, $\mathrm{Var}(Y|X)$

Задача 13.8. *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,* $N(X, 9)$*.*

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

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

    (a) *Найдите безусловный закон распределения вектора* Y_1*, …,* Y_n

    (b) *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* $cI + dS$*.*

    (c) *Найдите* $\mathbb{E}(X|Y_1, \ldots, Y_n)$

Задача 13.9. *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* $\mathbb{P}(X_1 > 0, X_2 > 0)$*.*

Задача 13.10. *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*.*

1. *Какое распределение имеет величина* Y*?*

2. *Верно ли, что пара* X *и* Y *имеет совместное нормальное распределение?*

3. *Чему равна корреляция* X *и* Y*?*

4. *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 Броуновское движение

Задача 14.1.

Задача 14.2. *Пусть* Z *— стандартная нормальная случайная величина. Определим случайный процесс* $X_t = \sqrt{t}Z$*.*

1. *Найдите* $\mathbb{E}(X_t)$

2. *Найдите* $\mathrm{Var}(X_t)$

3. *Верно ли, что у процесса* X_t *независимые приращения?*

4. *Нарисуйте три типичные траектории процесса X_t?*

5. *Будет ли X_t броуновским движением?*

Задача 14.3. *Пусть W_t — стандартное броуновское движение и $a > 0$. Являются ли следующие процессы броуновским движением?*

1. $X_t = -W_t$;

2. $X_t = W_{a+t} - W_a$;

3. $X_t = \frac{1}{a} W_{a^2 t}$

4. $X_t = W_t^3$

Задача 14.4.
Задача 14.5.

15 Статистика ноль

Задача 15.1. *Имеется пять чисел: x, 9, 5, 4, 7. При каком значении x медиана будет равна среднему?*
Задача 15.2. *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?*

А как могут измениться медиана и среднее, если истинный рост Маши равен 153?
Задача 15.3. *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?*
Задача 15.4. *Деканат утверждает, что если студента N перевести из группы A в группу B, то средний рейтинг каждой группы возрастет. Возможно ли это?*
Задача 15.5. *Есть три группы по 10 человек, две группы по 20 человек и одна группа по 40 человек. У каждой из групп свой преподаватель.*

1. *Каков средний размер группы, для которой читает лекции наугад выбранный профессор?*

2. *Каков средний размер группы, в которой учится наугад выбранный студент?*

3. *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?*

Задача 15.6. *Приведите примеры случайных величин, для которых:*

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

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

3. $\mathrm{Med}(X^k) = \mathrm{Med}(X)^k$ *для всех* k

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

Задача 15.7. *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.*

Задача 15.8. *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?

Задача 15.9. *Два лекарства испытывали на мужчинах и женщинах. Каждый человек принимал только одно лекарство. Общий процент людей, почувствовавших улучшение, больше среди принимавших лекарство А. Процент мужчин, почувствовавших улучшение, больше среди мужчин, принимавших лекарство В. Процент женщин, почувствовавших улучшение, больше среди женщин, принимавших лекарство В.*

1. *Возможно ли это?*
2. *Какое лекарству нужно порекомендовать больному, не зная его пола?*

Задача 15.10. *Из набора чисел* $\{2, 4, 10, 14\}$ *случайным образом равновероятно по очереди выбираются три числа с возможностью повторения.*

1. *Найдите закон расперделения (табличку с вероятностями) величины* X_1. *Найдите закон распределения величины* X_2.
2. *Найдите совместный закон распределения пары* X_1, X_2. *Найдите совместный закон распределения пары* X_1, X_3.
3. *Являются величины* X_1, X_2, X_3 *независимыми? Одинаково распределеннными?*
4. *Верно ли, что* $\mathbb{E}(X_1) = \mathbb{E}(X_2) = \mathbb{E}(X_3)$? *Верно ли, что* $\mathrm{Var}(X_1) = \mathrm{Var}(X_2) = \mathrm{Var}(X_3)$?
5. *Верно ли, что* $\mathrm{Cov}(X_1, X_2) = \mathrm{Cov}(X_1, X_3)$?
6. *Как изменятся ответы на предыдущие вопросы, если числа выбираются без возможности повторения?*

Задача 15.11. *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* $\mathrm{Var}(X_1) = \sigma^2$. *Обозначим среднее арифметическое выбранных* n *чисел с помощью* $\bar{X}_n$. *Чему равны* $\mathbb{E}(\bar{X}_n)$ *и* $\mathrm{Var}(\bar{X}_n)$, *если:*

1. *Мы выбираем* n *чисел из* N *с возвращениями*
2. *Мы выбираем* n *чисел из* N *без возвращений*
3. *Во что превращаются полученные формулы при* $n = 1$? *при* $n = N$? *при* $N \to \infty$?

Задача 15.12. *Исследовательница Мишель подбрасывает кубик* 100 *раз. Пусть* X_1 *— количество выпадений единицы, а* X_6 *— количество выпадений шестёрки.*

1. *Как распределена величина* X_1? *Величина* X_6? *Найдите* $\mathbb{E}(X_1)$, $\mathrm{Var}(X_1)$.
2. *Верно ли, что величины* X_1 *и* X_6 *независимы? Одинаково распередены?*
3. *Найдите* $\mathrm{Cov}(X_1, X_1 + X_2 + X_3 + X_4 + X_5 + X_6)$, $\mathrm{Cov}(X_1, X_6)$
4. *Найдите* $\mathrm{Corr}(X_1, X_6)$, *проинтерпретируйте эту величину*

Задача 15.13. *В множестве* A *всего два числ,* $A = \{24, 42\}$. *Случайным образом из множества* A *выбираются 3 числа с возможностью повторений. Явно найдите закон распределения выборочного среднего, выборочной медианы, выборочной моды, выборочного минимима и выборочного максимума.*

16 Случайная выборка

Задача 16.1. *Создайте случайную выборку объемом* $n = 1000$ *из равномерного на отрезке* $[0;1]$ *распределения.*

1. *Найдите выборочные характеристики: среднее, медиану, минимум и максимум, стандартную ошибку, 10%-ый и 95%-ый квантили.*
2. *Постройте гистограмму распределения, выборочную функцию распределения для первых 20 чисел из случайной выборки*
3. *Повторите данный опыт для нормального* $N(5,1)$ *распределения и для экспоненциального распределения с параметром* $\lambda = 1$
4. *Насколько сильно выборочные характеристики отличаются от истинных?*

Задача 16.2. *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?*

Задача 16.3. *Придумайте детерминистическую функцию, такую, которая бы превращала одну равномерную на* $[0;1]$ *случайную величину* X *в*

1. *случайную величину* Y*, принимающую значения* 1 *и* 0 *с вероятностями* 0.7 *и* 0.3 *соответственно*
2. *случайную величину* Z *с функцией плотности* $f(z) = 2z$ *на отрезке* $z \in [0;1]$
3. *пару независимых одинаково распределенных случайных величин* (Y_1, Y_2)*, принимающих значения* 1 *и* 0 *с вероятностями* 0.7 *и* 0.3 *соответственно*
4. *пару независимых равномерных на* $[0;1]$ *случайных величин*

Задача 16.4. *Постройте случайную выборку в* $n = 200$ *наблюдений из двумерного нормального распределения с параметрами:*

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

1. *Посчитайте выборочную ковариацию, выборочную корреляцию*
2. *Постройте диаграмму рассеяния, нанесите на диаграмму рассеяния линию* $y(x) = E(Y|X = x)$
3. *Насколько сильно выборочные характеристики отличаются от истинных?*

Задача 16.5. *Создайте* 500 *выборок объемом* $n = 20$ *каждая из равномерного на отрезке* $[0;1]$ *распределения и вычислите выборочное среднее для каждой из выборок.*

1. *Каково теоретическое математическое ожидание и дисперсия каждого из выборочных средних?*
2. *Постройте гистограмму выборочных средних*
3. *На фоне функции плотности стандартного нормального распределения изобразите в подходящем масштабе гистограмму стандартизированных выборочных средних*

17 Максимально правдоподобно — 1!

Задача 17.1. *Кот Матроскин каждый вечер ходит на рыбалку. Поймав одну «рыбку» кот Матроскин возвращается домой. 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. За ночь экосистема пруда успевает восстановиться от воздействия кота Матроскина.*

1. *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.*
2. *Постройте оценку* $\hat{p}_{KM}$ *методом максимального правдоподобия в общем виде. То есть, Кот Матроскин ходил на пруд n дней, поймал* $Y_к$ *карасей,* $Y_щ$ *щук и* $Y_б$ *бегемотов.*
3. *Зависимы ли величины* $Y_к$, $Y_щ$ *и* $Y_б$*? Как распределена величина* $Y_б$*? Найдите* $\mathbb{E}(Y_б)$, $\mathrm{Var}(Y_б)$.
4. *Найдите* $\mathbb{E}(\hat{p}_{KM})$, $\mathrm{Var}(\hat{p}_{KM})$
5. *Является ли оценка* $\hat{p}_{KM}$ *несмещенной, состоятельной?*
6. *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?*
7. *Какая из двух оценок является более эффективной? Почему?*

Задача 17.2. *«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.*

1. *Постройте ML и MM оценку для неизвестного параметра* n*, если оказалось, что* $s = 80$
2. *Постройте ML и MM оценку для неизвестного параметра* n *в общем случае*

Задача 17.3. *Вася и Петя независимо друг от друга прочитали всю Википедию. Вася всего нашёл 100 опечаток, Петя — 200 опечаток. При этом 80 опечаток оказались найдены и Петей, и Васей. С помощью ML и MM:*

1. *Оцените количество опечаток в Википедии*
2. *Оцените внимательность Васи, то есть вероятность, с которой Вася находит опечатки*

Задача 17.4. *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*

1. u_i *— независимы и* $N(0; b)$
2. u_i *— независимы и* $U[-b; b]$

Задача 17.5. *Задача о немецких танках* [2]

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*.*

1. *Постройте оценку количества танков методом моментов*
2. *Постройте оценку количества танков методом максимального правдоподобия*
3. *Постройте несмещённую оценку количества танков с наименьшей дисперсией*

18 Максимально правдоподобно — 2!

`Минитеория`

Метод моментов (MM, method of moments): найти θ из уравнения $\bar{X_n} = \mathbb{E}(X_i)|_{\theta=\hat{\theta}}$
Метод максимального правдоподобия (ML, maximum likelihood):
найти θ при котором вероятность получить имеющиеся наблюдения будет максимальной
Наблюдаемая информация Фишера: $\hat{I} = -\frac{\partial^2 l}{\partial^2\theta}(\hat{\theta})$, $\widehat{\mathrm{Var}}(\hat{\theta}_{ML}) = \hat{I}^{-1}$.

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 $\hat{\theta}_{ML} \sim N(\theta, \frac{1}{I(\theta)})$.

Задача 18.1. *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: $X_1 = 0$, $X_2 = 2$.

1. *Найдите оценки $\hat{\theta}_{ML}$ и $\hat{\theta}_{MM}$*
2. *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 $X_1 = 0$, $X_2 = 2$?*
3. *Постройте ML и MM оценки для произвольной выборки X_1, X_2, $\ldots X_n$*

Задача 18.2. *Пусть Y_1 и Y_2 независимы и распределены по Пуассону. Известно также, что $\mathbb{E}(Y_1) = e^a$ и $\mathbb{E}(Y_2) = e^{a+b}$.*

1. *Найдите ML оценки $\hat{a}$ и $\hat{b}$ для случая $y_1 = 7$ и $y_2 = 3$*

---

[2] Незадолго до высадки союзников в Нормандии, 6 июня 1944 года, в распоряжении союзников было всего два (!) немецких танка «Пантера V». По серийным номерам на шасси танков союзники оценили выпуск в феврале 1944 в 270 танков. Фактический выпуск «Пантер V» согласно немецким документам в феврале 1944 составил 276 танков,

2. *Найдите ML оценки $\hat{a}$ и $\hat{b}$ в общем виде*

Задача 18.3. *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:*

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

Задача 18.4. *Пусть $X_1, \ldots, X_n$ - независимы и экспоненциальны с параметром λ. Постройте MM и ML оценки параметра λ. Оцените дисперсию ML оценки.*

Задача 18.5. *Пусть $X_1, \ldots, X_n$ - независимы и $N(\mu;\sigma^2)$. Значение σ^2 известно. Постройте MM и ML оценки параметра μ.*

Задача 18.6. *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.

Задача 18.7. *Пусть $Y_1 \sim N(0;\frac{1}{1-\theta^2})$. Найдите ML оценку для θ. Оцените дисперсию ML оценки.*

Задача 18.8. *Пусть $X_1, X_2, \ldots, X_n$ 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.

Задача 18.9. *Пусть $X_1, X_2, \ldots, X_n$ независимы и равномерно распределены на отрезке $[0;\theta]$, $\theta > 1$*

1. *Постройте MM и ML оценки для неизвестного θ.*
2. *Как изменятся ответы на «а», если исследователь не знает значений самих X_i, а знает только количество X_i оказавшихся больше единицы?*

Задача 18.10. *В озере водятся караси, окуни, щуки и налимы. Вероятности их поймать занесены в табличку*

Рыба:	*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.

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

Задача 18.11. *Известно, что X_i — независимы и имеют закон распределения, заданный таблицей:*

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

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

Задача 18.12. *Пусть X_i независимы и имеют экспоненциальное распределение с параметром λ, т.е.* $p(t) = \lambda e^{-\lambda t}$.

1. *Найдите* $I(\lambda)$*, если наблюдаются* $X_1, \ldots, X_n$
2. *Пусть* $\lambda = 1/\theta$*, т.е.* $p(t) = \frac{1}{\theta}e^{-\frac{1}{\theta}t}$. *Найдите* $I(\theta)$*, если наблюдается* $X_1, \ldots, X_n$

Задача 18.13. *Пусть X_i — независимы и одинаково распределены. Пусть $I_{X_i}(\theta)$ - информация Фишера о θ, получаемая при наблюдении X_i.*

1. *Верно ли, что* $I_{X_1}(\theta) = I_{X_2}(\theta)$*?*
2. *Как найти* $I_{X_1,\ldots,X_n}(\theta)$ *зная* $I_{X_i}(\theta)$*?*

Задача 18.14. *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* $\sum \ln(\ln(X_i)) = -30$

1. *Найдите ML оценку параметра θ и ожидаемую и наблюдаемую информацию Фишера*
2. *Постройте 95% доверительный интервал для* θ

Задача 18.15. *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* $\sum 1/X_i = 12$

1. *Найдите ML оценку параметра θ и информацию Фишера* $I(\theta)$
2. *Пользуясь данными по выборке постройте оценку $\hat{I}$ для информации Фишера*
3. *Постройте 90% доверительный интервал для θ Hint: $\mathbb{E}(1/X_i) = 1/\theta^2$ (интеграл берется заменой $x = \theta^2 a^{-2}$)*

Задача 18.16. *Известно, что X_i независимы, $\mathbb{E}(X_i) = 5$, $\mathrm{Var}(X_i) = 4$ и n велико. Как примерно распределены следующие величины:*

1. $\bar{X}_n$,
2. $Y_n = (\bar{X}_n + 3)/(\bar{X}_n + 6)$,
3. $Z_n = \bar{X}_n^2$,
4. $W_n = 1/\bar{X}_n$

Задача 18.17. *Известно, что X_i независимы и равномерны на $[0;1]$.*

1. *Найдите* $\mathbb{E}(\ln(X_i))$, $\operatorname{Var}(\ln(X_i))$, $\mathbb{E}(X_i^2)$, $\operatorname{Var}(X_i^2)$
2. *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?*

Задача 18.18. *Величины X_i независимы и имеют функцию плотности $f(x) = a \cdot x^{a-1}$ на отрезке $[0;1]$.*

1. *Постройте оценку $\hat{a}$ методом моментов, укажите её примерный закон распределения*
2. *По 100 наблюдениям оказалось, что $\sum X_i = 25$. Посчитайте численное значение $\hat{a}$ и оцените дисперсию случайной величины $\hat{a}$.*

Задача 18.19. *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.*

1. *Найдите оценку p методом максимального правдоподобия*
2. *Найдите достаточную статистику T*
3. *Выразите $\widehat{\operatorname{Var}}(\hat{p})$ через достаточную статистику T*
4. *Найдите $\mathbb{P}(X_1 = 1 \mid T = t)$.*

Задача 18.20. *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:*

1. *Найдите оценку a методом максимального правдоподобия*
2. *Найдите достаточную статистику T*
3. *Выразите $\widehat{\operatorname{Var}}(\hat{a})$ через достаточную статистику T*
4. *Найдите $\mathbb{P}(X_1 < 10 \mid T = t)$.*

Задача 18.21. *Величины X_i равномерны на отрезке $[-a;3a]$ и независимы. Есть несколько наблюдений, $X_1 = 0.5$, $X_2 = 0.7$, $X_3 = -0.1$.*

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

5. *Найдите оптимальную теоретическую взвешивающую матрицу для обобщённого метода моментов*

6. *Постройте двухшаговую оценку обобщённого метода моментов, начав со взвешивающей матрицы* W

Задача 18.22. *Величины* X_i *имеют Пуассоновское распределение с параметром* λ *и независимы. Есть несколько наблюдений,* $X_1 = 5$, $X_2 = 7$, $X_3 = 1$.

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

2. *Постройте оценку метода моментов, используя* $\mathbb{E}(X_i)$

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

4. *Постройте оценку обобщёного метода моментов используя моменты* $\mathbb{E}(X_i)$, $\mathbb{E}((X_i - \bar{X})^2)$ *и взвешивающую матрицу*
$$W = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}$$

5. *Найдите оптимальную теоретическую взвешивающую матрицу для обобщённого метода моментов*

6. *Постройте двухшаговую оценку обобщённого метода моментов, начав со взвешивающей матрицы* W

19 Настоящий ценитель

Задача 19.1. *Случайные величины* X_1, X_2, $\ldots$ *независимы и одинаково распределены с неизвестными* $\mathbb{E}(X_i) = \mu$ *и* $\mathrm{Var}(X_i) = \sigma^2$. *Исследовательница Борислава хочет использовать оценку вида* $\hat{\mu} = c(X_1 + X_2 + \ldots + X_n)$ *для неизвестного параметра* μ.

1. *При каком c оценка Бориславы будет несмещённой? Возможно ли использовать такое c в практической задаче?*

2. *При каком c будет минимальной величина* $MSE = \mathbb{E}((\hat{\mu} - \mu)^2)$*? Возможно ли использовать такое c в практической задаче?*

3. *Святозар минимизирует по* $\hat{\mu}$ *штрафную функцию*
$$Q(\hat{\mu}) = \sum (y_i - \hat{\mu})^2 + \lambda \hat{\mu}^2.$$
При каком λ *оценка Святозара совпадёт с несмещённой оценкой Бориславы? С оценкой минимизирующей MSE?*

Задача 19.2. *Исследовательница Радомира размышляет о том, как оценить неизвестное математическое ожидание по выборке из независимых одинаково распределённых случайных величин. Она мучительно выбирает из нескольких оценок. Для каждой оценки определите, является ли она несмещённой, состоятельной и линейной по наблюдениям. Для линейных несмещённых оценок определите, являются ли они эффективными среди линейных несмещённых оценок.*

1. *Удалить из выборки наблюдение номер 13 и посчитать среднее арифметическое.*

2. *Удалить из выборки все нечётные наблюдения и посчитать среднее арифметическое.*

3. *Удалить из выборки все наблюдения после 13-го и посчитать среднее арифметическое.*
4. *Домножить наблюдение номер 13 на 13 и посчитать среднее арифметическое.*
5. *Прибавить число 13 к наблюдению номер 13 и посчитать среднее арифметическое.*
6. *Продублировать 13-ое наблюдение 13 раз и посчитать среднее арифметическое.*
7. *Продублировать каждое наблюдение 13 раз и посчитать среднее арифметическое.*
8. *Домножить первое наблюдения на 1, второе — на 2, третье — на 3, и так далее и посчитать среднее арифметическое.*
9. *Прибавить к первому наблюдению 1, ко второму — 2, к третьему — 3, и так далее и посчитать среднее арифметическое.*
10. *Продублировать первое наблюдения 1 раз, второе — 2 раза, третье — 3 раза, и так далее и посчитать среднее арифметическое.*

Задача 19.3. *Величины Y_i независимы и имеют функцию плотности*

$$f(y) = \begin{cases} 5y^4/\theta^5, & \text{если } y \in [0; \theta]; \\ 0, & \text{иначе.} \end{cases}$$

1. *Найдите ML оценку неизвестного параметра θ.*
2. *Устно, не производя вычислений, определите, является ли оценка $\hat{\theta} = \max\{Y_1, Y_2, \ldots, Y_n\}$ несмещённой.*
3. *Найдите функцию распределения Y_1, функцию распределения $\hat{\theta}$, функцию плотности $\hat{\theta}$.*
4. *Найдите $\mathbb{E}(\hat{\theta})$.*
5. *Если $\hat{\theta}$ смещённая, то скорректируйте оценку так, чтобы она стала несмещённой.*

Задача 19.4. *Величина Y имеет биномиальное распределение $Bin(n, p)$.*

1. *Является ли оценка $\hat{p} = Y/n$ для p несмещённой? Если является смещённой, то скорректируйте оценку так, чтобы она стала несмещённой.*
2. *Чему равна теоретическая дисперсия σ^2 величины Y?*
3. *Является ли оценка $\hat{\sigma}^2 = n\hat{p}(1-\hat{p})$ для σ^2 несмещённой? Если является смещённой, то скорректируйте оценку так, чтобы она стала несмещённой.*

Задача 19.5. *Величины X_i независимы и одинаково распределениы. Какая из приведенных оценок для $\mu = \mathbb{E}(X_i)$ является несмещенной? Обладает наименьшей дисперсией среди несмещённых оценок? Обладает наименьшей среднеквадратичной ошибкой MSE?*

1. $X_1 + 3X_2 - 2X_3$;
2. $(X_1 + X_2)/2$;
3. $(X_1 + X_2 + X_3)/3$;
4. $(X_1 + \ldots + X_{20})/21$;
5. $X_1 - 2X_2$.

Задача 19.6. *Величина X равномерна на $[0; a]$. Придумайте несмещённую оценку параметра a вида $\hat{a} = \alpha + \beta X$.*

Задача 19.7. *Величины X_i - независимы и одинаково распределены. При каком значении параметра β*

1. *оценка $2X_1 - 5X_2 + \beta X_3$ будет несмещённой для $\mathbb{E}(X_i)$?*
2. *оценка $\beta (X_1 + X_2 - 2X_3)^2$ будет несмещённойдля $\operatorname{Var}(X_i)$?*

Задача 19.8. *Величины X_1 и X_2 независимы и равномерны на $[0; a]$. При каком β оценка $Y = \beta \cdot \min\{X_1, X_2\}$ для параметра a будет несмещённой?*

Задача 19.9. *Величины X_i независимы и одинаково распределены. Какая из приведенных оценок для $\sigma^2 = \operatorname{Var}(X_i)$ является несмещённой?*

1. $X_1^2 - X_1X_2$;
2. $\frac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}{n}$;
3. $\frac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}{n-1}$;
4. $\frac{1}{2}(X_1 - X_2)^2$;
5. $X_1 - 2X_2$;
6. X_1X_2.

Задача 19.10. *Величина X равномерна на $[3a - 2; 3a + 7]$. При каких α и β оценка $Y = \alpha + \beta X$ неизвестного параметра a будет несмещённой?*

Задача 19.11. *Закон распределения величины X имеет вид*

1.

x_i	0	1	a
$P(X = x_i)$	1/4	1/4	2/4

;

2.

x_i	0	1	2
$P(X = x_i)$	1/4	a	$(3/4 - a)$

Постройте несмещённую оценку вида $Y = \alpha + \beta X$ для неизвестного параметра a.

Задача 19.12. *Время горения лампочки распределено экспоненциально с ожиданием равным θ . Вася включил одновременно 20 лампочек. Величина X обозначает время самого первого перегорания. Как с помощью X построить несмещенную оценку для θ ?*

Задача 19.13. *Величины X_i независимы и одинаково распределены, причем $\operatorname{Var}(X_i) = \sigma^2$, а $\mathbb{E}(X_i) = \frac{\theta}{\theta+1}$, где $\theta > 0$ — неизвестный параметр. С помощью $\bar{X}$ постройте состоятельную оценку для θ.*

Задача 19.14. *Величины $Y_i = \beta X_i + \epsilon_i$, константа β и случайные величины ϵ_i являются ненаблюдаемыми. Известно, что $\mathbb{E}(\epsilon_i) = 0$, $\operatorname{Var}(\epsilon_i) = \sigma^2$, ϵ_i являются независимыми. Константы X_i наблюдаемы, и известно, что $20 < X_i < 100$. У исследователя есть две оценки для β: $\hat{\beta}_1 = \frac{\bar{Y}}{\bar{X}}$ и $\hat{\beta}_2 = \frac{\sum X_i Y_i}{\sum X_i^2}$.*

1. *Проверьте несмещенность, состоятельность.*
2. *Определите, какая оценка является наиболее эффективной.*

Задача 19.15. *Исследователи Иван да Марья интересуются, какая доля населения берёт взятки. Они независимо друг от друга задают разным людям вопрос: «Берёте ли Вы взятки?»*

Иван использует следующий механизм: предлагает респонденту тайно подбросить правильную монетку, если монетка выпадает орлом, то предлагается ответить «да», если решкой — то предлагается ответить правду. Предположим, что все респонденты действую точно так, как предлагает Иван.

У Марьи кристально чистые голубые глаза, она невинна и чиста, и солгать ей просто невозможно.

Пусть p — доля берущих взятки, а $\hat{q}_I$ и $\hat{q}_M$ — доля ответивших «да» на вопросы Ивана да Марьи. Марья использует оценку $\hat{p}_M = \hat{q}_M$, а Иван — $\hat{p}_I = 2\hat{q}_I - 1$.

1. *Являются ли оценки $\hat{p}_M$, $\hat{p}_I$ несмещёнными? состоятельными?*
2. *Иван планирует опросить 100 человек. Сколько человек в зависимости от p нужно опросить Марье, чтобы $\operatorname{Var}(\hat{p}_I) = \operatorname{Var}(\hat{p}_M)$?*

Идея: Кирилл Фурманов

20 Классические интервальные оценки

Здесь интервальные оценки без гипотез :)

Задача 20.1. *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 $[0; kX]$.*

Задача 20.2. *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 $[0; kX]$.*

Задача 20.3. *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 $[\bar{X}_{10} - k; \bar{X}_{10} + k]$.*

Задача 20.4. *Вася наугад поймал 400 покемонов-девочек и 100 покемонов-мальчиков. Среди девочек 250 оказались ядовитыми, среди мальчиков — 60.*

1. *Найдите точечную оценку $\hat{p}$ для доли ядовитых покемонов среди мальчиков.*
2. *Найдите точечную оценку $\hat{\sigma}^2$ для дисперсии X_i, где X_i — индикатор ядовитости i-го покемона-мальчика.*
3. *Постройте 95%-ый доверительный интервал для доли ядовитых покемонов среди мальчиков.*
4. *Постройте 95%-ый доверительный интервал для доли ядовитых покемонов среди девочек.*
5. *Постройте 95%-ый доверительный интервал для разницы долей ядовитых покемонов среди девочек и мальчиков.*
6. *Нужно ли предположение о нормальности X_i и Y_i для решения предыдущих пунктов?*

Задача 20.5. *Вася наугад поймал 400 покемонов-девочек и 100 покемонов-мальчиков. Средний рост покемонов-девочек равен 0.9 метра, и сумма квадратов ростов равна $\sum Y_i^2 = 1000$. Для покемонов мальчиков средний рост равен 1 метру, а сумма квадратов ростов равна $\sum X_i^2 = 2000$.*

1. *Постройте точечную оценку для ожидания роста покемона-мальчика μ_X.*

2. *Постройте точечную оценку для дисперсии роста покемона-мальчика* σ_X^2.
3. *Постройте 95%-ый доверительный интервал для ожидаемого роста покемона-мальчика.*
4. *Постройте 95%-ый доверительный интервал для ожидаемого роста покемона-девочки.*
5. *Постройте 95%-ый доверительный интервал для разницы ожидаемых ростов покемона-мальчика и покемона-девочки.*
6. *Нужно ли предположение о нормальности* X_i *и* Y_i *для решения предыдущих пунктов?*

Задача 20.6. *В одном тропическом лесу длина удавов имеет нормальное распределение* $\mathcal{N}(\mu, \sigma^2)$. *По выборке из 10 удавов оказалось, что* $\sum Y_i = 20$ *метрам, а* $\sum Y_i^2 = 1000$.

1. *Постройте 95%-ый доверительный интервал для* μ.
2. *Постройте 95%-ый доверительный интервал для* σ^2.
3. *Важна ли предпосылка о нормальности при решении предыдущих пунктов?*

Задача 20.7. *В одном тропическом лесу водятся удавы и питоны. Длина удавов имеет нормальное распределение* $\mathcal{N}(\mu_X, \sigma_X^2)$. *По выборке из 10 удавов оказалось, что* $\sum X_i = 20$ *метрам, а* $\sum X_i^2 = 1000$. *Длина питонов имеет нормальное распределение* $\mathcal{N}(\mu_Y, \sigma_Y^2)$. *По выборке из 20 питонов оказалось, что* $\sum Y_i = 60$ *метрам, а* $\sum Y_i^2 = 4000$.

1. *Постройте точечные оценки для* μ_X, σ_X^2, μ_Y, σ_Y^2.
2. *Постройте 95%-ый доверительный интервал для* σ_X^2/σ_Y^2.
3. *Постройте 95%-ый доверительный интервал для разницы* $\mu_X - \mu_Y$ *предполагая равенство дисперсий* $\sigma_X^2 = \sigma_Y^2$.
4. *Важна ли предпосылка о нормальности при решении предыдущих пунктов?*

21 Проверка гипотез

Задача 21.1. *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.*

1. *Какова вероятность того, что Рао получит более короткий доверительный интервал при* $n = 10$?
2. *К чему стремится эта вероятность при* $n \to \infty$?

Задача 21.2. *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?*

Задача 21.3. *Величины* X_1 *и* X_2 *независимы и равномерны на отрезке* $[0; a]$. *Есть две гипотезы,* H_0: $a = 1$ *и* H_a: $a = 2$. *Мальвина отвергает* H_0 *в том случае, если* $X_1 + X_2 > 1.5$. *Найдите вероятность ошибок первого и второго рода.*

Задача 21.4. *Величины* X_1 *и* X_2 *независимы и нормальны* $\mathcal{N}(a; 1)$. *Есть две гипотезы,* H_0: $a = 1$ *и* H_a: $a = 2$. *Мальвина отвергает* H_0 *в том случае, если* $X_1 + X_2 > 1.5$. *Найдите вероятность ошибок первого и второго рода.*

Задача 21.5. *Величины X_1 и X_2 независимы и распределены экспоненциально с интенсивностью a. Есть две гипотезы, H_0: $a = 1$ и H_a: $a = 2$. Мальвина отвергает H_0 в том случае, если $\min\{X_1, X_2\} < 1$. Найдите вероятность ошибок первого и второго рода.*

Задача 21.6. *Величины X_1 и X_2 независимы и распределены по Пуассону с интенсивностью a. Есть две гипотезы, H_0: $a = 1$ и H_a: $a = 2$. Мальвина отвергает H_0 в том случае, если $X_1 + X_2 \geqslant 2$. Найдите вероятность ошибок первого и второго рода.*

Задача 21.7. *Бабушка Аксинья утверждает, что обладает сверхспособностями. У Аксиньи есть Внутренний Голос, который, стоит лишь Аксине глянуть на стакан, подсказывает, что налито в стакан, Бон-Аква или Аква Минерале.*

Исследователь Кирилл проводит с Аксиньей следующий опыт. Правила опыта Аксинье известны. Кирилл в тайне от Аксиньи наливает в 3 стакана Акву Минерале и 2 стакана Бон-Аквы. Затем предлагает их Аксинье в случайной порядке. Задача Аксиньи после осмотра всех стаканов определить, в каком порядке они предлагались.

Кирилл проверяет две гипотезы. Нулевую H_0: Аксинья не обладает сверхспособностями и её Внутренний Голос верно определяет содержимое каждого стакана с вероятностью $p = 0.5$ и альтернативную H_a: Аксинья обладает сверхспособностями и $p = 0.9$.

Критерий: если Аксинья ошиблась хотя бы один раз, то H_0 не отвергается; если не ошиблась ни разу, то H_0 отвергается.

1. *Предположим, Аксинья, не задумываясь, говорит ровно то, что подсказывает ей Внутренний Голос. Найдите вероятности ошибок первого и второго рода.*

2. *(*) Предположим, Аксинья осознаёт что, Внутренний Голос может ошибаться с вероятностью $1 - p$. И поэтому при очевидной ошибке Внутренного Голоса старается её исправить. Например, если Внутренний Голос говорит ААААБ, то, следуя ему, угадать все стаканы невозможно. Найдите вероятности ошибок первого и второго рода.*

Задача 21.8. *Пусть $X_1, \ldots, X_n$ — случайная выборка из нормального распределения с неизвестным математическим ожиданием μ и известной дисперсией $\sigma^2 = 4$.*

Объем выборки $n = 16$. Тестируются основная гипотеза $H_0 : \mu = 0$ против альтернативной гипотезы $H_a : \mu = 2$. С помощью леммы Неймана–Пирсона найдите наиболее мощный критерий, имеющий уровень значимости $\alpha = 0.05$.

Задача 21.9. *Величины X_1 и X_2 независимы и одинаково распределены на отрезке $[0;1]$. Есть две гипотезы, H_0: $X_i \sim U[0;1]$ и H_a: $f(x_i) = \begin{cases} 2x_i, & \text{если } x_i \in [0;1]; \\ 0, & \text{иначе} \end{cases}$. С помощью леммы Неймана–Пирсона найдите наиболее мощный критерий, имеющий уровень значимости $\alpha = 0.05$.*

Задача 21.10. *Величины X_1 и X_2 одинаково распределены на отрезке $[0;1]$. Есть две гипотезы, H_0: $X_i \sim U[0;1]$ и H_a: $f(x_1, x_2) = \begin{cases} x_1 + x_2, & \text{если } x_1, x_2 \in [0;1]; \\ 0, & \text{иначе} \end{cases}$. С помощью леммы Неймана–Пирсона найдите наиболее мощный критерий, имеющий уровень значимости $\alpha = 0.05$.*

Задача 21.11. *Пусть $X_1, \ldots, X_n$ — случайная выборка из нормального распределения с известным математическим ожиданием $\mu = 1$ и неизвестной дисперсией σ^2.*

Объем выборки $n = 16$. Тестируются основная гипотеза $H_0 : \sigma^2 = 4$ против альтернативной гипотезы $H_a : \sigma^2 = 9$. С помощью леммы Неймана–Пирсона найдите наиболее мощный критерий, имеющий уровень значимости $\alpha = 0.05$.

22 Многомерщина

Задача 22.1. *Известна ковариационная матрица и математическое ожидание вектора* y*:*

$$\mathbb{E}(y) = \begin{pmatrix} 2 \\ 1 \\ -3 \end{pmatrix}, \operatorname{Var}(y) = \begin{pmatrix} 2 & 1 & -1 \\ 1 & 4 & 2 \\ -1 & 2 & 9 \end{pmatrix}$$

Найдите $\mathbb{E}(Ay)$ *и* $\operatorname{Var}(Ay)$*, где* $A = \begin{pmatrix} 2 & 0 & 2 \\ -1 & 2 & 3 \end{pmatrix}$

Задача 22.2. *Пусть* $A = X(X^TX)^{-1}X^T$ *и матрица* X^TX *обратима.*

1. *Найдите* A^2*,* A^{2017}*,* A^T*.*
2. *В каком случае матрица* A *обратима?*

Задача 22.3. *Найдите матрицу* A *для каждой из ситуаций:*

1. *Матрица* A *проецирует* n*-мерные вектора на вектор* $\nVdash = (1, 1, 1, \ldots, 1)^T$*.*
2. *Матрица* A *проецирует* 3*-мерные вектора на вектор* $(1, 2, 9)^T$*.*
3. *Матрица* A *проецирует* 3*-мерные вектора на плоскость, порождённую векторами* $(1, 1, 1)^T$ *и* $(1, 2, 3)^T$*.*

Задача 22.4. *Про каждую из матриц проверьте, является ли она проектором. ... Для матриц-проекторов определите, на какие вектора они проецируют.*

Задача 22.5. *Известно, что ... Найдите закон распределения* $y^TB^{-1}y$*,* y^TP^y*....*

Задача 22.6. *Определим информацию Фишера как ковариационную матрицу вектора* $\frac{\partial \ell}{\partial \theta}$*,* $I = \operatorname{Var}\left(\frac{\partial \ell}{\partial \theta}\right)$.

1. *Найдите* $\mathbb{E}\left(\frac{\partial \ell}{\partial \theta}\right)$*.*
2. *Докажите, что* $I = -\mathbb{E}(\frac{\partial^2 \ell}{\partial \theta \partial \theta^T})$
3. *Докажите, что* $I = \mathbb{E}(\frac{\partial \ell}{\partial \theta} \frac{\partial \ell}{\partial \theta}^T)$
4. *Докажите, что* $\operatorname{Var}(\hat{\theta}) \operatorname{Var}(\frac{\partial \ell}{\partial \theta}) - I$ *является положительно определённой*

23 Самая главная компонента

Задача 23.1. *Найдите максимум функции* $2x^2+5y^2+3z^2+7w^2$ *при ограничении* $x^2+y^2+z^2+w^2 = 1$.

Задача 23.2. *Пусть* $A = \begin{pmatrix} 5 & 1 & 0 \\ 1 & 5 & 0 \\ 0 & 0 & 5 \end{pmatrix}$.

1. *Найдите собственные векторы и собственные числа матрицы* A
2. *Найдите максимум* v^TAv *при ограничении* $v^Tv = 1$*.*
3. *Найдите максимум* v^TAv *при ограничениях* $v^Tv = 1$ *и* $v \perp a$*, где* $a^T = (1, 1, 0)$*.*

Задача 23.3. *Вектора* z_1 *и* z_2 *имеют выборочную ковариационную матрицу*

$$M = \begin{pmatrix} 6 & 2 \\ 2 & 10 \end{pmatrix}.$$

1. *Как изменится коварицонная матрица, если центрировать вектора z_1 и z_2?*
2. *Как изменится коварицонная матрица, если центрировать и нормировать вектора z_1 и z_2?*

Задача 23.4. *Вектора-столбцы z_1 и z_2 содержат по пять наблюдений. Матрица X состоит из столбцов x_1 и x_2. Выборочная коварицонная матрица векторов z_1 и z_2 равна $\begin{pmatrix} 6 & 2 \\ 2 & 10 \end{pmatrix}$.*

1. *Найдите матрицу $X^T X$, если вектора x_i получены центрированием векторов z_i.*
2. *Найдите матрицу $X^T X$, если вектора x_i получены центрированием и нормирование векторов z_i.*

Задача 23.5. *Вениамин находит главные компоненты набора данных из трёх переменных. По каждой из переменных есть 100 наблюдений. Вениамин центрирует и не нормирует переменные, так как они изначально измеряются в одних и тех же единицах.*

Собственные числа выборочной ковариационной матрицы исходных переменных равны 5, 4 и 1.

1. *Найдите сумму выборочных дисперсий исходных переменных.*
2. *Найдите длины и выборочные дисперсии всех трёх главных компонент.*
3. *Какую долю от суммы выборочных дисперсий объясняют первые две главные компоненты?*

Задача 23.6. *Рассмотрим результаты пяти студентов за две контрольные работы:*

ФИО	*Контрольная 1*	*Контрольная 2*
Маша	*4*	*5*
Вася	*5*	*5*
Лена	*3*	*4*
Коля	*5*	*4*
Рита	*4*	*3*

1. *Найдите выборочную ковариационную матрицу.*
2. *Найдите собственные числа и собственные векторы единичной длины для ковариационной матрицы.*
3. *Выпишите первую и вторую главные компоненты.*
4. *Найдите сумму выборочных дисперсий исходных переменных.*
5. *Найдите длины и выборочные дисперсии всех главных компонент.*
6. *Какую долю от суммы выборочных дисперсий объясняет первая главная компоненты?*

24 И потребленье возрастает, а производство отстаёт!

Испить мудрость жадными глотками можно в источнике [1].

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

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

Задача 24.1. *Найдите коэффициент при* x^{17} *в многочлене* $(1+x^5+x^7)^{20}$.

Задача 24.2. *Найдите и запишите максимально просто производящие функции для последовательностей*

1. *1, 1, 1, 1, 0, 0, 0, ...*
2. *1, 1, 1, 1, 1, 1, 1, ...*
3. *1, 4, 6, 4, 1, 0, 0, ...*

Задача 24.3. *Найдите первые четыре коэффициента в многочленах* $(1+2x)^{-3}$ *и* $(1+x+x^2)^{-4}$.

Задача 24.4. *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* $X_1+X_2+X_3$.

Задача 24.5. *Производящая функция случайной величины* X *линейна и проходит через точки* $(1,1)$ *и* $(5,4)$. *Какие значения принимает величина* X *и с какими вероятностями?*

Задача 24.6. *На месте убийства преступники оставили улику — маленький-маленький кусочек графика производящей функции. Сможет ли Шерлок Холмс по этому кусочку восстановить все вероятности?*

Задача 24.7. *Леонардо Пизанский (Фибоначчи) приобрёл новорождённую пару кроликов. Каждая взрослая пара кроликов даёт в качестве приплода ежемесячно новую пару кроликов. Только что рождённая пара кроликов начинает приносить приплод со второго месяца жизни. Пусть* F_n *— количество пар кроликов в* n*-ый месяц.*

1. *Каким соотношением связаны* F_n, F_{n-1} *и* F_{n-2}*?*
2. *Пусть для удобства* $F_0=0$ *и* $F_1=1$. *Найдите производящую функцию для количества пар кроликов.*
3. *Используя хороший софт, поделите 1 на 99999999999999999999999899999999999999999999999. Здесь 23 девятки, восьмёрка и потом 24 девятки. Почему мы видим последовательность чисел Фибоннačи? Например, можно поделить в julia установив высокую точность:*

   ```
   setpreciștion(BigFloat, 10000)
   1 / 99999999999999999999999899999999999999999999999
   ```

4. *Разложите найденную функцию в сумму двух дробей и, вспомнив формулу для геометрической прогрессии, найдите явную формулу для чисел Фибоначчи.*

Задача 24.8. *Сева стартует в точке* 0 *на числовой прямой. За каждый час он перемещается вправо на единицу с вероятностью* 0.7 *или влево на единицу с вероятностью* 0.3. *Пусть* X_1 *— момент времени, когда Сева впервые достигнет точки* 1, *а* X_2 *— точки* 2.

1. *Как связаны производящие функции* $h_1(t)$ *и* $h_2(t)$ *величин* X_1 *и* X_2*?*
2. *Рассмотрев возможные положения Севы после первого шага, составьте квадратное уравнение на* $h_1(t)$.
3. *Найдите* $h_1(t)$.
4. *Найдите* $\mathbb{E}(X_1)$.
5. *Выпишите формулу для* $\mathbb{P}(X_1=2017)$.

Задача 24.9. *Ангела Меркель подбрасывает правильную монетку до выпадения трёх орлов подряд. Найдите производящую функцию для количества подбрасываний.*

25 Сэр Томас Байес

Байесовский подход (bayesian approach):

1. Сделать изначальное априорное предположение о распределении θ, $p(\theta)$;
2. Сформулировать модель для данных, $p(data|\theta)$;
3. Получить апостериорный закон распределения θ по формуле условной вероятности, $p(\theta|data) \propto p(\theta) \cdot p(data|\theta)$

Задача 25.1. *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. Мы не знаем вероятность выпадения орла α, поэтому считаем, что α равномерно распределена на отрезке $[0;1]$.

1. *Какова безусловная вероятность того, что α лежит в диапазоне $[0;0.5]$?*
2. *Какова условная вероятность того, что α лежит в диапазоне $[0;0.5]$, если монетка выпала 5 раз орлом и 7 раз решкой?*
3. *В байесовском подходе α — это константа или случайная величина?*
4. *В каком году умер Томас Байес?*

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

Задача 25.2. *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.*

1. *С помощью метода максимального правдоподобия определите, простая она или трудная.*
2. *С помощью байесовского подхода посчитайте вероятности того, что задача была простая, если на экзамене было 7 легких и 3 трудных задачи.*

26 Решения

1.1 $\mathbb{P}(X=1)=3/5$, $\mathbb{P}(X=2)=3/10$, $\mathbb{P}(X=3)=1/10$, $\mathbb{E}(X)=1.5$

1.2

1.3

1.4 N 3 4 5
2/8 3/8 3/8

1.5

1.6 5/11 и 6/11

1.7 N — количество подбрасываний, G — количество орлов, R — решек
E(N)=10/7, E(G)=1, E(R)=10/7-1=3/7

1.8

1.9 у Саши 3/4

1.10 стоп на 4-5-6 или стоп на 5-6

1.11 $\mathbb{P}(X = 4) = 1/8$, $\mathbb{E}(X) = 3$

1.12

2.1

2.2

2.3

2.4

2.5

2.6 1/2

2.7 $\mathbb{E}(X) = 0.25 \cdot 2 + 0.5 \cdot (\mathbb{E}(X) + 1) + 0.25 \cdot (\mathbb{E}(X) + 2)$
$\mathbb{E}(Y) = 0.25 \cdot 0 + 0.5 \cdot (\mathbb{E}(Y) + 1) + 0.25 \cdot (\mathbb{E}(Y) + 1)$

2.8 стратегия 1: говорить стоп, если на кону 200 рублей вне зависимости от числа набранных красных карточек

стратегия 2: если нет красных карточек или одна, то останавливаться при 200 рублях, а при двух карточках останавливаться на 100 рублях.

2.9

2.10

2.11

2.12

2.13

2.14

2.15

2.16

3.1 $\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{1/3}{1/6+1/3} = 2/3$

3.2 A – первый охотник попал в утку, B — в утку попала ровно одна пуля
$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{0.4 \cdot 0.3}{0.4 \cdot 0.3 + 0.6 \cdot 0.7} = 2/9$

3.3 $\mathbb{P}(A \mid B) = 1/6$

3.4 $\mathbb{P}(X \geqslant 2 \mid X \geqslant 1) = \frac{\mathbb{P}(X \geqslant 2)}{\mathbb{P}(X \geqslant 1)} = \approx 0.37$

3.5 $\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{7 \cdot 11 / C^2_{7+5+11}}{(7 \cdot 11 + 5 \cdot 11 + 7 \cdot 5)/C^2_{7+5+11}}$

3.6 B — событие, второй — черный, $\mathbb{P}(B) = 11/16$
$\mathbb{P}(A \mid B) = \mathbb{P}(A \cap B)/\mathbb{P}(B) = \frac{\frac{5}{16}\frac{11}{15}}{11/16}$

3.7 A — партия мяса заражена, B — партия мяса по тесту заражена
$\mathbb{P}(A \mid B) = \mathbb{P}(A \cap B)/\mathbb{P}(B) = \frac{0.04 \cdot 0.9}{0.04 \cdot 0.9 + 0.96 \cdot 0.1} \approx 0.27$

3.8 A — Петя из Б класса, B — Петя обожает географию
$\mathbb{P}(A \mid B) = \mathbb{P}(A \cap B)/\mathbb{P}(B) = \frac{0.4/3}{0.3/3 + 0.4/3 + 0.7/3} = 2/7$

3.9 1/3

3.10 да

3.11 нет, $(4/52) \cdot (3/51) = \mathbb{P}(A \cap B) \neq \mathbb{P}(A) \cdot \mathbb{P}(B) = (4/52)^2$

3.12

3.13 2/3, 1/2, 1/2, 14/27

4.1

4.2

4.3

4.4

4.5

5.1

5.2

5.3

5.4

5.5

5.6

5.7

5.8

5.9

5.10

5.11

5.12 при $c = 1/\mathbb{E}(X)$

5.13

5.14

5.15

5.16

6.1

6.2

6.3 1.75

6.4 $\mathbb{E}(X_0) = 100 \cdot \frac{99}{464} + \frac{100}{464} + \ldots + \frac{464}{464}$

6.5

6.6 $\mathbb{E}(X) = 1$, $\mathrm{Var}(X) = 1$

6.7 $\mathbb{E}(Y) = 20 - 20\left(\frac{19}{20}\right)^{10} \approx 8.025$

6.8 $\mathbb{E}(X) = 9$, $\mathrm{Var}(X) = 1.2$

6.9 $10 \cdot 0.9^7$

6.10 $\mathbb{E}(X) = \frac{30}{30} + \frac{30}{29} + \ldots + \frac{30}{1} \approx 119.85$
аппроксимация через логарифм $30 \ln 30 \approx 102$

6.11

6.12

6.13

6.14

6.15

6.16

6.17

6.18

6.19

6.20

6.21

6.22

6.23

6.24

7.1 $\mathbb{P}(X_8 = 13) = e^{-8.8} 8.8^{13}/13! \approx 0.046$
$\mathbb{P}(Y_{n+1} > 1) = e^{-1.1} \approx 0.33$

7.2

7.3

7.4

7.5 $\frac{a}{a+b}$, экспоненциально с параметром $a + b$, $\frac{1}{a+b}$

7.6 геометрическое распределение
$\mathbb{E}(N) = \frac{\lambda_{in}}{\lambda_{capacity} - \lambda_{in}}$

7.7 The probability of observing a taxi before a bus is given by $3/(3 + 2) = 3/5$ since the waiting times are independent and exponentially distributed. By the memoryless property both processes then restart and hence the probability of observing (at least) two taxis before the first bus is $(3/5)^2 = 9/25$. The probability of observing exactly two taxis before the first bus is $(3/5)^2 * (2/5) = 18/125$.

7.8

7.9

7.10

7.11

7.12

7.13

7.14

7.15

7.16

$$a(t + \Delta) = a(t)(1 - \Delta) + (1 - a(t))\Delta + o(\Delta)$$

сейчас четное = за секунду было четное и никто не зашел + за секунду было нечетное и один зашел $a'(t) = 1 - 2a(t)$, $a(0) = 1$ $a(t) = (1 + \exp(-2t))/2$

7.17

8.1

8.2

8.3

8.4

8.5

$$\mathbb{P}(X > 0.5) = 1 - \mathbb{P}(X \leqslant 0.5) = 1 - P(X \leqslant 0.5, Y \leqslant +\infty) = 1 - F(0.5, +\infty)$$

$F_X(x) = F(x, +\infty)$, $\mathbb{P}(X = Y + 0.2) = 0$ в силу непрерывности величин
X и Y независимы, так как $F(x, y) = F(x) \cdot F(y)$

8.6 можно через смешанную производную функции распределения $F(l, r)$ можно через разложение $f(l|r) \cdot f(r)$

$$f(l, r) = \begin{cases} 2, & \text{если } 0 < l < r < 1 \\ 0, & \text{иначе} \end{cases}$$

8.7

8.8 Например, $X \sim U[0; 1]$ и $Y = X$

8.9 $\mathrm{Corr}(X, Y) = 1$, $\mathrm{Corr}(X, Z) = -1$

8.10

8.11

8.12 $\mathrm{Cov}(X, Y) = 0$, зависимы

8.13 $\mathrm{Cov}(X, Y) = 0$, зависимы

8.14 $\mathrm{Cov}(X, Y) = 0$, зависимы

8.15 $-1/5$

8.16

8.17

8.18

8.19 нет, например, берем независимые X и Z и возьмём $Y = X + Z$.

8.20 нет

8.21 корреляция равна 0, зависимы, так как знание X несёт информацию об Y, например, при $X = 0$ можно утверждать, что $|Y| \in [1; 2]$.

9.1

9.2

9.3 $x_{max} = \mu$, $x = \mu \pm \sigma$

9.4

$$f_{|X|}(t) = \begin{cases} 2f_X(t), \ t > 0 \\ 0, \ \text{иначе} \end{cases}$$

9.5 медиана — $\exp(\mu)$ мода — $\exp(\mu{-}\sigma^2)$ $\mathbb{E}(Y) = \exp(\mu{+}\sigma^2/2)$ $\mathrm{Var}(Y) = \exp(2\mu{+}2\sigma^2){-}\exp(2\mu{+}\sigma^2)$

9.6

9.7 Применяем формулу интегрирования по частям.

10.1 b $4/20 \geqslant 1 - 100/12$
c $1 - e^{-3} \geqslant 0.75$

10.2 $\mathrm{Var}(X) \leqslant 5$

10.3 $\mathbb{P}(X < 20) \geqslant 0.5$

11.1

11.2

11.3

11.4 $S_{100} \sim \mathcal{N}(1000, 100)$, $\mathbb{P}(S_{100} > 1030) = \mathbb{P}(Z > 3) = 0.0013$

11.5

11.6

11.7

11.8 $4(\sqrt{2} - 1)/3$, 20 problems

11.9 $2/\pi$, 20 problems

12.1

12.2

12.3

12.4

12.5

12.6

12.7

12.8

12.9

12.10

12.11

12.12

12.13

12.14

12.15

12.16

12.17

13.1

13.2

13.3 $K = 170$, $M = 120$ (симметричный интервал) или $K = M = 168$ (площадь с одного края можно принять за 0)

Вариант: театр, два входа, два гардероба а) только пары, б) по одному

13.4

13.5 $\mathrm{Corr}(Y_1, Y_2) = 2/5$, $E(Y_t|Y_{t-1}) = 0.4Y_{t-1}$ В среднем Машка не становиться ни грустней, ни улыбчивей Представить $\mathbb{E}(Y_t|Y_{t-1}, Y_{t-2})$ в виде $\mathbb{E}(Y_t|Y_{t-1}, Y_{t-2}) = a_1 Y_{t-1} + a_2 Y_{t-2} + Z$

13.6

13.7

13.8

13.9

13.10 $Y \sim N(0; 1)$, нет, они нормальны только по отдельности, но не в совокупности, $\mathrm{Corr}(X, Y) = 0$. Взять $Y = XZ$, где Z принимает значение 1 с вероятностью $p = 3/4$ и -1 с вероятностью $1 - p = 1/4$

14.1

14.2

14.3 Являются: $X_t = -W_t$, $X_t = W_{a+t} - W_a$, $X_t = \frac{1}{a} W_{a^2 t}$

14.4

14.5

15.1 Среднее равно $(x + 25)/5$. Если $x < 5$, получаем $x = 0$. Если $x \in (5; 7)$, получаем $x = 25/4$. Если $x > 7$, получаем $x = 10$.

15.2 Медиана не изменится, среднее упадёт на $10/25 = 0.4$. Для случая роста 153: среднее упадёт на 0.8, медиана упадёт произвольно на некое число из отрезка $[0; 2]$.

15.3

15.4

15.5

15.6 два независимых симметричных распределения; практически любая сумма несимметричных распределений, например, два независимых с $p(x) = 2 - 2x$ на $[0; 1]$; неотрицательные случайные величины; симметричные около нуля случайные величины

15.7 Исключим те варианты, когда все пять наблюдений оказались или синхронно выше, или синхронно ниже медианы, получаем, $p = 1 - 2 \cdot 0.5^5 = 1 - 0.5^4$.

15.8

15.9

15.10

15.11

15.12

15.13

16.1

16.2

16.3

16.4

16.5

17.1 Метод максимального правдоподобия:

$$C_n^{Y_к} C_{n-Y_к}^{Y_щ} a^{Y_к} (2a)^{Y_щ} (1-3a)^{Y_б} \to \max_a$$

Решая задачу максимизации Кота Матроскина получаем

$$\hat{a} = \frac{Y_к + Y_щ}{3n}$$

Замечаем, что $Y_к + Y_щ \sim Bin(n, 3a)$. Отсюда $\mathbb{E}(\hat{a}) = a$, $\operatorname{Var}(\hat{a}) = \frac{a(1-3a)}{n}$. Оценка несмещённая и состоятельная.

С точки зрения Пса Шарика, неизвестными являются две вероятности, a и b. Он решает задачу максимизации по двум переменным. В результате получается вполне себе интуитивная оценка $\hat{a}_{ПШ} = Y_к/n$.

17.2 Метод правдоподобия: $\max_n \mathbb{P}(S = 80)$. Замечаем, что $S \sim Bin\left(100, p = \frac{100}{n}\right)$. Отсюда, $\hat{n}_{ML} = 125$.

Метод моментов. Рассмотрим Y_1, Y_2, …, Y_n. Величина Y_i равна 1 если при втором отлове i-ый заяц оказался с бантом и 0 иначе.

Метод моментов: $\mathbb{E}(Y_i)|_{n=\hat{n}} = \bar{Y}$:

$$\frac{100}{\hat{n}_{MM}} = \bar{Y}$$

Отсюда

$$\hat{n}_{MM} = \frac{100}{\bar{Y}} = \frac{100^2}{S}$$

17.3

17.4

17.5

18.1 $\hat{\theta}_{ML} = 0.25$, $\hat{\theta}_{MM} = 0.2$ $\hat{\theta}_{MM} = \frac{2{,}4-\bar{X}}{7}$

18.2 $\hat{a} = \ln(Y_1)$, $\hat{b} = \ln(Y_2) - \ln(Y_1)$

18.3

18.4

18.5

18.6 $\hat{a}_{ml} = \sum X_i^2/2n$, $\hat{a}_{mm} = \bar{X}$.

18.7

18.8

18.9

18.10

18.11

18.12

18.13

18.14

18.15

18.16

18.17

18.18

18.19

18.20

18.21

18.22

19.1

1. $c = 1/n$, да
2. $c = 1/(n + \sigma^2/\mu^2)$, нет, так как μ и σ неизвестны
3. $\lambda = 0$ и $\lambda = \sigma^2/\mu^2$

19.2

1. несмещённая, состоятельная, линейная, неэффективная
2. несмещённая, состоятельная, линейная, неэффективная
3. несмещённая, несостоятельная, линейная, неэффективная
4. смещённая, состоятельная, линейная
5. смещённая, состоятельная, нелинейная
6. несмещённая, состоятельная, линейная, неэффективная
7. несмещённая, состоятельная, линейная, эффективная
8. смещённая, несостоятельная, линейная
9. смещённая, несостоятельная, нелинейная
10. несмещённая, состоятельная, линейная, неэффективная

19.3

1. $\hat{\theta}_{ML} = \max\{Y_1, Y_2, \ldots, Y_n\}$.
2. Все Y_i меньше θ, значит и $\hat{\theta}$ всегда меньше θ, значит смещённая.
3. $F_{\hat{\theta}}(t) = \mathbb{P}(\hat{\theta} \leqslant t) = \mathbb{P}(Y_1 \leqslant t, Y_2 \leqslant t, \ldots) = (\mathbb{P}(Y_1 \leqslant t))^n$, $\mathbb{P}(Y_1 \leqslant t) = t^5/\theta^5$, $f_{\hat{\theta}}(t) = dF_{\hat{\theta}}(t)/dt = \frac{5nt^{5n-1}}{\theta^{5n}}$.
4. $\mathbb{E}(\hat{\theta}) = \frac{5n}{5n+1}\theta$.
5. $\hat{\theta}_{unbiased} = \frac{5n+1}{5n}\hat{\theta}$.

19.4

1. $\hat{p}$ несмещённая
2. $\sigma^2 = np(1-p)$.
3. $\mathbb{E}(\hat{\sigma}^2) = (n-1)p(1-p)$, смещённая, $\hat{\sigma}^2_{unbiased} = \frac{n}{n-1}\hat{\sigma}^2$.

19.5

19.6

19.7

19.8

19.9

19.10

19.11

19.12 Закон распределения X также экспоненциальный, но с другим λ. Честно находим $\mathbb{E}(X) = \theta/20$, отсюда $\hat{\theta}_{unbiased} = 20X$.

19.13

19.14

19.15

20.1 $\frac{1}{1.8} - \frac{1}{2.2}$, $[0; 10X]$.

20.2

20.3

20.4

20.5

20.6 Важна при обоих доверительных интервалах. Без предпосылки о нормальности интервал для дисперсии по данным формулам нельзя построить даже при больших n. При больших n можно отказаться от предпосылки о нормальности при построении интервала для μ.

20.7

21.1 к $1/2$

21.2 равномерно, $\alpha = 0.05$; нет, он резко увеличивает ошибку второго рода

21.3 $\alpha = 1/8$, $\beta = 9/32$

21.4 $\alpha = \mathbb{P}(\mathcal{N}(0;1) > -0.35) \approx 0.64$, $\beta = \mathbb{P}(\mathcal{N}(0;1) \leqslant -1.76) \approx 0.04$.

21.5

21.6

21.7

21.8 Критерий Неймана-Пирсона сводится к сравнению $\bar{X}$ с порогом. При верной H_0 величина $\bar{X}$ распределена $\mathcal{N}(0; \frac{4}{n})$. Отсюда искомый критерий имеет вид: Если $\bar{X} \geqslant 0.825$, то гипотеза H_0 отвергается в пользу гипотезы H_a.

21.9 Упрощая неравенство из леммы Неймана-Пирсона, получаем критерий: если $X_1 \cdot X_2 > t$, то H_0 отвергается. Величину t находим из уравнения

$$\int_0^t (1 - t/x)\,dx = 0.05$$

21.10

21.11

22.1 $\mathbb{E}(Ay) = A\mathbb{E}(y)$, $\operatorname{Var}(Ay) = A\operatorname{Var}(y)A^T$

22.2

22.3

22.4

22.5

22.6

23.1 Вычтем $2x^2 + 2y^2 + 2z^2 + 2w^2$. Получим, что оптимальное $x = 0$. Далее, $z = 0$, $y = 0$. В итоге $w = 1$ или $w = -1$.

23.2

23.3

23.4

23.5 $(5+4+1)=10$; длины — $\sqrt{5}\cdot\sqrt{99}$, $\sqrt{4}\cdot\sqrt{99}$, $\sqrt{1}\cdot\sqrt{99}$; выборочные дисперсии — 5, 4, 1; $(5+4)/10=0.9$.

23.6

24.1 $C_{20}^2\cdot 18$.

24.2 $1+x+x^2+x^3+x^4=(1-x^5)/(1-x)$ $1+x+x^2+x^3+\ldots=1/(1-x)$ $(1+x)^4$

24.3

24.4

24.5 0 с вероятностью $1/4$ и 1 с вероятностью $3/4$

24.6 да, сможет!

24.7 $F_n=F_{n-1}+F_{n-2}$. Запишем разложения для $g(x)$, $xg(x)$ и $x^2g(x)$ друг под другом. Вычитаем. Получаем, что $g(x)=x/(1-x-x^2)$.

Указанная дробь — это и есть производящая функция при маленьком x.

Производящая функция представима в виде суммы:

$$g(x)=\frac{1}{\sqrt{5}}\left(\frac{a}{x+a}-\frac{b}{x+b}\right),$$

где $a=(1-\sqrt{5})/2$, $b=(1+\sqrt{5})/2$.

24.8 Замечаем, что $h_2(t)=h_1(t)\cdot h_1(t)$. После первого шага:

$$h_1(t)=0.7t+0.3th_1^2(t)$$

24.9

25.1

25.2

27 Источники мудрости

[1] Herbert S Wilf. *generatingfunctionology*. Elsevier, 2013. URL: https://www.math.upenn.edu/~wilf/DownldGF.html. Шикарная книжка про производящие функции.