Подборка экзаменов по теории вероятностей. Факультет экономики, НИУ ВШЭ

Коллектив кафедры
математической экономики и эконометрики, талантливые студенты, фольклор

7 июля 2018 г.

Содержание

Описание

Свежую версию можно скачать с github-репозитория https://github.com/bdemeshev/probability_hse_exams.
Красные ссылки внутри pdf-файла кликабельны, и ведут на ответы и обратно.
Уникальное предложение для студентов факультета экономики НИУ-ВШЭ:
Найдите ошибки в этом документе или пришлите отсутствующие решения в техе и получите дополнительные бонусы! Найденные смысловые ошибки поощряются сильнее, чем просто опечатки. Замеченные ошибки и новые решения оформляйте в виде запросов на https://github.com/bdemeshev/probability_hse_exams/issues/. Перед публикацией запроса, пожалуйста, свертесь со свежей версией подборки.
В создании подборки храбро участвовали Андрей Зубанов, Кирилл Пономарёв, Александр Левкун, Оля Гнилова, Настя Жаркова, Гарик Варданян и другие :)

Доброе напутствие пишущим эту подборку :)

Здесь перечислены стилевые особенности коллекции и самые популярные ошибки. Узнать технические подробности по теху можно, например, в учебнике К.В. Воронцова.

1. Дробную часть числа отделяй от целой точкой: 3.14 — хорошо, $3,14$ — плохо. Это нарушает русскую традицию, но облегчает копирование-вставку в любой программный пакет.

2. Существует длинное тире, —, которое отличается от просто дефиса - и нужно, чтобы разделять части предложения, Инструкция в картинках по набору тире :)

3. Выключные формулы следует окружать \[...\]. Никаких $$...$$!

4. Про остальные окружения: для системы уравнений подойдёт cases, для формул на несколько строк – multline*, для нумерации – enumerate.

5. Русский текст внутри формулы нужно писать в \text{...}.

6. Для многоточий существует команда \ldots.

7. В преамбуле определены сокращения! Самые популярные: \P, \E, \Var, \Cov, \Corr, \cN.

8. Названия функций тоже идут со слэшем: \ln, \exp, \cos...

9. Таблицы нужно оформлять по стандарту booktabs. Самый удобный способ сделать это – зайти на tablesgenerator и выбрать там опцию booktabs table style вместо default table style.

10. Уважай букву ё – ставь над ней точки! :)

11. Начинай каждой предложение внутри теховского файла с новой строки. В готовом pdf предложения будут идти без разрыва, а читабельность теха повысится.

1. Минимумы

1.1. Контрольная работа 1

Теоретический минимум

1. Классическое определение вероятности
2. Определение условной вероятности
3. Определение независимости случайных событий
4. Формула полной вероятности
5. Формула Байеса
6. Функция распределения случайной величины. Определение и свойства.
7. Функция плотности. Определение и свойства.
8. Математическое ожидание. Определения для дискретного и абсолютно непрерывного случаев. Свойства.
9. Дисперсия. Определение и свойства.
10. Законы распределений. Определение, $\mathbb{E}(X)$, $\mathrm{Var}(X)$:
    а) Биномиальное распределение
    б) Распределение Пуассона
    в) Геометрическое распределение
    г) Равномерное распределение
    д) Экспоненциальное распределение

Задачный минимум

1. Пусть $\mathbb{P}(A) = 0.3, \mathbb{P}(B) = 0.4, \mathbb{P}(A \cap B) = 0.1$. Найдите

   а) $\mathbb{P}(A|B)$

   б) $\mathbb{P}(A \cup B)$

   в) Являются ли события A и B независимыми?

2. Пусть $\mathbb{P}(A) = 0.5, \mathbb{P}(B) = 0.5, \mathbb{P}(A \cap B) = 0.25$. Найдите

   а) $\mathbb{P}(A|B)$

   б) $\mathbb{P}(A \cup B)$

   в) Являются ли события A и B независимыми?

3. Карлсон выложил кубиками слово КОМБИНАТОРИКА. Малыш выбирает наугад четыре кубика и выкладывает их в случайном порядке. Найдите вероятность того, что при этом получится слово КОРТ.

4. Карлсон выложил кубиками слово КОМБИНАТОРИКА. Малыш выбирает наугад четыре кубика и выкладывает их в случайном порядке. Найдите вероятность того, что при этом получится слово РОТА.

5. В первой урне 7 белых и 3 черных шара, во второй урне 8 белых и 4 черных шара, в третьей урне 2 белых и 13 черных шаров. Из этих урн наугад выбирается одна урна. Какова вероятность того, что шар, взятый наугад из выбранной урны, окажется белым?

6. В первой урне 7 белых и 3 черных шара, во второй урне 8 белых и 4 черных шара, в третьей урне 2 белых и 13 черных шаров. Из этих урн наугад выбирается одна урна. Какова вероятность того, что была выбрана первая урна, если шар, взятый наугад из выбранной урны, оказался белым?

7. В операционном отделе банка работает 80% опытных сотрудников и 20% неопытных. Вероятность совершения ошибки при очередной банковской операции опытным сотрудником равна 0.01, а неопытным — 0.1. Найдите вероятность совершения ошибки при очередной банковской операции в этом отделе.

8. В операционном отделе банка работает 80% опытных сотрудников и 20% неопытных. Вероятность совершения ошибки при очередной банковской операции опытным сотрудником равна 0.01, а неопытным — 0.1. Известно, что при очередной банковской операции была допущена ошибка. Найдите вероятность того, что ошибку допустил неопытный сотрудник.

9. Пусть случайная величина X имеет таблицу распределения:

x	-1	0	1
$\mathbb{P}(X = x)$	0.25	c	0.25

   Найдите

   а) константу c

   б) $\mathbb{P}(X \geqslant 0)$

   в) $\mathbb{P}(X < -3)$

   г) $\mathbb{P}\left(X \in \left[-\frac{1}{2}; \frac{1}{2}\right]\right)$

   д) Функцию распределения случайной величины X

е) Имеет ли случайная величина X плотность распределения?

10. Пусть случайная величина X имеет таблицу распределения:

x	-1	0	1
$\mathbb{P}(X = x)$	0.25	c	0.25

Найдите

а) константу c

б) $\mathbb{E}(X)$

в) $\mathbb{E}(X^2)$

г) $\operatorname{Var}(X)$

д) $\mathbb{E}(|X|)$

11. Пусть случайная величина X имеет таблицу распределения:

x	-1	0	1
$\mathbb{P}(X = x)$	0.25	c	0.5

Найдите

а) константу c

б) $\mathbb{P}(X \geqslant 0)$

в) $\mathbb{P}(X < -3)$

г) $\mathbb{P}\left(X \in \left[-\frac{1}{2}; \frac{1}{2}\right]\right)$

д) Функцию распределения случайной величины X

е) Имеет ли случайная величина X плотность распределения?

12. Пусть случайная величина X имеет таблицу распределения:

x	-1	0	1
$\mathbb{P}(X = x)$	0.25	c	0.5

Найдите

а) константу c

б) $\mathbb{E}(X)$

в) $\mathbb{E}(X^2)$

г) $\operatorname{Var}(X)$

д) $\mathbb{E}(|X|)$

13. Пусть случайная величина X имеет биномиальное распределение с параметрами $n = 4$ и $p = 0.75$. Найдите

а) $\mathbb{P}(X = 0)$

б) $\mathbb{P}(X > 0)$

в) $\mathbb{P}(X < 0)$

г) $\mathbb{E}(X)$

д) $\operatorname{Var}(X)$

е) Наиболее вероятное значение, которое принимает случайная величина X

14. Пусть случайная величина X имеет биномиальное распределение с параметрами $n = 5$ и $p = 0.4$. Найдите
    а) $\mathbb{P}(X = 0)$
    б) $\mathbb{P}(X > 0)$
    в) $\mathbb{P}(X < 0)$
    г) $\mathbb{E}(X)$
    д) $\mathrm{Var}(X)$
    е) Наиболее вероятное значение, которое принимает случайная величина X

15. Пусть случайная величина X имеет распределение Пуассона с параметром $\lambda = 100$. Найдите
    а) $\mathbb{P}(X = 0)$
    б) $\mathbb{P}(X > 0)$
    в) $\mathbb{P}(X < 0)$
    г) $\mathbb{E}(X)$
    д) $\mathrm{Var}(X)$
    е) Наиболее вероятное значение, которое принимает случайная величина X

16. Пусть случайная величина X имеет распределение Пуассона с параметром $\lambda = 101$. Найдите
    а) $\mathbb{P}(X = 0)$
    б) $\mathbb{P}(X > 0)$
    в) $\mathbb{P}(X < 0)$
    г) $\mathbb{E}(X)$
    д) $\mathrm{Var}(X)$
    е) Наиболее вероятное значение, которое принимает случайная величина X

17. В лифт 10-этажного дома на первом этаже вошли 5 человек. Вычислите вероятность того, что на 6-м этаже выйдет хотя бы один человек.

18. В лифт 10-этажного дома на первом этаже вошли 5 человек. Вычислите вероятность того, что на 6-м этаже не выйдет ни один человек.

19. При работе некоторого устройства время от времени возникают сбои. Количество сбоев за сутки имеет распределение Пуассона. Среднее количество сбоев за сутки равно 3. Найти вероятность того, что в течение суток произойдет хотя бы один сбой.

20. При работе некоторого устройства время от времени возникают сбои. Количество сбоев за сутки имеет распределение Пуассона. Среднее количество сбоев за сутки равно 3. Найти вероятность того, что за двое суток не произойдет ни одного сбоя.

21. Пусть случайная величина X имеет плотность распределения

$$f_X(x) = \begin{cases} c, & \text{при } x \in [-1; 1] \\ 0, & \text{при } x \notin [-1; 1] \end{cases}$$

    Найдите

а) константу c

б) $\mathbb{P}(X \leqslant 0)$

в) $\mathbb{P}\left(X \in \left[\frac{1}{2}; \frac{3}{2}\right]\right)$

г) $\mathbb{P}(X \in [2; 3])$

д) $F_X(x)$

22. Пусть случайная величина X имеет плотность распределения

$$f_X(x) = \begin{cases} c, \text{ при } x \in [-1; 1] \\ 0, \text{ при } x \notin [-1; 1] \end{cases}$$

Найдите

а) константу c

б) $\mathbb{E}(X)$

в) $\mathbb{E}\left(X^2\right)$

г) $\mathrm{Var}(X)$

д) $\mathbb{E}(|X|)$

23. Пусть случайная величина X имеет плотность распределения

$$f_X(x) = \begin{cases} cx, \text{ при } x \in [0; 1] \\ 0, \text{ при } x \notin [0; 1] \end{cases}$$

Найдите

а) константу c

б) $\mathbb{P}\left(X \leqslant \frac{1}{2}\right)$

в) $\mathbb{P}\left(X \in \left[\frac{1}{2}; \frac{3}{2}\right]\right)$

г) $\mathbb{P}(X \in [2; 3])$

д) $F_X(x)$

24. Пусть случайная величина X имеет плотность распределения

$$f_X(x) = \begin{cases} cx, \text{ при } x \in [0; 1] \\ 0, \text{ при } x \notin [0; 1] \end{cases}$$

Найдите

а) константу c

б) $\mathbb{E}(X)$

в) $\mathbb{E}\left(X^2\right)$

г) $\mathrm{Var}(X)$

д) $\mathbb{E}(\sqrt{X})$

1.2. Контрольная работа 2

Теоретический минимум

1. Сформулируйте определение независимости событий, формулу полной вероятности.
2. Приведите определение условной вероятности случайного события, формулу Байеса.
3. Сформулируйте определение и свойства функции распределения случайной величины.
4. Сформулируйте определение и свойства функции плотности случайной величины.
5. Сформулируйте определение и свойства математического ожидания для абсолютно непрерывной случайной величины.
6. Сформулируйте определение и свойства математического ожидания для дискретной случайной величины.
7. Сформулируйте определение и свойства дисперсии случайной величины.
8. Сформулируйте определения следующих законов распределений: биномиального, Пуассона, шеометрического, равномерного, экспоненциального, нормального. Укажите математическое ожидание и дисперсию.
9. Сформулируйте определение функции совместного распределения двух случайных величин, независимости случайных величин. Укажите, как связаны совместное распределение и частные распределения компонент случайного вектора.
10. Сформулируйте определение и свойства совместной функции плотности двух случайных величин, сформулируйте определение независимости случайных величин.
11. Сформулируйте определение и свойства ковариации случайных величин.
12. Сформулируйте определение и свойства корреляции случайных величин.
13. Сформулируйте определение и свойства условной функции плотности.
14. Сформулируйте определение условного математического ожидания $\mathbb{E}(Y|X = x)$ для совместного дискретного и совместного абсолютно непрерывного распределений.
15. Сформулируйте определение математического ожидания и ковариационной матрицы случайного вектора и их свойства.
16. Сформулируйте неравенство Чебышёва и неравенство Маркова.
17. Сформулируйте закон больших чисел в слабой форме.
18. Сформулируйте центральную предельную теорему.
19. Сформулируйте теорему Муавра—Лапласа.
20. Сформулируйте определение сходимости по вероятности для последовательности случайных величин.

Задачный минимум

1. Пусть задана таблица совместного распределения случайных величин X и Y.

$X \backslash Y$	-1	0	1
-1	0.2	0.1	0.2
1	0.1	0.3	0.1

Найдите

а) $\mathbb{P}(X = -1)$

б) $\mathbb{P}(Y = -1)$

в) $\mathbb{P}(X = -1 \cap Y = -1)$

г) Являются ли случайные величины X и Y независимыми?

д) $F_{X,Y}(-1, 0)$

е) Таблицу распределения случайной величины X

ж) Функцию $F_X(x)$ распределения случайной величины X

з) Постройте график функции $F_X(x)$ распределения случайной величины X

2. Пусть задана таблица совместного распределения случайных величин X и Y.

$X \backslash Y$	-1	0	1
-1	0.2	0.1	0.2
1	0.2	0.1	0.2

Найдите

а) $\mathbb{P}(X = 1)$

б) $\mathbb{P}(Y = 1)$

в) $\mathbb{P}(X = 1 \cap Y = 1)$

г) Являются ли случайные величины X и Y независимыми?

д) $F_{X,Y}(1, 0)$

е) Таблицу распределения случайной величины Y

ж) Функцию $F_Y(y)$ распределения случайной величины Y

з) Постройте график функции $F_Y(y)$ распределения случайной величины Y

3. Пусть задана таблица совместного распределения случайных величин X и Y.

$X \backslash Y$	-1	0	1
-1	0.2	0.1	0.2
1	0.1	0.3	0.1

Найдите

а) $\mathbb{E}(X)$

б) $\mathbb{E}(X^2)$

в) $\text{Var}(X)$

г) $\mathbb{E}(Y)$

д) $\mathbb{E}(Y^2)$

е) $\text{Var}(Y)$

ж) $\mathbb{E}(XY)$

з) $\text{Cov}(X, Y)$

и) $\text{Corr}(X, Y)$

к) Являются ли случайные величины X и Y некоррелированными?

4. Пусть задана таблица совместного распределения случайных величин X и Y.

$X \setminus Y$	-1	0	1
-1	0.2	0.1	0.2
1	0.2	0.1	0.2

Найдите

а) $\mathbb{E}(X)$

б) $\mathbb{E}(X^2)$

в) $\text{Var}(X)$

г) $\mathbb{E}(Y)$

д) $\mathbb{E}(Y^2)$

е) $\text{Var}(Y)$

ж) $\mathbb{E}(XY)$

з) $\text{Cov}(X, Y)$

и) $\text{Corr}(X, Y)$

к) Являются ли случайные величины X и Y некоррелированными?

5. Пусть задана таблица совместного распределения случайных величин X и Y.

$X \setminus Y$	-1	0	1
-1	0.2	0.1	0.2
1	0.1	0.3	0.1

Найдите

а) $\mathbb{P}(X = -1 | Y = 0)$

б) $\mathbb{P}(Y = 0 | X = -1)$

в) Таблицу условного распределения случайной величины Y при условии $X = -1$

г) Условное математическое ожидание случайной величины Y при $X = -1$

д) Условную дисперсию случайной величины Y при условии $X = -1$

6. Пусть задана таблица совместного распределения случайных величин X и Y.

$X \backslash Y$	-1	0	1
-1	0.2	0.1	0.2
1	0.2	0.1	0.2

Найдите

а) $\mathbb{P}(X = 1 | Y = 0)$

б) $\mathbb{P}(Y = 0 | X = 1)$

в) Таблицу условного распределения случайной величины Y при условии $X = 1$

г) Условное математическое ожидание случайной величины Y при $X = 1$

д) Условную дисперсию случайной величины Y при условии $X = 1$

7. Пусть $\mathbb{E}(X) = 1$, $\mathbb{E}(Y) = 2$, $\text{Var}(X) = 3$, $\text{Var}(Y) = 4$, $\text{Cov}(X, Y) = -1$. Найдите

а) $\mathbb{E}(2X + Y - 4)$

б) $\text{Var}(3Y + 3)$

в) $\text{Var}(X - Y)$

г) $\text{Var}(2X - 3Y + 1)$

д) $\text{Cov}(X + 2Y + 1, 3X - Y - 1)$

е) $\text{Corr}(X + Y, X - Y)$

ж) Ковариационную матрицу случайного вектора $Z = (X \quad Y)$

8. Пусть $\mathbb{E}(X) = -1$, $\mathbb{E}(Y) = 2$, $\text{Var}(X) = 1$, $\text{Var}(Y) = 2$, $\text{Cov}(X, Y) = 1$. Найдите

а) $\mathbb{E}(2X + Y - 4)$

б) $\text{Var}(2Y + 3)$

в) $\text{Var}(X - Y)$

г) $\text{Var}(2X - 3Y + 1)$

д) $\text{Cov}(3X + Y + 1, X - 2Y - 1)$

е) $\text{Corr}(X + Y, X - Y)$

ж) Ковариационную матрицу случайного вектора $Z = (X \quad Y)$

9. Пусть случайная величина X имеет стандартное нормальное распределение. Найдите

а) $\mathbb{P}(0 < X < 1)$

б) $\mathbb{P}(X > 2)$

в) $\mathbb{P}(0 < 1 - 2X \leqslant 1)$

10. Пусть случайная величина X имеет стандартное нормальное распределение. Найдите

а) $\mathbb{P}(-1 < X < 1)$

б) $\mathbb{P}(X < -2)$

в) $\mathbb{P}(-2 < -X + 1 \leqslant 0)$

11. Пусть случайная величина $X \sim \mathcal{N}(1,4)$. Найдите $\mathbb{P}(1 < X < 4)$.

12. Пусть случайная величина $X \sim \mathcal{N}(2,4)$. Найдите $\mathbb{P}(-2 < X < 4)$.

13. Случайные величины X и Y независимы и имеют нормальное распределение, $\mathbb{E}(X) = 0$, $\mathrm{Var}(X) = 1$, $\mathbb{E}(Y) = 2$, $\mathrm{Var}(Y) = 6$. Найдите $\mathbb{P}(1 < X + 2Y < 7)$.

14. Случайные величины X и Y независимы и имеют нормальное распределение, $\mathbb{E}(X) = 0$, $\mathrm{Var}(X) = 1$, $\mathbb{E}(Y) = 3$, $\mathrm{Var}(Y) = 7$. Найдите $\mathbb{P}(1 < 3X + Y < 7)$.

15. Игральная кость подбрасывается 420 раз. При помощи центральной предельной теоремы приближенно найти вероятность того, что суммарное число очков будет находиться в пределах от 1400 до 1505?

16. При выстреле по мишени стрелок попадает в десятку с вероятностью 0.5, в девятку – 0.3, в восьмерку – 0.1, в семерку – 0.05, в шестерку – 0.05. Стрелок сделал 100 выстрелов. При помощи центральной предельной теоремы приближенно найти вероятность того, что он набрал не менее 900 очков?

17. Предположим, что на станцию скорой помощи поступают вызовы, число которых распределено по закону Пуассона с параметром $\lambda = 73$, и в разные сутки их количество не зависит друг от друга. При помощи центральной предельной теоремы приближенно найти вероятность того, что в течение года (365 дней) общее число вызовов будет в пределах от 26500 до 26800.

18. Число посетителей магазина (в день) имеет распределение Пуассона с математическим ожиданием 289. При помощи центральной предельной теоремы приближенно найти вероятность того, что за 100 рабочих дней суммарное число посетителей составит от 28550 до 29250 человек.

19. Пусть плотность распределения случайного вектора (X, Y) имеет вид

$$f_{X,Y}(x,y) = \begin{cases} x+y, & \text{при } (x,y) \in [0;1] \times [0;1] \\ 0, & \text{при } (x,y) \notin [0;1] \times [0;1] \end{cases}$$

Найдите

а) $\mathbb{P}(X \leqslant \frac{1}{2} \cap Y \leqslant \frac{1}{2})$,

б) $\mathbb{P}(X \leqslant Y)$,

в) $f_X(x)$,

г) $f_Y(y)$,

д) Являются ли случайные величины X и Y независимыми?

20. Пусть плотность распределения случайного вектора (X, Y) имеет вид

$$f_{X,Y}(x,y) = \begin{cases} 4xy, & \text{при } (x,y) \in [0;1] \times [0;1] \\ 0, & \text{при } (x,y) \notin [0;1] \times [0;1] \end{cases}$$

Найдите

а) $\mathbb{P}(X \leqslant \frac{1}{2} \cap Y \leqslant \frac{1}{2})$,

б) $\mathbb{P}(X \leqslant Y)$,

в) $f_X(x)$,

г) $f_Y(y)$,

д) Являются ли случайные величины X и Y независимыми?

21. Пусть плотность распределения случайного вектора (X, Y) имеет вид

$$f_{X,Y}(x,y) = \begin{cases} x+y, & \text{при } (x,y) \in [0;1] \times [0;1] \\ 0, & \text{при } (x,y) \notin [0;1] \times [0;1] \end{cases}$$

Найдите

а) $\mathbb{E}(X)$,

б) $\mathbb{E}(Y)$,

в) $\mathbb{E}(XY)$,

г) $\mathrm{Cov}(X, Y)$,

д) $\mathrm{Corr}(X, Y)$.

22. Пусть плотность распределения случайного вектора (X, Y) имеет вид

$$f_{X,Y}(x,y) = \begin{cases} 4xy, & \text{при } (x,y) \in [0;1] \times [0;1] \\ 0, & \text{при } (x,y) \notin [0;1] \times [0;1] \end{cases}$$

Найдите

а) $\mathbb{E}(X)$,

б) $\mathbb{E}(Y)$,

в) $\mathbb{E}(XY)$,

г) $\mathrm{Cov}(X, Y)$,

д) $\mathrm{Corr}(X, Y)$.

23. Пусть плотность распределения случайного вектора (X, Y) имеет вид

$$f_{X,Y}(x,y) = \begin{cases} x+y, & \text{при } (x,y) \in [0;1] \times [0;1] \\ 0, & \text{при } (x,y) \notin [0;1] \times [0;1] \end{cases}$$

Найдите

а) $f_Y(y)$,

б) $f_{X|Y}\left(x|\frac{1}{2}\right)$

в) $\mathbb{E}\left(X|Y = \frac{1}{2}\right)$

г) $\mathrm{Var}\left(X|Y = \frac{1}{2}\right)$

24. Пусть плотность распределения случайного вектора (X, Y) имеет вид

$$f_{X,Y}(x,y) = \begin{cases} 4xy, & \text{при } (x,y) \in [0;1] \times [0;1] \\ 0, & \text{при } (x,y) \notin [0;1] \times [0;1] \end{cases}$$

Найдите

а) $f_Y(y)$,

б) $f_{X|Y}\left(x|\frac{1}{2}\right)$

в) $\mathbb{E}\left(X|Y = \frac{1}{2}\right)$

г) $\mathrm{Var}\left(X|Y = \frac{1}{2}\right)$

1.3. Контрольная работа 3

Теоретический минимум

1. Дайте определение нормально распределённой случайной величины. Укажите диапазон возможных значений, функцию плотности, ожидание, дисперсию. Нарисуйте функцию плотности.

2. Дайте определение хи-квадрат распределения. Укажите диапазон возможных значений, выражение через нормальные распределения, математическое ожидание. Нарисуйте функцию плотности при разных степенях свободы.

3. Дайте определение распределения Стьюдента. Укажите диапазон возможных значений, выражение через нормальные распределения. Нарисуйте функцию плотности распределения Стьюдента при разных степенях свободы на фоне нормальной стандартной функции плотности.

4. Дайте определение распределения Фишера. Укажите диапазон возможных значений, выражение через нормальные распределенеия. Нарисуйте возможную функцию плотности.

Для следующего блока вопросов предполагается, что имеется случайная выборка $X_1, X_2, \ldots, X_n$ из распределения с функцией плотности $f(x, \theta)$, зависящей от от параметра θ. Дайте определение каждого понятия из списка или сформулируйте соответствующую теорему:

5. Выборочное среднее и выборочная дисперсия;

6. Формула несмещённой оценки дисперсии;

7. Выборочный начальный момент порядка k;

8. Выборочный центральный момент порядка k;

9. Выборочная функция распределения;

10. Несмещённая оценка $\hat{\theta}$ параметра θ;

11. Состоятельная последовательность оценок $\hat{\theta}_n$;

12. Эффективность оценки $\hat{\theta}$ среди множества оценок $\hat{\Theta}$;

13. Неравенство Крамера–Рао для несмещённых оценок;

14. Функция правдоподобия и логарифмическая функция правдоподобия;

15. Информация Фишера о параметре θ, содержащаяся в одном наблюдении;

16. Оценка метода моментов параметра θ при использовании первого момента, если $\mathbb{E}(X_i) = g(\theta)$ и существует обратная функция g^{-1};

17. Оценка метода максимального правдоподобия параметра θ;

Для следующего блока вопросов предполагается, что величины $X_1, X_2, \ldots, X_n$ независимы и нормальны $\mathcal{N}(\mu; \sigma^2)$.

18. Укажите закон распределения выборочного среднего, величины $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$, величины $\frac{\bar{X}-\mu}{\hat{\sigma}/\sqrt{n}}$, величины $\frac{\hat{\sigma}^2(n-1)}{\sigma^2}$;

19. Укажите формулу доверительного интервала с уровнем доверия $(1-\alpha)$ для μ при известной дисперсии, для μ при неизвестной дисперсии, для σ^2;

Задачный миннимум

1. Рост в сантиметрах (случайная величина X) и вес в килограммах (случайная величина Y) взрослого мужчины является нормальным случайным вектором $Z = (X, Y)$ с математическим ожиданием $\mathbb{E}(Z) = (175, 74)$ и ковариационной матрицей

$$\mathrm{Var}(Z) = \begin{pmatrix} 49 & 28 \\ 28 & 36 \end{pmatrix}$$

Лишний вес характеризуется случайной величиной $U = X - Y$. Считается, что человек страдает избыточным весом, если $U < 90$.

а) Определите вероятность того, что рост мужчины отклоняется от среднего более, чем на 10 см.

б) Укажите распределение случайной величины U. Выпишите её плотность распределения.

в) Найдите вероятность того, что случайно выбранный мужчина страдает избыточным весом.

2. Рост в сантиметрах, случайная величина X, и вес в килограммах, случайная величина Y, взрослого мужчины является нормальным случайным вектором $Z = (X, Y)$ с математическим ожиданием $\mathbb{E}(Z) = (175, 74)$ и ковариационной матрицей

$$\mathrm{Var}(Z) = \begin{pmatrix} 49 & 28 \\ 28 & 36 \end{pmatrix}$$

а) Найдите средний вес мужчины при условии, что его рост составляет 170 см.

б) Выпишите условную плотность распределения веса мужчины при условии, что его рост составляет 170 см.

в) Найдите условную вероятность того, что человек будет иметь вес, больший 90 кг, при условии, что его рост составляет 170 см.

3. Для реализации случайной выборки $x = (1, 0, -1, 1)$ найдите:

а) выборочное среднее,

б) неисправленную выборочную дисперсию,

в) исправленную выборочную дисперсию,

г) выборочный второй начальный момент,

д) выборочный третий центральный момент.

4. Для реализации случайной выборки $x = (1, 0, -1, 1)$ найдите:

а) вариационный ряд,

б) первый член вариационного ряда,

в) последний член вариационного ряда,

г) график выборочной функции распределения.

5. Пусть $X = (X_1, \ldots, X_n)$ — случайная выборка из дискретного распределения, заданного с помощью таблицы

x	-3	0	2
$\mathbb{P}(X_i = x)$	$2/3 - \theta$	$1/3$	θ

Рассмотрите оценку $\hat{\theta} = \dfrac{\bar{X} + 2}{5}$.

а) Найдите $\mathbb{E}\left(\hat{\theta}\right)$.

б) Является ли оценка $\hat{\theta}$ несмещенной оценкой неизвестного параметра θ?

6. Пусть $X = (X_1, \ldots, X_n)$ — случайная выборка из распределения с плотностью распределения

$$f(x, \theta) = \begin{cases} \dfrac{6x(\theta - x)}{\theta^3} & \text{при } x \in [0; \theta], \\ 0 & \text{при } x \notin [0; \theta], \end{cases}$$

где $\theta > 0$ — неизвестный параметр распределения и $\hat{\theta} = \bar{X}$.

а) Является ли оценка $\hat{\theta} = \bar{X}$ несмещенной оценкой неизвестного параметра θ?

б) Подберите константу c так, чтобы оценка $\tilde{\theta} = c\bar{X}$ оказалась несмещенной оценкой неизвестного параметра θ.

7. Пусть $X = (X_1, X_2, X_3)$ — случайная выборка из распределения Бернулли с неизвестным параметром $p \in (0, 1)$. Какие из следующих ниже оценкой являются несмещенными? Среди перечисленных ниже оценок найдите наиболее эффективную оценку:

- $\hat{p}_1 = \dfrac{X_1 + X_3}{2}$,
- $\hat{p}_2 = \frac{1}{4}X_1 + \frac{1}{2}X_2 + \frac{1}{4}X_3$,
- $\hat{p}_3 = \frac{1}{3}X_1 + \frac{1}{3}X_2 + \frac{1}{3}X_3$.

8. Пусть $X = (X_1, \ldots, X_n)$ — случайная выборка из распределения с плотностью

$$f(x, \theta) = \begin{cases} \frac{1}{\theta}\, e^{-\frac{x}{\theta}} & \text{при } x \geqslant 0, \\ 0 & \text{при } x < 0, \end{cases}$$

где $\theta > 0$ — неизвестный параметр. Является ли оценка $\hat{\theta}_n = \frac{X_1 + \ldots + X_n}{n+1}$ состоятельной?

9. Пусть $X = (X_1, \ldots, X_n)$ — случайная выборка из распределения с плотностью распределения

$$f(x, \theta) = \begin{cases} \dfrac{6x(\theta - x)}{\theta^3} & \text{при } x \in [0; \theta], \\ 0 & \text{при } x \notin [0; \theta], \end{cases}$$

где $\theta > 0$ — неизвестный параметр распределения. Является ли оценка $\hat{\theta}_n = \frac{2n+1}{n}\bar{X}_n$ состоятельной оценкой неизвестного параметра θ?

10. Пусть $X = (X_1, \ldots, X_n)$ — случайная выборка из распределения с плотностью распределения

$$f(x, \theta) = \begin{cases} \dfrac{6x(\theta - x)}{\theta^3} & \text{при } x \in [0; \theta], \\ 0 & \text{при } x \notin [0; \theta], \end{cases}$$

где $\theta > 0$ — неизвестный параметр распределения. Используя центральный момент второго порядка, при помощи метода моментов найдите оценку для неизвестного параметра θ.

11. Пусть $X = (X_1, \ldots, X_n)$ — случайная выборка. Случайные величины $X_1, \ldots, X_n$ имеют дискретное распределение, которое задано при помощи таблицы

x	-3	0	2
$\mathbb{P}(X_i = x)$	$2/3 - \theta$	$1/3$	θ

Используя второй начальный момент, при помощи метода моментов найдите оценку неизвестного параметра θ. Для реализации случайной выборки $x = (0, 0, -3, 0, 2)$ найдите числовое значение найденной оценки параметра θ.

12. Пусть $X = (X_1, \ldots, X_n)$ — случайная выборка из распределения с плотностью распределения

$$f(x, \theta) = \begin{cases} \frac{2x}{\theta}\, e^{-\frac{x^2}{\theta}} & \text{при } x > 0, \\ 0 & \text{при } x \leqslant 0, \end{cases}$$

где $\theta > 0$. При помощи метода максимального правдоподобия найдите оценку неизвестного параметра θ.

13. Пусть $X = (X_1, \ldots, X_n)$ – случайная выборка из распределения Бернулли с параметром $p \in (0; 1)$. При помощи метода максимального правдоподобия найдите оценку неизвестного параметра p.

14. Пусть $X = (X_1, \ldots, X_n)$ — случайная выборка из распределения с плотностью

$$f(x, \theta) = \begin{cases} \frac{1}{\theta}\, e^{-\frac{x}{\theta}} & \text{при } x \geqslant 0, \\ 0 & \text{при } x < 0, \end{cases}$$

где $\theta > 0$ — неизвестный параметр. Является ли оценка $\hat{\theta} = \bar{X}$ эффективной?

15. Стоимость выборочного исследования генеральной совокупности, состоящей из трех страт, определяется по формуле $TC = c_1 n_1 + c_2 n_2 + c_3 n_3$, где c_i — цена одного наблюдения в i-ой страте, а n_i — число наблюдений, которые приходятся на i-ую страту. Найдите n_1, n_2 и n_3, при которых дисперсия стратифицированного среднего достигает наименьшего значения, если бюджет исследования 8000 и имеется следующая информация:

Страта	1	2	3
Среднее значение	30	40	50
Стандартная ошибка	5	10	20
Вес	25%	25%	50%
Цена наблюдения	1	5	10

1.4. Контрольная работа 4

Теоретический минимум

1. Дайте определение ошибки первого и второго рода, критической области.

2. Укажите формулу доверительного интервала с уровнем доверия $(1-\alpha)$ для вероятности успеха, построенного по случайной выборке большого размера из распределения Бернулли $Bin(1, p)$.

Для следующего блока вопросов предполагается, что величины $X_1, X_2, \ldots, X_n$ независимы и нормальны $\mathcal{N}(\mu; \sigma^2)$. Укажите формулу для статистики:

3. Статистика, проверяющая гипотезу о математическом ожидании при известной дисперсии σ^2, и её распределение при справедливости основной гипотезы H_0: $\mu = \mu_0$.

4. Статистика, проверяющая гипотезу о математическом ожидании при неизвестной дисперсии σ^2, и её распределение при справедливости основной гипотезы H_0: $\mu = \mu_0$.

Для следующего блока вопросов предполагается, что есть две независимые случайные выборки: выборка $X_1, X_2, \ldots$ размера n_x из нормального распределения $\mathcal{N}(\mu_x; \sigma_x^2)$ и выборка $Y_1, Y_2, \ldots$ размера n_y из нормального распределения $\mathcal{N}(\mu_y; \sigma_y^2)$.
Укажите формулу для статистики или границ доверительного интервала:

5. Доверительный интервал для разницы математических ожиданий, когда дисперсии известны;

6. Доверительный интервал для разницы математических ожиданий, когда дисперсии не известны, но равны;

7. Статистика, проверяющая гипотезу о разнице математических ожиданий при известных дисперсиях, и её распределение при справедливости основной гипотезы H_0: $\mu_x - \mu_y = \Delta_0$;

8. Статистика, проверяющая гипотезу о разнице математических ожиданий при неизвестных, но равных дисперсиях, и её распределение при справедливости основной гипотезы H_0: $\mu_x - \mu_y = \Delta_0$;

9. Статистика, проверяющая гипотезу о равенстве дисперсий, и её распределение при справедливости основной гипотезы H_0: $\sigma_x^2 = \sigma_y^2$.

Задачный минимум

1. Пусть $X = (X_1, \ldots, X_n)$ — случайная выборка из нормального распределения с параметрами μ и $\sigma^2 = 4$. Используя реализацию случайной выборки,

$$x_1 = -1.11, \quad x_2 = -6.10, \quad x_3 = 2.42,$$

построите 90%-ый доверительный интервал для неизвестного параметра μ.

2. Пусть $X = (X_1, \ldots, X_n)$ — случайная выборка из нормального распределения с неизвестными параметрами μ и σ^2. Используя реализацию случайной выборки,

$$x_1 = -1.11, \quad x_2 = -6.10, \quad x_3 = 2.42,$$

постройте 90%-ый доверительный интервал для неизвестного параметра μ.

3. Пусть $X = (X_1, \ldots, X_n)$ — случайная выборка из нормального распределения с неизвестными параметрами μ и σ^2. Используя реализацию случайной выборки,

$$x_1 = 1.07, \quad x_2 = 3.66, \quad x_3 = -4.51,$$

постройте 80%-ый доверительный интервал для неизвестного параметра σ^2.

4. Пусть $X = (X_1, \ldots, X_n)$ и $Y = (Y_1, \ldots, Y_m)$ — независимые случайные выборки из нормального распределения с параметрами (μ_X, σ_X^2) и (μ_Y, σ_Y^2) соответственно, причем $\sigma_X^2 = 2$ и $\sigma_Y^2 = 1$. Используя реализации случайных выборок

$$x_1 = -1.11, \quad x_2 = -6.10, \quad x_3 = 2.42,$$
$$y_1 = -2.29, \quad y_2 = -2.91,$$

постройте 95%-ый доверительный интервал для разности математических ожиданий $\mu_X - \mu_Y$.

5. Пусть $X = (X_1, \ldots, X_n)$ и $Y = (Y_1, \ldots, Y_m)$ — независимые случайные выборки из нормального распределения с параметрами (μ_X, σ_X^2) и (μ_Y, σ_Y^2) соответственно. Известно, что $\sigma_X^2 = \sigma_Y^2$. Используя реализации случайных выборок

$$x_1 = 1.53, \quad x_2 = 2.83, \quad x_3 = -1.25$$
$$y_1 = -0.8, \quad y_2 = 0.06$$

постройте 95%-ый доверительный интервал для разности математических ожиданий $\mu_X - \mu_Y$.

6. Пусть $X = (X_1, \ldots, X_n)$ — случайная выборка из распределения Бернулли с параметром p. Используя реализацию случайной выборки $x = (x_1, \ldots, x_n)$, в которой 55 нулей и 45 единиц, постройте приближенный 95%-ый доверительный интервал для неизвестного параметра p.

7. Пусть $X = (X_1, \ldots, X_n)$ и $Y = (Y_1, \ldots, Y_m)$ — независимые случайные выборки из распределения Бернулли с параметрами $p_X \in (0; 1)$ и $p_Y \in (0; 1)$ соответственно. Известно, что $n = 100$, $\bar{x}_n = 0.6$, $m = 200$, $\bar{y}_m = 0.4$. Постройте приближенный 95%-ый доверительный интервал для отношения разности вероятностей успеха $p_X - p_Y$.

8. Дядя Вова (Владимир Николаевич) и Скрипач (Гедеван) зарабатывают на Плюке чатлы, чтобы купить гравицапу. Число заработанных за i-ый день чатлов имеет распределение Пуассона с неизвестным параметром λ. Заработки в различные дни независимы. За прошедшие 100 дней они заработали 250 чатлов.

   С помощью метода максимального правдоподобия постройте приближенный 95%-ый доверительн интервал для неизвестного параметра λ.

9. Пусть $X = (X_1, \dots, X_n)$ — случайная выборка из показательного (экспоненциального) распределени
с плотностью распределения

$$f(x, \lambda) = \begin{cases} \lambda e^{-\lambda x} \text{ при } x \geqslant 0 \\ 0 \text{ при } x < 0 \end{cases}$$

где $\lambda > 0$ — неизвестный параметр распределения. Известно, что $n = 100$ и $\bar{x}_n = 0.52$.

С помощью метода максимального правдоподобия постройте приближенный 95%-ый доверительн
интервал для параметра λ.

10. Пусть $X = (X_1, \dots, X_n)$ — случайная выборка из нормального распределения с неизвестным математическим ожиданием μ и известной дисперсией $\sigma^2 = 4$. Объем выборки $n = 16$. Для тестирования основной гипотезы $H_0 : \mu = 0$ против альтернативной гипотезы $H_1 : \mu = 2$ вы используете критерий: если $\bar{X} \leqslant 1$, то вы не отвергаете гипотезу H_0, в противном случае вы отвергаете гипотезу H_0 в пользу гипотезы H_1. Найдите

    а) вероятность ошибки 1-го рода;
    б) вероятность ошибки 2-го рода;
    в) мощность критерия.

11. На основе случайной выборки, содержащей одно наблюдение X_1, тестируется гипотеза $H_0 : X_1 \sim U[-0.7; 0.3]$ против альтернативной гипотезы $H_1 : X_1 \sim U[-0.3; 0.7]$. Рассматривается критерий вида: если $X_1 > c$, то гипотеза H_0 отвергается в пользу гипотезы H_1. Выберите константу c так, чтобы уровень значимости этого критерия составлял 0.1.

12. Пусть $X = (X_1, \dots, X_n)$ — случайная выборка из нормального распределения с параметрами μ и $\sigma^2 = 4$. Уровень значимости $\alpha = 0.1$. Используя реализацию случайной выборки $x_1 = -1.11, x_2 = -6.10, x_3 = 2.42$, проверьте следующую гипотезу:

$$\begin{cases} H_0 : \mu = 0, \\ H_1 : \mu > 0 \end{cases}$$

13. Пусть $X = (X_1, \dots, X_n)$ — случайная выборка из нормального распределения с параметрами μ и σ^2. Уровень значимости $\alpha = 0.1$. Используя реализацию случайной выборки $x_1 = -1.11, x_2 = -6.10, x_3 = 2.42$, проверьте следующую гипотезу:

$$\begin{cases} H_0 : \mu = 0, \\ H_1 : \mu > 0 \end{cases}$$

14. Пусть $X = (X_1, \dots, X_n)$ и $Y = (Y_1, \dots, Y_m)$ — независимые случайные выборки из нормального распределения с параметрами (μ_X, σ_X^2) и (μ_Y, σ_Y^2) соответственно, причем $\sigma_X^2 = 2$ и $\sigma_Y^2 = 1$. Уровень значимости $\alpha = 0.05$. Используя реализации случайных выборок

$$\begin{aligned} &x_1 = -1.11, \quad x_2 = -6.10, \quad x_3 = 2.42, \\ &y_1 = -2.29, \quad y_2 = -2.91, \end{aligned}$$

проверьте следующую гипотезу:

$$\begin{cases} H_0 : \mu_X = \mu_Y, \\ H_1 : \mu_X < \mu_Y \end{cases}$$

15. Пусть $X = (X_1, \ldots, X_n)$ и $Y = (Y_1, \ldots, Y_m)$ — независимые случайные выборки из нормального распределения с параметрами (μ_X, σ_X^2) и (μ_Y, σ_Y^2) соответственно. Известно, что $\sigma_X^2 = \sigma_Y^2$. Уровень значимости $\alpha = 0.05$. Используя реализации случайных выборок

$$x_1 = 1.53, \quad x_2 = 2.83, \quad x_3 = -1.25$$
$$y_1 = -0.8, \quad y_2 = 0.06$$

проверьте следующую гипотезу:

$$\begin{cases} H_0 : \mu_X = \mu_Y, \\ H_1 : \mu_X < \mu_Y \end{cases}$$

16. Пусть $X = (X_1, \ldots, X_n)$ и $Y = (Y_1, \ldots, Y_m)$ — независимые случайные выборки из нормального распределения с параметрами (μ_X, σ_X^2) и (μ_Y, σ_Y^2) соответственно. Уровень значимости $\alpha = 0.05$. Используя реализации случайных выборок

$$x_1 = -1.11, \quad x_2 = -6.10, \quad x_3 = 2.42,$$
$$y_1 = -2.29, \quad y_2 = -2.91,$$

проверьте следующую гипотезу:

$$\begin{cases} H_0 : \sigma_X^2 = \sigma_Y^2, \\ H_1 : \sigma_X^2 > \sigma_Y^2 \end{cases}$$

17. Пусть $X = (X_1, \ldots, X_n)$ — случайная выборка из распределения Бернулли с неизвестным параметром $p \in (0; 1)$.Имеется следующая информация о реализации случайной выборки, содержащей $n = 100$ наблюдений: $\sum_{i=0}^{n} x_i = 60$. На уровне значимости $\alpha = 0.05$ требуется протестировать следующую гипотезу:

$$\begin{cases} H_0 : p = 0.5, \\ H_1 : p > 0.5 \end{cases}$$

18. Пусть $X = (X_1, \ldots, X_n)$ и $Y = (Y_1, \ldots, Y_m)$ — две независимые случайные выборки из распределения Бернулли с неизвестными параметрами $p_X \in (0; 1)$ и $p_Y \in (0; 1)$. Имеется следующая информация о реализациях этих случайных выборок: $n = 100$, $\sum_{i=1}^{n} x_i = 60$, $m = 150$,$\sum_{j=1}^{m} y_j = 50$. На уровне значимости $\alpha = 0.05$ требуется протестировать следующую гипотезу:

$$\begin{cases} H_0 : p_X = p_Y, \\ H_1 : p_X \neq p_Y \end{cases}$$

19. Вася Сидоров утверждает, что ходит в кино в два раза чаще, чем в спортзал, а в спортзал в два раза чаще, чем в театр. За последние полгода он 10 раз был в театре, 17 раз – в спортзале и 39 раз в кино. На уровне значимости 5% проверьте утверждение Васи.

20. Вася очень любит тестировать статистические гипотезы. В этот раз Вася собирается проверить утверждение о том, что его друг Пётр звонит Васе исключительно в то время, когда Вася ест. Для этого Вася трудился целый год и провел серию из 365 испытаний. Результаты приведены в таблице ниже.

	Пётр звонит	Пётр не звонит
Вася ест	200	40
Вася не ест	25	100

На уровне значимости 5% протестируйте гипотезу о том, что Пётр звонит Васе независимо от момента приема пищи Васей.

21. Пусть $X = (X_1, \ldots, X_n)$ — случайная выборка из нормального распределения с математическим ожиданием $\mu \in \mathbb{R}$ и дисперсией $v > 0$, где μ и v — неизвестные параметры. Известно, что выборка состоит из $n = 100$ наблюдений, $\sum_{i=1}^{n} x_i = 30$, $\sum_{i=1}^{n} x_i^2 = 146$. При помощи теста отношения правдоподобия протестируйте гипотезу $H_0 : v = 1$ на уровне значимости 5%.

2. Контрольная работа 1

2.1. 2017-2018

Минимум

1. Функция распределения случайной величины: определения и свойства.

2. Экспоненциальное распределение: определение, математическое ожидание и дисперсия.

3. В операционном отделе банка работает 80% опытных сотрудников и 20% неопытных. Вероятность совершения ошибки при очередной банковской операции опытным сотрудником равна 0.01, а неопытным — 0.1. Известно, что при очередной банковской операции была допущена ошибка. Найдите вероятность того, что ошибку допустил неопытный сотрудник.

4. При работе некоторого устройства время от времени возникают сбои. Количество сбоев за сутки имеет распределение Пуассона. Среднее количество сбоев за сутки равно 3. Найдите вероятность того, что за двое суток не произойдет ни одного сбоя.

Задачи

1. Правильный кубик подбрасывают один раз. Событие A — выпало чётное число, событие B — выпало число кратное трём, событие C — выпало число, большее трёх.

   а) Сформулируйте определение независимости двух событий;

   б) Определите, какие из пар событий A, B и C будут независимыми.

2. Теоретический минимум (ТМ) состоит из 10 вопросов, задачный (ЗМ) — из 24 задач. Каждый вариант контрольной содержит два вопроса из ТМ и две задачи из ЗМ. Чтобы получить за контрольную работу оценку 4 и выше, необходимо и достаточно правильно ответить на каждый вопрос ТМ и задачу ЗМ доставшегося варианта. Студент Вася принципиально выучил только k вопросов ТМ и две трети ЗМ.

   а) Сколько всего можно составить вариантов, отличающихся хотя бы одним заданием в ТМ или ЗМ части? Порядок заданий внутри варианта не важен.

   б) Найдите вероятность того, что Вася правильно решит задачи ЗМ;

   в) Дополнительно известно, что Васина вероятность правильно ответить на вопросы ТМ, составляет $1/15$. Сколько вопросов ТМ выучил Вася?

3. Производитель молочных продуктов выпустил новый низкокалорийный йогурт Fit и утверждает, что он вкуснее его более калорийного аналога Fat. Четырем независимым экспертам предлагают выбрать наиболее вкусный йогурт из трёх, предлагая им в одинаковых стаканчиках в случайном порядке два Fat и один Fit. Предположим, что йогурты одинаково привлекательны. Величина ξ — число экспертов, отдавших предпочтение Fit.

   а) Какова вероятность, что большинство экспертов выберут Fit?

   б) Постройте функцию распределения величины ξ;

   в) Каково наиболее вероятное число экспертов, отдавших предпочтение йогурту Fit?

   г) Вычислите математическое ожидание и дисперсию ξ.

4. Дядя Фёдор каждую субботу закупает в магазине продукты по списку, составленному котом Матроскином. Список не изменяется, и в него всегда входит 1 кг сметаны, цена которого является равномерно распределённой величиной α, принимающей значения от 250 до 1000 рублей. Стоимость остальных продуктов из списка в тысячах рублей является случайной величиной ξ с функцией распределения

$$F(x) = \begin{cases} 1 - \exp(-x^2), \text{ если } x \geqslant 0 \\ 0, \text{ иначе.} \end{cases}$$

   а) Какую сумму должен выделить кот Матроскин дяде Фёдору, чтобы её достоверно хватало на покупку сметаны?

   б) Какую сумму должен выделить кот Матроскин дяде Фёдору, чтобы Дядя Фёдор с вероятностью 0.9 мог оплатить продукты без сметаны?

   в) Найдите математическое ожидание стоимости продуктов без сметаны;

   г) Найдите математическое ожидание стоимости всего списка.

   д) Какова вероятность того, что общие расходы будут в точности равны их математическому ожиданию?

   Подсказка: $\int_0^\infty \exp(-x^2)\,dx = \sqrt{\pi}/2$.

5. Эксперт с помощью детектора лжи пытается определить, говорит ли подозреваемый правду. Если подозреваемый говорит правду, то эксперт ошибочно выявляет ложь с вероятностью 0.1. Если подозреваемый обманывает, то эксперт выявляет ложь с вероятностью 0.95.

   В деле об одиночном нападении подозревают десять человек, один из которых виновен и будет лгать, остальные невиновны и говорят правду.

   а) Какова вероятность того, что детектор покажет, что конкретный подозреваемый лжёт?

   б) Какова вероятность того, что подозреваемый невиновен, если детектор показал, что он лжёт?

   в) Какова вероятность того, что эксперт точно выявит преступника?

   г) Какова вероятность того, что эксперт ошибочно выявит преступника, то есть покажет, что лжёт невиновный, а все остальные говорят правду?

2.2. 2016-2017

1. Из семей, имеющих двоих разновозрастных детей, случайным образом выбирается одна семья. Известно, что в семье есть девочка (событие A).

   а) Какова вероятность того, что в семье есть мальчик (событие B)?

   б) Сформулируйте определение независимости событий и проверьте, являются ли события A и B независимыми?

2. Система состоит из N независимых узлов. При выходе из строя хотя бы одного узла, система дает сбой. Вероятность выхода из строя любого из узлов равна 0.000001. Вычислите максимально возможное число узлов системы, при котором вероятность её сбоя не превышает 0.01.

3. Исследование состояния здоровья населения в шахтерском регионе «Велико-кротовск» за пятилетний период показало, что из всех людей с диагностированным заболеванием легких, 22% работало на шахтах. Из тех, у кого не было диагностировано заболевание легких, только 14% работало на шахтах. Заболевание легких было диагностировано у 4% населения региона.

   а) Какой процент людей среди тех, кто работал в шахте, составляют люди с диагностированным заболеванием легких?

   б) Какой процент людей среди тех, кто НЕ работал в шахте, составляют люди с диагностированным заболеванием легких?

4. Студент Петя выполняет тест (множественного выбора) проставлением ответов наугад. В тесте 17 вопросов, в каждом из которых пять вариантов ответов и только один из них правильный. Оценка по десятибалльной шкале формируется следующим образом:

$$\text{Оценка} = \begin{cases} \text{ЧПО} - 7, & \text{если ЧПО} \in [8;\ 17], \\ 1, & \text{если ЧПО} \in [0;\ 7], \end{cases}$$

   где ЧПО означает число правильных ответов.

   а) Найдите наиболее вероятное число правильных ответов.

   б) Найдите математическое ожидание и дисперсию числа правильных ответов.

   в) Найдите вероятность того, что Петя получит «отлично» (по десятибалльной шкале получит 8, 9 или 10 баллов).

   Студент Вася также выполняет тест проставлением ответов наугад.

   г) Найдите вероятность того, что все ответы Пети и Васи совпадут.

5. Продавец высокотехнологичного оборудования контактирует с одним или двумя потенциальными покупателями в день с вероятностями $1/3$ и $2/3$ соответственно. Каждый контакт заканчивается «ничем» с вероятностью 0.9 и покупкой оборудования на сумму в 50 000 у. е. с вероятностью 0.1. Пусть ξ — случайная величина, означающая объем дневных продаж в у. е.

   а) Вычислите $\mathbb{P}(\xi = 0)$.

   б) Сформулируйте определение функции распределения и постройте функцию распределения случайной величины ξ.

   в) Вычислите математическое ожидание и дисперсию случайной величины ξ.

6. Интервал движения поездов метро фиксирован и равен b минут, т.е. каждый следующий поезд появляется после предыдущего ровно через b минут. Пассажир приходит на станцию в случайный момент времени. Пусть случайная величина ξ, означающая время ожидания поезда, имеет равномерное распределение на отрезке $[0; b]$.

   а) Запишите плотность распределения случайной величины ξ.

   б) Найдите константу b, если известно, что в среднем пассажиру приходится ждать поезда одну минуту, т. е. $\mathbb{E}(\xi) = 1$.

   в) Вычислите дисперсию случайной величины ξ.

   г) Найдите вероятность того, что пассажир будет ждать поезд менее одной минуты.

   д) Найдите квантиль порядка 0.25 распределения случайной величины ξ.

   е) Найдите центральный момент порядка 2017 случайной величины ξ.

   ж) Постройте функцию распределения случайной величины ξ.

   Марья Ивановна из суеверия всегда пропускает два поезда и садится в третий.

   з) Найдите математическое ожидание и дисперсию времени, затрачиваемого Марьей Ивановной на ожидание «своего» поезда.

   Глафира Петровна не садится в поезд, если видит в нем подозрительного человека. Подозрительные люди встречаются в каждом поезде с вероятностью $3/4$.

   и) Найдите вероятность того, что Глафире Петровне придется ждать не менее пяти минут, чтобы уехать со станции.

   к) Найдите математическое ожидание времени ожидания «своего» поезда для Глафиры Петровны.

7. (Бонусная задача) На первом этаже десятиэтажного дома в лифт заходят 9 человек. Найдите математическое ожидание числа остановок лифта, если люди выходят из лифта независимо друг от друга.

3. Контрольная работа 1. ИП

3.1. 2017-2018

Ровно 272 года назад императрица Елизавета повелела завезти во дворцы котов для ловли мышей.

1. В отсутствии кота Леопольда мыши Белый и Серый подкидывают по очереди игральный додекаэдр . Сыр достаётся тому, кто первым выкинет число 6. Начинает подкидывать Белый.

   а) Какова вероятность того, что сыр достанется Белому?

   б) Сколько в среднем бросков продолжается игра?

   в) Какова дисперсия числа бросков?

2. Микки Маус, Белый и Серый решили устроить труэль из любви к мышки Мии. Сначала стреляет Микки, затем Белый, затем Серый, затем снова Микки и так до тех пор, пока в живых не останется только один.

   Прошлые данные говорят о том, что Микки попадает с вероятностью $1/3$, Белый — с вероятностью $2/3$, а Серый стреляет без промаха.

   Найдите оптимальную стратегию каждого мыша.

3. Микки Маус, Белый и Серый пойманый злобным котом Леопольдом до начала труэли. И теперь Леопольд будет играть с ними в странную игру.

   В комнате три закрытых внешне неотличимых коробки: с золотом, серебром и платиной. Общаться после начала игры мыши не могут, но могут заранее договориться о стратегии.

   Правила игры таковы. Кот Леопольд будет заводить мышей в комнату по очереди. Каждый из мышей может открыть две коробки по своему выбору. Перед следующим мышом коробки закрываются.

   Если Микки откроет коробку с золотом, Белый — с серебром, а Серый — с платиной, то они выигрывают. Если хотя бы один из мышей не найдёт свой металл, то Леопольд их съест.

   а) Какова оптимальная стратегия?

   б) Какова вероятность выигрыша при использовании оптимальной стратегии?

4. Накануне войны Жестокий Тиран Мышь очень большой страны издал указ. Отныне за каждого новорождённого мыша-мальчика семья получает денежную премию, но если в семье рождается вторая мышка-девочка, то всю семью убивают. Бедные жители страны запуганы и остро нуждаются в деньгах, поэтому в каждой семье мыши будут появляться до тех пор, пока не родится первая мышка-девочка.

   а) Каким будет среднее число детей в мышиной семье?

   б) Какой будет доля мышей-мальчиков в стране?

   в) Какой будет средняя доля мышей-мальчиков в случайной семье?

   г) Сколько в среднем мышей-мальчиков в случайно выбираемой семье?

5. Вальяжный кот Василий положил на счёт в банке на Гаити один гурд. Сумма на счету растёт непрерывно с постоянной ставкой в течение очень длительного промежутка времени. В случайный момент этого промежутка кот Василий закрывает свой вклад.

   Каков закон распределения первой цифры полученной Василием суммы?

4. Контрольная работа 2

4.1. 2017-2018

Минимум

1. Приведите определение условной вероятности случайного события, формулу Байеса.

2. Сформулируйте определение и свойства функции плотности случайной величины.

3. Сформулируйте определение условного математического ожидания $\mathbb{E}(Y|X = x)$ для совместного дискретного и совместного абсолютно непрерывного распределений.

4. Сформулируйте неравенство Чебышёва и неравенство Маркова.

5. Задана таблица совместного распределения случайных величин X и Y.

	$Y = -1$	$Y = 0$	$Y = 1$
$X = 0$	0.2	0.1	0.3
$X = 1$	0.2	0.1	0.1

   а) Найдите $F_{X,Y}(0, 0)$;
   б) Найдите $\mathbb{E}(X)$, $\mathbb{E}(X^2)$, $\mathbb{E}(Y)$, $\mathbb{E}(Y^2)$;
   в) Найдите $\text{Var}(X)$, $\text{Var}(Y)$;
   г) Найдите $\text{Cov}(X, Y)$, $\text{Corr}(X, Y)$

6. Плотность распределения случайного вектора (X, Y) имеет вид

$$f_{X,Y}(x, y) = \begin{cases} \frac{4x+10y}{7}, & \text{при } (x, y) \in [0; 1] \times [0; 1] \\ 0, & \text{при } (x, y) \notin [0; 1] \times [0; 1] \end{cases}$$

   а) Найдите $\mathbb{P}(X \leqslant Y)$;
   б) Найдите функцию плотности $f_X(x)$;
   в) Найдите $\mathbb{E}(X)$, $\mathbb{E}(Y)$ и $\text{Cov}(X, Y)$;
   г) Являются ли случайные величины X и Y независимыми?

Задачи

7. Статистика авиакомпании «А» за много лет свидетельствует о том, что 10% людей, купивших билет на самолет, не являются на рейс. Авиакомпания продала 330 билетов на 300 мест.

   а) Какова вероятность, что всем явившимся на рейс пассажирам хватит места?
   б) Укажите наибольшее число билетов, которое можно продавать на 300 мест, чтобы случаи переполнения случались не чаще, чем на одном из десяти рейсов.

8. Сегодня акция компании «Ух» стоит 1 рубль. Каждый день акция может с вероятностью 0.7 вырасти на 1%, с вероятностью 0.2999 упасть на 1% и с вероятностью 0.0001 обесцениться (упасть на 100%).

   а) Считая изменение цены акции независимыми, найдите математическое ожидание её стоимости через 20 торговых дней.

б) Найдите предел по вероятности среднего изменения цены акции в процентах на бесконечном промежутке времени (Ответ обоснуйте).

в) Найдите математическое ожидание цены акции на бесконечном промежутке времени.

г) Инвестор вложил все свои средства в акции компании «Ух». Найдите вероятность его разорения на бесконечном промежутке времени.

5. Контрольная работа 3

5.1. 2017-2018

Минимум

1. Дайте определение выборочной функции распределения.

2. Предположим, что величины X_1, X_2, ..., X_n независимы и нормальны $\mathcal{N}(\mu;\sigma^2)$. Укажите закон распределения выборочного среднего, величины $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$, величины $\frac{\bar{X}-\mu}{\hat{\sigma}/\sqrt{n}}$, величины $\frac{\hat{\sigma}^2(n-1)}{\sigma^2}$.

3. Рост в сантиметрах, случайная величина X, и вес в килограммах, случайная величина Y, взрослого мужчины является нормальным случайным вектором $Z=(X,Y)$ с математическим ожиданием $\mathbb{E}(Z)=(175,75)$ и ковариационной матрицей

$$\mathrm{Var}(Z)=\begin{pmatrix}49 & 28\\ 28 & 36\end{pmatrix}$$

   а) Найдите средний вес мужчины при условии, что его рост составляет 172 см.

   б) Выпишите условную плотность распределения веса мужчины при условии, что его рост составляет 172 см.

   в) Найдите условную вероятность того, что человек будет иметь вес, больший 92 кг, при условии, что его рост составляет 172 см.

4. Стоимость выборочного исследования генеральной совокупности, состоящей из трех страт, определяется по формуле $TC=c_1n_1+c_2n_2+c_3n_3$, где c_i — цена одного наблюдения в i-ой страте, а n_i — число наблюдений, которые приходятся на i-ую страту. Найдите n_1, n_2 и n_3, при которых дисперсия стратифицированного среднего достигает наименьшего значения, если бюджет исследования 8000 и имеется следующая информация:

Страта	1	2	3
Среднее значение	30	40	50
Стандартная ошибка	5	10	20
Вес	25%	25%	50%
Цена наблюдения	2	5	8

Задачи

5. Пусть $X_1,\ldots,X_n$ выборка из нормального распределения $N(\mu,1)$.

   а) Выпишите функцию правдоподобия;

   б) Методом максимального правдоподобия найдите оценку $\hat{\mu}$ математического ожидания μ;

   в) Проверьте состоятельность и несмещённость оценки $\hat{\mu}$;

   г) Вычислите информацию Фишера о параметре μ, содержащуюся во всей выборке;

   д) Для произвольной несмещённой оценки μ выпишите неравенство Рао-Крамера-Фреше;

   е) Проверьте свойство эффективности оценки $\hat{\mu}$;

   ж) Найдите оценку максимального правдоподобия $\hat{\theta}$ для второго начального момента;

   з) Проверьте свойства несмещенности и асимптотической несмещенности оценки $\hat{\theta}$;

и) С помощью дельта-метода вычислите, примерно, дисперсию оценки $\hat{\theta}$;

к) Проверьте состоятельность оценки $\hat{\theta}$.

6. Пусть X_1, ..., X_n выборка из распределения с функцией плотности:

$$f(x) = \begin{cases} \frac{2}{\theta^2}(\theta - x), & \text{при } x \in [0, \theta] \\ 0, & \text{при } x \notin [0, \theta] \end{cases}$$

а) Методом моментов найдите оценку параметра θ;

б) Приведите определение состоятельности оценки и проверьте, будет ли найденная оценка состоятельной.

7. В прихожей лежат четыре карты «тройка». На двух из них нет денег, на двух других 30 и 500 рублей. Вовочка не помнит, на какой из карт есть деньги, поэтому берет три карточки.

а) Найдите математическое ожидание и дисперсию средней по выбранным карточкам суммы денег;

б) Определите, какова вероятность того, что Вовочке удастся войти в метро, если стоимость проезда по тройке составляет 35 рублей.

8. По выборочному опросу студенческих семейных пар о расходах на ланч были получены следующие результаты:

Номер семьи	1	2	3	4
Расходы мужа	450	370	170	200
Расходы жены	210	350	250	180

Считая, что разница в расходах мужа и жены хорошо описываются нормальным распределением, постройте 95%-ый доверительный интервал для разницы математических ожиданий расходов супругов. Есть ли основания утверждать, что расходы одинаковы?

9. Наблюдатель Алексей Недопускальный решил проверить честность выборов. Ему удалось подглядеть, как проголосовали 60 избирателей. Из них 42 выбрали действующего президента.

а) Постройте 95%-ый доверительный интервал для истинной доли избирателей, проголосовавших «за» действующего президента.

б) По результатам ЦентрИзберКома «за» действующего президента проголосовало 76.67% населения. Согласуются ли эти данные с данными Алексея?

в) Сколько бюллетеней нужно подглядеть Алексею, чтобы с вероятностью 0.95 отклонение от выборочной доли проголосовавших «за» действующего президента от истинной не превышало 0.01?

6. Контрольная работа 3. ИП

6.1. 2017-2018

дата: 2018-03-24
24 марта 2018 года — Комоедица, день пробуждения медведя.

1. Медведь Михайло-Потапыч уснул в берлоге и ему снится сон про n-мерное пространство. Особенно ярко ему снится вектор $X = (X_1, X_2, \ldots, X_n)$ и вектор $e = (1, 1, 1, \ldots, 1)$.

   а) Изобразите векторы X и e в n-мерном пространстве;

   б) Изобразите проекцию X на $\mathrm{Lin}\{e\}$, обозначим её $\hat{X}$;

   в) Изобразите проекцию X на $\mathrm{Lin}^{\perp}\{e\}$, обозначим её $\hat{X}^{\perp}$;

   г) Выпишите явно вектора $\hat{X}$ и $\hat{X}^{\perp}$, и найдите их длины;

   д) Сформулируйте теорему Пифагора для нарисованного прямоугольного треугольника;

   е) Изобразите на рисунке такой угол α, что обычная t-статистика, используемая при построении доверительного интервала для μ, имела бы вид $t = \sqrt{n-1} \cdot \operatorname{ctg} \alpha$.

2. Исследователь Михаил предполагает, что все виды медведепришельцев встречаются равновероятно Отправившись на охоту в район Малой Медведицы Михаил поймал двух лиловых кальмаромедведе одного двурога медведеспинного и одного медведезавра ящероголового.

   Помогите Михаилу оценитель общее количество видов медведепришельцев с помощью метода максимального правдоподобия.

3. Помотавшись по просторам Вселенной Михаил изменил своё мнение. Никто кроме кальмаромедвед двурогов и медведезавров не попадается, однако попадаются они явно с разной вероятностью.
   Из 300 отловленных пришельцев оказалось 150 кальмаромедведей, 100 двурогов и 50 медведезавров.
   Михаил считает, что медведепришельцы встречаются независимо, p_1 — вероятность встретить кальмаромедведя, p_2 — двурога.

   а) Оцените вектор $p = (p_1, p_2)$ методом максимального правдоподобия;

   б) Оцените ковариационную матрицу $\mathrm{Var}(\hat{p})$;

   в) Оцените дисперсию $\mathrm{Var}(\hat{p}_1 - \hat{p}_2)$;

   г) Постройте доверительный интервал для разницы долей $p_1 - p_2$.

4. Винни-Пух лично измерил количество мёда (в кг) на 100 деревьях и обнаружил, что $\bar{X} = 10$ и $\hat{\sigma}^2 = 4$. По мнению Кролика, состоятельная оценка для параметра α правильности мёда имеет вид $\hat{\alpha} = \bar{X} + \sqrt{\bar{X} + 6}$.

   а) «Халява, сэр!» Найдите точечную оценку параметра α;

   б) Найдите 95%-ый доверительный интервал для α, симметричный относительно $\hat{\alpha}$.

5. Фотографы Андрей и Белла независимо друг от друга пытаются фотографировать кадьяков. Андрею удаётся сфотографировать одного кадьяка в неделю с вероятностью 0.5, а Белле — с вероятностью p, независимо друг от друга и от прошлого. За 100 недель они вместе сфотографировали 130 кадьяков.

   а) Оцените p и постройте 95%-ый доверительный интервал для p;

   б) Оцените p и постройте 95%-ый доверительный интервал для p, если дополнительно известно, что один фотограф опередил другого на 10 фото.

Просто красивая задачка. Эту задачу **не нужно** решать на кр :)
Медведю Мишутке никак не удаётся заснуть в берлоге, и потому он подбрасывает правильную монетку n раз. Обозначим вероятность того, что ни разу не идёт двух решек подряд буквой q_n.

а) Найдите $2^8 q_8$ и **назовите** это число;

б) Найдите $\lim 2q_{n+1}/q_n$ и **назовите** это число.

7. Контрольная работа 4

7.1. 2017-2018

Минимум

1. Пусть $X = (X_1, \ldots, X_n)$ — случайная выборка из нормального распределения с неизвестным математическим ожиданием μ и неизвестной дисперсией $\sigma^2 = 9$. Объем выборки $n = 20$. Для тестирования основной гипотезы $H_0 : \mu = 0$ против альтернативной гипотезы $H_1 : \mu = 5$ вы используете критерий: если $\overline{X} \leqslant 2$, то вы не отвергаете гипотезу H_0, в противном случае вы отвергаете гипотезу H_0 в пользу гипотезы H_1. Найдите

   а) Вероятность ошибки 1-го рода

   б) Вероятносто ошибки 2-го рода

2. Пусть $X = (X_1, \ldots, X_n)$ — случайная выборка из нормального распределения с неизвестными параметрами μ и σ^2. Используя реализацию случайной выборки: $x_1 = 2$, $x_3 = 8$, $x_3 = 5$, постройте 95%-ый доверительный интервал для неизвестного параметра σ^2.

3. Вася очень любит тестировать статистические гипотезы. В этот раз Вася собирается проверить утверждение о том, что его друг Пётр звонит Васе исключительно в то время, когда Вася ест. Для этого Вася трудится целый год и проводит серию из 365 испытаний. Результаты приведены в табилце ниже.

	Пётр не звонит	Пётр звонит
Вася ест	100	50
Вася не ест	125	90

   На уровне значимости 10% протестируйте гипотезу о том, что Пётр звонит Васе независимо от момента приема пищи.

4. Вася Сидоров утверждает, что ходит в кино в четыре раза чаще, чем в спортзал, а в спортзал в четыре раза чаще, чем в театр. За последние полгода он 105 раз был в театре, 63 раза - в спортазе и 42 раза в кино. На уровне значимости 10% проверьте утверждение Васи.

Квантили χ^2 распределения с 1, 2 и 3 степенями свободы

	0.025	0.05	0.1	0.9	0.95	0.975
1	0.001	0.004	0.016	2.706	3.841	5.024
2	0.051	0.103	0.211	4.605	5.991	7.378
3	0.216	0.352	0.584	6.251	7.815	9.348

Задачи

При решении задач пять–семь используйте данные обследования Росстата за первый квартал 2018 года:

	Число наблюдений	Среднее (тыс. руб.)	Выборочное отклонение (тыс. руб.)
Врачи	40	136	55
Преподаватели	60	139	60

Распределение заработной платы работников любой отрасли хорошо описывается нормальным законом.

5. На уровне значимости 5% проверьте гипотезу о том, что средняя зарплата врача составляет 100 т.р., против альтернативы, что она больше 100 т.р. Вычислите минимальный уровень значимости, при котором основная гипотеза отвергается (P–значение).

6. На уровне значимости 10% проверьте гипотезу о том, что разброс в зарплатах врачей и преподавателей одинаков, против двухсторонней альтернативы.

7. На уровне значимости 10% проверьте гипотезу о том, что средняя зарплата врачей и преподавателей совпадают, против альтернативы, что у преподавателей зарплата выше:

   а) Считая объемы выборок достаточно большими

   б) Считая дисперсии одинаковыми

8. Время в часах безотказной работы микронаушника, величина X, подчиняется экспоненциальному (показательному) закону распределения с неизвестным параметром λ:

$$f(x;\lambda) = \begin{cases} \lambda e^{-\lambda x}, \text{ при } x \geqslant 0 \\ 0, \text{ при } x < 0 \end{cases}$$

   По выборке из 100 независимых наблюдений $\bar{x} = 0.52$. С помощью асимптотических свойств оценок максимального правдоподобия постройте приближенный 95%-ый доверительный интервал:

   а) Для параметра λ

   б) Для вероятности того, что наушник проработает без сбоев весь тест — 45 минут

9. Приглашенный на Петербургский международный экономический форум Германом Грефом индийский мистик Садхгуру подарил Грефу древнюю шестигранную кость для принятия решений в сложных макроэкономических ситуациях. Служба безопасности Сбербанка провела серию из 100 испытаний и составила таблицу:

Грань	1	2	3	4	5	6
Число выпадений	10	10	15	15	25	25

   С помощью теста отношения правдоподобия на уровне значимости 5% проверьте гипотезу о том, что все грани равновероятны.

$$\ln(1/6) = -1.79, \ \ln(0.15) = -1.90, \ \ln(0.25) = -1.39, \ \ln(0.1) = -2.30$$

8. Контрольная работа 4. ИП

8.1. 2017-2018

Напутствие в добрый путь:

1. Работа сдаётся только в виде запроса pull-request на гитхаб-репозиторий.

2. Имя файла должно быть вида ivanov_ivan_161_kr_4.Rmd.

3. Также фамилию и имя нужно указать в шапке документа в поле author :)

4. Если нужно, то установите пакеты tidyverse, maxLik, nycflights13.

1. Симулируем бурную деятельность! В качестве параметра k в задаче используй число букв в своей фамилии в именительном падеже :)

   Каждый день Василий съедает случайное количество булочек, которое распределено по Пуассону с параметром 10. Логарифм затрат в рублях на каждую булочку распределён нормально $N(2, 1)$. Андрей каждый день съедает биномиальное количество булочек $Bin(2k, 0.5)$. Затраты Андрей на каждую булочку распределены равномерно на отрезке $[2; 20]$.

   а) Сколько в среднем тратит Василий на булочки за день?

   б) Чему равна дисперсия дневных расходов Василия?

   в) Какова вероятность того, что за один день Василий потратит больше денег, чем Андрей?

   г) Какова условная вероятность того, что Василий за день съел больше булочек, чем Андрей, если известно, что Василий потратил больше денег?

2. Сражаемся с реальностью! В пакете nycflights13 встроен набор данных weather о погоде в разные дни в разных аэропортах.

   а) Постройте гистограмму переменной влажность, humid. У графика подпишите оси!

   б) Постройте диаграмму рассеяния переменных влажность и количество осадков, precip. У графика подпишите оси!

      Посчитайте выборочное среднее и выборочную дисперсию влажности и количества осадков.

   в) С помощью максимального правдоподобия оцените параметр μ, предположив, что наблюдения за влажностью имеют нормальное $N(\mu, 370)$-распределение и независимы. Постройте 95%-ый доверительный интервал для μ.

   г) С помощью максимального правдоподобия оцените параметр σ^2, предположив, что наблюдения за влажностью имеют нормальное $N(60, \sigma^2)$-распределение и независимы. Постройте 95%-ый доверительный интервал для σ^2.

      Если при численной оптимизации параметр σ^2 становится отрицательным, можно задать параметры по-другому, например, $\sigma^2 = \exp(\gamma)$.

9. Финальные экзамены

9.1. 2017-2018

Вопрос 1. Дана случайная выборка из двух наблюдений, X_1 и X_2. Несмещённой и наиболее эффективной оценкой математического ожидания из предложенных является

- [A] $\frac{20}{20}X_1 + \frac{20}{20}X_2$
- [B] $\frac{8}{20}X_1 + \frac{12}{20}X_2$
- [C] $\frac{10}{20}X_1 + \frac{10}{20}X_2$
- [D] $\frac{1}{20}X_1 + \frac{19}{20}X_2$
- [E] $\frac{5}{20}X_1 + \frac{15}{20}X_2$

Вопрос 2. Величины X_1 и X_2 независимы и одинаково распределены. Оценка $\hat{\mu} = 3aX_1 + 4a^2X_2$ математического ожидания $\mu = \mathbb{E}(X_i)$ будет несмещённой при a равном

- [A] 1.2
- [B] 3
- [C] -1
- [D] 0
- [E] -3

Вопрос 3. Величины X_1, ..., X_5 представляют собой случайную выборку. Несмещённой оценкой дисперсии X_i является

- [A] $0.25\sum_{i=1}^{4}(X_i - \bar{X})^2$
- [B] $0.5\sum_{i=1}^{5}(X_i - \bar{X})^2$
- [C] $0.2\sum_{i=1}^{4}(X_i - \bar{X})^2$
- [D] $0.25\sum_{i=1}^{5}(X_i - \bar{X})^2$
- [E] $0.2\sum_{i=1}^{5}(X_i - \bar{X})^2$

Вопрос 4. Последовательность оценок $\hat{a}_n$ параметра a является состоятельной, если

- [A] $\mathbb{E}((\mathbb{E}(\hat{a}_n) - a)^2) \xrightarrow{P} \hat{a}_n$
- [B] $\mathbb{E}((\hat{a}_n - a)^2) \xrightarrow{P} \hat{a}_n$
- [C] $\hat{a}_n \xrightarrow{P} a$
- [D] $\mathbb{E}((\hat{a}_n - a)^2) \xrightarrow{P} 0$
- [E] $\mathbb{E}((\hat{a}_n - a)^2) \xrightarrow{P} a$

Вопрос 5. Оценка $\hat{a}$ называется эффективной оценкой параметра a в классе оценок K, если

- [A] $\mathbb{E}(\hat{a}^2) \geqslant \mathbb{E}(\tilde{a}^2)$ для всех $\tilde{a} \in K$
- [B] $\mathbb{E}((\hat{a} - a)^2) \leqslant \mathbb{E}((\tilde{a} - a)^2)$ для всех $\tilde{a} \in K$
- [C] $\mathbb{E}((\hat{a} - \tilde{a})^2) \leqslant \mathbb{E}((\tilde{a} - a)^2)$ для всех $\tilde{a} \in K$
- [D] $\mathbb{E}((\hat{a} - a)^2) \geqslant \mathbb{E}((\tilde{a} - a)^2)$ для всех $\tilde{a} \in K$
- [E] $\mathbb{E}((\hat{a} - \tilde{a})^2) \geqslant \mathbb{E}((\tilde{a} - a)^2)$ для всех $\tilde{a} \in K$

Вопрос 6. Апостериорная функция плотности пропорциональна

- [A] Отношению функции правдоподобия к априорной плотности
- [B] Отношению априорной плотности к функции правдоподобия
- [C] Разности априорной плотности и правдоподобия
- [D] Произведению априорной плотности и правдоподобия
- [E] Сумме априорной плотности и правдоподобия

Вопрос 7. Алгоритм Метрополиса-Гастингса порождает

- [A] Независимую выборку из апостериорного закона распределения
- [B] Независимую выборку из смеси априорного и апостериорного законов
- [C] Независимую выборку из априорного закона распределения
- [D] Зависимую выборку из априорного закона распределения
- [E] Зависимую выборку из апостериорного закона распределения

Вопрос 8. В алгоритме Метрополиса-Гастингса был предложен переход из точки $\theta^{(0)} = 4$ в точку $\theta^{(1)}_{prop} = 5$. Априорное распределение θ равномерное. Известны значения функций правдоподобия, $f(data|\theta = 4) = 0.7$, $f(data|\theta = 5) = 0.8$. Вероятность одобрения перехода равна

$\boxed{A}$ $0.8/5$

$\boxed{B}$ 1

$\boxed{C}$ $28/40$

$\boxed{D}$ $4/5$

$\boxed{E}$ $7/8$

Вопрос 9. Есть выборка $X_1, X_2, ..., X_5$ и выборка Y_1, Y_2, Y_3, Y_4. Исследовательница Ирина проводит тест суммы рангов Вилкоксона. У выборки X_i сумма рангов равна 7. Сумма рангов для выборки Y_j равна

$\boxed{A}$ 2

$\boxed{B}$ 1

$\boxed{C}$ 38

$\boxed{D}$ 43

$\boxed{E}$ 45

Вопрос 10. Вероятность того, что в случайной выборке три наблюдения подряд попадут в верхний теоретический квартиль равна

$\boxed{A}$ 0.01

$\boxed{B}$ 0.05

$\boxed{C}$ $1/2$

$\boxed{D}$ $1/4$

$\boxed{E}$ $1/64$

Вопрос 11. Величины $X_1, \ldots, X_n$ — случайная выборка с распределением

x	-4	0	3
$\mathbb{P}(X_i = x)$	$3/4 - \theta$	$1/4$	θ

Оценка неизвестного параметра θ, найденная с помощью первого начального момента, равна

$\boxed{A}$ $\frac{\bar{X}+7}{3}$

$\boxed{B}$ $\frac{\frac{1}{n}\sum_{i=1}^{n} X_i^2 - 12}{7}$

$\boxed{C}$ $\frac{12 - \frac{1}{n}\sum_{i=1}^{n} X_i^2}{7}$

$\boxed{D}$ $\frac{7 - \frac{1}{n}\sum_{i=1}^{n} X_i^2}{12}$

$\boxed{E}$ $\frac{\bar{X}+3}{7}$

Вопрос 12. Случайная выборка состоит из одного наблюдения X_1, которое имеет плотность распределени

$$f(x;\, \theta) = \begin{cases} \frac{1}{\theta} x^{-1+\frac{1}{\theta}} & \text{при } x \in (0;\, 1), \\ 0 & \text{при } x \notin (0;\, 1). \end{cases}$$

Оценка параметра θ, найденная с помощью метода максимального правдоподобия, равна

$\boxed{A}$ X_1

$\boxed{B}$ $\ln X_1$

$\boxed{C}$ $-\ln X_1$

$\boxed{D}$ $-X_1$

$\boxed{E}$ $\frac{1}{\ln X_1}$

Вопрос 13. Величины $X_1, \ldots, X_n$ — случайная выборка из распределения Бернулли с параметром $p \in (0;\, 1)$. Оценка максимального правдоподобия параметра p равна $\bar{X}$. Оценка максимального правдоподобия для $\sqrt{p}$ равна

$\boxed{A}$ $\frac{\sqrt{\sum_{i=1}^{n} X_i}}{n}$

$\boxed{B}$ $\sqrt{\sum_{i=1}^{n} X_i}$

$\boxed{C}$ $\frac{1}{n}\sum_{i=1}^{n} X_i$

$\boxed{D}$ $\frac{1}{n}\sum_{i=1}^{n} \sqrt{X_i}$

$\boxed{E}$ $\sqrt{\frac{1}{n}\sum_{i=1}^{n} X_i}$

Вопрос 14. Величины $X_1, \ldots, X_n$ — случайная выборка из распределения Бернулли с параметром $p \in (0;\, 1)$. Информация Фишера о параметре p, заключенная в одном наблюдении, равна

$\boxed{A}$ $\frac{1}{p(1-p)}$

$\boxed{B}$ $\frac{1}{p}$

$\boxed{C}$ $p(1-p)$

$\boxed{D}$ $1-p$

$\boxed{E}$ p

Вопрос 15. Известно истинное значение параметра, $\theta = 1$, и информация Фишера о параметре θ, заключенная в одном наблюдении случайной выборки, $I_1(\theta) = 8$. Оценка максимального правдоподобия $\hat{\theta}$ параметра θ, найденная по ста наблюдениям случайной выборки, имеет распределение, похожее на

$\boxed{A}$ $\mathcal{N}(1,\ 1/800)$

$\boxed{B}$ $\mathcal{N}(1,\ 1/\sqrt{800})$

$\boxed{C}$ $\mathcal{N}(1,\ 1/\sqrt{8})$

$\boxed{D}$ $\mathcal{N}(1,\ 1/8)$

$\boxed{E}$ $\mathcal{N}(1,\ 8)$

Вопрос 16. Дана реализация выборки: 1, 2, 0. Выборочный начальный момент второго порядка равен

$\boxed{A}$ 2.5

$\boxed{B}$ 5/3

$\boxed{C}$ 1/3

$\boxed{D}$ 1

$\boxed{E}$ 3

Вопрос 17. Математическое ожидание выборочного среднего, построенного по выборке из равномерного распределения на отрезке $[0, 2]$, равно

$\boxed{A}$ $1/\sqrt{n}$

$\boxed{B}$ 1

$\boxed{C}$ 0

$\boxed{D}$ 2

$\boxed{E}$ 1.5

Вопрос 18. Дана реализация выборки: 3, 2, 5, 4, 2. Выборочная функция распределения в точке $x = 2.5$ принимает значение

$\boxed{A}$ 0.25

$\boxed{B}$ 0.4

$\boxed{C}$ 0.5

$\boxed{D}$ 0.6

$\boxed{E}$ 0.2

Вопрос 19. Экзамен принимают два преподавателя: Злой и Добрый. Злой поставил оценки 2, 3, 10, 8, 1. А Добрый — оценки 6, 4, 7, 9. Значение статистики критерия Вилкоксона о совпадении распределений оценок может быть равно

$\boxed{A}$ 25

$\boxed{B}$ 23

$\boxed{C}$ 26

$\boxed{D}$ 22

$\boxed{E}$ 24

Вопрос 20. Датчик случайных чисел выдал два значения псевдослучайных чисел 0.1 и 0.8. Вычислите значение критерия Колмогорова и проверьте гипотезу о соответствии распределения равномерному на $(0, 1)$. Критическое значение статистики Колмогорова считайте равным 0.776.

$\boxed{A}$ 0.1, H_0 отвергается

$\boxed{B}$ 0.3, H_0 не отвергается

$\boxed{C}$ 0.8, H_0 отвергается

$\boxed{D}$ 0.2, H_0 не отвергается

$\boxed{E}$ 0.4, H_0 не отвергается

Вопрос 21. Случайные величины $X_1, \ldots, X_m$ — случайная выборка из нормального распределения. Величины $Y_1, \ldots, Y_n$ — независимая случайная выборка из нормального распределения. Для построения доверительного интервала для отношения дисперсий можно использовать статистику с распределением

- A $F_{m,n-2}$
- B $F_{m-1,n-1}$
- C χ^2_{m+n-2}
- D $F_{m+1,n+1}$
- E t_{m+n-2}

Вопрос 22. При построения доверительного интервала для разности математических ожиданий в двух нормальных выборках размером m и n в случае равных неизвестных дисперсий используется распределение

- A $\mathcal{N}(0, m+n-2)$
- B $F_{m,n}$
- C t_{m+n}
- D $F_{m-1,n-1}$
- E t_{m+n-2}

Вопрос 23. Для построения доверительного интервала для математического ожидания используется выборка из 100 наблюдений. Выборочное среднее равно 5. Дисперсия генеральной совокупности известна и равна 25. Минимальная длина 95%-доверительного интервала примерно равна

- A 5
- B 2.5
- C 0.98
- D 1.96
- E 10

Вопрос 24. При построении 90%-доверительного интервала для вероятности используется выборка из 25 наблюдений. Выборочная доля составляет 0.6. В симметричный доверительный интервал попадают значения

- A 0.6, 0.7, 0.85
- B 0.5, 0.6, 0.65
- C 0.35, 0.5, 0.65
- D 0.8, 0.9, 1.0
- E 0.7, 0.8, 0.9

Вопрос 25. При построении 90%-доверительного интервала для дисперсии используется выборка из 26 наблюдений. Несмещенная оценка дисперсии равна 100. Левая граница симметричного по вероятности доверительного интервала равна

- A 43.25
- B 106.32
- C 8.16
- D 32.8
- E 66.4

Вопрос 26. При проверке гипотезы о равенстве математических ожиданий оценок по математической статистике в двух группах, было получено P-значение 0.03. Тогда нулевая гипотеза

- A отвергается на уровне значимости 0.05 и на уровне значимости 0.01
- B не отвергается на уровне значимости 0.05 и отвергается на уровне значимости 0.01
- C не отвергается ни на уровне значимости 0.05, ни на уровне значимости 0.01
- D отвергается на уровне значимости 0.05 и не отвергается на уровне значимости 0.01
- E ответ зависит от альтернативной гипотезы

Вопрос 27. Имеются две случайных выборки $X_1, \ldots, X_{31}$ и $Y_1, \ldots, Y_{41}$ из нормальных распределений. Известно, что $\sum_{i=1}^{31}(X_i - \bar{X})^2 = 120$ и $\sum_{i=1}^{41}(Y_i - \bar{Y})^2 = 400$. При проверке гипотезы о равенстве дисперсий этих распределений значение тестовой статистики может быть равно

- A 0.3
- B 2.52
- C 2
- D 3.33
- E 2.5

Вопрос 28. Имеется выборка из одного наблюдения X_1. На основе этой выборки тестируется гипотеза H_0: $X_1 \sim U[0;2]$ против альтернативной гипотезы $X_1 \sim U[1,3]$. Используется критерий следующего вида: если $X_1 > a$, то H_0 отвергается. Минимальная вероятность ошибки первого рода достигается при a равном

$\boxed{A}$ 0.5 $\boxed{C}$ 2 $\boxed{E}$ 1

$\boxed{B}$ 1.5 $\boxed{D}$ 1.9

Вопрос 29. Исследовательница Алевтина подбросила кубик 12 раз и 12 раз на нём выпала шестёрка. Алевтина хочет проверить, выпадают ли все грани равновероятно, при помощи критерия χ^2 Пирсона. Значение тестовой статистики будет равно

$\boxed{A}$ 60 $\boxed{C}$ 5 $\boxed{E}$ 6

$\boxed{B}$ 12 $\boxed{D}$ 50

Вопрос 30. Исследовательница Глафира считает, что любовь к энергетическим напиткам и успешность сдачи экзамена по математической статистике должны быть как-то связаны. Опросив 200 своих однокурсников, она получила следующие результаты:

	пьёт энергетик	не пьёт энергетик
Успешно сдал	20	120
Завалил	40	20

Статистика χ^2 Пирсона для проверки независимости признаков с округлением до целых равна

$\boxed{A}$ 70 $\boxed{C}$ 65 $\boxed{E}$ 45

$\boxed{B}$ 55 $\boxed{D}$ 35

10. Ответы

Минимумы

Контрольная работа 1 — Задачный минимум

1. а) 0.25
   б) 0.6
   в) нет
2. а) 0.5
   б) 0.75
   в) нет
3. $\frac{4}{10\cdot 11\cdot 12\cdot 13}$
4. $\frac{4}{10\cdot 11\cdot 12\cdot 13}$
5. 0.5
6. 0.42
7. 0.028
8. $\frac{5}{7}$
9. а) 0.5
   б) 0.75
   в) 0
   г) 0.5
10. а) 0.5
    б) 0
    в) 0.5
    г) 0.5
    д) 0.5
11. а) 0.25
    б) 0.75
    в) 0
    г) 0.5
12. а) 0.25
    б) 0.25
    в) 0.75
    г) 0.5
    д) 0.75
13. а) $\left(\frac{1}{4}\right)^4$

б) $1-\left(\frac{1}{4}\right)^4$

в) 0

г) 3

д) 0.75

е) $2, 3$

14. а) $\left(\frac{3}{5}\right)^5$

б) $1-\left(\frac{3}{5}\right)^5$

в) 0

г) 2

д) 1.2

е) 2

15. а) e^{-100}

б) $1-e^{-100}$

в) 0

г) 100

д) 100

16. а) e^{-101}

б) $1-e^{-101}$

в) 0

г) 101

д) 101

17. $1-\frac{8^5}{9^5}$

18. $\frac{8^5}{9^5}$

19. $1-e^{-3}$

20. e^{-3}

21. а) 0.5

б) 0.25

в) 0.125

г) 1

22. а) 0.5

б) 0.5

в) $\frac{1}{3}$

г) $\frac{1}{12}$

д) 1

23. а) 2

б) 0.25

в) $\frac{3}{4}$

г) 1

24. а) 2

    б) 0.5

    в) 0.5

    г) 0

    д) 0.8

Контрольная работа 2 — Задачный минимум

1. а) 0.5

   б) 0.3

   в) 0.2

   г) нет

   д) 0.3

   е)

X	-1	1
$\mathbb{P}(\cdot)$	0.5	0.5

   ж) $F_X(x) = \begin{cases} 0, & \text{при } x < -1 \\ 0.5, & \text{при } x \in [-1; 1) \\ 1, & \text{при } x \geqslant 1 \end{cases}$

2. а) 0.5

   б) 0.4

   в) 0.2

   г) да

   д) 0.6

   е)

Y	-1	0	1
$\mathbb{P}(\cdot)$	0.4	0.2	0.4

   ж) $F_Y(y) = \begin{cases} 0, & \text{при } y < -1 \\ 0.4, & \text{при } y \in [-1; 0) \\ 0.6, & \text{при } y \in [0; 1) \\ 1, & \text{при } y \geqslant 1 \end{cases}$

3. а) 0

   б) 1

   в) 1

   г) 0

   д) 0.6

   е) 0.6

ж) 0

з) 0

и) 0

к) да, являются некоррелированными, но нельзя утверждать, что являются независимыми

4. а) 0

б) 1

в) 1

г) 0

д) 0.8

е) 0.8

ж) 0

з) 0

и) 0

к) да, являются некоррелированными, но нельзя утверждать, что являются независимыми

5. а) 0.25

б) 0.2

в)

$Y \mid \{X=-1\}$	-1	0	1
$\mathbb{P}(\cdot)$	0.4	0.2	0.4

г) 0

д) 0.8

6. а) 0.5

б) 0.2

в)

$Y \mid \{X=1\}$	-1	0	1
$\mathbb{P}(\cdot)$	0.4	0.2	0.4

г) 0

д) 0.8

7. а) 0

б) 36

в) 9

г) 60

д) -4

е) $\frac{-1}{3\sqrt{5}}$

ж) $\begin{pmatrix} 3 & -1 \\ -1 & 4 \end{pmatrix}$

8. а) -4

б) 8

в) 1

г) 10

д) -6

е) $\frac{-1}{\sqrt{5}}$

ж) $\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$

9. а) 0.3413

   б) 0.0228

   в) 0.1915

10. а) 0.6826

    б) 0.0228

    в) 0.1574

11. 0.4332

12. 0.8185

13. 0.4514

14. 0.5328

15. ≈ 0.8185

16. ≈ 0.9115

17. ≈ 0.6422

18. ≈ 0.9606

19. а) 0.125

    б) 0.5

    в) $f_X(x) = \begin{cases} x + \frac{1}{2}, & \text{при } x \in [0;1] \\ 0, & \text{при } x \notin [0;1] \end{cases}$

    г) $f_Y(y) = \begin{cases} y + \frac{1}{2}, & \text{при } y \in [0;1] \\ 0, & \text{при } y \notin [0;1] \end{cases}$

    д) нет

20. а) $\frac{1}{16}$

    б) $\frac{1}{2}$

    в) $f_X(x) = \begin{cases} 2x, & \text{при } x \in [0;1] \\ 0, & \text{при } x \notin [0;1] \end{cases}$

    г) $f_Y(y) = \begin{cases} 2y, & \text{при } y \in [0;1] \\ 0, & \text{при } y \notin [0;1] \end{cases}$

    д) да

21. а) $\frac{7}{12}$

б) $\frac{7}{12}$

в) $\frac{1}{3}$

г) $-\frac{1}{144}$

д) $-\frac{1}{11}$

22. а) $\frac{2}{3}$

б) $\frac{2}{3}$

в) $\frac{4}{9}$

г) 0

д) 0

23. а) $f_Y(y) = \begin{cases} y + \frac{1}{2}, & \text{при } y \in [0;1] \\ 0, & \text{при } y \notin [0;1] \end{cases}$

б) $f_{X|Y}(x|\frac{1}{2}) = \begin{cases} x + \frac{1}{2}, & \text{при } x \in [0;1] \\ 0, & \text{при } x \notin [0;1] \end{cases}$

в) $\frac{7}{12}$

г) $\frac{11}{144}$

24. а) $f_Y(y) = \begin{cases} 2y, & \text{при } y \in [0;1] \\ 0, & \text{при } y \notin [0;1] \end{cases}$

б) $f_{X|Y}(x|\frac{1}{2}) = \begin{cases} 2x, & \text{при } x \in [0;1] \\ 0, & \text{при } x \notin [0;1] \end{cases}$

в) $\frac{2}{3}$

г) $\frac{1}{18}$

Контрольная работа 3 — Задачный минимум

1. а) ≈ 0.15

б) $U \sim \mathcal{N}(101, 29)$, $f(u) = \frac{1}{\sqrt{2\pi \cdot 29}} e^{-\frac{1}{2}\frac{(u-101)^2}{29}}$

в) ≈ 0.02

2. а) 71.14

б) $f(y|x = 170) = \frac{1}{\sqrt{2\pi \cdot 20}} e^{-\frac{1}{2}\frac{(y-71.14)^2}{20}}$

в) ≈ 0

3. а) 0.25

б) 0.6875

в) $0.91(6)$

г) 0.75

д) -0.28125

4. а) $-1, 0, 1, 1$

б) -1

в) 1

г) $f(x) = \begin{cases} 0, & x < -1 \\ 0.25, & -1 \leqslant x < 0 \\ 0.5, & 0 \leqslant x < 1 \\ 1, & x \geqslant 1 \end{cases}$

5. а) θ

   б) да

6. а) нет, оценка смещена

   б) $c = 2$

7. а) все оценки несмещенные

   б) $\hat{p}_3$ наиболее эффективная

8. да

9. да

10. $\hat{\theta}_{MM} = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \overline{X})^2 \cdot 20}{n}}$

11. $\hat{\theta}_{MM} = \frac{1}{5}\left(6 - \frac{1}{n}\sum_{i=1}^{n} X_i^2\right)$, $\hat{\theta}_{MM} = 0.68$

12. $\hat{\theta}_{ML} = \frac{\sum_{i=1}^{n} x_i^2}{n}$

13. $\hat{p}_{ML} = \frac{\sum_{i=1}^{n} x_i}{n}$

14. да

15. $n_1 \approx 260$, $n_2 \approx 232$, $n_3 \approx 658$

Контрольная работа 4 — Задачный минимум

1. $\left[-1.6 - 1.65 \cdot \frac{2}{\sqrt{3}}; -1.6 + 1.65 \cdot \frac{2}{\sqrt{3}}\right]$

2. $\left[-1.6 - 2.92 \cdot \sqrt{\frac{18.33}{3}}; -1.6 + 2.92 \cdot \sqrt{\frac{18.33}{3}}\right]$

3. $\left[\frac{17.43 \cdot 2}{4.61}; \frac{17.43 \cdot 2}{0.21}\right]$

4. $\left[-1.6 - (-2.6) - 1.96 \cdot \sqrt{\frac{2}{3} + \frac{1}{2}}; -1.6 - (-2.6) + 1.96 \cdot \sqrt{\frac{2}{3} + \frac{1}{2}}\right]$

5. $\left[1.04 - (-0.37) - 3.18 \cdot \sqrt{3.02}\sqrt{\frac{1}{3} + \frac{1}{2}}; 1.04 - (-0.37) + 3.18 \cdot \sqrt{3.02}\sqrt{\frac{1}{3} + \frac{1}{2}}\right]$

6. $\left[0.45 - 1.96 \cdot \sqrt{\frac{0.45 \cdot 0.55}{100}}; 0.45 + 1.96 \cdot \sqrt{\frac{0.45 \cdot 0.55}{100}}\right]$

7. $\left[0.6 - 0.4 - 1.96 \cdot \sqrt{\frac{0.6 \cdot 0.4}{100} + \frac{0.4 \cdot 0.6}{200}}; 0.6 - 0.4 + 1.96 \cdot \sqrt{\frac{0.6 \cdot 0.4}{100} + \frac{0.4 \cdot 0.6}{200}}\right]$

8. $\left[2.5 - 1.96 \cdot \sqrt{\frac{1}{40}}; 2.5 + 1.96 \cdot \sqrt{\frac{1}{40}}\right]$

9. $\left[\frac{1}{0.52} - 1.96 \cdot \sqrt{\frac{1}{100 \cdot 0.52^2}}; \frac{1}{0.52} + 1.96 \cdot \sqrt{\frac{1}{100 \cdot 0.52^2}}\right]$

10. а) ≈ 0.02

    б) ≈ 0.02

    в) ≈ 0.98

11. 0.2

12. $z_{obs} \approx -1.39$, $z_{crit} = 1.28$, нет оснований отвергать H_0.

13. $t_{obs} \approx -0.65$, $t_{crit} = 1.89$, нет оснований отвергать H_0.

14. $z_{obs} \approx 0.93$, $z_{crit} = -1.65$, нет оснований отвергать H_0.

15. $t_{obs} \approx 0.89$, $t_{crit} = -2.35$, нет оснований отвергать H_0.

16. $F_{obs} \approx 95.37$, $F_{crit} = 199.5$, нет оснований отвергать H_0.

17. $z_{obs} \approx 2.04$, $z_{crit} = 1.65$, основная гипотеза отвергается.

18. $z_{obs} \approx 4.16$, $z_{crit} = 1.96$, основная гипотеза отвергается.

19. $\gamma_{obs} \approx 0.26$, $\gamma_{crit} = 5.99$, нет оснований отвергать H_0.

20. $\gamma_{obs} \approx 139.4$, $\gamma_{crit} = 3.84$, основная гипотеза отвергается.

21. $LR_{obs} \approx 5.5$, $LR_{crit} = 3.84$, основная гипотеза отвергается.

Решения контрольной номер 1

2017-2018

1. а) События называются независимыми, если $\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$

   б) Запасёмся всеми нужными вероятностями:
   $\mathbb{P}(A) = \frac{1}{2}$
   $\mathbb{P}(B) = \frac{1}{3}$
   $\mathbb{P}(C) = \frac{1}{2}$
   $\mathbb{P}(A \cap C) = \frac{1}{3}$ — выпадет чётое число больше трёх
   $\mathbb{P}(A \cap B) = \frac{1}{6}$ — выпадет чётное число, кратное трём
   $\mathbb{P}(A \cap C) = \frac{1}{6}$ — выпадет число, большее трёх и кратное трём
   Теперь можно проверять независимость:
   $\mathbb{P}(A \cap C) \neq \mathbb{P}(A) \cdot \mathbb{P}(C) \Rightarrow$ не являются независимыми
   $\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B) \Rightarrow$ являются независимыми
   $\mathbb{P}(B \cap C) = \mathbb{P}(B) \cdot \mathbb{P}(C) \Rightarrow$ являются независимыми

2. а) Количество возможных вариантов ТМ: C_{10}^2, количество возможных вариантов ЗМ: C_{24}^2. Количество их возможных сочетаний: $C_{10}^2 \cdot C_{24}^2$, где $C_n^k = \frac{n!}{k!(n-k)!}$.

   б) По классическому определению вероятностей, предполагая исходы равновероятными, искомая вероятность равна $\frac{C_{16}^2}{C_{24}^2}$

в) По тому же принципу:

$$\frac{C_k^2}{C_{10}^2} = \frac{1}{15} \Rightarrow \frac{\frac{k!}{2!(k-2)!}}{\frac{10!}{2!\cdot 8!}} = \frac{1}{15} \Rightarrow \frac{(k-1)k}{2}\frac{2}{9\cdot 10} = \frac{1}{15}$$

Получаем квадратное уравнение вида $k^2 - k - 6 = 0$ с корнями -2 и 3. Так как k не может быть отрицательным, ответ 3.

3. а) Если эксперт отдаёт предпочтение Fit, то это можно интерпретировать как «успех» в схеме Бернулли. Так как ξ - количество успехов, $k \in [0; 4]$, $p = \frac{1}{3}$, то

$$\mathbb{P}(\xi = k) = C_n^k (p)^k (1-p)^{n-k}$$

Большинство означает, что либо три, либо четыре эксперта выбрали Fit.

$$\mathbb{P}(\xi = 3) = C_4^3 \left(\frac{1}{3}\right)^3 \left(\frac{2}{3}\right)^1 = \frac{8}{81}$$

$$\mathbb{P}(\xi = 4) = C_4^4 \left(\frac{1}{3}\right)^4 \left(\frac{2}{3}\right)^0 = \frac{1}{81}$$

$$\mathbb{P}(\xi > 2) = \frac{9}{81}$$

б) Аналогично:

$$\mathbb{P}(\xi = 0) = C_4^0 \left(\frac{1}{3}\right)^0 \left(\frac{2}{3}\right)^4 = \frac{16}{81}$$

$$\mathbb{P}(\xi = 1) = C_4^1 \left(\frac{1}{3}\right)^1 \left(\frac{2}{3}\right)^3 = \frac{32}{81}$$

$$\mathbb{P}(\xi = 2) = C_4^2 \left(\frac{1}{3}\right)^2 \left(\frac{2}{3}\right)^2 = \frac{24}{81}$$

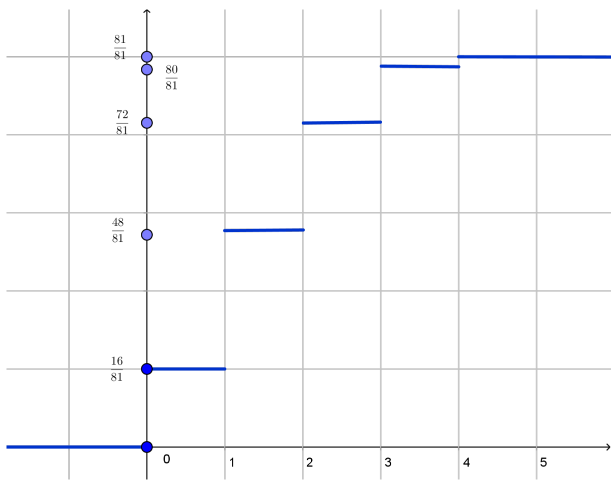


Рис. 1: Функция распределения

в) Все вероятности посчитаны, видим, что наибольшая достигается при $\xi = 1$.

г) $\mathbb{E}(X) = np = \frac{4}{3}$, $\text{Var}(X) = npq = \frac{8}{9}$

4. а) Так как указано, что цена сметаны распределена равномерно на отерзке $[250, 1000]$, максимальное значение цены — 1000, это и есть необходимая сумма.

   б) Вспомним, что функция распределения $F(x) = \mathbb{P}(X \leqslant x)$, нужно найти такой x, что $\mathbb{P}(X \leqslant x) = 0.9$:

$$0.9 = 1 - \exp(-x^2) \Rightarrow \exp(-x^2) = 0.1 \Rightarrow -x^2 = \ln(0.1) \Rightarrow x = \sqrt{-\ln(0.1)}$$

   в) Взяв производную от функции распределения списка без сметаны, получим функцию плотности:

$$f_X(x) = \begin{cases} 2x\exp(-x^2) & x \geqslant 0 \\ 0 & \text{иначе} \end{cases}$$

   Найдём математическое ожидание:

$$\int_0^{+\infty} 2x^2 \exp(-x^2) dx = -x\exp(-x^2)\big|_0^{+\infty} + \int_0^{+\infty} \exp(-x^2) dx = \frac{\sqrt{\pi}}{2}$$

   г) Математическое ожидание суммы случайных величин равно сумме математических ожиданий случайных влечин, если они существуют. Математическое ожидание от цены сметаны равно: $\frac{1000+250}{2} = 625$ Математическое ожидание списка без сметаны было найдено в предыдущем пункте, его осталось перевести в рубли. Получаем ответ: $625 + \frac{\sqrt{\pi}}{2} \cdot 1000$.

   д) Так как обе величины имеют абсолютно непрерывные распределения, вероятность попасть в конкретную точку равна нулю.

5. а) $\mathbb{P}(\text{детектор показал ложь и подозреваемый лжёт}) = 0.9 \cdot 0.1 + 0.1 \cdot 0.95 = 0.185$

   б) $\mathbb{P}(\text{невиновен}|\text{детектор показал ложь}) = \frac{0.9 \cdot 0.1}{0.185} = \frac{90}{185}$

   в) $\mathbb{P}(\text{эксперт точно выявит преступника}) = (0.9)^9 \cdot 0.95$

   г) $\mathbb{P}(\text{эксперт ошибочно выявит преступника}) = 9 \cdot 0.1 \cdot 0.9^8 \cdot 0.05$

2016-2017

1. а) Возможны четыре равновероятные ситуации:

$$\mathbb{P}(\text{ММ}) = \mathbb{P}(\text{МД}) = \mathbb{P}(\text{ДМ}) = \mathbb{P}(\text{ДД}) = 1/4$$

   Посчитаем условную вероятность:

$$\mathbb{P}(B \mid A) = \frac{\mathbb{P}(B \cap A)}{\mathbb{P}(A)} = \frac{\mathbb{P}(\text{МД, ДМ})}{\mathbb{P}(\text{ДМ, МД, ДД})} = \frac{2/4}{3/4} = \frac{2}{3}$$

   б) События A и B называются независимыми, если $\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$
   В нашем случае: $\mathbb{P}(A \cap B) = \mathbb{P}(\text{МД, ДМ}) = 2/4$, $\mathbb{P}(A) \cdot \mathbb{P}(B) = 3/4 \cdot 3/4$.
   Следовательно, $\mathbb{P}(A \cap B) \neq \mathbb{P}(A) \cdot \mathbb{P}(B)$, значит, события A и B не являются независимыми.

2. Пусть событие A_i означает, что i-ый узел системы дал сбой, а событие B_N, что вся система дала сбой.

   В условии сказано, что $\mathbb{P}(A_i) = 10^{-6}$, а найти нужно такое максимальное $N \in \mathbb{N}$, при котором

$$\mathbb{P}(B_N) \leqslant \frac{1}{10^2}$$

$$\begin{aligned}\mathbb{P}(B_N) &= \mathbb{P}\left(\cup_{i=1}^n A_i\right) = 1 - \mathbb{P}((\cup_{i=1}^n A_i)^c) \\ &\overset{\text{ф-ла де Моргана}}{=} 1 - \mathbb{P}\left(\cup_{i=1}^N A_i^c\right) \overset{A_1,\ldots,A_N - \text{ независ.}}{=} 1 - \mathbb{P}(A_1^c)\cdot\ldots\cdot\mathbb{P}(A_N^c) \\ &= 1 - \left(1 - \frac{1}{10^6}\right)^N\end{aligned}$$

Чтобы найти такое максимальное $N \in \mathbb{N}$, надо решить следующее неравенство

$$\begin{aligned}&1 - \left(1 - 10^{-6}\right)^N \leqslant 10^{-2} \\ &1 - 10^{-2} \leqslant \left(1 - 10^{-6}\right)^N \\ &\ln\left(1 - 10^{-2}\right) \leqslant N \ln\left(1 - 10^{-6}\right) \\ &N \leqslant \frac{\ln(1 - 10^{-2})}{\ln(1 - 10^{-6})} \approx 10050.33\end{aligned}$$

Значит, максимальное N равно 10050.

3. Введём обозначения для событий. Пусть A означает, что человек имеет заболевание лёгких, а B, что человек работал в шахте.

   В условии сказано, что $\mathbb{P}(B \mid A) = 0.22$, $\mathbb{P}(B \mid A^c) = 0.14$, $\mathbb{P}(A) = 0.04$.

   а) Нужно найти
   $$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B|A)\,\mathbb{P}(A)}{\mathbb{P}(B)}$$
   Для этого с помощью формулы полной вероятности посчитаем
   $$\mathbb{P}(B) = \mathbb{P}(B \mid A)\,\mathbb{P}(A) + \mathbb{P}(B \mid A^c)\,\mathbb{P}(A^c) = 0.22 \cdot 0.04 + 0.14 \cdot 0.96 = 0.1432$$
   Осталось подставить значения:
   $$\mathbb{P}(A \mid B) = \frac{0.22 \cdot 0.04}{0.1432} \approx 0.0615$$

   б) Все необходимые значения для второго пункта у нас есть, осталось применить формулу условной вероятности:
   $$\begin{aligned}\mathbb{P}(A \mid B^c) = \frac{\mathbb{P}(A \cap B^c)}{\mathbb{P}(B^c)} = \frac{\mathbb{P}(B^c \cap A)}{\mathbb{P}(A)} \cdot \frac{\mathbb{P}(A)}{\mathbb{P}(B^c)} = \mathbb{P}(B^c \mid A) \cdot \frac{\mathbb{P}(A)}{\mathbb{P}(B^c)} = \\ = (1 - \mathbb{P}(B \mid A)) \cdot \frac{\mathbb{P}(A)}{1 - \mathbb{P}(B)} = (1 - 0.22) \cdot \frac{0.04}{1 - 0.1432} \approx 0.0364\end{aligned}$$

4. Введём индикатор события «Петя дал верный ответ на i-ый вопрос»:
   $$X_i = \begin{cases} 1, & \text{если на } i\text{-ый вопрос теста Петя дал верный ответ} \\ 0, & \text{иначе} \end{cases}$$
   Заметим, что $X_i \sim Be\,(p = 1/5)$, $X_1, \ldots, X_{17}$ – независимы, $X = X_1 + \ldots + X_{17}$ – общее число верных ответов, $X \sim Bin\,(n = 17, p = 1/5)$.

   а) Наибольшее вероятное число правильных ответов m_0 может быть нвйдено по формуле:

1) если число $(n \cdot p - q)$ – не целое, где $q := 1 - p$, то

$$m_0 = [np - q] + 1,$$

2) если число $(n \cdot p - q)$ – целое, то наиболее вероятных значений m_0 два:

$$m_0' = np - q \text{ и } m_0'' = np - q + 1$$

Итак, поскольку $np - q = 17 \cdot \frac{1}{5} - \frac{4}{5} = 2.6$ – не целое, наиболее вероятное число верных ответов m_0 может быть найдено по формуле из пункта (1):

$$m_0 = [np - q] + 1 = [2.6] + 1 = 3$$

б)

$$\mathbb{E}(X) = np = 17 \cdot \frac{1}{5} = 3.4$$

$$\mathrm{Var}(X) = npq = 17 \cdot \frac{1}{5} \cdot \frac{4}{5} = 2.72$$

в)

$$\mathbb{P}(\text{Петя получит «отлично»}) = \mathbb{P}(X \geqslant 15) = \mathbb{P}(X = 15) + \mathbb{P}(X = 16) +$$
$$+ \mathbb{P}(X = 17) = C_{17}^{15} \cdot \left(\frac{1}{5}\right)^{15} \cdot \left(\frac{4}{5}\right)^{2} + C_{17}^{16} \cdot \left(\frac{1}{5}\right)^{16} \cdot \left(\frac{4}{5}\right)^{1} + C_{17}^{17} \cdot \left(\frac{1}{5}\right)^{17} \cdot \left(\frac{4}{5}\right)^{0} =$$
$$= 136 \cdot \frac{16}{5^{17}} + 17 \cdot \frac{4}{5^{17}} + \frac{1}{5^{17}} \approx 2.94 \cdot 10^{-9}$$

г) Рассмотрим первый вопрос теста. Петя может выбрать первый ответ с вероятностью $1/5$, и Вася может выбрать первый ответ с вероятностью $1/5$. Тогда они оба выберут одинаковый ответ с вероятностью $1/25$. Вариантов ответа в каждом вопросе 5, значит, вероятность совпадения ответа в одном вопросе равна $1/5$. Всего вопросов 17, тогда получаем

$$\mathbb{P}(\text{все ответы Пети и Васи совпадают}) = \left(\frac{1}{5}\right)^{17}$$

5. Введём случайную велчину η, которая означает число потенциальных покупателей, с которыми контактировал продавец оборудования. По условию задачи, η имеет таблицу распеределения:

η	1	2
$\mathbb{P}_\eta$	$1/3$	$2/3$

Случайная величина ξ может принимать значения $0, 50000$ и 100000

а) Найдём $\mathbb{P}(\xi = 0)$. По формуле полной вероятности, имеем:

$$\mathbb{P}(\xi = 0) = \mathbb{P}(\xi = 0 \mid \eta = 1) \cdot \mathbb{P}(\eta = 1) + \mathbb{P}(\xi = 0 \mid \eta = 2) \cdot \mathbb{P}(\eta = 2) =$$
$$= 0.9 \cdot \frac{1}{3} + 0.9 \cdot 0.9 \cdot \frac{2}{3} = 0.84$$

б) Найдём $\mathbb{P}(\xi = 50000)$ и $\mathbb{P}(\xi = 100000)$:

$$\mathbb{P}(\xi = 50000) = \mathbb{P}(\xi = 50000 \mid \eta = 1) \cdot \mathbb{P}(\eta = 1) + \mathbb{P}(\xi = 50000 \mid \eta = 2) \cdot \mathbb{P}(\eta = 2) = \\ = 0.1 \cdot \frac{1}{3} + 2 \cdot 0.1 \cdot 0.9 \cdot \frac{2}{3} = 0.15(3)$$

$$\mathbb{P}(\xi = 100000) = \mathbb{P}(\xi = 100000 \mid \eta = 1) \cdot \mathbb{P}(\eta = 1) + \mathbb{P}(\xi = 100000 \mid \eta = 2) \cdot \mathbb{P}(\eta = 2) = \\ = 0 \cdot \frac{1}{3} + 0.1 \cdot 0.1 \cdot \frac{2}{3} = 0.00(6)$$

Таблица распределения случайной величина ξ имеет вид:

ξ	0	5000	100000
$\mathbb{P}_\xi$	0.84	$0.15(3)$	$0.00(6)$

Тогда функция распределения случайной величины ξ имеет вид:

$$F_\xi(X) = \begin{cases} 0 & \text{при } x < 0 \\ 0.84 & \text{при } 0 \leqslant x < 50000 \\ 0.84 + 0.15(3) & \text{при } 50000 \leqslant x < 100000 \\ 1 & \text{при } x > 100000 \end{cases}$$

Опр.: $F_\xi = \mathbb{P}(\xi \leqslant x), x \in \mathbb{R}$

в)

$$\mathbb{E}(X) = 0 \cdot 0.84 + 50000 \cdot 0.15(3) + 100000 \cdot 0.00(6) = 8333.(3)$$

$$\mathrm{Var}(X) = (0 - 8333.(3))^2 \cdot 0.84 + (50000 - 8333.(3))^2 \cdot 0.15(3) + \\ + (100000 - 8333.(3))^2 \cdot 0.00(6) = 380555555.(5)$$

6. а) $f_\xi(x) = \begin{cases} \frac{1}{b} & \text{при } x \in [0, b] \\ 0 & \text{при } x \notin [0, b] \end{cases}$

б) Известно, что если $\xi \sim U[a, b]$, то $\mathbb{E}(\xi) = \frac{a+b}{2}$. Стало быть, из уравнения $\mathbb{E}(\xi) = 1$ получаем $\frac{b}{2} = 1$, то есть $b = 2$.

в) Известно, что если $\xi \sim U[a, b]$, то $\mathrm{Var}(\xi) = \frac{(b-a)^2}{12}$. Значит, $\mathrm{Var}(\xi) = \frac{2^2}{12} = \frac{1}{3}$

г) Воспользуемся формулой $\mathbb{P}(\xi \in B) = \int_B f_\xi(x)dx$. Имеем:

$$\mathbb{P}(\xi > 1) = \mathbb{P}(\xi \in (1, +\infty)) = \int_1^{+\infty} f_\xi(x)dx = \int_1^2 \frac{1}{2}dx = \frac{1}{2}$$

д) Требуется найти такое минимальное число $q_{0.25}$, что $\int_{-\infty}^{q_{0.25}} f_\xi(x)dx = 0.25$. Итак:

$$\int_{-\infty}^{q_{0.25}} f_\xi(x)dx = 0.25 \Leftrightarrow \int_{-\infty}^{q_{0.25}} \frac{1}{2}dx = 0.25 \Leftrightarrow \frac{1/2}{q_{0.25}} = 0.25 \Leftrightarrow$$

$$q_{0.25} = 2 \cdot 0.25 = 0.5$$

е)

$$\mathbb{E}[(\xi - \mathbb{E}(\xi))^{2017}] = \int_{-\infty}^{+\infty} (x - \mathbb{E}(\xi))^{2017} \cdot f_\xi(x)dx = \int_{-\infty}^{+\infty} (x-1)^{2017} f_\xi(x)dx =$$
$$= \int_0^2 (x-1)^{2017} \cdot \frac{1}{2}dx = \frac{(x-1)^{2018}}{2018} \cdot \frac{1}{2}\Big|_{x=0}^{x=2} = 0$$

ж) $F_\xi(x) = \begin{cases} 0 & \text{при } x < 0 \\ \frac{x}{2} & \text{при } 0 \leqslant x \leqslant 2 \\ 1 & \text{при } x > 2 \end{cases}$

з) Согласно условиям задачи, время до прихода 1-го поезда есть ξ; время до прихода 2-го поезда равно $\xi + b$; время до прихода 3-го (заветного) поезда есть $\xi + 2b$. Таким образом, Марья Ивановна в среднем ожидает «своего» поезда $\mathbb{E}(\xi + 2b) = 1 + 2b = 1 + 2 \cdot 2 = 5$ минут. При этом $\operatorname{Var}(\xi + 2b) = \operatorname{Var}(\xi) = 1/3$

к) Пусть τ – наименьший номер поезда без «подозрительных лиц». По условию задачи, таблица распределения случайной величины τ имеет вид:

τ	1	2	3	4	...
$\mathbb{P}_\tau$	$1/4$	$3/4 \cdot 1/4$	$(3/4)^2 \cdot 1/4$	$(3/4)^3 \cdot 1/4$	...

То есть случайная величина τ имеет геометрическое распределение с параметром $p = 1/4$ ($\tau \sim G(p = 1/4)$).

Несложно сообразить, что время ожидания Глафирой Петровной «своего» поезда составляет: $\eta := \xi + b(\tau - 1)$. Стало быть, $\mathbb{E}(\eta) = \mathbb{E}(\xi) + b \cdot (\mathbb{E}(\tau) - 1) = 1 + 2 \cdot (4 - 1) = 7$ минут.

Здесь мы воспользовались тем фактом, что если $\eta \sim G(p)$, то $\mathbb{E}(\eta) = 1/p$

и) Найдём теперь вероятность $\mathbb{P}(\eta \geqslant 5)$. Для нахождения искомой вероятности воспользуемся формулой полной вероятности:

$$\mathbb{P}(\eta \geqslant 5) = \mathbb{P}(\eta \geqslant 5, \tau < 3) + \mathbb{P}(\eta \geqslant 5, \tau = 3) + \mathbb{P}(\eta \geqslant 5, \tau > 3)$$

Если Глафира уехала на первом или втором поезде, то ждать больше 5 минут она не могла, то есть $\mathbb{P}(\eta \geqslant 5, \tau < 3) = 0$.

Если Глафира уехала на третьем поезде, то чтобы ждать больше пяти минут, ей нужно ждать первый поезд больше минуты, то есть $\mathbb{P}(\eta \geqslant 5, \tau = 3) = 0.5\,\mathbb{P}(\tau = 3)$.

Если Глафира уехала на четвертом поезде или позже, то она точно ждала больше 5 минут, $\mathbb{P}(\eta \geqslant 5, \tau > 3) = \mathbb{P}(\tau > 3)$.

$$\mathbb{P}(\eta \geqslant 5) = 0.5\,\mathbb{P}(\tau = 3) + \mathbb{P}(\tau > 3) = 0.5 \cdot (3/4)^2 \cdot (1/4) + (3/4)^3 = 63/128$$

7. Пусть ξ — случайная величина, обозначающая число остановок лифта. Предствим её в виде суммы $\xi = \xi_2 + \ldots + \xi_{10}$, где ξ_i — индикатор того, что лифт остановился на i-ом этаже, то есть

$$\xi_i = \begin{cases} 1 & \text{если лифт остановился} \\ 0 & \text{иначе} \end{cases} \qquad \forall i = 2, \ldots, 10$$

Найдём соответсвующие вероятности:

$$\mathbb{P}(\xi_i = 0) = \left(\frac{8}{9}\right)^9$$

$$\mathbb{P}(\xi_i = 1) = 1 - \mathbb{P}(\xi = 0) = 1 - \left(\frac{8}{9}\right)^9$$

Тогда $\mathbb{E}(\xi_i) = \mathbb{P}(\xi_i = 0) \cdot 0 + \mathbb{P}(\xi_i = 1) \cdot 1 = 1 - \left(\frac{8}{9}\right)^9$, и в итоге получаем:

$$\mathbb{E}(\xi) = 9 \cdot \mathbb{E}(\xi_i) = 9 \cdot \left(1 - \left(\frac{8}{9}\right)^9\right)$$

Решения контрольной номер 1. ИП

2017-2018

1. Обозначим вероятность того, что сыр достанется Белому за b, если игра начинается с его броска.

   а) Получаем уравнение
   $$b = \frac{1}{12} + \frac{11}{12} \cdot \frac{11}{12} b$$
   Пояснение: Как Белый может победить в исходной игре? Либо сразу выкинуть 6 с вероятностью $1/12$. Либо передать ход Серому $(11/12)$, получить ход снова $(11/12)$ и выиграть в продолжении игры. Продолжение игры по сути совпадает с исходной игрой.

   б) Игра продолжается до тех пор, пока кто-то не выкинет «6». Для нахождения среднего количества бросков воспользуемся методом первого шага.

   Обозначим среднее количество бросков нашей игры за S. Когда Белый бросает кубик, с вероятностью $\frac{1}{12}$ игра закончится за один бросок, а с вероятностью $\frac{11}{12}$ игра продолжится и ход перейдёт к Серому. Но та игра, которая начнётся, когда бросать будет Серый, ничем не отличается от предыдущей, поэтому среднее количество бросков в ней будет равно S. Однако мы попадём в эту игру, «потратив» один бросок. Таким образом мы получаем:
   $$S = \frac{1}{12} \cdot 1 + \frac{11}{12}(S + 1)$$
   Получается, что $S = 12$, значит игра длится в среднем 12 бросков.

2.

3. Для того, чтобы выжить, мышам нужно ещё до начала игры договориться о стратегии, которая позволит им с наибольшей вероятностью открыть нужные сундуки. Если хотя бы две мыши выберут одинаковый сундук, то их в любом случае съедят. Поэтому одной из оптимальных стратегий будет ещё до начала игры мышам договориться и назвать левый сундук золотым, сундук посередине серебряным, а правый — платиновым. Каждый мышонок должен открыть тот сундук, в честь которого назван необходимый ему металл. Если внутри он обнаруживает свой металл, то он выбирает этот сундук, если внутри находится не тот металл, мышонок открывает тот сундук, на который указывает лежащий внутри предмет.

   Например, первым заходит Микки Маус. Он открывает золотой (левый) ящик. Если внутри лежит золото, то он выходит из комнаты. Если же внутри лежит, например, серебро, то Микки Маус открывает сундук посередине. Путём перебора можно посчитать, что в 4 случаях из 6 мыши смогут найти нужный металл, поэтому вероятность выигрыша при данной стратегии равна $\frac{2}{3}$.

4.

5. Благосостояние кота Василия, положившего один гурд на вклад, равно $m_t = 1 \cdot e^{rt}$, где r — процентная ставка, а t — прошедшее время. Момент закрытия вклада T равномерно распределён на отрезке от 0 до a, поэтому сумма, которую получит Василий, представима в виде $Z = e^Y$, где $Y \sim U[0; ra]$. По условию, a очень велико, поэтому ra тоже очень велико.

   Вероятность того, что первая цифра будет равна 1, равна вероятности того, что доход Василия будет лежать в пределах от 1 до 2 гурдов, плюс вероятность того, что он лежит в пределах от 10 до 20 гурдов и т.д. Таким образом, можно представить эту вероятность, как:

$$\mathbb{P}(N = 1) = \mathbb{P}(e^Y \in [1; 2)) + \mathbb{P}(e^Y \in [10; 20)) + \ldots$$

   Это выражение можно преобразовать таким образом:

$$\mathbb{P}(N = 1) = \mathbb{P}(Y \in [\ln 1; \ln 2)) + \mathbb{P}(Y \in [\ln 10; \ln 20)) + \ldots$$

   Так как Y — равномерно распределённая величина, то $\mathbb{P}(Y \in [\ln 1; \ln 2)) = \frac{\ln 2 - \ln 1}{ra}$. Для последующих слагаемых вероятность рассчитывается таким же образом. Воспользовавшись свойством логарифма, можно заметить, что $\frac{\ln 20 - \ln 10}{ra} = \frac{\ln 2}{ra}$. Поэтому вероятность того, что на первом месте суммы вклада стоит единица, равна $n \cdot \frac{\ln 2}{ra}$, где n — количество слагаемых. Путём аналогичных рассуждений получаем, что вероятность того, что на первом месте стоит двойка, равна $n \cdot \frac{\ln 3 - \ln 2}{ra}$. Из-за того, что a велико, можно считать, что число слагаемых одинаково.

   На первом месте обязательно будет находиться какая-то цифра, поэтому сумма вероятностей будет равна 1. Получаем:

$$\frac{n}{ra}\left(\ln\frac{2}{1} + \ln\frac{3}{2} + \ldots + \ln\frac{10}{9}\right) = 1$$

   Таким образом $\frac{n}{ra} = \frac{1}{\ln 10}$. Получается, что вероятность того, что на первом месте стоит единица, равна:

$$\mathbb{P}(N = 1) = \frac{\ln 2}{\ln 10}$$

   Закон распределения первой цифры выводится сложением соответствующих вероятностей.

Решения контрольной номер 2

2017-2018

7. а) Всем хватит места, если число явившихся на рейс пассажиров (X) не превысит 300, то есть нужно найти $\mathbb{P}(X \leqslant 300)$. Найдём матожидание и дисперсию случайной величины X:

$$\mathbb{E}(X) = np = 330 \cdot 0.9 = 297$$
$$\text{Var}(X) = np(1-p) = 330 \cdot 0.9 \cdot 0.1 = 29.7$$

   Теперь посчитаем нужную вероятность:

$$\mathbb{P}(X \leqslant 300) = \mathbb{P}\left(\frac{X - 297}{\sqrt{29.7}} \leqslant \frac{300 - 297}{\sqrt{29.7}}\right) = \mathbb{P}(\mathcal{N}(0,1) \leqslant 0.55) \approx 0.709$$

б) Вероятность переполнения не должна превышать 0.1:

$$\begin{aligned}
&\mathbb{P}(X > 300) < 0.1\\
&\mathbb{P}\left(\frac{X - 0.9\cdot n}{\sqrt{0.9\cdot 0.1\cdot n}} > \frac{300 - 0.9\cdot n}{\sqrt{0.9\cdot 0.1\cdot n}}\right) < 0.1\\
&\frac{300 - 0.9\cdot n}{\sqrt{0.9\cdot 0.1\cdot n}} > 1.28\\
&300 - 0.9n > 1.28\cdot 0.3\sqrt{n}\\
&n < 325.6
\end{aligned}$$

8. а) Выпишем случайную величину X_i — цену акции после i-ого дня:

$$X_i = \begin{cases} 1.01, & p = 0.7\\ 0.99, & p = 0.2999\\ 0, & p = 0.0001\end{cases}$$

Нужно посчитать ожидание цены акциии после 20 дней:

$$\mathbb{E}(X_1\cdot\ldots\cdot X_{20}) \overset{\text{незав-ть}}{=} \mathbb{E}(X_1)\cdot\ldots\cdot\mathbb{E}(X_{20}) = 1.004^{20} \approx 1.083$$

б) По ЗБЧ:

$$\operatorname*{plim}_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n} X_i = \mathbb{E}(X_i) = 1.004$$

в) Аналогично пункту (а):

$$\mathbb{E}(X_1\cdot\ldots\cdot X_n) = (\mathbb{E}(X_1))^n = 1.004^n$$

И понятно, что $1.004^n \to_{n\to\infty} +\infty$.

г)

$$\begin{aligned}
\mathbb{P}(\text{разорения}) = 1 - \mathbb{P}(X_1 > 0,\ldots,X_n > 0) = 1 - \prod_{i=1}^{n}\mathbb{P}(X_i > 0)\\
= 1 - (1 - 0.0001)^n \to_{n\to\infty} 1
\end{aligned}$$

Решения контрольной номер 3

2017-2018

5. а) $L(X_1,\ldots,X_n,\mu) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\sum_{i=1}^n (X_i-\mu)^2}$

б) $\hat{\mu}_{ML} = \bar{X}$

в) $\mathbb{E}(\hat{\mu}_{ML}) = \mathbb{E}(\bar{X}) = \mu \Rightarrow$ оценка несмещённая

$\operatorname{plim}\hat{\mu}_{ML} = \operatorname{plim}\bar{X} = \mu \Rightarrow$ оценка состоятельная

г) $I(\mu) = n$

д) $\operatorname{Var}(\theta) \geqslant \frac{1}{I(\theta)}$

е) $\operatorname{Var}(\hat{\mu}_{ML}) = \frac{1}{n}$, так как неравенство Рао-Крамера выполнено как равенство, оценка является эффективной.

ж) $\theta = \mathbb{E}(X^2) = \operatorname{Var}(X) + \mu^2 = 1 + \mu^2$. Тогда в силу инвариантности оценок максимального правдоподобия: $\hat{\theta}_{ML} = 1 + \hat{\mu}^2$.

з) $\mathbb{E}(\hat{\theta}_{ML}) = 1 + \mathbb{E}(\hat{\mu}^2) = 1 + \mathbb{E}((\bar{X})^2)$
Пользуясь соотношением $\mathbb{E}((\bar{X})^2) = \operatorname{Var}(\bar{X}) + (\mathbb{E}(\bar{X}))^2$, получим: $\mathbb{E}(\hat{\theta}_{ML}) = 1 + \frac{1}{n} + \mu^2$, то есть оценка смещена.
Однако, $\lim_{n\to\infty}\left(1 + \frac{1}{n} + \mu^2\right) = 1 + \mu^2$, значит, оценка асимптотически несмещена.

и) $\hat{\theta}_{ML} \approx 1 + \mu^2 + 2\mu(\hat{\mu} - \mu)$
$\operatorname{Var}(\hat{\theta}_{ML}) \approx 4\mu^2 \operatorname{Var}(\hat{\mu}) = \frac{4\mu^2}{n}$

к) Так как $\hat{\theta}_{ML}$ асимптотически несмещена, то для проверки состоятельности достаточно показать, что $\operatorname{Var}(\hat{\theta}_{ML}) = \frac{4\mu^2}{n} \to_{n\to\infty} 0$.

6. а) $\mathbb{E}(X_1) = \int_0^\theta \frac{2}{\theta^2}(\theta - x)x dx = \frac{\theta}{3}$
$\frac{\hat{\theta}_{MM}}{3} = \bar{X} \Rightarrow \hat{\theta}_{MM} = 3\bar{X}$

б) Оценка $\hat{\theta}$ состоятельна. если $\operatorname{plim}\hat{\theta}_n = \theta$.
$\operatorname{plim}\hat{\theta}_{MM} = \operatorname{plim} 3\bar{X} = 3\,\mathbb{E}(X_1) = \theta \Rightarrow$ оценка состоятельна.

7. а) $\mathbb{E}\left(\frac{X_1+X_2+X_3}{3}\right) = \frac{1}{3}\cdot 3\,\mathbb{E}(X_1) = 132.5$
$\operatorname{Var}\left(\frac{X_1+X_2+X_3}{3}\right) = \frac{1}{9}\operatorname{Var}(X_1+X_2+X_3) = \frac{1}{9}(\operatorname{Var}(X_1)+\operatorname{Var}(X_2)+\operatorname{Var}(X_3)+2\operatorname{Cov}(X_1,X_2)+ 2\operatorname{Cov}(X_1,X_3) + 2\operatorname{Cov}(X_2,X_3)) = \frac{1}{9}(3\operatorname{Var}(X_1) + 6\operatorname{Cov}(X_1,X_2))$
$\operatorname{Var}(X_1) = \mathbb{E}(X_1^2) - \mathbb{E}(X_1)^2 = \frac{1}{4}\cdot 30^2 - \frac{1}{4}\cdot 500^2 - 132.5^2 = 45168.75$
$\operatorname{Cov}(X_1, X_1 + \ldots + X_4 = \operatorname{Var}(X_1) + 3\operatorname{Cov}(X_1,X_2) = 0 \Rightarrow \operatorname{Cov}(X_1,X_2) = -\frac{45168.75}{3} = -15056.25$
$\operatorname{Var}\left(\frac{X_1+X_2+X_3}{3}\right) = 5018.75$

б) 3/4

8. $\Delta_i = X_i - Y_i \sim \mathcal{N}(\mu_x - \mu_y, \sigma^2)$
$\bar{X} = 297.5, \bar{Y} = 247.5, \bar{\Delta} = \bar{X} - \bar{Y} = 50$
$\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^n(\Delta_i - \bar{\Delta})^2 = 18266.(6)$.
Критическое значение — $t_{0.975,3} = 3.182$ и доверительный интервал имеет вид:

$$50 - 3.182\sqrt{\frac{18266.(6)}{3}} < \mu_x - \mu_y < 50 + 3.182\sqrt{\frac{18266.(6)}{3}}$$

Так как 0 входит в доверительный интервал, нельзя отвергнуть предположение о равенстве расхожов.

9. а) $0.7 - 1.96\sqrt{\frac{0.7\cdot 0.3}{60}} < p < 0.7 + 1.96\sqrt{\frac{0.7\cdot 0.3}{60}}$

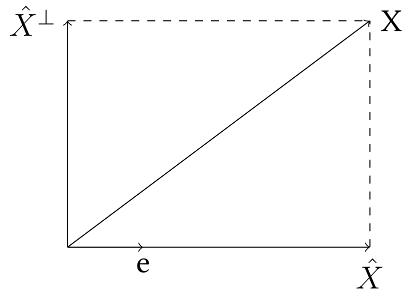
б) Да, так как 0.7667 входит в доверительный интервал.

в) $\mathbb{P}(|p - \hat{p}| \leqslant 0.01) = 0.95$
$\mathbb{P}\left(\frac{|0.7-p|}{\sqrt{\frac{0.7\cdot 0.3}{n}}} < \frac{0.01}{\sqrt{\frac{0.7\cdot 0.3}{n}}}\right) = 0.95$
$\frac{0.01}{\sqrt{\frac{0.7\cdot 0.3}{n}}} = 1.96 \Rightarrow n \approx 8068$

Решения контрольной номер 3. ИП

2017-2018

1а) - в) См. картинку :)



г) $\hat{X} = e \cdot \bar{X}$
   $\|\hat{X}\| = \sqrt{n} \cdot \bar{X}$
   $\hat{X}^{\perp} = X - e \cdot \bar{X} = (X_1 - \bar{X}, \ldots, X_n - \bar{X})$
   $\|\hat{X}^{\perp}\| = \sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2}$

д) $\|X\|^2 = \|\hat{X}^{\perp}\|^2 + \|\hat{X}\|^2$
   $\sum_{i=1}^{n} X_i^2 = \sum_{i=1}^{n}(X_i - \bar{X})^2 + n\bar{X}^2$

е) t-статистика для построения доверительного интервала для μ имеет вид:

$$\begin{aligned} t &= \frac{\bar{X} - \mu}{\sqrt{\bar{\sigma}^2/n}} = \frac{\bar{X} - \mu}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2/(n \cdot (n-1))}} \\ &= \sqrt{n-1} \cdot \frac{\sqrt{n} \cdot (\bar{X} - \mu)}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2}} = \sqrt{n-1} \cdot \frac{\|\hat{X}\| - \sqrt{n} \cdot \mu}{\|\hat{X}^{\perp}\|} \end{aligned}$$

Заметим, что $\operatorname{ctg} \alpha$ есть отношение прилежащего катета к противолежащему, таким образом, нужный нам угол α образуется между векторами X и $\hat{X}$. Зметим однако, что в нашем случае

$$t = \sqrt{n-1} \cdot \operatorname{ctg} \alpha = \frac{\|\hat{X}\|}{\|\hat{X}^{\perp}\|},$$

то есть наша статистика подойдёт только для проверки гипотезе о равенстве матожидания нулю.

Замечание. $t = \sqrt{n-1} \cdot \operatorname{ctg} \alpha$ будет t-статистикой только в том случае, если X_i будут н.о.р.с.в. с нормальным распределением, о чём в условие сказано не было.

2. Выпишем функцию правдоподобия для выборки из трёх видов, два из которых совпадают. Первый медведепришелец будет нового вида с вероятностью 1. Вероятность, что вид второго пойманного медведепришельца совпадёт с первмым, составляет $1/n$. После этого нужно поймать медведепришельца нового вида – это произойдёт с вероятностью $(n-1)/n$, и ещё одного нового вида – вероятность этого $(n-2)/n$. Поскольку медведепришелец, вид которого встречается дважды, мог встретить на любой из трёх позиций, функцию правдоподобия необходимо домножить на C_3^1. Таким образом, функция правдоподобия имеет вид:

$$L(n) = C_3^1 \cdot 1 \cdot \frac{1}{n} \cdot \frac{n-1}{n} \cdot \frac{n-2}{n}, n \geqslant 3.$$

Максимизируя её, внутри области определения получаем $\hat{n} = 5$.

Так как количество медведей велико и все они встречаются равновероятно, то $p_1 = p_2 = p_3 = 1/n$. Так же из выборки известно, что число видов космомедведей не меньше трёх. Потому $\hat{n} \geqslant 3$.

Найдите хитрую ошибку в предложенном решении:

$L(n) = \left(\frac{1}{n}\right)^2 \cdot \frac{1}{n} \cdot \frac{1}{n} = n^{-4}$

$\frac{\partial L(n)}{\partial n} = -4 \cdot n^{-5} = 0$

Данное уравнение не имеет решений при конечных n, но заметим, что при всех $n \geqslant 3$ выполняется $\frac{\partial L(n)}{\partial n} = -4 \cdot n^{-5} < 0$, таким образом максимальное значение находится в граничных точках.

$\lim\limits_{n \to \infty} \frac{1}{n^{-4}} = 0 < \frac{1}{3^{-4}}$

Таким образом получаем, что $\hat{n} = 3$.

3. а) $L(p_1, p_2) = p_1^{150} \cdot p_2^{100} \cdot (1 - p_1 - p_2)^{50}$

$\ell(p_1, p_2) = 150 \ln p_1 + 100 \ln p_2 + 50 \ln(1 - p_1 - p_2)$

$\begin{cases} \frac{\partial \ell(p_1,p_2)}{\partial p_1} = \frac{150}{p_1} - \frac{50}{1-p_1-p_2} = 0 \\ \frac{\partial \ell(p_1,p_2)}{\partial p_2} = \frac{100}{p_2} - \frac{50}{1-p_1-p_2} = 0 \end{cases}$

Откуда получаем:

$\begin{cases} \hat{p}_1 = 1/2 \\ \hat{p}_2 = 1/3 \end{cases}$

б) Найдём, какие значения должны стоять в теоретической ковариационной матрице. Заметим, что случайная величина найти кальмаромедведя (X) или двурога (Y) есть бернулевская случайная величина с параметром $p_{1+2} = p_1 + p_2$ и дисперсией $p_{1+2} \cdot (1 - p_{1+2})$, но тогда:

$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2 \cdot \mathrm{Cov}(X, Y)$

$\mathrm{Cov}(X,Y) = \frac{1}{2} \cdot (\mathrm{Var}(X+Y) - \mathrm{Var}(X) - \mathrm{Var}(Y)) = \frac{1}{2} \cdot ((p_1 + p_2) \cdot (1 - p_1 - p_2) - p_1 \cdot (1 - p_1) - p_2 \cdot (1 - p_2)) = -p_1 \cdot p_2$

Тогда подставляя в теоретическую ковариационную матрицу оценки параметров и домнажая всё на $1/300$, так как $\hat{p}_1$ и $\hat{p}_2$ являются средними, получим:

$$\mathrm{Var}(\hat{p}) = \frac{1}{n} \begin{pmatrix} \hat{p}_1 \cdot (1 - \hat{p}_1) & -\hat{p}_1 \cdot \hat{p}_2 \\ -\hat{p}_1 \cdot \hat{p}_2 & \hat{p}_1 \cdot (1 - \hat{p}_1) \end{pmatrix} = \frac{1}{300} \begin{pmatrix} 1/4 & -1/6 \\ -1/6 & 2/9 \end{pmatrix}$$

в) Для начала, найдём теоретическую дисперсию $\mathrm{Var}(X - Y)$.

$$\mathrm{Var}(X - Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) - 2 \cdot \mathrm{Cov}(X, Y) = p_1 \cdot (1 - p_1) + p_2 \cdot (1 - p_2) + 2 \cdot p_1 \cdot p_2$$

Тогда подставляя оценки для p_1 и p_2 и учитывая, что это оценки среднего, получим оценку:

$$\mathrm{Var}(\hat{p}_1 - \hat{p}_2) = 1/300 \cdot (1/4 + 2/9 + 2 \cdot 1/6) = 29/(36 \cdot 300)$$

г) Так как выборка достаточно велика, то статистика $\hat{p}_1 - \hat{p}_2$, являясь средним, будет иметь примерно нормальное распределение, и тогда:

$\frac{\hat{p}_1 - \hat{p}_2 - (p_1 - p_2)}{\sqrt{\mathrm{Var}(\hat{p}_1 - \hat{p}_2)}} \sim \mathcal{N}(0, 1)$

$\hat{p}_1 - \hat{p}_2 - z_{1-\frac{\alpha}{2}} \cdot \sqrt{\mathrm{Var}(\hat{p}_1 - \hat{p}_2)} \leqslant p_1 - p_2 \leqslant \hat{p}_1 - \hat{p}_2 - z_{\frac{\alpha}{2}} \cdot \sqrt{\mathrm{Var}(\hat{p}_1 - \hat{p}_2)}$

4. а) $\hat{\alpha} = \bar{X} + \sqrt{\bar{X} + 6} = 10 + \sqrt{10 + 6} = 14$

б) Так как $\bar{X}$ сходится по распределению к нормальному распределению и $\hat{\alpha} = g(\bar{X})$, где $g(\bar{X})$ гладкая по $\bar{X}$ функция при $\bar{X} \geqslant 0$, а также $\bar{X}$ сходится по вероятности к матожиданию, то можно абсолютно спокойно применить дельта-метод. Тогда:

$$(\alpha - g(\bar{X})) \sim N(0; \sigma^2 (g'(\mathbb{E}(X_1)))^2/n)$$

Но так как $\hat{\alpha}$ является состоятельной оценкой, то можно заменить $g'(\mathbb{E}(X_1))$ на $g'(\bar{X})$:

$$g'(\bar{X}) = 1 + \frac{1}{2 \cdot \sqrt{\bar{X}+6}} = 1 + \frac{1}{2 \cdot 4} = \frac{9}{8} = 1.125$$

и тогда можно построить ассимтотический доверительный интервал:

$$\begin{aligned}
\hat{\alpha} - z_{97.5\%} \cdot \sqrt{\sigma^2 \cdot (g'(\bar{X}))^2/n} \leqslant &\alpha \leqslant \hat{\alpha} - z_{2.5\%} \cdot \sqrt{\sigma^2 \cdot (g'(\bar{X}))^2/n} \\
16 - 1.96 \cdot 2 \cdot 9/(8 \cdot 10) \leqslant &\alpha \leqslant 16 + 1.96 \cdot 2 \cdot 9/(8 \cdot 10) \\
13.559 \leqslant &\alpha \leqslant 14.441
\end{aligned}$$

5. а) Так как не известно точно, кто сколько фотографий сделал, и так как метод оценки не указан, то воспользуемся методом моментов для построения оценки.

$$\begin{aligned}
N &= \mathbb{E}(\text{«фото Андрея»}) + \mathbb{E}(\text{«фото Беллы»}) \\
130 &= 100 \cdot 0.5 + p \cdot 100 \\
\hat{p} &= 0.8
\end{aligned}$$

Так как выборка достаточно велика, то $\frac{\hat{p}-p}{\sqrt{\hat{p}\cdot(1-\hat{p})/W}} \sim \mathcal{N}(0,1)$

$$\begin{aligned}
\hat{p} - z_{97.5\%}\sqrt{\hat{p}\cdot(1-\hat{p})/W} \leqslant &p \leqslant \hat{p} - z_{2.5\%}\sqrt{\hat{p}\cdot(1-\hat{p})/W} \\
0.8 - 1.96 \cdot \sqrt{0.8 \cdot 0.2/100} \leqslant &p \leqslant 0.8 + 1.96 \cdot \sqrt{0.8 \cdot 0.2/100} \\
0.72 \leqslant &p \leqslant 0.88
\end{aligned}$$

б) Так как неизвестно, кто больше снимков сделал, то рассмотрим два случая: Андрей сделал 60 фото и Белла — 70 фото, Андрей сделал 70 фото и Белла — 60 фото. В каждом случае при помощи метода максимального правдоподобия оценим вероятность p, после чего сравним значения функции правдоподобия с оценёнными параметрами для каждого случая.

$$\begin{aligned}
L(p) &= C_{100}^{60} \cdot 0.5^{60} \cdot 0.5^{40} \cdot C_{100}^{70} \cdot p^{70} \cdot (1-p)^{30} \\
\ell(p) &= const + 70 \ln p + 30 \ln(1-p) \\
\frac{\partial \ell(p)}{\partial p} &= \frac{70}{p} - \frac{30}{1-p} = 0 \\
\hat{p}_1 &= 0.7
\end{aligned}$$

Аналогично для второго случая получим оценку: $\hat{p}_2 = 0.6$.

Для простоты, будем сравнивать логарифмически функции правдоподобия $\ell_1(p_1)$ и $\ell_2(p_2)$ и тогда получим:

$$\begin{aligned}
\ell_1(p_1) &= const + 70 \ln 0.7 + 30 \ln 0.3 \approx const - 70 \cdot 0.357 - 30 \cdot 1.204 = const - 61.11 \\
\ell_2(p_2) &= const + 60 \ln 0.6 + 40 \ln 0.4 \approx const - 60 \cdot 0.511 - 40 \cdot 0.916 = const - 67.3
\end{aligned}$$

Так как $-67.3 < -61.11$, то более вероятно, что $\hat{p} = 0.7$

Тогда анологично предыдущему пункту получим доверительный интервал:

$$0.7 - 1.96 \cdot \sqrt{0.7 \cdot 0.3/100} \leqslant p \leqslant 0.7 + 1.96 \cdot \sqrt{0.7 \cdot 0.3/100}$$
$$0.61 \leqslant p \leqslant 0.79$$

Решения контрольной номер 4

2017-2018

1. Проверяем следующую гипотезу:

$$\begin{cases} H_0 : \mu_D = 100 \\ H_a : \mu_D > 100 \end{cases}$$

Считаем наблюдаемое значение статистики:

$$t_{obs} = \frac{\bar{X} - \mu_D}{\frac{\sigma_D}{\sqrt{n_D}}} = \frac{136 - 100}{\frac{55}{\sqrt{40}}} \approx 4.14$$

При верной H_0 t-статистика имеет распределение t_{40-1}, значит, $t_{crit} \approx 1.68$. Поскольку $t_{crit} > t_{obs}$, основная гипотеза отвергается, $p - value \approx 0$.

2. Проверяем следующую гипотезу:

$$\begin{cases} H_0 : \sigma_D^2 = \sigma_T^2 \\ H_a : \sigma_D^2 \neq \sigma_T^2 \end{cases}$$

Считаем наблюдаемое значение статистики:

$$F_{obs} = \frac{\hat{\sigma}_D^2}{\hat{\sigma}_T^2} = \frac{55^2}{60^2} \approx 0.84$$

При верной H_0 F-статистика имеет распределение $F_{40-1,60-1}$. Находим критические значения: $F_{left} \approx 0.6$, $F_{right} \approx 1.6$. Поскольку $F_{left} < F_{obs} < F_{right}$, нет оснований отвергать H_0.

3. Проверяем гипотезу

$$\begin{cases} H_0 : \mu_D = \mu_T \\ H_a : \mu_D < \mu_T \end{cases}$$

Когда n_D, n_T велики,

$$\frac{\bar{D} - \bar{T} - (\mu_D - \mu_T)}{\sqrt{\frac{\sigma_D^2}{n_D} + \frac{\sigma_T^2}{n_T}}} \overset{H_0}{\sim} \mathcal{N}(0, 1)$$

Считаем наблюдаемое значение статистики:

$$z_{obs} = \frac{136 - 139}{\sqrt{\frac{3025}{40} + \frac{3600}{60}}} \approx -0.25$$

По таблице находим $z_{crit} = -1.28$. Так как $z_{crit} < z_{obs}$, нет оснований отвергать H_0.

б) Когда считаем дисперсии одинаковыми, то:

$$\hat{\sigma}_0^2 = \frac{\hat{\sigma}_D^2(n_D - 1) + \hat{\sigma}_T^2(n_T - 1)}{n_D + n_T - 2} = \frac{3025 \cdot 39 + 3600 \cdot 59}{30 + 60 - 2} \approx 3371$$

и

$$\frac{\bar{D} - \bar{T} - (\mu_D - \mu_T)}{\hat{\sigma}_0^2 \sqrt{\frac{1}{n_D} + \frac{1}{n_T}}} \overset{H_0}{\sim} t_{n_D + n_T - 2}$$

Считаем наблюдаемое значение статистики:

$$t_{obs} = \frac{136 - 139}{\sqrt{3371}\sqrt{\frac{1}{40} + \frac{1}{60}}} \approx -0.25$$

По таблице находим критическое значение: $t_{crit} \approx -1.29$. Поскольку $t_{crit} < t_{obs}$, нет оснований отвергать H_0.

4. а) Сначала найдём оценку максимального правдоподобия параметра λ:

$$L = \prod_{i=1}^{n} \lambda e^{-\lambda x_i} = \lambda^n e^{-\lambda \sum_{i=1}^{n} x_i}$$

$$\ell = n \ln \lambda - \lambda \sum_{i=1}^{n} x_i$$

$$\frac{\partial \ell}{\partial \lambda} = \frac{n}{\lambda}\Big|_{\lambda = \hat{\lambda}} = 0$$

$$\frac{\partial^2 \ell}{\partial \lambda^2} = -\frac{n}{\lambda^2}$$

$$\hat{\lambda} = \frac{1}{0.52}$$

Так как

$$\frac{\hat{\lambda} - \lambda}{\sqrt{\frac{1}{I(\lambda)}}} \overset{as}{\sim} \mathcal{N}(0, 1),$$

доверительный интервал имеет вид

$$\frac{1}{0.52} - 1.96 \frac{1}{\frac{10}{0.52}} < \lambda < \frac{1}{0.52} + 1.96 \frac{1}{\frac{10}{0.52}}$$

б) Найдём вероятность того, что наушник проработает без сбоев 45 минут:

$$g(\lambda) = \mathbb{P}(X > 0.75) = 1 - F(0.75) = e^{-0.75\lambda}$$

Тогда

$$g(\hat{\lambda}) = e^{-0.75/0.52}$$

$$g'(\hat{\lambda}) = -0.75 e^{-0.75/0.52}$$

И доверительный интервал имеет вид:

$$e^{-0.75/0.52} - 1.96 \cdot \frac{0.75 \cdot 0.52}{10} \cdot e^{-0.75/0.52} < g(\lambda) < e^{-0.75/0.52} + 1.96 \cdot \frac{0.75 \cdot 0.52}{10} \cdot e^{-0.75/0.52}$$

5. Выпишем функцию правдоподобия:

$$
\begin{aligned}
&L = p_1^{10} \cdot p_2^{10} \cdot p_3^{15} \cdot p_4^{15} \cdot p_5^{25} \cdot (1 - p_1 - p_2 - p_3 - p_4 - p_5)^{25} \\
&\ell = 10 \ln p_1 + 10 \ln p_2 + 15 \ln p_3 + 15 \ln p_4 + 25 \ln p_5 + 25 \ln(1 - p_1 - p_2 - p_3 - p_4 - p_5)
\end{aligned}
$$

Максимизируя логарифмическую функцию правдоподобия по всем параметрам, получим следующие оценки для неограниченной модели:

$$
\begin{aligned}
&\hat{p}_1 = \hat{p}_2 = 0.1 \\
&\hat{p}_3 = \hat{p}_4 = 0.15 \\
&\hat{p}_5 = 0.25
\end{aligned}
$$

Подставив найденные значения в логарифмическую функцию правдоподобия, получим

$$\ell_{UR} \approx -172$$

В ограниченной модели $p_1 = \ldots = p_6 = 1/6$, и значение функции правдоподобия будет

$$\ell_R \approx -179$$

Теперь можно посчитать наблюдаемое значение:

$$LR = 2(\ell_{UR} - \ell_R) = 2(-172 - (-179)) = 14$$

Критическое значение $\chi_{0.95,5} =\approx 11 < 14$, значит, основная гипотеза отвергается.

Решения финальных экзаменов

2017-2018

Здесь табличка с ответами