СЕМ практикум, СИ, 2025/26

В следващите задачи намерете приближение на търсените вероятности чрез подходящи симулации.

Задача 1. В кутия има 8 топки, номерирани от 1 до 8. Вадим произволна топка и я връщаме в кутията. Отново вадим произволна топка. Каква е вероятността да извадим два пъти една и съща топка?

Задача 2. В кутия има 3 различни чифта чорапи. Вадим в тъмното 2 чорапа. Каква е вероятността извадените два чорапа да са чифт?

Задача 3. Иван има 4 ключа, но не знае кой е за неговата стая. Предполагаме, че ключовете са разбъркани по произволен начин. Иван пробва последователно с всеки от тях, като помни кой ключ е пробвал. Каква е вероятността да отключи с последния (четвъртия) ключ?

Задача 4. Студент се явява на изпит с конспект от 20 въпроса. От тях не знае само 3 въпроса. На изпита си тегли 2 въпроса от конспекта. Каква е вероятността да знае само един от изтеглените въпроси?

Задача 5. Каква е вероятността в група от 25 човека поне двама да имат рожден ден на един и същи ден от годината?

Задача 6. В отдел на фирма работят 20 човека. За Коледа те решават да си разменят подаръци. В кутия слагат 20 листчета, на всяко от които има едно име. Всеки тегли листче (без да го връща) и подарява на този, чието име е изтеглил. Каква е вероятността поне един да изтегли своето име?

Задача 7. На всеки от върховете на равностранен триъгълник има една мравка. Всяка мравка избира произволно един от другите два върха и тръгва към него. За единица време всяка мравка изминава разстоянието от един връх до друг. Две мравки могат да се разминат ако тръгнат една срещу друга. Каква е вероятността след единица време да има по една мравка на всеки връх?

Задача 8. В кутия има 6 сурови и 2 сварени яйца. Двама играчи, редувайки се, избират яйце докато извадят всички яйца. Намерете вероятностите на следните събития:

$A = \{$на един играч се падат двете сварени яйца$\}$;
$B = \{$пада се по едно сварено яйце на всеки играч$\}$;
$C = \{$падат се двете сварени яйца на този, който тегли първи$\}$;
$D = \{$падат се двете сварени яйца на този, който тегли втори$\}$.

Задача 9. На студенти е даден тест от 10 въпроса, всеки с по 4 възможни отговора, един от които е верен. Иван се явява на теста без да е учил и огражда произволно отговори. Каква е вероятността да е отговорил вярно на поне 5 от въпросите?

Задача 10. Авиокомпания е продала 143 билета за самолет, в който има 138 пътнически места. Вероятността пътник да дойде навреме за полета си е 0.92. Нека приемем, че даден пътник идва навреме независимо от останалите пътниците.

а) Каква е вероятността да има място за всички пътници, които са дошли навреме?

б) Каква е вероятността да остане едно незаето пътническо място?

Задача 11. В една кутия има 2 зелени и 2 червени топки. В друга кутия има 1 зелена и 4 червени топки. Хвърляме зар и ако се падне шестица, теглим топка от първата кутия, а ако не се падне шестица, теглим топка от втората кутия.

а) Каква е вероятността да извадим зелена топка?

б) Ако извадената топка е зелена, каква е вероятността да е извадена от втората кутия?

Задача 12. Разглеждаме три типа монети: тип T_{11} имат изписана единица от двете страни, тип T_{22} имат двойка от двете страни и тип T_{12} имат единица от едната страна и двойка от другата. В кутия има две монети T_{11}, една монета T_{22} и две монети T_{12}. Теглим произволна монета и я хвърляме.

а) Каква е вероятността да се падне единица?

б) Ако горната страна на хвърлената монета е единица, каква е вероятността другата страна да е двойка?

Задача 13. Имаме 3 карти: първата е бяла от двете страни, втората е черна от двете страни, а третата е бяла от едната и черна от другата страна. Всяка карта е поставена в затворена кутия. Избираме произволна кутия, отваряме я и виждаме, че горната страна на картата в нея е бяла. Каква е вероятността другата страна на картата също да е бяла?

Задача 14. В кутия има 99 топки номерирани от 1 до 99. Теглим без връщане 4 случайно избрани топки. Каква е вероятността първата извадена топка да е с най-голям номер от извадените?

Задача 15. Група от 20 човека, измежду които са Иван и Георги, е подредена по случаен начин в редица. Каква е вероятността Иван и Георги да са един до друг?

Задача 16. Тесте от 52 карти е разбъркано и е раздадено на 4 играчи. Каква е вероятността всеки играч да има едно асо?

Задача 17. На първия етаж на административна сграда 7 души чакат асансьора. Всеки от тях отива в някой от офисите в сградата. Сградата има 16 етажа и на всеки етаж има равен брой офиси (на първия етаж няма офиси).

а) Каква е вероятността поне двама от чакащите да отиват на един и същи етаж?

б) Ако Вие сте един от седемте, каква е вероятността поне един от останалите 6 да отива на Вашия етаж?

* * *

В следващите задачи намерете вероятностите като използвате вградените функции, свързани с дискретни разпределения.

Задача 18. Хвърляме зар 10 пъти.
а) Каква е вероятността да се паднат само 2 шестици?
б) Каква е вероятността да се паднат не повече от 2 шестици?
в) Каква е вероятността да се паднат 2 или повече шестици?
г) Каква е вероятността да се паднат между 3 и 8 шестици?

Задача 19. Хвърляме зар докато се падне шестица.
а) Каква е вероятността да хвърляме не повече от 10 пъти?
б) Каква е вероятността да хвърляме поне 6 пъти?

Задача 20. Хвърляме зар докато се паднат три шестици. Каква е вероятността да хвърляме не повече от 20 пъти?

Задача 21. Фенерче работи с 2 батерии. Иван има 8 батерии, от които 5 са нови и 3 са изтощени, но не знае кои точно. Ако пробва с 2 случайно избрани батерии, каква е вероятността фенерчето да не заработи?

Задача 22. Машинописка прави средно по една грешка на всеки 500 думи. На една страница има 300 думи.
а) Каква е вероятността да направи не повече от 2 грешки на 5 страници?
б) Каква е вероятността да направи между 1 и 3 грешки (включително) на 5 страници?

Задача 23. На студенти е даден тест от 10 въпроса, всеки с по 4 възможни отговора, един от които е верен. Иван се явява на теста без да е учил и огражда произволно отговори. Каква е вероятността да е отговорил вярно на поне 5 от въпросите?

Задача 24. Авиокомпания е продала 143 билета за самолет, в който има 138 пътнически места. Вероятността пътник да дойде навреме за полета си е 0.92. Нека приемем, че даден пътник идва навреме независимо от останалите пътниците.
а) Каква е вероятността да има място за всички пътници, които са дошли навреме?
б) Каква е вероятността да остане едно незаето пътническо място?

Задача 25. Батерия, произведена в даден завод, е дефектна с вероятност 0.03. Случайно избрани батерии се взети за проверка.
а) Каква е вероятността да се проверят не повече от 10 батерии докато се открие първата дефектна?
б) Каква е вероятността измежду първите 50 проверени батерии да има поне 2 дефектни?

Задача 26. В партида от 100 батерии има 3 дефектни. Избрани са 50 батерии за проверка. Каква е вероятността измежду избраните да има поне 2 дефектни?

Задача 27. В партида от 3000 батерии има 90 дефектни. Избрани са 50 батерии за проверка. Каква е вероятността измежду избраните да има поне 2 дефектни?

Задача 28. За клинично проучване са необходими доброволци, имащи определен ген, който се среща с вероятност $1/10$.

а) Каква е вероятността да се тестват 5 или повече доброволци докато се намери първия доброволец с въпросния ген?

б) Каква е вероятността да се тестват 50 или повече доброволци докато се намерят 10 доброволци с въпросния ген?

Задача 29. Средно веднъж на 90 дни в софийското метро възниква технически проблем, който води до спиране на движението на влаковете за поне 20 минути. Каква е вероятността за 360 дни да възникне такъв проблем повече от 3 пъти?

Задача 30. Теглим 10 случайно избрани карти от тесте с 52 карти (без връщане). Каква е вероятността да изтеглим поне 2 купи?

Задача 31. Теглим 10 пъти по една случайно избрана карта от тесте с 52 карти (с връщане). Каква е вероятността да изтеглим поне 2 купи?

* * *

Задача 32. Генерирайте 500 случайни числа от равномерно разпределение в интервала $(3,\, 5)$. Начертайте хистограма на генерираните числа и на същата картинка добавете графика на плътността на случайна величина $X \sim \mathrm{U}(3,\, 5)$. Повторете същото с 5000 случайни числа.

Задача 33. Генерирайте 500 случайни числа от експоненциално разпределение с параметър $\lambda = 1/7$. Начертайте хистограма на генерираните числа и на същата картинка добавете графика на плътността на случайна величина $X \sim \mathrm{Exp}(\lambda = 1/7)$. Повторете същото с 5000 случайни числа.

Задача 34. Генерирайте 500 случайни числа от нормално разпределение с параметри $\mu = 0$, $\sigma = 1$. Начертайте хистограма на генерираните числа и на същата картинка добавете графика на плътността на случайна величина $X \sim \mathcal{N}(\mu = 0,\, \sigma = 1)$. Повторете същото с 5000 случайни числа.

Задача 35. За $n = 200$ и $n = 1000$ генерирайте данни $x_1, x_2, \ldots, x_n$ от равномерно разпределение в интервала $(7,\, 9)$. Начертайте емпиричната функция на разпределение на данните. На същата картинка добавете функцията на разпределение на случайна величина $X \sim \mathrm{U}(7,\, 9)$.

Задача 36. За $n = 200$ и $n = 1000$ генерирайте данни $x_1, x_2, \ldots, x_n$ от експоненциално разпределение с параметър $\lambda = 3$. Начертайте емпиричната функция на разпределение на данните. На същата картинка добавете функцията на разпределение на случайна величина $X \sim \mathrm{Exp}(\lambda = 3)$.

Задача 37. За $n = 200$ и $n = 1000$ генерирайте данни $x_1, x_2, \ldots, x_n$ от нормално разпределение с параметри $\mu = 4$, $\sigma = 1.2$. Начертайте емпиричната функция на разпределение на данните. На същата картинка добавете функцията на разпределение на случайна величина $X \sim \mathcal{N}(\mu = 4,\, \sigma = 1.2)$.

Задача 38. Нека $X \sim \mathrm{U}(7,\, 9)$. Изобразете на графика плътността, функцията на разпределение и квантилната функция на X.

Задача 39. Нека $X \sim \mathrm{Exp}(\lambda = 3)$. Изобразете на графика плътността, функцията на разпределение и квантилната функция на X.

Задача 40. Нека $X \sim \mathcal{N}(\mu = 4,\, \sigma = 1.2)$. Изобразете на графика плътността, функцията на разпределение и квантилната функция на X.

Задача 41. Количеството (в мл) на портокалов сок в случайно избрана бутилка, произведена от дадена компания, е случайна величина, която има равномерно разпределение в интервала $(495,\, 502)$.

а) Намерете вероятността в случайно избрана бутилка да има повече от 500 мл сок.

б) Нека v е такова, че с вероятност 0.8 в случайно избрана бутилка има поне v мл сок. Намерете стойността на v.

Задача 42. Времето (в години) до повреда на пералня от даден модел е случайна величина X, която има експоненциално разпределение с параметър $\lambda = 1/4$.

а) Намерете вероятността пералня от този модел да не се повреди през първите 3 години.

б) Намерете вероятността пералня от този модел да се повреди през първите 2 години.

в) Ако пералня е работила без повреда 3 години, каква е вероятността да работи без повреда още 3 години?

г) Намерете t, за което $\mathbf{P}(X \le t) = 0.90$.

Задача 43. Количеството (в кг) кашкавал, което се изразходва в дадена пицария за една седмица, е нормално разпределено с параметри $\mu = 41$, $\sigma = 5$.

а) Каква е вероятността в произволно избрана седмица пицарията да изразходи над 51 кг кашкавал?

б) Каква е вероятността в произволно избрана седмица пицарията да изразходи между 45 и 50 кг кашкавал?

в) Колко кашкавал трябва да има в запас в дадена седмица, така че да е достатъчно с вероятност 0.99?

Задача 44. Намерете приближение на числото π като симулирате точки, попадащи равномерно във вътрешността на квадрат, и преброите каква част от тях попадат в кръга, вписан в квадрата.

Задача 45. Чрез подходящи симулации намерете приближение на интеграла

$$\int_{0.8}^{4} \frac{e^{-x^2/2}}{\sqrt{2\pi}}\, dx.$$

Задача 46. Генерирайте случайни числа $x_1, x_2, \ldots, x_{5000}$ от експоненциално разпределение с параметър $\lambda = 1/8$ и начертайте хистограма. Нека

$y_i = 1 - e^{-\lambda x_i}, \quad i = 1, \ldots, 5000.$

Начертайте хистограма на числата y_i.

Задача 47. Генерирайте случайни числа $x_1, x_2, \ldots, x_{5000}$ от равномерно разпределение в интервала (0, 1) и начертайте хистограма. Нека

$y_i = -(1/\lambda)\log(1 - x_i), \quad i = 1, \ldots, 5000, \quad \lambda = 1/8.$

Начертайте хистограма на числата y_i и на същата картинка добавете графика на плътността на експоненциално разпределение с параметър $\lambda = 1/8$.

Задача 48. Времето от зареждане до изтощаване на батерия на лаптоп при обичайна работа е нормално разпределено със средно 260 минути и стандартно отклонение 50 минути. Каква е вероятността батерия да се изтощи след повече от 4 часа работа? Каква е вероятността батерия да се изтощи след между 3 и 5 часа работа? Намерете t, такова че батерия се изтощава след повече от t минути с вероятност 0.9.

* * *

Задача 49. Разгледайте данните `survey` от пакета `MASS`. Представете таблично и графично данните за физическите упражнения (`Exer`).

Задача 50. (Данни `survey` от пакета `MASS`.) Представете чрез подходящи таблици и графики данните за пулса на студентите (`Pulse`).

Задача 51. (Данни `survey` от пакета `MASS`.) Представете чрез подходящи таблици и графики данните за възрастта на студентите (`Age`).

Задача 52. Генерирайте променливи `v1`, ..., `v5` като използвате следния код:

```
v1 <- rep(4, 30)
v2 <- rep(c(3.5,4.5), 15)
v3 <- rep(c(3,5), 15)
v4 <- rep(c(2:6), 6)
v5 <- rep(c(2,6), 15)
```

Представете графично всяка от променливите; намерете медианата, средната стойност и стандартното отклонение.

Задача 53. Във файла `cereals.RData` има данни за 77 зърнени закуски. Представете графично променливите `carbo`, `sodium` и `potass`; намерете медианата, средната стойност и стандартното отклонение.

Задача 54. (Данни `survey` от пакета `MASS`.) Чрез подходящи числови характеристики и графики покажете как пулсът се различава в зависимост от това дали студентът пише с лявата или с дясната ръка (`W.Hnd`).

Задача 55. (Данни `survey` от пакета `MASS`.) Намерете медианата, средната стойност и стандартното отклонение на:

а) пулса на студентите;
б) пулса на жените;
в) пулса на студентите на възраст не повече от 25 години;
г) пулса на студентите, правещи физически упражнения често;
д) пулса на студентите, които са непушачи и правят физически упражнения често.

Задача 56. (Данни `survey` от пакета `MASS`.) Чрез подходящи числови характеристики и графики покажете как пулсът се различава в зависимост от честотата на физически упражнения (`Exer`).

Задача 57. (Данни `survey` от пакета `MASS`.) Представете таблично и графично данните за пушенето (`Smoke`).

Задача 58. (Данни `survey` от пакета `MASS`.) Намерете медианата, средната стойност и стандартното отклонение на:

а) възрастта на студентите;
б) възрастта на пушачите;
в) възрастта на пишещите с дясната ръка;
г) възрастта на студентите с измерен пулс 70 и повече;
д) възрастта на студентите, които не правят физически упражнения.

Задача 59. (Данни `survey` от пакета `MASS`.) Представете графично височината на студентите. Чрез подходящи числови характеристики и графики сравнете височината на мъжете и жените.

* * *

Задача 60. За $n = 3, 7, 10, 30, 90, 200$ генерирайте данни $x_1, x_2, \ldots, x_n$ от експоненциално разпределение с параметър $\lambda = 1/5$ и пресметнете $\Sigma = x_1 + \ldots + x_n$. Повторете 10000 пъти.

$$
\begin{array}{llll}
x_1^{(1)} & x_1^{(2)} & \ldots & x_1^{(10000)} \\
x_2^{(1)} & x_2^{(2)} & \ldots & x_2^{(10000)} \\
\vdots & \vdots & & \vdots \\
x_n^{(1)} & x_n^{(2)} & \ldots & x_n^{(10000)} \\
\Sigma^{(1)} & \Sigma^{(2)} & \ldots & \Sigma^{(10000)}
\end{array}
$$

а) За всяко n начертайте хистограма на $\Sigma^{(1)}, \Sigma^{(2)}, \ldots, \Sigma^{(10000)}$.
б) За всяко n начертайте емпиричната функция на разпределение на $\Sigma^{(1)}, \Sigma^{(2)}, \ldots, \Sigma^{(10000)}$. На същата графика добавете функцията на разпределение на $\mathcal{N}(\mu = 5n,\ \sigma = 5\sqrt{n})$.

Задача 61. За $n = 3, 7, 10, 30, 90, 200$ генерирайте данни $x_1, x_2, \ldots, x_n$ от експоненциално разпределение с параметър $\lambda = 1/5$ и пресметнете $\overline{x} = (x_1 + \ldots + x_n)/n$. Повторете 10000 пъти.

$$
\begin{array}{llll}
x_1^{(1)} & x_1^{(2)} & \ldots & x_1^{(10000)} \\
x_2^{(1)} & x_2^{(2)} & \ldots & x_2^{(10000)} \\
\vdots & \vdots & & \vdots \\
x_n^{(1)} & x_n^{(2)} & \ldots & x_n^{(10000)} \\
\overline{x}^{(1)} & \overline{x}^{(2)} & \ldots & \overline{x}^{(10000)}
\end{array}
$$

а) За всяко n начертайте хистограма на $\overline{x}^{(1)}, \overline{x}^{(2)}, \ldots, \overline{x}^{(10000)}$.

б) За всяко n начертайте емпиричната функция на разпределение на $\overline{x}^{(1)}, \overline{x}^{(2)}, \ldots, \overline{x}^{(10000)}$. На същата графика добавете функцията на разпределение на $\mathcal{N}(\mu = 5,\ \sigma = 5/\sqrt{n})$.

Задача 62. За $n = 3, 7, 10, 30, 90, 200$ генерирайте данни $x_1, x_2, \ldots, x_n$ от Поасоново разпределение с параметър $\lambda = 3$ и пресметнете $\Sigma = x_1 + \ldots + x_n$. Повторете 10000 пъти.

а) За всяко n начертайте хистограма на $\Sigma^{(1)}, \Sigma^{(2)}, \ldots, \Sigma^{(10000)}$.

б) За всяко n начертайте емпиричната функция на разпределение на $\Sigma^{(1)}, \Sigma^{(2)}, \ldots, \Sigma^{(10000)}$. На същата графика добавете функцията на разпределение на $\mathcal{N}(\mu = 3n,\ \sigma = \sqrt{3n})$.

Задача 63. За $n = 3, 7, 10, 30, 90, 200$ генерирайте данни $x_1, x_2, \ldots, x_n$ от Поасоново разпределение с параметър $\lambda = 3$ и пресметнете $\overline{x} = (x_1 + \ldots + x_n)/n$. Повторете 10000 пъти.

а) За всяко n начертайте хистограма на $\overline{x}^{(1)}, \overline{x}^{(2)}, \ldots, \overline{x}^{(10000)}$.

б) За всяко n начертайте емпиричната функция на разпределение на $\overline{x}^{(1)}, \overline{x}^{(2)}, \ldots, \overline{x}^{(10000)}$. На същата графика добавете функцията на разпределение на $\mathcal{N}(\mu = 3,\ \sigma = \sqrt{3}/\sqrt{n})$.

Задача 64. Решете предходната задача като генерирате данни $x_1, \ldots, x_n$ от равномерно разпределение в интервала $(2,\ 8)$. Трябва да изберете подходящи стойности на μ и σ за б).

Задача 65. Времето на живот на електрическа крушка от даден тип има експоненциално разпределение със средно 900 часа. Измерено е времето на живот на 100 случайно избрани крушки. Каква е вероятността полученото средно време да е над 980 часа?

Задача 66. Времето на чакане (в секунди) на асансьор в сграда е равномерно разпределено в интервала $(0,\ 60)$. Ако засечем времето на чакане на 50 човека, каква е вероятността полученото средно време да е между 25 и 35 секунди?

Задача 67. Броят на стафидите в една кифличка, произведена в дадена пекарна, е случайна величина със следното разпределение:

брой	4	5	6	7
вероятност	0.2	0.4	0.3	0.1

Направена е извадка от 49 кифлички и са преброени стафидите във всяка. Каква е вероятността полученият среден брой стафиди да е над 5.5?

Задача 68. Регистрираният багаж на пътниците в даден самолет не трябва да надвишава общо 4000 кг. Количеството регистриран багаж на произволно избран пътник е случайна величина със средно 24 кг и стандартно отклонение 7 кг. Каква е вероятността общото количество регистриран багаж в самолет със 160 пътници да надвиши 4000 кг?

Задача 69. Броят на поръчките, постъпващи за ден в дадена фирма, има Поасоново разпределение със средно 5. Нека $X_1, \ldots, X_{80}$ са поръчките в 80 случайно избрани дни. Намерете $\mathbf{P}(4.5 < \overline{X} < 5.5)$.

* * *

Задача 70. Проведен е експеримент с монета. Според резултатите, при 100 хвърляния на монетата се паднало ези 58 пъти. Може ли въз основа на тези данни да се заключи, че вероятността да се падне ези при хвърляне на тази монета е повече от $1/2$?

Задача 71. Проведен е експеримент с монета. Според резултатите, при 100 хвърляния на монетата се паднало ези 61 пъти. Може ли въз основа на тези данни да се заключи, че вероятността да се падне ези при хвърляне на тази монета е повече от $1/2$?

Задача 72. Машина произвежда топки за тенис. Твърди се, че диаметърът на топка, произведена от машината, има нормално разпределение със средно $\mu = 6.7$ см и стандартно отклонение $\sigma = 0.12$ см. Измерен е диаметърът на 45 случайно избрани топки, произведени от машината. Средният диаметър на избраните топки е 6.73 см. Подкрепят ли тези данни, твърдението, че средният диаметър на топките, произведени от машината, е 6.7 см?

Задача 73. Машина произвежда топки за тенис. Твърди се, че диаметърът на топка, произведена от машината, има нормално разпределение със средно $\mu = 6.7$ см и стандартно отклонение $\sigma = 0.12$ см. Измерен е диаметърът на 45 случайно избрани топки, произведени от машината. Средният диаметър на избраните топки е 6.76 см. Подкрепят ли тези данни, твърдението, че средният диаметър на топките, произведени от машината, е 6.7 см?

Задача 74. Кадмият е тежък метал, който може да се натрупа до високи нива в гъбите. Измерена е концентрацията на кадмий в случайна извадка от 10 горски гъби събрани в дадена местност. Резултатите са следните (в мг/кг):

```
3.1  3.0  3.7  2.6  4.2  3.8  3.6  2.7  3.8  4.4
```

Може да предположим, че концентрацията на кадмий е нормално разпределена. Имаме ли основание да твърдим, че средната концентрация на кадмий в горските гъбите в дадената местност е по-малко от 4 мг/кг?

Задача 75. Проведен е експеримент, при който 58 домашни гълъба са пуснати на свобода от непознато място, което е на разстояние 106 км от дома им. От тях 32 намерили пътя към дома си. Преди провеждане на експеримента изследовател твърдял, че домашните гълъби намират пътя към дома си с вероятност над 51%. Подкрепят ли данните твърдението на изследователя?

Задача 76. Измерено е нивото на хемоглобин на 10 деца страдащи от специфична болест. Данните са:

```
12.3  11.2  14.2  15.3  14.8  13.5  11.1  15.1  15.4  13.2
```

Считаме, че нивото на хемоглобин има нормално разпределение.

а) Имаме ли основание да твърдим, че средното ниво на хемоглобин при децата страдащи от тази болест е различно от нормалното ниво от 14.6?

б) Имаме ли основание да твърдим, че средното ниво на хемоглобин при децата страдащи от тази болест е по-малко от 14.6?

Задача 77. Според проучване от 2009 година, около 7.5% от стоките в даден хипермаркет имат грешна цена на етикета. През 2010 е направено ново проучване и от 200 случайно избрани стоки, 14 били с грешна цена.

а) Имаме ли основание да твърдим, че през 2010 вероятността стока да е с грешна цена е различна от 7.5%?

б) Имаме ли основание да твърдим, че през 2010 вероятността стока да е с грешна цена е по-малка от 7.5%?

Задача 78. Измерена е дължината на дясното ухо на 66 случайно избрани жени на възраст между 18 и 30 години. Резултатите показват средна дължина 61.9 мм и стандартно отклонение 4 мм. Нека предположим, че популационната дисперсия е известна и $\sigma^2 = (4.1)^2$. Имаме ли основание да твърдим, че средната дължина на дясното ухо на жените на възраст между 18 и 30 години е различна от 60 мм?

Задача 79. Според производител на автомат за безалкохолни напитки, средното количество, което налива автоматът в една чаша е 170 грама със стандартно отклонение 3.9 грама. За да се провери това, е измерено количеството течност при 50 наливания. Резултатите показват средно 168 грама и същото стандартно отклонение. Може ли да се твърди, че в средно автоматът налива по-малко от 170 грама?

* * *

Задача 80. Преподавател дал два варианта (A и B) задачи на изпит. Резултатите от изпита са във файла `examAB.txt`. Може ли да се твърди, че в средно вариант B е по-труден от вариант A, т.е. в средно студентите получават по-малко точки ако им се е падне вариант B?

Задача 81. Проведен е експеримент, при който е измерено времето на реакция при 32 доброволци преди и след изпиване на 50 мл водка. Данните са във файла `reacttime.txt`. Може ли въз основа на данните да се заключи, че след употреба на 50 мл водка времето на реакция в средно се увеличава?

Задача 82. Направена е извадка от 200 болта, произведени от машина A и 200 болта, произведени от машина B. Дефектни се оказали 8 болта произведени от първата машина и 15 болта произведени от втората. Може ли да се твърди, че двете машини се различават по отношение на вероятността да произведат дефектен болт?

Задача 83. За всеки от следните случаи посочете дали данните представляват двойки наблюдения или са две независими извадки:

1) Избрани са 50 стоки, продавани в магазин за електроуреди, и за всяка стока е записана цената в магазин A и цената в магазин B.

2) За да се сравнят две марки електрически крушки е направен експеримент с по 65 крушки от всяка марка и е записано времето до изгаряне на всяка от крушките.

3) Оценките от първото и от второто контролно по *Увод в програмирането* на студентите от специалност КН.

4) Оценките по *Дискретни структури* на студентите от първа и от втора група от специалност КН.

Задача 84. Генерирайте данни $x_1, x_2, \ldots, x_n$ от $\mathcal{N}(\mu = 5,\ \sigma^2 = 1)$ и данни $y_1, y_2, \ldots, y_n$ от $\mathcal{N}(\mu = 5,\ \sigma^2 = 0.64)$. Проверете хипотезата за равенство на средните. Повторете 10000 пъти за $n = 20, 50, 100, 500$. В каква част от случаите нулевата хипотеза се отхвърля?

Задача 85. Генерирайте данни $x_1, x_2, \ldots, x_n$ от $\mathcal{N}(\mu = 5,\ \sigma^2 = 1)$ и данни $y_1, y_2, \ldots, y_n$ от $\mathcal{N}(\mu = 5.2,\ \sigma^2 = 1)$. Проверете хипотезата за равенство на средните. Повторете 10000 пъти за $n = 20, 50, 100, 500, 1000$. В каква част от случаите нулевата хипотеза се отхвърля?

Задача 86. При проучване направено през 2007 година от 500 анкетирани 26 били вегетарианци. През 2019 година е направено ново проучване, според което 43 от 540 анкетирани отговорили, че са вегетарианци. Дават ли ни тези данни основание да твърдим, че процентът на вегетарианците през 2019 година се е увеличил спрямо 2007 година?

Задача 87. За определяне на съдържанието на примеси в сплави на стомана съществуват два метода. В 8 проби е определено съдържанието на примеси по всеки от двата метода. Данните са следните:

Проба	*Метод 1*	*Метод 2*
1	1.2	1.4
2	1.3	1.7
3	1.5	1.5
4	1.4	1.3
5	1.7	2.0
6	1.8	2.1
7	1.4	1.7
8	1.3	1.6

Считаме, че данните са от нормално разпределение. Може ли да се приеме, че двата метода дават в средно едни и същи резултати?

Задача 88. За да се сравни скоростта на четене при различни шрифтове, на две групи от по 50 човека е даден един и същи текст, но на първата бил написан с един шрифт, а на втората с друг. Регистрирано е времето, за което всеки участник е прочел текста. Резултатите са: за първата група – средно 7.88 минути и стандартно отклонение 1.73; за втората група – средно 8.48 и стандартно отклонение 2.12. Може ли да се твърди, че средното време за четене е различно за двата шрифта?

* * *

Задача 89. Машина произвежда топки за тенис. Известно е, че диаметърът на топка, произведена от машината, е случайна величина със стандартно отклонение $\sigma = 0.12$ см. Измерен е диаметърът на 45 случайно избрани топки, произведени от машината.

а) Ако средният диаметър на избраните топки е 6.73 см, намерете 95-процентен доверителен интервал за средния диаметър на топките, произведени от машината.

б) Ако средният диаметър на избраните топки е 6.76 см, намерете 95-процентен доверителен интервал за средния диаметър на топките, произведени от машината.

Задача 90. Кадмият е тежък метал, който може да се натрупа до високи нива в гъбите. Измерена е концентрацията на кадмий в случайна извадка от 10 горски гъби събрани в дадена местност. Резултатите са следните (в мг/кг):

`3.1 3.0 3.7 2.6 4.2 3.8 3.6 2.7 3.8 4.4`

Може да предположим, че концентрацията на кадмий е нормално разпределена.

а) Намерете 95-процентен доверителен интервал за средната концентрация на кадмий в горските гъбите в дадената местност.

б) Намерете 90-процентен доверителен интервал за средната концентрация на кадмий в горските гъбите в дадената местност.

Задача 91. Проведен е експеримент с монета. При n хвърляния на монетата се паднало ези x пъти. Намерете 95-процентен доверителен интервал за вероятността да се падне ези при хвърляне на тази монета, ако

а) $n = 100$, $x = 58$;

б) $n = 200$, $x = 116$;

в) $n = 100$, $x = 61$.

Задача 92. Измерена е дължината на дясното ухо на 66 случайно избрани жени на възраст между 18 и 30 години. Резултатите показват средна дължина 61.9 мм и стандартно отклонение 4 мм. Нека предположим, че популационната дисперсия е известна и $\sigma^2 = (4.1)^2$.

а) Намерете 95-процентен доверителен интервал за средната дължина на дясното ухо на жените на възраст между 18 и 30 години.

б) Ако са направени измервания на 88 жени и са получени същите резултати за средната дължина и стандартното отклонение, намерете 95-процентен доверителен интервал за средната дължина на дясното ухо на жените на възраст между 18 и 30 години.

Задача 93. Проведен е експеримент, при който 58 домашни гълъба са пуснати на свобода от непознато място, което е на разстояние 106 км от дома им. От тях 32 намерили пътя към дома си. Намерете 95-процентен доверителен интервал за вероятността гълъб от дадения вид да намери пътя към дома си от непознато място на 106 км.

Задача 94. Генерирайте данни $x_1, x_2, \ldots, x_n$ от равномерно разпределение в интервала $(5,\ 9)$. Намерете 95-процентен доверителен интервал за средното μ. Повторете 10000 пъти за $n = 20, 50, 100, 500$. В каква част от случаите доверителният интервал съдържа 7?

Задача 95. Генерирайте данни $x_1, x_2, \ldots, x_n$ от равномерно разпределение в интервала $(5,\ 9)$. Проверете хипотезата $\mu = 7$ срещу $\mu \neq 7$. Повторете 10000 пъти за $n = 20, 50, 100, 500$. В каква част от случаите нулевата хипотеза не се отхвърля?

Задача 96. Генерирайте данни $x_1, x_2, \ldots, x_n$ от равномерно разпределение в интервала $(5,\ 9)$. Намерете 95-процентен доверителен интервал за средното μ. Проверете хипотезата $\mu = 7$ срещу $\mu \neq 7$. Повторете 10000 пъти за $n = 20, 50, 100, 500$. В каква част от случаите *доверителният интервал съдържа* 7 и едновременно с това *нулевата хипотеза не се отхвърля*?

Задача 97. Като са използвани данните от 50 измервания на дадена величина е намерен 95-процентен доверителен интервал [25.0128, 26.0212] за средното μ. Искаме да проверим хипотезата $H_0 : \mu = 25$ срещу двустранна алтернатива при ниво на значимост $\alpha = 0.05$. Какъв извод ще направим (отхвърляме ли нулевата хипотеза)?

* * *

Задача 98. Резултатите от 180 хвърляния на зар са дадени в таблицата:

1	2	3	4	5	6
28	36	36	30	27	23

Може ли въз основа на тези данни да се заключи, че зарът е балансиран?

Задача 99. Честотите на срещане на буквите в английския език са следните (в %):

E	T	A	O	I	N	S	R	H	D	other
12.02	9.10	8.12	7.68	7.31	6.95	6.28	6.02	5.92	4.32	26.28

В текст, състоящ се от 2004 букви, броят срещания на съответните букви е:

E	T	A	O	I	N	S	R	H	D	other
221	153	183	111	113	152	103	197	38	104	629

Подкрепят ли тези данни твърдението, че текстът е на английски?

Задача 100. Според законите на Мендел, даден сорт грах може да има бели, розови или червени цветове, с вероятности съответно 1/4, 1/2, 1/4. Направена е извадка от 564 растения от този сорт. От тях 141 цъфтели в бяло, 291 в розово и 132 в червено. В съгласие ли са тези резултати с теорията на Мендел?

Задача 101. В данните `pi2000` са първите 2000 цифри на числото π. Може ли да се твърди, че всяка цифра се среща с една и съща вероятност?

Задача 102. Използвайте данните `survey` от пакета `MASS` и направете двумерна таблица на пушенето по пола. Проверете хипотезата за независимост между пушенето и пола.

Задача 103. Във файла `ManWomanEye.txt` има данни за цвета на очите на 204 семейни двойки. Има ли връзка между цвета на очите на мъжа и цвета на очите на жената?

Задача 104. Разгледайте данните `HairEyeColor`. Има ли връзка между цвета на косата и цвета на очите?

Задача 105. В таблицата са представени данни за броя на пострадалите при катастрофи пътници, като са класифицирани по степен на нараняванията и по това дали пътникът е използвал препазен колан.

	Наранявания			
	Няма	Леки	Средни	Тежки
С колан	12813	647	359	42
Без колан	65963	4000	2642	303

Има ли връзка между използването на предпазен колан и степента на нараняванията?

Задача 106. Симулирайте n хвърляния на балансиран зар. Като използвате генерираните данни, проверете хипотезата, че всяка от страните се пада с една и съща вероятност. Повторете 10000 пъти за $n = 100, 200, 400$. Колко често заключението на теста е вярно?

* * *

Задача 107. Генерирайте данни (x_i, y_i), $i = 1, \ldots, 100$, като x_i са случайни числа от нормално разпределение $\mathcal{N}(\mu = 5,\, \sigma^2 = 1)$ и

а) $y_i = 2x_i$;

б) $y_i = 2x_i + \varepsilon_i$, където ε_i са случайни числа от $\mathcal{N}(\mu = 0,\, \sigma^2 = 1)$;

в) $y_i = 2x_i + \varepsilon_i$, където ε_i са случайни числа от $\mathcal{N}(\mu = 0,\, \sigma^2 = 4)$;

г) $y_i = 0.1x_i + \varepsilon_i$, където ε_i са случайни числа от $\mathcal{N}(\mu = 0,\, \sigma^2 = 4)$;

д) y_i са случайни числа от $\mathcal{N}(\mu = 5,\, \sigma^2 = 1)$;

е) $y_i = -2x_i + \varepsilon_i$, където ε_i са случайни числа от $\mathcal{N}(\mu = 0,\, \sigma^2 = 1)$.

За всеки от случаите представете данните графично и намерете корелацията.

Задача 108. Във файла `bac.txt` има данни за съдържанието на алкохол в кръвта (грамове алкохол на 100 мл кръв) в зависимост от броя на изпитите бутилки бира.

а) Постройте линеен модел. Напишете оцененото регресионно уравнение. Представете графично данните и построения линеен модел.

б) Интерпретирайте оценените коефициенти.

в) Имаме ли основание да твърдим, че има линейна връзка между броя на изпитите бутилки бира и съдържанието на алкохол в кръвта?

г) Може ли да се твърди, че при изпиването на още една бира съдържанието на алкохол в кръвта се увеличава средно с 0.02?

д) Намерете доверителен интервал за средното съдържание на алкохол в кръвта при 5 изпити бутилки бира.

е) За стойности на предиктора между 1 и 9, намерете доверителен интервал за средното съдържание на алкохол в кръвта и интервал за прогноза и ги илюстрирайте на графиката от а).

Задача 109. Генерирайте данни (x_i, y_i), $i = 1, \ldots, 50$, където:

x_i са случайни числа от равномерно разпределение в интервала $(1, 7)$;
ε_i са случайни числа от нормално разпределение с параметри $\mu = 0$ и σ;
$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$.

Разгледайте следните случаи за β_0, β_1, σ:

а) $\beta_0 = 2, \quad \beta_1 = 1.5, \quad \sigma = 2.$

б) $\beta_0 = 2, \quad \beta_1 = 1.5, \quad \sigma = 1.$

в) $\beta_0 = 2, \quad \beta_1 = 0.17, \quad \sigma = 1.$

Постройте линеен модел по данните (x_i, y_i), $i = 1, \ldots, 50$. Представете графично данните и построения линеен модел. Намерете R^2 и доверителен интервал за β_1.

Задача 110. Генерирайте данни $(x_i, y_i),\ i = 1, \ldots, 50$, където:

x_i са случайни числа от равномерно разпределение в интервала $(1, 7)$;
ε_i са случайни числа от нормално разпределение с параметри $\mu = 0$ и $\sigma = 2.5$;
$y_i = \beta_0 + \beta_1 x_i^2 + \varepsilon_i$;
$\beta_0 = 2, \quad \beta_1 = 1.1$.

Постройте линеен модел по данните $(x_i, y_i),\ i = 1, \ldots, 50$. Представете графично данните и построения линеен модел. Намерете R^2.

Задача 111. Във файла `satgpa.txt` има данни за резултатите на 1000 студенти. Разгледайте следните променливи:

sat_v резултат на Verbal SAT;
sat_m резултат на Math SAT;
sat_sum общ резултат на теста SAT;
hs_gpa успех от гимназията;
fy_gpa успех от първата година в колежа.

а) Постройте модел, който може да се използва за прогнозиране на успеха от първата година в колежа в зависимост от успеха от гимназията. Напишете оцененото регресионно уравнение. Направете подходяща графика.

б) Интерпретирайте коефициента пред предиктора. Може ли да се твърди, че има линейна връзка между двете променливи?

в) Прогнозирайте успеха от първата година в колежа на студент, чиито успех от гимназията е 3.5. Намерете доверителен интервал и интервал за прогноза.

г) Може ли да се получи по-добър модел, ако се включат и някои от останалите променливи?

Задача 112. Измерени са диаметъра (в инчове), височината (във футове) и обема (в кубични футове) на 31 черешови дървета. Диаметърът е измерен на височина 54 инча от земята. Данните са във файла `cherry.txt`.

а) Постройте модел, който може да се използва за прогнозиране на обема според диаметъра на дървото. Напишете оцененото регресионно уравнение. Направете подходяща графика.

б) Интерпретирайте коефициента пред диаметъра.

в) Постройте модел, който включва и височината. Интерпретирайте коефициента пред височината и коефициента пред диаметъра.

г) Намерете доверителен интервал за средния обем на дърво с диаметър 14 инча и височина 70 фута.

д) Постройте модел, в който участва диаметърът на квадрат. Намерете доверителен интервал за средния обем на дърво с диаметър 14 инча и височина 70 фута и сравнете с резултата от г).

Задача 113. Във файла `duke_forest.csv` има данни за 98 продадени къщи. Разгледайте следните променливи:

price цена (в долари);
bed брой спални;
bath брой бани;
area площ (в квадратни футове);
year_built година на построяване;
lot площ на дворното място (в акри).

Постройте модел, който може да се използва за прогнозиране на цената според характеристиките на къщата.

Задача 114. Генерирайте данни (x_i, y_i), $i = 1, \ldots, n$, където:

x_i са случайни числа от равномерно разпределение в интервала $(1, 10)$;
$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$;
$\beta_0 = 2, \quad \beta_1 = 5.$

Разгледайте следните случаи за ε_i:

а) ε_i са случайни числа от нормално разпределение с параметри $\mu = 0$, $\sigma = 5$;
б) ε_i са случайни числа от експоненциално разпределение с параметър $1/5$.

Постройте линеен модел по данните (x_i, y_i), $i = 1, \ldots, n$. Повторете 10000 пъти за $n = 30, 50, 100, 500$. За всеки от случаите и всяко n намерете:

- средното на $\widehat{\beta}_1$ на базата на 10000 повторения;
- колко често доверителният интервал за β_1 съдържа истинската стойност;
- средната дължина на доверителния интервал за β_1 на базата на 10000 повторения.

* * *

* * *

Задача 115. Нека $x_1, \ldots, x_{50}$ са наблюдения над случайна величина X и нека $\mu = \mathrm{E}(X)$. Наблюденията са записани във вектор `x` и е получен следният резултат от R:

```
One Sample t-test
data:  x
t = -1.7012, df = 49, p-value = 0.09524
alternative hypothesis: true mean is not equal to 1.1
95 percent confidence interval:
 0.7998512 1.1249405
```

а) Илюстрирайте P-стойността (P-value) на графика (на ръка).
б) Колко е P-стойността, ако алтернативната хипотеза е $H_1 : \mu < 1.1$?
в) Колко е P-стойността, ако алтернативната хипотеза е $H_1 : \mu > 1.1$?
г) Колко е стойността на $\overline{x} = (x_1 + \ldots + x_{50})/50$?
д) Искаме да проверим хипотезата $H_0 : \mu = 1.23$ срещу $H_1 : \mu \neq 1.23$. Имаме ли основание да отхвърлим нулевата хипотеза?

Задача 116. За всеки от 70 случайно избрани студенти от даден университет са измерени две променливи: X и Y. Данните са записани съответно във векторите `x` и `y`. Получен е следният резултат от R:

```
> m1 <- lm(y ~ x)
> confint(m1)
                  2.5 %    97.5 %
(Intercept)  1.02406175 4.7591657
x           -0.04332078 0.7999039
```

Може ли въз основа на тези данни да заключим, че има линейна връзка между X и Y?

Задача 117. Иван хвърля монета 200 пъти, за да провери дали е балансирана. Нека p е вероятността да се падне ези. Нека X е броят на падания на ези при 200 хвърляния. Той използва следното правило:

Ако $86 \leq X \leq 114$, *приема, че* $p = 0.5$.
Ако $X < 86$ *или* $X > 114$, *приема, че* $p \neq 0.5$.

а) Ако $p = 0.5$, каква е вероятността да приеме, че $p \neq 0.5$, използвайки правилото?

б) Ако $p = 0.61$, каква е вероятността да приеме, че $p \neq 0.5$, използвайки правилото?

Задача 118. За да се установи времето, необходимо на ученик да прочете даден текст, е направен тест с 30 ученици. Резултатите са следните (в минути):

25 29 18 29 22 20 27 24 20 29 18 20 31 25 21 24 24 21 18 24 24 29 25 24 27 22 25 22 27 25

Дават ли ни тези данни основание да твърдим, че средното време, необходимо за прочитане на текста, е по-малко от 25 минути?

Задача 119. Разполагаме с данни за земетресенията в Южна Калифорния по ден от седмицата, за определен период от време (включени са земетресения с магнитуд поне 4.4 по Рихтер):

Ден от седмицата	пн	вт	ср	чт	пт	сб	нд
Брой земетресения	144	170	158	172	148	152	156

Подкрепят ли тези данни твърдението, че земетресение е равновероятно да се случи в кой да е ден от седмицата?

Задача 120. Извършен е следният експеримент с цел да се сравни добива от два сорта домати. Избрани са случайно по 36 растения от всеки сорт и е записан добива от всяко растение. Данните са във файла `tomato2.txt`. Имаме ли основание да твърдим, че средният добив от втория сорт е по-голям в сравнение с първия сорт?

Задача 121. За да се сравни надеждността на две марки електромотори, са тествани по 30 мотора от всяка марка. От първата марка 22 мотора преминали теста успешно, докато от втората марка успешно преминали теста 16 мотора. Може ли да заключим, че вероятността моторите от първата марка да преминат успешно теста е по-голяма отколкото вероятността за моторите от втората марка?

Задача 122. За да се изследва доколко се различават цените в две книжарници, са избрани 73 книги и за всяка книга е записана цената в първата книжарница и цената във втората книжарница. Данните са във файла `books.txt`. Имаме ли основание да твърдим, че в средно в първата книжарница цените са по-високи отколкото във втората?

Задача 123. Според скорошно изследване, направено сред 500 финландци, 89 от участниците имали непоносимост към лактоза. Според по-стари данни, 17 на сто от финландците имат непоносимост към лактоза. Имаме ли основание да направим заключението, че делът на финландците с непоносимост към лактоза е повече от 17 на сто?

* * *