Теория вероятностей и статистика, МИРЭК, 2023-2024

Дедлайн: решение домашнего задания загружается в виде единого файла, имеющего pdf-формат, в систему SmartLMS в разделе с соответствующим размещенным заданием. При наличии сбоев в работе системы файл необходимо направить на почту mirectvis@gmail.com. Тема письма должна иметь следующий формат: “МИРЭК Фамилия Имя Группа Номер ДЗ”, например, “МИРЭК Потанин Богдан 200 ДЗ 1”.

Оформление: первый лист задания должен быть титульным и содержать лишь информацию об имени и фамилии, а также о номере группы студента и сдаваемого домашнего задания. Если pdf файл содержит фотографии, то они должны быть разборчивыми и повернуты правильной стороной.

Санкции: домашние задания, не удовлетворяющие требованиям к оформлению, выполненные не самостоятельно или сданные позже срока получают 0 баллов.

Проверка: при оценивании каждого задания проверяется не ответ, а весь ход решения, который должен быть описан подробно и формально, с использованием надлежащих определений, обозначений, теорем и т.д.

Самостоятельность: задания выполняются самостоятельно. С целью проверки самостоятельности выполнения домашнего задания студент может быть вызван на устное собеседование, по результатам которого оценка может быть либо сохранена, либо обнулена.

Домашнее задание №1

Задание №1. Строители (25 баллов)

Лаврентий приобрел новостройку в жилом комплексе Сказочная Статистика. По камерам он может наблюдать за строительством комплекса, которым занимается бригада из 10 строителей, один из которых является прорабом, а двое других – его заместителями. Лаврентий по камерам видит 6 строителей. Каждый из строителей мог с равной вероятностью попасть на камеру. Посчитайте вероятность, с которой среди увиденных Лаврентием строителей:

1. Есть прораб. **(2 балл)**
2. Есть хотя бы один заместитель прораба. **(3 балла)**
3. Есть прораб и хотя бы один заместитель прораба. **(5 балла)**
4. Есть заместитель прораба, если на камеру попало не менее 4 обычных строителей (не являющихся прорабом или его заместителями). **(10 балла)**
5. Запишите функцию распределения случайной величины, отражающей число заместителей прораба, попавших на камеру. **(5 балла)**

Решение:

1. Обозначим через G_i событие, в соответствии с которым i-й строитель оказался прорабом, где $i \in \{1, 2, ..., 10\}$. Через Z_i и R_i обозначим аналогичные события для заместителей прораба и обычных рабочих соответственно. Поскольку прораб всего один, то события R_i являются несовместными, откуда, обозначая рассматриваемое событие через G, получаем:

$$P(G) = P(G_1 \cup G_2 \cup G_3 \cup G_4 \cup G_5 \cup G_6) = P(G_1) + ... + P(G_6) = \frac{6}{10} = 0.6$$

2. Воспользуемся комбинаторикой. Обратим внимание, что рассматриваемому событию удовлетворяют случаи, когда имеется только один заместитель, либо когда их двое. Количество способов выбрать 6 рабочих таким образом, чтобы среди них оказался ровно один заместитель, составляет $C_2^1 C_8^5$. По аналогии при двух заместителях получаем $C_2^2 C_8^4$ способов. Поскольку всего имеется C_{10}^6 способов выбрать 6 строителей из 10, то искомая вероятность рассматриваемого события, которое для удобства обозначим как A, составляет:

$$P(A) = \frac{C_2^1 C_8^5 + C_2^2 C_8^4}{C_{10}^6} = \frac{13}{15}$$

   В качестве более простой альтернативы можно воспользоваться вероятностью обратного события:

$$P(A) = 1 - P(\overline{A}) = 1 - \frac{C_8^6}{C_{10}^6} = 1 - \frac{2}{15} = \frac{13}{15}$$

3. Обозначим рассматриваемое событие через B и найдем его вероятность по аналогии с предыдущим пунктом:

$$P\left(B\right)=\frac{C_1^1C_2^1C_7^4+C_1^1C_2^2C_7^3}{C_{10}^6}=0.5$$

В качестве альтернативы удобно воспользоваться формулой условной вероятности:

$$P(B)=P(A\cap G)=P(A|G)P(G)=\left(1-\frac{C_7^5}{C_9^5}\right)\times 0.6=0.5 \tag{1}$$

4. Через C обозначим событие, в соответствии с которым на камеру попали хотя бы 4 обычных рабочих. Воспользуемся формулой условной вероятности:

$$P\left(A|C\right)=\frac{P\left(A\cap C\right)}{P\left(C\right)}=$$
$$=\frac{C_2^2C_7^4+C_2^1C_7^4+C_2^1C_7^5}{C_{10}^6}/\frac{C_2^2C_7^4+C_2^1C_7^4+C_2^1C_7^5+C_2^0C_7^5+C_7^6}{C_{10}^6}=$$
$$=\frac{C_2^2C_7^4+C_2^1C_7^4+C_2^1C_7^5}{C_2^2C_7^4+C_2^1C_7^4+C_2^1C_7^5+C_2^0C_7^5+C_7^6}=0.84$$

5. Обозначим через X случайную величину, отражающую число попавших на камеру заместителей прораба. Поскольку на камеру могло попасть от 0 до 2 заместителей, то $\mathrm{supp}(X)=\{0,1,2\}$, откуда:

$$P(X=0)=\frac{C_7^6+C_7^5C_2^0}{C_{10}^6}=\frac{2}{15}$$
$$P(X=1)=\frac{C_7^5C_2^1+C_7^4C_2^1}{C_{10}^6}=\frac{8}{15}$$
$$P(X=2)=\frac{C_7^4C_2^2+C_7^3}{C_{10}^6}=\frac{5}{15}=\frac{1}{3}$$

Отсюда получаем:

$$F_X(x)=\begin{cases}0,\text{ если } x<0\\ \frac{2}{15},\text{ если } 0\leq x<1\\ \frac{2}{3},\text{ если } 1\leq x<2\\ 1,\text{ если } x\geq 2\end{cases}$$

Проверка в R:

```
workers <- c(rep("r", 7), "z", "z", "g")
n.workers <- length(workers)
n.sim <- 10000
event.G <- vector(mode = "logical", length = n.sim)
event.A <- event.G
```

```
event.B <- event.G
event.C <- event.G
n.A <- event.G
for (i in 1:n.sim)
{
  selected <- sample(workers, size = 6)
  event.G[i] <- "g" %in% selected
  event.A[i] <- "z" %in% selected
  event.B[i] <- event.G[i] & ("z" %in% selected)
  event.C[i] <- n.C[i] >= 4
  n.A[i] <- sum(selected == "z")
}
A <- function(n, k)
{
  factorial(n) / factorial(n - k)
}
C <- function(n, k)
{
  A(n, k) / factorial(k)
}
# пункт 1
mean(event.G)
# пункт 2
(C(2, 1) * C(8, 5) + C(2, 2) * C(8, 4)) / C(10, 6)
1 - C(8, 6) / C(10, 6)
mean(event.A)
# пункт 3
(C(2, 1) * C(7, 4) + C(2, 2) * C(7, 3)) / C(10, 6)
(1 - C(7, 5) / C(9, 5)) * 0.6
mean(event.B)
# пункт 4
(C(2, 2) * C(7, 4) + C(2, 1) * C(7, 4) + C(2, 1) * C(7, 5)) /
  (C(2, 2) * C(7, 4) + C(2, 1) * C(7, 4) + C(2, 1) * C(7, 5) +
     C(2, 0) * C(7, 5) + C(7, 6))
mean(event.A[event.C])
# пункт 5
p0 <- (C(7, 6) + C(7, 5) * C(2, 0)) / C(10, 6)
p1 <- (C(7, 5) * C(2, 1) + C(7, 4) * C(2, 1)) / C(10, 6)
p2 <- (C(7, 4) * C(2, 2) + C(7, 3)) / C(10, 6)
cbind(cumsum(c(p0, p1, p2)),
      c(mean(n.A <= 0), mean(n.A <= 1), mean(n.A <= 2)))
```

Задание №2. Два гнома (25 баллов)

Два гнома с разных концов шахты прорубают путь навстречу друг другу. Между гномами расположены комнаты с золотыми монетами (см. таблицу). Комнаты разделены между собой стенками. Гномы поочередно наносят удары по стенкам между комнатами. Первым удар наносит первый гном. От каждого удара первого гнома, независимо от результатов предыдущих ударов, стена ломается с вероятностью 0.6. Второй гном гораздо сильнее, поэтому каждый его удар гарантированно ломает стену. Гном, которому удалось сломать стену, попадает в комнату с сокровищами и забирает себе все ее содержимое (если этого еще не сделал другой гном).

Гном 1	1	2	0	Гном 2

Таблица 1: Числа указывают количество золота в комнате.

Мотивирующая музыка: www.youtube.com/watch?v=ytWz0qVvBZ0

1. Посчитайте, с какой вероятностью первый гном добудет ровно одну золоту монету. **(5 баллов)**

2. Составьте таблицу распределения золота, добытого первым гномом. **(5 баллов)**

3. Найдите дисперсию золота, добытого первым гномом. **(5 баллов)**

4. Посчитайте ковариацию между золотом, добытым первым и вторым гномами.**(10 баллов)**

   Подсказка: выразите золото, добытое вторым гномом, как функцию от количества золота, добытого первым гномом.

Решение:

1. Обозначим через X и Y количество золота, добытого первым и вторым гномами соответственно. Первый гном добудет 1 золотую монету, если ему удастся пробить первую стену с третьего раза, либо если он сразу успешно пробьет первую стену, но не сможет с первого раза пробить вторую, либо если он пробьет первую стену со второго раза. Пользуясь несовместностью данных событий получаем:

$$P(X = 1) = 0.4^2 \times 0.6 + 0.6 \times 0.4 + 0.4 \times 0.6 = 0.576$$

2. Первый гном не сможет добыть золота, если ему 3 раза подряд не удастся пробить первую стену, вероятность чего составляет:

$$P(X = 0) = (1 - 0.6)^3 = 0.064$$

   Первый гном добудет 3 золотые монеты, если ему удастся пробить обе стены с первого раза, вероятность чего составляет:

$$P(X = 3) = 0.6 \times 0.6 = 0.36$$

   В результате получаем таблицу распределения:

x	0	1	3
P(X = x)	0.064	0.576	0.36

3. Используя найденную ранее таблицу распределения получаем:

$$E(X) = 0 \times 0.064 + 1 \times 0.576 + 3 \times 0.36 = 1.656$$
$$E(X^2) = 0^2 \times 0.064 + 1^2 \times 0.576 + 3^2 \times 0.36 = 3.816$$
$$Var(X) = E(X^2) - E(X)^2 = 3.816 - 1.656^2 = 1.073664$$

4. Поскольку $Y = 3 - X$, то:

$$Cov(X, Y) = Cov(X, 3 - X) = -Var(X) = -1.073664$$

Задание №3. Пинокио (25 баллов)

Длина носа Пинокио увеличивается на 10% каждый раз, когда он говорит неправду. Пинокио отвечает на вопрос правду с вероятностью 0.2, независимо от того, как он отвечал на предыдущие вопросы. Изначально длина носа Пинокио составляла 5 сантиметров. Пинокио задали 8 вопросов.

1. Найдите математическое ожидание длины носа Пинокио. **(5 баллов)**

2. Посчитайте дисперсию длины носа Пинокио. **(5 баллов)**

3. Вычислите ковариацию между длиной носа Пинокио и числом правдивых ответов.**(5 баллов)**

4. Пинокио отвечает на вопросы до тех пор, пока не ответит правду. Найдите математическое ожидание длины его носа. **(10 баллов)**

Решение:

1. Обозначим через X_i случайную величину, отражающую то, во сколько раз увеличивается длина носа Пинокио после ответа на i-й вопрос, где $i \in \{1, 2, ..., 8\}$. Найдем математическое ожидание этой случайной величины:

$$E(X_i) = P(X_i = 1.1) \times 1.1 + P(X_i = 1) \times 1 = 0.8 \times 1.1 + 0.2 \times 1 = 1.08$$

Обратим внимание, что случайную величину X, отражающую длину носа Пинокио после 8 ответов, можно записать как:

$$X = 5 \times X_1 \times X_2 \times ... \times X_8$$

Вычислим математическое ожидание этой случайной величины, учитывая, что X_i независимы:

$$E(X) = E(5 \times X_1 \times X_2 \times ... \times X_8) = 5 \times E(X_1) \times ... \times E(X_8) =$$
$$= 5 \times \underbrace{1.08 \times 1.08 \times ... \times 1.08}_{8 \text{ раз}} = 5 \times 1.08^8 \approx 9.25$$

2. По аналогии найдем второй начальный момент:

$$E(X_i^2) = P(X_i = 1.1) \times 1.1^2 + P(X_i = 1) \times 1^2 = 0.8 \times 1.1^2 + 0.2 \times 1^2 = 1.168$$
$$E((5 \times X_1 \times X_2 \times ... \times X_8)^2) = E(5^2 \times X_1^2 \times X_2^2 \times ... \times X_8^2) =$$
$$= 5^2 \times E(X_1^2) \times ... \times E(X_8^2) = 5^2 \times 1.168^8 \approx 86.6$$

В итоге найдем дисперсию:

$$Var(X) = E(X^2) - E(X)^2 \approx 86.6 - 9.25^2 \approx 1.0375$$

3. Обратим внимание, что случайная величина $10(X_i - 1)$ принимает значение 1, если ответа на i-й вопрос длина носа Пинокио увеличилась и значение 0 – в противном случае. В результате число вопросов, после ответа на которые длина носа Пинокио увеличился, может быть представлено как:

$$10(X_1 - 1) + ... + 10(X_8 - 1) = 10(X_1 + ... + X_8) - 80$$

   Рассчитаем ковариацию:

$$\begin{gathered}
Cov\,(5 \times X_1 \times X_2 \times ... \times X_8, 10(X_1 + ... + X_8) - 80) = \\
= (5 \times 10)\, Cov(X_1 \times X_2 \times ... \times X_8, X_1 + ... + X_8) = \\
= 50Cov(X_1 \times X_2 \times ... \times X_8, X_1 + ... + X_8) = \\
= 50Cov(X_1 \times X_2 \times ... \times X_8, X_1) + ... + 50Cov(X_1 \times X_2 \times ... \times X_8, X_8) = \\
= \underbrace{8 \times 50Cov(X_1 \times X_2 \times ... \times X_8, X_1)}_{\text{поскольку очевидно, что все суммируемые ковариации одинаковы}} = \\
= \underbrace{8 \times 50\left(E(X_1^2 \times X_2 \times ... \times X_8) - E(X_1 \times X_2 \times ... \times X_8)E(X_1)\right)}_{\text{по формуле } Cov(X,Y)=E(XY)-E(X)E(Y)} = \\
= 400\left(E(X_1^2)E(X_2) \times ... \times E(X_8) - E(X_1)^2 \times E(X_2) \times ... \times E(X_8)\right) = \\
= 400 \times \left(1.168 \times 1.08^8 - 1.08^9\right) \approx 65.15
\end{gathered}$$

4. Воспользуемся методом первого шага. Обозначим через Y длину носа Пинокио через 1 день. Через A обозначим событие, при котором на первый же вопрос пинокио дает правдивый ответ. Отсюда получаем:

$$E(Y) = E(Y|A)P(A) + E(Y|\overline{A})P(\overline{A}) = 5 \times 0.2 + 1.1E(Y) \times 0.8$$

   Решая соответствующее равенство получаем $E(Y) \approx 8.33$.

Задание №4. Кубики (25 баллов)

Имеются два шестигранных кубиках. Один из них правильный, а другой - нет. У неправильного кубика вероятность выпадения четного числа вдвое больше вероятности выпадения нечетного. При этом все числа одной и той же четности выпадают с равной вероятностью. Лаврентий с равной вероятностью берет и бросает один из этих кубиков.

1. Найдите вероятность, с которой выпадет четное число. **(5 баллов)**
2. Найдите вероятность, с которой Лаврентий взял правильный кубик, если выпало четное число. **(5 баллов)**
3. Найдите математическое ожидание числа выпавших очков. **(5 баллов)**
4. Лаврентий усложняет правила игры. Сперва он как и ранее с равной вероятностью выбирает один из кубиков и бросает его. Затем, если выпадает четное число, то он бросает тот же самый кубик, а в противном случае – другой. Посчитайте вероятность, с которой при втором броске выпадет четное число. **(5 баллов)**
5. При использовании правил игры из предыдущего пункта посчитайте, с какой вероятностью Лаврентий бросил в первый раз правильный кубик, если при втором броске выпало четное число. **(5 баллов)**

Решение:

1. Через A обозначим событие, при котором Лаврентий выбрал правильный кубик. Через O обозначим событие, при котором выпадает четное число. При помощи формулы полной вероятности получаем:

$$P(O) = P(O|A)P(A) + P(O|\overline{A})P(\overline{A}) =$$
$$= \frac{1}{2} \times \frac{1}{2} + \frac{2}{3} \times \frac{1}{2} = \frac{7}{12}$$

2. Применим формулу условной вероятности

$$P(A|O) = \frac{P(O|A)P(A)}{P(O)} = \frac{\frac{1}{2} \times \frac{1}{2}}{\frac{7}{12}} = \frac{3}{7}$$

3. Обозначим через X случайную величину, отражающую число выпавших очков. Обратим внимание, что:

$$E(X|A) = \frac{1}{6} \times (1+2+3+4+5+6) = 3.5$$
$$E(X|\overline{A}) = \frac{1}{9} \times (1+3+5) + \frac{2}{9} \times (2+4+6) = \frac{11}{3}$$

Применим формулу полного математического ожидания:

$$E(X) = E(X|A)P(A) + E(X|\overline{A})P(\overline{A}) = 3.5 \times \frac{1}{2} + \frac{11}{3} \times \frac{1}{2} = \frac{43}{12}$$

4. Введем следующие события:

   (a) A_1 – первый бросок осуществлялся правильным кубиком.
   (b) A_2 – второй бросок осуществлялся правильным кубиком.
   (c) O_1 – при первом броске выпало четное число.
   (d) O_2 – при втором броске выпало четное число.

По формуле полной вероятности получаем:

$$P(O_2) = P(O_2|A_2)P(A_2) + P(O_2|\overline{A}_2)P(\overline{A}_2) = \frac{1}{2}P(A_2) + \frac{2}{3} \times P(\overline{A}_2)$$

Вновь воспользуемся формулой полной вероятности:

$$P(A_2) = P(A_2|O_1)P(O_1) + P(A_2|\overline{O}_1)P(\overline{O}_1) = \frac{1}{2} \times \frac{1}{2} + \frac{1}{3} \times \frac{1}{2} = \frac{5}{12}$$

В результате получаем:

$$P(O_2) = P(O_2|A_2)P(A_2) + P(O_2|\overline{A}_2)P(\overline{A}_2) = \frac{1}{2}\frac{5}{12} + \frac{2}{3} \times \frac{7}{12} = \frac{43}{72}$$

5. Применим формулу Байеса:
$$P(A_1|O_2) = \frac{P(O_2|A_1)P(A_1)}{P(O_2)} = \frac{P(O_2|A_1) \times \frac{1}{2}}{\frac{43}{72}}$$

Воспользуемся формулой полной вероятности:
$$P(O_2|A_1) = P(O_2|A_1 \cap O_1)P(O_1|A_1) + P(O_2|A_1 \cap \overline{O}_1)P(\overline{O}_1|A_1) =$$
$$= P(O_2|A_2) \times \frac{1}{2} + P(O_2|\overline{A}_2) \times \frac{2}{3} = \frac{1}{2} \times \frac{1}{2} + \frac{1}{2} \times \frac{2}{3} = \frac{7}{12}$$

В итоге получаем:
$$P(A_1|O_2) = \frac{\frac{7}{12} \times \frac{1}{2}}{\frac{43}{72}} = \frac{21}{43}$$

Проверка в R:

```
n <- 100000
# первый бросок правильного кубика
c1 <- sample(1:6, n, replace = TRUE)
# первый бросок неправильного кубика
nc1 <- sample(1:6, n, replace = TRUE, prob = c(1 / 9, 2 / 9, 1 / 9,
                                               2 / 9, 1 / 9, 2 / 9))
# второй бросок правильного кубика
c2 <- sample(1:6, n, replace = TRUE)
# второй бросок неправильного кубика
nc2 <- sample(1:6, n, replace = TRUE, prob = c(1 / 9, 2 / 9, 1 / 9,
                                               2 / 9, 1 / 9, 2 / 9))
# выбор кубика при первом броске
ind <- rbinom(n, 1, prob = 0.5)
# результат первого броска
x <- c1
x[ind == 0] <- nc1[ind == 0]
# выбор кубика при втором броске
ind2 <- ind
even <- x %in% c(2, 4, 6)
ind2[!even] <- 1 - ind[!even]
# результат второго броска
x2 <- c2
x2[ind2 == 0] <- nc2[ind2 == 0]
even2 <- x2 %in% c(2, 4, 6)
# первый пункт
mean(x %in% c(2, 4, 6))
# второй пункт
mean((ind == 1)[x %in% c(2, 4, 6)])
# третий пункт
mean(x)
# четвертый пункт
mean(even2)
# пятый пункт
mean((ind == 1)[even2])
```