

Фамилия:.....

Имя:.....

Группа:.....

Задача №1

Посетители коворкинга, независимо друг от друга¹, арендуют места по цене 180 рублей в час. Время (в часах), которое случайно взятый посетитель проводить в коворкинге, является непрерывной случайной величиной со следующей функцией распределения:

$$F_X(t) = \begin{cases} 0, & \text{если } t < 0 \\ \alpha t^2, & \text{если } t \in (0, 2) \\ 1, & \text{если } t > 2 \end{cases}$$

1. Найдите параметр α . (2 балла)
2. Запишите функцию плотности суммы, которую платит за коворкинг случайно взятый посетитель. Вычислите значения функции плотности в точках 1 и 3. (3 балла)
3. Посчитайте, к чему стремится по вероятности средняя сумма², которую выплачивают все посетители за время, проведенное в коворкинге (предполагается, что посетителей очень много). Укажите, какую теорему (закон) вы использовали для решения данного пункта (необходимо также указать условия применения данной теоремы). (3 балла)
4. Посчитайте дисперсию суммы, которую выплачивает за время, проведенное в коворкинге, случайно взятый посетитель. (2 балла)
5. За день в коворкинге побывали 512 посетителей. При помощи центральной предельной теоремы (ЦПТ) рассчитайте, приблизительно, вероятность, с которой выручка коворкинга от аренды мест посетителям превысила 123840 рублей. Предварительно объясните, почему в данном случае применима ЦПТ. (10 баллов)
6. Посчитайте корреляцию между суммарными затратами первых двух и первых трех посетителей (то есть первые два посетителя входят в число первых трех). (3 балла)
7. Найдите условный второй начальный момент затрат посетителя, если известно, что он провел в коворкинге от полчаса до часа. (2 балла)

¹Предполагается, что коворкинг очень большой и может вмещать одновременно неограниченное количество посетителей. При этом время, которое проводит в коворкинге один посетитель, никак не влияет на время, проведенное в коворкинге остальными посетителями.

²Обычное арифметическое среднее всех выплат.

Решение:

1. Поскольку функция распределения непрерывной случайной величины непрерывна, то:

$$F_X(2) = 1 \implies \alpha \times 2^2 = 1 \implies \alpha = 0.25$$

2. Дифференцируя функцию распределения получаем функцию плотности:

$$f_X(t) = \frac{dF(t)}{dt} = \begin{cases} 0.5t, & \text{если } t \in (0, 2) \\ 0, & \text{в противном случае} \end{cases}$$

Исходя из полученного выражения рассчитываем искомые значения $f_X(1) = 0.5 \times 1 = 0.5$ и $f_X(3) = 0$.

3. Пользуясь найденной функцией плотности рассчитаем математическое ожидание:

$$E(X) = \int_0^2 t \times 0.5t dt = \frac{4}{3}$$

Обозначая через $Y = 180X$ сумму, выплаченную за аренду коворкинга, получаем:

$$E(Y) = E(180X) = 180E(X) = 180 \times \frac{4}{3} = 240$$

Через Y_i обозначим выплату i -го посетителя. Поскольку посетители арендуют коворкинг независимо друг от друга, то Y_i и Y_j независимы для любых $i \neq j$. Следовательно, применяя закон больших чисел получаем:

$$\sum_{i=1}^n Y_i \xrightarrow{p} 240$$

4. По аналогии последовательно вычислим второй начальный момент и дисперсию:

$$E(X^2) = \int_0^2 t^2 \times 0.5t dt = 2Var(X) = 2 - \left(\frac{4}{3}\right)^2 = \frac{2}{9}$$

Отсюда получаем:

$$Var(Y) = 180^2 Var(X) = 180^2 \times \frac{2}{9} = 7200$$

5. Обозначим через Y_i сумму, заплаченную i -м посетителем, где $i \in \{1, \dots, 512\}$. Поскольку соответствующие суммы независимы и одинаково распределены, то в силу ЦПТ:

$$\sum_{i=1}^{512} Y_i \sim \mathcal{N}(512 \times 240, 512 \times 7200) = \mathcal{N}(122880, 1920^2)$$

Отсюда получаем, что:

$$P\left(\sum_{i=1}^{122} Y_i > 123840\right) \approx 1 - \Phi\left(\frac{123840 - 122880}{1920}\right) = 1 - \Phi(0.5) \approx 0.309$$

6. Сперва вычислим ковариацию, учитывая лишь незануляющиеся вследствие независимости элементы:

$$\text{Cov}(Y_1 + Y_2, Y_1 + Y_2 + Y_3) = \text{Var}(Y_1) + \text{Var}(Y_2) = 7200 + 7200 = 14400$$

Пользуясь найденной ковариацией посчитаем корреляцию:

$$\begin{aligned} \text{Cor}(Y_1 + Y_2, Y_1 + Y_2 + Y_3) &= \frac{14400}{\sqrt{(\text{Var}(Y_1) + \text{Var}(Y_2)) \times (\text{Var}(Y_1) + \text{Var}(Y_2) + \text{Var}(Y_3))}} = \\ &= \frac{14400}{\sqrt{(7200 + 7200) \times (7200 + 7200 + 7200)}} \approx 0.816 \end{aligned}$$

7. Найдем вероятность условия:

$$P(X \in [0.5, 1]) = F_X(1) - F_X(0.5) = 0.25 \times 1^2 - 0.25 \times 0.5^2 = 0.1875$$

Отсюда получаем условную функцию плотности проведенного в коворкинге времени:

$$f_X(t) = \begin{cases} 0.5t/0.1875, & \text{если } t \in [0.5, 1) \\ 0, & \text{в противном случае} \end{cases} = \begin{cases} \frac{8t}{3}, & \text{если } t \in [0.5, 1) \\ 0, & \text{в противном случае} \end{cases}$$

Используя условную функцию плотности рассчитываем условный второй начальный момент:

$$E(X^2 | X \in [0.5, 1]) = \int_{0.5}^1 t^2 \times \frac{8t}{3} dt = 0.625$$

Наконец, применяя свойство линейности математического ожидания получаем:

$$E((180X)^2 | X \in [0.5, 1]) = 180^2 E(X^2 | X \in [0.5, 1]) = 180^2 \times 0.625 = 20250$$

Проверка в R:

```
n <- 512 * 100000
u <- runif(n)
q_X <- function(t)
{
  return(2 * sqrt(t))
}
x <- q_X(u)
y <- 180 * x
# пункт 3
c(true = 240, est = mean(y))
# пункт 4
c(true = 7200, est = var(y))
# пункт 5
library("Thermimage")
y_sum <- meanEveryN(y, 512, lag = 0) * 512
c(clt = 0.309, est = mean(y_sum > 123840))
# пункт 6
y1 <- y[seq(1, length(y), by = 512)] +
```

```
y[seq(2, length(y), by = 512)] +  
y[seq(3, length(y), by = 512)]  
y2 <- y[seq(1, length(y), by = 512)] +  
y[seq(2, length(y), by = 512)]  
c(true = 0.816, est = cor(y1, y2))  
# пункт 7  
c(true = 20250, est = mean(y[(x > 0.5) & (x < 1)] ^ 2))
```


Фамилия:.....

Имя:.....

Группа:.....

Задача №2

Коварный преподаватель на лекции дал 144 студентам проверочную работу, включающую 5 тестовых заданий с 12 вариантами ответов в каждом (необходимо выбрать один верный ответ). Поскольку студенты в тот же день писали контрольную работу по другому предмету, то очень устали и выбирали варианты ответов в тесте наугад (каждый вариант ответа мог быть выбран с равной вероятностью). При помощи центральной предельной теоремы посчитайте, приблизительно, вероятность, с которой:

1. Студенты (суммарно) решат верно более 65 заданий. **(5 балла)**
2. Хотя бы треть студентов справилась хотя бы с одним заданием (дали верный ответ хотя бы на один вопрос). **(5 баллов)**
Подсказка: $\left(\frac{11}{12}\right)^5 \approx 0.65$
3. Общее число успешно решенных заданий (всеми студентами в сумме) окажется хотя бы на 20% (в 1.2 раза) больше числа студентов, успешно решивших по крайней мере одно задание. **(5 баллов)**

Решение:

1. Через $X_i \sim B\left(5, \frac{1}{12}\right)$ обозначим случайную величину, отражающую число верных ответов в тесте, которые дал i -й студент.

Обратим внимание, что поскольку студенты решают задачи независимо друг от друга и с равной вероятностью приходят к верному ответу, то можно воспользоваться ЦПТ:

$$\begin{aligned} E(X_i) &= 5 \times \frac{1}{12} = \frac{5}{12} \\ \text{Var}(X_i) &= 5 \times \frac{1}{12} \times \frac{11}{12} = \frac{55}{144} \\ (X_1 + \dots + X_{144}) &\sim \mathcal{N}\left(\frac{5}{12} \times 144, \frac{55}{144} \times 144\right) = \mathcal{N}(60, 55) \end{aligned}$$

В результате получаем:

$$P(X_1 + \dots + X_{144} > 65) = 1 - P(X_1 + \dots + X_{144} \leq 65) = 1 - \Phi\left(\frac{65 - 60}{\sqrt{55}}\right) \approx 0.25$$

2. Через Y_i обозначим случайную величину, принимающую значение 1, если студент решил хотя бы одну задачу и 0 – в противном случае. При этом обратим внимание, что $Y_i = 1$ когда $X_i \geq 1$ и $Y_i = 0$, если $X_i = 0$, откуда нетрудно показать, что $Y_i \sim \text{Ber}\left(1 - \left(\frac{11}{12}\right)^5\right)$, то есть $Y_i \sim \text{Ber}(0.35)$, а значит в силу ЦПТ:

$$\begin{aligned} E(Y_i) &\approx 0.35 \\ \text{Var}(Y_i) &= 0.35 \times (1 - 0.35) = 0.2275 \\ (Y_1 + \dots + Y_n) &\sim \mathcal{N}(0.35 \times 144, 0.2275 \times 144) = \mathcal{N}(50.4, 32.76) \end{aligned}$$

Посчитаем искомую вероятность:

$$\begin{aligned} P\left(\frac{1}{144}(Y_1 + \dots + Y_{144}) \geq \frac{1}{3}\right) &= P(Y_1 + \dots + Y_{144} \geq 48) \approx \\ &\approx 1 - \Phi\left(\frac{48 - 50.4}{\sqrt{32.76}}\right) \approx 0.66 \end{aligned}$$

3. Необходимо рассчитать следующую вероятность:

$$\begin{aligned} P(X_1 + \dots + X_{144} \geq 1.2(Y_1 + \dots + Y_{144})) &= \\ = P((X_1 - 1.2Y_1) + (X_2 - 1.2Y_2) + \dots + (X_{144} - 1.2Y_{144}) \geq 0) \end{aligned}$$

Обратим внимание, что случайные величины $(X_i - 1.2Y_i)$ независимы и одинаково распределены, а значит в отношении их суммы применима центральная предельная теорема.

$$\begin{aligned} E(X_i - 1.2Y_i) &= \frac{5}{12} - 1.2 \times 0.35 = -\frac{1}{300} \\ E(X_i Y_i) &= E(X_i | Y_i = 1)P(Y_i = 1) + E(X_i | Y_i = 0)P(Y_i = 0) = \end{aligned}$$

$$= E(X_i | X_i \geq 1)P(X_i \geq 1) + E(X_i | X_i = 0) \times P(X_i = 0) = E(X_i) = \frac{5}{12}$$

$$Cov(X_i, Y_i) = \frac{5}{12} - \frac{5}{12} \times 0.35 \approx 0.27$$

$$Var(X_i - 1.2Y_i) = \frac{5}{12} + 1.2^2 \times 0.2275 - 2 \times 1.2 \times 0.27 = 0.096$$

В результате получаем:

$$\sum_{i=1}^{144} (X_i - 1.2Y_i) \sim \mathcal{N}\left(-\frac{1}{300} \times 144, 0.096 \times 144\right) = \mathcal{N}(-0.48, 13.824)$$

Воспользуемся найденным приближением распределения для расчета искомой вероятности:

$$\begin{aligned} P((X_1 - 1.2Y_1) + (X_2 - 1.2Y_2) + \dots + (X_{144} - 1.2Y_{144}) \geq 0) = \\ = 1 - \Phi\left(\frac{0 + 0.48}{\sqrt{13.824}}\right) \approx 0.45 \end{aligned}$$

Проверка в R:

```
n.sim <- 10000
n.student <- 144
yes <- rep(NA, n.sim)
works65 <- rep(NA, n.sim)
works33prc <- rep(NA, n.sim)
a <- rep(NA, n.sim)
for (i in 1:n.sim)
{
  works <- rbinom(n = n.student, size = 5, prob = 1 / 12)
  s_works <- sum(works)
  s_works_greater <- sum(works >= 1)
  works65[i] <- s_works > 65
  works33prc[i] <- s_works_greater >= (n.student / 3)
  yes[i] <- s_works >= (1.2 * s_works_greater)
}
# пункт 1
mean(works65)
1 - pnorm((65 - n.student * (5 / 12)) /
  sqrt(n.student * 5 * (1 / 12) * (11 / 12)))
# пункт 2
mean(works33prc)
p <- 1 - (11 / 12) ^ 5
1 - pbinom(p = p, size = n.student, n.student / 3 - 1)
1 - pnorm((n.student * (1 / 3) - p * n.student) /
  sqrt(144 * p * (1 - p)))
# пункт 3
mean(yes)
1 - pnorm(0.48 / sqrt(13.824))
```


Фамилия:.....

Имя:.....

Группа:.....

Задача №3

Имеется бесконечная последовательность случайных величин X_n со следующими функциями плотности:

$$f_{X_n}(x) = \begin{cases} \frac{2(\theta_n - x)}{\theta_n^2}, & \text{если } x \in (0, \theta_n) \\ 0, & \text{в противном случае} \end{cases}, \text{ где } \theta_n = \frac{2n}{n+1}$$

Эта последовательность сходится по вероятности к некоторой случайной величине Y .

1. Запишите функцию распределения Y . (5 баллов)
2. Укажите, к чему сходится по вероятности $\sqrt[n]{e^{X_n}}$. Приведите формальное доказательство. (3 балла)
3. Определите, к чему сходится по вероятности последовательность $\sqrt[n]{e^{X_n}}Y$. Приведите формальное доказательство. (2 балла)

Подсказка:

$$E(X_n) = \frac{\theta_n}{3} \quad \text{Var}(X_n) = \frac{\theta_n^2}{18}$$

Решение:

1. Найдем функцию распределения X_n :

$$F_{X_n}(x) = \begin{cases} 0, & \text{если } x < 0 \\ \frac{x(2\theta_n - x)}{\theta_n^2}, & \text{если } x \in (0, \theta_n) \\ 1, & \text{если } x > \theta_n \end{cases}$$

Поскольку X_n сходится к Y по вероятности, то должна соблюдаться и сходимост по распределению. Исходя из формы функции распределения элементов последовательности нетрудно предположить, что Y имеет следующую функцию распределения:

$$F_Y(x) = \begin{cases} 0, & \text{если } x < 0 \\ \frac{x(4-x)}{4}, & \text{если } x \in (0, 2) \\ 1, & \text{если } x > 2 \end{cases}$$

Обращая внимание на то, что $\lim_{n \rightarrow \infty} \theta_n = 2$, убедимся, что в данном случае соблюдается сходимост по распределению. Полагая $x \in (0, 2)$ получаем:

$$F_{X_n}(x) = \lim_{n \rightarrow \infty} \frac{x(2\theta_n - x)}{\theta_n^2} = \frac{x(4 - x)}{4}$$

2. Покажем, что последовательность $\sqrt[n]{e^{X_n}}$ сходится по вероятности к 1. Для этого сперва убедимся, что последовательность $\ln(\sqrt[n]{e^{X_n}}) = \frac{X_n}{n}$ сходится по вероятности к 0:

$$\begin{aligned} \lim_{n \rightarrow \infty} E\left(\frac{X_n}{n}\right) &= \lim_{n \rightarrow \infty} \frac{\theta_n}{3n} = \lim_{n \rightarrow \infty} \frac{2n}{3(n+1)n} = \lim_{n \rightarrow \infty} \frac{2}{3(n+1)} = 0 \\ \lim_{n \rightarrow \infty} Var\left(\frac{X_n}{n}\right) &= \lim_{n \rightarrow \infty} \frac{\theta_n^2}{18n^2} = \lim_{n \rightarrow \infty} \frac{4n^2}{18n^2(n+1)^2} = \lim_{n \rightarrow \infty} \frac{2}{9(n+1)^2} = 0 \end{aligned}$$

По теореме Манна-Вальда отсюда следует, что $\sqrt[n]{e^{X_n}}Y = e^{\ln(\sqrt[n]{e^{X_n}})}Y$ сходится по вероятности к $e^0 = 1$.

3. Пользуясь полученным ранее результатом заключаем, что по теореме Слуцкого $\sqrt[n]{e^{X_n}}Y$ сходится по вероятности к $1 \times Y = Y$.

