Теория вероятностей и статистика, МИРЭК, 2022-2023

Дедлайн: домашнее задание отправляется в **pdf** формате на почту семинариста. В копию письма необходимо поставить ассистента группы.

Почты, на которые следует отправлять домашние задания, в зависимости от вашего семинариста:

1. Погорелова Полина Вячеславовна – tvis.we.2021@gmail.com

2. Потанин Богдан Станиславович – tvismirec@gmail.com

3. Слаболицкий Илья Сергеевич – tvis.fweia.hse@gmail.com

Домашнее задание должно быть отправлено на указанные почты в **pdf** формате до конца дня **04.12.2022** включительно (по московскому времени). Тема письма должна иметь следующий формат: “МИРЭК Фамилия Имя Группа Номер ДЗ”, например, “МИРЭК Потанин Богдан 200 ДЗ 2”.
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. Байкеры и дальнобойщик (80 баллов)

Количество литров, заправляемых в мотоцикл подъехавшим на заправку байкером, является случайной величиной X со следующей функцией плотности:

$$f_X(t) = \begin{cases} \frac{(t-1)}{18}, \text{ если } t \in [1, 5) \\ \frac{\alpha - 0.8t}{18}, \text{ если } t \in [5, 10) \\ 0, \text{ в противном случае} \end{cases}$$

Байкеры заправляют мотоциклы независимо друг от друга. Стоимость одного литра бензина составляет 90 рублей.

1. Найдите параметр α. (**10 баллов**)
2. Посчитайте математическое ожидание суммы, которую выплачивает за бензин случайно взятый байкер. (**5 баллов**)
3. Повторите предыдущий пункт для дисперсии. (**5 баллов**)
4. Запишите функцию распределения суммы, которую платит за бензин случайно взятый байкер. (**10 баллов**)
5. Рассчитайте, какую сумму не превысит выплата байкера в половине случаев. (**10 баллов**)
6. За день на заправке побывали 122 байкера. При помощи центральной предельной теоремы (ЦПТ) рассчитайте, приблизительно, вероятность, с которой выручка заправки от продажи бензина байкерам превысила 60480 рублей. Предварительно объясните, почему в данном случае применима ЦПТ. (**10 баллов**)
7. В этот же день на заправку подъехал дальнобойщик на огромном грузовике. Дальнобойщик попросил заправить ему ровно в 122 раз больше литров бензина, чем только что воспользовавшемуся услугами заправки байкеру. Посчитайте, с какой вероятностью дальнобойщику придется заплатить более 60480 рублей. (**10 баллов**)
8. Посчитайте корреляцию между выручкой заправки от продажи бензина дальнобойщику и байкерам в соответствующий день. (**10 баллов**)
   Подсказка: без потери общности предположите, что дальнобойщик заправил себе в 122 раз больше литров, чем первый из воспользовавшихся услугой заправки в этот день байкеров.
9. Найдите условное математическое ожидание затрат байкера на покупку бензина, если известно, что он купил более 7 литров бензина. (**10 баллов**)

Решение:

1. Поскольку интеграл функции плотности на носителе должен равняться единице, получаем:

$$\int_1^5 \frac{(t-1)}{18} dt + \int_5^{10} \frac{\alpha - 0.8t}{18} dt = 1 \implies \frac{5\alpha - 22}{18} = 1 \implies \alpha = 8$$

2. Пользуясь найденной функцией плотности рассчитаем математическое ожидание:

$$E(X) = \int_1^5 t\frac{(t-1)}{18}dt + \int_5^{10} t\frac{8-0.8t}{18}dt = \frac{16}{3}$$

Обозначая через $Y = 90X$ сумму, выплаченную за бензин, получаем:

$$E(Y) = E(90X) = 90E(X) = \frac{16}{3} \times 90 = 480$$

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

$$E(X^2) = \int_1^5 t^2\frac{(t-1)}{18}dt + \int_5^{10} t^2\frac{8-0.8t}{18}dt = \frac{191}{6}$$

$$Var(X) = \frac{191}{6} - \left(\frac{16}{3}\right)^2 = \frac{61}{18}$$

Отсюда получаем:

$$Var(Y) = 90^2 Var(X) = 90^2 \times \frac{61}{18} = 27450$$

4. Сперва найдем функцию распределения при $t \in (1,5)$:

$$F_X(t) = \int\limits_1^t \frac{(x-1)}{18}dx = \frac{(t-1)^2}{36}$$

По аналогии находим функцию распределения на участке $t \in [5,10)$:

$$F_X(t) = P(X \le t) = P(X \in [1,t]) = P(X \in [1,5)) + P(X \in [5,t]) =$$
$$= F_X(5) + \int_5^t \frac{8-0.8x}{18}dx = \frac{(5-1)^2}{36} + \frac{-t^2+20t-75}{45} = \frac{-t^2+20t-55}{45}$$

В результате получаем:

$$F_X(t) = \begin{cases} 0, \text{ если } t < 1 \\ \frac{(t-1)^2}{36}, \text{ если } t \in [1,5) \\ \frac{-t^2+20t-55}{45}, \text{ если } t \in [5,10) \\ 1, \text{ если } t \ge 10 \end{cases}$$

Следовательно:

$$F_Y(t) = P(90X \le t) = P\left(X \le \frac{t}{90}\right) = F_X\left(\frac{t}{90}\right) =$$

$$= \begin{cases} 0, \text{ если } t < 90 \\ \frac{\left(\frac{t}{90}-1\right)^2}{36}, \text{ если } t \in [90,450) \\ \frac{-\left(\frac{t}{90}\right)^2+20\left(\frac{t}{90}\right)-55}{45}, \text{ если } t \in [450,900) \\ 1, \text{ если } t \ge 900 \end{cases}$$

5. Необходимо найти медиану. Обратим внимание, что $F_X(5) = \frac{4}{9} < 0.5$, а значит медиана больше 5 и искать ее нужно на втором участке функции распределения, откуда:

$$F_X(t_{0.5}) = 0.5 \implies \frac{-t^2 + 20t - 55}{45} = 0.5 \implies t_{0.5} \approx 5.257$$

Отсюда следует, что медиана Y будет равняться $90 \times 5.257 = 473.1$

В качестве альтернативы также можно было бы сперва найти квантильную функцию:

$$F_X^{-1}(t) = \begin{cases} 1 + 6\sqrt{t}, \text{ если } t \in [0, \frac{4}{9}) \\ 10 - 3\sqrt{5}\sqrt{1-t}, \text{ если } t \in [\frac{4}{9}, 1) \end{cases}$$

Следовательно:

$$F_Y^{-1}(t) = 90F_X^{-1}(t) = \begin{cases} 90 + 540\sqrt{t}, \text{ если } t \in [0, \frac{4}{9}) \\ 900 - 270\sqrt{5}\sqrt{1-t}, \text{ если } t \in [\frac{4}{9}, 1) \end{cases}$$

В результате получаем:

$$F_X^{-1}(0.5) = 900 - 270\sqrt{5}\sqrt{1-0.5} \approx 473.1$$

6. Обозначим через Y_i сумму, заплаченную i-м байкером, где $i \in \{1, ..., 122\}$. Поскольку соответствующие суммы независимы и одинаково распределены, то в силу ЦПТ:

$$\sum_{i=1}^{122} Y_i \dot\sim \mathcal{N}\left(122 \times 480, 27450 \times 122\right) = \mathcal{N}\left(58560, 3348900\right)$$

Отсюда получаем, что:

$$P\left(\sum_{i=1}^{122} Y_i > 60480\right) \approx 1 - \Phi\left(\frac{60480 - 58560}{\sqrt{3348900}}\right) \approx 1 - \Phi(1.049) \approx 0.147$$

7. Обозначим через $Z = 122Y_1$ то, сколько заплатил дальнобойщик, откуда:

$$\begin{gathered} P((Z > 60480) = P\left(122Y_1 > 60480\right) = \\ = P\left(Y_1 > \frac{60480}{122}\right) = 1 - F_Y\left(\frac{60480}{122}\right) = \\ = 1 - \frac{-\left(\frac{\frac{60480}{122}}{90}\right)^2 + 20\left(\frac{\frac{60480}{122}}{90}\right) - 55}{45} \approx 0.448 \end{gathered}$$

Мораль: $122Y_1$ и $\sum\limits_{i=1}^{122} Y_i$ распределены по разному. В общем случае нельзя заменять сумму случайных величин на умножение одной случайной величины на количество элементов суммы. В последнем случае, если речь идет о сумме независимых и одинаково распределенных случайных величин, следует пользоваться ЦПТ, в то время как в первом случае удобно применять исходное распределение.

8. Сперва вычислим ковариацию:

$$Cov(122Y_1, Y_1 + ... + Y_{122}) = Cov(122Y_1, Y_1) = 122Var(Y_1) = 122 \times 27450 = 3348900$$

Пользуясь найденной ковариацией посчитаем корреляцию:

$$Cor(122Y_1, Y_1 + ... + Y_{122}) = \frac{3348900}{\sqrt{122^2 Var(Y_1) \times Var(Y_1 + ... + Y_{122})}} =$$
$$= \frac{3348900}{\sqrt{122^2 \times 27450 \times 3348900}} \approx 0.090536$$

9. Найдем вероятность условия:

$$P(X > 7) = 1 - F_X(7) = 1 - \frac{-7^2 + 20 \times 7 - 55}{45} = 0.2$$

Отсюда получаем условную функци плотности купленных литров бензина:

$$f_X(t) = \begin{cases} \frac{8-0.8t}{18}/0.2, \text{ если } t \in [7, 10) \\ 0, \text{ в противном случае} \end{cases} = \begin{cases} \frac{40-4t}{18}, \text{ если } t \in [7, 10) \\ 0, \text{ в противном случае} \end{cases}$$

Используя условную функцию плотности рассчитываем условное математическое ожидание:

$$E(X|X > 7) = \int_7^{10} t\frac{40 - 4t}{18} dt = 8$$

Наконец, применяя свойство линейности математического ожидания получаем:

$$E(Y|X > 7) = E(90X|X > 7) = 90E(X|X > 7) = 720$$

Проверка в R:

```
n <- 122 * 100000
u <- runif(n)
q_X <- function(t)
{
  val <- rep(NA, length(t))
  cond <- t < (4 / 9)
  val[cond] <- 1 + 6 * sqrt(t[cond])
  val[!cond] <- 10 - (3 * sqrt(5)) * sqrt(1 - t[!cond])
  return(val)
}
x <- q_X(u)
y <- 90 * x
# пункт 2
c(true = 480, est = mean(y))
# пункт 3
c(true = 27450, est = var(y))
# пункт 5
```

```
c(true = 473.1, est = median(y))
# пункт 6
library("Thermimage")
y_sum <- meanEveryN(y, 122, lag = 0) * 122
c(clt = 0.159, est = mean(y_sum > 60480))
# пункт 7
c(true = 0.448, est = mean((122 * y) > 60480))
# пункт 8
y1 <- y[seq(1, length(y), by = 122)]
c(true = 0.090536, est = cor(122 * y1, y_sum))
# пункт 9
c(true = 720, est = mean(y[x > 7]))
```

Задание №2. Последовательность (20 баллов)

Имеется бесконечная последовательность независимых равномерных случайных величин $X_n \sim U(0,1)$. Также, имеется последовательность $Y_n = X_1 \times X_2 \times ... \times X_n$.

1. Докажите, что последовательность $\cos(Y_n)$ сходится по вероятности к 1. (**10 баллов**)
2. Определите, к чему по вероятности сходится последовательность $\sqrt[n]{Y_n}$ и приведите соответствующее доказательство. (**10 баллов**)

Решение:

1. Пользуясь тем, что $P(Y_n < 0) = 0$, применим неравенство Маркова и покажем, что для каждого $\varepsilon > 0$:

$$\lim_{n\to\infty} P\left(|Y_n - 0| > \varepsilon\right) = \lim_{n\to\infty} P\left(|X_1 \times X_2 \times ... \times X_n - 0| > \varepsilon\right) =$$
$$= \lim_{n\to\infty} P\left(X_1 \times X_2 \times ... \times X_n > \varepsilon\right) \leq \lim_{n\to\infty} \frac{E(X_1 \times X_2 \times ... \times X_n)}{\varepsilon} =$$
$$= \lim_{n\to\infty} \frac{E(X_1) \times ... \times E(X_n)}{\varepsilon} = \lim_{n\to\infty} \frac{0.5^n}{\varepsilon} = 0$$

   Мы показали, что последовательность Y_n стремится к 0, а значит, по теореме Манна-Вальда, последовательность $\cos(Y_n)$ стремится к $\cos(0) = 1$.

2. Сперва рассмотрим логарифм исходной последовательности:

$$\ln(\sqrt[n]{Y_n}) = \frac{1}{n}\ln(Y_n) = \frac{1}{n}\ln\left(X_1 \times X_2 \times ... \times X_n\right) = \frac{1}{n}\sum_{i=1}^{n}\ln(X_i)$$

В итоге мы получили арифметическое среднее логарифмов равномерных случайных величин. Согласно закону больших чисел оно стремится к:

$$E(\ln(X_1)) = \int_0^1 \ln(x)dx = -1$$

Таким образом $\ln(\sqrt[n]{Y_n}) \xrightarrow{p} -1$. Следовательно, в силу теоремы Манна-Вальда получаем:

$$\sqrt[n]{Y_n} = e^{\ln(\sqrt[n]{Y_n})} \xrightarrow{p} e^{-1}$$