

Домашнее задание 1

Дедлайн: 2024-09-16

1. Вася решает три задачи по теории вероятностей. Вероятности решить эти задачи равны 0.1, 0.2 и 0.3. Решения задач никак не связаны между собой, знание ни одной из задач не помогает решить ни одну другую. Обозначим буквой N общее количество решенных задач.

а) Найдите все значения N и их вероятности.

б) Найдите $\mathbb{P}(N > 1)$, $\mathbb{E}(N)$ и $\mathbb{E}(N^2)$.

2. За работу Вася получает случайное количество ξ баллов, равновероятно распределённое от 1 до n . Найдите $\mathbb{E}(\xi)$, $\mathbb{E}(\xi^2)$, $\mathbb{E}(\xi^3)$.

3. Берём набор данных по ссылке

https://github.com/bdemeshev/hse_panda_metrics_2024_2025/raw/main/home_assignments/ha01_data.csv.

Здесь две переменных: y_i — количество просмотренных Васей рилзов в день i и бинарная переменная x_i ($x_i = A$ — обычный день, $x_i = B$ — день дедлайна по теории вероятностей).

Рассмотрим две гипотезы. Нулевая гипотеза H_0 : приближение дедлайна по вероятностям никак не влияет на количество просмотренных рилзов. Альтернативная гипотеза H_1 : приближение дедлайна в среднем снижает количество просмотренных рилзов.

а) Посчитайте фактическое значение статистики $S = \bar{y}_B - \bar{y}_A$.

б) Предполагая, что H_0 верна, сгенерируйте 10000 случайных перестановок меток x и для каждой перестановки посчитайте значение статистики $S^{\text{new}} = \bar{y}_B^{\text{new}} - \bar{y}_A^{\text{new}}$.

в) Оцените p -значение, в данном случае p -значение — это вероятность $\mathbb{P}(S^{\text{new}} \leq S \mid S, H_0)$.

г) Для принятия решения, отвергать или нет H_0 , мы используем уровень значимости $\alpha = 0.05$. Отвергаем ли мы H_0 ?
