

ВШБ Бизнес-информатика: ТВиМС 2025.
Экзаменационный вариант 2

1. (8 баллов) В ресторане "Вкусно и вопросительный знак" поток клиентов моделируется Пуассоновским процессом. Известно, что вероятность, что в определенное время суток за час в ресторан никто не придет составляет 0.07. Какова вероятность, что в случайный момент в течение этого времени суток ждать следующего вошедшего клиента мы будем больше 15 минут?

2. (12 баллов) Кредитные риски при выдаче кредита в компании "ВопросБанк" моделируются распределением Лапласа с параметрами $\alpha = 0.5$ и $\beta = -0.1$, которое имеет следующую функцию плотности:

$$f(x) = \frac{\alpha}{2} e^{-\alpha|x-\beta|}, \quad -\infty < x < +\infty,$$

где $\alpha > 0$ - параметр масштаба, $-\infty < \beta < +\infty$ - параметр сдвига.

Начальный момент k -го порядка для распределения Лапласа может быть рассчитан по следующей формуле:

$$E[X^k] = \int_{-\infty}^{+\infty} x^k f(x) dx = \sum_{i=0}^{\lfloor k/2 \rfloor} \frac{\beta^{k-2i}}{\alpha^{2i}} \frac{k!}{(k-2i)!},$$

где $\lfloor k/2 \rfloor$ - целая часть $k/2$.

Каждый день ВопросБанк обрабатывает 1000 независимых заявок на получение кредита. Какова вероятность, что средний дневной риск по всем клиентам за день поднимется выше отметки 0.1 менее чем 3 раз за год?

3. (14 баллов) Предположим, что у нас есть реализация случайной выборки $\mathcal{X} = (x_1, \dots, x_n)$ неизвестной случайной величины X с плотностью

$$f_X(x; \theta) = \begin{cases} \frac{4x^3}{\theta^4}, & \text{при } x \in [0, \theta] \\ 0, & \text{иначе.} \end{cases}$$

Реализация, которая была получена: $(x_1, \dots, x_6) = (1.5, 2.8, 5.4, 0.8, 6.1, 7.2)$.

- (a) (7 баллов) Найдите оценку параметра θ - функцию от выборки $\hat{\theta}_{ML} = \hat{\theta}_{ML}(\mathcal{X})$ - методом максимального правдоподобия.

Посчитайте её реализацию на предоставленных данных.

- (b) (7 баллов) Найдите оценку параметра θ - функцию от выборки $\hat{\theta}_{MM} = \hat{\theta}_{MM}(\mathcal{X})$ - методом моментов.

Посчитайте её реализацию на предоставленных данных.

4. (14 баллов) На дисциплине «Теория вероятностей» 25% студентов списывают. Поэтому преподаватели помимо письменной части экзамена ввели ещё и обязательную устную защиту для всех студентов. Известно, что студенты, которые списывали на письменной части экзамена, на устной защите на каждый вопрос по решённой ими задаче независимо отвечают с вероятностью 0.7. Студенты, которые решали письменный экзамен самостоятельно, на каждый вопрос по своей работе независимо отвечают с вероятностью 0.85. Студентам на устной защите задаётся 8 вопросов.

Какой максимальный порог отсечения K нужно ввести (ответил хотя бы на K вопросов — защитился; не ответил хотя бы на K вопросов — обнуление), чтобы при количестве ответов меньше K вероятность того, что студент списал, была бы не ниже 80%?

5. (6 баллов) 55 студентов университета на определённом курсе были случайным образом распределены в две учебные группы размером 30 и 25 студентов соответственно. В конце года все студенты сдали экзамен, и их оценки обобщены в таблице ниже.

	Размер выборки	Выборочное среднее	Выборочное стандартное отклонение
Группа 1	30	75.33	7.61
Группа 2	25	71.40	6.37

- (a) (2 балла) Посчитайте 95% доверительный интервал для математического ожидания экзаменационных оценок студентов группы 1.
- (b) (4 балла) Используйте подходящий тест гипотез, чтобы определить, больше ли средний балл студентов группы 1, чем средний балл студентов группы 2.

Сформулируйте гипотезы и ваши предположения касательно свойств и характеристик случайных величин, которые вы исследуете. Укажите используемую статистику и её распределение при нулевой гипотезе. Оформите ваши результаты и сделайте выводы.

6. (10 баллов) Студенческий совет университета собрал случайную выборку из 525 студентов, чтобы определить, поддерживают ли они новое расписание экзаменов. Результаты обобщены в таблице ниже.

Область обучения	Размер выборки	Поддерживают новое расписание экзаменов
Гуманитарные науки	325	230
Естественные науки	200	110

- (a) (2 балла) Посчитайте 92% доверительный интервал для истинной доли студентов гуманитарных наук, которые поддерживают новое расписание экзаменов.
- (b) (2 балла) Посчитайте 96% доверительный интервал для истинной разности долей положительных ответов между студентами гуманитарных и естественных наук.
- (c) (5 баллов) Проведите двусторонний тест на уровне значимости 4%, чтобы проверить гипотезу о том, что доля студентов гуманитарных наук, которые поддерживают новое расписание экзаменов, равна доле студентов естественных наук, которые поддерживают новое расписание экзаменов. Сформулируйте гипотезы, укажите используемую статистику и её распределение при нулевой гипотезе. Проведите тестирование через: **score** (критерий) и **p**-значение.
- (d) (1 балл) Оформите ваши результаты, сравните результаты пункта (c) с доверительным интервалом из пункта (b) и сделайте выводы.

7. (21 балл) Предположим, что у нас есть реализации случайных выборок: $\mathcal{X} = \{x_1, \dots, x_n\}$ и $\mathcal{Y} = \{y_1, \dots, y_m\}$, случайных величин $X \sim \mathcal{N}(\mu_x, \sigma_x^2)$ и $Y \sim \mathcal{N}(\mu_y, \sigma_y^2)$ соответственно. В следующих пунктах исследуются различные аспекты этих выборок.

(a) (3 балла) Пусть \bar{y} — реализация выборочного среднего на выборке \mathcal{Y} . Найдите m такое, что интервал:

$$(\bar{y} - 0.55 \sigma_y, \bar{y} + 0.55 \sigma_y)$$

является приблизительно 95% доверительным интервалом для μ_y .

(b) (5 баллов) Мы привыкли строить доверительные интервалы для математического ожидания вокруг выборочного среднего. Но ничто не мешает нам забросить эту "рыболовную сеть" вокруг одной реализации y случайной величины Y в попытке поймать математическое ожидание μ_y . Каким будет доверительный уровень интервала такой же ширины, как в предыдущем пункте, но построенного на основе всего лишь одной реализации y ?

(c) (6 баллов) Пусть \bar{X} и \bar{Y} — выборочные средние двух независимых *случайных* выборок случайных величин X и Y , каждая размера n ($m = n$), где истинные дисперсии известны $\sigma_x^2 = 2\sigma^2$ и $\sigma_y^2 = 2\sigma^2$ соответственно. Найдите n такое, что:

$$P\left(\bar{X} + \bar{Y} - \frac{2\sigma}{5} < \mu_x + \mu_y < \bar{X} + \bar{Y} + \frac{2\sigma}{5}\right) = 0.86.$$

(d) (7 баллов) Мы хотим проверить гипотезу о том, что $\mu_x = \mu_y + \Delta$ против двусторонней альтернативной гипотезы. Предположим, что известны данные: $n = 5$, $m = 9$, $\bar{x} = 9.5$, $\bar{y} = 4.5$, $\Delta = 2$, $\sigma_x = 2.2$, $\sigma_y = 3.2$. Используйте данные и проведите тест, используя уровни значимости $\alpha = 2\%$, 5% .

Для полного оценивания этого пункта недостаточно просто сказать, отклоняем ли мы гипотезу или нет. Вам нужно как-то обосновать ваши заключения: укажите используемую статистику и её распределение при нулевой гипотезе, процесс принятия решения (что с чем сравниваем).

8. (18 баллов) Рассмотрим две случайные величины X и Y . Они обе принимают значения $-1, 0$ и 1 . Совместные вероятности для каждой пары заданы следующей таблицей, где $\theta \in \mathbb{R}$ — параметр:

	$X = -1$	$X = 0$	$X = 1$
$Y = -1$	0.1	$0.1 + \theta$	$0.3 + 3\theta$
$Y = 0$	$0.1 + \theta$	$0.2 - 6\theta$	0
$Y = 1$	$0.1 + \theta$	0.1	0

- (a) (2 балла) Какой диапазон значений может принимать параметр θ , чтобы приведённая выше таблица соответствовала таблице вероятностей?

В дальнейших пунктах предполагается, что θ находится в диапазоне, найденном в пункте 1, однако вы должны проводить все вычисления для произвольного θ .

- (b) (1 балл) Вычислите

$$P(X = 0 \mid X + Y = 0).$$

- (c) (2 балла) Постройте таблицу вероятностей условного распределения X при условии $Y = 1$.

- (d) (2 балла) Вычислите $\text{Corr}(X, Y)$.

- (e) (5 баллов) Предположим, что у вас есть реализация случайной выборки: $\mathcal{Y} = (y_1, \dots, y_m)$, где каждая y_i получена из закона распределения случайной величины Y из таблицы выше.

Найдите оценку параметра θ - функцию от выборки $\hat{\theta}_{MM} = \hat{\theta}_{MM}(\mathcal{Y})$ - методом моментов.

- (f) (6 баллов) Рассмотрите $\hat{\theta}_1 = X$ и $\hat{\theta}_2 = \frac{X+Y}{2}$ как оценки для неизвестного параметра θ . Какую из них вы предпочтёте и почему? (Посмотрите свойства этих точечных оценок).