ВШБ Бизнес-информатика: ТВиМС 2025.
Экзаменационный вариант 4

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 0.23. Какова вероятность, что в случайный момент в течение этого времени суток ждать следующего вошедшего клиента мы будем менее 45 минут?

2. (12 баллов) Кредитные риски при выдаче кредита в компании "ВопросБанк"моделируются распределением Лапласа с параметрами $\alpha = 1.5$ и $\beta = -0.5$, 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:

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

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 опустится ниже отметки -0.55 от 18 до 20 раз за год?

3. (14 баллов) На дисциплине «Теория вероятностей» 60% 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 0.5. Студенты, которые решали письменный экзамен самостоятельно, на каждый вопрос по своей работе независимо отвечают с вероятностью 0.9. Студентам на устной защите задаётся 7 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 95%?

4. (14 баллов) Предположим, что у нас есть реализация случайной выборки $\mathcal{X} = (x_1, \ldots, x_n)$ неизвестной случайной величины X с плотностью

$$f_X(x;\theta) = \begin{cases} \frac{\theta}{x^{1+\theta}}, & \text{при } x \geqslant 1 \\ 0, & \text{при } x < 1 \end{cases}$$

Реализация, которая была получена: $(x_1, \ldots, x_6) = (1.2, 2.5, 3.1, 3.8, 10.4, 4.9)$.

(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.

5. (6 баллов) 61 студент университета на определённом курсе были случайным образом распределены в две учебные группы размером 33 и 28 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	33	72.0	7.0
Группа 2	28	73.5	6.5

(a) (2 балла) Посчитайте 95% доверительный интервал для математического ожидания экзаменационных оценок студентов группы 2.

(b) (4 балла) Используйте подходящий тест гипотез, чтобы определить, больше ли средний балл студентов группы 2, чем средний балл студентов группы 1.

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	221
Естественные науки	200	120

(a) (2 балла) Посчитайте 98% доверительный интервал для истинной доли студентов естественных наук, которые поддерживают новое расписание экзаменов.

(b) (2 балла) Посчитайте 97% доверительный интервал для истинной разности долей положительных ответов между студентами гуманитарных и естественных наук.

(c) (5 баллов) Проведите двусторонний тест на уровне значимости 3%, 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: $\mathcal{X} = \{x_1, \ldots, x_n\}$ и $\mathcal{Y} = \{y_1, \ldots, y_m\}$, 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. Следующие вопросы исследуют различные аспекты этих выборок.

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

$$(\bar{y} - 0.42\,\sigma_y, \bar{y} + 0.42\,\sigma_y)$$

   является приблизительно 94% 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 $\sigma_x^2 = 1.44\sigma^2$ и $\sigma_y^2 = \sigma^2$ соответственно. Найдите n такое, что:

$$P\left(\bar{X} + \bar{Y} - 0.28\sigma < \mu_x + \mu_y < \bar{X} + \bar{Y} + 0.28\sigma\right) = 0.75.$$

   (d) (7 баллов) Мы хотим проверить гипотезу о том, что $\mu_x = \mu_y + \Delta$ против двусторонней альтернативной гипотезы. Предположим, что известны данные: $n = 7$, $m = 10$, $\bar{x} = 11$, $\bar{y} = 4.5$, $\Delta = 2.5$, $\sigma_x = 2.8$, $\sigma_y = 3.2$. Используйте данные и проведите тест, используя уровни значимости $\alpha = 1\%,\ 4\%$.

   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	$0.05+\theta$	0.05
$Y=0$	0	$0.15-6\theta$	$0.05+\theta$
$Y=1$	$0.2+3\theta$	0.4	$0.1+\theta$

(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

$$P(Y=-1 \mid X+Y=0).$$

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

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

(e) (5 баллов) Предположим, что у вас есть реализация случайной выборки: $\mathcal{X}=(x_1,\ldots,x_n)$, где каждая x_i получена из закона распределения случайной величины X из таблицы выше.

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

(f) (6 баллов) Рассмотрите $\hat{\theta}_1 = Y$ 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).