ВШБ Бизнес-информатика: ТВиМС 2025.
Лист задач для самостоятельного решения #11.
Точечные оценки. Метод моментов. Интервальные оценки.

1. Независимые случайные величины X_1, X_2 и X_3 имеют одинаковое матожидание m но разные стандартные отклонения $\sigma, 2\sigma$ и 3σ соответственно. В качестве оценки матожидания мы рассматриваем три варианта: $\hat{\theta}_1 = \frac{1}{3}X_1 + \frac{1}{3}X_2 + \frac{1}{3}X_3$, $\hat{\theta}_2 = \frac{1}{6}X_1 + \frac{1}{5}X_2 + \frac{1}{4}X_3$, $\hat{\theta}_3 = \frac{1}{6}X_1 + \frac{1}{3}X_2 + \frac{1}{2}X_3$. Какая из этих оценок лучше? Указание - проверить несмещенность, у несмещенных сравнить дисперсии.

2. Пусть X_1, X_2, X_3 — случайная выборка из генеральной совокупности с математическим ожиданием μ и дисперсией σ^2. Рассмотрим следующую оценку дисперсии σ^2:

   $$\hat{\theta} = c(X_1 - X_2)^2.$$

   Найти константу c такую, что $\hat{\theta}$ является несмещенной оценкой для σ^2.

3. Случайные величины X_1 и X_2 распределены по одному закону и независимы. Среди всех несмещенных оценок матожидания вида $c_1X_1 + c_2X_2$ найти оценку с наименьшей дисперсией.

4. Количество покупок, совершаемых клиентами в интернет-магазине за день, подчиняется распределению Пуассона. У нас есть выборка данных по количеству покупок за несколько дней, результаты записаны в таблице. Необходимо определить параметр λ по этой выборке, используя метод моментов.

Количество покупок за день x_i	0	1	2	3	4	5
Количество дней с количеством покупок x_i	10	37	38	22	12	6

5. Найти методом моментов по выборке $x_1, x_2, \ldots, x_n$ точечную оценку параметра p биномиального распределения.

6. При условии равномерного распределения случайной величины X произведена выборка:

3	5	7	9	11	13	15	17	19	21	21	16	15	26	22	14	21	22	18	25

   Найти оценку параметров a и b по методу моментов.

Для определения среднего возраста своих клиентов крупный производитель одежды провёл случайную выборку из 50 клиентов и получил $\bar{x} = 36$. Известно, что стандартное отклонение генеральной совокупности $\sigma = 12$:

1. Постройте 95% доверительный интервал для среднего возраста μ всех клиентов.

2. Предположим, что требуется, чтобы 95% доверительный интервал был строго равен $[\bar{X} - 2, \bar{X} + 2]$. Какой размер выборки для этого потребуется?

Проведена случайная выборка из 200 студентов. 30 из них говорят, что им "очень нравится"статистика.

1. Вычислите долю студентов в этой выборке, которые говорят, что им "очень нравится"статистика, и затем постройте 95% доверительный интервал для доли генеральной совокупности.

2. Теперь вы берёте ещё одну случайную выборку в другом учебном заведении. На этот раз в выборке 40 студентов, и 16 из них говорят, что им "очень нравится"статистика. Постройте 95% доверительный интервал для этого значения. Подумайте, почему два доверительных интервала могут отличаться.

Рассмотрим случайную выборку размера 20 из распределения $\mathcal{N}(\mu, \sigma^2)$. Наблюдаемые значения выборочного среднего $\bar{X}$ и выборочной дисперсии S^2 равны $\bar{x} = 81.2$ и $s^2 = 26.5$. Найдите соответственно 90%, 95% и 99% доверительные интервалы для среднего генеральной совокупности μ. Отметьте и прокомментируйте, как увеличивается длина доверительных интервалов при увеличении уровня доверия.

Есть опасения по поводу скорости автомобилей, проезжающих по определённому участку шоссе. Для случайной выборки из 7 автомобилей радар зафиксировал следующие скорости (в милях в час):

$$79 \quad 73 \quad 68 \quad 77 \quad 86 \quad 71 \quad 69$$

1. Найдите выборочное среднее и выборочную дисперсию.
2. Указав все необходимые предположения, постройте 90% доверительный интервал для средней скорости всех автомобилей, проезжающих по этому участку шоссе.

Рассмотрим оригинальное клиническое исследование вакцины Солка от полиомиелита, проведённое в 1954 году. Случайным образом одна группа детей получила вакцину (группа лечения), а другая группа получила плацебо (контрольная группа). Пусть p_c и p_T обозначают истинные доли заболевших полиомиелитом в контрольной группе и группе лечения соответственно. Результаты исследования представлены в таблице:

Группа	Количество детей	Количество случаев полиомиелита
Лечение	$200,745$	57
Контроль	$201,229$	199

Постройте 95% доверительный интервал для разности $(p_c - p_T)$. (Не округляйте слишком сильно, оставьте хотя бы $4-5$ знаков в дробной части). Что можно сказать об эффективности вакцины на основе построенного доверительного интервала?