

Практика 2.

Задания по темам:

Таблица исходных данных.

Первичные описательные статистики:

центральная тенденция.

Меры изменчивости.

Перевод измерений в разные шкалы.

Блок 1. Таблица исходных данных

Задание 1 Построение таблицы исходных данных

В исследовании участвовало 10 человек. Для каждого зафиксировали:

- пол (М/Ж);
- возраст (полных лет);
- время выполнения теста (в секундах).

Исходные данные:

М, 19, 120	Ж, 20, 115	М, 20, 130
Ж, 19, 125	М, 21, 118	Ж, 20, 122
М, 22, 140	Ж, 21, 128	М, 19, 116
Ж, 20, 121		

- Представьте данные в виде стандартной таблицы «объект–признак» (строки – испытуемые, столбцы – признаки).
- Сколько всего признаков? Какие из них качественные, какие количественные?

Задание 2 Пропуски и выбросы в таблице

В таблице ниже приведены результаты измерения роста (см) учеников:

162, 170, 168, 175, 180, 999, 165, 172, 169.

- Какое значение явно является **выбросом**? Предложите возможную причину его появления.
- Как следует поступить с выбросом при дальнейшем анализе? (Удалить, заменить, оставить – аргументируйте.)
- Постройте «чистую» таблицу исходных данных без выброса.

Блок 2. Первичные описательные статистики

Центральная тенденция: среднее и медиана

Задание 3 Вычисление среднего и медианы

Дана выборка оценок за контрольную работу:

4, 5, 3, 4, 4, 5, 5, 2, 3, 4, 5, 5, 4, 4, 3.

- (a) Найдите выборочное среднее $\bar{x}$.
- (b) Найдите медиану.
- (c) Сравните их. Что можно сказать о симметрии распределения?

Задание 4 Влияние выброса на среднее и медиану

Исходная выборка: 10, 12, 11, 9, 13, 12, 11.

- (a) Вычислите среднее и медиану.
- (b) Замените последнее значение (11) на 100 (выброс). Как изменились среднее и медиана?
- (c) Какой показатель центральной тенденции устойчивее к выбросам? Почему?

Задание 5 Нахождение пропущенного значения по среднему

Средний балл пяти студентов равен 4.2. Известны оценки четверых: 4, 5, 3, 4. Найдите пятую оценку.

Меры изменчивости: размах, дисперсия, стандартное отклонение

Задание 6 Размах и дисперсия (ручной расчёт)

Выборка: 2, 5, 7, 10, 6.

- (a) Найдите размах.
- (b) Вычислите выборочную дисперсию по формуле $s^2 = \frac{1}{n-1} \sum (x_i - \bar{x})^2$.
- (c) Найдите стандартное отклонение s .
- (d) Поясните, в каких единицах измеряется s по сравнению с дисперсией.

Задание 7 Сравнение изменчивости двух наборов

Две группы спортсменов показали результаты в прыжках в длину (см):

Группа А: 210, 215, 220, 225, 230

Группа Б: 200, 220, 220, 220, 240

- (a) Для каждой группы найдите среднее и стандартное отклонение.
- (b) Какая группа более однородна по результатам? (Сравните коэффициенты вариации $CV = \frac{s}{\bar{x}} \cdot 100\%$.)

Задание 8 Свойства дисперсии

Пусть $x_1, x_2, \dots, x_n$ – выборка, s_x^2 – её дисперсия. Что произойдёт с дисперсией, если:

- (a) К каждому значению прибавить одно и то же число c ?
- (b) Каждое значение умножить на постоянную k ?

Проверьте на примере: выборка 1, 2, 3. Вычислите исходную дисперсию, затем $x_i + 10$ и $5x_i$.

Блок 3. Перевод измерений в разные шкалы

Задание 9 Линейное преобразование шкалы (температура)

Температура в градусах Цельсия переводится в градусы Фаренгейта по формуле $F = 1.8C + 32$.

- (а) Имеется выборка температур в °C: 0, 10, 20, 30, 40. Найдите среднее и стандартное отклонение в °C.
- (б) Переведите каждое значение в °F и заново вычислите среднее и стандартное отклонение.
- (с) Как связаны средние и стандартные отклонения в двух шкалах? Проверьте формулами.

Задание 10 Стандартизация (z-оценка)

По выборке успеваемости: 70, 75, 80, 65, 90, 85, 75.

- (а) Вычислите выборочное среднее $\bar{x}$ и стандартное отклонение s .
- (б) Для каждого значения найдите z -оценку: $z_i = \frac{x_i - \bar{x}}{s}$.
- (с) Чему равны среднее и стандартное отклонение полученного набора z -оценок? Проверьте.
- (д) Объясните, зачем применяют стандартизацию при сравнении данных из разных шкал.

Задание 11 Перевод в процентиля и ранжирование

Даны результаты теста (баллы): 45, 52, 60, 48, 55, 62, 59, 61, 57, 50.

- (а) Упорядочите выборку и найдите 30-й百分иль (значение, ниже которого находятся 30% данных).
- (б) Переведите баллы в **ранги** (наименьшему баллу – ранг 1, наибольшему – ранг 10). При совпадении используйте средний ранг.

Задание 12 Преобразование шкалы для устранения асимметрии

Предположим, вы измерили время реакции (в мс) и получили сильно скошенное распределение: 250, 260, 270, 280, 500, 600, 1200.

- (а) Какое преобразование шкалы (логарифмическое, квадратный корень, обратное) может сделать распределение более симметричным? Почему?
- (б) Примените логарифмическое преобразование (натуральный логарифм) к данным. Сравните среднее и медиану исходных и преобразованных данных.

Комплексное задание (по всем темам)

Задание 13 Полный анализ выборки

Исследователь измерил уровень тревожности по шкале (от 0 до 50) у 12 пациентов:

12, 15, 14, 18, 22, 17, 13, 19, 20, 21, 16, 14

- (a) Представьте данные в виде таблицы исходных данных (пронумеруйте пациентов от 1 до 12).
- (b) Вычислите основные описательные статистики:
 - среднее арифметическое;
 - медиану;
 - моду;
 - размах;
 - дисперсию и стандартное отклонение.
- (c) Преобразуйте шкалу: переведите исходные баллы в z -оценки.
- (d) Если к каждому баллу прибавить 5 пунктов, как изменятся среднее и стандартное отклонение? Проверьте вычислением.
- (e) Сделайте вывод: является ли распределение симметричным? (Сравните среднее и медиану.)

Задание 14 Исследование двух групп и масштабирование

Даны две выборки результатов тестирования по математике (макс. 100 баллов): Группа 1: 85, 90, 78, 92, 88

Группа 2: 65, 70, 68, 72, 75

- (a) Для каждой группы найдите среднее, стандартное отклонение и коэффициент вариации.
- (b) Чтобы сравнить группы, исследователь переводит баллы в шкалу от 0 до 1 по формуле $x' = \frac{x - \min}{\max - \min}$ (min-max нормализация). Выполните это преобразование для каждой группы отдельно.
- (c) Найдите среднее и стандартное отклонение для нормализованных данных в каждой группе. Что изменилось?
- (d) Какой метод приведения к единой шкале (z -оценки или min-max) лучше подходит для сравнения разброса групп? Аргументируйте.