Министерство науки и высшего образования Российской Федерации

Федеральное государственное бюджетное образовательное учреждение

высшего образования «Кубанский государственный университет»

Кафедра вычислительных технологий

**ОТЧЕТ**

о выполнении лабораторной работы №6

по дисциплине «Обработка больших данных»

Тема**:** Проверка статистических гипотез.

Выполнил: ст. гр. 39/1

Козлов Э.Д.

Проверил: преподаватель

Яхонтов А.А.

Краснодар

2025

**Цель:** Ознакомиться с некоторыми статистическими тестами, принципами их работы. Научиться

оценивать нормальность распределения выборки, а также выполнять оценку статистических

гипотез.

### **Задание**

1. Провести дескриптивный анализ данных.

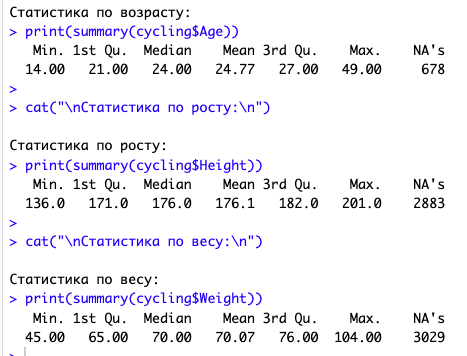
2. Провести проверку на нормальность и дисперсию. Из чего сделать вывод о требуемом тесте.

3. Проверьте гипотезу о среднем весе спортсменов выбранного вида спорта (вид спорта остается по ЛР4).

4. Проверьте гипотезу о равенстве среднего веса женщин (мужчин) в двух разных выбранных видах спорта (сравнение двух независимых выборок – двухвыборочный критерий.).

**Ход работы**

1. **Дескриптивный анализ**



1. **Нормальность и дисперсия**

Нулевая гипотеза — данные имеют нормальное распределение.

Альтернативная гипотеза — данные не имеют нормального распределения.

Нормальность веса для женщин:

Изображение выглядит как текст, Шрифт, белый, чек

Автоматически созданное описание

Нормальность веса для мужчин:

Изображение выглядит как текст, Шрифт, белый, чек

Автоматически созданное описание

Из тестов можно сделать вывод, что нулевая гипотеза – неверна. Распределение веса мужчин и женщин в велоспорте не соответствуют нормальному.

Тест на равенство дисперсия полов:

Изображение выглядит как текст, Шрифт, белый, чек

Автоматически созданное описание

Нулевая гипотеза: дисперсии веса мужчин и женщин равны.

Альтернативная гипотеза: дисперсии не равны.

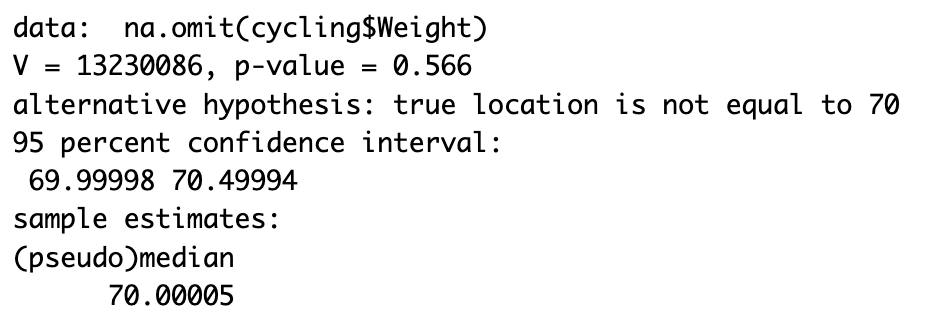
Т.к. p-value < 0.05 делаем вывод, что нулевая гипотеза неверна.

**3. Гипотеза о среднем весе велосипедистов.**

Нулевая гипотеза: медиана веса велогонщиков равна 70 кг.

Альтернативная гипотеза: медиана веса велогонщиков не равна 70 кг.

Исходя из того, что p-value равен 0.566 > 0.05, то нет оснований отвергать нулевую гипотезу. Следовательно, медиана веса велогонщиков статистически не отличается от 70 кг. Доверительный интервал для медианы включает значение 70, что подтверждает это заключение.



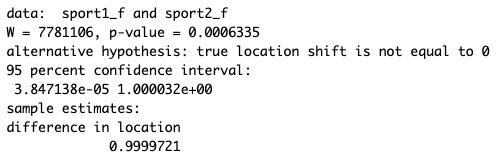
**4. Сравнение веса велосипедистов и легко-атлетов.**

Для женщин атлетов тест нормальности распределения показал:

W = 0.8733, p-value < 2.2e-16

Из чего следует, что данные там распределены тоже ненормально.

Тест Манна-Уитни:



p-value = 0.0006335 — меньше 0.05, значит есть статистически значимая разница в распределении веса между женщинами велоспорта и лёгкой атлетики.

Оценка разницы = 1 кг.