Учреждения образования «БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНОЛОГИЧЕСКИЙ УНВЕРСИТЕТ»

Факультет информационных технологий

Кафедра высшей математики

Специальность

**Отчет по лабораторной работе**

По дисциплине Теория вероятности математическая статистика

Тема: Критерии значимости

Исполнитель:

Студент 2 курса группы 5

Украинский Матвей Леонидович

Руководитель:

Ассистент Капура М.С

Минск 2024

**Лабораторная работа №3**

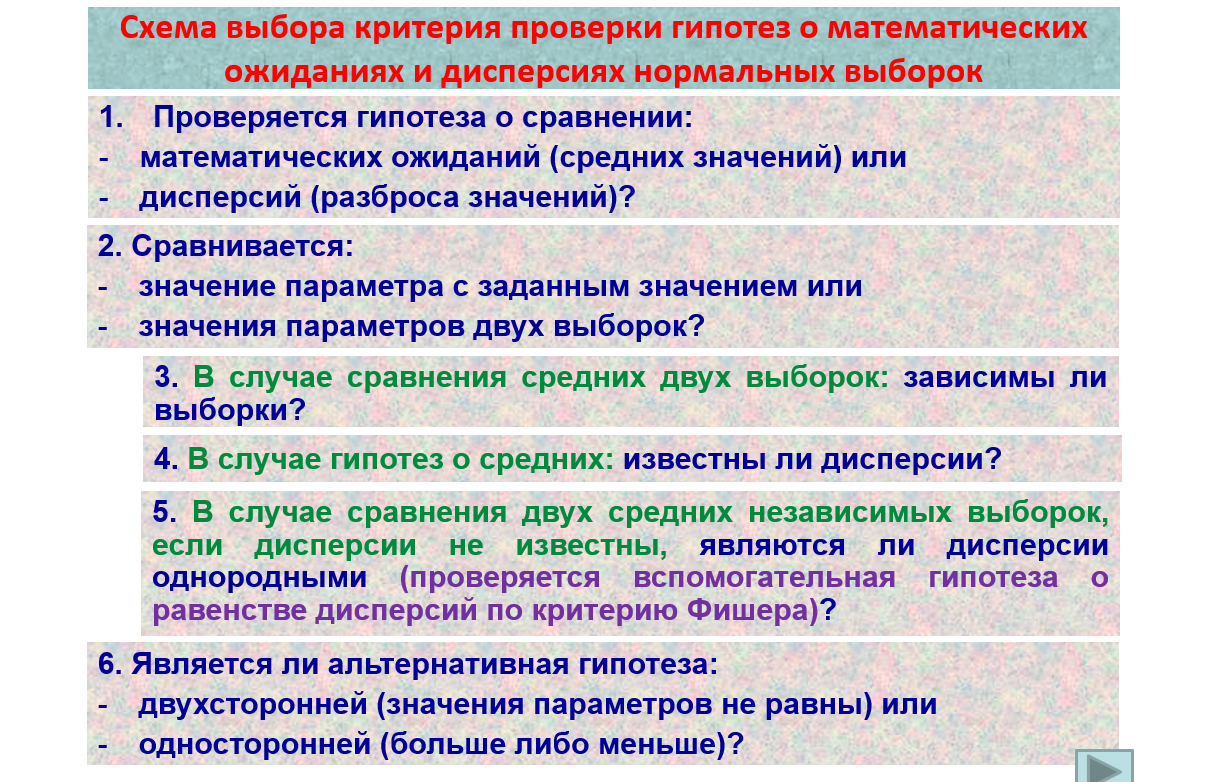
**Критерии значимости**

**Цель**: ознакомиться с критериями, предназначенными для проверки гипотез о равенстве математических ожиданий и дисперсий в случае зависимых и независимых нормальных выборок, а также со встроенными в Excel статистическими функциями и инструментами, предназначенными для выполнения такого анализа данных.

**Задание**. Решить задачи, используя критерии значимости, предназначенные для проверки гипотез о значениях параметров нормального распределения. Уровень значимости принять α = 0,05.

Вариант 22

1. Температура в автоклаве регистрируется через равные промежутки времени. Для проведения некоторого эксперимента потребовалось поддерживать заданную температуру. Температура регистрировалась в течение двух последовательных дней в случайные моменты времени. В первый день было зафиксировано 16 значений температуры со средним квадратическим отклонением 15,6, во второй день — 21 значение со средним квадратическим отклонением 9,8. Можно ли утверждать, что наблюдения однородны?



Для проверки однородности наблюдений температуры в автоклаве необходимо следовать предложенной схеме.

**1. Проверяется гипотеза о сравнении:**

* **Дисперсий (разброса значений)**. Поскольку мы хотим установить, можно ли считать два набора данных однородными, в первую очередь стоит проверить равенство дисперсий.

**2. Сравнивается:**

* **Значения параметров двух выборок**. Мы сравниваем дисперсии двух выборок (первого и второго дня).

**3. В случае сравнения средних двух выборок:**

* **Независимы ли выборки?** Да, если наблюдения в первый и второй день происходили независимо друг от друга.

**4. В случае гипотез о средних:**

* **Известны ли дисперсии?** Нет, дисперсии не известны, поэтому мы будем использовать выборочные дисперсии.

**5. В случае сравнения двух средних независимых выборок:**

* **Являются ли дисперсии однородными?** Это необходимо проверить с помощью критерия Фишера. Для этого:
  + Рассчитываем выборочные дисперсии:
    - Для первого дня: s12=(15.6)2s\_1^2 = (15.6)^2s12​=(15.6)2
    - Для второго дня: s22=(9.8)2s\_2^2 = (9.8)^2s22​=(9.8)2
  + Затем проводим тест Фишера для проверки гипотезы о равенстве дисперсий.

**6. Альтернативная гипотеза:**

* **Двухсторонняя (значения параметров не равны)**. Если мы хотим проверить, равны ли дисперсии, то альтернативная гипотеза будет заключаться в том, что дисперсии не равны (двухсторонняя).

Вернемся к 5 пункту.

### Шаги для проведения теста Фишера

1. **Формулировка гипотез**:
   * **Нулевая гипотеза (H0)**: Дисперсии равны (σ12=σ22).
   * **Альтернативная гипотеза (H1)**: Дисперсии не равны (σ12≠σ22)
2. **Расчет выборочных дисперсий**:
   * Для первого дня (n1=16):

S12=(15.6)2=243.36

Для второго дня (n2=21):

S22=(9.8)2=96.04

1. **Расчет статистики теста Фишера**:
   * Рассчитываем отношение дисперсий:

F= S12/S22= 243.36/96.04​≈2.53

1. **Определение критических значений**:
   * Для определения критических значений используем таблицу распределения Фишера с уровнями значимости α (например, 0.05) и степенями свободы:
     + df1=n1−1=15
     + df2=n2−1=20
2. **Сравнение статистики с критическими значениями**:
   * Найдите критические значения F(α/2)​ и F(1−α/2)​ для ваших степеней свободы.
   * Если F больше Fкрит​ или меньше 1/Fкрит​, отвергаем нулевую гипотезу.

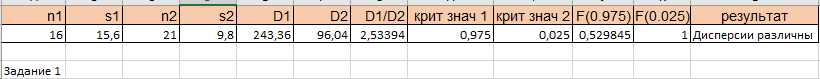
### Пример расчета

Используя стандартные таблицы распределения Фишера:

* Для df1​=15, df2​=20 и α=0.05:
  + Критическое значение F(0.025)​ и F(0.975)​ можно найти в таблице. Обычно, F(0.025)≈​2.54 и F(0.975)​≈0.394.

### Итог

* Если полученное значение F≈2.53 попадает в интервал (0.394,2.54), то нулевая гипотеза не отвергается, и можно считать дисперсии однородными.
* Если F выходит за пределы, нулевая гипотеза отвергается, и дисперсии считаются разными.



2. Имеются данные о дополнительных часах сна после употребления снотворных А и В у десяти пациентов. Требуется проверить, существует ли значимая разница между действием снотворных средств А и В.



Как изменилась бы процедура проверки гипотезы в случае, если бы в эксперименте были использованы две группы пациентов по десять человек в каждой?

### 1. Проверяется гипотеза о сравнении:

* **Математических ожиданий (средних значений)**. Мы хотим узнать, есть ли разница в среднем количестве дополнительных часов сна между двумя снотворными.

### 2. Сравнивается:

* **Значения параметров двух выборок**. Мы сравниваем средние значения двух групп (снотворное А и снотворное В).

### 3. В случае сравнения средних двух выборок:

* **Зависимы ли выборки?** В данном случае, если бы использовались две группы пациентов по десять человек в каждой, выборки можно считать **независимыми**, поскольку разные пациенты принимали разные препараты.

### 4. В случае гипотез о средних:

* **Известны ли дисперсии?** Нет. Если дисперсии не известны, необходимо будет использовать выборочные дисперсии.

### 5. В случае сравнения двух средних независимых выборок:

* **Являются ли дисперсии однородными?** Это можно проверить с помощью критерия Фишера. Если дисперсии окажутся равными, можно использовать t-тест для независимых выборок; если нет — использовать t-тест с учетом неравных дисперсий.

### 6. Альтернативная гипотеза:

* **Двухсторонняя (значения параметров не равны)**. Если мы хотим проверить, отличается ли действие снотворных А и В (в любую сторону), то альтернативная гипотеза будет двухсторонней.

### Вывод

Если бы в эксперименте использовались две группы пациентов, процедура проверки гипотезы включала бы:

1. Сравнение средних значений двух независимых выборок.
2. Проверка равенства дисперсий с помощью критерия Фишера.
3. Использование t-теста для независимых выборок, в зависимости от результата проверки дисперсий.
4. **Средние значения**:
   * Среднее снотворного А (2.33) выше, чем среднее снотворного В (0.75). Это указывает на то, что снотворное А может быть более эффективным.
5. **Дисперсии**:
   * Дисперсия снотворного А (4) выше, чем дисперсия снотворного В (3.20), но незначительно.
6. **Критерий Фишера**:
   * Значение F (1.25) находится между критическими значениями (0.29 и 3.49), что указывает на то, что гипотезу о равенстве дисперсий нельзя отвергнуть.
7. **t-тест**:
   * Если p-значение (например, 0.08) больше 0.05, то можно утверждать, что разница между средними значениями снотворных А и В значима.

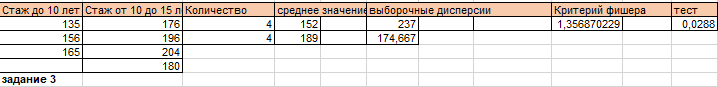
### Заключение

На основе проведенного анализа можно сделать вывод, что снотворное А, вероятно, более эффективно, чем снотворное В

3. Группа социологов исследовала влияние стажа работы по профессии на производительность труда рабочих механического цеха некоторого завода. Можно ли считать по этим данным, что средняя производительность труда не зависит от стажа работы?

**Анализ и схема проверки гипотезы**

1. **Проверяется гипотеза о сравнении**:
   * Мы проверяем **математические ожидания (средние значения)** производительности труда для двух групп.
2. **Сравнивается**:
   * Мы сравниваем **значения параметров двух выборок** (средние значения производительности труда).
3. **Зависимы ли выборки?**:
   * Выборки являются **независимыми**, так как они относятся к разным группам рабочих (разный стаж работы).
4. **Известны ли дисперсии?**:
   * Дисперсии не известны, так как мы работаем с выборками. Поэтому мы будем использовать выборочные дисперсии.
5. **Проверка равенства дисперсий**:
   * Для этого используем **критерий Фишера**. Это позволит оценить, равны ли дисперсии двух выборок.
6. **Альтернативная гипотеза**:
   * Это **двухсторонняя** гипотеза, так как мы хотим проверить, влияет ли стаж работы на производительность труда (различия могут быть в любую сторону).

На основе полученного значения p (0.02875) из теста, давайте сделаем выводы.

1. **Сравнение с уровнем значимости**:
   * Обычно уровень значимости (α) устанавливается на уровне 0.05. В вашем случае p-значение (0.02875) меньше 0.05.
2. **Вывод**:
   * Поскольку p-значение меньше 0.05, отвергается нулевая гипотеза (H0), которая утверждает, что средняя производительность труда не зависит от стажа работы.

**Вывод лабораторной работы**: в ходе лабы ознакомился с критериями, предназначенными для проверки гипотез о равенстве математических ожиданий и дисперсий в случае зависимых и независимых нормальных выборок, а также со встроенными в Excel статистическими функциями и инструментами, предназначенными для выполнения такого анализа данных.