

3.3.3 Сравнение групп в данных

Межгрупповые сравнения

□ Предмет сравнения.

- Сравнение среднего значения с эталонным;
- сравнение средних;
- сравнение дисперсий;
- сравнение функций распределения (проверка однородности распределений).

Выполняется путем проверки соответствующих статистических гипотез

Межгрупповые сравнения

□ Задачи.

- Одновыборочные (выборочное значение сравнивается с эталонным);
- двухвыборочные (сравниваются параметры или распределения двух выборок);
- k -выборочные, $k > 2$ (сравниваются параметры или распределения трех и более выборок).

Межгрупповые сравнения

□ Выборки.

- Связанные;
- несвязанные (независимые).

Статистические критерии для межгрупповых сравнений

- *Параметрические* ;
- *непараметрические*.

Параметрические критерии дают более достоверные результаты, но они и более требовательны к данным

Параметрические критерии

Требования к данным:

- **тип шкалы** – количественная (шкала интервалов);
- **распределение** – близкое к нормальному (нет выбросов, унимодальное, в идеале – симметричное).

Непараметрические критерии

Применяются в случаях, когда не выполнены требования для параметрических критериев:

- **тип шкалы** – не метрическая (порядковая или номинальная);
- либо
- **распределение** отлично от нормального.

Параметрические критерии

Требования к данным:

- **тип шкалы** – количественная (шкала интервалов);
- **распределение** – близкое к нормальному (нет выбросов, унимодальное, в идеале – симметричное).

Непараметрические критерии

Применяются в случаях, когда не выполнены требования для параметрических критериев:

- **тип шкалы** – не метрическая (порядковая или номинальная);
- либо
- **распределение** отлично от нормального.

3.3.3.1 Сравнение двух дисперсий

Задача сравнения дисперсий возникает, если требуется сравнить точность приборов, инструментов, методов измерений и т. п.

Ясно, что предпочтительнее тот инструмент (метод), который обеспечивает меньшее рассеивание результатов измерений, т. е. меньшую дисперсию.

**Параметрические критерии
(для нормально распределенных
совокупностей)**

Постановка задачи

Пусть две генеральные совокупности X и Y имеют нормальное распределение.

Из этих совокупностей независимо извлечены выборки объемов n_1 и n_2 соответственно.

По выборкам найдены исправленные выборочные дисперсии

$$s_X^2 \text{ и } s_Y^2.$$

На практике обычно $s_X^2 \neq s_Y^2$.

Вопрос: значимо или незначимо различие исправленных дисперсий?

Постановка задачи

Требуется при заданном уровне значимости α проверить нулевую гипотезу о равенстве генеральных дисперсий X и Y :

$$H_0: D(X) = D(Y).$$

Постановка задачи

Учитывая, что

$$M(s_X^2) = D(X), \quad M(s_Y^2) = D(Y),$$

Исправленные дисперсии являются несмещенными оценками генеральных дисперсий

нулевую гипотезу можно переформулировать:

$$H_0: M(s_X^2) = M(s_Y^2).$$

В случае справедливости нулевой гипотезы различие в значениях исправленных дисперсий незначимо и объясняется случайными причинами, в частности, случайным отбором объектов выборки.

Следует сделать вывод, что инструменты (методы) имеют одинаковую точность.

Если нулевая гипотеза отвергнута, то различие исправленных дисперсий значимо и не может быть объяснено случайными причинами, а является следствием того, что генеральные дисперсии различны.

Следует сделать вывод, что точность инструментов (методов) различна.

F-критерий (критерий Фишера)

Обозначим:

S_o^2 и S_m^2 – соответственно, бóльшая и меньшая из исправленных дисперсий двух выборок;
 n_o и n_m – объемы выборок, по которым найдены, соответственно, бóльшая и меньшая исправленная дисперсия.

Построим статистику $F = \frac{S_o^2}{S_m^2}$.

F-критерий (критерий Фишера)

Доказано:

при условии справедливости нулевой гипотезы СВ F в пределе имеет распределение Фишера (Фишера–Снедекора) со степенями свободы $k_1 = n_o - 1$ и $k_2 = n_m - 1$.

СВ F используется в качестве статистического критерия проверки гипотезы о равенстве двух дисперсий.

Критическая область (и определяемое ею правило проверки гипотезы) строится по-разному в зависимости от вида альтернативной гипотезы.

F-критерий (критерий Фишера)

□ Первый случай.

Нулевая гипотеза $H_0: D(X) = D(Y)$,
альтернативная гипотеза $H_1: D(X) > D(Y)$.

В этом случае строится правосторонняя критическая область, исходя из требования

$$P(F > F_{кр}(\alpha, k_1, k_2)) = \alpha.$$

СВ F табулирована, поэтому значение $F_{кр}(\alpha, k_1, k_2)$ может быть найдено с помощью статистических таблиц или встроенных функций.

F-критерий (критерий Фишера)

Алгоритм проверки нулевой гипотезы.

1. По данным двух выборок найти наблюдаемое значение критерия

$$F_{набл} = \frac{S_o^2}{S_m^2}.$$

2. По таблице критических точек распределения Фишера найти критическую точку $F_{кр}(\alpha, k_1, k_2)$.
3. Если $F_{набл} \leq F_{кр}$, то нет оснований отвергнуть нулевую гипотезу;
если $F_{набл} > F_{кр}$, то нулевая гипотеза должна быть отвергнута как противоречащая данным наблюдений.

F-критерий (критерий Фишера)

□ Второй случай.

Нулевая гипотеза $H_0: D(X) = D(Y)$,
альтернативная гипотеза $H_1: D(X) \neq D(Y)$.

В этом случае строится двусторонняя критическая область, исходя из требования

$$P(F < F_1(\alpha, k_1, k_2)) = P(F > F_2(\alpha, k_1, k_2)) = \frac{\alpha}{2}.$$

При этом достаточно найти только правую критическую точку F_2 . Для этого можно использовать таблицу критических точек распределения Фишера, выбрав уровень значимости равный $\alpha/2$.

F-критерий (критерий Фишера)

Алгоритм проверки нулевой гипотезы.

1. По данным двух выборок найти наблюдаемое значение критерия

$$F_{\text{набл}} = \frac{s_o^2}{s_m^2}.$$

2. По таблице критических точек распределения Фишера найти критическую точку $F_{кр}(\alpha/2, k_1, k_2)$.
3. Если $F_{\text{набл}} \leq F_{кр}$, то нет оснований отвергнуть нулевую гипотезу;
если $F_{\text{набл}} > F_{кр}$, то нулевая гипотеза должна быть отвергнута как противоречащая данным наблюдений.

F-критерий (критерий Фишера)

Вычисления с помощью пакета Anaconda.

Метод `ppf` класса `f` позволяет вычислить значение квантили распределения Фишера для заданного значения α .

С его помощью можно определить значение $F_{кр}$.

Пример $F_{кр}$ для правосторонней критической области.

```
fcr = scipy.stats.f.ppf(1-alfa, k1, k2)
```

В случае правосторонней критической области требуется выполнение неравенства, противоположного определяющему α -квантиль

Сравнение точности измерений

Пример.

Длина образца измерена при помощи двух движущихся микроскопов. Результаты измерений приведены в таблице:

Номер измерения		1	2	3	4	5	6	7	8	9	10
Результаты измерения, 10^{-9} м	Прибор № 1	100	101	103	98	97	98	102	101	99	101
	Прибор № 2	97	102	103	96	100	101	100			

Проверим с уровнем значимости $\alpha = 0,1$ гипотезу о равенстве генеральных дисперсий результатов измерений.

Сравнение точности измерений

Пример (продолжение).

По данным первой выборки найдем $s_X^2 = 3,778$;
по данным второй выборки – $s_Y^2 = 6,476$.

Поэтому в данном случае

$$s_o^2 = 6,476; \quad s_m^2 = 3,778;$$

$$n_o = 7; \quad n_m = 10;$$

$$k_1 = 6; \quad k_2 = 9.$$

Наблюдаемое значение критерия равно

$$F_{\text{набл}} = \frac{s_o^2}{s_m^2} = 1,714.$$

Сравнение точности измерений

Пример (продолжение).

В иллюстративных целях рассмотрим оба варианта построения альтернативной гипотезы.

1. Нулевая гипотеза $H_0: D(X) = D(Y)$,
альтернативная гипотеза $H_1: D(Y) > D(X)$.

Критическое значение критерия равно

$$F_{кр}(0,1; 6; 9) = 2,55.$$

$F_{\text{набл}} < F_{кр}$, поэтому нет оснований отвергнуть нулевую гипотезу, следует считать ее не противоречащей данным наблюдений.

Нет оснований считать, что первый прибор обеспечивает более высокую точность измерений, чем второй

Сравнение точности измерений

Пример (продолжение).

2. Нулевая гипотеза $H_0: D(X) = D(Y)$,
альтернативная гипотеза $H_1: D(Y) \neq D(X)$.

Критическое значение критерия следует находить для уровня значимости, равного $\alpha/2 = 0,05$:

$$F_{кр}(0,05; 6; 9) = 3,37.$$

$F_{\text{набл}} < F_{кр}$, поэтому нет оснований отвергнуть нулевую гипотезу, следует считать ее не противоречащей данным наблюдений.

Нет оснований считать, что точность измерений первым и вторым прибором различна

Сравнение точности измерений

Пример (продолжение).

В иллюстративных целях рассмотрим оба варианта построения альтернативной гипотезы.

1. Нулевая гипотеза $H_0: D(X) = D(Y)$,
альтернативная гипотеза $H_1: D(Y) > D(X)$.

Критическое значение критерия равно

$$F_{кр}(0,1; 6; 9) = 2,55.$$

$F_{\text{набл}} < F_{кр}$, поэтому нет оснований отвергнуть нулевую гипотезу, следует считать ее не противоречащей данным наблюдений.

Нет оснований считать, что первый прибор обеспечивает более высокую точность измерений, чем второй

Сравнение точности измерений

Пример (продолжение).

2. Нулевая гипотеза $H_0: D(X) = D(Y)$,
альтернативная гипотеза $H_1: D(Y) \neq D(X)$.

Критическое значение критерия следует находить для уровня значимости, равного $\alpha/2 = 0,05$:

$$F_{кр}(0,05; 6; 9) = 3,37.$$

$F_{набл} < F_{кр}$, поэтому нет оснований отвергнуть нулевую гипотезу, следует считать ее не противоречащей данным наблюдений.

Нет оснований считать, что точность измерений первым и вторым прибором различна

Сравнение точности измерений

Замечание.

Во втором случае можно было сразу сделать вывод, что не следует отвергать нулевую гипотезу, поскольку эта гипотеза не была отвергнута в случае односторонней критической области.

Рассмотрение второго случая приведено в иллюстративных целях.

Особенности F -критерия

Критерий Фишера – наиболее популярный параметрический критерий для сравнения дисперсий.

При этом он очень чувствителен к предположению о нормальности распределения совокупностей.

Если данные не описываются нормальным законом, то распределение СВ F не сходится к распределению Фишера → это распределение нельзя использовать для проверки нулевой гипотезы.

Критерий Левене

Критерий Левене – второй по популярности параметрический критерий.

Статистика критерия рассчитывается заметно более сложно, чем в случае критерия Фишера.

Считается, что критерий Левене менее чувствителен к отклонению данных от нормальности.

Критерий Левене

Статистика критерия Левене имеет вид

$$W = (n_1 + n_2 - 2) \cdot \frac{n_1(\bar{Z}_{1.} - \bar{Z}_{..})^2 + n_2(\bar{Z}_{2.} - \bar{Z}_{..})^2}{\sum_{j=1}^{n_1} (Z_{1j} - \bar{Z}_{1.})^2 + \sum_{j=1}^{n_2} (Z_{2j} - \bar{Z}_{2.})^2},$$

где

$$Z_{1j} = |X_j - \bar{X}|, \quad Z_{2j} = |Y_j - \bar{Y}|,$$

$$\bar{Z}_{i.} = \frac{1}{n_i} \sum_{j=1}^{n_i} Z_{ij} \quad -$$

среднее Z_{ij} по i -й выборке, $i = 1, 2$,

$$\bar{Z}_{..} = \frac{1}{n_1 + n_2} \left(\sum_{j=1}^{n_1} Z_{1j} + \sum_{j=1}^{n_2} Z_{2j} \right) \quad -$$

среднее Z_{ij} по всем выборкам.

Непараметрические критерии
(для совокупностей, распределение
которых отлично от нормального)

Критерий Муда

Для построения статистики критерия следует

- объединить обе выборки в одну и построить общий вариационный ряд;
- ранжировать – присвоить членам полученного ряда соответствующие ранги;
- для построения статистики – использовать ранги выборки, имеющей меньший объем.

При ранжировании следует иметь в виду:
если значения нескольких членов ряда совпадают,
то всем этим членам присваивается один и тот же
ранг, равный среднему арифметическому их
порядковых номеров

Критерий Муда

Статистика критерия Муда имеет вид:

$$M^* = \frac{M - \frac{n \min \cdot (n_1 + n_2 + 1)(n_1 + n_2 - 1)}{12} + \frac{1}{2}}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)(n_1 + n_2 + 2)(n_1 + n_2 - 2)}{180}}}.$$

При $n_1 > 10$, $n_2 > 10$ статистика M^* хорошо описывается стандартным (нормированным) нормальным распределением.

Критерий Муда

Алгоритм проверки нулевой гипотезы.

1. По данным двух выборок найти наблюдаемое значение статистики M^* .
2. По таблице критических точек распределения стандартного нормального закона найти критическую точку $u_{кр}(1-\alpha/2)$.
3. Если $|M^*_{набл}| \leq u_{кр}$, то нет оснований отвергнуть нулевую гипотезу;
если $|M^*_{набл}| > u_{кр}$, то нулевая гипотеза должна быть отвергнута как противоречащая данным наблюдений.

Критерий Муда

Вычисления с помощью пакета Anaconda.

Функция `mood()` модуля `stats` библиотеки `scipy` позволяет проверить гипотезу о равенстве дисперсий двух выборок с помощью критерия Муда.

Возвращает наблюдаемое значение критерия и величину *p-value*.

<https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.mood.html#scipy.stats.mood>

Критерий Муда

Замечание.

Статистика, вычисляемая функцией `mood()` для аппроксимации нормального распределения, немного отличается от приведенной выше формулы. См. программный код (кнопка `source` в описании функции).

Сравнение точности измерений

Пример (продолжение).

```
x = [100, 101, 103, 98, 97, 98, 102, 101, 99, 101, 100,  
      101, 103, 98, 97, 98, 102, 101, 99, 101]  
y = [97, 102, 103, 96, 100, 101, 100, 97, 102, 103, 96,  
      100, 101, 100]
```

```
stat_m, p = scipy.stats.mood(x, y)  
print(stat_m, p)
```

Вывод: -1.36160567916 0.173322362006

Задание.

Нет оснований отвергнуть нулевую гипотезу

Самостоятельно выполнить проверку гипотезы о равенстве дисперсий в соответствии с приведенным выше алгоритмом.