# Практическая работа № 6.
# Проверка гипотез о значениях параметров распределения генеральной совокупности

## 6.1. О содержании и задачах практической работы

Предположим, нам известна функциональная форма закона распределения случайной величины, но неизвестны значения параметров, от которых она зависит. Имеются основания считать, что один из параметров (назовем его $\theta$) имеет значение $\theta_0$. Требуется выяснить, подтверждают или не подтверждают эту гипотезу опытные данные. Или, как говорят, «проверить» гипотезу о значении параметра распределения.

Напомним общепринятый подход к решению поставленной задачи. Положим, у нас есть полученная в результате многократного проведения экспериментов выборка значений исследуемой случайной величины. Процесс проверки есть некоторое правило, согласно которому множество всех результатов наблюдений разбивается на две непересекающиеся части: критическую область и область принятия. Попадание результатов в критическую область считают отрицанием проверяемой гипотезы, а попадание в область принятия – ее подтверждением.

При подтверждении или отрицании гипотезы можно совершить ошибки двух видов. Если мы отвергаем гипотезу в то время, когда она в действительности верна, то мы совершаем ошибку первого рода. Если мы принимаем гипотезу в то время, когда она неверна, то совершаем ошибку второго рода. Если критическая область выбрана, то можно рассчитать вероятность $\alpha$ ошибки первого рода. С ошибкой второго рода сложнее: ее числовое значение $\alpha_2$ можно получить, если конкретизировать «неверность» (выбрать конкретное альтернативное значение параметра).

Величины $\alpha$ и $\alpha_2$ зависят от выбора критической области и при этом опосредовано взаимосвязаны: уменьшение одной из них за счет выбора критической области влечет за собой увеличение другой. Общепринят следующий принцип выбора критической области: при заданном значении $\alpha$ критическая область выбирается так, чтобы $\alpha_2$ было минимальным.

Вышеизложенные соображения нашли отражение в традиционной методике (схеме) «проверки» параметрических гипотез (см. п. 2 подраздела 6.2).

В настоящей практической работе обсуждаются две задачи, связанные с ее использованием:

1. Практическое применение методики: принятие или отклонение гипотезы на основании опытных данных при заданном уровне значимости.
2. Анализ ошибок, к которым может привести принятое решение.

*О первой задаче*. На самом деле, поскольку мы собираемся применять готовую методику, то принятие или отклонение гипотезы является не столько проблемой, сколько рутинной процедурой. Однако в последнее время эту процедуру принято использовать в несколько измененном виде. Поясним, чем вызваны эти изменения и в чем они состоят.

Традиционная процедура «проверки» статистической гипотезы позволяет принять решение о принятии или отклонении основной гипотезы по результатам сравнения наблюдаемого значения с границей критической области. Эти границы определяются через квантили распределения статистики и зависят от заданного уровня значимости $\alpha$. Если будет задано новое значение уровня значимости, то процедуру проверки придется проводить заново.

Поскольку выбор уровня значимости не диктуется теорией и $\alpha$ выбирается из неоднозначно трактуемых практических соображений, то необходимость заново проводить расчеты, если мы решим изменить $\alpha$, может расцениваться как некоторое неудобство использования методики. Чтобы этот недостаток исправить, в настоящее время используют модификацию традиционной схемы, основанную на использовании так называемого $p$-значения (она приведена в п. 3 подраздела 6.2). Использование модифицированной методики по существу позволяет указать диапазон значений ошибок первого рода (уровней значимости $\alpha$), при котором опытные данные не будут противоречить основной гипотезе о равенстве параметра $\theta$ значению $\theta_0$.

*О второй задаче*. Выделим ряд вопросов, относящихся к проблеме анализа ошибок, и далее обсудим их в примерах и заданиях настоящей практической работы.

(1) Пусть из практических соображений максимально приемлемой вероятностью совершить ошибку, состоящую в отклонении верной гипотезы, является $\alpha$, и критическая область выбирается в соответствии с методикой (см. п. 2 подраздела 6.2). Если на основании опытных данных гипотеза будет принята, то с какой вероятностью решение окажется ошибочным (при условии справедливости той или иной альтернативной гипотезы)?

(2) Положим, критическая область определяется практическими соображениями (а не по ошибкам первого и второго рода). К каким ошибкам первого и второго рода может привести ее использование при решении принять или отвергнуть гипотезу о равенстве параметра $\theta$ значению $\theta_0$?

(3) Объем выборки оказывает существенное влияние на вероятность ошибок, возникающих при принятии решения. Как определить число опытных данных, при котором величины ошибок первого и второго рода имеют приемлемые для экспериментатора значения?

## 6.2. Математические понятия и утверждения: краткая информация и ссылки на источники

**1.** Параметрические гипотезы. Основная и альтернативная гипотезы. Уровень значимости. Статистика критерия. Выбор критической области. Решающее правило. Ошибка первого рода. Ошибка второго рода. Мощность критерия [1, с. 178 – 185; 2, с. 247 – 253, 268 – 269].

**2.** Традиционная схема проверки гипотезы о сравнении параметров с эталоном.

*1-й шаг. Формулируем основную и альтернативную гипотезы.*

Основная гипотеза $H_0$: $\theta = \theta_0$.

Альтернативную гипотезу $H_1$ выбираем из следующих вариантов:

$\theta \neq \theta_0$ (двусторонняя альтернатива);

$\theta > \theta_0$ (правосторонняя альтернатива);

$\theta < \theta_0$ (левосторонняя альтернатива).

*2-й шаг. Задаем уровень значимости* $\alpha$.

Выбираем $\alpha$ из множества $\{0{,}05;\ 0{,}0025;\ 0{,}01;...\}$.

*3-й шаг. Выбираем подходящую статистику* $Z(X_1, X_2, ..., X_n)$.

*4-й шаг. Разбиваем множество возможных значений статистики* $Z$ *(область* $G$*) на две непересекающиеся подобласти:* $G_\alpha$ *(критическую) и* $G \setminus G_\alpha$ *(допустимую).*

Способ разбиения зависит от альтернативы (таблица).

*Таблица*

| Альтернатива | Критическая область |
|---|---|
| $H_1$: $\theta \neq \theta_0$ | $G_\alpha = \{Z < z_{\frac{\alpha}{2}}\} \cup \{Z > z_{1-\frac{\alpha}{2}}\}$, где $z_{\frac{\alpha}{2}}$, $z_{1-\frac{\alpha}{2}}$ – квантили распределения статистики $Z$ |
| $H_1$: $\theta > \theta_0$ | $G_\alpha = \{Z > z_{1-\alpha}\}$, где $z_{1-\alpha}$ – квантиль порядка $1-\alpha$ распределения статистики $Z$ |
| $H_1$: $\theta < \theta_0$ | $G_\alpha = \{Z < z_\alpha\}$, где $z_\alpha$ – квантиль порядка $\alpha$ распределения статистики $Z$ |

*5-й шаг. Вычисляем выборочное значение статистики* $z_{\text{выб}} = Z(x_1, x_2, ..., x_n)$.

*6-й шаг. Принимаем решение: принять или отклонить гипотезу* $H_0$.

Принятие решения происходит в соответствии со следующим решающим правилом (критерием проверки):

если $z_{\text{выб}} \in G_\alpha$, то гипотезу $H_0$ отвергаем;

если $z_{\text{выб}} \in V \setminus V_{\text{кр}}$, то гипотезу $H_0$ принимаем.

**3.** Модификация схемы проверки статистических гипотез с использованием $p$-значения.

Метод, основанный на использовании так называемого $p$-значения критерия, позволяет решить для всех уровней значимости, принять или отклонить основную гипотезу одновременно.

*Определение. $p$-значением* $p(x_1, x_2, ..., x_n)$ нулевой гипотезы, проверяемой по выборке с помощью статистики критерия $Z$ и критической области $G_\alpha$, называется наименьший уровень значимости, при котором основная гипотеза при имеющейся выборке отклоняется:

$$p(x_1, x_2, ..., x_n) = \min\{\alpha \mid z_{\text{выб}} \in G_\alpha\}.$$

Здесь $z_{\text{выб}} = Z(x_1, x_2, ..., x_n)$ – выборочное значение статистики.

Для всех значений уровня значимости, таких, что $\alpha \le p(x_1, x_2, ..., x_n)$, основная гипотеза принимается, при всех $\alpha > p(x_1, x_2, ..., x_n)$ – отклоняется. Чем меньше $p$-значение, тем больше оснований отклонить нулевую гипотезу.

Вид формул, по которым вычисляются $p$-значения, зависит от вида критической области.

Справедливо следующее утверждение:

1. Если критическая область правосторонняя, т.е. имеет вид $(z_{1-\alpha}; +\infty)$, где $z_{1-\alpha}$ – квантиль порядка $1-\alpha$, то $p$-значение находится по формуле $p(x_1, x_2, ..., x_n) = P\{Z > z_{\text{выб}} | H_0\}$.

2. Если критическая область левосторонняя, т.е. имеет вид $(-\infty; z_\alpha)$, где $z_\alpha$ – квантиль порядка $\alpha$, то $p$-значение находится по формуле $p(x_1, x_2, ..., x_n) = P\{Z < z_{\text{выб}} | H_0\}$.

3. Если критическая область двусторонняя, т.е. имеет вид $(-\infty; z_{\alpha/2}) \cup (z_{1-\alpha/2}; +\infty)$, то $p$-значение находится по формуле $p(x_1, x_2, ..., x_n) = 2 \cdot \min\{p, 1-p\}$, где $p = P\{Z < z_{\text{выб}} | H_0\}$.

4. Функция мощности критерия [2, с. 266 – 268].

Пусть проверяется гипотеза $H_0$: $\theta = \theta_0$ и $V_{\text{кр}}$ – критическая область. Функцией мощности называется вероятность отклонения гипотезы $H_0$ при различных значениях параметра $\theta$: $M\left(V_{\text{кр}}, \theta\right) = P\left\{Z \in V_{кр}\right\}$. Из определения непосредственно следует, что при значении $\theta = \theta_0$ функция мощности совпадает с ошибкой первого рода: $M\left(V_{\text{кр}}, \theta_0\right) = \alpha$. При простой альтернативной гипотезе $H_1$: $\theta = \theta_1$ значение функции мощности при $\theta = \theta_1$ непосредственно связано с ошибкой второго рода $\alpha_2$, а именно $M(V_{\text{кр}}, \theta_1) = 1 - \alpha_2$.

## 6.3. Библиотечные инструменты языка программирования Python

**Загрузка основных модулей**

```
import numpy as np
import scipy.stats as sts
import matplotlib.pyplot as plt
%matplotlib inline
```

**Проверка гипотезы о значении математического ожидания нормально распределенной генеральной совокупности при неизвестной дисперсии при двусторонней альтернативе**

Функция `ttest_1samp(a, popmean, axis = 0, nan_policy = 'propagate')` модуля `scipy.stats`

Возвращает: выборочное значение статистики $W = \dfrac{\overline{X} - m_0}{S / \sqrt{n}}$; достигаемый уровень значимости – $p$-значение.

Параметры: `a` – выборка; `popmean` – гипотетическое значение математического ожидания; `nan_policy` – задает способ обработки пропущенных значений.

## 6.4. Примеры для совместного обсуждения

Рассмотрим на простых примерах понятия ошибок первого и второго рода.

**Пример 0.** После проведения серии экспериментов была получена выборка объема $n$ из нормально распределенной генеральной совокупност с неизвестными параметрами. Требовалось проверить гипотезу $H_0:\ m = m_0$. По выборке были найдены выборочное среднее $\overline{x}$ и несмещенная оценка дисперсии $s^2$.

1-й случай. Оказалось, что $\overline{x}$ меньше $m_0$, поэтому в качестве альтернативной гипотезы решили использовать гипотезу $H_1: m < m_0$. При каких уровнях значимости гипотеза $H_0$ будет принята, а при каких отклонена? Сопоставьте ваши выводы с формулой нахождения $p$-*значения* п. 3 раздела 6.2.

2-й случай. Оказалось, что $\overline{x}$ больше $m_0$, поэтому в качестве альтернативной гипотезы решили использовать гипотезу $H_1: m > m_0$. При каких уровнях значимости гипотеза $H_0$ будет принята, а при каких отклонена? Сопоставьте ваши выводы с формулой нахождения $p$-*значения* п. 3 раздела 6.2.

3-й случай. В качестве альтернативной гипотезы решили использовать гипотезу $H_1: m \neq m_0$. При каких уровнях значимости гипотеза $H_0$ будет принята, а при каких отклонена? Сопоставьте ваши выводы с формулой нахождения $p$-*значения* п. 3 раздела 6.2.

Смоделируйте выборку объема 100 из нормально распределенной генеральной совокупности с параметрами 2 и 3 и далее используйте ее в качестве опытных данных. Воспользуйтесь полученными при рассмотрении 3-го случая результатами, чтобы найти диапазон значений уровня значимости, при которых при двусторонней альтернативе будет принята гипотеза $H_0: m = 2$.

Для проверки гипотезы $H_0: m = 2$ используйте функцию `ttest_1samp`. Сопоставьте полученные результаты.

**Пример 1.** Предположим, у нас есть выборка объема 25 нормально распределенной, генеральной совокупности со средним квадратичным отклонением, равным 2, и неизвестным математическим ожиданием. Проверятся гипотеза о равенстве математического ожидания значению 10 на уровне значимости $\alpha$. Выборочное среднее, рассчитанное по выборке, оказалось бóльшим 10, поэтому для проверки гипотезы выберем правостороннюю альтернативу. Критическую область определим, действуя строго по схеме (п. 2 подраздела 6.2). Нас интересует опосредованная зависимость $\alpha_2$ от $\alpha$ (значений ошибок второго рода от ошибок первого рода). Найдем аналитическое выражение для этой зависимости и построим ее график (вариант выполнения задания приведен в приложении к практической работе № 6, пример 1).

**Пример 2.** Пусть исходные данные те же, что и в примере 1, но задача поставлена иначе. Зафиксируем уровень значимости $\alpha$ и посмотрим, как меняется ошибка второго рода при изменении альтернативных значений математического ожидания. Найдем аналитическое выражение для этой зависимости и построим ее график (вариант выполнения задания приведен в приложении к практической работе № 6, пример 2).

**Пример 3.** По-прежнему будем рассматривать выборку из нормально распределенной генеральной совокупности со средним квадратичным отклонением, равным 2, и неизвестным математическим ожиданием, только объем выборки конкретизировать не будем (обозначим его $n$). Как и в предыдущих примерах, проверяется гипотеза о равенстве математического ожидания десяти на уровне значимости $\alpha$ (его значение для определенности возьмем равным 0,01). Критическую область выбираем, исходя из правосторонней альтернативы. Будем искать ошибки второго рода при некотором альтернативном значении математического ожидания (например, одиннадцати). Собственно, нас будет интересовать вопрос, как ошибка второго рода зависит от объема выборки. Для того, чтобы прояснить его, найдем аналитическое выражение для этой зависимости и построим ее график (вариант выполнения задания приведен в приложении к практической работе № 6, пример 3).

## 6.5. Задания для самостоятельного выполнения

**Задание 1.** Смоделируйте выборку объема $n=90$ из нормально распределенной генеральной совокупности с параметрами $m_0, \sigma_0$ (значение параметров $m_0, \sigma_0$ возьмите равными длине вашего имени и фамилии соответственно) и далее используйте ее в качестве опытных данных.

Задача состоит в проверке гипотезы $H_0: \sigma=\sigma_0$ при неизвестном математическом ожидании. Исходя из значения выборочной дисперсии, сформулируйте подходящую одностороннюю альтернативу, выберите статистику и укажите закон ее распределения.

1) Не используя имеющиеся в языке программирования Python функции проверки параметрических гипотез, вычислите $p$-значение и найдите диапазон значений $\alpha$, при которых основная гипотеза принимается.

2) Выберите одно из значений $\alpha$ из найденного диапазона и, исходя из соответствующей этому значению критической области (согласно традиционной методике), вычислите ошибку второго рода (предварительно получив аналитическое выражение).

**Задание 2.** В урне содержатся неразличимые на ощупь черные и белые шары. Есть предположение, что число черных шаров в урне вдвое превышает число белых и, значит, вероятность $p$ извлечь из урны черный шар равна $\frac{2}{3}$. Эта гипотеза отклоняется, если при извлечении $n$ шаров (с возвращением) доля черных шаров отличается от $\frac{2}{3}$ более чем на 15%.

1) Найдите аналитическое выражение для функции мощности критерия (зависимости вероятности отклонения гипотезы $H_0: p=\frac{2}{3}$ в условиях заданной выше критической области). Постройте график зависимости функции мощности от параметра $p$ для трех случаев: из урны извлекали 60, 90 и 120 шаров. Для каждого случая вычислите вероятность ошибки первого рода, а также вероятность ошибки второго при альтернативной гипотезе $p=0{,}5$.

2) Постройте графики зависимости вероятностей ошибок первого и второго рода от числа извлеченных шаров (ошибки второго рода рас-

считывайте для альтернативной гипотезы $p = 0{,}5$). Прикиньте графически, сколько шаров нужно извлечь, чтобы величина вероятностей ошибок первого и второго рода не превышала 0,05.