Тема 6. Тестирование гипотез. Часть 2

Пример: как вычисляются вероятности ошибок I и II рода

Рассмотрим снова пример "Завышенная зарплата", в котором проверялась гипотеза о среднемесячной зарплате:

$$H_0 : \mu = 55\,000 \qquad H_a : \mu < 55\,000$$

Ошибка I рода

Предположим, что кто-то задал **следующее правило принятия решения**:

> **Отклонить H_0, если $\bar{X} < 50\,969.$**

Что это значит?

- Если в выборке зарплата окажется **меньше 50 969**, то мы **отвергаем** H_0.
- Но если на самом деле μ=55 000, то при таком правиле возможна **ошибка I рода** — отклонение **истинной** гипотезы.

Вероятность ошибки I рода:

Это вероятность того, что $\bar{X} < 50\,969$, **при условии, что H_0 верна**, т.е. μ=55 000:

$$\alpha = P(\bar{X} < 50\,969 \mid \mu = 55\,000) = P\left(Z < \frac{50\,969 - 55\,000}{\frac{15\,500}{\sqrt{40}}}\right) = P(Z < -1.645) = \boxed{0.05}$$

Вывод:

- Ошибка I рода происходит **в 5% случаев**, если H_0 верна.
- Это и есть заданный **уровень значимости** α=0.05.
- Здесь порог 50 969 был выбран **искусственно**, чтобы показать, **как находить α**, если известны правило отклонения и параметры.

Ошибка II рода

Теперь предположим, что **истинное значение средней зарплаты — 40 000**, и мы всё ещё используем **то же самое правило отклонения**: отвергать H_0, если $\bar{X} < 50\,969$.

Что произойдёт?

- При такой ситуации возможна **ошибка II рода**: H_0 **не будет отклонена**, хотя она **ложна**, и верна альтернативная гипотеза μ < 55 000.

Вероятность ошибки II рода:

Обозначим её через β:

$$\beta = P(\bar{X} > 50\,969 \mid \mu = 40\,000) = P\left(Z > \frac{50\,969 - 40\,000}{\frac{15\,500}{\sqrt{40}}}\right) = P(Z > 4.476) \approx \boxed{3.8 \cdot 10^{-6}}$$

Это крайне малая вероятность — около 0.0000038.

Мощность критерия (power of the test):

Мощность показывает, насколько хорошо тест **обнаруживает отклонение от H_0**, когда оно действительно есть.

$$\text{Power} = 1 - \beta = 1 - 0.0000038 = \boxed{0.9999962}$$

То есть, **в 99.9996% случаев тест верно отклонит H_0**, если зарплата действительно ниже 55 000.

Вывод и переход

Таким образом, мы:

- **Рассчитали α**: вероятность **отклонения истинного H_0**;
- **Рассчитали β**: вероятность **не отвергнуть ложный H_0**;
- Посчитали **мощность теста** =1−β.

В следующих разделах мы будем рассматривать **четыре типа статистических тестов**:

1. 🔷 **Тесты для среднего** μ
2. 🔷 **Тесты для доли** p
3. 🔷 **Тесты для разности средних** μ1−μ2

1. Проверка гипотез для параметров генеральной совокупности μ и p

Тестирование гипотез о **среднем значении** μ

Случай 1: **Известное стандартное отклонение** σ (sigma)

Если $X_1, \ldots, X_n \sim \mathcal{N}(\mu, \sigma^2)$, и σ известна, то:

- Выборочное среднее: $\bar{X} \sim \mathcal{N}(\mu, \sigma^2/n)$,
- Тогда:

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,1)$$

- Это и есть z-статистика — она следует стандартному нормальному распределению.

Теоретическое обоснование:

Это частный случай более общей формулы:

$$\text{Тестовая статистика} = \frac{\text{Оценка} - \text{Значение из } H_0}{\text{Стандартная ошибка оценки}}$$

- **Оценка** — выборочная статистика (например, $\bar{X}$),
- **Параметр** — гипотетическое значение, утверждаемое в H_0,
- **Стандартная ошибка** — стандартное отклонение выборочной статистики.

📌 В нашем случае:

- Оценка = $\bar{X}$,
- Параметр = μ_0,
- SE = $\sigma/\sqrt{n}$ (по закону распределения выборочного среднего при известной σ).

Случай 2: **Неизвестное стандартное отклонение** σ (sigma)

В реальной жизни σ почти никогда не известна, поэтому её заменяют на выборочное стандартное отклонение s. Тогда z-распределение **больше не применимо**.

Вместо него используется **t-статистика**, и она подчиняется **распределению Стьюдента с n−1 степенью свободы**:

$$\boxed{t_{\text{ст}} = \frac{\bar{X} - \mu_0}{s/\sqrt{n}} \sim t(n-1)}$$

Теоретическое обоснование:

Если $X_1, \ldots, X_n \sim \mathcal{N}(\mu, \sigma^2)$, но σ неизвестна, и мы используем s, то:

- $\bar{X} \sim \mathcal{N}(\mu, \sigma^2/n)$,
- s^2 — несмещённая оценка дисперсии,
- Тогда случайная величина:

$$T = \frac{\bar{X} - \mu}{s/\sqrt{n}} \sim t(n-1)$$

Это классический результат, вытекающий из свойства того, что числитель — нормально распределён, а знаменатель (включающий s) — корень из независимой χ2-распределённой величины.

Условия применимости:

Условие	**Пояснение**
SRS	Выборка должна быть простой случайной
$n \geq 30$	Тогда по ЦПТ можно считать $\bar{X} \sim \mathcal{N}$
$n < 30$	Нужно предполагать, что $X \sim \mathcal{N}$
σ известна	Используется **z-статистика**
σ неизвестна	Используется **t-статистика**

Проверка гипотез для **доли в генеральной совокупности p**

Основная идея

Методика тестирования гипотез для доли аналогична тесту для среднего значения μ. Разница — в **оценке параметра** и **формуле стандартной ошибки**.

Шаги проверки гипотезы для доли

(1) Формулируем гипотезы

- **Нулевая гипотеза**:

$$H_0 : p = p_0$$

— утверждает, что доля в популяции равна какому-то конкретному значению.

- **Альтернативная гипотеза**:

- $H_a : p > p_0$ — односторонняя (вправо),
- $H_a : p < p_0$ — односторонняя (влево),
- $H_a : p \neq p_0$ — двусторонняя.

(2) Задаём уровень значимости α

- Стандартные значения: 0.05, 0.01, 0.1

(3) Собираем выборку, рассчитываем выборочную долю $\hat{p}$

- Объём выборки: n
- Количество успехов: X
- Выборочная доля:

$$\hat{p} = \frac{X}{n}$$

(4) Проверяем условия нормального приближения

Для применения нормального распределения (через ЦПТ), выборка должна быть достаточно большой:

$$n \cdot \hat{p} \geq 5, \quad n \cdot (1 - \hat{p}) \geq 5$$

Если это выполнено — используем z-распределение.

(5) Формируем тестовую статистику

Если нулевая гипотеза верна, то $\hat{\boldsymbol{p}}$ распределена нормально:

$$\hat{p} \sim \mathcal{N}\left(p_0, \frac{p_0(1 - p_0)}{n}\right)$$

Следовательно, **z-статистика** будет:

$$\boxed{z_{\text{ст}} = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}}$$

Интерпретация:

- $z_{\text{ст}}$ показывает, **на сколько стандартных отклонений** выборочная доля $\hat{p}$ отличается от заявленной доли p_0.
- Если $|z_{\text{ст}}|$ **слишком велико**, результат считается **маловероятным**, и H_0 **отвергается**.

Решение задачи: два подхода

1. Сравнение с критическим значением:

- Если H_a односторонняя:

$$\text{Отклоняем } H_0 \text{ если } z_{\text{ст}} > z_{\text{кр}} \text{ (или } < z_{\text{кр}})$$

- Если H_a двусторонняя:

$$\text{Отклоняем } H_0 \text{ если } |z_{\text{ст}}| > z_{1-\alpha/2}$$

2. Сравнение с p-value:

- Если $\text{p-value} < \alpha$, то H_0 **отвергается**.