ПРОВЕРКА СТАТИСТИЧЕСКИХ ГИПОТЕЗ

ОСНОВНЫЕ ВИДЫ ГИПОТЕЗ

H_0 - основная альтернатива
H_1 - конкурирующая альтернатива

Параметрические гипотезы

Пусть $x = \left(x_1, x_2, \ldots, x_n\right)$ - выборка объема n из распределения $F\left(x/\theta\right) = P\left\{\xi < x/\theta\right\}$. Пусть θ – скалярный неизвестный параметр.

1. **Простая гипотеза**

$$H_0 : \theta = \theta_0 \quad H_1 : \theta = \theta_1 \tag{1}$$

2. **Сложная односторонняя гипотеза**

$$H_0 : \theta = \theta_0 \quad H_1 : \theta > \theta_0 \tag{2}$$

3. **Сложная двусторонняя гипотеза**

$$H_0 : \theta = \theta_0 \quad H_1 : \theta \neq \theta_0 \tag{3}$$

Непараметрические гипотезы

1. **Гипотеза согласия**

Выборка $x=\left(x_1,x_2,\ldots,x_n\right)$ получена из неизвестного распределения $F(x)$, $F_0(x)$ некоторая заданная функция распределения.

$$H_0: F(x)=F_0(x) \qquad H_1: F(x)\neq F_0(x) \qquad (4)$$

2. **Гипотеза однородности**

Имеются *k* независимых выборок

$$x^{(1)}=\left(x_1^{(1)},x_2^{(1)},\ldots,x_n^{(1)}\right), \quad \ldots \quad ,x^{(k)}=\left(x_1^{(k)},x_2^{(k)},\ldots,x_n^{(k)}\right),$$

$F_i(x)$, $i=1,2,\ldots,k$ – неизвестные функции распределения для этих выборок.

$$H_0: F_1(x)=F_2\left(X\right)=\ldots=F_k(x) \qquad H_1: F_1(x)\neq F_2\left(X\right)\neq\ldots\neq F_k(x) \qquad (5)$$

3. Гипотеза независимости

Выборка $X = \{(x_1 y_1), (x_2 y_2), ..., (x_n y_n)\}$ получена из наблюдений двумерной случайной величины $\xi = (\xi_1, \xi_2)$. $F_\xi(x, y) = P\{\xi_1 < x, \xi_2 < y\}$ - неизвестная функция распределения случайной величины ξ. $F_{\xi_i}(x) = P(\xi_i < x),\ i = 1,2.$

$$H_0 : F_\xi(x, y) = F_{\xi_1}(x) F_{\xi_2}(y) \quad H_1 : F_\xi(x, y) \neq F_{\xi_1}(x) F_{\xi_2}(y) \qquad (6)$$

4. Гипотеза случайности

Выборка $x = (x_1, x_2, ..., x_n)$ получена из неизвестного распределения $F(x)$.

$F_x(x_1, x_2, ..., x_n)$ - функция распределения вектора $x = (x_1, x_2, ..., x_n)$.

$$H_0 : F_x(x_1, x_2, ..., x_n) = F(x_1) F(x_2) .. F(x_n)$$
$$H_1 : F_x(x_1, x_2, ..., x_n) \neq F(x_1) F(x_2) .. F(x_n) \qquad (7)$$

Определение
Множество значений выборки χ_0, по которым принимается альтернатива H_0 называется **областью принятия** H_0. Множество значений выборки χ_1, по которым принимается альтернатива H_1 называется **критической областью** для H_0.

Пусть d_0 - решение в пользу H_0, d_1 - решение в пользу H_1.

Определение
Ошибкой первого рода называется ошибка, когда принимается H_1, а на самом деле верна гипотеза H_0.

$$P\left(d_1 / H_0\right) = \alpha \tag{8}$$

Определение
Ошибкой второго рода называется ошибка, когда принимается H_0, а на самом деле верна гипотеза H_1.

$$P\left(d_0 / H_1\right) = \beta \tag{9}$$

Определение Мощностью критерия для проверки гипотезы называется вероятность

$$P\left(d_1 / H_1\right) = W, \ \ W = 1 - \beta \tag{10}$$

Определение
Значение вероятности α называют **уровнем значимости** критерия

$$\alpha = (0{,}001;\ 0{,}01;\ 0.05;\ 0.1)$$

Основные этапы задачи проверки гипотез

1. Задание H_0, H_1. (H_0 – наиболее важное утверждение)
2. Выбор уровня значимости α.
3. Выбор статистической характеристики $Z(x)$, на основании которой проверяется гипотеза.
4. Построение критической области χ_1 и области принятия χ_0.
5. Построение решающего правила для проверки гипотезы

$$Z(x) \in \chi_0, \ \Rightarrow H_0$$
$$Z(x) \in \chi_1, \ \Rightarrow H_1$$