Testowanie hipotez - podstawowe informacje

Niech $\mathbf{X} = (X_1, \ldots, X_n)$ będzie próbą losową na przestrzeni $\mathcal{X}$, zaś $\mathcal{P} = \{P_\theta, \theta \in \Theta\}$ rodziną rozkładów prawdopodobieństwa określonych na przestrzeni prób $\mathcal{X}$.

Będziemy zakładać, że mamy do czynienia z *hipotezami parametrycznymi*, tzn. z hipotezami dotyczącymi wartości nieznanego parametru $\theta \in \Theta$ rozkładu (np. wartości oczekiwanej, wariancji itd).

Definicja 1. *Hipotezą zerową* nazywamy hipotezę $\theta \in \Theta_0 \subset \Theta$, której prawdziwość chcemy zweryfikować na podstawie obserwacji. Hipoteza alternatywna jest postaci $\theta \in \Theta_1 = \Theta \setminus \Theta_0$.

Hipoteza prosta to hipoteza, której odpowiada jedna wartość, np. $H_0: \theta = 2$, *hipoteza złożona* to hipoteza, ktorej odpowiada więcej niż jedna wartość, np. $H_0: \theta > 4$.

Definicja 2. *Obszar krytyczny testu* jest obszarem odrzucenia hipotezy zerowej. Najczęściej ma on postać $K = \{\mathbf{X} \in \mathcal{X} : T(\mathbf{X}) > c\}$, gdzie c jest *poziomem krytycznym testu*, wyznaczonym przez kwantyl rozkładu, z jakiego pochodzi statystyka testowa przy założeniu prawdziwości hipotezy zerowej (zależy on od przyjętego poziomu istotności testu α).

Podsumowując, aby przeprowadzić test statystyczny, musimy mieć:

1. hipotezę zerową H_0 i hipotezę alternatywną H_1,
2. poziom istotności $\alpha \in (0, 1)$ - bardzo mała liczba, np. standardowo 0,05,
3. statystykę testową $T(\mathbf{X})$,
4. obszar krytyczny K.

Decyzja: jeżeli $T(\mathbf{X}) \in K$, to odrzucamy hipotezę H_0, jeżeli $T(\mathbf{X}) \notin K$, to nie mamy podstaw do odrzucenia hipotezy H_0.

Definicja 3. *P-wartość (p-value)* to graniczny poziom istotności - najmniejszy, przy którym zaobserwowana wartość statystyki testowej prowadzi do odrzucenia hipotezy zerowej.

Jest to więc taki poziom istotności, przy którym zmienia się decyzja testu (zwiększając poziom istotności testu α od zera: najpierw dla najmniejszych poziomów istotności nie mamy podstaw do odrzucenia H_0, a później, po przekroczeniu p-wartości, zaczynamy odrzucać H_0).

P-wartość pozwala bezpośrednio ocenić wiarygodność hipotezy: im p-wartość jest większa, tym bardziej hipoteza H_0 jest prawdziwa. Mała p-wartość świadczy przeciwko hipotezie zerowej.

Znajomość p-wartości pozwala przeprowadzić testowanie dla dowolnej wartości α:
-odrzucamy H_0, gdy

$$p\text{-wartość} \leqslant \alpha,$$

-nie mamy podstaw do odrzucenia H_0, gdy

$$p\text{-wartość} > \alpha.$$

Testy t-Studenta

Są to parametryczne testy istotności dla jednej lub dwóch prób, polegające na testowaniu hipotez dotyczących **wartości oczekiwanych**.

- W przypadku jednej próby zawierającej realizacje zmiennej losowej X testujemy równość jej wartości oczekiwanej $\mathbb{E}X = \mu$ pewnej ustalonej liczbie μ_0.
- W przypadku dwóch prób zawierających realizacje zmiennych X oraz Y testujemy równość ich wartości oczekiwanych $\mathbb{E}X = \mu_1$ oraz $\mathbb{E}Y = \mu_2$.

Zakładamy, że pomiary podlegają rozkładowi normalnemu, lub mają dowolny rozkład o ile ich liczebności są dość duże (z reguły $n > 30$ w przypadku jednej próby oraz $n_1, n_2 \geqslant 100$ w przypadku dwóch prób). Czasami też się zakłada, że wariancje w próbach nie różnią się od siebie istotnie, ale to założenie można ominąć wprowadzając korektę.

1. Test t-Studenta dla jednej próby

- Hipotezy

$$\begin{aligned} H_0 :&\quad \mu = \mu_0, \\ H_1 :&\quad \mu > \mu_0, \qquad (1)\\ &\quad \mu < \mu_0, \qquad (2)\\ &\quad \mu \neq \mu_0 \qquad (3) \end{aligned}$$

- Statystyka testowa

$$T = \sqrt{n}\frac{\bar{x} - \mu_0}{s},$$

  gdzie $s^2 = \frac{1}{n-1}\sum\limits_{i=1}^{n}(x_i - \bar{x})^2$ to próbkowe odchylenie standardowe.

- Obszar krytyczny zależy od postaci hipotezy alternatywnej w następujący sposób:

$$\begin{aligned} K_1 &= \left(F^{-1}_{t_{n-1}}\left(1-\alpha\right), +\infty\right), \qquad (1)\\ K_2 &= \left(-\infty, -F^{-1}_{t_{n-1}}\left(1-\alpha\right)\right), \qquad (2)\\ K_3 &= \left(-\infty, -F^{-1}_{t_{n-1}}\left(1-\tfrac{\alpha}{2}\right)\right) \cup \left(F^{-1}_{t_{n-1}}\left(1-\tfrac{\alpha}{2}\right), +\infty\right), \qquad (3) \end{aligned}$$

  gdzie $F^{-1}_{t_{n-1}}(a)$ to kwantyl rzędu a rozkładu t-Studenta z $(n-1)$ stopniami swobody. Jeżeli wariancja rozkładu jest znana, wówczas s zastępujemy przez odchylenie standardowe σ rozkładu, zaś $F^{-1}_{t_{n-1}}(a)$ zastępujemy przez $\Phi^{-1}(a)$ (Φ - dystrybuanta rozkładu standardowego normalnego $N(0,1)$).

2. Test t-Studenta dla dwóch prób niezależnych

- Hipotezy

$$H_0 : \mu_1 = \mu_2, \quad H_1 : \mu_1 \neq \mu_2$$

- Statystyka testowa

$$T = \frac{\bar{x_1} - \bar{x_2}}{s_{\bar{x}_1 - \bar{x}_2}},$$

gdzie

$$s_{\bar{x}_1-\bar{x}_2} = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)},$$

s_1, s_2 to odchylenia standardowe z próbek, zaś n_1, n_2 to liczebności próbek.

- Obszar krytyczny

$$K = \left(-\infty, -F^{-1}_{t_{n_1+n_2-2}}\left(1-\frac{\alpha}{2}\right)\right) \cup \left(F^{-1}_{t_{n_1+n_2-2}}\left(1-\frac{\alpha}{2}\right), +\infty\right)$$

3. Test t-Studenta dla dwóch prób zależnych

- Hipotezy

$$H_0 : \mu_1 = \mu_2, \quad H_1 : \mu_1 \neq \mu_2$$

- Statystyka testowa

$$T = \frac{\bar{d}}{s_{\bar{d}}},$$

gdzie

$$\bar{d} = \frac{1}{n}\sum_{i=1}^{n} d_i,$$

$$d_i = x_{1i} - x_{2i}, \quad i = 1, \ldots, n,$$

$$s_{\bar{d}} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(d_i - \bar{d})^2},$$

zaś x_{1i}, x_{2i} oznaczają wartości cechy X dla i-tego obiektu w pierwszym i drugim badaniu.

- Obszar krytyczny

$$K = \left(-\infty, -F^{-1}_{t_{n-1}}\left(1-\frac{\alpha}{2}\right)\right) \cup \left(F^{-1}_{t_{n-1}}\left(1-\frac{\alpha}{2}\right), +\infty\right)$$

Uwaga: Gdy liczebność próby jest duża (większa od 30), kwantyl rozkładu t-Studenta można zastąpić odpowiednim kwantylem rozkładu standardowego normalnego.

W przypadku dwóch prób możliwe jest testowanie hipotez jednostronnych postaci $H_1 : \mu_1 > \mu_2$, $H_1 \colon \mu_1 < \mu_2$. Wówczas, analogicznie jak w przypadku 1, zmianie ulegają obszary krytyczne.