5 Testy statystyczne

Hipotezy statystyczne

Tradycyjnie, niech θ oznacza parametr modelu statystycznego.

Dotychczasowe rozważania dotyczyły metod estymacji tego parametru (punktowej lub przedziałowej). Teraz zamiast szacować nieznaną wartość parametru będziemy weryfikowali hipotezę mówiącą, że jego “prawdziwa” wartość nie różni się istotnie od zadanej wartości, co zapisujemy: $\theta = \theta_0$, gdzie θ_0 jest ustalone.

Poza samą hipotezą (nazywać ją będziemy hipotezą zerową) musimy jeszcze podać hipotezę alternatywną, czyli ustalić jaka jest nasza decyzja w przypadku odrzucenia hipotezy zerowej.

Przykładowo, dla hipotezy zerowej $H_0 :\ \theta = \theta_0$, możliwe są następujące alternatywy: $H_1 :\ \theta \neq \theta_0$, $H_1 :\ \theta > \theta_0$ lub $H_1 :\ \theta < \theta_0$.

Przykłady układów hipotez

1.

- Hipoteza zerowa: wartość oczekiwana (średnia) badanej cechy **nie różni się istotnie** od 20.
- Hipoteza alternatywna: wartość oczekiwana (średnia) badanej cechy **jest istotnie większa** od 20.

2.

- Hipoteza zerowa: wartości oczekiwane (średnie) badanej cechy w dwóch grupach **nie różnią się istotnie**.
- Hipoteza alternatywna: wartości oczekiwane (średnie) badanej cechy w dwóch grupach **różnią się istotnie**.

3.

- Hipoteza zerowa: **nie ma istotnej zależności** pomiędzy dwoma badanymi cechami.
- Hipoteza alternatywna: **istnieje istotna zależność** pomiędzy dwoma badanymi cechami.

Obszary krytyczne

Konstruując procedurę testową wyznaczamy tzw. **obszar krytyczny** (obszar odrzuceń hipotezy zerowej). Najbardziej typowym jest **prawostronny obszar krytyczny** postaci:

$$R = \{\mathbf{x} : \ T(\mathbf{x}) \geq k\},$$

gdzie T jest statystyką testową, a k oznacza wartość krytyczną.

Stąd jeśli wartość statystyki testowej jest duża (przekracza wartość krytyczną), to odrzucamy hipotezę zerową.

Inne postaci obszarów krytycznych:

- Lewostronny obszar krytyczny:

$$R = \{ \ \mathbf{x} \in \mathcal{X} : \ T(\mathbf{x}) \leq k \ \},$$

- Dwustronny obszar krytyczny:

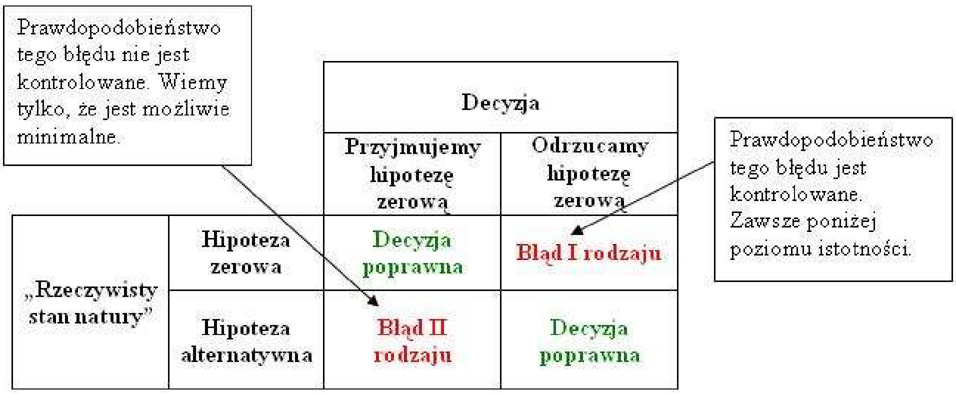
$$R = \{ \ \mathbf{x} \in \mathcal{X} : \ T(\mathbf{x}) \geq k_1 \text{ lub } T(\mathbf{x}) \leq k_2 \ \},$$

Błędy pierwszego i drugiego rodzaju

Przyjmując lub odrzucając hipotezę zerową podejmujemy decyzję, która może być poprawna lub błędna.

Podczas testowania hipotezy zerowej możemy popełnić jeden z dwóch następujących błędów.

1. Odrzucamy hipotezę zerową gdy jest ona prawdziwa - błąd I rodzaju.
2. Przyjmujemy hipotezę zerową gdy jest ona fałszywa - błąd II rodzaju.



Wynik testowania hipotez

Ponieważ decyzja przyjęcia hipotezy zerowej może pociągnąć za sobą popełnienie błędu II rodzaju (prawdopodobieństwo tego błędu nie jest kontrolowane i nawet w najlepszych testach może być bardzo duże), to wynikiem testowania hipotez jest jedna z dwóch decyzji:

1. “odrzucamy hipotezę zerową” tzn. stwierdzamy występowanie istotnych statystycznie różnic (zależności), na poziomie istotności α,
2. “nie ma podstaw do odrzucenia hipotezy zerowej”, tzn. nie stwierdzamy występowania istotnych statystycznie różnic (zależności), na poziomie istotności α.

Testy ilorazu wiarogodności

Załóżmy, że dysponujemy n-elementową próbą, a θ oznacza parametr modelu statystycznego. Rozważamy zagadnienie testowania układu hipotez:

$$H_0:\ \theta \in \Theta_0,$$

$$H_1:\ \theta \in \Theta_1,\ \Theta_0 \cup \Theta_1 = \Theta,\ \Theta_0 \cap \Theta_1 = \varnothing.$$

Ponadto załóżmy, że populacja, z której pochodzi nasza próba ma rozkład absolutnie ciągły.

DEFINICJA

Testem ilorazu wiarogodności nazywamy test z obszarem krytycznym postaci:

$$R = \{\mathbf{x}:\ \frac{\sup_{\theta\in\Theta} L(\theta;\mathbf{x})}{\sup_{\theta\in\Theta_0} L(\theta;\mathbf{x})} \geq k_\alpha\},$$

gdzie wartość krytyczną k_α wyznaczamy tak, aby prawdopodobieństwo błędu pierwszego rodzaju było równe α.

Przykład 1.

Załóżmy, że dysponujemy n-elementową próbą prostą $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)'$ z populacji o rozkładzie wykładniczym z parametrem λ.

Obszar krytyczny testu ilorazu wiarogodności dla układu hipotez:

$$H_0 : \lambda = \lambda_0$$

$$H_1 : \lambda < \lambda_0$$

ma następującą postać:

$$R = \{\mathbf{x}:\ 2\lambda_0 n\bar{x} \leq \chi^2(1-\alpha, 2n)\}.$$

Testy jednostajnie najmocniejsze

DEFINICJA

Funkcję β określoną następująco

$$\beta(\theta) = P_\theta(\boldsymbol{X} \in R)$$

nazywamy **funkcją mocy testu**.

Rozważamy zagadnienie testowania układu hipotez:

$$H_0: \ \theta \in \Theta_0,$$

$$H_1: \ \theta \in \Theta_1, \ \Theta_0 \cap \Theta_1 = \varnothing.$$

DEFINICJA

Test statystyczny nazywamy **testem jednostajnie najmocniejszym na poziomie istotności** α, gdy

1. $$\forall \theta \in \Theta_0: \ \beta(\theta) \leq \alpha,$$
2. $$\forall \theta \in \Theta_1: \ \beta(\theta) = \max.$$

Przykład 2.

Załóżmy, że dysponujemy n-elementową próbą prostą $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)'$ z populacji o rozkładzie wykładniczym z parametrem λ.

Weryfikujemy hipotezę zerową

$$H_0: \ \lambda = \lambda_0.$$

Dla hipotez alternatywnych $H_1: \ \lambda > \lambda_0$ oraz $H_1: \ \lambda < \lambda_0$ **istnieją** testy jednostajnie najmocniejsze.

Dla hipotezy alternatywnej $H_1: \ \lambda \neq \lambda_0$, jednostajnie najmocniejszy test **nie istnieje!!!**

p-wartość

DEFINICJA

p-wartość jest najmniejszym poziomem istotności testu, przy którym odrzucamy hipotezę zerową.

WNIOSEK

1. Jeżeli p–wartość $$, to odrzucamy H_0.
2. Jeżeli p–wartość $>$, to nie ma podstaw do odrzucenia H_0.

Sposób obliczania p-wartości

- Prawostronny obszar krytyczny:

$$\mathrm{P}_0(T \geq T(\mathbf{x})).$$

- Lewostronny obszar krytyczny:

$$\mathrm{P}_0(T \leq T(\mathbf{x})).$$

- Dwustronny obszar krytyczny:

$$2\min\{\ \mathrm{P}_0(T \geq T(\mathbf{x})),\ \mathrm{P}_0(T \leq T(\mathbf{x}))\ \}.$$