Elementy statystyki

STA - Wykład 5

dr hab. Waldemar Wołyński
Wydział Matematyki i Informatyki
Uniwersytet im. Adama Mickiewicza

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

- 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.
- 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**.
- 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.

Błędy pierwszego i drugiego rodzaju

Prawdopodobieństwo tego błędu nie jest kontrolowane. Wiemy tylko, że jest możliwie minimalne.

		Decyzja	
		Przyjmujemy hipotezę zerową	Odrzucamy hipotezę zerową
„Rzeczywisty stan natury”	Hipoteza zerowa	Decyzja poprawna	Błąd I rodzaju
	Hipoteza alternatywna	Błąd II rodzaju	Decyzja poprawna

Prawdopodobieństwo tego błędu jest kontrolowane. Zawsze poniżej poziomu istotności.

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 statystycznych 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. stwierdzamy brak istotnych statystycznie różnic (zależności), na poziomie istotności α.

Metoda ilorazu wiarogodności

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.

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

Załóżmy, że dysponujemy n-elementową próbą prostą $\mathbf{X} = (X_1, X_2, \ldots, X_n)$, $n > 1$ z populacji o rozkładzie normalnym z nieznanymi parametrami μ i σ^2. Obszar krytyczny testu ilorazu wiarogodności dla układu hipotez:

$$H_0 : \mu = \mu_0$$

$$H_1 : \mu < \mu_0$$

ma następującą postać:

$$R = \{\mathbf{x} : \ \frac{\bar{x} - \mu_0}{s}\sqrt{n} \leq t(\alpha, n-1)\}.$$

Uwaga: W modelu jednej próby prostej z populacji o rozkładzie normalnym: funkcja wiarogodności ma postać:

$$L(\mu, \sigma^2; \mathbf{x}) = (2\pi)^{-\frac{n}{2}}(\sigma^2)^{-\frac{n}{2}} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2\right],$$

Ponadto, estymatorami największej wiarogodności parametrów μ i σ^2 są statystyki $\hat{\mu} = \bar{X}$, $\hat{\sigma}^2 = \tilde{S}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$.

p-wartość

p–wartość

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

Wniosek

- *Jeżeli p–wartość* $\leq \alpha$*, to odrzucamy* H_0*.*
- *Jeżeli p–wartość* $> \alpha$*, 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:

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

Sposób podejmowania decyzji