STATYSTYKA dla ZOM II

dr inż Krzysztof Bryś

Wykład 4

Weryfikacja hipotez statystycznych za pomocą testów istotności.

hipoteza statystyczna- przypuszczenie dotyczące nieznanego rozkładu badanej cechy populacji.

hipoteza parametryczna- hipoteza statystyczna dotycząca wartości parametru rozkładu badanej cechy.

weryfikacja- odpowiedź na pytanie czy hipoteza statystyczna jest prawdziwa.

test statystyczny- reguła postępowania, która danej próbie przyporządkowuje decyzję przyjęcia lub odrzucenia badanej hipotezy

H_0- hipoteza zerowa (podlega badaniu)

H_1- hipoteza alternatywna

test istotności- test statystyczny, w którym wnioskowanie odbywa się przy założeniu, że hipoteza H_0 jest prawdziwa. Pozwala jedynie odrzucić H_0 (tzn. przyjąć H_1).

W przypadku weryfikacji hipotez za pomocą testów istotności wskazane jest stawianie jako H_0 hipotez co do których zachodzi podejrzenie o ich fałszywości!

Typy błędów popełnianych przy weryfikacji hipotez:

błąd 1-go rodzaju - odrzucenie prawdziwej hipotezy H_0

błąd 2-go rodzaju - przyjęcie fałszywej hipotezy H_0

poziom istotności α - prawdopodobieństwo popełnienia błędu 1-go rodzaju

β - prawdopodobieństwo popełnienia błędu 2-go rodzaju

moc testu $= 1 - \beta$ - prawdopodobieństwo odrzucenia fałszywej hipotezy H_0.

Jedyny błąd jaki można popełnić weryfikując hipotezę za pomocą testu istotności to błąd 1-go rodzaju!

Zbiór krytyczny W - zbiór wartości taki, że przy założeniu, że H_0 jest prawdziwa: $P(u_n \in W) = \alpha$, gdzie u_n-obliczona wartość statystyki testowej

W praktyce $\alpha \in < 0.01; 0.1 >$.

Uwagi:

1) Przy założeniu, że H_1 prawdziwa: $P(u_n \in W) > \alpha$

2) Jeśli na poziomie istotności α_1 odrzucamy H_0, to na poziomie $\alpha_2 < \alpha_1$ może nie być podstaw do odrzucenia H_0.

Algorytm weryfikacji hipotez za pomocą testu istotności:

1. Wybieramy model.
2. Obliczamy wartość statystyki testowej u_n.
3. Budujemy zbiór krytyczny W (w zależności od postaci H_1).
4. Jeśli $u_n \in W$, to odrzucamy H_0 na poziomie istotności α. W przeciwnym przypadku mówimy, że nie ma podstaw do odrzucenia H_0.

krytyczny poziom istotności α_k - poziom: istotności, przy którym następuje zmiana decyzji weryfikacyjnej:

jeśli $\alpha < \alpha_k$ to mówimy, że nie ma podstaw do odrzucenia H_0 na poziomie istotności $alpha$

jeśli $\alpha > \alpha_k$ to odrzucamy H_0 na poziomie istotności α.

Testy zgodności

Służą do weryfikacji zgodności pomiędzy rozkładem zbioru wartości w próbie a pewnym teoretycznym rozkładem prawdopodobieństwa o dystrybuancie F_0 (gęstości prawdopodobieństwa f_0).

Weryfikowana hipoteza ma postać:
$H_0 : F = F_0$ albo $H_0 : f = f_0$
przeciw
$H_1 : F \neq F_0$ albo $H_1 f \neq f_0$,
gdzie F - nieznana dystrybuanta (f - nieznana gęstość prawdopodobieństwa) zmiennej losowej X reprezentującej badaną cechę.

Test zgodności chi-kwadrat Pearsona

Dzielimy zbiór wartości danej próby na rozłączne przedziały $I_1, \ldots, I_k$. Przy założeniu, że hipoteza H_0 jest prawdziwa,
$p_j = P(X \in I_j) = F_0(\alpha_j) - F_0(\alpha_{j-1})$, gdzie $I_j = (\alpha_{j-1}; \alpha_j)$ dla $j = 1, \ldots, k$.

Obliczamy wartość statystyki testowej:

$$\chi^2 = \sum_{i=1}^{k} \frac{(n_j - np_j)^2}{np_j},$$

gdzie n_j jest liczbą obserwacji należących do przedziału I_j, które zaobserwowano w próbie, $n = \sum_{j=1}^{k} n_j$ jest liczbą wszystkich obserwacji w próbie, np_j nazywamy hipotetyczną liczbą obserwacji z przedziału I_j (jest to liczba obserwacji, które powinny należeć do I_j gdyby H_0 była prawdziwa).

Jeśli obliczona wartość statystyki χ^2 należy do zbioru krytycznego $W = (\chi^2(\alpha, k-1); +\infty)$, to odrzucamy $H_0 : F = F_0$ i przyjmujemy $H_1 : F \neq F_0$. W przeciwnym przypadku mówimy, że nie ma podstaw do odrzucenia H_0.