124. Procedura testowania hipotez – etapy, statystyka testowa i obszar krytyczny, błąd pierwszego i drugiego rodzaju, poziom istotności testu, test istotności dla średniej, test zgodności chi-kwadrat

Weryfikacja hipotez statystycznych odbywa się za pomocą testu statystycznego w oparciu o realizację próby losowej.

Test statystyczny (df) to reguła, która przyporządkowuje każdej realizacji próby losowej jedną z dwóch decyzji dotyczącej hipotezy (albo przyjąć albo odrzucić).

H_0 – hipoteza zerowa

H_1 – hipoteza alternatywna

Test statystyczny wykorzystuje **statystykę testową**:

$$T_n = h(X_n - X_q)$$

Statystyka testowa powinna być tak dobrana by T_n określało prawidłowość lub nieprawidłowość hipotezy.

Obszar krytyczny (K_0)- zbiór tych wartości elementów statystyki testowej dla których ***odrzuca się*** hipotezę.

Obszar przyjęć (K_1)- zb. tych wart. elem. stat. testowej dla których ***przyjmuje się*** hipotezę

Rodzaje błędów:

- **pierwszego rodzaju-** odrzucamy hipotezę gdy jest prawdziwa
- **drugiego rodzaju**- przyjmujemy hipotezę gdy jest fałszywa

Poziom istotności testu

- $P(T_n \in K_0 \mid H_0) = \alpha$ **to** prawdopodobieństwo błędu 1szego rodzaju
- $P(T_n \in K_1 \mid H_1) = \beta$ **to** prawdopodobieństwo błędu 2giego rodzaju

Założenie: statystyka testowa musi mieć znany rozkład

Nie można zbudować takiego testu, który jednocześnie zminimalizowałby α i β

Testy istotności (df) to szczególny przypadek testów w których weryfikuje się tylko jedną hipotezę zerową (nie stawiamy innych hipotez). Konsekwencja: Nie posługujemy się prawdopodobieństwem β, tylko wykorzystujemy poziom istotności testu α.

Tok postępowania posługując się testem istotności:

1. Mając postawioną hipotezę H_0
2. buduje się T_n i wyznacza się jej rozkład, przy założeniu, że H_0 jest prawdziwa
3. Wybór poziomu istotności testu α (im przyjmuje się mniejszy poziom istotności w teście tym trudniej jest odrzucić hipotezę H_0; najczęściej przyjmuje się α=0,05, a mniejsze wartości przyjmuje się w wyjątkowo ważnych badaniach, np. medycznych)

4. Wyznacza się obszar krytyczny K_0
5. Losujemy próbę $t_n = h(x_1, \dots, x_n)$
6. Na podstawie próby obliczamy wartość statystyki testowej
7. Jeśli $t_n \in K_0$ to H_0 odrzucamy na poziomie istotności α.

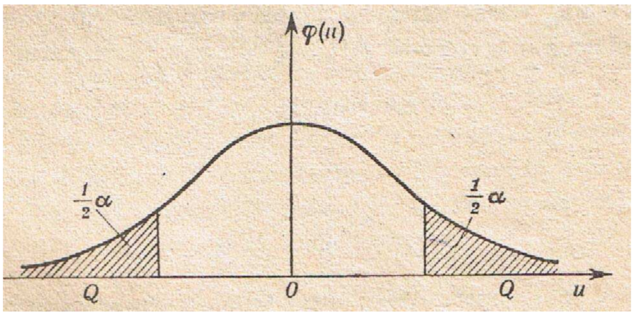
   Jeśli $t_n \notin K_0$ to stwierdzamy brak podstaw do przyjęcia H_0.

Test istotności dla wartości średniej

model 1. (omawiany na wykładzie) Populacja generalna ma rozkład normalny N(m, δ), przy czym odchylenie standardowe δ populacji jest znane. Należy na podstawie wyników próby losowej n-elementowej sprawdzić hipotezę H_0: $m = m_0$ (gdzie m_0 jest konkretną wartością hipotetyczną średniej) wobec hipotezy alternatywnej H_1: $m \neq m_0$.

Na podstawie wyników próby oblicza się wartość statystyki $\bar{x}$, tj. średniej z próby, a następnie wartość zmiennej normalnej standaryzowanej U, wg wzoru: $U = \frac{\bar{x} - m_0}{\delta}\sqrt{n}$

Z tablicy rozkładu N(0,1) wyznacza się następnie taką wartość krytyczną u_α, by dla założonego z góry małego prawdopodobieństwa α (poziomu istotności) zachodziła równość $P\{|U| \geq u_\alpha\} = \alpha$ (rys)



rysunek *Dwustronny obszar krytyczny*

Zbiór wartości U określony nierównością $|U| \geq u_\alpha$ jest obszarem krytycznym tego testu, tzn. gdy z próby otrzymamy taką wartość u, że $|u| \geq u_\alpha$, to hipotezę H_0 odrzucamy. Gdy zaś $|u| < u_\alpha$ to nie ma podstaw do odrzucenia hipotezy H_0.

$U_{kryt} = U\left(1 - \frac{\alpha}{2}\right)$ kwanty rozkładu normalnego standardowego (0,1) rzędu $\left(1 - \frac{\alpha}{2}\right)$

$$K_0 = \left(-\infty, -u\left(1 - \frac{\alpha}{2}\right)\right\rangle \cup \left\langle u\left(1 - \frac{\alpha}{2}\right), +\infty\right)$$

Powyższy test z tzw. dwustronnym obszarem krytycznym stosujemy jedynie dla takiego przypadku hipotezy alternatywnej, w której występuje nierówność $m \neq m_0$. Gdy hipoteza alternatywna ma postać H_1: $m < m_0$, to stosujemy testy istotności z tzw. lewostronnym obszarem krytycznym określonym nierównością $U \leq u_\alpha$. Dla hipotezy alternatywnej postaci H_1: $m > m_0$, stosujemy test ist. z tzw. prawostronnym obszarem krytycznym określonym nierównością $U \geq u_\alpha$.

model 2.

Populacja generalna ma rozkład normalny N(m, δ), przy czym odchylenie standardowe **δ** populacji jest ***nie znane*.**

Weryfikacja hipotezy H_0 w oparciu o ***małe***, n-elementowe próby losowe. *(Wzór do obliczania- odwołuję do literatury.)*

model 3. Populacja generalna ma rozkład normalny N(m, δ) lub dowolny inny rozkład o średniej wartości m i o skończonej, ale ***nieznanej wartości wariancji*** **δ^2**.

Na podstawie dużej próby losowej (n co najmniej rzędu kilku dziesiątków) z tej populacji należy zweryfikować hipotezę H_0. Test istotności jest analogiczny jak w modelu 1, z tą różnicą, że zamiast wartości δ przyjmuje się wyznaczoną z dużej próby wartość s.

<u>Test zgodności Chi-kwadrat</u>

- najczęściej stosowany test nieparametryczny.
- służy on do weryfikowania hipotezy, że obserwowana cecha X w zbiorowości generalnej ma określony typ rozkładu, np. dwumianowy, Poissona, normalny itd. Może to być typ rozkładu skokowego lub ciągłego.
- statystyka, jakiej używa się przy weryfikacji hipotezy o zgodności próby wyników z rozkładem populacji, ma rozkład asymptotyczny χ^2

Test zgodności chi-kwadrat stosuje się:

- gdy dane pochodzą z dużej n-elementowej próby wyznaczonej w sposób niezależny
- gdy dane są przedstawione w postaci szeregu rozdzielczego o r przedziałach klasowych, o liczebnościach przedziałów n, …, n spełniających warunek
  $n_1 + n_2 + \ldots + n_r = n$. Na ogół przyjmuje się, że $n_i > 5$, $i = 1, 2, \ldots, r$
- gdy rozkład hipotetyczny może być zarówno rozkładem typu ciągłego, jak i skokowego.

Postać statystyki sprawdzającej hipotezy H0:

$$\chi^2 = \sum_{i=1}^{r} \frac{(ni - npi)^2}{npi}$$

gdzie:

- pi – prawdopodobieństwo, że cecha X przyjmuje wartość należącą do i-tego przedziału klasowego
- npi – liczba jednostek, które powinny znaleźć się w i-tym przedziale przy założeniu, że cecha ma rozkład zgodny z hipotezą (***liczebność hipotetyczna***)

statystyka ta ma rozkład χ^2 o *k=(r–s–1)* </div>

gdzie:

- k – ilość stopni swobody
- s – liczba parametrów do wyznaczenia na podstawie próby
- r – liczba przedziałów klasowych

χ^2 oznacza wartość empiryczną statystyki

Statystka ta stanowi rozbieżność pomiędzy rozkładem empirycznym a teoretycznym, co oznacza że zbyt duże wartości χ^2 powodują odrzucenie hipotezy zerowej.

Postać zbioru krytycznego:

$K_0 = \left\langle \chi^2(1-\alpha, r-s-1), +\infty \right\rangle$ odczytujemy z tablic

Literatura

1. *wykłady 3.12.2008r. oraz 10.12.2008r.*
2. *„Statystyka matematyczna modele i zadania" Jerzy Greń, Wyd. VII, Warszawa 1982 PWN*
3. *http://mfiles.pl/pl/index.php/Test_zgodno%C5%9Bci_chi-kwadrat dostęp 29.12.2010r.*