Metody matematyczne w transporcie

Testy istotności różnic między parametrami cechy dla dwóch populacji

Marzena Nowakowska
Wydział Zarządzania i Modelowania Komputerowego

Podstawowe pojęcia teorii estymacji

Przedmiotem teorii estymacji jest wnioskowanie o wartości parametru lub parametrów, od których zależy rozkład badanej zmiennej losowej.

Estymacja polega na oszacowaniu wartości nieznanego parametru *PA* populacji, na podstawie *n* wyników zaobserwowanych w próbie. Przykłady parametrów: średnia, wariancja, proporcja, parametry strukturalne w zależności regresyjnej.

Zadaniem **estymacji przedziałowej** jest konstrukcja przedziału liczbowego, zwanego przedziałem ufności, który z określonym prawdopodobieństwem (jak największym) będzie zawierał prawdziwą wartość parametru. Ww. prawdopodobieństwo nosi nazwę współczynnika ufności i określa się je w procentach jako wartości wyrażoną zależnością (1-α), α – prawdopodobieństwo popełnienia błędu I rodzaju. Im większa tego współczynnika, tym szerszy przedział ufności, więc mniejsza dokładność estymacji parametru.

Zadaniem **estymacji punktowej** jest znalezienie wartości liczbowej U_n zwanej estymatorem, którą będzie można uznać za najlepsze przybliżenie parametru *PA* populacji.

Estymator = statystyka szacująca wartość parametru.
Statystyka = zmienna losowa będąca funkcją zbioru wartości z próby.

Cechy najlepszego estymatora

Estymator U_n parametru *PA* jest najlepszy jeśli posiada trzy własności:

- **Jest nieobciążony, czyli $E(U_n) = PA$**

  Oznacza to, że jeśli dla serii prób wyznaczy się wartości estymatora U_n, to średnia z tych wartości jest równa szacowanemu parametrowi.

- **Jest zgodny, czyli jest stochastycznie zbieżny do *PA***

  $$\forall \varepsilon > 0 \ \lim_{n \to \infty} P(|U_n - PA| < \varepsilon) = 1$$

  Oznacza to, że gdy liczebność próby rośnie, to prawdopodobieństwo, że wartość estymatora U_n różni się dowolnie mało od wartości parametru *PA* zbliża się do jedności.

- **Jest najefektywniejszy, czyli ma najmniejszą wariancję.**

  Oznacza to, że jego wartości są bardziej skupione wokół jego wartości średniej niż wartości innych estymatorów.

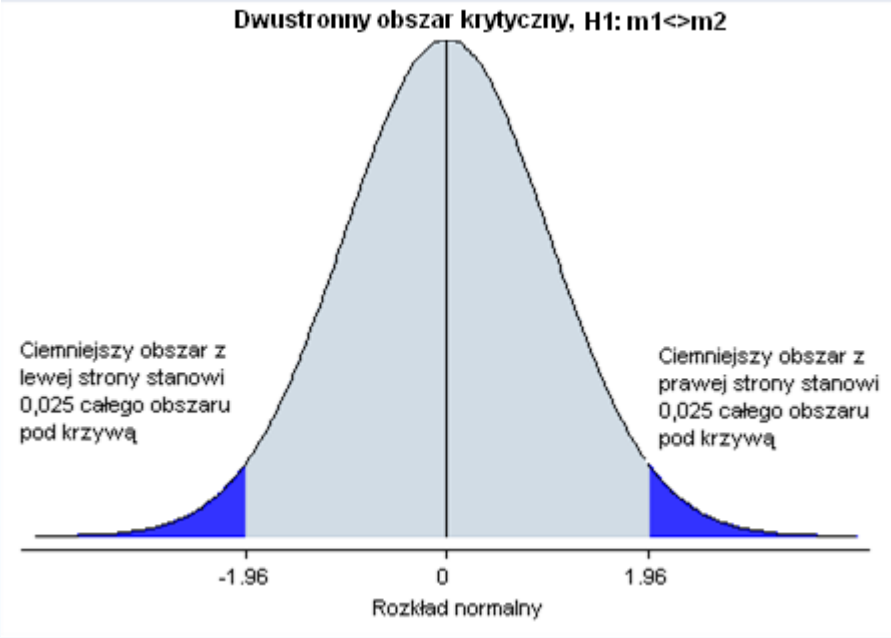
Obszary krytyczne

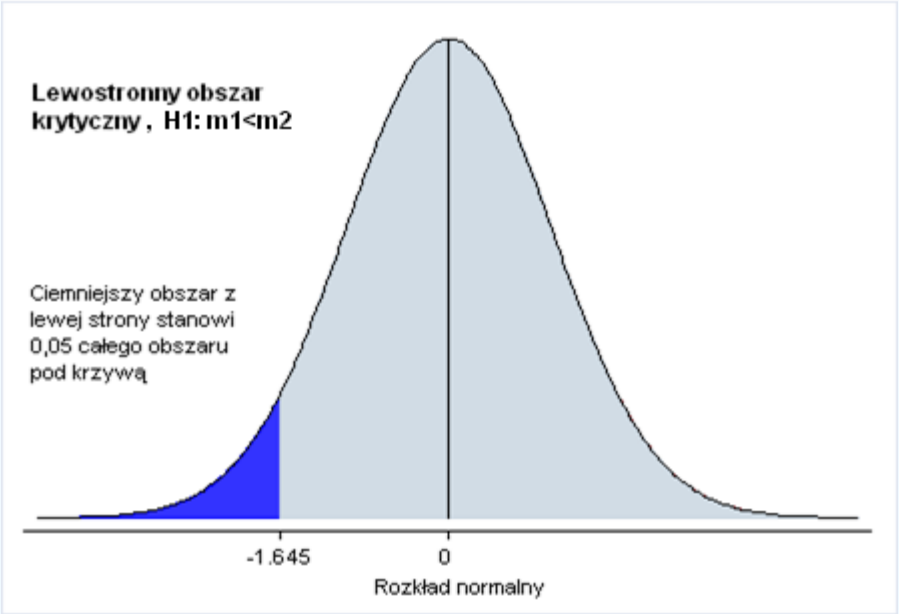
Treść H0: m1 = m2

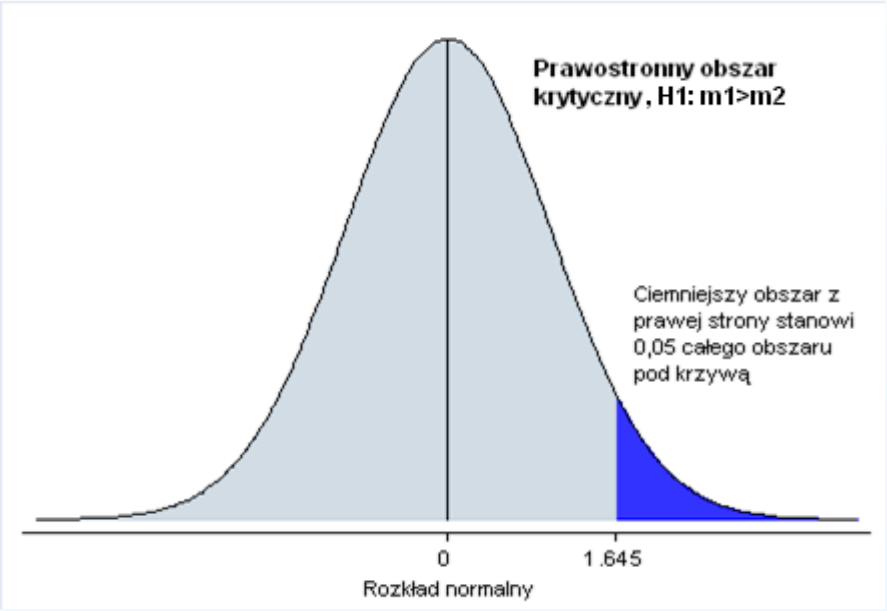
m1 –nieznana wartość parametru

m2 –przypuszczenie na temat wartości parametru (jaką wartość badacz chce sprawdzić)

Objaśnienie obszarów krytycznych dla α = 0,05 i statystyki testowej o określonym rozkładzie; tu – o rozkładzie normalnym.







Podstawowe pojęcia

- Dwie zbiorowości statystyczne
  populacje zależne i niezależne

- Test różnic między parametrami dla dwóch populacji:
  - założenie o normalności rozkładu cechy
  - postaci hipotez: zerowej i alternatywnej

Testy dla dwóch średnich populacji niezależnych

Próby niezależne są tworzone z danych pobieranych z dwóch niezależnych (nie związanych ze sobą) populacji.
Przykładem wykorzystania testu jest sprawdzenie istotności różnic w średniej płacy kobiet i mężczyzn na wybranych stanowiskach. Obie grupy są niezależne – nie ma powiązań między zarobkami dla którejkolwiek kobiety i któregokolwiek mężczyzny.

H0: $m_1 = m_2$ **weryfikowana hipoteza zerowa**

Możliwe hipotezy alternatywne

H1: $m_1 \neq m_2$ **dwustronny obszar krytyczny**

H1: $m_1 < m_2$ **lewostronny obszar krytyczny**

H1: $m_1 > m_2$ **prawostronny obszar krytyczny**

Jeżeli wariancje obu populacji są równe, to test jest oparty o statystykę t-Studenta.

Jeżeli wariancje obu populacji są różne, to test jest oparty o aproksymowaną statystykę t-Studenta.

Liczba stopni swobody statystyki testowej jest równa $n_1 + n_2 - 2$, gdzie n_1 jest liczebnością jednej próby, a n_2 jest liczebnością drugiej próby.

Testy dla dwóch wariancji (populacje niezależne)

Jeżeli przyjąć założenie, że w obu populacjach rozkład badanej cechy jest normalny, to dla weryfikacji hipotezy zerowej o równości wariancji cechy dla tych populacji:

$$\mathbf{H0:}\ \sigma_1^2 = \sigma_2^2$$

wykorzystuje się statystykę F, która ma rozkład F-Snedecora o $(n_1-1)(n_2-1)$ stopniach swobody.

Testy dla dwóch średnich populacji zależnych

Próby zależne są tworzone z danych pochodzących jednej próby pobranej z jednej populacji. Dla tej próby wartość cechy jest zazwyczaj mierzona dwukrotnie (każdy element próby ma przypisane dwie wartości cechy – badania „przed i po”). W ten sposób powstają dwa szeregi statystyczne tworzące dwie próby.

Rodzaj ingerencji zależy od przedmiotu badań.
Można w ten sposób sprawdzać efektywność statystyczną leku (w badaniach medycznych), nawozu (przy opracowywaniu środków podnoszących efektywność produkcji roślinnej), sposobu karmienia, odzwyczajania od używek.

H0: $m_1 = m_2$ **weryfikowana hipoteza zerowa**

Możliwe hipotezy alternatywne

H1: $m_1 \neq m_2$ **dwustronny obszar krytyczny**

H1: $m_1 < m_2$ **lewostronny obszar krytyczny**

H1: $m_1 > m_2$ **prawostronny obszar krytyczny**

Statystyka testowa ma rozkład t-Studenta o *n-1* stopniach swobody, gdzie *n* jest liczebnością próby.

Przekształcenie testu dwustronnego w test jednostronny

Majac p-wartość dla hipotezy dwustronnej można łatwo (o ile to ma sens) wyznaczyć p-wartość dla hipotezy jednostronnej.

Jeżeli przez p0 oznaczy się p-wartość dla hipotezy dwustronnej, to dla hipotezy jednostronnej p-wartości można wyznaczyć jako p0/2, jeżeli hipoteza **alternatywna** jednokierunkowa jest „w tym kierunku, w którym są obserwacje” i 1 - p0/2 w przeciwnym przypadku.

Jeżeli średnia_próby(X) > średnia_próby(Y), to treść H1 jest taka, że $\mu(X) > \mu(Y)$.
Jeżeli średnia_próby(X) < średnia_próby(Y), to treść H1 jest taka, że $\mu(X) < \mu(Y)$.

Zapis symboliczny

Dla hipotezy zerowej H0: $\mu_1 = \mu_2$ i alternatywnej H1: $\mu_1 \neq \mu_2$ dane jest prawdopodobieństwo testowe:

$$p_0 = \text{p-wartość}(H1: \mu_1 - \mu_2 \neq 0)$$

Dla testu jednostronnego hipotezę alternatywną formułuje się tak, aby otrzymać test jednostronny w kierunku wskazywanym przez dane. W takim przypadku wartość jest dana zaleznością:

$$\text{p-value}(H1: \mu_1 - \mu_2 \ (\textit{właściwa relacja: } < \textit{ albo } >) \ 0) = p_0/2$$

Procedura testowania - Enterprise Guide

Przygotować projekt: biblioteka, dane – zwrócić uwagę na strukturę zbioru

Sprawdzić normalność rozkładu cech

Wybrać rodzaj testu dla dwóch prób (dwupróbkowy – test dla średnich z populacji niezależnych; parzysty – test dla średnich z populacji zależnych)

Sformułować

- Hipotezę zerową
- Hipotezę alternatywną

Wykonać test

Wyniki:

- Miary statystyczne policzone dla prób
- Wyniki testów normalności
- Wyniki testu istotności różnic między średnimi, zawierające:
  - ✓ statystykę testową oraz liczbę stopni swobody
  - ✓ prawdopodobieństwo testowe

Wnioski