

Testowanie rozmytych hipotez w oparciu o rozmytą statystykę testową

Zuzanna Rachuba

Wydział matematyki i informatyki
Uniwersytet im. Adama Mickiewicza, Poznań

- 1 *Wstęp*
- 2 *Algorytm analizowania "rozmytej" tezy*
- 3 *Dlaczego warto badać rozmyte tezy?*
- 4 *Podstawowe pojęcia matematyczne związane ze zbiorami rozmytymi*
- 5 *Testowanie hipotez rozmytych*
- 6 *Testowanie hipotez rozmytych w rozkładzie normalnym*

Wyobraźmy sobie, że chcemy sprawdzić tezę bez jednoznacznych danych. Gdy informacje są nieprecyzyjne, np. „około 5” lub „raczej małe”, tradycyjne metody statystyczne zawodzą. Potrzebujemy więc podejścia uwzględniającego tę niepewność.

- Przekształcamy klasyczne hipotezy na wersje „rozmyte” – zamiast mówić „średnia wynosi dokładnie 10”, możemy powiedzieć, że jest „około 10”.
- Sprawdzamy, na ile to możliwe – wyznaczamy zakresy wartości (przedział ufności), które są zgodne z naszymi danymi i hipotezą.
- Budujemy test – używamy narzędzi matematycznych, żeby określić, czy nasze dane wspierają hipotezę, czy nie.
- Podejmujemy decyzję – wprowadzamy „poziom pewności” i na jego podstawie mówimy, czy akceptujemy hipotezę, czy ją odrzucamy.

Dlaczego warto badać rozmyte tezy?

Tradycyjnie hipotezy mają precyzyjny charakter, np. „średnia wynosi dokładnie 0.15”. W rzeczywistości używamy jednak stwierdzeń typu "średnia jest mała" lub "średnia wynosi około $1/2$ ".

Rozwiązanie? Hipotezy rozmyte

Definicja: Zbiór rozmyty A w przestrzeni X jest definiowany przez funkcję przynależności:

$$A : X \rightarrow [0, 1],$$

gdzie każdemu elementowi x z przestrzeni X przyporządkowana jest liczba z przedziału $[0, 1]$, określająca stopień przynależności x do zbioru A .

Definicja: Cięcie α zbioru rozmytego A , oznaczane jako $A[\alpha]$, to klasyczny zbiór elementów $x \in X$, które mają stopień przynależności co najmniej równy α :

$$A[\alpha] = \{x \in X \mid A(x) \geq \alpha\}, \quad \text{dla } 0 < \alpha \leq 1.$$

Przykład: Jeśli $A(x)$ oznacza stopień "wysokości" budynku, to:

- $A[0.5]$: budynki, które są "średnio wysokie" lub wyższe,
- $A[0.8]$: budynki, które są "bardzo wysokie".

Definicja: Liczba rozmyta to wartość, którą opisujemy z pewnym stopniem niepewności. Zamiast mówić, że coś ma dokładnie wartość x , mówimy, że jest „około x ” (gdzie stopień "około" jest określony)

Liczby rozmyte są opisane funkcją przynależności, która przypisuje wartość x stopień, w jakim należy do danej kategorii.

1. Rozmyta hipoteza prosta:

$$H : \theta \approx \theta_0$$

- θ jest w przybliżeniu równe θ_0 .

2. Rozmyta hipoteza dwustronna:

$$H : \theta \not\approx \theta_0$$

- θ jest nie w przybliżeniu równe θ_0 .

Czym różni się testowanie rozmyte od klasycznego?

- W klasycznym podejściu testujemy **ostre hipotezy**, np. $H_0 : \theta = \theta_0$.
- W podejściu rozmytym wprowadzamy niepewność:

$$H_0 : \theta \text{ jest bliskie } \theta_0$$

- Rozmyte wartości są reprezentowane jako przedziały:

$$H_0[\alpha] = [a_1 + (\theta_0 - a_1)\alpha, a_3 - (a_3 - \theta_0)\alpha].$$

Jak działają α -cięcia w rozmytym testowaniu?

- Średnia z próby ($\tilde{X}[\alpha]$):

$$\tilde{X}[\alpha] = \left[\bar{x} - z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}, \bar{x} + z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}} \right]$$

- Hipoteza rozmyta ($H_0[\alpha]$):

$$H_0[\alpha] = [a_1 + (\theta_0 - a_1)\alpha, a_3 - (a_3 - \theta_0)\alpha].$$

- Uwzględniamy niepewność w obu zakresach!

Statystyka testowa w podejściu rozmytym:

- Wyznaczamy przedział dla statystyki $Z[\alpha]$:

$$\tilde{Z}[\alpha] = \frac{\tilde{X}[\alpha] - H_0[\alpha]}{\sigma/\sqrt{n}}$$

- Przedział $Z[\alpha]$ zmienia się wraz z poziomem α .

Jak przeprowadzić test rozmyty?

- 1 Wyznacz przedziały dla $\tilde{X}[\alpha]$ i $H_0[\alpha]$.
- 2 Oblicz przedziały dla statystyki $Z[\alpha]$.
- 3 Wyznacz obszar akceptacji (AR) i obszar odrzucenia (AT):

Jeśli $\frac{AR}{AT} \leq \phi$, odrzucamy H_0 .

- 4 Zinterpretuj wynik: rozmytość pozwala na bardziej elastyczną decyzję!

Rozważmy testowanie hipotez rozmytych dla średniej.

Założenia:

- Próbkę pochodzi z rozkładu normalnego $N(\theta, \sigma^2)$, gdzie σ^2 jest znane.
- Wielkość próbki: $n = 50$
- Średnia próbki: $\bar{x} = 2.1$
- Odchylenie standardowe: $\sigma = 0.5$
- Testowane hipotezy:

$$H_0 : \theta = (1.8, 2.0, 2.2), \quad H_1 : \theta \in (2.0, 2.5)$$

- Poziom istotności: $\alpha = 0.05$

Statystyka testowa:

Dla hipotezy $H_0 : \theta = \theta_0$, statystyka testowa ma postać:

$$z_0 = \frac{\bar{x} - \theta_0}{\frac{\sigma}{\sqrt{n}}} = \frac{2.1 - 2}{\frac{0.5}{\sqrt{50}}} = 1.414$$

Wartość krytyczna dla $\alpha = 0.05$: $z_{1-\alpha/2} = 1.645$.

Przedział rozmyty dla H_0 :

$$H_0[0.5] = [1.9, 2.1]$$

Granice przedziału dla Z :

$$Z[0.5] = \left[2.1 - \frac{0.5}{\sqrt{50}} - 0.2(1 - 0.5) \cdot \frac{0.5}{\sqrt{50}}, \right. \\ \left. 2.1 + \frac{0.5}{\sqrt{50}} + 0.2(1 - 0.5) \cdot \frac{0.5}{\sqrt{50}} \right]$$

Po podstawieniu:

$$Z[0.5] = [2.021, 2.264]$$

Testowanie hipotezy:

- Obliczamy pole akceptacji (AR) i pole odrzucenia (AT).
- Stosujemy metodę czwórkową, obliczając stosunek $\frac{AR}{AT}$.

Wynik testu:

$$\frac{AR}{AT} \leq \phi \quad \text{dla} \quad \phi \in (0, 0.48]$$

Zatem odrzucamy hipotezę H_0 .

Interpretacja: Dzięki testowaniu hipotez rozmytych uwzględniamy niepewność w definiowaniu granic hipotezy H_0 oraz w rozróżnianiu, czy θ jest "w przybliżeniu równe" θ_0 , czy "istotnie większe" od θ_0 .



S. Mahmoud Taheri, Mohsen Arefi, *Testing fuzzy hypotheses based on fuzzy test statistic*, Published online: 5 July 2024.

