Параметричні тести значущості

Формули

Параметричні тести значущості - поточна блок-схема

1. Сформулюємо основну гіпотезу H_0 щодо деякого параметра в генеральній сукупності.
   Найпоширенішою формою гіпотези H_0 є $parametr = liczba$
   Сформулюємо альтернативну гіпотезу H_1. Він може приймати форму
   $parametr \neq liczba$, $parametr > liczba$, $parametr < liczba$.
2. Розраховуємо відповідну статистику.
3. Створюємо і рисуємо критичну зону (двох або односторонніх залежностей від H_1)
4. Перевіряємо, чи статистика в критичній області. Якщо так - ми відхиляємо гіпотезу H_0 на користь альтернативної гіпотези H_1. Якщо ні, ми робимо висновок, що немає підстав для відхилення гіпотези H_0.

I. Тести значущості для одного середнього значення в загальній сукупності з нормальним розподілом

I.a Ми знаємо стандартне відхилення в генеральній сукупності σ .

1. Формулюємо гіпотезу:

$H_0 : m = m_0$

$H_1 : m \neq m_0$ або $H_1 : m > m_0$ або $H_1 : m < m_0$

де m то средне значення генеральної сукупності, а m_o це параметр, який ми вибираємо

2. Розраховуємо статистику: $Z = \frac{\overline{X} - m_0}{\sigma} \cdot \sqrt{n}$

3. Створюємо та малюємо критичну область для нормального розподілу (дво або односторонний залежно від H_1)

4. Перевіряємо, чи статистика в критичній області. Якщо так - ми відхиляємо гіпотезу H_0 на користь альтернативної гіпотези H_1. Якщо ні, ми робимо висновок, що немає підстав для відхилення гіпотези H_0.

I.b Ми не знаємо стандартного відхилення в генеральній сукупності σ , розмір вибірки n великий.

1. Формулюємо гіпотезу:

$H_0 : m = m_0$

$H_1 : m \neq m_0$ або $H_1 : m > m_0$ або $H_1 : m < m_0$

де m то середне значення генеральної сукупності, а m_o це обране нами значення

2. Розраховуємо статистику: $Z = \frac{\overline{X} - m_0}{S} \cdot \sqrt{n}$

3. Створюємо та малюємо критичну область для нормального розподілу (дво або односторонний залежно від H_1)

4. Перевіряємо, чи статистика в критичній області. Якщо так - ми відхиляємо гіпотезу H_0 на користь альтернативної гіпотези H_1. Якщо ні, ми робимо висновок, що немає підстав для відхилення гіпотези H_0.

I.c Ми не знаємо стандартного відхилення в генеральній сукупності σ , розмір вибірки n невеликий.

1. Формулюємо гіпотезу:

$H_0 : m = m_0$

$H_1 : m \neq m_0$ або $H_1 : m > m_0$ або $H_1 : m < m_0$

де m то середне значення генеральної сукупності, а m_o це обране нами значення

2. Розраховуємо статистику: $t = \frac{\overline{X} - m_0}{S}\sqrt{n-1}$

3. Ми створюємо та рисуємо критичну область для t-розподілу Стьюдента (в залежності від двох або одностороннього H_1) для $n-1$ступенів свободи. Ми пам'ятаємо про подвоєння рівня значущості в односторонніх областях.

4. Перевіряємо, чи статистика в критичній області. Якщо так - ми відхиляємо гіпотезу H_0 на користь альтернативної гіпотези H_1. Якщо ні, ми робимо висновок, що немає підстав для відхилення гіпотези H_0 .

II. Порівняння середніх показників двох нормально розподілених популяцій.

II.a Ми знаємо стандартні відхилення в загальних сукупностях σ_1 і σ_2 .

1. Формулюємо гіпотезу:

$H_0 : m_1 = m_2$

$H_1 : m_1 = m_2$ або $H_1 : m_1 < m_2$ або $H_1 : m_1 > m_2$

де m_1 , m_2 це середне в двох полуляціях

2. Розраховуємо статистику: $Z = \frac{\overline{X_1} - \overline{X_2}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$

3. Створюємо та малюємо критичну область для нормального розподілу (двох або односторонний залежно від H_1)

4. Перевіряємо, чи статистика в критичній області. Якщо так - ми відхиляємо гіпотезу H_0 на користь альтернативної гіпотези H_1. Якщо ні, ми робимо висновок, що немає підстав для відхилення гіпотези H_0.

II.b Нам не відомі стандартні відхилення в загальних сукупностях σ_1 та σ_2, але розміри вибірки n_1 та n_2 є великими.

1. Формулюємо гіпотезу:

$H_0 : m_1 = m_2$

$H_1 : m_1 = m_2$ або $H_1 : m_1 < m_2$ або $H_1 : m_1 > m_2$

де m_1, m_2 це середне в двох популяціях

2. Розраховуємо статистику: $Z = \dfrac{\overline{X_1} - \overline{X_2}}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}}$

3. Створюємо та малюємо критичну область для нормального розподілу (двох або односторонний залежно від H_1)

4. Перевіряємо, чи статистика в критичній області. Якщо так - ми відхиляємо гіпотезу H_0 на користь альтернативної гіпотези H_1. Якщо ні, ми робимо висновок, що немає підстав для відхилення гіпотези H_0.

II.c Нам не відомі стандартні відхилення в загальних сукупностях σ_1 та σ_2, а розміри вибірки n_1 і n_2 малі.

1. Формулюємо гіпотезу:

$H_0 : m_1 = m_2$

$H_1 : m_1 = m_2$ або $H_1 : m_1 < m_2$ або $H_1 : m_1 > m_2$

де m_1, m_2 це середне в двох популяціях

2. Розраховуємо статистику: $t = \dfrac{\overline{X_1} - \overline{X_2}}{\sqrt{\dfrac{n_1 S_1^2 + n_2 S_2^2}{n_1 + n_2 - 2}\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}}$

3. Ми створюємо та рисуємо критичну область для t-розподілу Стьюдента (в залежності від H_1 двох або одностороннього) для n_1+n_2-2 ступенів свободи. Ми пам'ятаємо про подвоєння рівня значущості в односторонніх областях.

4. Перевіряємо, чи статистика в критичній області. Якщо так - ми відхиляємо гіпотезу H_0 на користь альтернативної гіпотези H_1. Якщо ні, ми робимо висновок, що немає підстав для відхилення гіпотези H_0.

II. Тести значущості для однієї дисперсії в загальній сукупності з нормальним розподілом

III.a Розмір вибірки n великий

1. Формулюємо гіпотезу:

$H_0 : \sigma^2 = \sigma_0^2$ або $H_0 : \sigma = \sigma_0$

$H_1 : \sigma^2 > \sigma_0^2$ або $H_1 : \sigma > \sigma_0$

де σ^2 це дисперсія в генеральній сукупності, а σ_0^2 це параметр, який ми вибираємо.

2. Розраховуємо статистику: $Z = \sqrt{2\chi^2} - \sqrt{2n-3}$, де $\chi^2 = \dfrac{nS^2}{\sigma_0^2}$.

3. Створюємо і малюємо критичну область для нормального (правого) розподілу.

4. Перевіряємо, чи статистика в критичній області. Якщо так - ми відхиляємо гіпотезу H_0 на користь альтернативної гіпотези H_1. Якщо ні, ми робимо висновок, що немає підстав для відхилення гіпотези H_0.

III.b Розмір вибірки n невеликий

1. Формулюємо гіпотезу:

$H_0 : \sigma^2 = \sigma_0^2$ або $H_0 : \sigma = \sigma_0$

$H_1 : \sigma^2 > \sigma_0^2$ або $H_1 : \sigma > \sigma_0$

де σ^2 це дисперсія в генеральній сукупності, а σ_0^2 це параметр, який ми вибираємо.

2. Розраховуємо статистику: $\chi^2 = \dfrac{nS^2}{\sigma_0^2}$.

3. Створюємо і малюємо критичну область для розподілу хі-квадрат (правосторонній) зі $n-1$ ступенями свободи.

4. Перевіряємо, чи статистика в критичній області. Якщо так - ми відхиляємо гіпотезу H_0 на користь альтернативної гіпотези H_1. Якщо ні, ми робимо висновок, що немає підстав для відхилення гіпотези H_0.

III. Порівняння дисперсії двох нормально розподілених популяцій.

1. Формулюємо гіпотезу:

$H_0 : \sigma_1^2 = \sigma_2^2$ або $H_0 : \sigma_1 = \sigma_2$

$H_1 : \sigma_1^2 > \sigma_2^2$ або $H_1 : \sigma_1 > \sigma_2$

де σ_1^2 і σ_2^2 – це дисперсії в популяціях, пронумерованих таким чином, що $\hat{S}_1 > \hat{S}_2$.

2. Розраховуємо статистику: $F = \frac{\hat{S}_1^2}{\hat{S}_2^2}$, для $\hat{S}_1 > \hat{S}_2$.

3. Ми створюємо та малюємо критичну область для розподілу Шнедерсона (правостронний) з $n_1 - 1$ та $n_2 - 1$ ступенями свободи.

4. Перевіряємо, чи статистика в критичній області. Якщо так - ми відхиляємо гіпотезу H_0 на користь альтернативної гіпотези H_1. Якщо ні, ми робимо висновок, що немає підстав для відхилення гіпотези H_0.

IV. Тести значущості для однієї ймовірності (відсотка, частки) у загальній сукупності

1. Формулюємо гіпотезу:

$H_0: p = p_0$

$H_1: p \neq p_0$ або $H_1: p > p_0$ або $H_1: p < p_0$

де p - це відсоток загальній популяції, а p_o це вибраний нами параметр.

2. Розраховуємо статистику: $Z = \dfrac{\frac{m}{n} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}$

де m to кількість одиниць у вибірці з перевіреною ознакою, а $\frac{m}{n}$ це відсоток одиниць у вибірці, які мають цю ознаку.

3. Створюємо та малюємо критичну область для нормального розподілу (двох або односторонній залежно від H_1)

4. Перевіряємо, чи статистика в критичній області. Якщо так - ми відхиляємо гіпотезу H_0 на користь альтернативної гіпотези H_1. Якщо ні, ми робимо висновок, що немає підстав для відхилення гіпотези H_0.

V. Порівняння двох ймовірностей (відсотків, часток) у двох генеральних сукупностях

1. Формулюємо гіпотезу:

$H_0 : p_1 = p_2$

$H_1 : p_1 \neq p_2$ або $H_1 : p_1 > p_2$ або $H_1 : p_1 < p_2$

де p_1, p_2 це відсотки в генеральній сукупності.

2. Розраховуємо статистику: $Z = \dfrac{\dfrac{m_1}{n_1} - \dfrac{m_2}{n_2}}{\sqrt{\dfrac{\dfrac{m_1+m_2}{n_1+n_2}\left(1 - \dfrac{m_1+m_2}{n_1+n_2}\right)}{\dfrac{n_1 \cdot n_2}{n_1+n_2}}}}$

де m_1, m_2 – кількість одиниць у вибірці з ознакою, що перевіряється, а $\dfrac{m_1}{n_1}, \dfrac{m_2}{n_2}$ – вибіркові відсотки відповідно.

3. Створюємо та малюємо критичну область для нормального розподілу (обох або односторонній залежно від H_1)

4. Перевіряємо, чи статистика в критичній області. Якщо так - ми відхиляємо гіпотезу H_0 на користь альтернативної гіпотези H_1. Якщо ні, ми робимо висновок, що немає підстав для відхилення гіпотези H_0.