Statystyka i rachunek prawdopodobieństwa

Wykład 12: Hipotezy statystyczne—testy parametryczne

Mariusz Tarnopolski

Instytut Astronomii UMK

Parametryczne testy istotności na wartość przeciętną

Model 1

$H_0 : m = m_0$—hipoteza, że wartość przeciętna cechy X w badanej populacji jest równa wartości m_0. Wykorzystujemy statystykę $\bar{X}$ z próbki o liczności n.

$H_0 : m = m_0$—hipoteza, że wartość przeciętna cechy X w badanej populacji jest równa wartości m_0. Wykorzystujemy statystykę $\bar{X}$ z próbki o liczności n.

Niech badana cecha X populacji generalnej ma rozkład $\mathcal{N}(m, \sigma)$ przy znanym σ.

Badamy H_0 wobec:

1. $H_1 : m < m_0$
2. $H_1 : m > m_0$
3. $H_1 : m \neq m_0$

na poziomie istotności α.

Statystyka testowa (TS): $U = \dfrac{\bar{X} - m_0}{\sigma/\sqrt{n}}$ przy prawdziwości H_0 ma rozkład $\mathcal{N}(0, 1)$.

Parametryczne testy istotności na wartość przeciętną

Model 1

$H_0 : m = m_0$—hipoteza, że wartość przeciętna cechy X w badanej populacji jest równa wartości m_0. Wykorzystujemy statystykę $\bar{X}$ z próbki o liczności n.

Niech badana cecha X populacji generalnej ma rozkład $\mathcal{N}(m, \sigma)$ przy znanym σ.

Badamy H_0 wobec:

1. $H_1 : m < m_0$
2. $H_1 : m > m_0$
3. $H_1 : m \neq m_0$

na poziomie istotności α.

Statystyka testowa (TS): $U = \dfrac{\bar{X} - m_0}{\sigma/\sqrt{n}}$ przy prawdziwości H_0 ma rozkład $\mathcal{N}(0, 1)$.

Jeśli obliczona z próbki wartość $U \in K_\alpha$, to H_0 **odrzucamy** (na poziomie α lub też: z błędem α).

Jeśli $U \notin K_\alpha$, to **nie ma podstaw do odrzucenia** H_0.

Dla

1. $H_1 : m < m_0$
2. $H_1 : m > m_0$
3. $H_1 : m \neq m_0$

obszar krytyczny K_α (tj. taki, który minimalizuje błąd II rodzaju) jest:

1. lewostronny: $(-\infty, u_\alpha]$, gdzie u_α to kwantyl rzędu α rozkładu $\mathcal{N}(0,1)$—test jednostronny. Zachodzi $u_\alpha = -u_{1-\alpha}$
2. prawostronny: $[u_{1-\alpha}, +\infty)$, gdzie $u_{1-\alpha}$ to kwantyl rzędu $1-\alpha$ rozkładu $\mathcal{N}(0,1)$—test jednostronny
3. dwustronny: $(-\infty, -u_{1-\alpha/2}] \cup [u_{1-\alpha/2}, +\infty)$—test dwustronny

Badana cecha X populacji ma rozkład $\mathcal{N}(m, \sigma)$ o **nieznanych** m oraz σ.

Statystyka testowa: $t = \dfrac{\bar{X} - m_0}{S/\sqrt{n}}$, która przy prawdziwości H_0 ma rozkład t Studenta o $(n-1)$ stopniach swobody; S to nieobciążony estymator odchylenia standardowego (zob. wykład 9).

Parametryczne testy istotności na wartość przeciętną
Model 2

Badana cecha X populacji ma rozkład $\mathcal{N}(m, \sigma)$ o **nieznanych** m oraz σ.
Statystyka testowa: $t = \dfrac{\bar{X} - m_0}{S/\sqrt{n}}$, która przy prawdziwości H_0 ma rozkład t Studenta o $(n-1)$ stopniach swobody; S to nieobciążony estymator odchylenia standardowego (zob. wykład 9).
K_α wyznacza się analogicznie jak w modelu 1, tyle, że przy użyciu kwantyli rozkladu t Studenta:

1. lewostronny: $(-\infty, -t_{1-\alpha}]$, gdzie $t_{1-\alpha}$ to kwantyl rzędu $1-\alpha$ rozkładu t Studenta o $n-1$ stopniach swobody—test jednostronny.
2. prawostronny: $[t_{1-\alpha}, +\infty)$—test jednostronny
3. dwustronny: $(-\infty, -t_{1-\alpha/2}] \cup [t_{1-\alpha/2}, +\infty)$—test dwustronny

Parametryczne testy istotności na wartość przeciętną
Model 2

Badana cecha X populacji ma rozkład $\mathcal{N}(m,\sigma)$ o **nieznanych** m oraz σ.

Statystyka testowa: $t = \dfrac{\bar{X}-m_0}{S/\sqrt{n}}$, która przy prawdziwości H_0 ma rozkład t Studenta o $(n-1)$ stopniach swobody; S to nieobciążony estymator odchylenia standardowego (zob. wykład 9).

K_α wyznacza się analogicznie jak w modelu 1, tyle, że przy użyciu kwantyli rozkladu t Studenta:

1. lewostronny: $(-\infty,-t_{1-\alpha}]$, gdzie $t_{1-\alpha}$ to kwantyl rzędu $1-\alpha$ rozkładu t Studenta o $n-1$ stopniach swobody—test jednostronny.
2. prawostronny: $[t_{1-\alpha},+\infty)$—test jednostronny
3. dwustronny: $(-\infty,-t_{1-\alpha/2}]\cup[t_{1-\alpha/2},+\infty)$—test dwustronny

Jeśli obliczona z próbki wartość $t\in K_\alpha$, to H_0 **odrzucamy** na poziomie istotności α.

Jeśli $t\notin K_\alpha$, to **nie ma podstaw do odrzucenia** H_0.

Badana cecha X populacji ma **dowolny** rozkład o nieznanych wartości przeciętnej i wariancji, ale próbka jest dostatecznie duża ($n \gtrsim 30$).

Badana cecha X populacji ma **dowolny** rozkład o nieznanych wartości przeciętnej i wariancji, ale próbka jest dostatecznie duża ($n \gtrsim 30$).
Test przeprowadza się podobnie jak w modelu 1—korzystamy z CLT, zatem $\bar{X}$ ma w przybliżeniu rozkład normalny. Nieznaną wartość σ estymujemy z próbki jako S. Statystyka testowa to $U = \dfrac{\bar{X} - m_0}{S/\sqrt{n}}$

Niech badana cecha X populacji generalnej ma rozkład $\mathcal{N}(m, \sigma)$ o nieznanych m oraz σ. Weryfikujemy hipotezę $H_0 : \sigma^2 = \sigma_0^2$ wobec hipotezy alternatywnej:

1. $H_1 : \sigma^2 < \sigma_0^2$
2. $H_1 : \sigma^2 > \sigma_0^2$
3. $H_1 : \sigma^2 \neq \sigma_0^2$

na poziomie istotności α.

Niech badana cecha X populacji generalnej ma rozkład $\mathcal{N}(m, \sigma)$ o nieznanych m oraz σ. Weryfikujemy hipotezę $H_0 : \sigma^2 = \sigma_0^2$ wobec hipotezy alternatywnej:

1. $H_1 : \sigma^2 < \sigma_0^2$
2. $H_1 : \sigma^2 > \sigma_0^2$
3. $H_1 : \sigma^2 \neq \sigma_0^2$

na poziomie istotności α.

Statystyka testowa: $\chi^2 = \dfrac{(n-1)S^2}{\sigma_0^2}$ przy prawdziwości H_0 ma rozkład χ^2 o $(n-1)$ stopniach swobody.

Parametryczne testy istotności na wariancję

Model 1

Niech badana cecha X populacji generalnej ma rozkład $\mathcal{N}(m, \sigma)$ o nieznanych m oraz σ. Weryfikujemy hipotezę $H_0 : \sigma^2 = \sigma_0^2$ wobec hipotezy alternatywnej:

1. $H_1 : \sigma^2 < \sigma_0^2$
2. $H_1 : \sigma^2 > \sigma_0^2$
3. $H_1 : \sigma^2 \neq \sigma_0^2$

na poziomie istotności α.

Statystyka testowa: $\chi^2 = \dfrac{(n-1)S^2}{\sigma_0^2}$ przy prawdziwości H_0 ma rozkład χ^2 o $(n-1)$ stopniach swobody.

K_α analogicznie jak poprzednio:

1. lewostronny: $(0, \chi^2_{\alpha,n-1}]$, gdzie $\chi^2_{\alpha,n-1}$ to kwantyl rzędu α rozkładu χ^2 o $(n-1)$ stopniach swobody—test jednostronny
2. prawostronny: $[\chi^2_{1-\alpha,n-1}, \infty)$—test jednostronny
3. dwustronny: $(0, \chi^2_{\alpha/2,n-1}] \cup [\chi^2_{1-\alpha/2,n-1}, \infty)$—test dwustronny

Niech badana cecha X populacji ma rozkład $\mathcal{N}(m, \sigma)$ o nieznanych m oraz σ, ale próba jest duża ($n \gtrsim 50$).
Statystyka testowa jak w modelu 1; korzystamy z faktu, że ma ona w przybliżeniu rozkład[1] $\mathcal{N}\left(\sqrt{2n-3}, 1\right)$.

[1] Uwzględniając poprawkę Edgewortha; por. wykład 10, slajd 21.

Niech badana cecha X populacji ma rozkład $\mathcal{N}(m, \sigma)$ o nieznanych m oraz σ, ale próba jest duża ($n \gtrsim 50$).
Statystyka testowa jak w modelu 1; korzystamy z faktu, że ma ona w przybliżeniu rozkład[1] $\mathcal{N}\left(\sqrt{2n-3}, 1\right)$. Wtedy statystyka ustandaryzowana:

$$U = \sqrt{\frac{2(n-1)S^2}{\sigma_0^2}} - \sqrt{2n-3}$$

ma w przybliżeniu rozkład $\mathcal{N}(0, 1)$.

[1] Uwzględniając poprawkę Edgewortha; por. wykład 10, slajd 21.

Parametryczne testy istotności na wariancję

Model 2

Niech badana cecha X populacji ma rozkład $\mathcal{N}(m, \sigma)$ o nieznanych m oraz σ, ale próba jest duża ($n \gtrsim 50$).
Statystyka testowa jak w modelu 1; korzystamy z faktu, że ma ona w przybliżeniu rozkład[1] $\mathcal{N}\left(\sqrt{2n-3}, 1\right)$. Wtedy statystyka ustandaryzowana:

$$U = \sqrt{\frac{2(n-1)S^2}{\sigma_0^2}} - \sqrt{2n-3}$$

ma w przybliżeniu rozkład $\mathcal{N}(0, 1)$.
Obszary krytyczne to (por. slajd 3):

1. lewostronny: $(-\infty, u_\alpha]$—test jednostronny. Przypomnijmy, że $u_\alpha = -u_{1-\alpha}$
2. prawostronny: $[u_{1-\alpha}, +\infty)$—test jednostronny
3. dwustronny: $(-\infty, -u_{1-\alpha/2}] \cup [u_{1-\alpha/2}, +\infty)$—test dwustronny

[1] Uwzględniając poprawkę Edgewortha; por. wykład 10, slajd 21.

Niech badana cecha X ma dowolny rozkład (o skończonej kurtozie), ale próba jest duża ($n \gtrsim 100$). Weryfikacja opiera się na statystyce testowej

$$U = \frac{S^2 - \sigma_0^2}{\sigma_0^2}\sqrt{\frac{n}{2}}$$

która—przy prawdziwości H_0—ma w przybliżeniu rozkład $\mathcal{N}(0, 1)$. Obszary krytyczne K_α są identyczne jak w modelu 2 (slajd 7).

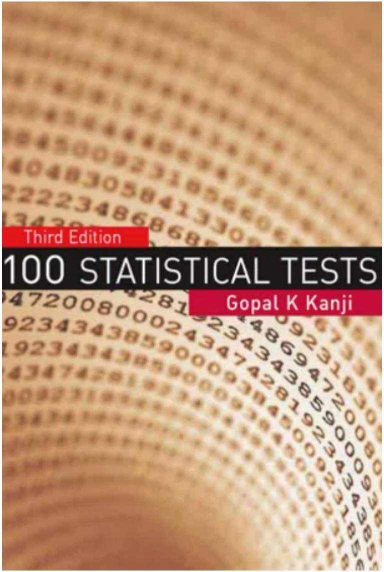
Badamy dwie populacje i chcemy porównać ich wartości przeciętne.

1. Rozkłady $\mathcal{N}(m_i, \sigma_i)$, znane σ_i, nieznane m_i. $H_0 : m_1 = m_2$. Niech $n_i > 5$. TS to $U = \left(\bar{X}_1 - \bar{X}_2\right)\left(\sigma_1^2/n_1 + \sigma_2^2/n_2\right)^{-1/2} \sim \mathcal{N}(0,1)$.
2. J.w., nieznane m_i i σ_i, zaś $n_i > 100$. TS to $U = \left(\bar{X}_1 - \bar{X}_2\right)\left(S_1^2/n_1 + S_2^2/n_2\right)^{-1/2} \sim \mathcal{N}(0,1)$
3. J.w., ale $n_i > 5$. Statystyka Cochrana–Coxa (o skomplikowanym rozkładzie; przybliżenie Welcha rozkładem t Studenta o odpowiedniej liczbie stopni swobody): $C = \left(\bar{X}_1 - \bar{X}_2\right)\left[S_1^2/n_1 + S_2^2/n_2\right]^{-1/2}$
4. J.w., ale $\sigma_1 = \sigma_2$. TS to $U = \left(\bar{X}_1 - \bar{X}_2\right)\left\{\left[\left(n_1 S_1^2 + n_2 S_2^2\right)/\left(n_1 + n_2 - 2\right)\right]\left(n_1 + n_2\right)/n_1 n_2\right\}^{-1/2}$, która ma rozkład t Studenta o $(n_1 + n_2 - 2)$ stopniach swobody.

Badamy dwie populacje i chcemy porównać ich wariancje.

1. Rozkłady $\mathcal{N}(m_i, \sigma_i)$, $H_0 : \sigma_1^2 = \sigma_2^2$ oraz $n_i \gtrsim 30$. Statystyka Snedecora:
$$F = \frac{S_1^2}{S_2^2} = \frac{\frac{1}{n_1-1}\sum_{i=1}^{n_1}\left(X_{1i} - \bar{X}_1\right)^2}{\frac{1}{n_2-1}\sum_{i=1}^{n_2}\left(X_{2i} - \bar{X}_2\right)^2}$$
ma rozkład F Snedecora o $(n_1 - 1, n_2 - 1)$ stopniach swobody (przyjmuje się bez straty ogólności, że $S_1^2 \geqslant S_2^2$).
2. Test Levene'a—odporniejszy na odchylenia od normalności; używa odchyleń od średniej. Definiujemy dla dwóch grup $i = 1, 2$ zmienne losowe $Z_{ij} = |X_{ij} - \bar{X}_i|$ i testujemy równość wartości przeciętnych zmiennych w grupach (zob. poprzednie testy). Statystyka testowa ma w przybliżeniu rozkład F.
3. Test Browna-Forsythe'a—jeszcze bardziej odporny na odchylenia od normalności (można stosować przy silnej skośności); zamiast średniej (jak w teście Levene'a) używa odchyleń od mediany.

Mnogość testów statystycznych



This isn't even my final form...