Wykład VI: Wnioskowanie statystyczne - testowanie hipotez

Alicja Janic

Politechnika Wrocławska *alicja.janic@pwr.edu.pl*

7 styczeń 2021

Testy dla wartości średniej - wstęp

Hipotezy

Dokładne toczenie tłoka pompy paliwa silnika samochodowego ma dawać średnicę pewnej części tłoka równą $7,5$ mm. Celem eksperymentu jest sprawdzenie, czy zużycie noża tokarki nie spowodowało zwiększenia wartości średniej θ interesujących nas średnic. Pożądaną wartością średnia tych średnic jest oczywiście wartość $\theta = 7,5$ mm Mamy dwie hipotezy, z których pierwsza podlega weryfikacji i może zostać odrzucona na korzyść drugiej hipotezy

$$\begin{aligned} \mathrm{H}_0 &: \quad \theta = \theta_0 \\ \mathrm{H}_1 &: \quad \theta > \theta_0 \end{aligned}$$

Problem taki nazywamy problemem testowania hipotez

Testy dla wartości średniej - wstęp

Statystyka testowa

Celem zweryfikowania hipotezy zerowej dokonano 50 pomiarów odpowiedniej średnicy tłoków. Pomiary można uznać za niezależne i pochodzące z tego samego rozkładu normalnego ze znanym odchyleniem standardowym: $\sigma = 0.05$. Zatem dysponujemy realizacją prostej próby losowej $X_1, X_2, ..., X_{50}$ z rozkładu $\mathcal{N}(\theta, 0.05)$. W rozważanym przypadku statystyka $\overline{X}$ ma pod warunkiem zachodzenia hipotezy H_0 rozkład normalny o wartości średniej $\theta_0 = 7.5$ mm, przy czym $\sigma = 0.05$. Możemy statystyką $\overline{X}$, zastąpić jej standaryzowaną wersją

$$Z = \frac{\overline{X} - \theta_0}{\sigma / \sqrt{n}}$$

Testy dla wartości średniej - wstęp

Statystyka testowa

Ponieważ

$$Z = \frac{\overline{X} - \theta_0}{\sigma/\sqrt{n}} = \frac{\overline{X} - \theta}{\sigma/\sqrt{n}} + \frac{\theta - \theta_0}{\sigma/\sqrt{n}}$$

pod warunkiem H_1 statystyka testowa Z ma rozkład normalny przesunięty względem rozkładu standardowego o $\frac{\theta-\theta_0}{\sigma/\sqrt{n}}$ W przypadku gdy hipoteza zerowa jest fałszywa statystyka Z powinna mieć tendencję do przyjmowania „dużych" wartości

Testy dla wartości średniej - wstęp

Zbiór krytyczny

Ustaliwszy statystykę testową, dzielimy zbiór wszystkich możliwych wartości tej statystyki na dwa dopełniające się podzbiory:

1. zbiór wartości statystyki testowej, prowadzących do odrzucenia hipotezy H_0 na korzyść hipotezy H_1 (jest to zbiór „nietypowych" wartości statystyki testowej pod warunkiem prawdziwości H_0), zbiór ten nazywamy **zbiorem krytycznym** i oznaczamy literą C

W badanym przykładzie zbior krytyczny wyznaczamy z warunku

$$P_{\mathrm{H}_0}(Z \geqslant z_{1-\alpha}) = \alpha$$

i wyraża się wzorem $C = \{z : z \geqslant z_{1-\alpha}\}$

Testy dla wartości średniej - wstęp

Zbiór przyjęć hipotezy H_0

1. zbiór wartości statystyki testowej, prowadzących do nieodrzucenia hipotezy H_0 (będziemy mówić krótko **zbiór przyjęć** hipotezy H_0), stanowiący dopełnienie zbioru krytycznego (zgodnie z naszą konwencją zbiór przyjęć możemy oznaczyć C')

Wartości brzegowe zbioru C graniczące ze zbiorem C' nazywamy wartościami krytycznymi testu

Testy dla wartości średniej - wstęp

Błąd pierwszego rodzaju i poziom istotności

Odrzucenie hipotezy zerowej, gdy ta jest prawdziwa, nazywamy **błędem pierwszego rodzaju**. Jeżeli hipoteza zerowa jest hipotezą prostą, prawdopodobieństwo α popełnienia błędu pierwszego rodzaju nazywamy **poziomem istotności testu**. Jeżeli w wyniku przeprowadzenia testu otrzymano wartość statystyki testowej należącą do zbioru krytycznego, to mówimy, że dane okazały się statystycznie istotne na poziomie α

Testy dla wartości średniej

Przykład 1

Po dokonaniu pomiarów średnic 50 tłoków, $x_1, x_2..., x_{50}$, okazało się, że ich średnia $\overline{x} = 7,515$. Stąd

$$z = \frac{\overline{x} - \theta_0}{\sigma / \sqrt{n}} = 20\sqrt{50}(7,515 - 7,5) = 2,121$$

Zatem dla $\alpha = 0.01$ oraz $z_{0,99} = 2,326$ i zbiór krytyczny $C = \{z \geqslant 2,326\}$, czyli zaobserwowana wartość statystyki testowej Z należy do zbioru przyjęć. Zwiększenie poziomu istotności np. do wartości $\alpha = 0,05$, powoduje odrzucenie hipotezy zerowej i przyjęcie hipotezy alternatywnej, ponieważ $Z_{0,95} = 1,645$ (innymi słowy, dane są statystycznie istotne na poziomie 0,05)

Testy dla wartości średniej

Zbiory krytyczne dla pozostałych typów hipotez alternatywnych

$H_1: \ \theta > \theta_0$	$P_{H_0}(Z \geqslant z_{1-\alpha}) = \alpha$	$C = \{z : z \geqslant z_{1-\alpha}\}$
$H_1: \ \theta < \theta_0$	$P_{H_0}(Z \leqslant -z_{1-\alpha}) = \alpha$	$C = \{z : z \leqslant -z_{1-\alpha}\}$
$H_1: \ \theta \neq \theta_0$	$P_{H_0}(\lvert Z\rvert \geqslant z_{1-\alpha/2}) = \alpha$	$C = \{z : z \leqslant -z_{1-\alpha/2}$ lub $z \geqslant z_{1-\alpha/2}\}$

Testy dla wartości średniej - wstęp

Definicja p-wartości

Najmniejszy poziom istotności, przy którym zaobserwowana wartość statystyki testowej prowadzi do odrzucenia hipotezy zerowej, nazywamy ***p*-wartością** przeprowadzonego testu

Zauważmy, że jeżeli testujemy hipotezę $\mathrm{H}_0 : \theta = \theta_0$ przy hipotezie alternatywnej $\mathrm{H}_1 : \theta > \theta_0$ i zaobserwowaliśmy wartość z statystyki testowej, to p-wartość jest równa

$$p - \text{wartość} = P_{\mathrm{H}_0}(Z \geqslant z) = 1 - \Phi(z)$$

$$\begin{array}{l|l} \mathrm{H}_1 : \theta < \theta_0 & p - \text{wartość} = P_{\mathrm{H}_0}(Z \leqslant z) = \Phi(z) \\ \mathrm{H}_1 : \theta \neq \theta_0 & p - \text{wartość} = 2 \cdot P_{\mathrm{H}_0}(Z \geqslant |z|) = 2(1 - \Phi(|z|) \end{array}$$

Testy dla wartości średniej - wstęp

p-wartość

Im mniejsza jest *p*-wartość, tym mocniejsze staje się przekonanie testującego o fałszywości hipotezy zerowej i prawdziwości hipotezy alternatywnej. Nikt nie odrzuci hipotezy zerowej otrzymawszy *p*-wartość rzędu 0,4 - zaobserwowana wartość statystyki testowej należy w takiej sytuacji uznać za zdecydowanie typową przy zachodzeniu hipotezy zerowej. Jeżeli zależy nam na bardzo „pewnym" spełnianiu hipotezy zerowej możemy ją odrzucić otrzymawszy *p*-wartość równą np. 0,12 - jeżeli np. hipoteza zerowa oznacza, że nowy konserwant nie zagraża zdrowiu Praktycznie zawsze odrzucimy hipotezę zerowa otrzymawszy *p*-wartość rzędu 0,001

Testy dla wartości średniej

Przykład 2

Specjaliści sieci supermarketów sprzedających między innymi produkty spożywcze podejrzewają, że mleko pochodzące od jednego z producentów kooperujących z siecią ma niższą zawartość tłuszczu niż nominalna wartość 3,2%. Specjaliści zakładają przy tym, że deklarowane przez producenta odchylenie standardowe zawartości tłuszczu w mleku nie zmieniło się i wynosi 0,05%. Ponadto zakładają, że faktyczna procentowa zawartość tłuszczu jest wielkością losową o rozkładzie normalnym. Postanowiono zatem poddać testowi hipotezę $\mathrm{H}_0 : \theta = 3,2$ gdzie θ oznacza procentową zawartość tłuszczu w mleku, przy alternatywie

$$\mathrm{H}_1 : \ \theta < 3,2$$

Testy dla wartości średniej

Rozwiązanie

Uzyskano następujące zawartości tłuszczu:

$$3,26,\ 3,12,\ 3,24,\ 3,16,\ 3,08,\ 3,14,\ 3,23,\ 3,11,\ 3,09,\ 3,24$$

Średnia w otrzymanej próbie wynosi 3,167, skąd statystyka testowa przyjmuje wartość

$$z = \frac{3,167 - 3,3}{0,05/\sqrt{10}} = -2,087$$

Zatem

$$p - \text{wartość} = P_{H_0}(Z \leqslant -2,087) = 1 - \Phi(2,087) = 0,0185$$

Przypadek rozkładu normalnego o nieznanym odchyleniu standardowym

Przypadek 2

Właściwą statystyką testową dla testowania hipotezy o średniej θ przy hipotezie alternatywnej jednostronnej lub dwustronnej jest oczywiście statystyka T dana wzorem

$$T = \frac{\overline{X} - \theta_0}{S/\sqrt{n}},$$

gdzie S jest odchyleniem standardowym w próbie oraz n jest licznością próby

Zbiory krytyczne dla różnych typów hipotez alternatywnych

$\mathrm{H}_1 : \theta > \theta_0$	$P_{\mathrm{H}_0}(T \geqslant t_{1-\alpha,n-1}) = \alpha$	$C = \{t : t \geqslant t_{1-\alpha,n-1}\}$
$\mathrm{H}_1 : \theta < \theta_0$	$P_{\mathrm{H}_0}(T \leqslant -t_{1-\alpha,n-1}) = \alpha$	$C = \{t : t \leqslant -t_{1-\alpha,n-1}\}$
$\mathrm{H}_1 : \theta \neq \theta_0$	$P_{\mathrm{H}_0}(\lvert T\rvert \geqslant t_{1-\alpha/2,n-1}) = \alpha$	$C = \{t : t \leqslant -t_{1-\alpha/2,n-1}$ lub $t \geqslant t_{1-\alpha/2,n-1}\}$

Przykład 3 - rozwiązanie

Testy dla dwóch prób w rodzinie rozkładów normalnych

Testy dla dwóch niezależnych prób

Rozważmy najpierw problem porównania wartości średnich dwóch różnych populacji, w przypadku gdy dysponujemy niezależnymi próbami losowymi z tych populacji, a mianowicie próbą o liczności n_1 z pierwszej populacji $X_1, X_2, ..., X_{n_1}$ o rozkładzie $\mathcal{N}(m_1, \sigma_1)$ oraz próbą o liczności n_2 z drugiej populacji, $Y_1, Y_2, ..., Y_n$ o rozkładzie $\mathcal{N}(m_2, \sigma_2)$. Hipoteza zerowa ma postać: $\mathrm{H}_0: \quad m_1 = m_2$
Hipoteza alternatywna może mieć jedną z następujących postaci:

$$\mathrm{H}_1: \quad m_1 > m_2$$
$$\mathrm{H}_1: \quad m_1 < m_2$$
$$\mathrm{H}_1: \quad m_1 \neq m_2$$

Testy dla dwóch prób w rodzinie rozkładów normalnych

Przypadek1: znane odchylenia standardowe obydwu populacji

Załóżmy, że są znane odchylenia standardowe obydwu populacji. Niech $\overline{X}$ i $\overline{Y}$ oznaczają, odpowiednio, średnią w pierwszej i drugiej próbie losowej. Wiemy już, że statystyka

$$Z = \frac{(\overline{X}-\overline{Y})-(m_1-m_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}$$

ma standardowy rozkład normalny. Jeżeli jest spełniona hipoteza zerowa powyższa statystyka przyjmuje postać

$$Z = \frac{(\overline{X}-\overline{Y})}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}$$

Zatem ostatecznie można testowanie równości średnich sprowadzić do wcześniej zbadanego problemu testowania pojedynczej średniej

Testy dla dwóch prób w rodzinie rozkładów normalnych

Przypadek2: nieznane odchylenia standardowe obydwu populacji

Przypadek nieznanych odchyleń standardowych σ_1 i σ_2 rozważymy jedynie przy założeniu równości obydwu odchyleń standardowych $\sigma_1 = \sigma_2$. Procedurę testową możemy oprzeć na statystyce, która dla hipotezy zerowej przyjmuje postać

$$T = \frac{(\overline{X} - \overline{Y})}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

i która ma wówczas rozkład t Studenta z $n_1 + n_2 - 2$ stopniami swobody

Testy dla dwóch prób w rodzinie rozkładów normalnych

Przypadek2: nieznane odchylenia standardowe obydwu populacji

Przypomnijmy, że oparta na obydwu próbach statystyka

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2},$$

gdzie S_i^2, $i = 1, 2$, jest wariancją w i-tej próbie, jest nieobciążonym estymatorem wariancji σ^2 tzn. $E(S_p^2) = \sigma^2$.

Testy dla dwóch prób w rodzinie rozkładów normalnych

Zbiory krytyczne dla różnych typów hipotez alternatywnych

Przypadek1: znane odchylenia standardowe obydwu populacji:

$\mathrm{H}_1: \ m_1 > m_2$	$P_{\mathrm{H}_0}(Z \geqslant z_{1-\alpha}) = \alpha$	$C = \{z : z \geqslant z_{1-\alpha}\}$
$\mathrm{H}_1: \ m_1 < m_2$	$P_{\mathrm{H}_0}(Z \leqslant -z_{1-\alpha}) = \alpha$	$C = \{z : z \leqslant -z_{1-\alpha}\}$
$\mathrm{H}_1: \ m_1 \neq m_2$	$P_{\mathrm{H}_0}(\lvert Z\rvert \geqslant z_{1-\alpha/2}) = \alpha$	$C = \{z : z \leqslant -z_{1-\alpha/2}$ lub $z \geqslant z_{1-\alpha/2}\}$

Testy dla dwóch prób w rodzinie rozkładów normalnych

Zbiory krytyczne dla różnych typów hipotez alternatywnych

Przypadek2: nieznane odchylenia standardowe obydwu populacji:

$$
\begin{array}{l|l}
\mathrm{H}_1 : \; m_1 > m_2 & C = \{t : t \geqslant t_{1-\alpha, n_1+n_2-2}\} \\
\mathrm{H}_1 : \; m_1 < m_2 & C = \{t : t \leqslant -t_{1-\alpha, n_1+n_2-2}\} \\
\mathrm{H}_1 : \; m_1 \neq m_2 & C = \{t : t \leqslant -t_{1-\alpha/2, n_1+n_2-2} \\
 & \text{lub} \;\; t \geqslant t_{1-\alpha/2, n_1+n_2-2}\}
\end{array}
$$

Testy dla dwóch prób w rodzinie rozkładów normalnych

Pary obserwacji

Jakościowo inna sytuacja, gdy mamy do czynienia z **parami obserwacji**

$$(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n),$$

gdzie pary mają taki sam dwuwymiarowy rozkład normalny i są wzajemnie niezależne, ale zmienne w parze mogą być zależne. Różnice $D_i = X_i - Y_i$ tworzą próbę niezależnych zmiennych losowych o rozkładzie normalnym z nieznaną wartością średnią m_D. Hipoteza zerowa przyjmuje wówczas postać: $\mathrm{H}_0 : \ m_D = 0$, natomiast możliwe hipotezy alternatywne: $\mathrm{H}_1 : \ m_D < 0$, $\mathrm{H}_1 : \ m_D > 0$, $\mathrm{H}_1 : \ m_D \neq 0$

Testy dla dwóch prób w rodzinie rozkładów normalnych

Pary obserwacji

Zauważmy dalej, że różnice D_i tworzą próbę niezależnych zmiennych losowych o rozkładzie normalnym z nieznaną wartością średnią $m_D = m_1 - m_2$, i że statystyka:

$$T = \frac{\overline{D} - m_D}{S_D/\sqrt{n}}$$

ma rozkład t Studenta z $n-1$ stopniami swobody. Pamiętamy, że S_D^2 jest estymatorem warianci różnic D_i, tzn.

$$S_D^2 = \frac{1}{n-1}\sum_{i=1}^{n}(D_i - \overline{D})^2,$$

gdzie $\overline{D} = \frac{1}{n}\sum_{i=1}^{n} D_i$

Testy dla dwóch prób w rodzinie rozkładów normalnych

Pary obserwacji

Przy H_0 statystyka:

$$T = \frac{\overline{D}}{S_D/\sqrt{n}}$$

ma rozkład t Studenta z $n-1$ stopniami swobody.
W ten sposób zadanie konstrukcji testów dla porównania wartości średnich par obserwacji sprowadza się do analogicznego zadania dla pojedynczej wartości średniej (mianowicie wartości średniej różnic D_i przy nieznajomości ich standardowego odchylenia)

Przykład 4

Jednym z testów, którymi rozpoczęto analizę nowego leku na nadciśnienie tętnicze było zaaplikowanie go próbie 22 chorych pacjentów, u których ciśnienie skurczowe było bliskie wartości 144 mmHg. Ponieważ górna granica normy tego ciśnienia wynosi 140, chciano sprawdzić, czy zastosowanie określonej terapii badanym lekiem daje obniżenie ciśnienia o około 5 mmHg. Każdemu pacjentowi zmierzono ciśnienie skurczowe przed rozpoczęciem terapii i po jej zakończeniu. W ten sposób dla i-tego pacjenta dysponowano parą wyników (x_i, y_i). Dla próby 22 pacjentów otrzymano $\overline{d} = 5,3$ oraz $s_D = 0,4$

Rozwiązanie

Celem zadania jest poddanie testowi hipotezy zerowej

$$H_0: \ m_D = 5 \quad \text{przy hipotezie alternatywnej} \quad H_1: \ m_D \neq 5$$

Należy zastosować powyższą statystykę, która przy hipotezie zerowej przyjmuje wówczas postać

$$T = \frac{\overline{D} - d_0}{S_D/\sqrt{n}},$$

gdzie w naszym przypadku $d_0 = 5$. Wartość statystyki $t = 3,518$, co dało p-wartość (dla rozkładu t Studenta z 21 stopniami swobody i przy dwustronnej hipotezie alternatywnej) 0,002. Zatem zdecydowanie odrzucamy hipotezę zerową - terapia nie spełnia nałożonych założeń

Testy dla wariancji w rodzinie rozkładów normalnych

Testy dla wariancji

Niech dana będzie próba losowa o liczności n z rozkładu normalnego o nieznanej wariancji σ^2. Do testowania hipotezy zerowej $\mathrm{H}_0: \ \sigma^2 = \sigma_0^2$ użyjemy statystyki

$$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2},$$

która ma przy H_0 rozkład χ^2 z $n-1$ stopniami swobody

Zbiory krytyczne dla różnych typów hipotez alternatywnych

$\mathrm{H}_1 : \sigma^2 > \sigma_0^2$	$P_{\mathrm{H}_0}(\chi^2 \geqslant \chi^2_{1-\alpha,n-1}) = \alpha,$ $C = \{x^2 : x^2 \geqslant \chi^2_{1-\alpha,n-1}\}$
$\mathrm{H}_1 : \sigma^2 < \sigma_0^2$	$P_{\mathrm{H}_0}(\chi^2 \leqslant \chi^2_{\alpha,n-1}) = \alpha,$ $C = \{x^2 : x^2 \leqslant \chi^2_{\alpha,n-1}\}$
$\mathrm{H}_1 : \sigma^2 \neq \sigma_0^2$	$P_{\mathrm{H}_0}(\chi^2 \leqslant \chi^2_{\alpha/2,n-1}) + P_{\mathrm{H}_0}(\chi^2 \geqslant \chi^2_{1-\alpha/2,n-1}) = \alpha,$ $C = \{x^2 : x^2 \leqslant \chi^2_{\alpha/2,n-1} \text{ lub } x^2 \geqslant \chi^2_{1-\alpha/2,n-1}\}$

Testy dla proporcji

Problem testowania i statystyka testowa

Ograniczymy się do przypadku dostatecznej liczności próby, by móc skorzystać z przybliżenia normalnego statystyki

$$Z = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}},$$

gdzie p jest prawdziwą wartością prawdopodobieństwa sukcesu, n jest licznością próby, na podstawie której obliczamy częstość $\hat{p}$

Testy dla proporcji

Problem testowania i statystyka testowa

Przy założeniu prawdziwości hipotezy zerowej $\mathrm{H}_0: \; p = p_0$ statystyka

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

ma w przybliżeniu standardowy rozkład normalny i zadanie testowania hipotezy zerowej przy hipotezie alternatywnej

$$\mathrm{H}_1: \; p > p_0 \quad \text{lub} \quad p < p_0 \quad \text{lub} \quad p \neq p_0$$

sprowadza się do zadania testowania hipotez o wartości średniej rozkładu normalnego przy znanym odchyleniu standardowym

Przykład 5

Pewne ugrupowanie polityczne było przekonane, że poparcie Polaków dla wejścia ich kraju do UE nigdy nie przekroczy 53%. Przeprowadzona w czerwcu 2000r. ankieta wśród 1000 dorosłych Polaków dała 57% poparcie starań Polski do UE. Przetestować hipotezę wspomnianego ugrupowania politycznego. Przyjąć poziom istotności $\alpha = 0.025$

Rozwiązanie

Problem testowania:

$$H_0: \; p = 0,53 \quad \text{przy hipotezie alternatywnej} \quad H_1: \; p > 0,53$$

Statystyka ma wartość

$$z = \frac{0,57 - 0,53}{\sqrt{\frac{0,53(1-0,53)}{1000}}} = 2,534$$

co daje p-wartość:

$$p = P_{H_0}(Z \geqslant 2,534) = 1 - \Phi(2,534) = 0,006$$

Na poziomie istotności 0,025 ($> p = 0,006$) odrzucamy hipotezę wspomnianego ugrupowania politycznego