Statystyczna analiza danych SAD

Wykład 10

Testowanie hipotez statystycznych

Hipoteza statystyczna jest to przypuszczenie dotyczące nieznanej własności rozkładu prawdopodobieństwa badanej cechy populacji.

Testowanie hipotez statystycznych

Przykłady

(a) Producent opon twierdzi, że nowy typ opony ma trwałość większą niż 60000 km. Jeśli μ (km) oznacza wartość średnią trwałości opon, to hipotezą producenta jest $\boxed{H : \mu > 60000}$.

(b) Producent twierdzi, że średni czas bezawaryjnej pracy drukarki to więcej niż 200 godzin. Wówczas: $\boxed{H : \mu > 200}$

(c) Socjolog twierdzi, że dzieci w miastach mają lepsze wyniki w nauce niż dzieci poza ośrodkami miejskimi. Niech p_1 (p_2) oznacza proporcję dzieci w miastach (poza miastami) o średnich ocenach rocznych co najmniej dobrych. Hipotezą socjologa jest $\boxed{H : p_1 > p_2}$.

(d) Fizycy przypuszczają, że ilość cząstek emitowanych przez substancję radioaktywną w przedziałach czasu o danej długości jest zmienną losową o rozkładzie Poissona. Wówczas: $\boxed{H : X \sim P(\lambda), \lambda > 0.}$

Testowanie hipotez statystycznych

Hipotezę nazywamy **parametryczną**, jeśli jest stwierdzeniem dotyczącym nieznanego parametru liczbowego lub wektorowego rozkładu cechy populacji, np. hipotezy (a), (b), (c).

W przeciwnym przypadku hipoteza jest **nieparametryczna,** np. hipoteza (d).

W zadaniach testowania hipotez formułowane są 2 hipotezy:

Hipoteza zerowa to hipoteza testowana w celu jej ewentualnego odrzucenia, oznaczana przez $\boxed{H_0}$.

Hipoteza alternatywna to hipoteza, która będzie przyjęta, jeśli odrzucimy hipotezę zerową, oznaczana przez $\boxed{H_1}$.

Testowanie hipotez statystycznych

Hipotezy H_0 i H_1 muszą się nawzajem wykluczać (nie mogą być jednocześnie prawdziwe), np. niech $p \in (0,1)$ oznacza prawdopodobieństwo sukcesu w doświadczeniu Bernoulli'ego. Możliwe są hipotezy:

$$\boxed{H_0 : p = \frac{1}{2}, \quad H_1 : p \neq \frac{1}{2}},$$

lub

$$\boxed{H_0 : p = \frac{1}{2}, \quad H_1 : p > \frac{1}{2}},$$

ale niemożliwe

$$\boxed{H_0 : p = \frac{1}{2}, \quad H_1 : p > \frac{1}{3}},$$

ponieważ wartość $p = \frac{1}{2}$ jest parametrem z zakresu H_0 i H_1 jednocześnie.

Zbiory parametrów wymieniane w obu hipotezach nie są więc rozłączne.

Testowanie hipotez statystycznych

Rola hipotez H_0 i H_1 nie jest symetryczna:

- Hipoteza alternatywna H_1 to ta, którą zaakceptujemy, jeśli próbka dostarczy nam dostatecznych dowodów jej prawdziwości, ta o której sądzimy, że jest prawdziwa i szukamy potwierdzenia w próbce (ta na której nam zależy aby była prawdziwa).

- Hipoteza zerowa H_0 to ta, o której prawdziwości nie jesteśmy przekonani w sytuacji gdy nie możemy zaakceptować na podstawie próbki hipotezy alternatywnej (ta którą poddajemy w wątpliwość).

Testowanie hipotez statystycznych

Przykład. Załóżmy, że skuteczność pewnej terapii medycznej wynosi 80%. Zaproponowano nową terapię, której oczekiwana skuteczność $p \times 100\%$ powinna być lepsza (tzn. oczekujemy, że $p \geq 0{,}8$).
Nowa terapia będzie szeroko stosowana, jeśli po badaniach wstępnych uzyskamy „pewność", że $p > 0{,}8$. Zatem:

$$H_0\colon\ p = 0{,}8 \qquad\qquad H_1\colon\ p > 0{,}8$$

Testowanie hipotez statystycznych

Przykład. Nowa technologia produkcji może zmniejszyć dobowy poziom emisji zanieczyszczeń do atmosfery. Chcielibyśmy wiedzieć, czy zmniejsza ona poziom zanieczyszczeń? Wówczas:

H_0: Nowa technologia <u>nie zmniejsza</u> dobowego poziomu emisji zanieczyszczeń atmosfery (nie jest lepsza od starej).

H_1: Nowa technologia <u>zmniejsza</u> dobowy poziom emisji zanieczyszczeń atmosfery (tzn. jest lepsza).

Możliwe decyzje weryfikacyjne:

- Nie ma dostatecznych dowodów aby odrzucić H_0, (i tym samym. przyjąć H_1), tzn. na podstawie obserwacji nie możemy stwierdzić, że nowa technologia zmniejsza poziom zanieczyszczeń.
- Obserwacje dostarczają dostatecznych dowodów, aby przyjąć H_1, równoważnie odrzucić H_0, tzn. stwierdzamy, iż można uznać, że nowa technologia zmniejsza poziom zanieczyszczeń.

Testowanie hipotez statystycznych

Model matematyczny

(a) μ_0 - znany średni poziom dobowy emisji przy starej technologii

(b) μ - nieznany średni poziom dobowy emisji przy nowej technologii

(c) wiemy, że $\mu \leq \mu_0$. Chcielibyśmy stwierdzić, że nowa technologia zmniejsza poziom emisji. Zatem:

$$H_0\colon \mu = \mu_0 , \qquad H_1\colon \mu < \mu_0 .$$

(d) w ciągu *n* losowo wybranych dni obserwujemy dobowe poziomy emisji przy nowej technologii: $X_1, X_2, ..., X_n$.

(e) zmienne losowe $X_1, X_2, ..., X_n$ są <u>niezależne o jednakowym rozkładzie</u> $N(\mu, \sigma)$, gdzie σ jest znane.

Testowanie hipotez statystycznych

Decyzje: „**przyjąć** $\boldsymbol{H_1}$”, lub „**nie można odrzucić** $\boldsymbol{H_0}$” oprzemy na podstawie realizacji **średniej z próby losowej** $\bar{X}$, tzn. **średniej z próbki** $\bar{x}$**.**

Uzasadnienie

Średnia z próby losowej $\bar{X}$ ma rozkład $N(\mu, \frac{\sigma}{\sqrt{n}})$ skoncentrowany wokół μ. Zatem dostatecznie małe wartości $\bar{X}$ sugerują, że hipoteza $H_1 : \mu < \mu_0$ jest prawdziwa, ponieważ:

(1) jeśli $H_0 : \mu = \mu_0$ jest prawdziwa, to wartości $\bar{X}$ skupiają się wokół μ_0, a statystyka

$$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} \approx N(0,1)$$

(2) jeśli $H_1 : \mu < \mu_0$ jest **prawdziwa**, tzn. nieznane $\mu = \mu_1 < \mu_0$, to wartości $\bar{X}$ skupiają się wokół μ_1. Wówczas Z jest sumą zmiennej o rozkładzie $N(0,1)$ oraz stałej ujemnej

$$\boxed{Z = \frac{\bar{X} - \mu_1}{\sigma / \sqrt{n}} + \frac{\mu_1 - \mu_0}{\sigma / \sqrt{n}}}$$

Testowanie hipotez statystycznych

Punkty (1) i (2) sugerują **sposób testowania.**

Niech c będzie odpowiednio dobraną stałą, a $\bar{x}$ wartością statystyki $\bar{X}$ obliczoną dla konkretnej próbki, wówczas

(i) jeśli $z = \dfrac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} \leq c$, to <u>przyjmujemy H_1</u>:.

(ii) jeśli $z = \dfrac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} > c$, to <u>nie ma podstaw do odrzucenia H_0</u>.

Wybór stałej c

Niech α będzie małą liczbą z przedziału (0,1), np. α = 0,01, 0,05,

zaś $c = z_\alpha = -z_{1-\alpha}$.

Wówczas jeśli $H_0 : \mu = \mu_0$ prawdziwa, to $P_{H_0}(Z \leq z_\alpha) = \alpha$.

Stąd α jest prawdopodobieństwem podjęcia błędnej decyzji (polegającej na przyjęciu hipotezy H_1) w przypadku gdy hipoteza H_0 jest prawdziwa.

Testowanie hipotez statystycznych

Prawdopodobieństwo α podjęcia błędnej decyzji (polegającej na przyjęciu hipotezy H_1, w przypadku gdy hipoteza H_0 jest prawdziwa) nazywamy **prawdopodobieństwem błędu I rodzaju,** lub **poziomem istotności testu.**

Zbiór $C = \{z : z \leq z_\alpha\}$ nazywamy **zbiorem krytycznym -** jest to zbiór wartości statystyki testowej Z dla których **odrzucamy** H_0 na korzyść H_1.

I. Testowanie hipotez o wartości średniej rozkładu normalnego, gdy znana jest wariancja.

Założenie: Zmienne losowe $X_1, X_2, \dots, X_n$ tworzą prostą próbą losową z rozkładu $N(\mu, \sigma)$, σ - znane.

$$\boxed{H_0 : \mu = \mu_0}$$

Statystyka testowa:

$$\boxed{Z = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}}} = \frac{\overline{X} - \mu}{\sigma / \sqrt{n}} + \frac{\mu - \mu_0}{\sigma / \sqrt{n}}.$$

Jeśli H_0 **prawdziwa,** to $Z \sim N(0,1)$.

Testowanie hipotez statystycznych

Model 1. H_0: $\mu = \mu_0$ przeciw H_1: $\mu > \mu_0$.

Wówczas przyjmujemy $\boxed{C = \{z : z \geq z_{1-\alpha}\}}$ = zbiór krytyczny testu hipotezy H_0 przeciw H_1 na poziomie istotności α, gdzie

$$P_{H_0}(Z \in C) = P_{H_0}(Z \geq z_{1-\alpha}) = \alpha .$$

Model 2. H_0: $\mu = \mu_0$ przeciw H_1: $\mu < \mu_0$.

Wówczas przyjmujemy $\boxed{\{z : z \leq -z_{1-\alpha}\}}$ - zbiór krytyczny, gdzie

$$P_{H_0}(Z \in C) = P_{H_0}(Z \leq -z_{1-\alpha}) = \alpha .$$

Model 3. H_0: $\mu = \mu_0$ przeciw H_1: $\mu \neq \mu_0$.

Wówczas przyjmujemy $\boxed{C = \{z : |z| \geq z_{1-\alpha/2}\}}$ - zbiór krytyczny, gdzie

$$P_{H_0}(Z \leq -z_{1-\alpha/2}) = \alpha / 2$$

$$P_{H_0}(Z \in C) = P_{H_0}(|Z| \geq z_{1-\alpha/2}) = \alpha .$$

Testowanie hipotez statystycznych

Przykład. Dotychczasowa dzienna wartość sprzedaży pewnego artykułu miała rozkład normalny o średniej 1000 (\$) i standardowym odchyleniu 100 (\$). Po serii reklam telewizyjnych, w ciągu 9 losowo wybranych dni uzyskano następujące wartości sprzedaży:

1280, 1250, 990, 1100, 880, 1300, 1100, 950, 1050.

Czy, na poziomie istotności $\alpha = 0{,}01$, można twierdzić, że reklamy spowodowały zwiększenie sprzedaży, jeśli można założyć, że wartości dziennych sprzedaży są niezależnymi zmiennymi losowymi o jednakowym rozkładzie normalnym?

Testowanie hipotez statystycznych

Błędy testowania,
(przy symetrycznym traktowaniu hipotez)

		Podjęta decyzja	
		Akceptacja H_0	Odrzucenie H_0 (Akceptacja H_1)
Stan rzeczywisty	H_0 prawdziwa	Decyzja prawidłowa	Błąd I rodzaju
	H_1 prawdziwa	Błąd II rodzaju	Decyzja prawidłowa

Testowanie hipotez statystycznych

Błędy testowania,
(przy **„niesymetrycznym”** traktowaniu hipotez)

		Podjęta decyzja	
		Nie odrzucamy H_0 (H_0?)	Odrzucenie H_0 (Akceptacja H_1)
Stan rzeczywisty	H_0 prawdziwa	(?)	Błąd I rodzaju
	H_1 prawdziwa	(?) Błąd II rodzaju	Decyzja prawidłowa

Testowanie hipotez statystycznych

Rozwiązanie:

1. H_0: $\mu = 1000$ przeciw H_1: $\mu > 1000$

2. Statystyka testowa: $Z = \dfrac{\overline{X} - 1000}{\sigma / \sqrt{n}}$

3. $\alpha = 0{,}01$, $1 - \alpha = 0{,}99$, $z_{0,99} = 2{,}33$.

Zbiór krytyczny $C = \{ z: z \geq 2{,}33\}$

4. $\sigma = 100$, $n = 9$, średnia próbkowa $\overline{x} = 1100$, stąd wartość statystyki testowej

$$z = \frac{\overline{x} - 1000}{\sigma / \sqrt{n}} = \frac{1100 - 1000}{100/3} = 3.$$

5. Ponieważ $3 \geq 2{,}33$ ($z = 3 \in C$) <u>odrzucamy</u> H_0) .

Odpowiedź: Na poziomie istotności $\alpha = 0{,}01$ stwierdzamy, że średnia wartość sprzedaży wzrosła po serii reklam.

Testy parametryczne dla średniej populacji

Testowanie hipotez o wartości średniej rozkładu normalnego, gdy nieznana jest wariancja.

Niech $X_1, X_2, ..., X_n$ będzie prostą próbą losowa z rozkładu $N(\mu, \sigma)$, σ - nieznane.

$$\boldsymbol{H_0: \mu = \mu_0}$$

Statystyka testowa:

$$\boxed{T = \frac{\overline{X} - \mu_0}{S/\sqrt{n}}} = \frac{\overline{X} - \mu}{S/\sqrt{n}} + \frac{\mu - \mu_0}{S/\sqrt{n}}.$$

Jeśli $\boldsymbol{H_0}$ **prawdziwa,** to $\boldsymbol{T \sim t_{n-1}}$ = rozkład t Studenta o $n-1$ stopniach swobody

Testy parametryczne dla średniej populacji

Model 1. H_0: $\mu = \mu_0$ **przeciw** H_1: $\mu > \mu_0$.

Wówczas przyjmujemy **C** = $\{t : t \geq t_{1-\alpha,n-1}\}$ = **zbiór krytyczny** testu hipotezy H_0 przeciw H_1 **na poziomie istotności** α, gdzie

$$P_{H_0}(T \in C) = P_{H_0}(T \geq t_{1-\alpha,n-1}) = \alpha,$$

$t_{1-\alpha,n-1}$ = kwantyl rzędu **1 -** α rozkładu t Studenta z n - 1 stopniami swobody.

Model 2. H_0: $\mu = \mu_0$ **przeciw** H_1: $\mu < \mu_0$.

Wówczas przyjmujemy **C** = $\{t : t \leq -t_{1-\alpha,n-1}\}$ - **zbiór krytyczny**, gdzie

$$P_{H_0}(T \in C) = P_{H_0}(T \leq -t_{1-\alpha,n-1}) = \alpha.$$

Testy parametryczne dla średniej populacji

Model 3. H_0: $\mu = \mu_0$ **przeciw** H_1: $\mu \neq \mu_0$ **.**

Wówczas $C = \{t : |t| \geq t_{1-\alpha/2,n-1}\}$ - **zbiór krytyczny**, gdzie

$$P_{H_0}(T \leq -t_{1-\alpha/2,n-1}) = \alpha/2, \quad \text{skąd} \quad P_{H_0}(|T| \geq t_{1-\alpha/2,n-1}) = \alpha.$$

Zadanie. Producent twierdzi, że nowy model samochodu ma wartość średnią przebiegu nie wymagającą żadnej interwencji 120000 (km). W teście dla 4 losowo wybranych samochodów uzyskano następujące przebiegi nie wymagające żadnego serwisu: 110000, 120000, 118000, 112000. Czy można twierdzić, że producent zawyża wartość średniego przebiegu nowego modelu? Przyjmij α = **0,05** oraz rozkład normalny przebiegu.

Test parametryczny dla średniej populacji

1. H_0: μ **= 120000 przeciw** H_1: μ **< 120000**

2. Statystyka testowa:

$$\boxed{T = \frac{\bar{X} - 120000}{S/\sqrt{n}}}$$

3. α **= 0,05, 1 –** α **= 0,95, liczba stopni swobody** = $n-1$ **= 4 –1 = 3**,

$\boxed{t_{0,95,3}}$ = 2,353. Stąd **zbiór krytyczny C** = { t: $t \leq -$ 2,353}.

Test parametryczny dla średniej populacji

4. $n = 4$, z obliczeń $\bar{x} = 115000$, $s^2 = \dfrac{680000}{4-1} = 226667$, stąd wartość statystyki testowej

$$t = \frac{\bar{x} - 120000}{s/\sqrt{n}} = \frac{115000 - 120000}{\sqrt{226667}/\sqrt{4}} = -2{,}10.$$

5. –2,10 > – 2,353, więc nie ma podstaw do odrzucenia $\boldsymbol{H_0}$ na poziomie istotności 0,05.

Odpowiedź: Na poziomie istotności α **= 0,05** stwierdzamy, że nie można odrzucić twierdzenia producenta

P – wartość

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

Np. w ostatnim zadaniu

$$t = -2{,}10, \qquad P_{H_0}(T \le -t_{1-\alpha,n-1}) = \alpha$$

$$P_{H_0}(T \le -2{,}10) = 0{,}063.$$

Im mniejsza jest ***p*-wartość**, tym mocniejsze staje się przekonanie testującego o fałszywości hipotezy zerowej i prawdziwości hipotezy alternatywnej.

Testy parametryczne dla wariancji

Testowanie hipotez o wariancji rozkładu normalnego, gdy nieznana jest wartość średnia.

Niech $X_1, X_2, \ldots, X_n$ będzie prostą próbą losowa z rozkładu $N(\mu, \sigma)$, μ, σ - nieznane.

$$\boxed{H_0 : \sigma^2 = \sigma_0^2}.$$

Statystyka testowa:

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

Jeśli $\boxed{H_0}$ **prawdziwa,** to $\chi^2 \sim \chi^2_{n-1}$.

Testy parametryczne dla wariancji

Model 1. $\boxed{H_0 : \sigma^2 = \sigma_0^2}$, $\boxed{H_1 : \sigma^2 > \sigma_0^2}$.

Wówczas ***C* =** $\{\chi^2_{obl} : \dfrac{(n-1)s^2}{\sigma_0^2} \geq \chi^2_{1-\alpha,n-1}\}$ **= zbiór krytyczny** testu hipotezy H_0 przeciw H_1 na **poziomie istotności** α, gdzie

$$P_{H_0}(\chi^2 \in C) = P_{H_0}(\chi^2 \geq \chi^2_{1-\alpha,n-1}) = \alpha,$$

$\chi^2_{1-\alpha,n-1}$ = **kwantyl rzędu 1 - α rozkładu** χ^2_{n-1}.

Testy parametryczne dla wariancji

Model 2. $\boxed{H_0 : \sigma^2 = \sigma_0^2}$, $\boxed{H_1 : \sigma^2 < \sigma_0^2}$.

Wówczas ***C*** = $\{\chi^2_{obl} : \frac{(n-1)s^2}{\sigma_0^2} \le \chi^2_{\alpha,n-1}\}$ - **zbiór krytyczny**, gdzie

$$P_{H_0}(\chi^2 \in C) = P_{H_0}(\chi^2 \le \chi^2_{\alpha,n-1}) = \alpha.$$

Testy parametryczne dla wariancji

Model 3. $\boxed{H_0 : \sigma^2 = \sigma_0^2}$, $\boxed{H_1 : \sigma^2 \neq \sigma_0^2}$.

Wówczas **zbiór krytyczny** ***C*** =

$$\{\chi^2_{obl} : \frac{(n-1)s^2}{\sigma_0^2} \leq \chi^2_{\alpha/2}\} \cup \{\chi^2_{obl} : \frac{(n-1)s^2}{\sigma_0^2} \geq \chi^2_{1-\alpha/2}\},$$

gdzie $\chi^2_{\alpha/2} = \chi^2_{\alpha/2, n-1}$, $\chi^2_{1-\alpha/2} = \chi^2_{1-\alpha/2, n-1}$.

Test parametryczny dla wariancji

Zadanie. Zmierzono czas życia 15 losowo wybranych żarówek z bieżącej produkcji. Policzono standardowe odchylenie próbkowe *s* = 13 (godz.). Czy na poziomie istotności α = 0,05 (5%) można twierdzić, że odchylenie standardowe czasu życia losowo wybranej żarówki jest różne od 10 (godz.)

Rozwiązanie.

1. $\boxed{H_0 : \sigma = 10}$ **przeciw** $\boxed{H_1 : \sigma \neq 10}$

2. Statystyka testowa: $\chi^2 = \dfrac{(n-1)S^2}{10^2}$

3. $\alpha = 0{,}05$, $\alpha/2 = 0{,}025$, $1-\alpha/2 = 0{,}975$, $n = 15$, liczba stopni swobody $n-1 = 15-1 = 14$,

$$\chi^2_{\alpha/2,n-1} = \chi^2_{0{,}025,14} = 5{,}629, \quad \chi^2_{1-\alpha/2,n-1} = \chi^2_{0{,}975,14} = 26{,}119.$$

Reguła decyzyjna (na podstawie **zbioru krytycznego**): **odrzuć H_0, jeśli obliczona wartość statystyki**

$$\chi^2_{obl} \leq 5{,}629 \quad \text{lub} \quad \chi^2_{obl} \geq 26{,}119.$$

Test parametryczny dla wariancji

Rozwiązanie (c.d.)

4. s =13, stąd wartość statystyki testowej

$$\boxed{\chi^2_{obl}} = \frac{(n-1)s^2}{100} = \frac{(14)(13^2)}{100} = 23{,}66.$$

5. 5,629 < 23,66 < 26,119, więc nie ma podstaw do odrzucenia H_0.

Odpowiedź. Na poziomie istotności 0,05, brak jest dostatecznych dowodów aby twierdzić, że $\sigma \neq 10$.

Testowanie hipotez statystycznych

Błędy testowania,
(przy symetrycznym traktowaniu hipotez)

		Podjęta decyzja	
		Akceptacja H_0	Odrzucenie H_0 (Akceptacja H_1)
Stan rzeczywisty	H_0 prawdziwa	Decyzja prawidłowa	Błąd I rodzaju
	H_1 prawdziwa	Błąd II rodzaju	Decyzja prawidłowa