

Metody probabilistyczne i statystyka - wykład dwunasty¹

Testowania hipotez: testy istotności

dr Jarosław Kotowicz

Instytut Informatyki Uniwersytet w Białymstoku

wersja z roku ak. 2020/21

¹©J.Kotowicz, 2021

Spis treści

1 Parametryczne testy istotności

- Ogólne pojęcie testów istotności
- Testy istotności dla wartości średniej
 - Populacja normalna ze znanym odchyleniem standardowym
 - Populacja normalna z nieznanym odchyleniem standardowym (mała próba)
 - Populacja generalna o dowolnym rozkładzie
- Testy istotności dla wartości dwóch średnich
 - Populacje normalne
 - Populacje dowolne (duże próby)
- Testy istotności dla wariancji
- Testy istotności dla dwóch wariancji

2 Nieparametryczne testy istotności

Test istotności. I

Definicja 1

Mówimy, że δ jest testem na poziomie istotności α jeżeli

$$\sup_{\theta \in \Theta_0} P_{\theta}(\{\delta(X) = 1\}) \leq \alpha. \quad (1)$$

Uwaga 1

W przypadku hipotezy prostej warunek (1) ma postać

$$P_{\theta_0}(\{\delta(X) = 1\}) \leq \alpha.$$

Do najczęściej stosowanych w praktyce testów statystycznych należą testy istotności, nazywane tak ze względu na to, że w testach tych uwzględnia się tylko poziom istotności. Testy te mają taką własność, że dla zadanego z góry poziomu istotności α zapewniają możliwie najmniejszą

Test istotności. II

wartość błędu II-go rodzaju, czyli możliwie największą moc. Testy te są również testami zgodnymi.

W testach istotności zdecydowanie odrzuca się hipotezę zerową, gdy wartość sprawdzianu wpada do zbioru krytycznego, w przeciwnym wypadku orzeka się tylko, że nie ma podstaw do odrzucenia tej hipotezy.

Zakładać będziemy, że mamy do dyspozycji wyniki obserwacji dla prób prostych.

Uwaga 2

*W praktyce np. w testach dostępnych w programach statystycznych podawana jest **p-wartość** (ang. *p-value*). Przedstawmy jej definicję.*

Test istotności. III

Definicja 2

Założmy, że test ma postać $\delta(X) = \mathbb{I}_{\{T(X) > c\}}$. Poziom krytyczny testu lub inaczej *p-wartość* jest to najmniejszy poziom istotności przy którym odrzucamy hipotezę H_0 tzn. jeśli zaobserwujemy $x \in \mathcal{X}$, to

$$P = \sup_{\theta \in \Theta_0} P_{\theta}(\{T(X) > T(x)\}).$$

Zasady konstrukcji testów istotności.

- 1 Formułujemy hipotezę zerową i alternatywną.
- 2 Na podstawie próby losowej (X_1, \dots, X_n) wyznaczamy statystykę T_n (sprawdzian hipotezy), której rozkład określamy przy założeniu, że prawdziwa jest hipoteza zerowa.

Test istotności. IV

- 3 Dla ustalonego z góry „małego” prawdopodobieństwa α wyznaczamy obszar krytyczny Λ tak, aby

$$P(\{T_n \in \Lambda\}) = \alpha.$$

- 4 Jeżeli konkretna realizacja próby należy do Λ , to hipotezę zerową odrzucamy, w przeciwnym wypadku nie ma podstaw do odrzucenia hipotezy zerowej.

Uwaga 3

α jest poziom istotności (prawdopodobieństwem popełnienia błędu I-go rodzaju). W praktyce $\alpha \in [0,01; 0,1]$.

Założenia dotyczące populacji i próby

Zakładamy, że rozkład cechy w zbiorowości generalnej jest rozkładem normalnym $\mathcal{N}(m, \sigma)$. Wybór sprawdzianu hipotezy zależy od liczebności próby n oraz od tego, czy parametr σ w zbiorowości generalnej jest znany. I tak, jeśli:

- σ jest znane i $n \leq 30$ albo
- σ jest znane i $n > 30$ lub
- σ jest nieznanne i $n > 30$, ale wówczas $\sigma \approx s$,

to sprawdzianem hipotezy $H_0 : m = m_0$ jest statystyka

$$U = \frac{\bar{X} - m}{\sigma} \sqrt{n}$$

o rozkładzie $\mathcal{N}(0,1)$.

Natomiast, gdy σ jest nieznanne i $n \leq 30$, sprawdzianem tej hipotezy jest:

$$T = \frac{\bar{X} - m}{S} \sqrt{n-1} \quad \text{lub} \quad T = \frac{\bar{X} - m}{\tilde{S}} \sqrt{n}$$

Założenia w przypadku populacji normalnej ze znanym odchyleniem

W przypadku znanego odchylenia standardowego lub też nieznanego odchylenia standardowego, lecz dużej próby

- 1 Estymator - średnia z próby \bar{X} .
- 2 Hipoteza zerowa $H_0 : m = m_0$.
- 3 Jeżeli hipoteza zerowa jest prawdziwa, to \bar{X} ma rozkład $\mathcal{N}\left(m_0, \frac{\sigma}{\sqrt{n}}\right)$.
- 4 Sprawdzian hipotezy - statystyka $U = \frac{\bar{X} - m_0}{\sigma} \sqrt{n}$, która ma rozkład normalny standardowy.
- 5 Poziom istotności α .
- 6 Parametrem opisującym rodzinę rozkładów jest $\theta \equiv m$. Będziemy też pisać θ_0 zamiast m_0 .

Przypadek I. I

Testujemy hipotezę H_0 przeciwko hipotezie H_1 , gdzie

$$H_0 : m = m_0$$

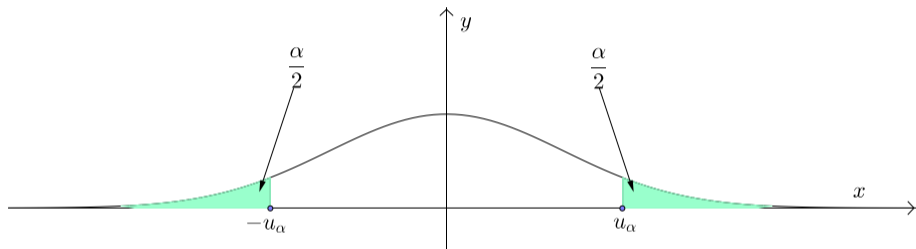
$$H_1 : m \neq m_0.$$

Dla poziomu istotności α mamy

$$P_{\theta_0}(\{|U| \geq u_\alpha\}) = \alpha.$$

Obszar krytyczny $\{u : |u| \geq u_\alpha\}$, u_α wartość krytyczna, $\Phi(u_\alpha) = 1 - \frac{\alpha}{2}$.

Przypadek I. II



Rysunek: Obustronny obszar krytyczny testu istotności dla średniej populacji normalnej ze znanym σ .

Przypadek II. I

Testujemy hipotezę H_0 przeciwko hipotezie H_1 , gdzie

$$H_0 : m = m_0$$

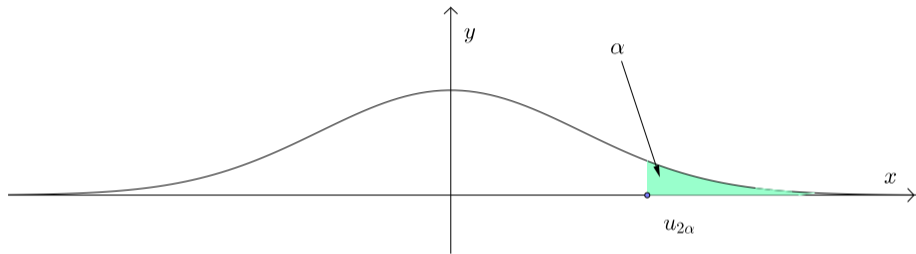
$$H_1 : m > m_0.$$

Dla poziomu istotności α mamy

$$P_{\theta_0}(\{U \geq u_{2\alpha}\}) = \alpha.$$

Obszar krytyczny $\{u : u \geq u_{2\alpha}\}$, $u_{2\alpha}$ wartość krytyczna, $\Phi(u_{2\alpha}) = 1 - \alpha$.

Przypadek II. II



Rysunek: Prawostronny obszar krytyczny testu istotności dla średniej populacji normalnej ze znanym σ .

Przypadek III. I

Testujemy hipotezę H_0 przeciwko hipotezie H_1 , gdzie

$$H_0 : m = m_0$$

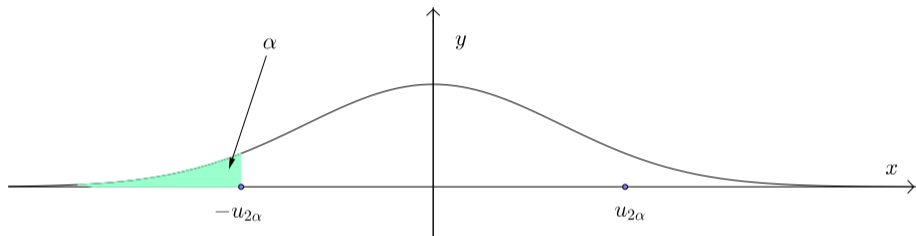
$$H_1 : m < m_0.$$

Dla poziomu istotności α mamy

$$P_{\theta_0}(\{U \leq -u_{2\alpha}\}) = \alpha.$$

Obszar krytyczny $\{u : u \leq -u_{2\alpha}\}$, $-u_{2\alpha}$ wartość krytyczna, $\Phi(u_{2\alpha}) = 1 - \alpha$.

Przypadek III. II



Rysunek: Lewostronny obszar krytyczny testu istotności dla średniej populacji normalnej ze znanym σ .

Przykłady. I

Przykład 1 ([1, Przykład 11.2])

Czas montowania elementu T w automatycznej pralce bębnekowej jest zmienną losową o rozkładzie normalnym. Norma techniczna przewiduje na tę czynność 6 minut. Natomiast wśród jej wykonawców panuje pogląd, że ten czas jest zbyt krótki. Zweryfikujemy tą hipotezę na poziomie istotności $\alpha = 0,05$, przy założeniu, że standardowe odchylenie czasu montowania wynosi $\sigma = 1,5$ minuty. Badanie przeprowadzono w grupie 25 robotników i ich średni czas montowania przez nich wynosił $\bar{X} = 6\frac{1}{3}$ minuty.

ROZWIĄZANIE.

Testujemy hipotezę H_0 przeciwko H_1 , gdzie

$$H_0 : \quad m = 6$$

$$H_1 : \quad m > 6.$$

Przykłady. II

$$U = \frac{\bar{X} - m_0}{\sigma} \sqrt{n} = \frac{6\frac{1}{3} - 6}{1,5} \sqrt{25} \approx 1,1.$$

Mamy w tym wypadku prawostronny obszar krytyczny

$$P(\{U \geq u_{0,1}\}) = 0,05,$$

gdzie $\Phi(u_{0,1}) = 1,65$, a więc nie ma podstaw do odrzucenia hipotezy zerowej. □

Założenia dla populacji normalnej z nieznanym odchyleniem dla małej próby

W przypadku nieznanego odchylenia standardowego i małej próby

- 1 Estymator – średnia z próby \bar{X} .
- 2 Hipoteza zerowa $H_0 : m = m_0$.
- 3 Sprawdzian hipotezy – statystyka $T = \frac{\bar{X} - m_0}{S} \sqrt{n - 1}$.
- 4 Jeżeli hipoteza zerowa jest prawdziwa, to \bar{X} ma rozkład t -Studenta o $n - 1$ stopniach swobody.
- 5 Poziom istotności α .
- 6 Parametrem opisującym rodzinę rozkładów jest $\theta \equiv m$. Będziemy też pisać θ_0 zamiast m_0 .

Przypadek I. I

Testujemy hipotezę H_0 przeciwko hipotezie H_1 , gdzie

$$H_0 : m = m_0$$

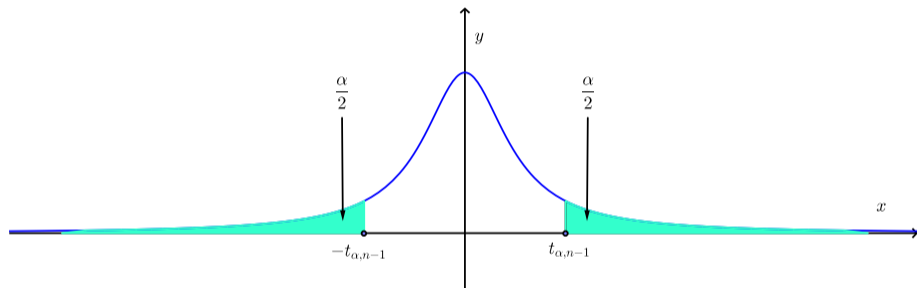
$$H_1 : m \neq m_0.$$

Dla poziomu istotności α mamy

$$P(\{|T| \geq t_\alpha\}) = \alpha.$$

Obszar krytyczny $\{t : |t| \geq t_\alpha\}$, t_α wartość krytyczna.

Przypadek I. II



Rysunek: Obustronny obszar krytyczny testu istotności dla średniej populacji normalnej z nieznanym σ .

Przypadek II.

Testujemy hipotezę H_0 przeciwko hipotezie H_1 , gdzie

$$H_0 : m = m_0$$

$$H_1 : m > m_0.$$

Dla poziomu istotności α mamy

$$P(\{T \geq t_{2\alpha}\}) = \alpha.$$

Obszar krytyczny $\{t : t \geq t_{2\alpha}\}$, $t_{2\alpha}$ wartość krytyczna.

Przypadek III.

Testujemy hipotezę H_0 przeciwko hipotezie H_1 , gdzie

$$H_0 : m = m_0$$

$$H_1 : m < m_0.$$

Dla poziomu istotności α mamy

$$P(\{T \leq -t_{2\alpha}\}) = \alpha.$$

Obszar krytyczny $\{t : t \leq -t_{2\alpha}\}$, $-t_{2\alpha}$ wartość krytyczna.

Przykład. I

Przykład 2 ([1, Przykład 11.3])

Plony żyta na powierzchniach uprawianych w pewnym województwie mają rozkład normalny o nieznanym parametrach. Przyjmując, że średni plon z tych powierzchni wynosi 28 kwintali. Sprawdzimy słuszność hipotezy, przy założeniu, że dla 20 powierzchni otrzymano średni plon 25 kwintali z odchyleniem standardowym 4,5 kwintala. Przyjmijemy poziom istotności $\alpha = 0,01$.

ROZWIĄZANIE.

Testujemy hipotezę H_0 przeciwko hipotezie H_1

$$H_0 : m = 28$$

$$H_1 : m \neq 28.$$

$$T = \frac{\bar{X} - m_0}{S} \sqrt{n-1} = \frac{25 - 28}{4,5} \sqrt{19} \approx -2,906.$$

Przykład. II

Mamy tutaj obustronny obszar krytyczny

$$P(\{|t| \geq t_{0,01}\}) = 0,01,$$

gdzie jest 19 stopni swobody i $t_{0,01;19} = 2,861$ oraz $|t| = 2,906 > t_{0,01;19}$. Wartość statystyki z próby znalazła się w obszarze krytycznym, więc odrzucamy hipotezę zerową na korzyść hipotezy alternatywnej. □

Założenia w populacji o dowolnym rozkładzie.

- 1 Hipoteza zerowa $H_0 : m = m_0$.
- 2 Średnia z próby \bar{X} ma asymptotyczny rozkład normalny $\mathcal{N}\left(m_0, \frac{\tilde{s}}{\sqrt{n}}\right)^2$.
- 3 Sprawdzianem hipotezy jest statystyka $U = \frac{\bar{X} - m_0}{s} \sqrt{n}$, która ma asymptotyczny rozkład normalny standardowy.
- 4 Poziom ufności α .

² \tilde{s} jest asymptotycznie zbliżone do odchylenia standardowego z populacji.

Testowanie hipotez

Testujemy hipotezę H_0 przeciwko hipotezie H_1 , gdzie

$$H_0 : m = m_0$$

$$H_1 : m \neq m_0$$

Dla poziomu istotności α mamy

$$P(\{|U| \geq u_\alpha\}) = \alpha.$$

Obszar krytyczny testu $\{u : |u| \geq u_\alpha\}$, u_α wartość krytyczna.

Uwaga 4

Podobnie, jak w przypadku, gdy znane jest odchylenie standardowe, można wyznaczyć obszary krytyczne dla dwóch kolejnych hipotez alternatywnych.

Ogólne założenia

- 1 Badamy dwie populacje normalne o rozkładach $\mathcal{N}(m_1, \sigma_1)$ i $\mathcal{N}(m_2, \sigma_2)$.
- 2 Z każdej populacji losujemy próbę losową o liczebnościach równych odpowiednio n_1 i n_2 .
- 3 Hipoteza zerowa $H_0 : m_1 = m_2$.
- 4 Estymatorem jest różnica średnich $\bar{X}_1 - \bar{X}_2$ o rozkładzie $\mathcal{N}\left(m_1 - m_2, \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right)$.
- 5 Poziom ufności α .
- 6 W zależności od założeń dotyczących zbiorowości generalnych oraz od liczebności prób – sprawdzian hipotezy H_0 ma różną postać i jest związany z rozkładem normalnym lub rozkładem t -Studenta.

Sprawdzian hipotezy. I

Jeśli

- σ_1, σ_2 - znane i $n_1 \leq 30, n_2 \leq 30$,
- σ_1, σ_2 - znane i $n_1 > 30, n_2 > 30$,
- σ_1, σ_2 - nieznane i $n_1 > 30, n_2 > 30$ to $\sigma_1^2 \approx s_{(1)}^2, \sigma_2^2 \approx s_{(2)}^2$,

wówczas sprawdzian hipotezy ma postać

$$U = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}},$$

który, przy założeniu prawdziwości H_0 , ma rozkład normalny standardowy.

W przypadku, gdy

Sprawdzian hipotezy. II

- σ_1, σ_2 - nieznane i $\sigma_1 = \sigma_2$, $n_1 \leq 30, n_2 \leq 30$,

to korzystamy ze sprawdzianu hipotezy

$$T = \frac{(\bar{X}_1 - \bar{X}_2)}{\sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

gdzie s_p^2 jest wariancją prób połączonych, który, przy założeniu prawdziwości H_0 , ma rozkład t -Studenta o $n_1 + n_2 - 2$ stopniach swobody. Ponadto wyraża się wzorem

$$s_p^2 = \frac{(n_1-1)\tilde{s}_1^2 + (n_2-1)\tilde{s}_2^2}{n_1+n_2-2} = \frac{n_1 s_1^2 + n_2 s_2^2}{n_1+n_2-2}.$$

Znane odchylenia standardowe w populacjach

Testujemy hipotezę H_0 przeciwko hipotezie H_1 , gdzie

$$H_0 : m_1 = m_2$$

$$H_1 : m_1 \neq m_2.$$

Przy założeniu prawdziwości hipotezy zerowej sprawdzianem testu jest statystyka

$$U = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}.$$

Dla poziomu ufności α mamy

$$P(\{|U| \geq u_\alpha\}) = \alpha.$$

Obszar krytyczny testu $\{u : |u| \geq u_\alpha\}$.

Odchylenia standardowe nieznane, ale równe

Testujemy hipotezę H_0 przeciwko hipotezie H_1 , gdzie

$$H_0 : m_1 = m_2$$

$$H_1 : m_1 \neq m_2.$$

Przy założeniu równości odchyłek standardowych $\sigma_1 = \sigma_2 = \sigma$ i hipotezy zerowej sprawdzianem testu jest statystyka

$$T = \frac{(\bar{X}_1 - \bar{X}_2)}{\sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}.$$

Mamy $P(\{|t| \geq t_\alpha\}) = \alpha$.

Ogólne założenia. I

Populacje dowolne (duże próby)

- 1 Z każdej populacji losujemy próbę losową o liczebnościach równych odpowiednio n_1 i n_2 .
- 2 Formułujemy hipotezę zerową i alternatywną

$$H_0 : m_1 = m_2$$

$$H_1 : m_1 \neq m_2,$$

i przyjmujemy poziom ufności α .

- 3 Estymatorem jest różnica średnich $\bar{X}_1 - \bar{X}_2$ o rozkładzie $\mathcal{N}\left(m_1 - m_2, \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right)$.
- 4 Wykorzystujemy statystykę $U = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$, która przy założeniu prawdziwości hipotezy zerowej, ma asymptotyczny rozkład normalny $\mathcal{N}(0,1)$.
- 5 Tak więc jest $P(\{|U| \geq u_\alpha\}) = \alpha$ i obszar krytyczny testu $\{u : |u| \geq u_\alpha\}$.

Przykład. I

Przykład 3 ([1, Przykład 11.4])

Przypuszcza się, że młodsze osoby łatwiej decydują się na zakup nowych, nieznanych produktów. Badanie przeprowadzone wśród przypadkowych 20 nabywców nowego produktu i 22 nabywców znanego już wyrobu pewnej firmy dostarczyło następujących informacji o wieku klientów

- kupujący nowy produkt: średnia 27,7; odchylenie 5,5,
- kupujący znany produkt: średnia 32,1; odchylenie 6,3.

Zweryfikujemy hipotezę, że średni wiek kupujących nowy produkt (m_1) jest równy średniemu wiekowi (m_2) kupujących znany produkt, przy poziome istotności $\alpha = 0,05$.

$$H_0 : m_1 = m_2$$

$$H_1 : m_1 < m_2.$$

Przykład. II

ROZWIĄZANIE.

Zakładamy, że rozkład wieku obu zbiorowości jest normalny i charakteryzuje się tym samym odchyleniem standardowym. Stosując statystykę $t = \frac{(\bar{X}_1 - \bar{X}_2)}{\sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$ otrzymujemy $s_p^2 = 35,206$,

$$t = \frac{27,7 - 32,1}{\sqrt{35,206 \left(\frac{1}{20} + \frac{1}{22} \right)}} = -2,4.$$

Lewostronny obszar krytyczny określa równość $P(\{t \leq -t_{2\alpha}\}) = \alpha$. W naszym wypadku $t_{2\alpha} = t_{0,1;40} = -1,684$.

Wartość statystyki z próby znalazła się w obszarze krytycznym, więc odrzucamy hipotezę zerową na korzyść hipotezy alternatywnej. □

Populacja normalna o nieznanym parametrach. I

Przyjmijmy, że parametr $\theta = (m, \sigma)$.

Należy zweryfikować hipotezę, że wariancja σ^2 w tej populacji ma ustaloną wartość σ_0^2 .

$$H_0 : \quad \sigma^2 = \sigma_0^2,$$

$$H_1 : \quad \sigma^2 > \sigma_0^2 \quad (\text{większe zróżnicowanie}).$$

Stosujemy statystykę $\chi^2 = \frac{nS^2}{\sigma^2}$, która ma rozkład $\chi^2(n-1)$. Obszar krytyczny, dla poziomu istotności α , wyznacza równość $P_{\theta}(\{\chi^2 \geq \chi_{\alpha, n-1}^2\}) = \alpha$.

Jeżeli wartość statystyki z próby przekroczy wartość krytyczną, to odrzucamy hipotezę zerową na korzyść alternatywnej. W przeciwnym wypadku nie ma podstaw do odrzucenia hipotezy zerowej.

Populacja normalna o nieznanym parametrach. II

Przykład 4 ([1, Przykład 11.6])

Chcemy sprawdzić czy odchylenie standardowe w rozkładzie czasu montowania elementu T w prasie automatycznej rzeczywiście wynosi $\sigma = 1,5$. Przyjąć poziom istotności $\alpha = 0,1$.

ROZWIĄZANIE.

$$H_0 : \sigma^2 = 2,25,$$

$$H_1 : \sigma^2 > 2,25.$$

W badanej grupie 25 robotników otrzymano wariancję z próby $s^2 = 2,8$. Obliczając statystykę otrzymujemy $\chi^2 = \frac{(25-1) \cdot 2,8}{2,25} = 29,87$. Otrzymujemy $\chi_{0,1;24}^2 = 33,196$. Tak więc wartość statystyki z próby znalazła się poza obszarem krytycznym, więc nie ma podstaw do odrzucenia hipotezy zerowej. □

Populacja normalna o nieznanym parametrach. III

Uwaga 5

- 1 Dla $n - 1 > 30$ obszar krytyczny testu wariancji należy budować na podstawie rozkładu normalnego.
- 2 Zamiast statystyki $\chi^2 = \frac{nS^2}{\sigma^2}$ można stosować statystykę $\chi^2 = \frac{(n-1)\tilde{S}^2}{\sigma^2}$.

Populacje normalne $\mathcal{N}(m_1, \sigma_1)$ i $\mathcal{N}(m_2, \sigma_2)$. I

Nieznane są wartości oczekiwane i odchylenia standardowe.

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

$$H_1 : \sigma_1^2 \neq \sigma_2^2.$$

Alternatywnie hipotezę zerową i alternatywą można sformułować:

$$H_0 : \frac{\sigma_1^2}{\sigma_2^2} = 1$$

$$H_1 : \frac{\sigma_1^2}{\sigma_2^2} \neq 1.$$

Populacje normalne $\mathcal{N}(m_1, \sigma_1)$ i $\mathcal{N}(m_2, \sigma_2)$. II

Do weryfikacji hipotezy zerowej używamy wariancji $\tilde{\xi}_1^2$ i $\tilde{\xi}_2^2$, obliczanych z dwóch niezależnych prób, o liczebnościach równych odpowiednio n_1 i n_2 . Stosujemy statystykę $F = \frac{\frac{\tilde{\xi}_1^2}{\sigma_1^2}}{\frac{\tilde{\xi}_2^2}{\sigma_2^2}}$ o rozkładzie $F(n_1 - 1, n_2 - 1)$.

Przy założeniu prawdziwości hipotezy zerowej mamy $F = \frac{\tilde{\xi}_1^2}{\tilde{\xi}_2^2}$. Prawostronna część obszaru krytycznego opisana jest zależnością

$$P(\{F \geq F_{\frac{\alpha}{2}}\}) = \frac{\alpha}{2}.$$

Natomiast lewostronna część obszaru krytycznego ma postać

$$P(\{F \leq F_{\frac{\alpha}{2}}\}) = \frac{\alpha}{2}.$$

Populacje normalne $\mathcal{N}(m_1, \sigma_1)$ i $\mathcal{N}(m_2, \sigma_2)$. III

Uwaga 6

W tablicach jest tylko $F_{\frac{\alpha}{2}}$, zatem w statystyce F w liczniku umieszczamy większą z wariancji obu prób. Obliczoną tak wartość F porównujemy z $F_{\frac{\alpha}{2}}$. Jeżeli jest spełniona nierówność $F \geq F_{\frac{\alpha}{2}}$, to odrzucamy hipotezę zerową.

Jeżeli

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

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

to

$$F = \frac{\tilde{S}_1^2}{\tilde{S}_2^2}.$$

Populacje normalne $\mathcal{N}(m_1, \sigma_1)$ i $\mathcal{N}(m_2, \sigma_2)$. IV

Natomiast jeżeli

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

$$H_1 : \sigma_1^2 < \sigma_2^2,$$

to

$$F = \frac{\tilde{S}_2^2}{\tilde{S}_1^2}.$$

W obu przypadkach mamy prawostronny obszar krytyczny $P(\{F \geq F_\alpha\}) = \alpha$.

Przykład 5 ([1, Przykład 11.7])

Sprawdzić, czy słuszne jest założenie o równości odchyłeń standardowych wieku w populacji kupujących wyroby nowe i znane, dla poziomu istotności $\alpha = 0,05$.

Populacje normalne $\mathcal{N}(m_1, \sigma_1)$ i $\mathcal{N}(m_2, \sigma_2)$. V

ROZWIĄZANIE. Wtedy $\frac{\alpha}{2} = 0,025$. Ponadto mamy

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

$$H_1 : \sigma_1^2 \neq \sigma_2^2.$$

Ponadto mamy

$$F = \frac{\tilde{s}_2^2}{\tilde{s}_1^2} = \frac{39,69}{30,25} = 1,31 \quad \text{oraz} \quad F_{0,025;21;19} = 2,49.$$

Tak więc brak jest podstaw do odrzucenia hipotezy zerowej ($1,31 < 2,49$). □

Spis treści

1 Parametryczne testy istotności

- Ogólne pojęcie testów istotności
- Testy istotności dla wartości średniej
 - Populacja normalna ze znanym odchyleniem standardowym
 - Populacja normalna z nieznanym odchyleniem standardowym (mała próba)
 - Populacja generalna o dowolnym rozkładzie
- Testy istotności dla wartości dwóch średnich
 - Populacje normalne
 - Populacje dowolne (duże próby)
- Testy istotności dla wariancji
- Testy istotności dla dwóch wariancji

2 Nieparametryczne testy istotności

Klasyfikacja nieparametrycznych testów istotności

Nieparametryczne testy istotności dzielimy na

- 1 testy losowości – weryfikują hipotezę, że próba ma charakter losowy,
- 2 testy niezależności – sprawdzają hipotezę o niezależności dwóch zmiennych losowych,
- 3 testy zgodności – weryfikują hipotezę o postaci funkcyjnej rozkładu populacji generalnej,
- 4 inne.

Test zgodności sprawdza zgodność rozkładu empirycznego z próby z rozkładem hipotetycznym lub też zgodność dwóch lub więcej rozkładów empiryczny z próby.

Test zgodności chi-kwadrat. I

Najczęściej stosowanymi nieparametrycznymi testami są testy zgodności chi-kwadrat.

Test ten zbudowany jest na podstawie statystyki χ^2 . Mamy

$$H_0 : F = F_0 \quad (\text{populacja generalna ma rozkład określony} \\ \text{pewną dystrybuantą } F_0),$$

$$H_1 : F \neq F_0.$$

Zasady przeprowadzania testu chi-kwadrat

- 1 Losujemy z populacji dużą próbę (będziemy wykorzystywać rozkład graniczny statystyki).
- 2 Budujemy szereg rozdzielczy – tworzymy r rozłącznych klas wartości badanej cechy (zmiennej) w próbie (liczebność i -tej klasy wynosi n_i).
- 3 Zakładamy prawdziwość hipotezy zerowej.

Test zgodności chi-kwadrat. II

- Obliczamy prawdopodobieństwa p_i tego, że badana cecha przyjmie wartości z i -tej klasy. Przy założeniu prawdziwości hipotezy zerowej liczebności poszczególnych klas powinny wynosić np_i dla $i \in \overline{1, r}$, gdzie n liczebność próby.

Uwagi

- Podstawą konstrukcji miary zgodności rozkładu empirycznego z hipotetycznym jest różnica między liczebnościami zaobserwowanymi n_i , a liczebnościami hipotetycznymi np_i .
- Do oceny zgodności stosujemy statystykę

$$\chi^2 = \sum_{i=1}^r \frac{(n_i - np_i)^2}{np_i},$$

która przy założeniu prawdziwości hipotezy zerowej ma asymptotyczny rozkład $\chi^2(r - k - 1)$, gdzie k jest liczbą parametrów rozkładu, które zostały oszacowane na podstawie próby metodą największej wiarygodności.

Test zgodności chi-kwadrat. III

- 3 Jeżeli prawdziwa jest hipoteza zerowa, to dla poziomu istotności α mamy

$$P(\{\chi^2 \geq \chi_\alpha^2\}) = \alpha.$$

- 4 Jeżeli wartość statystyki z próby jest nie mniejsza niż wartość krytyczna (odpowiada to temu, że różnica między rozkładem empirycznym, a hipotetycznym jest statystycznie istotna), to hipotezę zerową odrzucamy.

Test zgodności chi-kwadrat. IV

Przykład 6

[1, Przykład 11.10] Przez 300 dni obserwowano pracę pewnej maszyny, rejestrując liczbę uszkodzeń w ciągu dnia. Otrzymano następujące dane

liczba uszkodzeń (x_i)	liczba dni (n_i)
0	140
1	110
2	30
3	10
4	10

Zweryfikujemy hipotezę, że liczba uszkodzeń ma rozkład Poissona.

ROZWIĄZANIE. Określamy wartość parametru λ dla rozkładu Poissona.

Test zgodności chi-kwadrat. V

Estymatorem parametru λ otrzymanym MNW³ jest średnia arytmetyczna z próby \bar{X} .

W naszym przypadku wynosi ona $\bar{x} = 0,8$. Mamy

$$P(\{X = k\}) = \frac{(0,8)^k}{k!} e^{-0,8}, \text{ dla } k \in \overline{0, 4}.$$

Konstrukcja tablicy roboczej

i	x_i	n_i	p_i	np_i	$(n_i - np_i)^2$	$\frac{(n_i - np_i)^2}{np_i}$
1	0	140				
2	1	110				
3	2	30				
4	3	10				
5	≥ 4	10				
Σ		300				

Test zgodności chi-kwadrat. VI

i	x_i	n_i	p_i	np_i	$(n_i - np_i)^2$	$\frac{(n_i - np_i)^2}{np_i}$
1	0	140	0,449	134,7	28,09	0,209
2	1	110	0,359	107,7	5,29	0,049
3	2	30	0,144	43,2	174,24	4,033
4	3	10	0,038	11,4	1,96	0,1719
5	≥ 4	10	0,010			
Σ		300	1			

Test zgodności chi-kwadrat. VII

i	x_i	n_i	p_i	np_i	$(n_i - np_i)^2$	$\frac{(n_i - np_i)^2}{np_i}$
1	0	140	0,449	134,7	28,09	0,209
2	1	110	0,359	107,7	5,29	0,049
3	2	30	0,144	43,2	174,24	4,033
4	3	10	0,038	11,4	1,96	0,1719
5	≥ 4	10	0,010	3,0		
Σ		300	1			

Test zgodności chi-kwadrat. VIII

Z rozkładu asymptotycznego musi być tak, że $np_i \geq 5$, więc łączymy dwie ostatnie klasy tworząc nową tablicę roboczą.

i	x_i	n_i	p_i	np_i	$(n_i - np_i)^2$	$\frac{(n_i - np_i)^2}{np_i}$
1	0	140	0,449	134,7	28,09	0,209
2	1	110	0,359	107,7	5,29	0,049
3	2	30	0,144	43,2	174,24	4,033
4	≥ 3	20	0,048	14,4	31,36	2,178
Σ		300	1	300		$\chi^2 = 6,469$

Przyjmując poziom istotności $\alpha = 0,05$ i mając $r - k - 1 = 4 - 1 - 1 = 2$ więc $\chi_{0,05;2}^2 = 5,991$. Na podstawie danych odrzucamy hipotezę zerową.

Test zgodności chi-kwadrat. IX

Uwaga 7

Gdyby dany był parametr λ , to wartość krytyczną należałoby określić dla $r - 1$ stopni swobody.

³Metoda największej wiarygodności.

Test zgodności λ -Kolmogorowa. I

Uwaga 8

Test ten może być stosowany tylko w przypadku ciągłej dystrybuanty teoretycznej.

Definicja 3 (Przypomnienie)

Dystrybuantą empiryczną F_n nazywamy funkcję określoną na podstawie danych (x_i, ω_i) dla $i \in \overline{1, k}$ wzorem

$$F_n(x) = \sum_{i: x_i \leq x} \omega_i, \quad (2)$$

gdzie $\omega_i = \frac{n_i}{n}$ są wagami x_i .

Zasady/etapy konstrukcji testu λ -Kolmogorowa

Test zgodności λ -Kolmogorowa. II

- 1 Formułujemy hipotezę zerową i alternatywną

$$H_0 : F = F_0$$

$$H_1 : F \neq F_0.$$

- 2 Losujemy dużą próbę n elementową oraz budujemy szereg przedziałowy z dużą ilością wąskich klas.
- 3 Wyznaczamy dystrybuantę empiryczną.

Test zgodności λ -Kolmogorowa. III

- Przy założeniu prawdziwości hipotezy zerowej różnice między dystrybuanta empiryczna, a hipotetyczną nie powinny być duże.

Miarą zgodności obu dystrybuant jest statystyka $\lambda = D\sqrt{n}$, gdzie

$$D = \sup_x |F_n(x) - F_0(x)|.$$

Przy założeniu prawdziwości hipotezy zerowej statystyka λ ma asymptotyczny rozkład λ -Kolmogorowa z obszarem krytycznym wyznaczonym przez równość $P(\{\lambda \geq \lambda_\alpha\}) = \alpha$ z poziomem istotności α .

Jeżeli na podstawie próby otrzymamy wartość nie mniejszą niż wartość krytyczna ($\lambda \geq \lambda_\alpha$) to hipotezę zerową odrzucamy.

Test zgodności λ -Kolmogorowa. IV

Przykład 7 ([1, Przykład 11.11])

Przypuśćmy, że jednostkowe koszty produkcji danego wyrobu mają rozkład normalny. W celu weryfikacji tego przypuszczenia zbadano próbę 200 zakładów produkujących ten wyrób otrzymując następujące dane

i	koszty jednostkowe	liczba zakładów (n_i)
1	2,50 – 3,50	5
2	3,50 – 4,50	10
3	4,50 – 5,50	35
4	5,50 – 6,50	80
5	6,50 – 7,50	50
6	7,50 – 8,50	10
7	8,50 – 9,50	10

Test zgodności λ -Kolmogorowa. V

Zweryfikujemy hipotezę, że rozkład jednostkowego kosztu produkcji tego wyrobu ma rozkład normalny.

ROZWIĄZANIE. Standaryzujemy górne granice klas, gdyż potrzebujemy m i σ , które oszacujemy na podstawie próby $m = \bar{x} = 6,15$ i $\sigma = s = 1,216$.

Etapy obliczania statystyki D (budowanie pomocniczej tablicy).

Do konstrukcji potrzebna będzie nam standaryzowana zmienna losowa.

Definicja 4

Wartość standaryzowana odpowiadająca obserwacji x wyraża się wzorem

$$u = \frac{x - \bar{x}}{s}. \quad (3)$$

Test zgodności λ -Kolmogorowa. VI

Budowania pomocniczej tablicy

i	x	u	$\Phi(u) = F_0(x)$	ω_i	$F_n(x)$	$ F_n(x) - F_0(x) $
1						
2						
3						
4						
5						
6						
7						
Σ						

Test zgodności λ -Kolmogorowa. VII

i	x	u	$\Phi(u) = F_0(x)$	ω_i	$F_n(x)$	$ F_n(x) - F_0(x) $
1	3,50	-2,18	0,015	0,025	0,025	0,010
2	4,50					
3	5,50					
4	6,50					
5	7,50					
6	8,50					
7	9,50					
Σ						

Test zgodności λ -Kolmogorowa. VIII

i	x	u	$\Phi(u) = F_0(x)$	ω_i	$F_n(x)$	$ F_n(x) - F_0(x) $
1	3,50	-2,18	0,015	0,025	0,025	0,010
2	4,50	-1,36	0,087	0,050	0,075	0,012
3	5,50	-0,54	0,295	0,175	0,250	$D = 0,045$
4	6,50	0,29	0,614	0,400	0,650	0,036
5	7,50	1,11	0,867	0,250	0,900	0,033
6	8,50	1,93	0,973	0,050	0,950	0,023
7	9,50	2,76	0,997	0,050	1,000	0,003
Σ				1		

Mamy $D = 0,045$ oraz $\lambda = 0,045 \cdot \sqrt{200} = 0,637$. Dla poziomu istotności $\alpha = 0,05$ mamy $\lambda_\alpha = 1,36$. Tak więc nie ma podstaw do odrzucenia hipotezy zerowej. \square

Test zgodności λ -Kołmogorowa. IX

Uwaga 9

Tablice rozkładu λ -Kołmogorowa podają wartość dystrybuant $K(\lambda)$. λ_α otrzymujemy z warunku $K(\lambda_\alpha) = 1 - \alpha$.

Test zgodności Kołmogorowa-Smirnowa. I

Test zgodności Kołmogorowa-Smirnowa służy do weryfikacji hipotezy, że dwie populacje mają jednakowy rozkład (dwie próby pochodzą z tej samej populacji).

Niech dwie populacje mają rozkłady opisane dystrybuantami F_1 i F_2 .

Zasady konstrukcji testu Kołmogorowa-Smirnowa

- 1 Formułujemy hipotezę zerową i alternatywną

$$H_0 : F_1 = F_2,$$

$$H_1 : F_1 \neq F_2.$$

- 2 Losujemy duże próby n_1 elementową z populacji pierwszej i n_2 elementową z populacji z populacji drugiej.
- 3 Wyznaczamy dystrybuanty empiryczne F_{n_1} i F_{n_2} .

Test zgodności Kołmogorowa-Smirnowa. II

- 4 Przy założeniu prawdziwości hipotezy zerowej różnica między dystrybuantami nie powinny być duże. Miarą zgodności obu dystrybuant jest statystyka $\lambda = D^* \sqrt{n}$, gdzie $D^* = \sup_x |F_{n_1}(x) - F_{n_2}(x)|$, a $n = \frac{n_1 n_2}{n_1 + n_2}$.
- 5 Przy założeniu prawdziwości hipotezy zerowej statystyka λ ma asymptotyczny rozkład λ -Kołmogorowa z obszarem krytycznym wyznaczonym przez równość $P(\{\lambda \geq \lambda_\alpha\}) = \alpha$ z poziomem istotności α . Jeżeli na podstawie próby otrzymamy wartość nie mniejszą niż wartość krytyczna λ_α , to hipotezę zerową odrzucamy.

Test Craméra-von Misesa. I

Test Craméra-von Misesa – test zgodności rozkładu z zadany rozkładem wzorcowym lub drugą próbą. Zwykle stosuje się go do sprawdzenia zgodności z rozkładem normalnym. Został zaproponowany przez Harald Craméra i Richarda von Mises w latach 1928-1930.

Statystyka Craméra-von Misesa

1 Wersja dla jednej próby i rozkładu wzorcowego (teoretycznego).

$$W^2 = n \int_{-\infty}^{+\infty} (F_n(x) - F(x))^2 dF(x),$$

gdzie F_n to dystrybuanta empiryczna, F dystrybuanta rozkładu teoretycznego, a n to liczebność próby.

Test Craméra-von Misesa. II

Zwykle do obliczeń używany jest wzór

$$W^2 = \frac{1}{12n} + \sum_{i=1}^n \left(F(X_{i:n}) - \frac{2i-1}{2n} \right)^2,$$

gdzie $X_{i:n}$ to i -ta statystyka pozycyjna, F i n mają te same znaczenie, jak poprzednio.

- 2 **Wersja dla dwóch prób.** Niech x_1, x_2, \dots, x_n i y_1, y_2, \dots, y_m będą obserwowanymi wartościami w pierwszej i drugiej próbie posortowane rosnąco. Niech r_1, r_2, \dots, r_n będą rangami obserwacji x_i w połączonej próbie (X i Y rangowane razem) i niech s_1, s_2, \dots, s_m będą rangami obserwacji y_i w połączonej próbie. Wówczas

$$W^2 = \frac{U}{nm(n+m)} - \frac{4mn-1}{6(m+n)},$$

Test Craméra-von Misesa. III

gdzie

$$U = n \sum_{i=1}^n (r_i - i)^2 + m \sum_{j=1}^m (s_j - j)^2.$$

Test Andersona-Darlinga. I

Test Andersona-Darlinga – test zgodności rozkładu z zadaniem rozkładem teoretycznym (wzorcowym). Zwykle stosuje się go do sprawdzenia zgodności z rozkładem normalnym. Jest modyfikacją testu Craméra-von Misesa dokonaną w celu poprawy jego czułości w „ogonach” testowanego rozkładu.

Statystyka Andersona-Darlinga

$$A^2 = n \int_{-\infty}^{+\infty} \frac{(F_n(x) - F(x))^2}{F(x)(1 - F(x))} dF(x),$$

gdzie F_n to dystrybuanta empiryczna, F to dystrybuanta rozkładu wzorcowego, a n to liczebność próby.

Jest to wersja testu Craméra-von Misesa ważona czynnikiem $\frac{1}{F(x)(1-F(x))}$.

Test Andersona-Darlinga. II

Zwykle do obliczeń używany jest prostszy wzór

$$A^2 = -n - \frac{1}{n} \sum_{i=1}^n ((2i-1) \ln F(X_{i:n}) + (2n+1-2i) \ln(1-F(X_{i:n})))$$

lub (inna wersja)

$$A^2 = -n - \sum_{i=1}^n \frac{2i-1}{n} \left[\ln F(X_{i:n}) + \ln \left(1 - F(X_{(n+1-i):n}) \right) \right],$$

gdzie $X_{i:n}$ to i -ta statystyka pozycyjna, F i n mają te same znaczenie, jak poprzednio.

Dla rozkładu normalnego stosuje się czasem poprawkę na wielkość próby

$$A^{2*} = A^2 \left(1 + \frac{0,75}{n} + \frac{2,25}{n^2} \right).$$

Test Andersona-Darlinga. III

Dla rozkładu normalnego, gdy A^{2*} przekracza 0.752 to hipoteza o normalności rozkładu w populacji jest odrzucana na poziomie 5%. Dla innych rozkładów test także może być stosowany, ale ma inne wartości krytyczne.

Bibliografia

- [1] J. Józwiak i J. Podgórski. *Statystyka od podstaw*. Wyd. 5 zmienione. Warszawa: PWE, 2000.