Statystyka matematyczna - wykład siódmy[1] Testowanie hipotez – część II. kierunek: matematyka I° specjalność: matematyka finansowa

dr Jarosław Kotowicz

Uniwersytet w Białymstoku

Spis treści

Populacja normalna o nieznanych parametrach. I

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

$$\begin{aligned} H_0: &\quad \sigma^2 = \sigma_0^2, \\ H_1: &\quad \sigma^2 > \sigma_0^2 \quad \text{(większe zróżnicowanie).} \end{aligned}$$

Stosujemy statystykę $\chi^2 = \frac{nS^2}{\sigma^2}$, która ma rozkład $\chi^2(n-1)$. Obszar krytyczny, dla poziomu istotności α, wyznacza równość
$P_{(m,\sigma^2)}(\{\chi^2 \geqslant \chi^2_{\alpha,n-1}\}) = \alpha$.
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 nieznanych parametrach. II

Przykład 1 ([1, Przykład 11.6])

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

ROZWIĄZANIE.

$$\begin{aligned} H_0 &: \quad \sigma^2 = 2{,}25, \\ H_1 &: \quad \sigma^2 > 2{,}25. \end{aligned}$$

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 $\chi^2_{0{,}1;24} = 33{,}196$. Tak więc wartość statystyki z próby znalazła się poza obszarem krytycznym, więc nie ma podstaw do odrzucenia hipotezy zerowej. □

Populacja normalna o nieznanych parametrach. III

Uwaga 1

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.

$$\begin{aligned} H_0 : &\quad \sigma_1^2 = \sigma_2^2, \\ H_1 : &\quad \sigma_1^2 \neq \sigma_2^2. \end{aligned}$$

Alternatywnie hipotezę zerową i alternatywą można sformułować:

$$\begin{aligned} H_0 : &\quad \frac{\sigma_1^2}{\sigma_2^2} = 1, \\ H_1 : &\quad \frac{\sigma_1^2}{\sigma_2^2} \neq 1. \end{aligned}$$

Do weryfikacji hipotezy zerowej używamy wariancji $\tilde{s}_1^2$ i $\tilde{s}_2^2$, obliczanych z dwóch niezależnych prób, o liczebnościach równych odpowiednio n_1 i n_2.

Stosujemy statystykę $F = \frac{\frac{\tilde{s}_1^2}{\sigma_1^2}}{\frac{\tilde{s}_2^2}{\sigma_2^2}}$ o rozkładzie $F(n_1 - 1, n_2 - 1)$.

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

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

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

Natomiast lewostronna część obszaru krytycznego ma postać

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

Uwaga 2

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* $F \geqslant F_{\frac{\alpha}{2}}$*, to odrzucamy hipotezę zerową.*

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

Jeżeli

$$\begin{aligned} H_0 : &\quad \sigma_1^2 = \sigma_2^2, \\ H_1 : &\quad \sigma_1^2 > \sigma_2^2, \end{aligned}$$

to

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

Natomiast jeżeli

$$\begin{aligned} H_0 : &\quad \sigma_1^2 = \sigma_2^2, \\ H_1 : &\quad \sigma_1^2 < \sigma_2^2, \end{aligned}$$

to

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

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

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

Przykład 2 ([1, Przykład 11.7])

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

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

$$\begin{aligned} H_0 : &\quad \sigma_1^2 = \sigma_2^2, \\ H_1 : &\quad \sigma_1^2 \neq \sigma_2^2. \end{aligned}$$

Ponadto mamy

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

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

Tak wiec brak jest podstaw do odrzucenia hipotezy zerowej $(1{,}31 < 2{,}49)$.
$\square$

Spis treści

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

$$\begin{aligned} H_0 : F = F_0 \qquad & \text{(populacja generalna ma rozkład określony} \\ & \text{pewną dystrybuantą } F_0\text{)}, \\ H_1 : F \neq F_0. & \end{aligned}$$

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

4. 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 $i = 1, \ldots r$, gdzie n liczebność próby.

Uwagi

1. Podstawą konstrukcji miary zgodności rozkładu empirycznego z hipotetycznym jest różnica miedzy liczebnościami zaobserwowanymi n_i, a liczebnościami hipotetycznymi np_i.
2. 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

Test zgodności chi-kwadrat. III

rozkładu, które zostały oszacowane na podstawie próby metodą największej wiarygodności.

3. Jeżeli prawdziwa jest hipoteza zerowa, to dla poziomu istotności α mamy
$$P(\{\chi^2 \geqslant \chi^2_\alpha\}) = \alpha.$$
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 3

[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. Estymatorem λ otrzymanym MNW[2] jest średnia arytmetyczna z próby $\overline{X}$. W naszym przypadku wynosi ona $\overline{x} = 0{,}8$.

Test zgodności chi-kwadrat. V

Mamy

$$P(\{X = k\}) = \frac{(0{,}8)^k}{k!} e^{-0{,}8}, \text{ dla } k = 0, 1, 2, 3, 4.$$

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	$\geqslant 4$	10				
$\sum$		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	$\geqslant 4$	10	0,010			
$\sum$		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	$\geqslant 4$	10	0,010	3,0		
$\sum$		300	1			

Test zgodności chi-kwadrat. VIII

Z rozkładu asymptotycznego musi być tak, że $np_i \geqslant 5$, 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	$\geqslant 3$	20	0,048	14,4	31,36	2,178
$\sum$		300	1	300		$\chi^2 = 6{,}469$

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

Test zgodności chi-kwadrat. IX

Uwaga 3

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

[2] Metoda największej wiarygodności.

Test zgodności λ-Kołmogorowa. I

Uwaga 4

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

Definicja 1 (Przypomnienie)

Dystrybuantą empiryczną F_n nazywamy funkcję określoną na podstawie danych (x_i, ω_i) dla $i = 1, \ldots, k$ wzorem

$$F_n(x) = \sum_{i: x_i \leqslant x} \omega_i, \tag{1}$$

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

Zasady/etapy konstrukcji testu λ-Kołmogorowa

Test zgodności λ-Kołmogorowa. II

1. Formułujemy hipotezę zerową i alternatywną
$$\begin{aligned} H_0 : \quad & F = F_0, \\ H_1 : \quad & F \neq F_0. \end{aligned}$$
2. Losujemy dużą próbę n elementową oraz budujemy szereg przedziałowy z dużą ilością wąskich klas.
3. Wyznaczamy dystrybuantę empiryczną.
4. 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 $D = \sup\limits_{x} |F_n(x) - F_0(x)|$.
   Przy założeniu prawdziwości hipotezy zerowej statystyka λ ma asymptotyczny rozkład λ - Kołmogorowa z obszarem krytycznym

Test zgodności λ-Kołmogorowa. III

wyznaczonym przez równość $P(\{\lambda \geqslant \lambda_\alpha\}) = \alpha$ z poziomem istotności α.

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

Test zgodności λ-Kołmogorowa. IV

Przykład 4 ([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-Kołmogorowa. 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 $m = \overline{x} = 6,15$ i $\sigma = s = 1,216$.
Etapy obliczania statystyki D (budowanie pomocniczej tablicy).
Do konstrukcji potrzebna będzie nam standaryzowana zmienna losowa.

Definicja 2

Wartość standaryzowana odpowiadająca obserwacji x wyraża się wzorem

$$u = \frac{x - \overline{x}}{s}. \tag{2}$$

Test zgodności λ-Kołmogorowa. VI

Budowania pomocniczej tablicy

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

Test zgodności λ-Kołmogorowa. VII

i	x	u	$\Phi(u) = F_0(x)$	ω_i	$F_n(x)$	$\lvert F_n(x) - F_0(x)\rvert$
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-Kołmogorowa. VIII

i	x	u	$\Phi(u) = F_0(x)$	ω_i	$F_n(x)$	$\lvert F_n(x) - F_0(x)\rvert$
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	$\lambda = 0{,}045$
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. □

Test zgodności λ-Kołmogorowa. IX

Uwaga 5

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

$$\begin{aligned} H_0 : &\qquad F_1 = F_2, \\ H_1 : &\qquad F_1 \neq F_2. \end{aligned}$$

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 $\lambda = D^\star\sqrt{n}$, gdzie $D^\star = \sup\limits_x |F_{n_1}(x) - F_{n_2}(x)|$, 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 $P(\{\lambda \geqslant \lambda_\alpha\}) = \alpha$ 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 zadanym rozkładem wzorcowym lub drugą próbą. Zwykle stosuje się go do sprawdzenia zgodności z rozkładem normalnym. Został zaproponowany przez Haralda Craméra i Richarda von Mises w latach 1928-1930.

Statystyka Craméra-von Misesa

1. **Wersja dla jednej próby i rozkładu wzorcowego.**

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

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

Test Craméra-von Misesa. II

Zwykle do obliczeń używany jest prostszy wzór:

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

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

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

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

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

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

Test Andersona-Darlinga. III

Dla rozkładu normalnego, gdy $A^{2\star}$ 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

J. Jóźwiak i J. Podgórski. *Statystyka od podstaw*. Wyd. 5 zmienione. Warszawa: PWE, 2000.