# STATYSTYKA MATEMATYCZNA

WYKŁAD 5

# WERYFIKACJA HIPOTEZ NIEPARAMETRYCZNYCH

# Test zgodności $\chi^2$

Hipoteza zerowa $H_0$ (Cecha X populacji ma rozkład o dystrybuancie F).

Hipoteza alternatywna $H_1$ (Cecha $X$ populacji nie ma rozkładu o dystrybuancie F).

Weryfikacja powyższych hipotez za pomocą tzw. testu $\chi^2$ przebiega następująco:

1. Pobieramy liczną próbę ($n$ >80). Prezentujemy ją w szeregu rozdzielczym klasowym w $r$ klasach.
2. Obliczamy na podstawie próby estymatory największej wiarygodności nieznanych parametrów.
3. Przyjmujemy, że cecha $X$ ma rozkład o dystrybuancie F.
4. Dla każdego przedziału klasowego $A_i =< a_i ; a_{i+1})$ obliczamy prawdopodobieństwo

$$p_i = P(X \in A_i) = P(a_i \leq X < a_{i+1}) = F(a_{i+1}) - F(a_i)$$

5. Obliczamy

$$u_n = \sum_{i=1}^{r} \frac{(n_i - np_i)^2}{np_i} = \sum_{i=1}^{r} \frac{(n_i - \hat{n}_i)^2}{\hat{n}_i}$$

gdzie $n_i$ jest liczebnością (empiryczną) klasy $A_i$.

$\hat{n}_i = np_i$ jest liczebnością teoretyczną klasy $A_i$

6. Wyznaczamy zbiór krytyczny prawostronny

$K =< k\ ; \infty)$, gdzie $k$ wyznaczamy z tablicy rozkładu $\chi^2$ dla $r - l - 1$ stopniami swobody

gdzie $l$ – liczba nieznanych parametrów rozkładu X,

i dla prawdopodobieństwa $\alpha$ (równemu poziomowi istotności).

7. Podejmujemy decyzję:

odrzucamy hipotezę $H_0$, gdy $u_n \in K$

przyjmujemy hipotezę $H_0$, gdy $u_n \notin K$

**Uwaga**. Do obliczania prawdopodobieństw $p_i$, pierwsza i ostatnia klasa szeregu rozdzielczego powinny mieć postać $A_1 = (-\infty ; a_2)$, $A_r =< a_r ; \infty)$ i do każdej z nich powinno należeć co najmniej 5 elementów próby. Do pozostałych klas powinno należeć co najmniej 10 elementów próby. Klas nie może być mniej niż 4.

**Przykład.**
Badano rozkład liczby awarii systemu komputerowego (cecha X populacji). W ciągu 100 tygodni zarejestrowano następujące ilości awarii:

| Liczba awarii | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| Liczba tygodni | 24 | 32 | 23 | 12 | 9 |

Na poziomie istotności $\alpha = 0{,}05$ sprawdź czy liczba awarii ma rozkład Poissona.
hipotezy:

$H_0$(Cecha X populacji ma rozkład Poissona)
$H_1$(Cecha $X$ populacji nie ma rozkładu Poissona).

| $i$ | $n_i$ | $i \cdot n_i$ | $p_i$ | $n\,p_i$ | $\frac{(n_i - np_1)^2}{np_i}$ |
|---|---|---|---|---|---|
| 0 | 24 | 0 | 0,223 | 22,3 | 0,13 |
| 1 | 32 | 32 | 0,33 | 33 | 0,06 |
| 2 | 23 | 46 | 0,251 | 25,1 | 0,18 |
| 3 | 12 | 36 | 0,13 | 13 | 0,02 |
| 4 | 9 | 36 | 0,066 | 6,6 | 0,9 |
| | | 150 | 1,00000 | 100 | 1,29 |

Przyjmujemy $\lambda \approx 1{,}5$ $u_{100} = 1{,}29$.
Wyznaczamy zbiór krytyczny prawostronny

$$K = < k; \infty ).$$

Liczbę *k* odczytujemy z tablicy rozkładu $\chi^2$ dla

r – 1 -1 = 5 – 2 = 3 stopni swobody i

prawdopodobieństwa $\alpha = 0{,}05$.

Mamy $k = 7{,}815$, więc

$$K = < 7{,}815 ; \infty ).$$

Ponieważ $u_{100} = 1{,}29 \notin K$,

więc hipotezę, że cecha ma rozkład Poissona

przyjmujemy.

**Test normalności (test Shapiro-Wilka)**
Wysuwamy dwie hipotezy:
$H_0$ – $X$ ma rozkład normalny,
$H_1$ – $X$ nie ma rozkładu normalnego.

Dane statystyczne porządkujemy niemalejąco: $x_{(1)}, x_{(2)}, \ldots, x_{(n)}$
Stosujemy statystykę

$$U_n = \frac{\left[\sum_{i=1}^{[n/2]} a_{n,i}\left(x_{(n-i+1)} - x_{(i)}\right)\right]^2}{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2}$$

gdzie $[n/2]$ jest częścią całkowitą liczby $n/2$,
$a_{n,i}$ – współczynniki Shapiro-Wilka odczytane z tablicy:

| $n$ \ $i$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| 8 | 0,6052 | 0,3164 | 0,1743 | 0,0561 | — | — | — | — | — | — |
| 10 | 0,5739 | 0,3291 | 0,2141 | 0,1224 | 0,0399 | — | — | — | — | — |
| 12 | 0,5475 | 0,3325 | 0,2347 | 0,1586 | 0,0922 | 0,0303 | — | — | — | — |
| 14 | 0,5251 | 0,3318 | 0,2460 | 0,1802 | 0,1240 | 0,0727 | 0,0240 | — | — | — |
| 15 | 0,5150 | 0,3306 | 0,2495 | 0,1878 | 0,1353 | 0,0880 | 0,0433 | 0 | — | — |
| 16 | 0,5056 | 0,3290 | 0,2521 | 0,1939 | 0,1447 | 0,1005 | 0,0593 | 0,0196 | — | — |
| 18 | 0,4886 | 0,3253 | 0,2553 | 0,2027 | 0,1587 | 0,1197 | 0,0837 | 0,0496 | 0,0163 | — |
| 20 | 0,4734 | 0,3211 | 0,2565 | 0,2085 | 0,1686 | 0,1334 | 0,1013 | 0,0711 | 0,0422 | 0,0140 |

Rozpatrujemy zbiór krytyczny: $K = < 0; k >$
gdzie $k$ odczytujemy dla poziomu istotności $\alpha$ i danego $n$
z tablicy testu Shapiro-Wilka:

(tablica testu Shapiro-Wilka dla $\alpha = 0{,}05$)

| $n$ | 8 | 10 | 12 | 14 | 15 | 16 | 18 | 20 |
|---|---|---|---|---|---|---|---|---|
| $k$ | 0,818 | 0,842 | 0,859 | 0,874 | 0,881 | 0,887 | 0,897 | 0,905 |

Decyzje:
Jeśli $u_n \in K$ to $H_0$ odrzucamy.
Jeśli $u_n \notin K$ to nie ma podstaw do odrzucenia $H_0$.

**Przykład**

Dana jest uporządkowana próba 18 elementowa: 124, 142, 149, 156, 161, 168, 173, 179, 182, 193, 197, 204, 219, 228, 237, 252, 259, 274. Na poziomie istotności 0,05 sprawdzić testem Shapiro-Wilka hipotezę o normalności rozkładu badanej cechy.

Rozwiązanie

Średnia wynosi 194,3.

Suma kwadratów odchyleń od średniej $\sum_{i=1}^{n}(x_i - \bar{x})^2 = 31375,6$.

$$u_n = \frac{\left[0,4886(274-124) + 0,3253(259-142) + ... + 0,0163(193-182)\right]^2}{31375,6} = 0,97$$

$K = <0;\ 0,897>$, zatem $u_n \notin K$ i hipotezę o normalności rozkładu badanej cechy należy przyjąć.

# TEST NIEZALEŻNOŚCI $\chi^2$

Rozpatrujemy badane równocześnie dwie cechy X i Y (nie muszą być mierzalne).

Sprawdzamy hipotezę: $H_0$(X, Y są niezależne),

$\alpha$ - poziom istotności.

Próbę losową n elementową ($n \geq 80$) zapisujemy w postaci tablicy (podział na warianty powinien być taki aby $n_{ij} \geq 8$):

| | | Y | | | | |
|---|---|---|---|---|---|---|
| | | $Y_1$ | $Y_2$ | ... | $Y_l$ | $n_{i\bullet}$ |
| X | $X_1$ | $n_{11}$ | $n_{12}$ | ... | $n_{1l}$ | $n_{1\bullet}$ |
| | $X_2$ | $n_{21}$ | $n_{22}$ | ... | $n_{2l}$ | $n_{2\bullet}$ |
| | ... | ... | ... | ... | ... | ... |
| | $X_k$ | $n_{k1}$ | $n_{k2}$ | ... | $n_{kl}$ | $n_{k\bullet}$ |
| | $n_{\bullet j}$ | $n_{\bullet 1}$ | $n_{\bullet 2}$ | ... | $n_{\bullet l}$ | n |

Na podstawie próby obliczamy wartość statystyki

$$u_n = \sum_{i=1}^{k} \sum_{j=1}^{l} \frac{\left(n_{ij} - \hat{n}_{ij}\right)^2}{\hat{n}_{ij}} \quad (*)$$

(rozpatrywana statystyka ma rozkład $Y_{(k-1)(l-1)}$)
gdzie

$$\hat{n}_{ij} = \frac{n_{i\bullet} n_{\bullet j}}{n} = \frac{(\text{suma i} - \text{tego wiersza}) \times (\text{suma j} - \text{tej kolumny})}{\text{liczebność próby}}$$

Zbiór krytyczny ma postać

$$K = \langle k; \infty) ;$$

gdzie

$$P(Y_{(k-1)(l-1)} \geq \text{k}) = \alpha$$

Jeśli $u_n \in K$ to $H_0$ odrzucamy, w przeciwnym przypadku nie ma podstaw do odrzucenia $H_0$.

**Uwaga 1.**
W przypadku gdy cechy X i Y mają tylko po dwa warianty to rozpatrywana tablica ma postać (tzw. tablica czteropolowa):

| | | **Y** | | |
|---|---|---|---|---|
| | | 1 | 2 | |
| **X** | 1 | A | B | A+B |
| | 2 | C | D | C+D |
| | | A+C | B+D | n |

Statystyka $U_n$ ma wtedy postać:

$$U_n = \frac{n(AD - BC)^2}{(A+B)(A+C)(B+D)(C+D)}$$

i ma rozkład $Y_1$.

Jeśli w tablicy jest komórka o małej liczebności, to zalecana jest poprawka Yatesa i statystyka $U_n$ ma wtedy postać:

$$U_n = \frac{n(\left|AD - BC\right| - 0{,}5n)^2}{(A+B)(A+C)(B+D)(C+D)}$$

Dla tablicy 2x3:

| | | **Y** | | | |
|---|---|---|---|---|---|
| | | 1 | 2 | 3 | |
| **X** | 1 | $n_{11}$ | $n_{12}$ | $n_{13}$ | $N_1$ |
| | 2 | $n_{21}$ | $n_{22}$ | $n_{23}$ | $N_2$ |
| | | A | B | C | |

Statystyka $U_n$ ma postać:

$$U_n = N_1 N_2 \left( \frac{\frac{n_{11}}{N_1} - \frac{n_{21}}{N_2}}{A} + \frac{\frac{n_{12}}{N_1} - \frac{n_{22}}{N_2}}{B} + \frac{\frac{n_{13}}{N_1} - \frac{n_{23}}{N_2}}{C} \right)$$

i ma rozkład $Y_2$.

**Uwaga 2.**
Wielkość

$$T = \sqrt{\frac{U_n}{n\sqrt{(k-1)(l-1)}}}$$

nazywamy współczynnikiem Czuprowa
$(T \in < 0; 1 > )$.

Wielkość

$$V = \sqrt{\frac{U_n}{n(m-1)}}$$

gdzie m = min(k, l)

nazywamy współczynnikiem Cramera
$(V \in < 0;1 >)$.
Zauważmy, że dla tablic kwadratowych T = V.

Współczynniki te mogą służyć do oceny siły zależności między cechami (nawet w przypadku cech niemierzalnych).

**Uwaga 3.**

Jeśli mamy tablice wielkości $n_{ij}$ oraz $\hat{n}_{ij}$ to stosując funkcję CHI.TEST(tablica1; tablica2) programu EXCEL możemy wyznaczyć krytyczny poziom istotności i rozstrzygnąć niezależność rozpatrywanych cech.

**Przykład**

W celu zweryfikowania hipotezy, że studentki pewnej uczelni lepiej zdają egzaminy niż studenci, wylosowano próbę $n = 180$ studentek i studentów i otrzymano następujące wyniki zaliczenia letniej sesji egzaminacyjnej:

| SESJA | STUDENTKI | STUDENCI |
|---|---|---|
| ZALICZONA | 75 | 25 |
| NIEZALICZONA | 55 | 25 |

Na poziomie istotności $\alpha = 0{,}1$ sprawdzić hipotezę o niezależności wyników egzaminacyjnych od płci.

Rozwiązanie

$$u_n = 0{,}84 \qquad K = \langle 2{,}706;\infty)$$

zatem nie ma podstaw do odrzucenia hipotezy o niezależności.

# Badanie losowości próby - test serii.

W wielu zagadnieniach wnioskowania statystycznego istotnym założeniem jest losowość próby. Prostym testem do weryfikacji tej własności jest test serii.

Dla rozpatrywanego ciągu danych statystycznych obliczamy medianę $m_e$ (wartość środkowa).

Jeśli $x_1 \leq x_2 \leq \ldots\ldots \leq x_n$ dane uporządkowane to

$$m_e = \begin{cases} x_{\frac{n+1}{2}} & \text{dlannieparzystych} \\ \frac{1}{2}\left( x_{\frac{n}{2}} + x_{\frac{n+2}{2}} \right) & \text{dlanparzystych} \end{cases}$$

Przykład.
Dla danych (po uporządkowaniu)
2, 2, 3, 3, 4, 5, 5, 5, 5 medianą jest 4.
Dla danych (po uporządkowaniu)
2, 2, 2, 3, 3, 4, 5, 5, 5, 5
medianą jest 3,5.

Elementom próby przypisujemy symbol *a* lub *b*:

*a* - gdy $x_i > m_e$,

*b* - gdy $x_i < m_e$

(elementów $x_i = m_e$ nie rozpatrujemy).

**Serie** to podciągi złożone z jednakowych symboli.

Rozpatrujemy hipotezy
$H_0$(elementy próby mają charakter losowy),
$H_1$(elementy próby nie mają charakteru losowego),

Stosujemy statystykę:

$$U_n = \text{liczba serii}$$

Zbiór krytyczny:

$$K = (-\infty; k_1> \cup < k_{2;} \infty)$$

gdzie $k_1$ odczytujemy z tablicy dla poziomu istotności $\alpha/2$ i liczb $n_1$ oraz $n_2$ ,
gdzie $k_2$ odczytujemy z tablicy dla poziomu istotności $1 - \alpha/2$ i liczb $n_1$ oraz $n_2$ ,
gdzie $n_1$ - liczba symboli *a*, $n_2$ - liczba symboli *b*,
Decyzje:
Jeśli $U_n \in K$ to $H_0$ odrzucamy ,
Jeśli $U_n \notin K$ to nie ma podstaw do odrzucenia $H_0$ .

**Uwaga.**
Gdy $n_1$ lub $n_2$ jest większe od 20, to liczba serii ma w przybliżeniu rozkład

$$N\left(\frac{2n_1n_2}{n}+1;\sqrt{\frac{2n_1n_2\left(2n_1n_2-n\right)}{n^2(n-1)}}\right)$$

Dla dużych n można stosować rozkład

$$N\left(\frac{n}{2};\frac{\sqrt{n}}{2}\right)$$

**Tablica rozkładu serii**

Tablica dla $\alpha = 0{,}025$: (tablica jest symetryczna)

| $n_1$ | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 5 | | | 2 | 2 | | | | | | | | | | | | | | | |
| 6 | | 2 | 2 | 3 | 3 | | | | | | | | | | | | | | |
| 7 | | 2 | 2 | 3 | 3 | 3 | | | | | | | | | | | | | |
| 8 | | 2 | 3 | 3 | 3 | 4 | 4 | | | | | | | | | | | | |
| 9 | | 2 | 3 | 3 | 4 | 4 | 5 | 5 | | | | | | | | | | | |
| 10 | | 2 | 3 | 3 | 4 | 5 | 5 | 5 | 6 | | | | | | | | | | |
| 11 | | 2 | 3 | 4 | 4 | 5 | 5 | 6 | 6 | 7 | | | | | | | | | |
| 12 | 2 | 2 | 3 | 4 | 4 | 5 | 6 | 6 | 7 | 7 | 7 | | | | | | | | |
| 13 | 2 | 2 | 3 | 4 | 5 | 5 | 6 | 6 | 7 | 7 | 8 | 8 | | | | | | | |
| 14 | 2 | 2 | 3 | 4 | 5 | 5 | 6 | 7 | 7 | 8 | 8 | 9 | 9 | | | | | | |
| 15 | 2 | 3 | 3 | 4 | 5 | 6 | 6 | 7 | 7 | 8 | 8 | 9 | 9 | 10 | | | | | |
| 16 | 2 | 3 | 4 | 4 | 5 | 6 | 6 | 7 | 8 | 8 | 9 | 9 | 10 | 10 | 11 | | | | |
| 17 | 2 | 3 | 4 | 4 | 5 | 6 | 7 | 7 | 8 | 9 | 9 | 10 | 10 | 11 | 11 | 11 | | | |
| 18 | 2 | 3 | 4 | 5 | 5 | 6 | 7 | 8 | 8 | 9 | 9 | 10 | 10 | 11 | 11 | 12 | 12 | | |
| 19 | 2 | 3 | 4 | 5 | 6 | 6 | 7 | 8 | 8 | 9 | 10 | 10 | 11 | 11 | 12 | 12 | 13 | 13 | |
| 20 | 2 | 3 | 4 | 5 | 6 | 6 | 7 | 8 | 9 | 9 | 10 | 10 | 12 | 12 | 13 | 13 | 13 | 13 | 14 |

## Tablica rozkładu serii

Tablica dla $\alpha = 0{,}975$: (tablica jest symetryczna)

| $n_2$ \ $n_1$ | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2 | 4 | | | | | | | | | | | | | | | | | | |
| 3 | 5 | 6 | | | | | | | | | | | | | | | | | |
| 4 | 5 | 7 | 8 | | | | | | | | | | | | | | | | |
| 5 | 5 | 7 | 8 | 9 | | | | | | | | | | | | | | | |
| 6 | 5 | 7 | 8 | 9 | 10 | | | | | | | | | | | | | | |
| 7 | 5 | 7 | 9 | 10 | 11 | 12 | | | | | | | | | | | | | |
| 8 | 5 | 7 | 9 | 10 | 11 | 12 | 13 | | | | | | | | | | | | |
| 9 | 5 | 7 | 9 | 11 | 12 | 13 | 13 | 14 | | | | | | | | | | | |
| 10 | 5 | 7 | 9 | 11 | 12 | 13 | 14 | 15 | 15 | | | | | | | | | | |
| 11 | 5 | 7 | 9 | 11 | 12 | 13 | 14 | 15 | 16 | 16 | | | | | | | | | |
| 12 | 5 | 7 | 9 | 11 | 12 | 13 | 15 | 15 | 16 | 17 | 18 | | | | | | | | |
| 13 | 5 | 7 | 9 | 11 | 13 | 14 | 15 | 16 | 17 | 18 | 18 | 19 | | | | | | | |
| 14 | 5 | 7 | 9 | 11 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 19 | 20 | | | | | | |
| 15 | 5 | 7 | 9 | 11 | 13 | 14 | 15 | 17 | 17 | 18 | 19 | 20 | 21 | 21 | | | | | |
| 16 | 5 | 7 | 9 | 11 | 13 | 15 | 16 | 17 | 18 | 19 | 20 | 20 | 21 | 22 | 22 | | | | |
| 17 | 5 | 7 | 9 | 11 | 13 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 22 | 23 | 24 | | | |
| 18 | 5 | 7 | 9 | 11 | 13 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 24 | 25 | | |
| 19 | 5 | 7 | 9 | 11 | 13 | 15 | 16 | 17 | 19 | 20 | 21 | 22 | 22 | 23 | 24 | 25 | 25 | 26 | |
| 20 | 5 | 7 | 9 | 11 | 13 | 15 | 16 | 17 | 19 | 20 | 21 | 22 | 23 | 24 | 24 | 25 | 26 | 26 | 27 |

**Przykład**

W celu zbadania rozkładu wydajności pracy zarejestrowano czas wykonania detalu przez 15 wylosowanych pracowników i otrzymano wyniki (min):

16, 20, 25, 34, 22, 33, 47, 30, 28, 19, 22, 40, 36, 31, 38.

Sprawdzimy na poziomie istotności 0,05 hipotezę, że wybór próby był losowy.

Rozwiązanie.
Wyznaczamy medianę ( po uporządkowaniu danych niemalejąco) i otrzymujemy $m_e = 30$.
Kolejnym danym przyporządkowujemy symbole a i b:

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 16 | 20 | 25 | 34 | 22 | 33 | 47 | 30 |
| b | b | b | a | b | a | a | - |

| | | | | | | |
|---|---|---|---|---|---|---|
| 28 | 19 | 22 | 40 | 36 | 31 | 38 |
| b | b | b | a | a | a | a |

Liczba serii wynosi u = 6

Z tablic rozkładu serii odczytujemy

$$K = (-\infty; 3> \cup < 12_{;} \infty)$$

Ponieważ $u \notin K$ to nie ma podstaw do odrzucenia hipotezy $H_0$ , zatem możemy sądzić, że próba ma charakter losowy.

# Badanie zgodności rozkładów - test serii.

Mamy dwie próby pochodzące z dwóch populacji. Na podstawie tych prób chcemy sprawdzić czy rozkłady obu populacji nie różnią się (czyli w szczególności czy dwie próby pochodzą z jednej populacji).Prostym testem do weryfikacji tej własności jest również test serii.

Wyniki obu prób porządkujemy w jeden niemalejący ciąg.

Elementom tego ciągu przypisujemy symbol *a* lub *b*:

*a* - gdy element pochodzi z I próby,
*b* - gdy element pochodzi z II próby

**Serie** to podciągi złożone z jednakowych symboli.

Rozpatrujemy hipotezy

$H_0$(rozkłady populacji są takie same),

$H_1$(rozkłady populacji różnią się istotnie),

Stosujemy statystykę:

$$U_n = \text{liczba serii}$$

Zbiór krytyczny:

$$K = (0; k>$$

gdzie k odczytujemy z tablicy dla poziomu istotności $\alpha$
i liczb $n_1$ oraz $n_2$ ,
gdzie $n_1$ - liczba symboli *a*, $n_2$ - liczba symboli *b*,

Tablica dla $\alpha = 0{,}05$: (tablica jest symetryczna)

| $n_1$ / $n_2$ | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 4 | | | **2** | | | | | | | | | | | | | | | | |
| 5 | | **2** | **2** | **3** | | | | | | | | | | | | | | | |
| 6 | | **2** | **3** | **3** | **3** | | | | | | | | | | | | | | |
| 7 | | **2** | **3** | **3** | **4** | **4** | | | | | | | | | | | | | |
| 8 | **2** | **2** | **3** | **3** | **4** | **4** | **5** | | | | | | | | | | | | |
| 9 | **2** | **2** | **3** | **4** | **4** | **5** | **5** | **6** | | | | | | | | | | | |
| 10 | **2** | **3** | **3** | **4** | **5** | **5** | **6** | **6** | **6** | | | | | | | | | | |
| 11 | **2** | **3** | **3** | **4** | **5** | **5** | **6** | **6** | **7** | **7** | | | | | | | | | |
| 12 | **2** | **3** | **4** | **4** | **5** | **6** | **6** | **7** | **7** | **8** | **8** | | | | | | | | |
| 13 | **2** | **3** | **4** | **4** | **5** | **6** | **6** | **7** | **8** | **8** | **9** | **9** | | | | | | | |
| 14 | **2** | **3** | **4** | **5** | **5** | **6** | **7** | **7** | **8** | **8** | **9** | **9** | **10** | | | | | | |
| 15 | **2** | **3** | **4** | **5** | **6** | **6** | **7** | **8** | **8** | **9** | **9** | **10** | **10** | **11** | | | | | |
| 16 | **2** | **3** | **4** | **5** | **6** | **6** | **7** | **8** | **8** | **9** | **10** | **10** | **11** | **11** | **11** | | | | |
| 17 | **2** | **3** | **4** | **5** | **6** | **7** | **7** | **8** | **9** | **9** | **10** | **10** | **11** | **11** | **12** | **12** | | | |
| 18 | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **8** | **9** | **10** | **10** | **11** | **11** | **12** | **12** | **13** | **13** | | |
| 19 | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **8** | **9** | **10** | **10** | **11** | **12** | **12** | **13** | **13** | **14** | **14** | |
| 20 | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **9** | **10** | **11** | **11** | **12** | **12** | **13** | **13** | **14** | **14** | **15** |

Decyzje:

Jeśli $U_n \in K$ to $H_0$ odrzucamy ,

Jeśli $U_n \notin K$ to nie ma podstaw do odrzucenia $H_0$ .

**Przykład**

W celu porównania rozkładu wydajności pracy w dwóch filiach przedsiębiorstwa, zarejestrowano wydajność pracy 10 wylosowanych pracowników z każdej filii i otrzymano wyniki:

Filia I:

4,9 7,9 8,1 6,1 4,7 3,9 3,2 5,8 4,5 6,3

Filia II:

8,8 18,7 15,5 9,5 7,1 6,5 6,8 10,4 7,8 16,3

Sprawdzimy na poziomie istotności 0,05 hipotezę, że rozkład wydajności pracy w tych filiach jest taki sam.

Rozwiązanie.

Po uporządkowaniu danych w jeden ciąg niemalejąco i przyporządkowaniu symboli a i b:

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 3,2 | 3,9 | 4,5 | 4,7 | 4,9 | 5,8 | 6,1 | 6,3 | 6,5 | 6,8 | |
| a | a | a | a | a | a | a | a | b | b | |
| 7,1 | 7,8 | 7,9 | 8,1 | 8,8 | 9,5 | 10,4 | 15,5 | 16,3 | | 18,7 |
| b | b | a | a | b | b | b | b | b | | b |

Liczba serii wynosi u = 4

Z tablic rozkładu serii odczytujemy

$$K = (0; 6>$$

Ponieważ $u \in K$ to odrzucamy hipotezę $H_0$, zatem możemy sądzić, że wydajność pracy w tych filiach ma różny rozkład.