

# Testowanie hipotez statystycznych

## Definicje

**Hipoteza** - sąd o zbiorowości generalnej (populacji) wydany na podstawie próby statystycznej.

**Rodzaje hipotez** - parametryczne (o wartości przeciętnej, o wskaźniku struktury, o wariancji, itp.) oraz nieparametryczne (o rozkładzie cechy, o niezależności cech X i Y, itp.).

**Hipoteza zerowa ( $H_0$ )** - hipoteza sprawdzana.

**Hipoteza alternatywna ( $H_1$ )** - hipoteza, którą jesteśmy skłonni przyjąć gdy odrzucimy hipotezę zerową ( $H_0$ ).

**Test statystyczny** - reguła postępowania w wyniku której odrzucimy hipotezę zerową ( $H_0$ ).

## Rodzaje błędów w testowaniu hipotez

	<u>przyjąć</u> $H_0$	<u>odrzucić</u> $H_0$
$H_0$ <u>prawdziwa</u>	<b>O.K.</b> $1 - \alpha$	<b>błąd I-rodzaju</b> $\alpha$
$H_0$ <u>falszywa</u>	<b>błąd II-rodzaju</b> $\beta$	<b>O.K.</b> $1 - \beta$

**$\alpha$**  - jest to prawdopodobieństwo popełnienia błędu I-rodzaju i nazywane jest **poziomem istotności**. Zwykle przyjmuje się:

**$\alpha=0,05$**  (używane są również poziomy: 0,1; 0,02; 0,01)

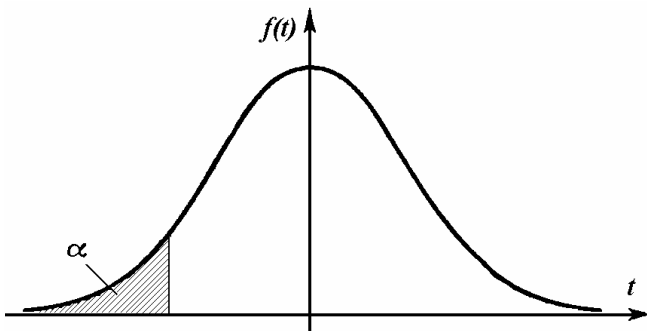
**Dobry test:** test w którym  $\alpha \approx \beta$

**Testy istotności** - testy, w których dla z góry ustalonego poziomu prawdopodobieństwa błędu I-rodzaju ( $\alpha$ ) poziom prawdopodobieństwa błędu II-rodzaju ( $\beta$ ) jest minimalny.

**Sprawdzian (hipotezy)** - statystyka, której wartość policzona na podstawie próby pozwala podjąć decyzję o odrzuceniu hipotezy zerowej ( $H_0$ ).

**Zbiór (obszar) krytyczny** - zbiór wartości sprawdzianu, które przemawiają za odrzuceniem hipotezy zerowej ( $H_0$ ).

## Rodzaje zbiorów (obszarów) krytycznych

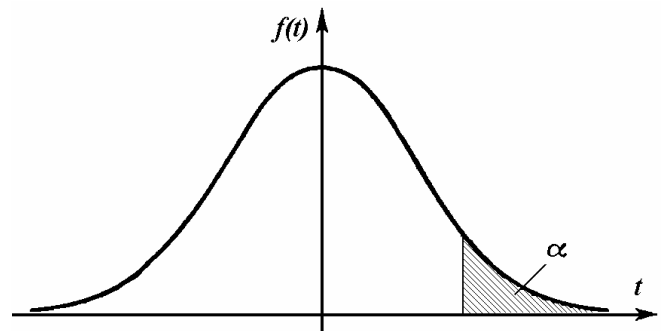


**Lewostronny**

obszar odrzucenia  $H_0$

$$H_0: Q = Q_0$$

$$H_1: Q < Q_0$$

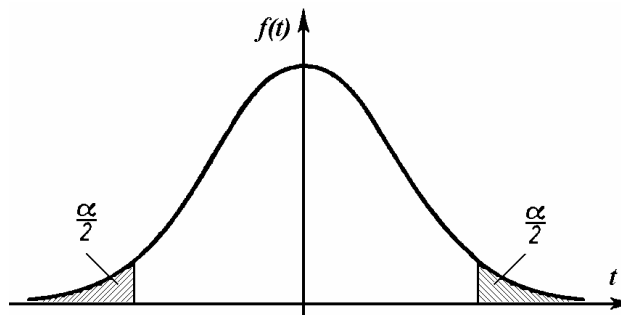


**prawostronny**

obszar odrzucenia  $H_0$

$$H_0: Q = Q_0$$

$$H_1: Q > Q_0$$



**obustronny**

obszar odrzucenia  $H_0$

$$H_0: Q = Q_0$$

$$H_1: Q \neq Q_0$$

# Testy parametryczne

Elementarnymi testami są tutaj następujące testy:

1. Testowanie hipotezy o wartości przeciętnej ( $m$ ) oraz
2. Testowanie hipotezy o wskaźniku struktury ( $p$ ).

W celu porównywania obu wymienionych parametrów w dwóch zbiorowościach stosuje się następujące testy (tylko na ćwiczeniach):

1. Testowanie hipotezy o równości dwóch wartości przeciętnych ( $m_1 = m_2$ ) oraz
2. Testowanie hipotezy o równości dwóch wskaźników struktury ( $p_1 = p_2$ ).

## Testowanie hipotezy o wartości przeciętnej ( $m$ )

**Założenie:** Cecha ma w populacji rozkład normalny  $N(m; \sigma)$ .  
Założenie to można weryfikować nieparametrycznymi testami zgodności (np. test zgodności chi-kwadrat).

### Formułowanie hipotez

Hipoteza zerowa ( $H_0$ ) jest hipotezą „o równości” i brzmi:

$$H_0: m = m_0$$

gdzie  $m_0$  jest konkretną wartością (liczbą).

Hipoteza alternatywna ( $H_1$ ) może być sformułowana trojako (najczęściej w zależności od wyniku uzyskanego w próbie):

$$H_1: m \neq m_0 \text{ (albo } H_1: m < m_0 \text{ albo też } H_1: m > m_0)$$

Wybór hipotezy alternatywnej ( $H_1$ ) ma decydujące znaczenie dla sformułowania obszaru odrzucenia.

## Konstruowanie sprawdzianu

Wybór sprawdzianu hipotezy zerowej ( $H_0$ ) zależy od liczebności próby  $n$  oraz od znajomości odchylenia standardowego  $\sigma$  w populacji.

Jeżeli:

- $\sigma$  jest znane i  $n \leq 30$  albo
  - $\sigma$  jest znane i  $n > 30$  albo
  - $\sigma$  jest nieznane i  $n > 30$  ale wówczas możemy przyjąć  $\sigma \approx S$
- to sprawdzianem hipotezy zerowej  $H_0$  jest statystyka:

$$(9.1) \quad T = \frac{\bar{X} - m}{\sigma} \sqrt{n}$$

która ma rozkład normalny  $N(0 ; 1)$

Jeżeli:

- $\sigma$  jest nieznane i  $n \leq 30$

to sprawdzianem hipotezy zerowej  $H_0$  jest statystyka:

$$(9.2) \quad T = \frac{\bar{X} - m}{S} \sqrt{n-1}$$

która ma rozkład Studenta o  $n-1$  stopniach swobody.

## Wnioskowanie

Jeżeli wartość sprawdzianu  $T$  znajdzie się:

1. w obszarze odrzucenia, to odrzucaamy  $H_0$  i przyjmujemy  $H_1$ .
2. poza obszarem odrzucenia, to nie mamy podstaw do odrzucenia  $H_0$ .

**UWAGA !!!** Nigdy nie mówimy o przyjęciu hipotezy  $H_0$ .

## **Jak odczytać z tablic wartość krytyczną $t_{kryt}$ , tj. granicę (granice) dla obszaru odrzucenia**

- Przyjmujemy **poziom istotności** czyli prawdopodobieństwo  $\alpha$  popełnienia błędu I-rodzaju.
- Rodzaj obszaru krytycznego określamy wstępnie na podstawie hipotezy alternatywnej  $H_1$  (wyjaśniają to rysunki na stronie 2).

### **Rozkład normalny $N(0 ; 1)$ (rozdane 2-stonicowe tablice)**

1. Dla obszaru lewostronnego odczytujemy taką wartość  $-t_{kryt}$ , dla której  $\Phi(-t_{kryt}) = \alpha$
2. Dla obszaru prawostronnego przyjmujemy wartość odczytaną dla obszaru lewostronnego i bierzemy ją ze znakiem dodatnim:  $+t_{kryt}$ .
3. Dla obszaru obustronnego odczytujemy taką wartość  $-t_{kryt}$ , dla której  $\Phi(-t_{kryt}) = \alpha/2$ . Granicami będą wartości:  $\pm t_{kryt}$

### **Rozkład Studenta (rozdane tablice)**

1. Dla obszaru lewostronnego lub prawostronnego odczytujemy taką wartość  $t_{kryt}$ , dla której  $P\{|T_{n-1}| > t_{kryt}\} > 2\alpha$  i przyjmujemy  $-t_{kryt}$  dla obszaru lewostronnego lub  $+t_{kryt}$  dla prawostronnego.
2. Dla obszaru obustronnego odczytujemy taką wartość  $t_{kryt}$ , dla której  $P\{|T_{n-1}| > t_{kryt}\} > \alpha$ . Granicami obszarów odrzucenia będą wartości:  $\pm t_{kryt}$

**PRZYKŁAD**

W 100 losowo wybranych gospodarstwach domowych średnia miesięczna opłata za energię elektryczną wyniosła 68 złotych, a odchylenie standardowe 14 złotych. Zweryfikuj panującą opinię, że przeciętne miesięczne wydatki na energię elektryczną w całej populacji ( $m$ ) wynoszą 75 złotych przyjmując poziom istotności 0,05.

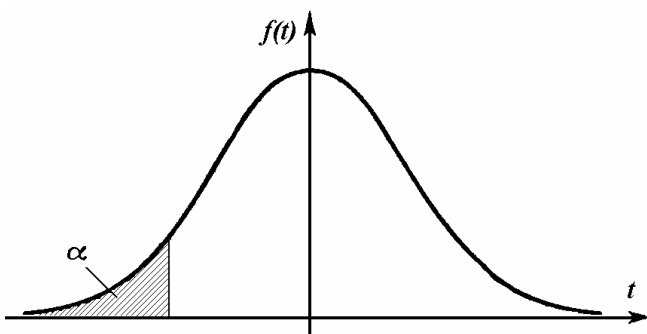
**Dane:**  $n = 100$        $\bar{x} = 68$        $S = 14$   
 $\alpha = 0,05$        $m_0 = 75$        $\sigma \approx S$

**Hipotezy:**       $H_0: m = 75$   
                           $H_1: m < 75$       (obszar lewostronny)

**Sprawdzian:**       $T = \frac{\bar{X} - m}{\sigma} \sqrt{n}$   
 $T = \frac{68 - 75}{14} \sqrt{100}$   
 $T = -5$

**Wartość krytyczna:** odczyt z rozkładu normalnego  $N(0;1)$

$$\alpha = 0,05 \quad \rightarrow \quad t_{kryt} = -1,64$$



$$T = -5 < -1,64$$

Wartość sprawdzianu  $T = -5$  leży w obszarze odrzucenia:

**WNIOSKOWANIE:** Należy odrzucić  $H_0$  i przyjąć  $H_1$ , tzn. że nieznane przeciętne wydatki na energię w całej populacji ( $m$ ) są mniejsze od 75 złotych.

**PRZYKŁAD** (czas dojazdu pracowników firmy DINO)

Dla 17 losowo wybranych pracowników firmy DINO otrzymano średni czas dojazdu 26 minut, a odchylenie standardowe 6 minut. Zweryfikuj panującą opinię, że przeciętny czas dojazdu w całej populacji ( $m$ ) wynosi 25 minut przyjmując poziom istotności 0,05.

Dane:  $n = 17$     $\bar{x} = 26$     $S = 6$

$$\alpha = 0,05 \quad m_0 = 25$$

Hipotezy:    $H_0: m = 25$

$H_1: m \neq 25$    (obszar obustronny)

Sprawdzian:

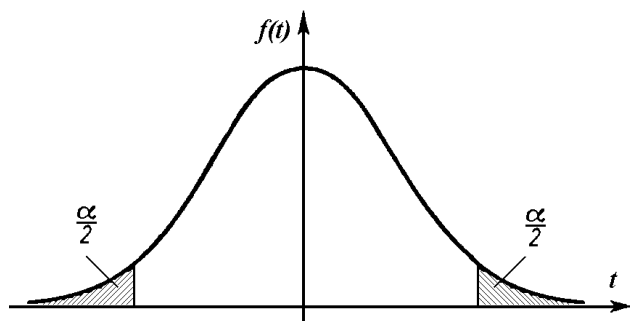
$$T = \frac{\bar{X} - m}{S} \sqrt{n-1}$$

$$T = \frac{26 - 25}{6} \sqrt{17-1}$$

$$T = \frac{2}{3}$$

Wartość krytyczna:   odczyt z rozkładu Studenta  
o  $17-1=16$  stopniach swobody.

$$\alpha = 0,05 \quad \rightarrow \quad t_{kryt} = \pm 2,1199$$



$$-2,1199 < T = 2/3 < +2,1199$$

Wartość sprawdzianu  $T = 2/3$  nie leży w obszarze odrzucenia.

**WNIOSKOWANIE:** Nie ma podstaw do odrzucenia  $H_0$ , tzn. że nieznaną przeciętną czas dojazdu w całej populacji ( $m$ ) jest być może równy 25 minut; test tego nie rozstrzyga.

## Testowanie hipotezy o wskaźniku struktury ( $p$ )

**Założenie:** Cecha ma w populacji rozkład dwupunktowy z parametrem  $p$  oznaczającym prawdopodobieństwo, że cecha przyjmie wyróżnioną wartość. Próba musi być duża ( $n > 100$ ).

### **Formułowanie hipotez**

**Hipoteza zerowa** ( $H_0$ ) jest hipotezą „o równości” i brzmi:

$$H_0: p = p_0$$

gdzie  $p_0$  jest konkretną wartością (liczbą).

**Hipoteza alternatywna** ( $H_1$ ) może być sformułowana trojako (najczęściej w zależności od wyniku uzyskanego w próbie):

$$H_1: p \neq p_0 \text{ (albo } H_1: p < p_0 \text{ albo też } H_1: p > p_0)$$

Wybór hipotezy alternatywnej ( $H_1$ ) ma decydujące znaczenie dla sformułowania obszaru odrzucenia.

### **Sprawdzian**

$$(9.5) \quad T = \frac{\frac{X}{n} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

która ma w przybliżeniu **rozkład normalny**  $N(0 ; 1)$

### **Wnioskowanie**

Jeżeli wartość sprawdzianu  $T$  znajdzie się:

1. w obszarze odrzucenia, to odrzucaamy  $H_0$  i przyjmujemy  $H_1$ .
2. poza obszarem odrzucenia, to nie mamy podstaw do odrzucenia  $H_0$ .



**PRZYKŁAD**

Panuje opinia, że w 40 % rodzin poważniejsze decyzje finansowe podejmuje małżonek. Zapytano 200 losowo wybranych przedstawicieli rodzin: „Kto podejmuje poważniejsze decyzje finansowe w domu?” W 72 przypadkach otrzymano odpowiedź, że podejmuje je małżonek. Zweryfikuj powszechnie panującą opinię na temat odsetka rodzin ( $p$ ), w których poważniejsze decyzje finansowe podejmuje małżonek przyjmując poziom istotności  $\alpha=0,02$ .

**Dane:**     $n = 200$      $X = 72$      $p_0 = 0,4$   
                   $\alpha = 0,02$

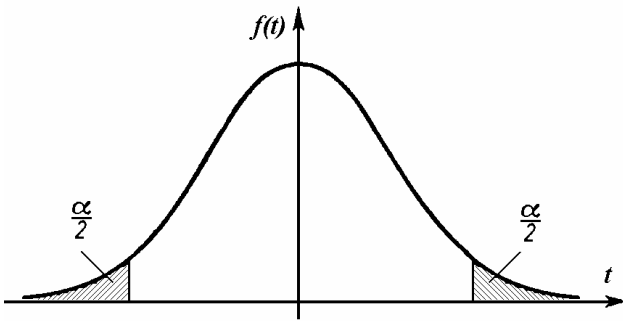
**Hipotezy:**         $H_0: p = 0,4$   
                           $H_1: p \neq 0,4$         (obszar obustronny)

**Sprawdzian:**

$$T = \frac{\frac{X}{n} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$
$$T = \frac{\frac{72}{200} - 0,4}{\sqrt{\frac{0,4(1-0,4)}{200}}}$$
$$T = -1,15$$

**Wartość krytyczna:** odczyt z rozkładu normalnego  $N(0;1)$

$$\alpha = 0,02 \rightarrow \alpha/2 = 0,01 \rightarrow t_{kryt} = \pm 2,33$$



$$-2,33 < T = -1,15 < +2,33$$

Wartość sprawdzianu  $T = -1,15$  nie leży w obszarze odrzucenia.

**WNIOSKOWANIE:** Nie ma podstaw do odrzucenia  $H_0$ , tzn. że nieznaną odsetek rodzin w całej populacji ( $p$ ), w których małżonek podejmuje poważniejsze decyzje finansowe jest być może równy 40%; test tego nie rozstrzyga.

# Testy nieparametryczne

Omówimy tutaj dwa spośród wielu testów nieparametrycznych:

1. test niezależności chi-kwadrat (testowanie niezależności cechy  $X$  i cechy  $Y$ ) oraz
2. test zgodności chi-kwadrat (testowanie zgodności rozkładu badanej cechy  $X$  z wybranym rozkładem teoretycznym).

## Test niezależności $\chi^2$ (*chi-kwadrat*)

Test służy badaniu zależności dwóch cech:  $X$  i  $Y$ . Obie cechy mogą być dowolne (jakościowe lub ilościowe).

Dla obu cech zbudowana jest tablica korelacyjna o  $r$  wierszach i  $s$  kolumnach (sposób przypisania cech  $X$  i  $Y$  do wierszy i kolumn jest dowolny).

### Formułowanie hipotez

$H_0$ : cecha  $Y$  NIE ZALEŻY od cechy  $X$

$H_1$ : cecha  $Y$  ZALEŻY od cechy  $X$

Oznaczmy:

$n_{ij}$  – liczebności empiryczne (liczba jednostek charakteryzujących się  $i$ -tym wariantem jednej cechy oraz  $j$ -tym wariantem drugiej cechy).

$n$  – liczba badanych jednostek

$n_{i\bullet}$  – liczebności brzegowa  $i$ -tego wiersza

$n_{\bullet j}$  – liczebności brzegowa  $j$ -tej kolumny

$n'_{ij}$  – liczebności teoretyczne (liczone przy założeniu, że hipoteza  $H_0$  jest prawdziwa). Liczebności teoretyczne wyliczamy następująco:

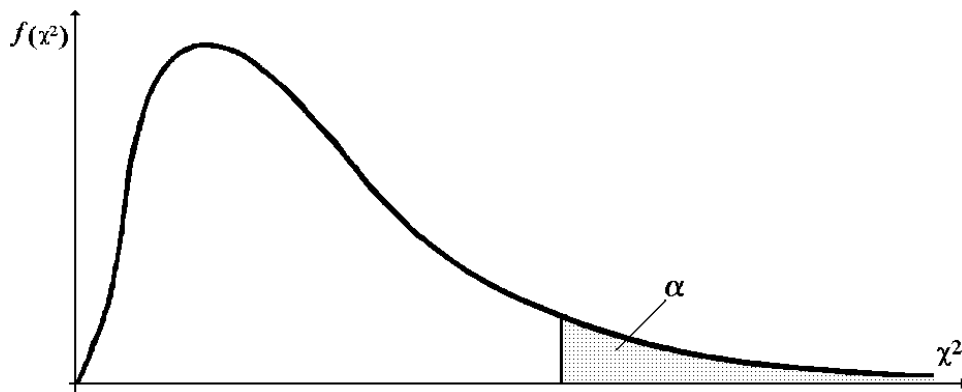
$$n'_{ij} = \frac{n_{i\bullet} \times n_{\bullet j}}{n}$$

## Sprawdzian

$$(10.7) \quad \chi^2 = \sum_{i=1}^r \sum_{j=1}^s \frac{(n_{ij} - n'_{ij})^2}{n'_{ij}}$$

która ma rozkład  $\chi^2$  o  $k = (r - 1)(s - 1)$  stopniach swobody.

Obszar odrzucenia jest tutaj obszarem prawostronnym.



## Wnioskowanie

Jeżeli wartość sprawdzianu  $\chi^2$  znajdzie się:

1. w obszarze odrzucenia, to odrzucaamy  $H_0$  i przyjmujemy  $H_1$ .
2. poza obszarem odrzucenia, to nie mamy podstaw do odrzucenia  $H_0$ .

## **Pomiar siły współzależności cech X i Y**

- Jeżeli obie cechy są cechami mierzalnymi możemy wykorzystać współczynnik korelacji  $r_{XY}$  Pearsona.
- W przeciwnym wypadku możemy zastosować jedną z miar opartych na wartości sprawdzianu  $\chi^2$ . Są to współczynniki współzależności:

1. C – Pearsona

$$C = \sqrt{\frac{\chi^2}{\chi^2 + n}}$$

2.  $\phi$  -Yula'a (10.8)

3.  $T$  -Czuprowa (10.9)

4.  $V$  -Cramera (10.10)

### **PRZYKŁAD**

Przeprowadzono szkolenie kilkuset kursantów. Podzielono ich losowo na cztery grupy i każdą z nich szkolono odrębną metodą. Na zakończenie kursu sprawdzono wiedzę kursantów za pomocą testu. Informacje o wynikach zestawiono w tablicy korelacyjnej.

Na poziomie istotności  $\alpha = 0,05$  zweryfikuj zastrzeżenie, że wynik testu zależał od metody szkolenia.

Wyniki testu – liczebności empiryczne  $[n_{ij}]$

Wynik testu (Y)	metoda nauczania (X)				$n_{i\bullet}$
	A	B	C	D	
<i>mierny</i>	30	40	40	20	130
<i>dostateczny</i>	30	40	20	40	130
<i>dobry</i>	40	20	40	40	140
$n_{\bullet j}$	100	100	100	100	400

# HIPOTEZY

$H_0$ : wynik testu **NIE ZALEŻY** od metody nauczania

$H_1$ : wynik testu **ZALEŻY** od metody nauczania

## Liczebności teoretyczne $[n'_{ij}]$

Wynik testu (Y)	metoda nauczania (X)				$n_{i\bullet}$
	A	B	C	D	
mierny					
dostateczny	32,5	32,5	32,5	32,5	
dobry	35,0	35,0	35,0	35,0	
$n_{\bullet j}$					

## Obliczanie wartości sprawdzianu $\chi^2$

Wynik testu (Y)	metoda nauczania (X)				$\Sigma$
	A	B	C	D	
mierny					
dostateczny	0,19	1,73	4,81	1,73	
dobry	0,71	6,43	0,71	0,71	
$\Sigma$					

wartości sprawdzianu  $\chi^2 =$

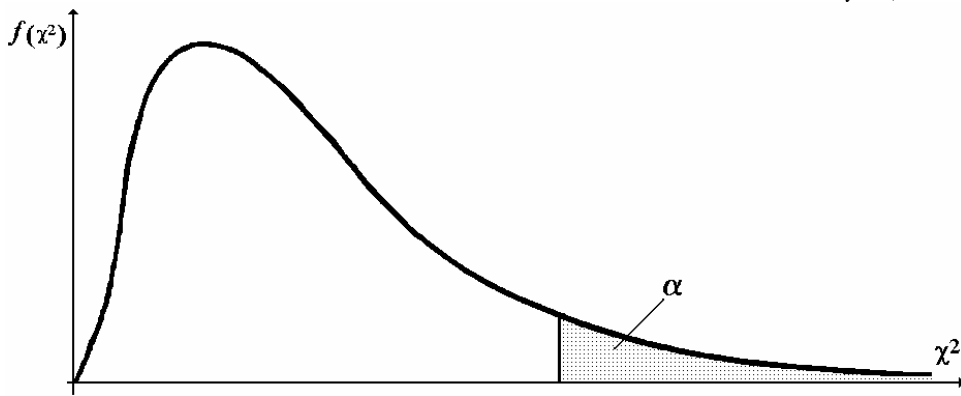
Liczba wierszy ( $r$ ) = 3

Liczba kolumn ( $s$ ) = 4

Liczba stopni swobody ( $k$ ) =  $(3-1)(4-1) =$

Poziom istotności  $\alpha =$

Wartość krytyczna odczytana z tablic:  $\chi^2_{kryt} = 12,592$



$$\chi^2_{kryt} = 12,592$$

<

$$\chi^2 = 25,48$$

Wartość sprawdzianu  $\chi^2 = 25,48$  leży w obszarze odrzucenia.

**WNIOSKOWANIE:** Należy odrzucić  $H_0$  i przyjąć  $H_1$ , tzn. że wynik testu (Y) zależał od metody nauczania (X).

### **Siła współzależności obu cech**

Obie cechy są niemierzalne (jakościowe).

Użyjemy zatem współczynnika współzależności C – Pearsona.

$$C = \sqrt{\frac{\chi^2}{\chi^2 + n}} = \sqrt{\frac{25,48}{25,48 + 400}} = 0,245$$

Współzależność obu cech jest wyraźna lecz niska.

## Test zgodności $\chi^2$ (*chi-kwadrat*)

Test służy badaniu czy rozkład cechy  $X$  podlega określonemu rozkładowi teoretycznemu.

Analogicznie jak w poprzednim teście sprawdzian  $\chi^2$  oparty jest na porównywaniu liczebności empirycznych z teoretycznymi wyliczonymi przy założeniu prawdziwości hipotezy  $H_0$ .

Ponieważ każdy rozkład wymaga odmiennej techniki wyliczania liczebności teoretycznych, to test zgodności  $\chi^2$  zilustrujemy na przykładzie sprawdzania wybranego rozkładu.

### PRZYKŁAD

Badaną cechą  $X$  jest odszkodowanie z tytułu kradzieży sprzętu komputerowego [tys. zł]. Pobrano próbę losową 168 wypłat odszkodowań. Wyniki zestawiono w postaci szeregu rozdzielczego z przedziałami klasowymi.

Na poziomie istotności  $\alpha = 0,05$  zweryfikuj założenie, że kwota odszkodowania  $X$  podlega rozkładowi normalnemu  $N(m; \sigma)$ .

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
				<i>środek klasy</i>				
$i$	$x_{0i}$	$x_{1i}$	$n_i$	$x_i$	$x_i * n_i$	$x_i - x_{\bar{s}r}$	$(7)*(7)$	$(8)*(4)$
1	3	5	16					
2	5	7	30					
3	7	9	34					
4	9	11	40					
5	11	13	30					
6	13	15	18					
<b>Razem</b>	<b>x</b>	<b>x</b>	<b>168</b>	<b>x</b>		<b>x</b>	<b>x</b>	

Dokończ samodzielnie obliczenia, a przekonasz się, że średnia z próby wynosi 9,1 tys. zł, a odchylenie standardowe 2,95 tys. zł.



## Formułowanie hipotez

$H_0$ : cecha  $X$  MA rozkład normalny

$H_1$ : cecha  $X$  NIE MA rozkładu normalnego

## Obliczanie wartości sprawdzianu $\chi^2$

Dane:  $\bar{x} = 9,1$        $S = 2,95$        $n = 168$

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$i$	$x_{0i}$	$x_{1i}$	$n_i$	$u_{1i}$	$\Phi(u_{1i})$	$p_i$	$n'_i$	$\chi^2$
1	3	5	16					
2	5	7	30					
3	7	9	34					
4	9	11	40					
5	11	13	30					
6	13	15	18	x	x			
Razem	x	x	168	x	x	1,00000	168,00	

Kolumna (5) - obliczanie  $u_{1i}$

Standaryzujemy prawe krańce przedziału klasowego (3),  
tj. standaryzujemy wartości  $x_{1i}$  według wzoru:

$$u_{1i} = \frac{x_{1i} - \bar{x}}{S} = \frac{x_{1i} - 9,1}{2,95}$$

Kolumna (6) - odczyt wartości dystrybuanty  $\Phi(u_{1i})$  z tablic  $N(0;1)$

Kolumna (7) - obliczanie prawdopodobieństw  $p_i$  dla klas przedziałowych

Klasa 1  $p_1 = \Phi(u_{11})$

Klasa 2 - 5  $p_i = \Phi(u_{1i}) - \Phi(u_{1i-1})$

Klasa 6 (ostatnia)  $p_6 = 1 - \Phi(u_{15})$

Kolumna (8) - obliczanie liczebności teoretycznych  $n'_i$  dla klas

$$n'_i = n \times p_i = 168 \times p_i$$

**Kolumna (9)** - wartość sprawdzianu  $\chi^2$

$$(10.1) \quad \chi^2 = \sum_{i=1}^r \frac{(n_i - n'_i)^2}{n'_i}$$

który ma **rozkład  $\chi^2$**  o  $k = r - s - 1$  stopniach swobody, gdzie:

$r$  - liczba klas w szeregu rozdzielczym,

$s$  - liczba parametrów, które należało wstępnie oszacować na podstawie próby (tutaj: średnia i odchylenie standardowe)

**Wartość sprawdzianu** wynosi w przykładzie:

$$\begin{aligned} \chi^2 &= \sum_{i=1}^6 \frac{(n_i - n'_i)^2}{n'_i} = \\ &= 0,344 + 0,518 + 1,476 + 0,110 + 0,119 + 0,340 = \\ &= 2,907 \end{aligned}$$

Wyznaczanie obszaru odrzucenia oraz wnioskowanie jest tutaj analogiczne jak w teście niezależności chi-kwadrat.

W przykładzie:

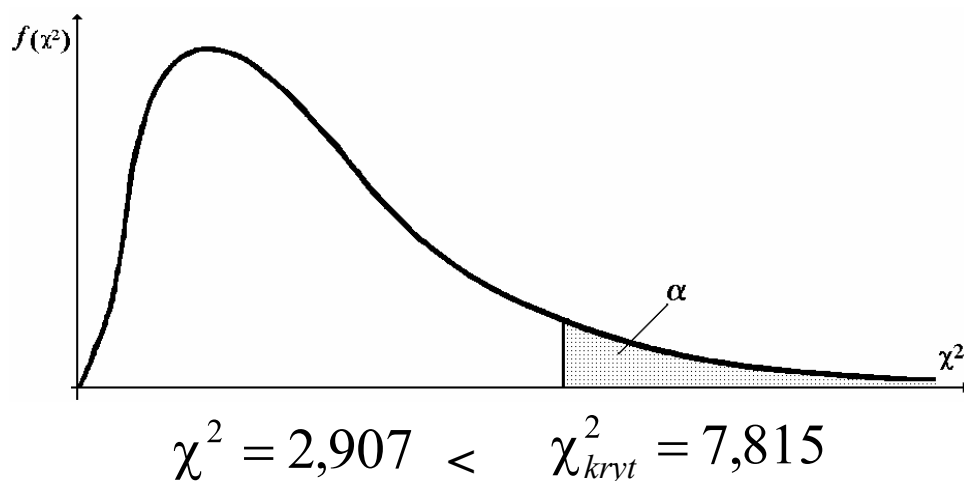
Liczba przedziałów klasowych ( $r$ ) = 6

Liczba oszacowanych wstępnie parametrów ( $s$ ) = 2

Liczba stopni swobody ( $k$ ) =  $6 - 2 - 1 =$

Poziom istotności  $\alpha =$

Wartość krytyczna odczytana z tablic:  $\chi^2_{kryt} = 7,815$



Wartość sprawdzianu  $\chi^2 = 2,907$  nie leży w obszarze odrzucenia.

**WNISKOWANIE:** Nie ma podstaw do odrzucenia hipotezy zerowej ( $H_0$ ) brzmiącej, że kwota odszkodowań z tytułu kradzieży sprzętu komputerowego ma rozkład normalny  $N(m; \sigma)$ .