Testowanie średnich - dwie średnie

Maciej Beręsewicz

03.12.2014

Zadanie 9. Zbadano 200 rodzin w miastach liczących 5 – 10 tys. mieszkańców i 100 rodzin wiejskich pod względem liczby posiadanych dzieci. Okazało się, ze rodziny miejskie miały łącznie 440 dzieci, a typowy obszar zmienności dla tej cechy wynosił od 1 do 3,4. Z kolei rodziny wiejskie posiadały średnio 2,5 dzieci przy odchyleniu standardowym wynoszącym 1. Traktując otrzymane wyniki jako wyniki badań próbnych, sprawdź hipotezę, ze średnia liczba dzieci w rodzinach wiejskich jest większa niż w rodzinach miejskich ($\alpha = 0, 05$).

W pierwszym kroku zapisujemy hipotezy:

- H0: $\mu_{1:miejskie} \neq \mu_{2:wiejskie}$
- H1: $\mu_{1:miejskie} < \mu_{2:wiejskie}$

Aby zweryfikować hipotezę musimy obliczyć statystykę krytyczną określoną wzorem:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}}.$$

gdzie:

$$s_p^2 = \frac{(n_1 - 1) \times s_1^2 + (n_2 - 1) \times s_2^2}{n_1 + n_2 - 2}$$

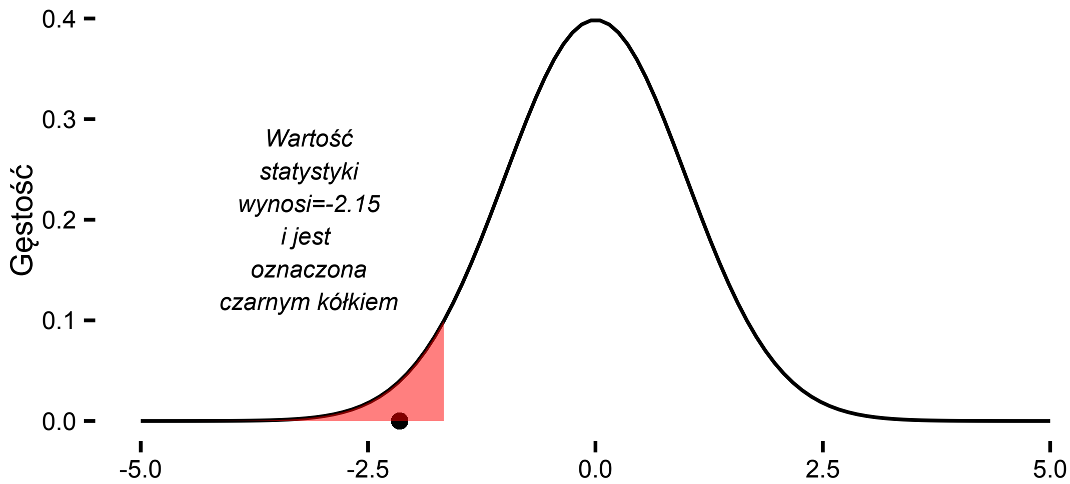
Natomiast weryfikujemy hipotezę porównując wartość statystyki testowej do wartości krytycznej określonej następująco.

$$t_{kr} = t_{\alpha, n_1 + n_2 - 2}$$

Szukamy wartości dla α dlatego, że badana hipoteza alternatywna jest jednostronna (pojawia się większe). Podstawiając wartości do wzoru dla dla statystyki t otrzymujemy:

$$t = \frac{2.2 - 2.5}{\sqrt{\frac{1.2938255}{200} + \frac{1.2938255}{100}}} = -2.1534648.$$

W związku z tym porównując do wartości krytycznej $t < t_{kr}$ odpowiadamy, że należy odrzucić hipotezę zerową ponieważ wartość statystyki jest mniejsza od wartości krytycznej. Możemy przedstawić to również graficznie.


0.4
0.3
0.2
0.1
0.0
Gęstość
Wartość
statystyki
wynosi=-2.15
i jest
oznaczona
czarnym kółkiem
-5.0
-2.5
0.0
2.5
5.0