Statystyka i modelowanie w naukach o środowisku

Wykład 3

Testowanie hipotez

Test istotności dla dwóch populacji

Test istotności dla dwóch wariancji służący do wnioskowania o równości dwóch wariancji w dwóch populacjach normalnych

Badamy dwie populacje generalne, mające rozkłady normalne $N(\mu_1,\sigma_1)$ i $N(\mu_2, \sigma_2)$.

Warunek normalności rozkładów jest spełniony - korzystamy z testu F na równość dwóch wariancji.

Test istotności dla dwóch wariancji służący do wnioskowania o równości dwóch wariancji w dwóch populacjach normalnych

Sprawdzamy, czy $\sigma^2_1 = \sigma^2_2$:

H_0: $\sigma^2_1 = \sigma^2_2$

H_1: $\sigma^2_1 \neq \sigma^2_2$

σ^2_1 wariancja pierwszej populacji

σ^2_2 wariancja drugiej populacji

Test istotności dla dwóch wariancji

Z obydwu populacji losuje się próby o liczebności n_1 i n_2. Następnie oblicza się wariancję dla każdej z prób $s^2{}_1$ i $s^2{}_2$.

W przypadku testu dla dwóch wariancji jako statystykę testową stosuje się statystykę F- Snedecora.

Test istotności dla dwóch wariancji

W liczniku statystyki F musi znaleźć się większa z wariancji w związku z założoną hipoteza zerową, a stopnie swobody n_L i n_M dotyczą odpowiednio wariancji z licznika oraz mianownika.

jeśli $s^2{}_1 > s^2{}_2$

$$F = \frac{s_1^2}{s_2^2}$$

jeśli $s^2{}_2 > s^2{}_1$

$$F = \frac{s_2^2}{s_1^2}$$

Test istotności dla dwóch wariancji

Następnie odczytujemy z tablic wartość statystyki $\mathbf{F}_{\alpha, nL, nM}$, (statystyki F przy założonym poziomie istotności oraz n_L i n_M stopniach swobody $n_L = n_1 - 1$, $n_M = n_2 - 1$ lub $n_L = n_2 - 1$, $n_M = n_1 - 1$).

Jeżeli $F > F_{a, nL, nM}$ to istnieją podstawy do odrzucenia hipotezy zerowej, w przeciwnym wypadku nie ma podstaw do odrzucenia hipotezy zerowej.

Test istotności dla dwóch wariancji - przykład

Przy założeniu, że dzienne wydatki studentów na rozrywkę w Krakowie (x_1) są takie same jak w Rzeszowie(x_2), wylosowano z Krakowa 45 studentów, a z Rzeszowa 36 studentów. Na podstawie zebranych informacji obliczono średnią arytmetyczną i wariancję wydatków studentów na rozrywkę w obu miastach.

$$\bar{x_1} = 22\ zł$$

$$\bar{x_2} = 18\ zł$$

$$s_1^2 = 2{,}5\ (zł)^2$$

$$s_2^2 = 1{,}4\ (zł)^2$$

Zweryfikować hipotezę, że wariancje wydatków na rozrywkę obu badanych grup studentów są takie same, przy założeniu, że poziom istotności wynosi 0,05, wobec hipotezy alternatywnej zakładającej, że wariancja wydatków w grupie krakowskiej jest większa.

Test istotności dla dwóch wariancji - przykład

$H_0: \sigma^2{}_1 = \sigma^2{}_2$

$H_1: \sigma^2{}_1 \neq \sigma^2{}_2$

$$F = \frac{s_1^2}{s_2^2} = \frac{2,5}{1,4} = 1,79$$

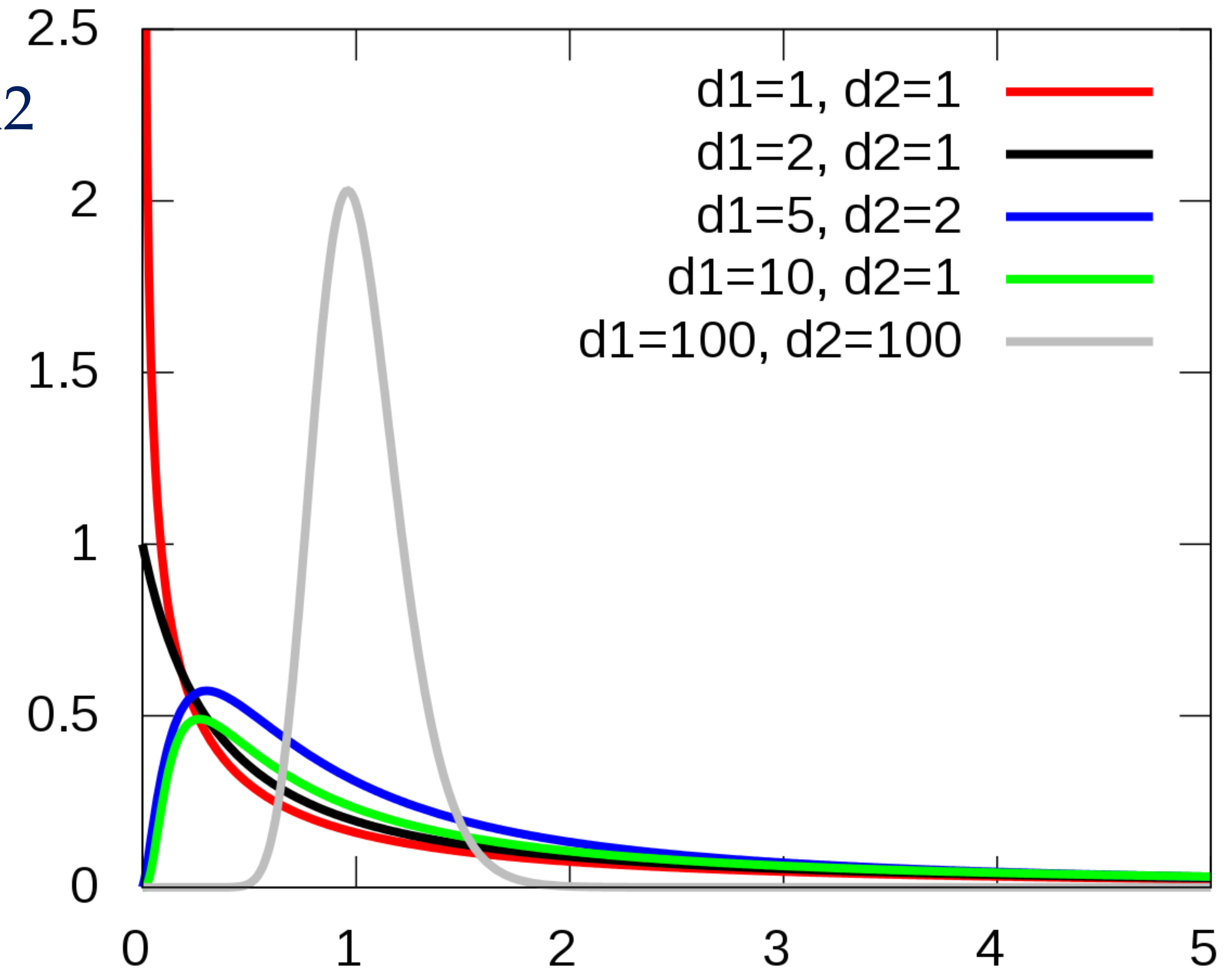
Wartość krytyczna $F_{0,05,\ 44,\ 35}$

wartość statystyki F przy poziomie istotności równym 0,05 oraz 44 i 35 stopniach swobody

Excel:

=ROZKŁAD.F.ODW(0.05;44;35)

Rozkład F Snedecora
dwa parametry na rys.
stopnie swobody d1 i d2



Tablice dla rozkładu F Snedecora
dwa parametry - stopnie swobody v1 i v2

Wartości krytyczne rozkładu F-Snedecora

$X \sim F_{\nu 1, \nu 2}$ - X zmienna losowa o rozkładzie F- Snedecora z liczbami stopni swobody (ν1, ν2)

poziom istotności α =0,05,

$F_{\alpha, \nu 1, \nu 2}$ - wartość krytyczna - liczba taka, że $P(X > F_{\alpha, \nu 1, \nu 2}) = \alpha$

	v1														
v2	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
1	161,446	199,499	215,707	224,583	230,160	233,988	236,767	238,884	240,543	241,882	242,981	243,905	244,690	245,363	245,949
2	18,513	19,000	19,164	19,247	19,296	19,329	19,353	19,371	19,385	19,396	19,405	19,412	19,419	19,424	19,429
3	10,128	9,552	9,277	9,117	9,013	8,941	8,887	8,845	8,812	8,785	8,763	8,745	8,729	8,715	8,703
4	7,709	6,944	6,591	6,388	6,256	6,163	6,094	6,041	5,999	5,964	5,936	5,912	5,891	5,873	5,858
5	6,608	5,786	5,409	5,192	5,050	4,950	4,876	4,818	4,772	4,735	4,704	4,678	4,655	4,636	4,619
6	5,987	5,143	4,757	4,534	4,387	4,284	4,207	4,147	4,099	4,060	4,027	4,000	3,976	3,956	3,938
7	5,591	4,737	4,347	4,120	3,972	3,866	3,787	3,726	3,677	3,637	3,603	3,575	3,550	3,529	3,511
8	5,318	4,459	4,066	3,838	3,688	3,581	3,500	3,438	3,388	3,347	3,313	3,284	3,259	3,237	3,218
9	5,117	4,256	3,863	3,633	3,482	3,374	3,293	3,230	3,179	3,137	3,102	3,073	3,048	3,025	3,006
10	4,965	4,103	3,708	3,478	3,326	3,217	3,135	3,072	3,020	2,978	2,943	2,913	2,887	2,865	2,845
11	4,844	3,982	3,587	3,357	3,204	3,095	3,012	2,948	2,896	2,854	2,818	2,788	2,761	2,739	2,719
12	4,747	3,885	3,490	3,259	3,106	2,996	2,913	2,849	2,796	2,753	2,717	2,687	2,660	2,637	2,617

Test istotności dla dwóch wariancji - przykład

Odpowiedź:

$F > F_{0,05,\ 44,\ 35}$

H_0 należy odrzucić.
Wariancja wydatków w grupie krakowskiej jest większa.

Test istotności dla dwóch wariancji – przykład Statistica

W Statistice test F wykonuje się tak jak test t dla porównania średnich dwóch populacji. Wynik znajduje się na końcu w tabeli wyników.

Wartość p>α oznacza, że nie możemy odrzucić hipotezy o równych wariancjach.

	iloraz F Wariancje	p Wariancje
68	2.766598	0.072702

Test istotności dla dwóch wariancji – przykład kalkulator online

http://www.statskingdom.com/220VarF2.html

Proszę pamiętać o zamianie przecinków na kropki.

Test istotności dla dwóch frakcji

cecha X_1 ma rozkład dwupunktowy z nieznanym parametrem p_1,
cecha X_2 ma rozkład dwupunktowy z nieznanym parametrem p_2,
pobrano n_1 – elementową próbę z pierwszej populacji oraz
n_2 – elementową próbę z drugiej populacji,
k_i – liczba elementów wyróżnionych w i -tej próbie;

Test istotności dla dwóch frakcji

Sprawdzamy, czy $p_1 = p_2$:

H_0: $p_1 = p_2$

H_1: $p_1 \neq p_2$

p_1 frakcja pierwszej populacji

p_2 frakcja drugiej populacji

Test istotności dla dwóch frakcji

cecha X_1 ma rozkład dwupunktowy z nieznanym parametrem p_1,
cecha X_2 ma rozkład dwupunktowy z nieznanym parametrem p_2,
pobrano n_1 – elementową próbę z pierwszej populacji oraz
n_2 – elementową próbę z drugiej populacji,
k_i – liczba elementów wyróżnionych w i -tej próbie

$$\overline{p_1} = \frac{k_1}{n_1}$$

$$\overline{p_2} = \frac{k_2}{n_2}$$

$$\bar{p} = \frac{k_1 + k_2}{n_1 + n_2}$$

Test istotności dla dwóch frakcji

$$u_{emp} = \frac{\overline{p_1} - \overline{p_2}}{\sqrt{\bar{p}(1-\bar{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

Test istotności dla dwóch frakcji wartości krytyczne u_{kryt}

dla α=0,05

=ROZKŁAD.NORMALNY.ODW(0,975;0;1)=1,9599≈1,96

dla α=0,01

=ROZKŁAD.NORMALNY.ODW(0,995;0;1)=2,5758≈2,58

Test istotności dla dwóch frakcji

Sprawdzamy, czy $p_1 = p_2$:

H_0: $p_1 = p_2$

H_1: $p_1 \neq p_2$

jeśli $|u_{emp}| > u_{kryt}$ nie ma podstaw do przyjęcia H_0

Test istotności dla dwóch frakcji - przykład

W dwóch dzielnicach miasta przeprowadzono ankietę na temat sortowania odpadków w gospodarstwach domowych. Otrzymano następujące wyniki: w pierwszej na 210 ankietowanych gospodarstw w 55 sortowano odpadki, natomiast w drugiej na 130 gospodarstw w 51 sortowano odpadki. Na poziomie istotności 0,01 zweryfikuj hipotezę o jednakowej frakcji gospodarstw sortujących odpadki w obu miastach.

Test istotności dla dwóch frakcji - przykład

$$\overline{p_1} = \frac{k_1}{n_1} = \frac{55}{210}$$

$$\overline{p_2} = \frac{k_2}{n_2} = \frac{51}{130}$$

$$\bar{p} = \frac{k_1 + k_2}{n_1 + n_2} = \frac{106}{340}$$

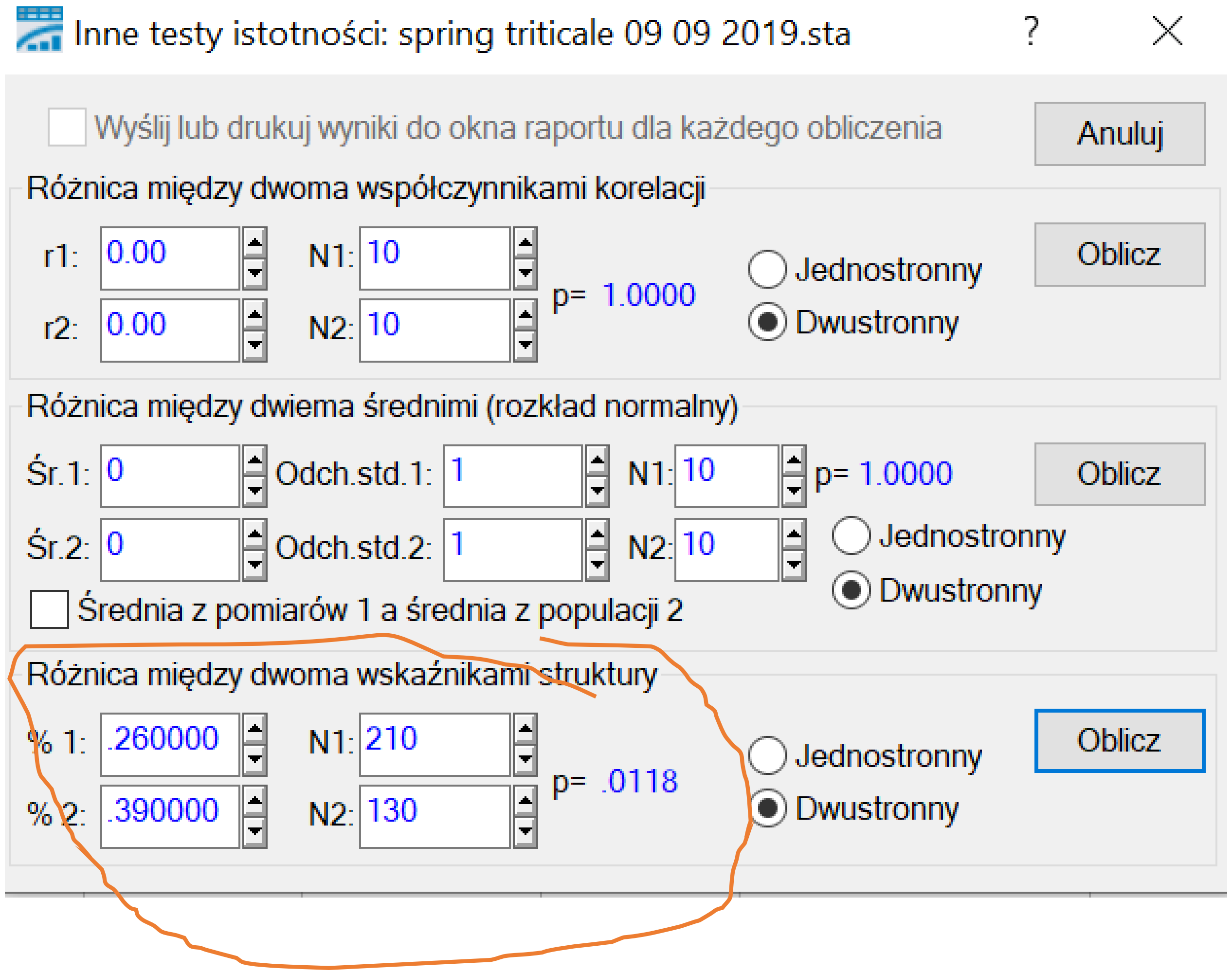
α=0,01

$$u_{emp} = 2,52$$

$$u_{kryt} = 2,58$$

$$|u_{emp}| < u_{kryt}$$

Test istotności dla dwóch frakcji – przykład Statistica



Test istotności dla dwóch frakcji - przykład

p>α możemy przyjąć H_0

Frakcje gospodarstw sortujących odpadki w miastach są równe.

Test istotności dla dwóch frakcji – przykład online kalkulator

https://www.socscistatistics.com/tests/ztest/default2.aspx