Wiesława Malska

Wykorzystanie testu Levene'a i testu Browna-Forsythe'a w badaniach jednorodności wariancji

Edukacja - Technika - Informatyka nr 4(18), 365-370

2016

Artykuł został opracowany do udostępnienia w internecie przez Muzeum Historii Polski w ramach prac podejmowanych na rzecz zapewnienia otwartego, powszechnego i trwałego dostępu do polskiego dorobku naukowego i kulturalnego. Artykuł jest umieszczony w kolekcji cyfrowej bazhum.muzhp.pl, gromadzącej zawartość polskich czasopism humanistycznych i społecznych.

Tekst jest udostępniony do wykorzystania w ramach dozwolonego użytku.

MUZEUM HISTORII POLSKI

© Wydawnictwo UR 2016 „Edukacja – Technika – Informatyka" nr 4/18/2016
ISSN 2080-9069 www.eti.rzeszow.pl
ISSN 2450-9221 online DOI: 10.15584/eti.2016.4.48

WIESŁAWA MALSKA

Wykorzystanie testu Levene'a i testu Browna-Forsythe'a w badaniach jednorodności wariancji

The use of Levene test and Brown-Forsythe test in the analysis of homogeneity of variance

Doktor inżynier, Politechnika Rzeszowska, Wydział Elektrotechniki i Informatyki, Katedra Energoelektroniki, Elektroenergetyki i Systemów Złożonych, Polska

Streszczenie

W artykule przedstawiono analizę możliwości zastosowania testu Levene'a i testu Browna-Forsythe'a do weryfikacji hipotezy o jednorodności wariancji dla dwóch lub więcej populacji, dostępne w programie STATISTICA. Wybierając odpowiedni test, należy zwrócić uwagę na liczebności prób losowych. Dla współczynnika istotności α podjęcie decyzji weryfikującej sprowadza się do interpretacji wartości prawdopodobieństwa testowego p, jaki otrzymuje się w wynikach odpowiednich testów. Wybór właściwego testu jest podstawowym wymogiem prawidłowego przebiegu procesu weryfikacji hipotezy o jednorodności wariancji dla dwóch lub więcej populacji.

Słowa kluczowe: hipoteza statystyczna, test istotności, test Levene'a, test Browna-Forsythe'a, poziom istotności, weryfikacja hipotez statystycznych.

Abstract

The paper presents an analysis of the applicability of the test Levene and Brown-Forsythe test to verify the hypothesis of homogeneity variance for two or more populations, which are available in STATISTICA. Choosing the right test, it is necessary to note the number of random samples. For the factor of significance α decision verifying boils down to the interpretation of probability test p, which receive the results of relevant tests. Choosing the correct test is a basic requirement for the proper conduct of the verification process statistical hypothesis of homogeneity of variance for two or more of the population.

Key words: statistical hypothesis, Levene tests, Brown-Forsythe tests, the level of significance, verification of statistical hypotheses.

Wstęp

Weryfikacja hipotez statystycznych jest obok estymacji działem wnioskowania o populacji generalnej na podstawie jedynie wyników próby losowej. Dowolny sąd (opinia, przypuszczenie) dotyczący parametrów lub też postaci

rozkładu cechy statystycznej (zmiennej) w populacji generalnej jest nazywany hipotezą statystyczną.

W testowaniu hipotez statystycznych prawdziwość lub fałszywość hipotezy ocenia się na podstawie wyników próby losowej, a reguła postępowania służąca sprawdzaniu prawdziwości hipotezy statystycznej jest nazywana testem istotności.

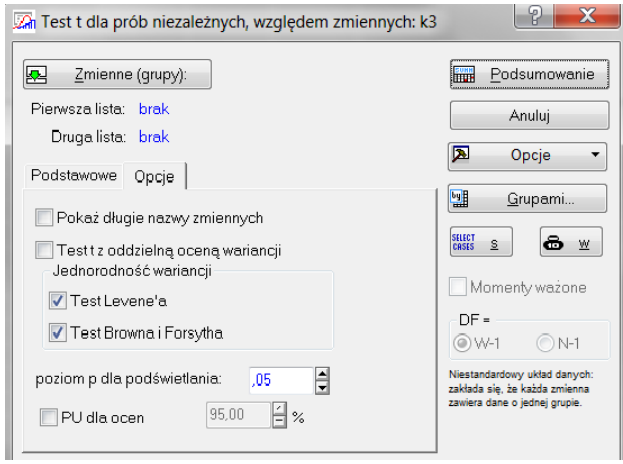
Wariancja jako miara rozrzutu (rozproszenia, dyspersji, zróżnicowania) jest często wykorzystywana w badaniach statystycznych do oceny powtarzalności wyników uzyskiwanych w eksperymentach naukowych lub do oceny ich jednorodności [Sobczyk 2010]. Wykorzystywany podczas weryfikacji hipotezy dla dwóch wariancji test istotności wymaga spełnienia warunku normalności rozkładu rozważanej cechy statystycznej w dwóch zbiorowościach statystycznych. Do najczęściej stosowanych testów należy test F-Snedecora, test Levene'a i test Browna-Forsythe'a [Kot i in. 2011; Rabiej 2012]. W artykule ograniczono się do wykorzystania testu Levene'a i testu Browna-Forsythe'a, które dostępne są w programie STATISTICA i pozwalają na weryfikację hipotezy parametrycznej o jednorodności wariancji.

Test Levene'a i test Browna-Forsythe'a

Test Levene'a można zastosować do badania równości (jednorodności) wariancji dla dwóch lub więcej zmiennych. Podstawowym wymogiem testu jest normalność rozkładu rozważanych zmiennych (cech statystycznych) oraz porównywalna liczebność (liczność) w każdej próbie losowej. Niestety w teście Levene'a to założenie nie zawsze jest spełnione. Jeżeli każda ze zmiennych X_i (i – numer populacji generalnej) ma rozkład normalny, to moduły różnicy odchyleń od średniej arytmetycznej w odpowiedniej grupie mają rozkłady silnie prawostronnie asymetryczne [Kot i in. 2011; Ostasiewicz i in. 2006]. Odstępstwo to osłabia moc testu, szczególnie w przypadku, gdy liczebności prób losowych są różnoliczne.

Dla dwóch wariancji wykorzystywany jest test Levene'a dostępny w module *Statystyka/Statystyki podstawowe i tabele/Test t dla prób niezależnych* na karcie *Opcje* (rys. 1). Natomiast w przypadku porównywania więcej niż dwóch wariancji należy zastosować testy *ANOVA* (Analysis of Variance) dostępne w module *Statystyka/Statystyki podstawowe i tabele/Przekroje, prosta ANOVA* (rys. 2) [www.statsoft.pl].

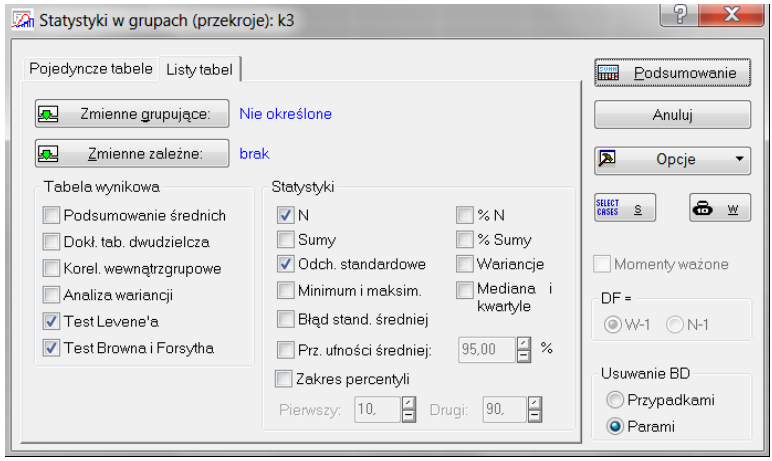
Procedura testu istotności Levene'a sprowadza się do porównywania wariancji poprzez testowanie równości wartości bezwzględnych odchyleń od średnich w poszczególnych grupach. W teście Levene'a wykonywana jest analiza wariancji wartości bezwzględnych odchyleń od średniej arytmetycznej w odpowiedniej grupie [Jóźwiak i in. 2009].



Rys. 1. Widok karty *Opcje* z testami jednorodności dla dwóch wariancji dostępne w module *Statystyka/Statystyki podstawowe i tabele/Test t dla prób niezależnych*

Konkluzję końcową podejmuje się poprzez porównanie wartości tzw. prawdopodobieństwa testowego p (p – poziom prawdopodobieństwa testowego w programie STATISTICA) z przyjętym poziomem istotności α [α – poziom istotności, $\alpha \in (0{,}01 \div 0{,}1)$].

Test Browna-Forsythe'a służy do weryfikacji hipotezy o jednorodności wariancji dla wielu zmiennych. Wymogiem jest spełnienie tak jak w teście Levene'a normalności rozkładu każdej z cech (zmiennych). Nie jest wymagana porównywalna liczebność prób losowych.



Rys. 2. Widok karty *Listy tabel* z testami jednorodności dla więcej niż dwóch wariancji dostępne w module *Statystyka/Statystyki podstawowe i tabele/Przekroje, prosta ANOVA/Statystyki w grupach*

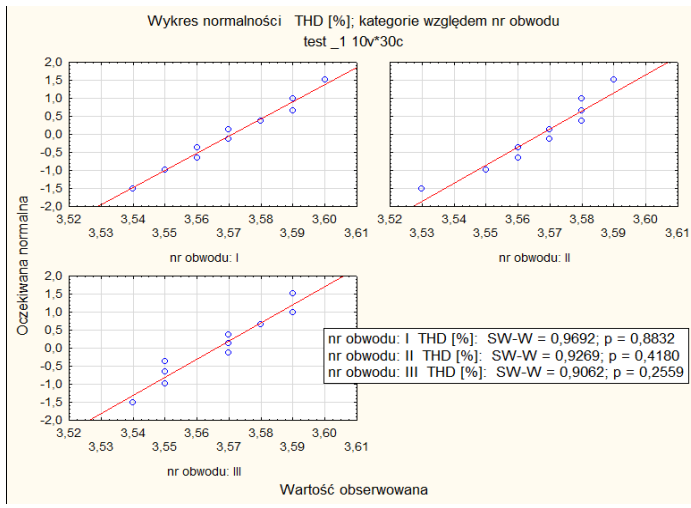
Test Browna-Forsythe'a jest zmodyfikowanym testem Levene'a. Modyfikacja polega głównie na obliczaniu odchyleń od median prób losowych, a nie od średnich arytmetycznych.

W testach statystycznych dla wielu wariancji w hipotezie zerowej zakłada się, że wariancje we wszystkich populacjach generalnych są równe, ale nie musimy dokładnie znać ich wielkości.

	1 nr obwodu	2 THD [%]	3 nr obwodu	4 THD [%]	5 nr obwodu	6 THD [%]
1	I	3,56	II	3,57	III	3,59
2	I	3,59	II	3,59	III	3,55
3	I	3,55	II	3,53	III	3,54
4	I	3,57	II	3,58	III	3,57
5	I	3,59	II	3,58	III	3,55
6	I	3,58	II	3,55	III	3,59
7	I	3,54	II	3,57	III	3,57
8	I	3,56	II	3,56	III	3,57
9	I	3,57	II	3,58	III	3,55
10	I	3,6	II	3,56	III	3,58

Rys. 3. Arkusz z wynikami pomiarów

Jako przykład wykorzystania testu Levene'a i testu Browna-Forsythe'a wykonano pomiary współczynnika zawartości wyższych harmonicznych w trzech obwodach. Dla każdego obwodu wykonano po dziesięć pomiarów współczynnika zawartości wyższych harmonicznych napięcia THD [%]. Wyniki pomiarów przedstawiono na rys. 3.



Rys. 4. Skategoryzowane wykresy normalności wraz z wynikami testu W Shapiro-Wilka

Analizę rozpoczyna się od sprawdzenia założeń normalności rozkładu pomiarów w każdym z obwodów. Sprawdzenia dokonuje się za pomocą testu W Shapiro-Wilka i skategoryzowanych wykresów normalności.

Skategoryzowane wykresy normalności wraz wynikami testu normalności W Shapiro-Wilka przedstawiono na rys. 4. Wartości prawdopodobieństwa testowego p dla każdego z obwodów są większe od współczynnika istotności równego 0,05.

Także skategoryzowane wykresy normalności świadczą o spełnieniu warunku normalności. Można zatem przystąpić do sprawdzenia założenia jednorodności wariancji. Do realizacji tego wykorzystano moduł *Statystyka/Statystyki podstawowe i tabele/Przekroje, prosta ANOVA*. Na rys. 5 przedstawiono wyniki testu Levene'a, a na rys. 6 dla porównania wyniki testu Browna-Forsythe'a.

	Test Levene'a jednorodności wariancji (test _1) Zaznaczone efekty są istotne z p < ,05000							
Zmienna	**SS Efekt**	df Efekt	MS Efekt	SS Błąd	df Błąd	MS Błąd	F	p
THD [%]	0,000014	2	0,000007	0,002590	27	0,000096	0,072289	0,930441

Rys. 5. Wynik testu Levene'a

	Test jednorod. wariancji Browna-Forsythe'a (test _1) Zaznaczone efekty są istotne z p < ,05000							
Zmienna	**SS Efekt**	df Efekt	MS Efekt	SS Błąd	df Błąd	MS Błąd	F	p
THD [%]	0,000020	2	0,000010	0,003300	27	0,000122	0,081818	0,921667

Rys. 6. Wynik testu Browna-Forsythe'a

Z testów Levene'a i Browna-Forsythe'a wynika, że spełnione jest założenie jednorodności wariancji. W obu przypadkach p > 0,05, czyli większe od przyjętego współczynnika istotności α równego 0,05.

Podsumowanie

Postępowanie badawcze związane z weryfikacją hipotez statystycznych powinno przebiegać w określonej sekwencji czynności. Programy komputerowe umożliwiają analizę i obliczenia różnych zagadnień statystycznych, czego dowodem jest wykorzystany do obliczeń w niniejszym artykule program STATISTICA. Test Levene'a i test Browna-Forsythe'a pozwala na weryfikację hipotez o jednorodności wariancji, a w tabelach wyników testów wystarczy interpretować wartość prawdopodobieństwa testowego. Konkluzja końcowa uzależniona jest od wartości prawdopodobieństwa testowego p i przyjętego poziomu istotności α. Zaprezentowany w artykule przykład hipotezy o jednorodności wariancji z wykorzystaniem testu Levene'a i testu Browna-Forsythe'a wymaga znajomości tematyki dotyczącej wnioskowania statystycznego i właściwej interpretacji otrzymanych wyników testów.

Literatura

Jóźwiak J., Podgórski J. (2009), *Statystyka od podstaw*, Warszawa.

Kot S., Jakubowski J., Sokołowski A. (2011), *Statystyka*, Warszawa.

Luszniewicz A., Słaby T. (2008), *Statystyka z programem komputerowym STATISTICA PL*, Warszawa.

Ostasiewicz S., Rusnak Z., Siedlecka U. (2006), *Statystyka. Elementy teorii i zadania*, Wrocław.

Rabiej M. (2012), *Statystyka z programem STATISTICA*, Gliwice.

Sobczyk M. (2010), *Statystyka matematyczna*, Warszawa.

StatSoft (2008), *Zastosowania metod statystycznych w badaniach naukowych III*, Kraków.

www.statsoft.pl.