

Sylwia KOSZNIK-BIERNACKA

Analiza wyników matur z języka polskiego w wybranych województwach

Odkąd obowiązuje tzw. nowa matura¹, czyli od 2005 r., i wyniki egzaminów maturalnych są wyrażone w skali procentowej, można dokonać dokładniejszej analizy uzyskiwanych wyników przez abiturientów. Poprzednia sześciostopniowa skala oceny matur uniemożliwiała np. badania postaci rozkładu uzyskiwanych wyników, czego podjęto się właśnie w obecnym opracowaniu.

Na podstawie danych udostępnionych przez Okręgową Komisję Egzaminacyjną w Gdańsku dokonano analizy wyników pisemnych matur z języka polskiego na poziomie podstawowym w województwach pomorskim i kujawsko-pomorskim w 2011 r. Liczba osób zdających maturę wynosiła 38493. Wielkość tej próby podzielono według następujących kategorii, tj. na uzyskane wyniki:

- przez maturzystów ogółem,
- w poszczególnych województwach,
- z uwzględnieniem szkół publicznych i niepublicznych²,
- z uwzględnieniem typu szkoły³.

W artykule podano wartości podstawowych parametrów charakteryzujących otrzymane wyniki. Na podstawie tak licznej zbiorowości podjęto próbę dopasowania do danych empirycznych odpowiedniego rozkładu prawdopodobieństwa, opisującego wyniki matur. Z literatury (Jóźwiak, Podgórski, 2001; Plucińska, Pluciński, 2000) wiadomo, że jeżeli na dane zjawisko oddziałuje duża liczba niezależnych czynników, których wpływ (traktowany odrębnie) jest mało znaczący, to takie zmienne mają często rozkład zbliżony do normalnego, np.: waga oraz wzrost osobników jednorodnych populacji ludzkich lub zwierzęcych, losowe błędy pomiarów czy cechy elementów pochodzących z masowej produkcji. Zatem w opracowaniu postawiono hipotezę, że wyniki pisemnych matur z języ-

¹ Rozporządzenie Ministra Edukacji Narodowej z 30 kwietnia 2007 r. w sprawie warunków i sposobu oceniania, klasyfikowania i promowania uczniów i słuchaczy oraz przeprowadzania sprawdzianów i egzaminów w szkołach publicznych (Dz. U. z 11.05.2007 r., Nr 83, poz. 562) oraz rozporządzenie Ministra Edukacji Narodowej z 17 listopada 2010 r. zmieniające rozporządzenie w sprawie warunków i sposobu oceniania, klasyfikowania i promowania uczniów i słuchaczy oraz przeprowadzania sprawdzianów i egzaminów w szkołach publicznych (Dz. U. z 2.12.2010 r., Nr 228, poz. 1491).

² Szkoły niepubliczne z uprawnieniami publicznymi.

³ Technikum, liceum ogólnokształcące, liceum profilowane, liceum uzupełniające, technikum uzupełniające.

ka polskiego na poziomie podstawowym można opisać rozkładem normalnym. W celu weryfikacji postawionej hipotezy przeprowadzono stosowne analizy i odpowiednie testy zgodności. Ostatecznie okazało się, że nie ma podstaw do odrzucenia postawionej hipotezy dla wszystkich rozpatrywanych kategorii danych. Przeprowadzono także szczegółową analizę wariancji (Józwiak, Podgórski, 2001; Stanisław, 2007), aby porównać średnie wyniki osiągnięte przez maturzystów w zakresie poszczególnych kategorii.

WYNIKI MATUR OGÓŁEM Z JĘZYKA POLSKIEGO W WOJEWÓDZTWACH POMORSKIM I KUJAWSKO-POMORSKIM

Wyznaczono wartości podstawowych parametrów statystycznych.

ZESTAWIENIE PODSTAWOWYCH WARTOŚCI PARAMETRÓW STATYSTYCZNYCH WYNIKÓW MATUR OGÓŁEM Z JĘZYKA POLSKIEGO W BADANYCH WOJEWÓDZTWACH

<i>N</i> ważnych	Średnia	Mediana	Moda	Wariancja	Odchylenie standardowe	Skośność	Kurtoza
38493	53,38	53,00	54,00	271,95	16,49	−0,10	0,02

Źródło: opracowanie własne.

Otrzymane wyniki matury z języka polskiego przedstawiono też na histogramie z dopasowanym wykresem rozkładu normalnego.

Sporządzono również wykres normalności.

Na podstawie wykrr. 1 i 2 oraz obliczonych wartości skośności i kurtozy bliskich zero (zestawienie) można postawić hipotezę, że rozkład normalny $N(53,38; 16,49)$ będzie dobrze opisywał badane zjawisko (Drapella, 2004).

Ponieważ wybrana zbiorowość jest dość liczna, a dane mają wiele wariantów, wydaje się, że test zgodności Mizesa-Smirnowa jest najodpowiedniejszy do zweryfikowania postawionej hipotezy. Obliczona wartość statystyki testowej w tym teście wynosi 0,006697. Porównując ją na poziomie istotności 0,05 z wartościami krytycznymi testu Mizesa-Smirnowa, którego implementację komputerową można znaleźć w pracy A. Drapelli (2000) okazuje się, że nie ma podstaw do odrzucenia postawionej hipotezy, czyli wyniki matur z języka polskiego na poziomie podstawowym mogą być opisane rozkładem normalnym $N(53,38; 16,49)$.

Drugim aspektem opracowania było badanie średniej w poszczególnych, wyodrębnionych kategoriach wyników matur. Dla danych ogółem została ona przedstawiona na wykrr. 3 wraz z uwzględnieniem typowej zmienności, o której można mówić w przypadku rozkładów symetrycznych zbliżonych do rozkładu normalnego.

Jak prezentuje wyk. 3, ok. 66% maturzystów zdało maturę z języka polskiego na poziomie 37—70%. Niestety więcej niż 5% zdających nie uzyskało wymaganego progu 30%.

*WYNIKI MATUR Z JĘZYKA POLSKIEGO
W BADANYCH WOJEWÓDZTWACH*

Obliczono podstawowe parametry statystyczne badanych danych.

**TABL. 1. PODSTAWOWE WARTOŚCI PARAMETRÓW STATYSTYCZNYCH
WYNIKÓW MATUR Z JĘZYKA POLSKIEGO WEDŁUG WOJEWÓDZTW**

Województwa	<i>N</i> waż- nych	Średnia	Mediana	Moda	Wariancja	Odchylenie standardowe	Skośność	Kurtoza
Kujawsko-pomor- skie	18434	53,34057	53	50	272,5063	16,50776	−0,014934	0,018193
Pomorskie	20059	53,41996	54	54	271,4421	16,47550	−0,174153	0,024298

Ź r ó d ł o: opracowanie własne.

Wyniki egzaminu z języka polskiego uzyskane przez maturzystów w województwach pomorskim i kujawsko-pomorskim przedstawiono też na histogramie z dopasowanym wykresem rozkładu normalnego.

Na podstawie wykresów 4 i 5 oraz bliskich zeru wartości skośności i kurtozy (tabl. 1) postawiono hipotezę, że wyniki matur z języka polskiego na poziomie podstawowym w województwach kujawsko-pomorskim i pomorskim można opisać rozkładem normalnym odpowiednio z parametrami $N(53,34; 16,51)$ i $N(53,42; 16,48)$. Hipotezę tę zweryfikowano przeprowadzając dla obu przypadków test zgodności Mizesa-Smirnowa. Otrzymane wartości statystyki testowej to odpowiednio 0,009272 i 0,005536. Porównując je z wartościami krytycznymi w obu przypadkach nie ma podstaw do odrzucenia hipotezy zerowej, że wyniki matur można opisać rozkładem normalnym odpowiednio $N(53,34; 16,51)$ i $N(53,42; 16,48)$.

Rozważono także wartości średnie wyników maturalnych z języka polskiego w badanych województwach, które porównano na wyk. 6.

Na podstawie wyk. 6 można sądzić, że otrzymane wyniki średnie w badanych województwach są zbliżone. Aby sprawdzić, czy rzeczywiście tak jest przeprowadzono analizę wariancji. Można tego dokonać, bo jak wykazano wcześniej, wyniki z obu województw mogą być opisane rozkładem normalnym, więc podstawowe założenie klasycznej analizy wariancji jest spełnione. Zweryfikowano hipotezę, że średnie wyniki matur z języka polskiego na poziomie

podstawowym w obu badanych województwach są równe, tzn. nie zależą od województwa. Za pomocą modułu Excela *Analiza danych* przeprowadzono stosowne obliczenia, których wyniki przedstawiono w tabl. 2.

TABL. 2. WYNIKI ANALIZY WARIANCJI WEDŁUG WOJEWÓDZTW
A. PODSUMOWANIE WYNIKÓW

Województwa	Licznik	Suma	Średnia	Wariancja
Kujawsko-pomorskie	18434	983280	53,34057	272,506262
Pomorskie	20059	1071551	53,41996	271,442130

B. ANALIZA WARIANCJI

Źródła wariacji	SS	df	MS	F	Wartość p	Test F
Pomiędzy województwa- mi	60,55233347	1	60,55233	0,22265838	0,637024461	3,841700205
W obrębie województw	10467694,17	38491	271,9517			
Razem	10467754,72	38492				

U w a g a. SS — suma kwadratów, df — stopnie swobody, MS — średnia kwadratów, F — empiryczna wartość statystyki testowej F, Wartość p — wartość progowa testu F, test F — wartość krytyczna testu F.

Ź r ó d ł o: jak przy tabl. 1.

Ponieważ wartość statystyki F (tabl. 2) jest znacznie mniejsza od statystyki testowej, czyli na poziomie istotności 5%, to nie ma podstaw do odrzucenia postawionej hipotezy zerowej. Zatem średnie wyniki uzyskane przez maturzystów w obu województwach nie różnią się statystycznie.

Przeprowadźmy jeszcze test z (Mynarski, 2003, 2006) w zakresie różnicy między średnimi wynikami matur z języka polskiego na poziomie podstawowym w badanych województwach. Przeprowadzono stosowne obliczenia, a otrzymane wyniki podano w tabl. 3.

TABL. 3. TEST Z Z DWIEMA PRÓBAMI ŚREDNICH WEDŁUG WOJEWÓDZTW

Test z	Kujawsko-pomorskie	Pomorskie
Średnia	53,34056634	53,4199611
Znana wariancja	272,5	271,44
Obserwacje	18434	20059
Różnica średnich według hipotezy	0	
z	-0,471831663	
Test z dwustronny	1,959963985	

Ź r ó d ł o: jak przy tabl. 1.

Uzyskana wartość statystyki z jest mniejsza od krytycznej, zatem i tutaj nie ma podstaw do odrzucenia hipotezy zerowej o równości średnich. Ten test potwierdza, że średnie wyniki matur w obu województwach nie różnią się statystycznie.

WYNIKI MATUR Z JĘZYKA POLSKIEGO Z UWZGLĘDNIENIEM SZKÓŁ PUBLICZNYCH I NIEPUBLICZNYCH

Podstawowe parametry statystyczne danych dotyczących szkół publicznych i niepublicznych przedstawiono w tabl. 4.

**TABL. 4. PODSTAWOWE WARTOŚCI PARAMETRÓW STATYSTYCZNYCH WYNIKÓW
MATUR Z JĘZYKA POLSKIEGO WEDŁUG SZKÓŁ PUBLICZNYCH I NIEPUBLICZNYCH**

Szkoły	N waż- nych	Średnia	Mediana	Moda	Wariancja	Odchylenie standardowe	Skośność	Kurtoza
Publiczne	35404	54,25782	54	54	257,7918	16,05590	-0,054065	-0,036167
Niepubliczne	3089	43,34315	43	30	324,7054	18,01958	0,020942	0,038122

Ź r ó d ł o: jak przy tabl. 1.

Wyniki matury z języka polskiego z uwzględnieniem rodzaju szkoły przedstawiono również na histogramie z dopasowanym wykresem rozkładu normalnego.

Również w tej kategorii wyników matur (wykr. 7 i 8) zastosowanie krzywej normalnej powinno dawać dobrą zgodność z rozkładem empirycznym w zakresie poszczególnych rodzajów szkół. Sprawdzono zatem formalnie stosując test zgodności Mizesa-Smirnowa, czy wyniki matur na poziomie podstawowym w szkołach publicznych i niepublicznych można opisać rozkładem normalnym odpowiednio z parametrami $N(54,26; 16,06)$ i $N(43,34; 18,02)$. Obliczone wartości statystyki testowej to odpowiednio 0,007786 i 0,011775. Tak więc w obu przypadkach wartości statystyki testowej są zdecydowanie mniejsze niż wartości krytyczne w teście zgodności Mizesa-Smirnowa, nie ma więc podstaw do odrzucenia postawionej hipotezy.

Dane empiryczne można opisać rozkładem normalnym, czyli spełnione jest podstawowe założenie analizy wariancji. Przeprowadzono zatem badanie wartości średniej wyników uzyskanych w szkołach publicznych i niepublicznych, co przedstawiono wraz z typową zmiennością na wykr. 9.

Z wykr. 9 można odczytać, że w szkołach publicznych ponad 66% abiturientów zdało egzamin maturalny na poziomie dostatecznym i dobrym. W szkołach niepublicznych było nieco gorzej. Typowy obszar zmienności cechy był tam poniżej progowych 30%, czyli odsetek niezdanych matur z języka polskiego w szkołach niepublicznych był większy niż w szkołach publicznych.

Przeprowadzając klasyczną analizę wariancji wyników uzyskanych przez uczniów szkół publicznych i niepublicznych, otrzymano rezultaty, które podano w tabl. 5.

**TABL. 5. WYNIKI ANALIZY WARIANCJI Z UWZGLĘDNIENIEM
SZKÓŁ PUBLICZNYCH I NIEPUBLICZNYCH
A .PODSUMOWANIE WYNIKÓW**

Szkoły	Licznik	Suma	Średnia	Wariancja
Publiczne	35404	1920944	54,25782	257,7918
Niepubliczne	3089	133887	43,34315	324,7054

B. ANALIZA WARIANCJI

Źródła wariacji	SS	df	MS	F	Wartość p	Test F
Pomiędzy szkołami publicznymi i niepublicznymi	338461,9	1	338461,9	1286,145	0	3,8417
W obrębie szkół publicznych i niepublicznych	10129293	38491	263,16			
R a z e m	10467755	38492				

U w a g a. Jak przy tabl. 2.
Ź r ó d ł o: jak przy tabl. 1.

Otrzymana wartość statystyki testowej F jest bardzo duża, można więc powiedzieć, że rodzaj szkoły zdecydowanie wpływa na średnie wyniki maturalne uzyskiwane z języka polskiego. Ze statystycznego punktu widzenia wyniki średnie osiągane przez maturzystów zdających egzamin z języka polskiego w szkołach publicznych i niepublicznych są różne.

Potwierdzono ten fakt przeprowadzając test z porównujący uzyskane wyniki średnie w obu badanych grupach. Wyniki obliczeń podano w tabl. 6.

TABL. 6. TEST Z Z DWIEMA PRÓBAMI DLA ŚREDNICH W SZKOŁACH PUBLICZNYCH I NIEPUBLICZNYCH

Test z	Publiczne	Niepubliczne
Średnia	54,257824	43,34315312
Znana wariancja	257,7918	324,7054
Obserwacje	35404	3089
Różnica średnich według hipotezy	0	
z	32,556008	
Test z dwustronny	1,959964	

Źródło: jak przy tabl. 1.

Wartość statystyki z jest bardzo duża w porównaniu z wartością krytyczną tego testu. Potwierdza to wcześniejszy wniosek, że wyników średnich uzyskanych w obu rozważanych grupach nie można uznać za porównywalne statystycznie.

WYNIKI MATUR Z JĘZYKA POLSKIEGO Z UWZGLĘDNIENIEM TYPU SZKOŁY

Najpierw obliczono podstawowe parametry statystyczne dla tej kategorii danych.

TABL. 7. PODSTAWOWE WARTOŚCI PARAMETRÓW STATYSTYCZNYCH WYNIKÓW MATUR Z JĘZYKA POLSKIEGO Z UWZGLĘDNIENIEM TYPU SZKOŁY

Typ szkoły	N	Średnia	Mediana	Moda	Wariancja	Odchylenie standardowe	Skośność	Kurtoza
Liceum ogólnokształcące	22482	58,82208	59	59	246,0199	15,68502	-0,260278	0,275152
profilowane	1008	44,49603	44	30	219,6802	14,82161	-0,073727	0,435019
uzupełniające	1495	36,96187	37	30	228,8533	15,12790	-0,031856	-0,160606
Technikum	13174	47,09579	47	30	191,1237	13,82475	-0,018072	0,251815
uzupełniające	334	35,45808	36	30	241,0598	15,52610	-0,061662	0,052492

Źródło: jak przy tabl. 1.

Wyniki matury z języka polskiego z uwzględnieniem typu szkoły przedstawiono też na histogramie z dopasowanym wykresem rozkładu normalnego.

Oczekiwaną normalną w zakresie poszczególnych typów szkół pokazano również na wykr. 11.

Na podstawie danych zilustrowanych na wyk. 10 i 11 postawiono hipotezę, że wyniki matur z języka polskiego na poziomie podstawowym w poszczególnych typach szkół można również opisać rozkładem normalnym z następującymi parametrami dla: liceum ogólnokształcącego $N(58,82208; 15,68502)$, liceum profilowanego $N(44,49603; 14,82161)$, liceum uzupełniającego $N(36,96187; 15,12790)$, technikum $N(47,09579; 13,82475)$, technikum uzupełniającego $N(35,45808; 15,52610)$. Zweryfikowano tę hipotezę przeprowadzając test zgodności Mizesa-Smirnowa, otrzymane wartości statystyki testowej to odpowiednio 0,005423243, 0,012769; 0,012935; 0,011917; 0,019955. W przypadku wszystkich typów szkół porównując otrzymane wartości z wartościami krytycznymi testu Mizesa-Smirnowa nie ma podstaw do odrzucenia hipotezy zerowej. Wyniki matur z języka polskiego na poziomie podstawowym można opisać rozkładem normalnym z podanymi wcześniej parametrami. Ponieważ spełnione jest podstawowe założenie analizy wariancji, to zbadano też wartości średnie uzyskane w poszczególnych typach szkół. Otrzymane wyniki wraz z typową zmiennością podano na wyk. 12.

Przeprowadzając klasyczną analizę wariancji wyników uzyskanych w poszczególnych typach szkół otrzymano rezultaty przedstawione w tabl. 8.

TABL. 8. WYNIKI PRZEPROWADZONEJ ANALIZY WARIANCJI WEDŁUG TYPU SZKOŁY
A. PODSUMOWANIE WYNIKÓW

Typ szkoły	Licznik	Suma	Średnia	Wariancja
Liceum ogólnokształcące	22482	1322438	58,82208	246,0199
profilowane	1008	44852	44,49603	219,6802
uzupełniające	1495	55258	36,96187	228,8533
Technikum	13174	620440	47,09579	191,1237
uzupełniające	334	11843	35,45808	241,0598

B. ANALIZA WARIANCJI

Źródła wariancji	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	Wartość <i>p</i>	Test <i>F</i>
Pomiędzy typami szkół	1775910	4	443977,4	1965,958	0	2,372163
W obrębie typu szkoły ...	8691845	38488	225,8326			
R a z e m	10467755	38492				

U w a g a. Jak przy tabl. 2.

Ź r ó d ł o: jak przy tabl. 1.

Ponieważ wartość statystyki testowej *F* jest bardzo duża można stwierdzić, że typ szkoły zdecydowanie wpływa na uzyskiwane wyniki maturalne z języka polskiego.

W przypadku tej kategorii danych nie przeprowadzono testu z równości średnich, bo dotyczy on porównywania średnich tylko w dwóch populacjach danych.

Wnioski

Na podstawie zbiorowości liczącej 38493 uczniów zdających maturę z języka polskiego na poziomie podstawowym stwierdzono, że uzyskane wyniki można opisać rozkładem normalnym bez względu na analizowane w opracowaniu województwo, szkoły publiczne i niepubliczne, jak i bez względu na typ szkoły.

Z przeprowadzonych analiz wynika też, że średnie wyniki maturalne z języka polskiego na poziomie podstawowym nie zależą od województwa, z którego pochodzili zdający, ale zdecydowanie zależą od typu szkoły oraz tego, czy była to szkoła publiczna czy niepubliczna.

Trzeba jednak dodać, że „zdawalność” matur zależy też w dużej mierze od realizowanego przez szkołę poziomu programu kształcenia. Może być on realizowany na poziomie rozszerzonym i jest to często praktykowane w liceach ogólnokształcących, co odzwierciedlają uzyskiwane nieco lepsze wyniki maturalne. Z kolei w pozostałych typach szkół program kształcenia jest realizowany zazwyczaj na poziomie podstawowym, a zatem może być to przyczyną uzyskiwania nieco gorszych wyników maturalnych z języka polskiego przez abiturientów techników i liceów profilowanych.

LITERATURA

- Drapella A. (2000), *Inżynieria niezawodnościowa*, Warszawa
- Drapella A. (2004), *Wnioskowanie statystyczne na podstawie skośności i kurtozy*, WPAP, Słupsk
- Jóźwiak J., Podgórski J. (2001), *Statystyka od podstaw*, PWE, Warszawa
- Mynarski S. (2003), *Analiza danych rynkowych i marketingowych z wykorzystaniem programu STATISTICA*, WAE, Kraków
- Mynarski S. (2006), *Analiza danych rynkowych i marketingowych z wykorzystaniem programu Excel*, WAE, Kraków
- Plucińska A., Pluciński E. (2000), *Rachunek prawdopodobieństwa. Statystyka matematyczna. Procesy stochastyczne*, WNT, Warszawa
- Stanisz A. (2007), *Przystępny kurs statystyki z zastosowaniem STATISTICA PL na przykładach z medycyny*, „Modele liniowe i nieliniowe”, t 2., StatSoft, Kraków

SUMMARY

The paper presents an analysis of the matriculation exam results of the Polish language at a basic level in Pomorskie and Kujawsko-Pomorskie voivodships in 2011. Values of the basic parameters characterizing the results are presented as well as a detailed analysis of variance between the different categories of data was made. The Author hypothesized that the matriculation exam results of the written Polish at a basic level can be described by a normal distribution. In order to verify the appropriate analysis was carried out and appropriate compliance testing. Ultimately, it turned out that there is no reason to reject the hypothesis for all categories of data concerned.

РЕЗЮМЕ

В статье был представлен анализ результатов экзаменов на аттестат зрелости по польскому языку на основном уровне в поморском и куявско-поморском воеводствах в 2011 г. Были указаны значения основных параметров характеризующих полученные результаты. Был также проведен подробный анализ дисперсии по отдельным категориям данных. Была представлена гипотеза, что результаты письменного варианта экзамена на аттестат зрелости по польскому языку на основном уровне можно охарактеризовать нормальным распределением. Для ее проверки были проведены соответствующие анализы и тесты на соответствие. В конечном итоге показалось, что нет никаких оснований отвергнуть гипотезу для всех рассматриваемых категорий данных.