

Analiza i interpretacja wyników

Elżbieta Ostaficzuk

EGZAMINU DONIOSŁEGO

*Wolno dojrzewa człowiek. Wiele mu trzeba męstwa,
aby się nie dać wiatrowi, i rosnąć w górę i w głąb,
by tryumfalną zielenią zaszumieć w końcu jak dąb
zwycięstwa...*

K. I. Gałczyński

Wyniki egzaminów zewnętrznych – gimnazjalnego i maturalnego, a także sprawdzianu w szóstej klasie szkoły podstawowej – są wykorzystywane do podejmowania ważnych decyzji względem uczniów lub szkoły¹. Ważne decyzje dla uczniów, to – na przykład – wybór szkoły lub kierunku studiów. Wprawdzie wyniki egzaminów maturalnych z przedmiotów wybranych przez ucznia w sposób znaczący zależą od jego aspiracji życiowych, to jednak analiza i interpretacja wyników egzaminów doniosłych również w tym przypadku może być cenną wskazówką do pracy dydaktycznej z uczniami, którzy dopiero będą przystępować do egzaminów.

Okręgowe Komisje Egzaminacyjne przekazują do szkół podstawowych wyniki sprawdzianu przeprowadzonego w szóstej klasie zawierające:

- **indywidualne wyniki sprawdzianu**, to znaczy sumę punktów za wszystkie zadania (tak zwany surowy wynik testowania – SWT) oraz liczby punktów uzyskane za zadania sprawdzające

¹ *Egzamin doniosły* – według Bolesława Niemierki (*Ocenianie szkolne bez tajemnic*, Warszawa, WSiP, 2002) – to egzamin, w którym znaczenie informacji o wyniku uczenia się jest większe niż znaczenie komentarza dydaktycznego, a więc sumowanie osiągnięć przeważa nad ich kształtowaniem.

umiejętności opisane w standardach wymagań egzaminacyjnych²;

- **tabele z wynikami szkoły** i poszczególnych klas w kontekście gmin, powiatów, województwa;
- *Sprawozdanie ze sprawdzianu w szóstej klasie*³, to znaczy podstawowe informacje o organizacji sprawdzianu, o ogólnych wynikach sprawdzianu wraz z charakterystyką arkusza – kartoteką zestawu zadań oraz **analizą łatwości zadań** zawartych w arkuszu.

Okręgowe Komisje Egzaminacyjne przekazują do gimnazjów wyniki egzaminu zawierające:

- **indywidualne wyniki gimnazjalistów**, to znaczy sumę punktów za wszystkie zadania (SWT) oraz liczby punktów uzyskane za zadania sprawdzające umiejętności opisane

² Standardy wymagań na sprawdzianie w szóstej klasie szkoły podstawowej, na egzaminie gimnazjalnym z zakresu przedmiotów humanistycznych oraz z zakresu przedmiotów matematyczno-przyrodniczych oraz standardy wymagań na egzaminach maturalnych z poszczególnych przedmiotów są opublikowane, między innymi, na stronach Centralnej Komisji Egzaminacyjnej – www.cke.edu.pl

³ Raporty OKE w Warszawie na temat wyników egzaminów doniosłych to – między innymi – *Sprawozdanie ze sprawdzianu w szóstej klasie szkoły podstawowej* oraz *Sprawozdanie z egzaminu gimnazjalnego*. 2005 województwo mazowieckie. OKE, Warszawa 2005.

w standardach wymagań egzaminacyjnych;

- **tabele z wynikami szkoły** i poszczególnych klas w kontekście gmin, powiatów, województwa;
- *Sprawozdanie z egzaminu gimnazjalnego*, to znaczy podstawowe informacje o organizacji egzaminu gimnazjalnego, o ogólnych wynikach egzaminu w części humanistycznej i matematyczno-przyrodniczej wraz z charakterystyką arkuszy egzaminacyjnych – kartotekami zestawów zadań oraz **analizą łatwości zadań** zawartych w arkuszach.

Okręgowe Komisje Egzaminacyjne przekazują szkołom ponadgimnazjalnym wyniki egzaminu zawierające:

- **indywidualne wyniki abiturientów** ze wszystkich przedmiotów maturalnych zdawanych w zakresie podstawowym lub rozszerzonym, to znaczy sumę punktów za wszystkie zadania arkusza (SWT);
- **tabele z wynikami szkoły** i poszczególnych klas w kontekście gmin, powiatów, województwa;
- *Sprawozdanie z egzaminu maturalnego*⁴, to znaczy podstawowe informa-

⁴ *Sprawozdanie z egzaminu maturalnego. Analiza wyników egzaminu maturalnego. Język polski* (Ze-

cje o organizacji egzaminu maturalnego, o ogólnych wynikach egzaminu oraz charakterystykę arkuszy egzaminacyjnych – kartoteki zestawów zadań wraz z **analizą łatwości zadań** zawartych w arkuszach.

Bezpośrednie rozpatrywanie nieusystematyzowanych w żaden sposób wyników egzaminów zewnętrznych nie jest działaniem ani sensownym, ani nawet możliwym. Konieczna jest klasyfikacja, opis i zastosowanie reguł prawomocnego wnioskowania. Statystyka – dziedzina wiedzy zajmująca się właśnie zbieraniem, klasyfikowaniem, opisem oraz interpretacją danych uzyskiwanych w badaniach eksperymentalnych – dostarcza odpowiedniego warsztatu pojęciowego i praktycznej metodologii. Dzięki temu statystyka umożliwia poznanie bogactwa informacji zawartych w wynikach egzaminów zewnętrznych. Jest to dziedzina wiedzy może trudna, ale jednak nie tajemna!

W prezentowanym artykule – w celu ułatwienia wydobycia interesujących informacji i pełniejszego wykorzystania wyników egzaminów – podjęto próbę przedstawienia podstawowych pojęć statystycznych. W części pierwszej, poświęconej analizie indywidualnych wyników egzaminacyjnych, przedstawiono różne sposoby prezentacji danych empirycznych (1.1) oraz opisu ich struktury, a mianowicie – miary tendencji centralnej (1.2), a także miary rozproszenia (1.3). Część druga jest poświęcona interpretacji wyników egzaminu zewnętrznego, w tym interpretacji wyników pod kątem poziomu kompetencji egzaminowanego zespołu uczniów (2.1). Prognozowanie rozwoju uczniów, na podstawie wyników egzaminów zewnętrznych, stanowi kolejne ważne wyzwanie badawcze stojące przed środowiskiem oświatowym⁵.

szyt 1); *Języki obce* (Zeszyt 2); *Historia, Historia sztuki, Historia muzyki, Wiedza o społeczeństwie, Wiedza o tańcu* (Zeszyt 3); *Matematyka, Fizyka i astronomia, Informatyka* (Zeszyt 4); *Biologia, Chemia, Geografia* (Zeszyt 5). 2005 województwo mazowieckie. OKE Warszawa 2005.

⁵ W niniejszym numerze „Meritum” opisany został polski projekt badawczy nad edukacyjną wartością

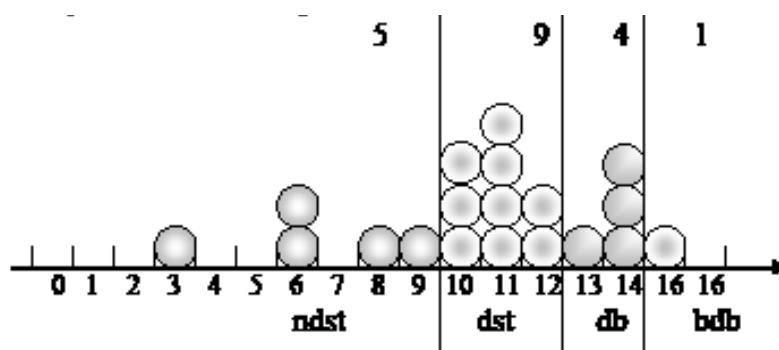
1. Analiza indywidualnych wyników egzaminacyjnych

1.1. Prezentacja danych empirycznych

Zbieranie, klasyfikowanie, opis oraz interpretacja danych uzyskiwanych w badaniach sondażowych oraz eksperymentalnych stanowią główne zagadnienia, którymi zajmuje się statystyka opisowa⁶.

Dawniej statystycy poświęcali wiele czasu na mozolny i systematyczny spis zbiorów osób lub rzeczy. Napoleon Bonaparte uważał, że *Statystyka to buchalteria rzeczy*⁷. Jeszcze w latach siedemdziesiątych ubiegłego wieku sporządzanie rozkładu wyników testowania wymagało biegłości rzemieślniczej. Zalecane wówczas działania ułatwiające prezentację wyników punktowych testowania i zamianę ich na stopnie szkolne są przedstawione na rys. 1. Indywidualne wyniki uczniowskie – zilustrowane w postaci kuleczek – grupowano w określonych przedziałach punktowych. W ten sposób uzyskiwano automatyzację zliczania i przyporządkowywania wynikom stopni szkolnych.

Rysunek 1. Rozkład wyników testowania⁸



Rozkład wyników testowania można było przedstawić również za pomocą tabelki, gdzie do zliczania poszczególnych klas wyników posługiwano się ukośnymi kreskami.

Tabela 1. Rozkład liczebności wyników testowania⁹

Granice klas	Środki klas (x)	Kreski	Liczebności klas (f)	Skumulowane liczebności klas
16–19	17,5	/	1	1
20–23	21,5	///	3	4
24–27	25,5	///	3	7
28–31	29,5	//	2	9
32–35	33,5	/	1	10
			$n=10$	

dodaną, czyli mierzaniem jakości pracy szkoły na podstawie wyników egzaminów zewnętrznych.

⁶ G. A. Ferguson, Y. Takane, *Analiza statystyczna w psychologii i pedagogice*, Wyd. Naukowe PWN, Warszawa 1997, s. 24.

⁷ A. Dąbrowski, *Statystyka – kilka ważnych przykładów*, MATEMATYKA 1'98, s. 15.

⁸ B. Niemierko, *ABC testów osiągnięć szkolnych*, WSiP, Warszawa 1975, s. 67.

⁹ B. Niemierko, *Testy osiągnięć szkolnych. Podstawowe pojęcia i techniki obliczeniowe*. WSiP, Warszawa 1975, s. 50.

Indywidualne wyniki egzaminacyjne (SWT) stanowią materiał statystyczny, który należy uporządkować według określonych cech, na przykład według przedmiotów egzaminacyjnych czy zakresu egzaminu. Uporządkowany – zazwyczaj rosnąco – materiał statystyczny jest prezentowany przeważnie w postaci tabeli i zwany **szeregiem rozdzielczym jedno- lub wielostopniowym**¹⁰.

Przykład 1. W szkole im. *Diagnostyki Edukacyjnej* do egzaminu przystąpiło $N=100$ uczniów. Za poprawne rozwiązanie wszystkich zadań zawartych w *Arkuszu egzaminacyjnym* jeden uczeń mógł osiągnąć maksymalnie 50 punktów. Wyniki wszystkich uczniów zostały uporządkowane rosnąco.

Tabela 2. Szereg rozdzielczy jednostopniowy – wyniki uzyskane na egzaminie przez uczniów szkoły im. *Diagnostyki Edukacyjnej*

SWT x_i	Liczebność f_i
(1)	(2)
0	1
1	2
2	3
3	2
4	2
5	2
6	3
7	2
8	2
9	2
10	2
11	2
12	2
13	2
14	2
15	3
16	3
17	2
18	3
19	3
20	3
21	2
22	0
23	0
24	0

(1)	(2)
25	4
26	0
27	0
28	0
29	5
30	0
31	4
32	6
33	2
34	3
35	2
36	2
37	2
38	2
39	2
40	2
41	0
42	2
43	0
44	0
45	0
46	2
47	2
48	3
49	3
50	2

Szereg rozdzielczy wielostopniowy przyporządkowuje klasom (grupom) obserwowanych wyników odpowiednie liczebności tych klas. Liczbę klas (k) ustala się tak, aby $5 \leq k \leq 20$. Przy ustalaniu klas można skorzystać z reguły:

$$k \approx \sqrt{N}$$

lub w przypadku licznych zbiorów obserwowanych wyników:

$$k \approx \sqrt{\frac{N}{2}}$$

Można też, ustalając liczbę klas, posłużyć się tabelką:

N	10	20	50	100	500	1000	10000
k	6	7	8	10	13	15	20

Długość przedziału klasowego (d) ustala się wykorzystując rozstęp (R) badanych wyników:

$$R = x_{max} - x_{min}$$

$$d \approx \frac{R}{k}$$

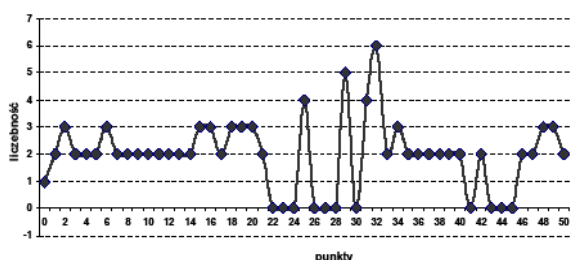
Tabela 3. Szereg rozdzielczy wielostopniowy – wyniki uzyskane na egzaminie przez uczniów szkoły im. *Diagnostyki Edukacyjnej*

Lp.	SWT	Liczebność f_i
1	0 – 4	10
2	5 – 9	11
3	10 – 14	10
4	15 – 19	14
5	20 – 24	5
6	25 – 29	9
7	30 – 34	15
8	35 – 39	10
9	40 – 44	4
10	45 – 50	12

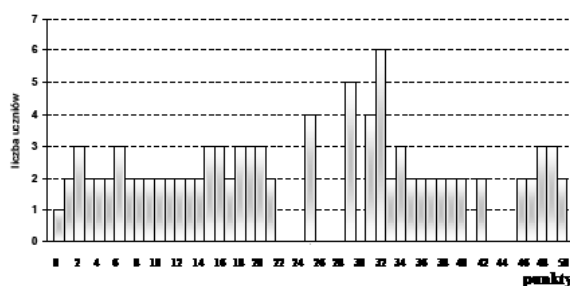
Dane statystyczne można również przedstawiać **graficznie** za pomocą krzywej liczebności lub histogramu.

¹⁰ Cz. Lewicki, Zbiór zadań ze statystyki dla pedagogów. Wydawnictwa Uczelniane WSP w Rzeszowie, 1989, s. 11.

Rysunek 2. Krzywa liczebności – wyniki uzyskane na egzaminie przez uczniów szkoły im. *Diagnostyki Edukacyjnej* (szereg rozdzielczy jednostopniowy z tab. 2.)

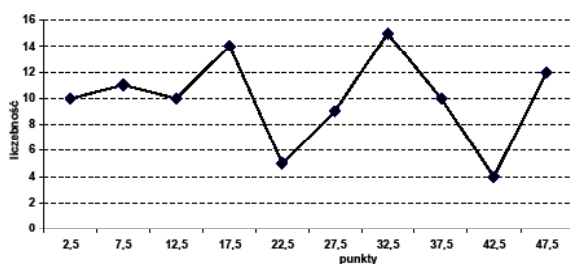


Rysunek 3. Histogram – wyniki uzyskane na egzaminie przez uczniów szkoły im. *Diagnostyki Edukacyjnej* (szereg rozdzielczy jednostopniowy z tab. 2.)

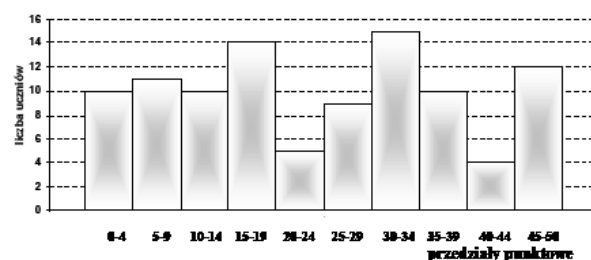


W przypadku krzywej liczebności przedstawiającej szereg rozdzielczy wielostopniowy na osi argumentów zaznacza się środki przedziałów klasowych.

Rysunek 4. Krzywa liczebności – wyniki uzyskane na egzaminie przez uczniów szkoły im. *Diagnostyki Edukacyjnej* (szereg rozdzielczy wielostopniowy z tab. 3.)



Rysunek 5. Histogram – wyniki uzyskane na egzaminie przez uczniów szkoły im. *Diagnostyki Edukacyjnej* (szereg rozdzielczy wielostopniowy z tab. 3.)



1.2. Miary tendencji centralnej

Uporządkowane wyniki egzaminu – w postaci szeregu rozdzielczego lub przedstawione graficznie – implikują dalsze analizy. Istotę rozważań stanowi problem – czy istnieją pewne wielkości, które w najlepszy sposób reprezentują obserwowany wycinek rzeczywistości? Niepokoję badawcze uspokaja statystyka, oferując miary tendencji centralnej, czyli takie wielkości, za pomocą których można natychmiast poznać **wielkość przeciętną** obserwowanego zjawiska.

Miary tendencji centralnej – to średnie klasyczne, wśród których na gruncie edukacji króluje średnia arytmetyczna, a także – średnie pozycyjne: mediana i modalna.

1. Średnia arytmetyczna dla grupy egzaminowanych uczniów:

- x – wyniki punktowe uczniów
- N – liczba egzaminowanych uczniów
- \bar{x} – średnia arytmetyczna

$$\bar{x} = \frac{\sum x}{N}$$

2. Mediana (Me) – środkowy wynik punktowy z uporządkowanego malejąco (lub rosnąco) ciągu wyników egzaminu. W przypadku parzystej liczby uczniów jest to średnia arytmetyczna dwóch środkowych wyników.

3. Modalna (Mo) – wynik punktowy występujący najczęściej wśród grupy egzaminowanych uczniów. Może jednocześnie istnieć kilka wielkości modalnych.

Tabela 4. Praktyczne zasady stosowania miar tendencji centralnej¹¹

Miara tendencji centralnej	Szczególnie przydatna, gdy	Nie należy się posługiwać, gdy
Średnia arytmetyczna	<ul style="list-style-type: none"> mają być wykonane inne obliczenia rozkład jest symetryczny względem środka 	<ul style="list-style-type: none"> jest bardzo mała próba losowa rozkład posiada na końcach klasy otwarte rozkład jest wielowierzchołkowy lub wybitnie asymetryczny
Mediana	<ul style="list-style-type: none"> jest mała próba losowa w rozkładzie występują klasy otwarte jest duża asymetria rozkładu 	<ul style="list-style-type: none"> mają być na tej podstawie wykonane dalsze obliczenia statystyczne
Modalna	<ul style="list-style-type: none"> jest wymagana jak najszybsza ocena wartości centralnej wystarczy przybliżona ocena wartości centralnej 	<ul style="list-style-type: none"> mają być na tej podstawie wykonane dalsze obliczenia statystyczne

1.3. Miary dyspersji

Posługiwanie się w opisie wyników egzaminów wyłącznie miarami tendencji centralnej jest nie tylko niewystarczające, ale może wręcz prowadzić do mylnych wniosków. Czy z równych średnich arytmetycznych można wnioskować o identycznych wynikach egzaminu różnych klas, szkół? Różnica rozkładów może przecież polegać na rozproszeniu (dyspersji) wyników. Informacja o stopniu rozproszenia wyników jest równie ważna, jak informacja o wielkości przeciętnej danego zjawiska. Różnice w zmianach wyników osiąganych przez uczniów są nawet bardziej interesujące niż średnie tych wyników!

Miary dyspersji są wskaźnikami jakościowymi. Informują o poziomie jednorodności badanej populacji. Im mniejsza wartość miar dyspersji, tym większa jednorodność badanych

¹¹ Praktyczne zasady stosowania miar tendencji centralnej określił Czesław Nowaczyk w *Podstawach metod statystycznych dla pedagogów*, PWN, Warszawa-Poznań 1985, s. 77-78.

wyników. W edukacji dyspersja jest niepożądana, na przykład gdy chodzi o poziom nauczania w różnych szkołach, o poziom osiągnięć uczniów w tej samej klasie.

Przykładem bezwzględnych miar dyspersji są wariancja oraz odchylenie standardowe. Wariancja i odchylenie standardowe są wielkościami mianowanymi (związanymi z tą samą jednostką, co wielkość badana). Powoduje to ograniczenie stosowania tych miar rozproszenia.

1. Wariancja (S^2) wyników egzaminu:

$$S^2 = \frac{\sum (x - \bar{x})^2}{N}$$

x – wyniki punktowe uczniów,
 N – liczba egzaminowanych uczniów,
 \bar{x} – średnia arytmetyczna.

2. Odchylenie standardowe (S) wyników egzaminu:

$$S = \sqrt{\frac{\sum (x - \bar{x})^2}{N}}$$

3. Rozstęp (R) wyników egzaminu:

$R = \text{max. wynik ucznia} - \text{min. wynik ucznia.}$

Przykład 2. Uczniowie szkoły im. *Diagnostyki Edukacyjnej* (przykład 1.) osiągnęli z egzaminu wyniki opisane w tab. 2. Miary tendencji centralnej oraz rozrzutu wyników są opisane w tab. 5.

Tabela 5. Miary tendencji centralnej i rozrzutu¹² wyników egzaminu egzaminacyjnych osiągniętych przez uczniów szkoły im. *Diagnostyki Edukacyjnej* (tab. 2.)

Miary tendencji centralnej		Miary rozrzutu	
Średnia arytmetyczna	24	Odchylenie standardowe	10,5
Mediana	23	Rozstęp	50-0
Modalna	32		

Wartość mediany została wyznaczona jako średnia arytmetyczna środkowych wyników: pięćdziesiątego i pięćdziesiąte-

¹² *Sprawozdanie z egzaminu maturalnego...*, op. cit., s. 128.

go pierwszego z uporządkowanego rosnąco zbioru wyników egzaminacyjnych (tab. 2.), to znaczy:

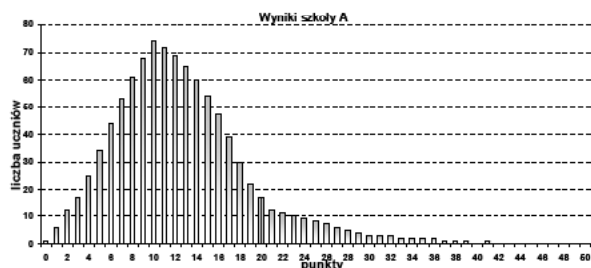
$$M_s = \frac{I_{39} + I_{41}}{2}$$

Średnia arytmetyczna w wysokości 24,0 oraz odchylenie standardowe równe 10,5 oznaczają, że prawie 70% wyników egzaminacyjnych uczniów mieści się w przedziale: $\langle 24,0-10,5; 24,0+10,5 \rangle$, czyli jest z zakresu od 13 do 35 punktów.

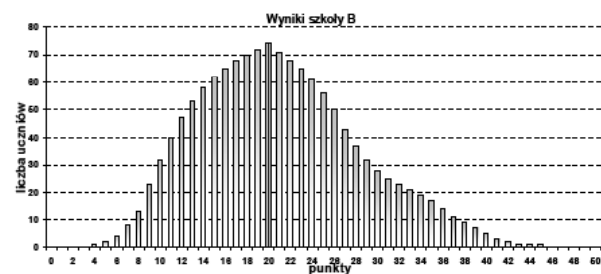
1.4. Porównywanie wyników egzaminu doniosłego

W egzaminie doniosłym uczestniczą jednocześnie uczniowie wielu szkół. Rozkłady indywidualnych wyników egzaminacyjnych uczniów z hipotetycznych szkół A, B i C przedstawiono na rys. 6. – 8.

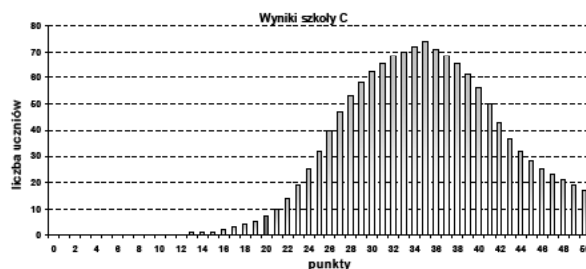
Rysunek 6. Rozkład wyników egzaminacyjnych uczniów szkoły A



Rysunek 7. Rozkład wyników egzaminacyjnych uczniów szkoły B



Rysunek 8. Rozkład wyników egzaminacyjnych uczniów szkoły C

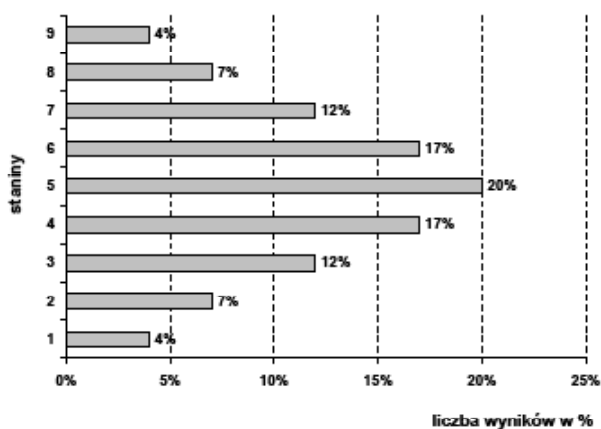


Z histogramu przedstawiającego osiągnięcia egzaminacyjne szkoły A (rys. 6.) widać, że dwudziestopunktowy wynik ucznia tej szkoły ma rangę wysoką. W szkole B uczeń z identycznym wynikiem punktowym (rys. 7.) plasuje się w grupie uczniów średnich. Natomiast w szkole C (rys. 8.) wynik dwudziestopunktowy – na tle osiągnięć uczniów tej szkoły – jest słaby. W jaki więc sposób można porównać wyniki tych uczniów?

Porównywanie indywidualnych wyników egzaminacyjnych, niezależnie od tego jak wygląda kształt histogramu, umożliwia **normalizacja histogramu wyników**. Jest to możliwe przy założeniu, że wyniki egzaminu – jak większość zjawisk i procesów zachodzących w otaczającym nas świecie – zależą od wielu czynników. Wynika stąd, że rozkład wyników egzaminu powinien mieć kształt krzywej normalnej (tak zwanej krzywej Gaussa). Wobec tego, by znormalizować wyniki egzaminu, należy je tak pogrupować, aby uzyskany w ten sposób histogram miał kształt najbardziej zbliżony do krzywej normalnej. Najczęściej przyjmuje się podział wyników egzaminu na 9 przedziałów¹³ zwanych staninami. Na rys. 9. przedstawiono koncepcję normalizacji. Kolejne staniny zawierają odpowiednio po 4, 7, 12, 17, 20, 17, 12, 7, 4 procent wyników egzaminu. Charakter dydaktyczny staninów wyrażają ich nazwy: najwyższy, bardzo wysoki... Oznacza to, że uczeń z wynikiem egzaminu plasującym się w IX staninie osiągnął najwyższy poziom umiejętności badanych na egzaminie.

¹³ B. Niemierko, *Pomiar wyników kształcenia*. WSiP, Warszawa 1999.

Rysunek 9. Normy staninowe. Wyniki testowania zgrupowane w 9 znormalizowanych przedziałach



Nazwy staninów:

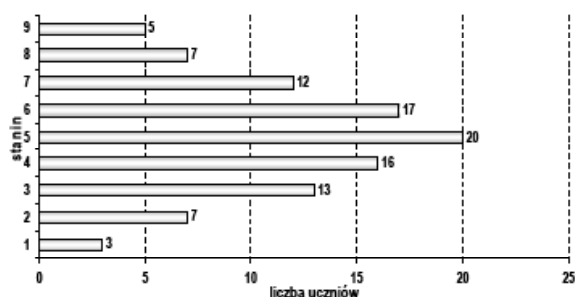
- 9 – najwyższy
- 8 – bardzo wysoki
- 7 – wysoki
- 6 – wyżej średni
- 5 – średni
- 4 – niżej średni
- 3 – niski
- 2 – bardzo niski
- 1 – najniższy

Przykład 3. Uczniowie szkoły im. *Diagnostyki Edukacyjnej* (przykład 1.) osiągnęli z egzaminu wyniki opisane w tab. 2. Normalizacja tych wyników za pomocą staninowej „dziwiałki” jest opisana w tab. 6., a rozkład staninowy wyników przedstawiono na rys. 10.

Tabela 6. Przyporządkowanie norm staninowych surowym wynikom egzaminacyjnym osiągniętym przez uczniów szkoły im. *Diagnostyki Edukacyjnej*

SWT x_i	Liczebność f_i	Liczba uczniów w przedziale	Stanin
(1)	(2)	(3)	(4)
0	1	3	I
1	2		
2	3	7	II
3	2		
4	2		
5	2		
6	3	13	III
7	2		
8	2		
9	2		
10	2		
11	2	16	IV
12	2		
13	2		
14	2		
15	3		
16	3		
17	2	20	V
18	3		
19	3		
20	3		
21	2		
22	0		
23	0		
24	0		
25	4	3	I
26	0		
27	0		
28	0		
29	5	7	II
30	0		
31	4		
32	6	17	VI
33	2		
34	3		
35	2		
36	2	12	VII
37	2		
38	2		
39	2		
40	2		
41	0		
42	2		
43	0		
44	0		
45	0	7	VIII
46	2		
47	2		
48	3	5	IX
49	3		
50	2		

Rysunek 10. Znormalizowany rozkład wyników egzaminacyjnych osiągniętych przez uczniów szkoły im. *Diagnostyki Edukacyjnej*



Objaśnienia:

Stanin	Przedział punktowy (max. 50 p.)
9 – najwyższy	49 – 50
8 – bardzo wysoki	46 – 48
7 – wysoki	36 – 45
6 – wyżej średni	31 – 35
5 – średni	18 – 30
4 – niżej średni	11 – 17
3 – niski	5 – 10
2 – bardzo niski	2 – 4
1 – najniższy	0 – 1

Analizę surowych wyników egzaminów doniosłych można przeprowadzić dla poszczególnych klas i całej szkoły. W przypadku wyników uczniów kończących szkołę – analiza będzie miała charakter sumujący; w przypadku wyników uczniów rozpoczynających dopiero naukę w szkole – analiza będzie miała raczej charakter kształtujący.

Analiza surowych wyników egzaminu może zawierać:

- rozkład wyników egzaminu;
- ustalenie miar tendencji centralnej i określenie, która z miar najlepiej reprezentuje charakter wyników osiągniętych w klasie bądź szkole;
- porównanie – za pomocą rozkładu staninowego – wyników uczniów poszczególnych klas lub ogólnie szkoły z rozkładem wyników w województwie i objaśnienie różnic w rozkładach;
- refleksje na temat zgodności ocen szkolnych z wynikami egzaminu doniosłego.

W ostatnich latach powszechna dostępność komputerów o znacznej mocy obliczeniowej, łatwość wprowadzania danych i otrzymywania gotowych wyników, spowodowała – paradoksalnie – utratę sensu badanych zjawisk i zmniejszyła odpowiedzialność badacza za właściwe rozumienie i stosowanie pojęć i metod statystycznych.

W szczególności wielką odpowiedzialnością moralną są obciążone raporty opisujące wyniki egzaminów zewnętrznych.

Zdarza się, że nawet obliczenia procentowe powodują utratę spostrzegawczości statystycznej. Łatwość wykonywania w arkuszu kalkulacyjnym obliczeń powoduje, że użytkownik komputera wielokrotnie komunikuje te same dane – poprzez wartości bezwzględne i wartości procentowe. Na przykład w tab. 7. przedstawiono kolumny: *Liczba uczniów* i *% uczniów*. Zdarza się również, że użytkownik komputera nie zwraca uwagi na dokumentację niepełnego zbioru danych – prawdopodobnie jest to wynik opcji formatowania liczb, na przykład zaokrąglania do jednego miejsca po przecinku. Tego typu błędy nieuwagi przedstawiono w tab. 7. oraz na rys. 11. – po zsumowaniu nie uzyskuje się wyników 100% uczniów uczestniczących w egzaminie.

W raportach CKE opisujących krajowe wyniki sprawdzianu 2006 *Pszczoly i miody* oraz krajowe wyniki egzaminów gimnazjalnych 2006 dostrzeżono już i zasygnalizowano problem niedokładności wyników przedstawianych w dokumentacji i opublikowano ostrzeżenie: *zaokrąglenia powodują, że suma odsetków w poszczególnych staninach nie zawsze jest równa 100%*¹⁴.

Tabela 7. Wyniki uczniowskie z egzaminu gimnazjalnego z zakresu przedmiotów matematyczno-przyrodniczych ($N=65590$)¹⁵

Stanin	Opis wyniku	Przedziały punktowe	Liczba uczniów	% uczniów
1	najniższy	0–9	2909	4,4
2	bardzo niski	10–12	4094	6,2
3	niski	13–16	7502	11,4
4	niżej średni	17–21	10537	16,1
5	średni	22–28	14538	22,1
6	wyżej średni	29–34	10408	15,9
7	wysoki	35–40	8647	13,2
8	bardzo wysoki	41–44	4279	6,5
9	najwyższy	45–50	2676	4,1

99,9%

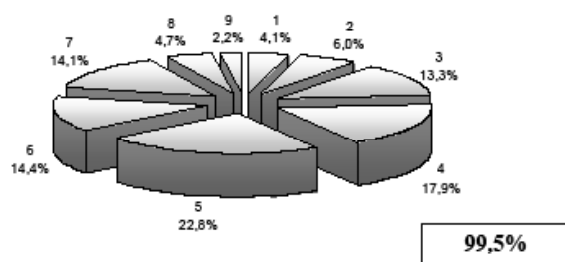
Oprogramowanie komputerowe automatycznie oferuje kilkadziesiąt form wykonywanych wykresów. W związku z tym zdarza się, że użytkownik wybiera wykres będący niewłaściwą ilustracją opisywanego zjawiska. Na przykład, do przedstawiania rangowanych wyników sprawdzianu na pewno nie należy stosować popularnego rozkładu kołowego, którego zasadniczą cechą jest podawanie wszystkich wyników na

¹⁴ Raport CKE: *Wyniki egzaminu gimnazjalnego 2006 (informacja dla uczniów, ich rodziców i nauczycieli)*. Warszawa, 13 czerwca 2006.

¹⁵ *Sprawozdanie z egzaminu gimnazjalnego. 2005 województwo mazowieckie*. OKE, Warszawa, s. 16.

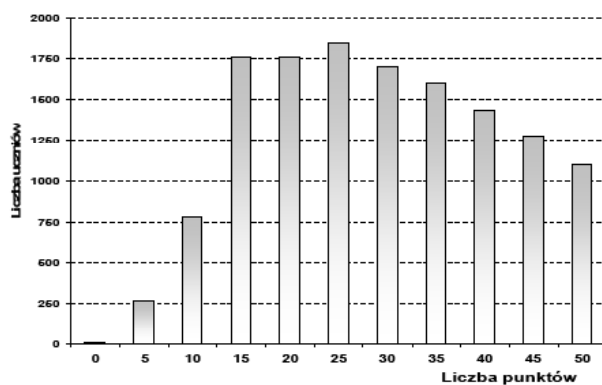
jednym, wspólnym poziomie. Rozkład kołowy nie umożliwia więc w żadnym przypadku przedstawiania wyników w uporządkowanej – rosnącej lub malejącej – kolejności!

Rysunek 11. Znormalizowany rozkład wyników ze sprawdzianu w szóstej klasie szkoły podstawowej ($N=61705$)¹⁶



Zdarza się także, że użytkownik komputera przedstawia rozkład wyników pogrupowanych w klasy za pomocą diagramu słupkowego i podaje niejasny, wprowadzający w błąd opis osi argumentów. Na przykład, z opisu osi argumentów przedstawionego na rys. 12. wynika, że na egzaminie uzyskano jedynie wyniki będące wielokrotnościami 5 punktów, to znaczy: 0, 5, 10... Prawdopodobnie w tym przypadku chodziło jednak o opis wyników pogrupowanych w przedziały: <0; 4>, <5; 9>, <10; 14>...

Rysunek 12. Rozkład wyników egzaminu maturalnego z matematyki ($N=21930$)¹⁷

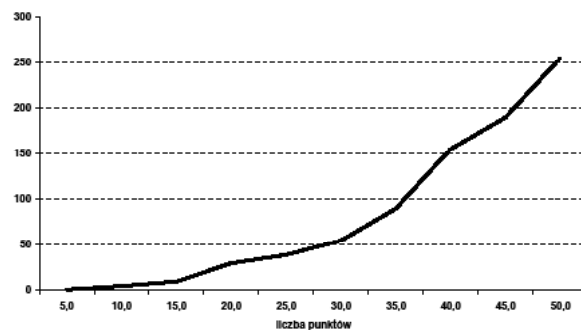


Punktowe wyniki egzaminów zewnętrznych są wartościami zmiennej dyskretnej (to znaczy, że nie każda liczba rzeczywista w zakresie od 0 do 50 może być wynikiem egzaminu – na przykład liczba 10,152 nie może być w żadnym przypadku wynikiem punktowym z egzaminu). Wobec tego przedstawianie rozkładu wyników za pomocą wykresu funkcji ciągłej jest chyba efektem nieuwagi użytkownika komputera.

¹⁶ Sprawozdanie ze sprawdzianu dla uczniów w szóstej klasie szkoły podstawowej. 2005 województwo mazowieckie. OKE, Warszawa, s. 11.

¹⁷ Sprawozdanie z egzaminu maturalnego. Analiza wyników egzaminu maturalnego. Zeszyt 4, OKE, Warszawa 2005, s. 16.

Rysunek 13. Rozkład wyników egzaminu maturalnego z języka francuskiego ($N=832$)¹⁸

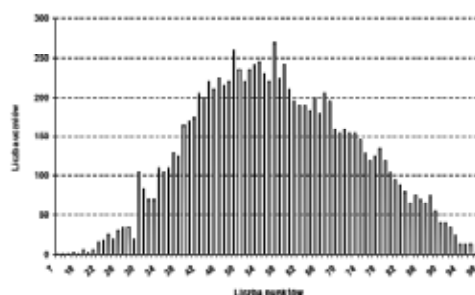


Kolejną grupę usterek edytorskich powoduje – zapewne – łatwość przekazywania lub kopiowania danych za pomocą komputera. Sytuację tego typu, powodującą niezgodność opisu danych w tabelce i na diagramie, przedstawiono w tab. 8. i na rys. 14. Wyniki punktowe: 52 i 53 zostały pominięte w zakresie rangowania w tab. 8, tymczasem z wykresu (rys. 14.) wynika, że istniała grupa uczniów, którzy uzyskali na egzaminie 52 lub 53 punkty.

Tabela 8. Normy staninowe dla wyników indywidualnych z historii ($N=10023$)¹⁹

Stanin	Opis wyniku	Przedziały punktowe	Przedziały procentowe	Procent zdających
1	najniższy	7–31	7–31	4,2
2	bardzo niski	32–37	32–37	6,1
3	niski	38–44	38–44	12,8
4	niżej średni	45–51	45–51	15,9
5	średni	54–60	54–60	21,1
6	wyżej średni	61–69	61–69	16,8
7	wysoki	70–77	70–77	11,2
8	bardzo wysoki	78–85	78–85	7,6
9	najwyższy	86–98	86–98	4,3

Rysunek 14. Rozkład wyników egzaminu maturalnego z historii²⁰



¹⁸ Sprawozdanie z egzaminu maturalnego. Analiza wyników egzaminu maturalnego. Zeszyt 2. OKE, Warszawa 2005, s. 128.

¹⁹ Sprawozdanie z egzaminu maturalnego. Analiza wyników egzaminu maturalnego. Zeszyt 3. OKE, Warszawa 2005, s. 22.

²⁰ Ibidem, s. 16.

Jak się okazuje, rzemieślniczy trud dawnego statystyka ostatnio często bywa zastąpiony błędami nieuwagi wynikającymi z automatycznego korzystania z oprogramowania komputerowego.

2. Interpretacja wyników egzaminu doniosłego

Zestaw zadań zawartych w arkuszu egzaminacyjnym jest tak skonstruowany, aby sprawdzał opanowanie wiadomości i umiejętności opisanych w standardach wymagań egzaminacyjnych i podstawie programowej. W *Sprawozdaniach OKE* są podane kartoteki poszczególnych zestawów zadań, zawierające dokładny opis sprawdzanych czynności. Szczegółowa analiza wyników osiągniętych za rozwiązanie poszczególnych zadań egzaminu bądź grup zadań tworzących podtesty może być przeprowadzona w odniesieniu do wymagań egzaminacyjnych. Najczęściej stosowaną do badania poziomu kompetencji miarą jakościową jest współczynnik łatwości:

$$p = \frac{\sum x}{Nk}$$

Przyjęto oznaczenia:

- p – łatwość zadania lub badanego podtestu
 N – liczba uczniów egzaminowanych
 k – maksymalna liczba punktów możliwych do uzyskania przez jednego ucznia za rozwiązanie zadania lub podtestu
 $\sum x$ – suma punktów uzyskanych za rozwiązanie zadania lub podtestu przez N uczniów.

2.1. Interpretacja wyników pomiaru poziomu kompetencji zespołu uczniów

Jeżeli wartość współczynnika łatwości p dla danej kompetencji (badanej w określonym zadaniu lub podteście) w rozpatrywanym zespole uczniów:

- **przekroczyła wartość 0,75**, to kompetencję uważamy za opanowaną zadowalająco;
- **mieści się w przedziale 0,30 – 0,75**, to kompetencja jest opanowana słabo;
- **była niższa niż 0,30**, to kompetencja jest całkowicie nieopanowana²¹.

W przypadku gdy interpretacja wyników egzaminu może jeszcze mieć wpływ na kształtowanie edukacji zespołu uczniów, wnioski diagnostyczne są warunkowane dwoma poziomami opanowania kompetencji: $p=0,75$ lub $p=0,30$. Mianowicie, jeżeli zespół uczniów opanował daną kompe-

tencję zadowalająco ($p \geq 0,75$), to należy ją rozwijać, stawiając przed uczniami trudne i złożone zadania. Jeżeli uczniowie opanowali kompetencję słabo ($0,3 \leq p < 0,75$), to ćwiczeniami utrwalającymi znajomość i sprawność posługiwania się daną kompetencją należy objąć właściwie cały zespół. W przypadku całkowitego nieopanowania kompetencji ($p < 0,3$) pracę nad jej opanowaniem należy rozpocząć od podstaw²².

Podsumowanie

Wyniki egzaminu gimnazjalnego w większym stopniu umożliwiają interpretację i formułowanie wniosków diagnostycznych niż wyniki egzaminów maturalnych. Wyniki gimnazjalne przesyłane do szkół zawierają bowiem wartości współczynników łatwości osiąganych przez uczniów poszczególnych klas i szkoły w zakresie umiejętności badanych standardami egzaminacyjnymi. Z drugiej strony – wyniki egzaminu maturalnego są bardziej „doniosłe” – w zasadzie decydują o przyszłości maturzysty. Surowe wyniki egzaminu maturalnego, którymi szkoły dysponują, uświadamiają konieczność uzyskania dodatkowych informacji, aby można było wyniki egzaminów wykorzystać do interpretacji osiągnięć szkoły, do planowania doskonalenia procesu nauczania.

Prognozowanie rozwoju uczniów na podstawie wyników egzaminów zewnętrznych umożliwi wyznaczenie tak zwanej edukacyjnej wartości dodanej:

edukacyjna wartość dodana = wynik uzyskany – wynik przewidywany²³

Edukacyjna wartość dodana stworzy możliwość określenia postępu względnego uczniów. Porównywanie edukacyjnej wartości dodanej wydaje się być bardziej przydatne, mniej krzywdzące dla szkół, niż porównywanie – na przykład – liczby egzaminów zdanych.

Komunikowanie i interpretacja wyników egzaminów doniosłych wywiera istotny wpływ na środowisko dydaktyczne szkoły. Ważna jest również interpretacja wyników egzaminów w ujęciu kontekstowym, uwzględniającym czynniki indywidualne, środowiskowe i pedagogiczne.

System egzaminów zewnętrznych to niezwykle istotna zmiana w reformowanej polskiej oświacie. Zmiana ważna, ale trudna, wymagająca czasu także na upowszechnienie umiejętności niezbędnych do prawidłowego korzystania z wyników egzaminów. ■

**Autorka jest nauczycielem konsultantem
w Mazowieckim Samorządowym Centrum
Doskonalenia Nauczycieli
Wydział w Warszawie**

²² M. Sobczak, *Jakościowa analiza wyników egzaminu zewnętrznego a jego funkcja kształtująca*, [w:] *Materiały z IV Ogólnopolskiej Konferencji z cyklu „Diagnostyka edukacyjna”*. Wyd. PANDIT, Kraków 2001 oraz M. Sobczak, *Testy sprawdzające z matematyki dla klasy I. Liceum ogólnokształcące, liceum profilowane, technikum. Zakres podstawowy i rozszerzony*. Wyd. NOWIK, Opole 2003.

²³ H. Szaleniec, *Zanim Twoi uczniowie zasiądą do matury*. WSiP, Warszawa 2005, s. 36 i nast.

²¹ M. Sobczak, *Wykorzystanie wyników oceniania zewnętrznego i wewnątrzszkolnego do podnoszenia jakości pracy szkoły*. OKE, Kraków 2005 www.oke.krakow.pl