



MIECZYŚLAW SŁOWIK

Politechnika Poznańska  
mieczyslaw.slowik@put.  
poznan.pl



MIKOŁAJ BARTKOWIAK

Politechnika Poznańska  
mikolaj.bartkowiak@put.  
poznan.pl

# Elementy statystycznej analizy wyników pomiarów na przykładzie badań wybranych cech mieszanek mineralno-asfaltowych

Pomiary od tysięcy lat są niezmiernie ważnym elementem działalności człowieka. Według [1] *pomiar to zbiór operacji mających na celu wyznaczenie wartości wielkości*. Wyznaczone w wyniku pomiaru wartości wielkości mierzonej są wyrażone za pomocą iloczynu liczby i jednostki miary. Badaczom najpewniej od samego początku wykonywania badań towarzyszyło pytanie o dokładność wykonanego oznaczenia. Pytanie dlaczego oznaczona wielkość ma właśnie taką wartość, a nie

np. o pół jednostki większą lub mniejszą. Takie wątpliwości doprowadziły zapewne do obliczania niepewności pomiaru. Dziś wynikiem oznaczenia nazywamy średnią wraz z przedziałem niepewności. Sama wartość średniej nie pozwala ocenić wiarygodności i użyteczności pomiaru [1]. Należy pamiętać, że *obliczenie niepewności pomiaru nie jest ani zadaniem rutynowym, ani zadaniem czysto matematycznym i zależy ono od szczegółowej wiedzy o naturze wielkości mierzonej i pomiaru* [1].

## Zbiory danych

Do prezentacji metod wnioskowania statystycznego posługują następujące zbiory (próby):

- zbiór nr I, wartości wymiarów planety Wenus, z publikacji [5], liczebność próby: 15 wartości, w tym wartości odbiegające: 2 lub 1 ( $P = 95\%$ ), w zależności od typu przeprowadzanego testu;
- zbiór nr II, wartości z publikacji [11], liczebność próby: 54, w tym wartości odbiegające: 3 ( $P = 95\%$ );
- zbiór nr III, pomiar ciężaru, z publikacji [13], liczebność próby: 11, zbiór według publikacji nie ma rozkładu normalnego;
- zbiór nr IV, losowe wartości, z publikacji [13], liczebność próby: 10, zbiór według publikacji ma rozkład normalny;
- zbiór nr V, wartości otrzymane z pomiaru gęstości objętościowej mieszanki mineralno-asfaltowej metodą B – SSD, dane własne wykorzystane w publikacji [2], liczebność próby: 40, w tym wartości odbiegające: 3;
- zbiór nr VI, wartości otrzymane z pomiaru modułu sztywności mieszanki mineralno-asfaltowej do warstwy wiążącej metodą rozciągania pośredniego w  $10^{\circ}\text{C}$ , dane własne

wykorzystane w publikacji [15], liczebność próby: 36, w tym wartości odbiegające: 3;

- zbiór nr VII, wartości otrzymane z pomiaru modułu sztywności mieszanki mineralno-asfaltowej do warstwy podbudowy metodą rozciągania pośredniego w  $20^{\circ}\text{C}$ , liczebność próby: 40, zbioru nie weryfikowano dotychczas pod względem występowania wartości odbiegających;
- zbiór nr VIII, wartości otrzymane z pomiaru modułu sztywności mieszanki mineralno-asfaltowej do warstwy podbudowy metodą rozciągania pośredniego w  $20^{\circ}\text{C}$ , liczebność próby: 80, zbioru nie weryfikowano dotychczas pod względem występowania wartości odbiegających;
- zbiór nr IX, wartości otrzymane z pomiaru modułu sztywności mieszanki mineralno-asfaltowej do warstwy podbudowy metodą rozciągania pośredniego, liczebność próby: 335, zbioru nie weryfikowano dotychczas pod względem występowania wartości odbiegających.

Zbiory od I do IV są to zbiory, które były wykorzystywane w wielu ważnych dla statystyki artykułach naukowych. Zbiory te zostały wykorzystane do potwierdzenia poprawności używanych algorytmów. Zbiory od V do IX to zbiory wartości pomiarów cech mieszanek mineralno-asfaltowych wykonanych przez autorów. Zbiory V, VII, VIII, XI to zbiory uzyskane w wyniku badania zestawu tych samych próbek wykonanych z tej samej mieszanki mineralno-asfaltowej.

## Testy statystyczne eliminujące wartości odbiegające

Podczas wykonywania oznaczenia można popełnić szereg błędów, które ogólnie dzieli się na: błędy przypadkowe, błędy systematyczne, błędy nadmierne [1]. Błędy nadmierne określane również jako błędy grube lub omyłki to błędy powstałe wskutek *jednorazowego wpływu istotnej przyczyny działającej przejściowo, występują tylko dla niektórych pomiarów (np. pomyłka przy odczycie lub zapisie wskazania przyrządu pomiarowego, niewłaściwy sposób pobrania, przechowywania lub przygotowania próbki do badań laboratoryjnych, itp.)*. Wartość obarczona błędem grubym jest z definicji, skrajną wartością w uporządkowanej według rosnących lub malejących wartości serii pomiarowej. Szczególnie łatwy jest do wykrycia i identyfikacji w odniesieniu do serii pomiarowych zawierających wyniki pomiarów dokonywanych w warunkach powtarzalności (oznaczenia takich samych jednostek otrzymane tą samą metodą, w tym samym laboratorium, przez tego samego operatora

z użyciem tego samego wyposażenia w krótkich odstępach czasu). [19].

Problem wątpliwych wartości oznaczeń w praktyce laboratoryjnej występuje stosunkowo często. Można przyjąć, iż za każdym razem kiedy wykonuje się oznaczenie, wartości uzyskane w wyniku jego przeprowadzenia powinny zostać poddane testom eliminującym wartości odbiegające. Na podstawie tych testów laborant może podjąć decyzję o włączeniu lub wykluczeniu poszczególnych wartości ze zbioru wartości oznaczeń. Według niektórych publikacji [19], w razie stwierdzenia, że przyczyną anormalnego wyniku była ewidentna pomyłka operatora lub zaburzenia w pracy aparatury, taki wynik powinien być natychmiast odrzucony. Niestety ustalenie zewnętrznej przyczyny uzyskania anormalnej wartości pomiaru jest trudne. W przypadku pomiarów laboratoryjnych wymaga to szczegółowej wiedzy o pomiarze, znacznego doświadczenia w jego wykonywaniu oraz szczegółowej obserwacji pomiaru podczas jego trwania, co jest zajęciem bardzo czasochłonnym. Jeżeli pomiary wykonywane są w znacznych odstępach czasu, większego znaczenia nabiera właściwa kalibracja przyrządów pomiarowych. Wyszukiwanie argumentów przemawiających za odrzuceniem wartości pomiaru jest wysoce ryzykowne, gdyż istnieje wówczas możliwość podjęcia decyzji tendencyjnych. Decyzja o odrzuceniu danej wartości bez przeprowadzenia stosownych testów statystycznych nosi charakter mniej lub bardziej subiektywny i może narazić badacza, który ją podejmuje, na zarzut niezgodnej ze sztuką prezentacji swoich wyników. *Sytuację pogarsza fakt, że taka nietypowa wartość może odzwierciedlać jakiś ważny rzeczywisty efekt fizyczny. Faktem jest, że wiele ważnych odkryć naukowych w przeszłości po raz pierwszy objawiło się jako wartość rażąco odbiegająca od pozostałych* [19].

Najlepszym sposobem postępowania w odniesieniu do wątpliwych wartości jest powtarzanie pomiaru niekiedy wielokrotnie. Jednak z różnych względów nie zawsze jest to możliwe, dlatego racjonalnym sposobem podejmowania decyzji dotyczących wyników wątpliwych jest wnioskowanie oparte na statystycznych metodach odrzucania wartości odbiegających.

W literaturze dostępnych jest wiele testów statystycznych dotyczących wykluczania wartości odbiegających. Postęp jaki dokonał się w ciągu lat w technice komputerowej spowodował, że obecnie raczej nie stosuje się już tablic statystycznych, gdyż nie ma większego problemu z wyliczeniem wartości krytycznych ze wzoru, o ile został podany. Do przeprowadzania testów służą dedykowane temu programy statystyczne, bądź napisane własnoręcznie skrypty programów obliczeniowych. Po przeanalizowaniu części literatury, autorzy na podstawie własnych doświadczeń wybrali kilka testów eliminujących wartości odbiegające, podając też wzory na wartości krytyczne do każdego z testów. Wszystkie zaprezentowane testy eliminujące wartości odbiegające zakładają, że rozkład prawdopodobieństwa wartości uzyskanych z pomiarów powinien być rozkładem normalnym.

#### 1) Test Grubbsa [5, 16, 17]

Frank E. Grubbs w swoich publikacjach, np. [5], przedstawił wiele rodzajów testów, w tej pracy przedstawiono najpopularniejszy z nich. W zbiorze danych występuje wartość odbiegająca jeśli:

$$G > \lambda_G \quad [1]$$

$$\frac{\max |x_i - m|}{SD} > \frac{t \cdot (n-1)}{\sqrt{(t^2 + n-2) \cdot n}} \quad [2]$$

gdzie:  $x_i$  – i-ta wartość uzyskana z pomiaru,  
 $n$  – liczba wartości w zbiorze,  
 $m$  – średnia arytmetyczna,  
 $SD$  – odchylenie standardowe (metoda „n-1”),  
 $t$  – wartość rozkładu t-Studenta dla liczby swobody  $n-2$  i poziomu istotności  $p = \frac{\alpha}{2n}$  (test dwustronny).

#### 2) Test Rosnera [11, 9, 17]

Test Rosnera, znany bardziej jako *extreme studentized deviate test* (ESD), jest modyfikacją testu Grubbsa. Test Rosnera można jednak wykonywać iteracyjnie analizując kolejno najbardziej odbiegające wartości, natomiast w teście Grubbsa należy a priori określić liczbę wątpliwych wartości. W zbiorze danych występuje wartość odbiegająca jeśli:

$$R_i > \lambda_{Ri} \quad [3]$$

$$\frac{|x_i - m_i|}{SD_i} > \sqrt{\frac{t \cdot (n-1)}{\sqrt{(t^2 + n-i-1) \cdot (n-i+1)}}} \quad [4]$$

Wartości średniej arytmetycznej i odchylenia standardowego oblicza się według wzorów, które wyłączają wartość wątpliwą:

$$m_i = \frac{1}{n+1-i} \sum_{j=1}^{n+1-i} x_j \quad [5]$$

$$SD_i = \sqrt{\frac{1}{n-i} \sum_{j=1}^{n+1-i} (x_j - m_i)^2} \quad [6]$$

#### 3) Test Hampela [8, 9]

Zaletą testu Hampela jest łatwość jego przeprowadzania, która wynika z faktu, iż w celu jego wykonania nie są potrzebne żadne wartości rozkładów prawdopodobieństwa. W celu przeprowadzenia testu należy wykonać następujące czynności:

- obliczyć wartość mediany  $Me$  ze wszystkich wartości  $x_i$ ,
- obliczyć bezwzględną wartość odchylenia  $H_i$  od mediany w przypadku każdej wartości, korzystając z wzoru:

$$H_i = |x_i - Me| \quad [7]$$

- obliczyć wartość mediany zbioru odchylenia  $H_i$  czyli  $Me_{H_i}$ ,
- sprawdzić obecność wartości odbiegających na podstawie kryterium:

$$H_i \geq 4,5 \cdot Me_{H_i} \quad [8]$$

Jeśli podane kryterium jest spełnione to wartość  $x_i$ , dla której obliczono  $H_i$  zostaje uznana za wartość odbiegającą.

#### 4) Kryterium Chauveneta [3]

Kryterium Chauveneta wykorzystuje rozkład normalny. Analizowana wartość  $x_i$  jest uznawana za wartość odbiegającą jeśli:

$$CH > z \quad [9]$$

$$\frac{\max |x_i - m|}{SD} > z \quad [10]$$

gdzie:  $z$  – wartość rozkładu normalnego dla poziomu istotności  $p = \frac{1}{2n}$ ,

#### 5) Test Grafa [19]

Analizowana wartość  $x_i$  jest uznawana za wartość odbiegającą jeśli:

$$F > 4 \quad [11]$$

$$\frac{|x_i - m_{n-1}|}{SD_{n-1}} > 4 \quad [12]$$

gdzie:  $m_{n-1}$  – średnia arytmetyczna z próby po wyłączeniu wartości wątpliwej,

$SD_{n-1}$  – odchylenie standardowe z próby po wyłączeniu wartości wątpliwej.

#### 6) Reguła trzysigmowa [19, 1]

Prawdopodobieństwo, że jakaś wartość uzyskana z pomiaru znajdzie się poza przedziałem ( $m - 3 \cdot SD$ ,  $m + 3 \cdot SD$ ) wynosi 0,0027. Można przyjąć, iż to prawdopodobieństwo jest tak znikomo małe, że jeśli jakaś wartość znajdzie się poza tym przedziałem to z dużym prawdopodobieństwem jest błędem grubym, czyli wartością odstającą.

## Eliminacja wartości odbiegających w zbiorach danych

Skoro istnieje wiele testów eliminujących wartości odbiegające, to można analizować ich skuteczność. Porównywaniu testów eliminujących wartości odbiegające poświęcono wiele publikacji, np. [9, 19]. Jednak w różnych publikacjach jako najskuteczniejsze przedstawiane są różne testy eliminujące wartości odbiegające. Według autorów wynika to z różnic jakie mogą wystąpić w algorytmie przeprowadzania testów, przyjmowaniu różnych poziomów istotności oraz w niektórych wypadkach błędom popełnionym przy wyznaczaniu wartości krytycznych.

Dlatego zaprezentowane wcześniej zbiory danych poddano opisanym testom eliminującym wartości odstające. Rezultat przeprowadzonych testów zamieszczono w tabeli 1.

W przypadku zbiorów I, II, IV otrzymano takie same rezultaty jak pomysłodawcy metod statystycznych w artykułach źródłowych [5, 11, 13], co potwierdza prawidłowość napisanych przez autorów skryptów z algorytmami. Na podstawie przedstawionych danych można stwierdzić, iż najwięcej wartości odbiegających wykryto stosując kryterium Chauveneta i test Hampela. Można również rozważyć, czy w wyniku zastosowania prezentowanych testów nie zostaną wyeliminowane poprawne wartości. Należy jednak podkreślić, iż wszystkie zaprezentowane testy są zaakceptowane przez środowisko naukowe i stosowane od wielu lat. Ponadto, autorzy stwierdzają, że w znanych im zbiorach danych, czyli zbiorach V i VI, testy nie wyeliminowały żadnej wartości otrzymanej dla poprawnych warunków badania, czyli we wszystkich wypadkach eliminowały wartości otrzymane dla niewłaściwych warunków przeprowadzania badania, co zostało opisane w publikacjach [2, 15]. W przypadku IV zbioru zgodnie z publikacją [14] nie wykryto wartości odbiegających. W najliczniejszym IX zbiorze kryterium Chauveneta wykazało dziesięć wartości odbiegających, natomiast test Hampela dwadzieścia trzy. W przypadku III zbioru kryterium Chauveneta wykazało trzy wartości odbiegające, a test Hampela tylko dwie. Dla analizowanych zbiorów wykonano także dwa testy normalności rozkładu (Lillieforsa oraz Shapiro-Wilka). W tabeli 2 przedstawiono rezultaty testów normalności rozkładu dla zbiorów zawierających wszystkie wartości, zaś w tabeli 3 dla zbiorów po wyeliminowaniu wartości odbiegających testem Hampela i kryterium Chauveneta.

W przypadku wszystkich zbiorów zawierających wartości odbiegające, z wyjątkiem I, testy nie potwierdziły rozkładu normalnego zbiorów (tablica nr 2). Po wyeliminowaniu wartości uznanych za odbiegające testy potwierdziły normalność rozkładu dla wszystkich zbiorów z wyjątkiem IX. Rezultaty testów na normalność rozkładu dla V i VI zbioru przedstawiono w publikacjach [2, 15]. W przypadku dwóch zbiorów o rozkładzie normalnym (I, III) kryterium Chauveneta i test Hampela wykazały odmienne rezultaty. Więcej wartości odbiegających wykazało kryterium Chauveneta, również wartość statystyki  $p$ , określająca stopień podobieństwa rozkładu zbioru do

Tabela 1. Rezultaty różnych testów eliminujących wartości odbiegające

Test	Zbiór (próba)								
	I	II	III	IV	V	VI	VII	VIII	IX
Liczność próby	15	54	11	10	40	36	40	80	335
Wartości uznane za odbiegające	2 lub 1	3	brak danych	0	3	3	brak danych	brak danych	brak danych
Test Grubbsa	1	0	1	0	2	0	2	2	0
Test Rosnera	1	3	1	0	2	3	2	2	10
Test Hampela	1	6	2	0	3	3	2	5	23
Kryterium Chauveneta	2	6	3	0	3	3	2	5	10
Test Grafa	0	0	1	0	2	0	2	2	0
reguła trzysigmowa	0	1	0	0	2	0	2	5	18

Tabela 2. Rezultaty testów normalności rozkładu w przypadku zbiorów zawierających wszystkie wartości

Parametr statystyczny	Zbiór (próbka)								
	I	II	III	IV	V	VI	VII	VIII	IX
Liczność próby	15	54	11	10	40	36	40	80	335
Średnia	0,018	2,321	172,00	449,50	2573	16190	9522	9526	9492
Odchylenie standardowe	0,551	1,183	24,95	82,08	22,24	1063	585,61	632,00	674,83
Rozstęp	2,410	6,260	88,00	280,00	124,00	4946	3047	4111	4225
RSD	3061	50,97	14,51	18,26	0,864	6,566	6,150	6,634	7,110
Normalność rozkładu według testu Shapiro-Wilka	W=0,940 P=38,0% TAK	W=0,906 P=0,04% nie	W=0,789 P=1,0% nie	W=0,943 P=62,1% TAK	W=0,828 P=0,00% nie	W=0,880 P=0,1% nie	W=0,831 P=0,00% nie	W=0,895 P=0,00% nie	W=0,909 P=0,00% nie
Normalność rozkładu według testu Lilliefors'a	TAK	nie	nie	TAK	nie	nie	nie	nie	nie

Tabela 3. Rezultaty testów normalności rozkładu w przypadku wybranych zbiorów po wyeliminowaniu wartości odbiegających

Parametr statystyczny	Zbiór (próbka)								
	I	I	II	III	III	VII	VIII	IX	IX
Zastosowanie testu:	Hampela	Chauveneta	Chauveneta Hampela	Hampela	Chauveneta	Chauveneta Hampela	Chauveneta Hampela	Chauveneta	Hampela
Liczność próby	14	13	48	9	8	38	75	325	312
Liczba wartości odrzuconych	1	2	6	2	3	2	5	10	23
Średnia	0,119	0,051	2,080	162,3	159,9	9622	9606	9563	9585
Odchylenie standardowe	0,401	0,322	0,702	9,747	6,813	381,0	420,7	543,8	457,7
Rozstęp	1,45	1,07	3,0	34	22	1396	1658	3236	2803
RSD	336,6	633,4	33,73	6,004	4,26	3,96	4,38	5,687	4,776
Normalność rozkładu według testu Shapiro-Wilka	W=0,956 P=66,2% TAK	W=0,973 P=92,2% TAK	W=0,977 P=47,2% TAK	W=0,958 P=84,5% TAK	W=0,982 P=99,7% TAK	W=0,961 P=19,8% TAK	W=0,971 P=8,29% TAK	W=0,973 P=0,00% nie	W=0,986 P=0,38% nie
Normalność rozkładu według testu Lilliefors'a	TAK	TAK	TAK	TAK	TAK	TAK	TAK	nie	nie

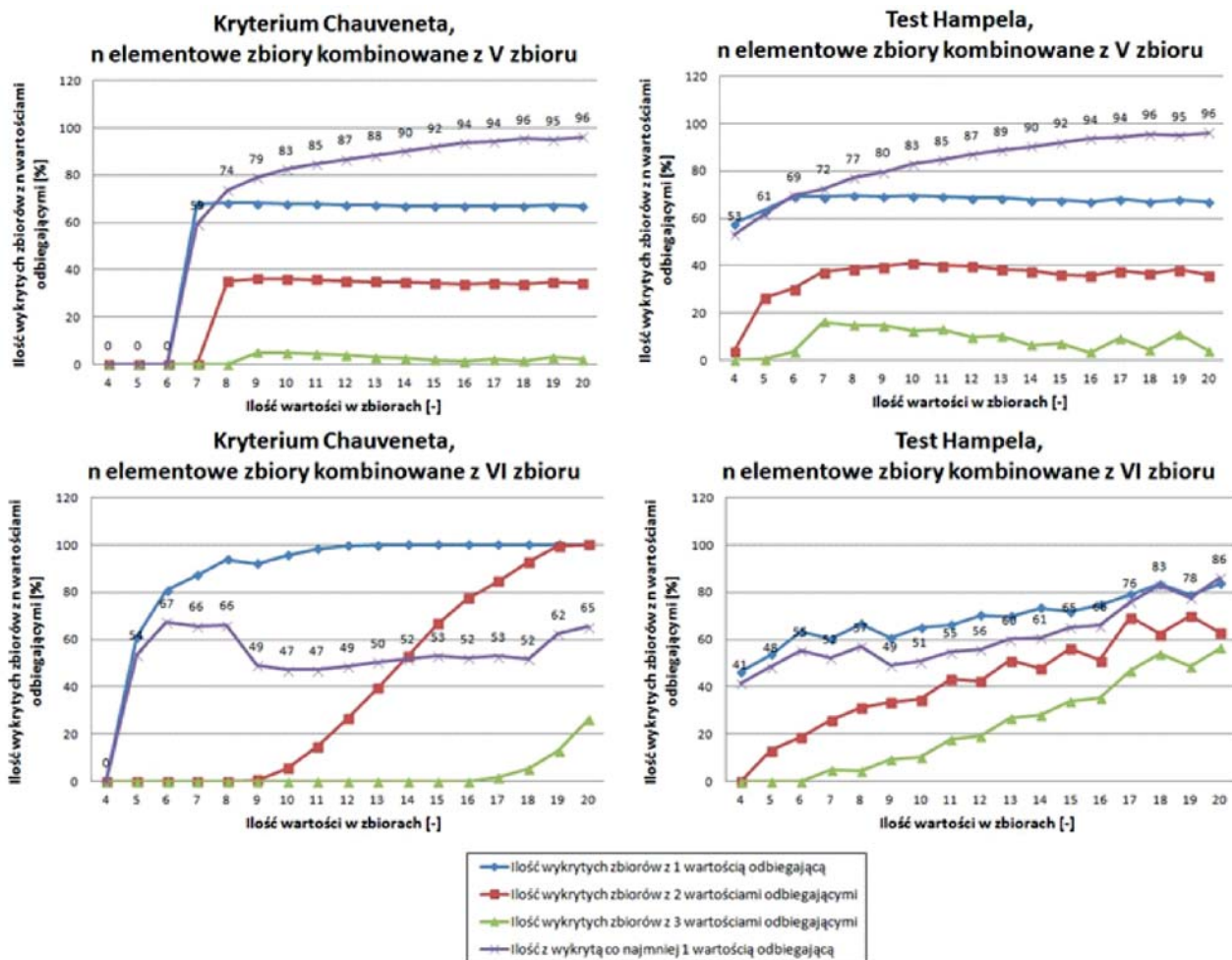
rozkładu normalnego, w przypadku zastosowania kryterium Chauveneta jest większa.

## Eliminacja wartości odbiegających w próbach o małej liczności

Istotnym zagadnieniem w przypadku prowadzenia badań laboratoryjnych, w tym także badań materiałów drogowych (zazwyczaj prowadzi się badania tych materiałów na bardzo małych próbach) jest ustalenie skuteczności testów eliminujących wartości odstające w próbach bardzo małych ( $n \leq 10$ ) i małych ( $10 < n \leq 30$ ). W tym celu spośród przedstawionych wcześniej zbiorów danych wybrano 2 zbiory (V i VI), z których losowano wartości za pomocą kombinacji bez powtórzeń. Każda próba wylosowana za pomocą kombinacji zawierała jedną, dwie lub trzy wartości odbiegające. Próby te były poddawane testom na wartości odbiegające. Liczba wartości w kombinowanych próbach wynosiła od 4 do 20. Dwa zbiory wybrane do kombinowania to zbiory wartości

otrzymane z przeprowadzonych badań, w trakcie realizacji których dołożono wyjątkowych starań do zachowania jednakowych warunków przeprowadzania badania oraz szczegółowo kontrolowano i notowano warunki przeprowadzania badań. Liczba zbiorów poddawanych testom eliminującym wartości odbiegające wahała się od 14562 zbiorów (dla  $n=4$ ) do 3948373 (dla  $n=8$ ).

Rezultaty testowania zbiorów z wartościami odbiegającymi pokazano na rysunku 1. Zbiór V jest zbiorem o mniejszym rozproszeniu wartości niż zbiór VI. Dlatego w zbiorze V kryterium Chauveneta wykrywa wartości odbiegające dopiero przy zbiorach zawierających siedem wartości, natomiast w zbiorze VI przy zbiorach zawierających pięć wartości. Zbiory z dwoma wartościami odbiegającymi kryterium Chauveneta wykrywa dopiero od zbiorów zawierających odpowiednio osiem i dziewięć wartości. Znacznie lepiej eliminuje wartości odstające test Hampela. Jedną lub dwie wartości odbiegające wykrywa już w czteroelementowych zbiorach. Trzy wartości odbiegające mogą być wykryte już w siedmioelementowym zbiorze.



Rys. 1. Liczba wykrytych zbiorów z wartościami odlegającymi wyrażona w procentach

## Normalność rozkładu

Jedno z podstawowych twierdzeń statystyki matematycznej, tzw. centralne twierdzenie graniczne (CTG), określa [1], że wypadkowa wielu niezależnych równorzędnych zmiennych losowych, niezależnie od rozkładu tych zmiennych ma rozkład zbliżony do normalnego. W pracy [1] Arendarski wskazuje, że jeśli zmienne składowe są niezależne, ich rozkład może nawet odstępować od normalnego w sposób znaczący, to już od czterech składowych można uznać zasadność stosowania CTG. Na podstawie tego twierdzenia należy stwierdzić, iż większość poprawnie wykonanych pomiarów, z wyjątkiem tak prostych pomiarów, jak np. pomiar długości, powinna mieć rozkład normalny.

W celu określenia rozkładu analizowanego parametru, należy się posłużyć testami zgodności, które służą do weryfikacji hipotezy dotyczącej zgodności pomiędzy rozkładem zbioru wartości w próbie i rozkładem teoretycznym. Szczególnym przypadkiem testów zgodności są tzw. testy na normalność rozkładu (ang. *normality tests*), w których rozkładem teoretycznym jest rozkład normalny. W literaturze można się spotkać z takimi testami na normalność rozkładu jak: test Shapiro-Wilka, test Kołmogorowa-Smirnowa, test Lillieforsa oraz inne. Wnioskując za [10] najlepszym testem (testem o największej mocy) jest test Shapiro-Wilka, co potwierdzają doświadczenia autorów.

Istotnym zagadnieniem związanym ze stosowaniem testów na normalność rozkładu, podobnie jak w przypadku testów eliminujących wartości odlegające, jest minimalna liczba wartości w próbie, od której testowanie zbioru jest skuteczne. Według niektórych publikacji normalność rozkładu można badać nawet od trzech wartości [12]. Jednak według dotychczasowych doświadczeń autorów, opartych także na badaniu normalności rozkładów zbiorów zaprezentowanych w rozdziale drugim, test Shapiro-Wilka można stosować do próby o liczności  $n=6$  elementów. Gdyż przy takiej liczbie elementów test Shapiro-Wilka wykazywał, iż rozkład analizowanych wartości nie jest rozkładem normalnym w zbiorach zawierających jedną wartość odlegającą.

Przed przystąpieniem do badań laboratoryjnych dobrze jest wykonać możliwie dużą serię pomiarów. Jeżeli badanie laboratoryjne nie jest zbyt czasochłonne, można wykonać dwie takie serie, przy możliwie najmniejszej i największej wartości badanej cechy. W przypadku przeprowadzenia wielu takich serii pomiarowych, można starać się o określenie równania określającego niepewność pomiaru w zależności właściwości badanych próbek lub parametrów badania, jak zostało to wykonane dla oznaczania gęstości objętościowej mieszanki mineralno-asfaltowej w normie PN-EN 12697-6:2012. Takie serie pomiarów pozwalają przede wszystkim ocenić źródła

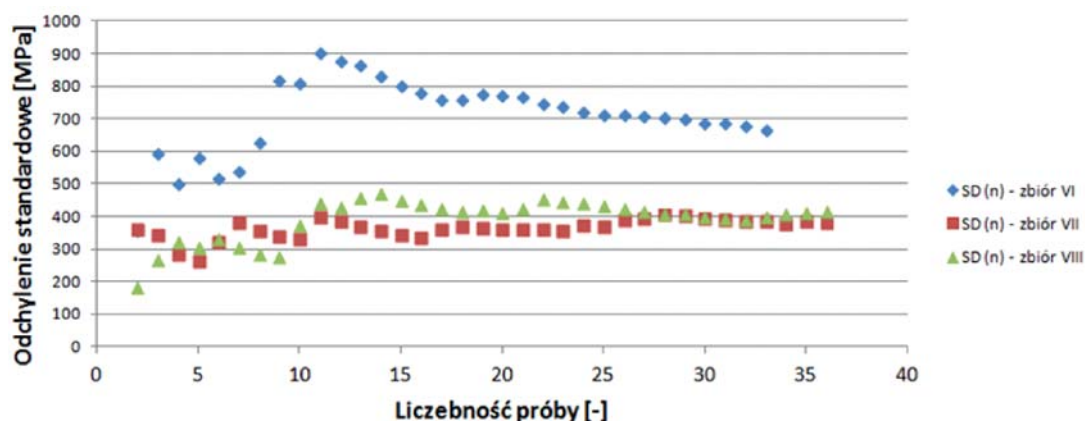
niepewności pomiaru oraz określić rozkład uzyskanych wartości, który jak już wspomniano najczęściej jest rozkładem normalnym. Właśnie do sprawdzania rozkładu wartości badanej wielkości należy wykorzystać testy zgodności. Testy zgodności zaleca się także stosować do każdego zbioru wartości otrzymanych z pomiaru, o ile ma dostateczną liczebność. Można je też stosować po testach eliminujących wartości odbiegające, aby sprawdzić czy po eliminacji próba ma rozkład normalny.

Dotychczasowe doświadczenia autorów potwierdzają, iż większość wartości uzyskanych z pomiarów ma rozkład normalny. Jeżeli w pomiarze wykonywanym metodą pośrednią, otrzymany zbiór wartości nie jest rozkładem normalnym, to może to świadczyć o niejednorodności próby statystycznej pobranej do badań lub zmiany warunków przeprowadzania oznaczenia. Takie wnioskowanie jest zgodne z twierdzeniami zawartymi w artykule [19], ponadto podkreśla się w nim, iż rozkłady wielomodalne (najczęściej bimodalne) są wynikiem połączenia dwóch lub więcej odrębnych statystycznie, jednorodnych zbiorów w jeden niejednorodny.

## Zmienność odchylenia standardowego

Odchylenie standardowe  $\sigma$  rozkładu obserwacji według [1] jest parametrem teoretycznym, którego wiarygodną wartość można oszacować na podstawie odpowiednio dużej ( $n > 30$ ) serii pomiarów tej samej wartości wielkości mierzonej. Dlatego w wielu książkach dotyczących statystyki [4] odchylenie standardowe z próby (mniejszej niż 30 wartości) oznacza się jako  $s$ , natomiast odchylenie standardowe uzyskane dla dużej próby zwane odchyleniem standardowym populacji jest oznaczone zwyczajowym symbolem  $\sigma$ .

W próbach do około piętnastu elementów wartość odchylenia standardowego wzrasta. Przy zwiększaniu liczebności próby wartość odchylenia standardowego zaczyna się stabilizować z trendem malejącym, by około trzydziestu elementów jego wartość nie zmieniała się więcej niż o 15% jego wartości. Jeżeli następują większe zmiany odchylenia można domniemywać, iż warunki pomiaru zmieniły się istotnie. Wartość średniej od liczebności próby równej czterem elementom, według dotychczasowych doświadczeń autorów, zazwyczaj nie zmienia się więcej niż o  $\pm 2\%$  wartości średniej.



Rys. 2. Zmienność wartości odchylenia standardowego

## Obliczanie niepewności pomiaru

Jak wielokrotnie podkreślano *obliczanie niepewności nie jest ani zadaniem rutynowym, ani zadaniem czysto matematycznym, zależy ono od szczegółowej wiedzy o naturze wielkości mierzonej i procedurze pomiaru oraz od właściwego zdefiniowania wielkości mierzonej (mezurandu)* [1]. Dlatego jeżeli operuje się niepewnościami, należy znać sposób i warunki ich wyznaczenia. Pierwszym krokiem procesu pomiarowego jest szczegółowe zdefiniowanie wielkości mierzonej (mezurandu) [1]. Na przykład przy wyznaczaniu niepewności pomiaru w pomiarach wykonywanych metodami nieniszczącymi należy szczegółowo określić czy oznaczono niepewność w przypadku jednej próbki badanej wielokrotnie, czy w przypadku wielu próbek wykonanych z tego samego materiału. Wiąże się to z definicją mezurandu, można np. oznaczać zawartość wolnych przestrzeni dla mieszanki mineralno-asfaltowej lub ułożonej warstwy asfaltowej, a można również oznaczać niepewność dla jednej próbki. Różnice między tak wyznaczonymi niepewnościami są niezaprzeczone i oczywiste, jednak czasami dążąc do maksymalnego zawężenia przedziału niepewności, są one pomijane.

Istotna jest również metoda wyznaczenia niepewności, tj. czy niepewność została wyznaczona metodą A czy B. Przy metodzie statystycznej (metoda A), jak podano w [1], określono dwa sposoby wyznaczania niepewności. Pierwszy z nich zakłada obliczenie niepewności korzystając ze znanej wartości odchylenia standardowego w populacji, które w tym wypadku można nazwać połączonym odchyleniem standardowym. W drugim niepewność oznacza się korzystając z rozkładu t-Studenta. Obie metody mają swoje wady i zalety, jednak biorąc pod uwagę ograniczony zakres dostępnego czasu częściej korzysta się z sposobu drugiego. Z doświadczeń autorów wynika, że dla małej liczby wartości niepewność obliczona za pomocą rozkładu t-Studenta jest większa, zaś dla powyżej dziesięciu wartości w próbce obliczone niepewności są zbliżone do siebie, różnice między nimi zazwyczaj nie przekraczają średnio 10% większej wartości.

## Podsumowanie

Istotnym zagadnieniem związanym z wykonywaniem pomiarów jest określenie minimalnej liczby oznaczeń, którą powinno się określać przed wykonywaniem badań, na których będziemy opierać jakieś tezy, a po badaniach wstępnych. W badaniach wstępnych można określić wartość odchylenia standardowego, dzięki której będzie można wyznaczyć, za pomocą wzorów znanych z literatury [4], niezbędną liczbę pomiarów, na podstawie tzw. maksymalnego błędu szacunku średniej, czyli akceptowalnej niepewności pomiaru.

Określając minimalną liczbę pomiarów warto wziąć pod uwagę również minimalną liczbę prób, jaka jest wymagana, aby przeprowadzić omówione wcześniej testy statystyczne. Test Hampela pozwala wyeliminować wartości odbiegające już w czteroelementowych zbiorach. Test Shapiro-Wilka na normalność rozkładu można, według autorów, stosować już od 6 elementów w próbie. W artykułach dotyczących pomiarów można znaleźć takie oznaczenia, w których minimalna liczba pomiarów została określona jako sześć [18]. Można również znaleźć takie pomiary, dla których minimalna liczba elementów w próbie została określona jako cztery, jak np. w normie PN-EN 12697-26:2012 „Mieszanki mineralno-asfaltowe – Metody badań mieszanek mineralno-asfaltowych na gorąco – Część 26: Sztynność”, według której były wykonywane pomiary, z których wartości testowane były w tym artykule. Istnieją również w drogownictwie badania, w których uzyskanie choćby czterech elementów w próbie jest bardzo czasochłonne, jak np. właściwości niskotemperaturowych, a konkretnie badanie TSRST wg normy PN-EN 12697-46:2012 „Mieszanki mineralno-asfaltowe – Metody badań mieszanek mineralno-asfaltowych na gorąco – Część 46: Pękanie niskotemperaturowe i właściwości w badaniach osiowego rozciągania”. W badaniach TSRST jedynym sposobem na zapewnienie jakiegokolwiek większej liczby oznaczeń wydaje się być tylko zakup wielostanowiskowego urządzenia pomiarowego. Jednak minimalna liczba oznaczeń wykonywanych w celach badawczych nie powinna być mniejsza od czterech, o ile jest to możliwe.

#### Bibliografia

- [1] *Arendarski J.*: Niepewność pomiarów. Oficyna Wydawnicza Politechniki Warszawskiej, Warszawa, 2013.
- [2] *Bartkowiak M., Słowik M.*: Niepewność pomiaru gęstości objętościowej mieszanki mineralno – asfaltowej. *Drogownictwo* 11/2015, s. 380-386.

- [3] *Firlej J.*: Wykorzystanie kryterium Chauveneta do oceny wyników badań, *Drogownictwo* 6-7/2008, s.201-204, Warszawa 2008.
- [4] *Greń J.*: Statystyka matematyczna. Modele i zadania., Państwowe Wydawnictwo Naukowe, Warszawa 1984.
- [5] *Grubbs F. E.*: Procedures for Detecting Outlying Observations in Samples. *Technometrics*, vol. 11, no. 1, 1969, p. 1-21.
- [6] Guide to the expression of uncertainty in measurement, Joint Committee for Guides in Metrology (JCGM), 2008.
- [7] Guide to the expression of uncertainty in measurement, Joint Committee for Guides in Metrology (JCGM), 2008.
- [8] *Konieczka P., Namieśnik J.*: Ocena i kontrola jakości wyników pomiarów analitycznych, Wydawnictwa Naukowo – Techniczne, Warszawa 2007.
- [9] *Manoj K., Senthamarai Kannan K.*: Comparison of methods for detecting outliers. *International Journal of Scientific & Engineering Research*, vol. 4, number 9, p. 709-714.
- [10] *Razali N.M., Wah Y.B.*: Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. *Journal of Statistical Modeling and Analytics*, 2011.
- [11] *Rosner B.*: Percentage Points for a Generalized ESD Many-Outlier Procedure. *Technometrics*, vol. 25, no. 2, 1983, p. 165-172.
- [12] *Royston P.*: Approximating the Shapiro-Wilk W-test for non-normality, *Statistics and Computing* 1992, Vol. 2, pp. 117-119.
- [13] *Shapiro S. S., Wilk M. B.*: An Analysis of Variance Test for Normality. *Biometrika*, vol. 52, no. 3/4, 1965, p. 591-611.
- [14] *Shapiro S. S., Wilk M. B.*: An analysis of variance test for normality (complete samples). *Biometrika*, Vol. 52, No. 3/4, Dec. 1965, pp. 591-611.
- [15] *Słowik M., Bartkowiak M.*: Calculation of Measurement Uncertainty for Stiffness Modulus of Asphalt Mixture. *Journal of Civil Engineering and Architecture*, vol. 9, number 11, p. 1325-1333.
- [16] *Słowik M.*: Wybrane zagadnienia lepkosprężystości drogowych asfaltów modyfikowanych zawierających elastomer SBS. Wydawnictwo Politechniki Poznańskiej, Poznań 2013.
- [17] Strona internetowa National Institute of Standards and Technology (Narodowy Instytut Standaryzacji i Technologii).
- [18] *Sztukiewicz R., Rydzewski P.*: Opracowanie statystyczne wyników w metodzie ultradźwiękowej przy małej ilości pomiarów. Konferencja naukowa nt: Diagnostyka konstrukcji nawierzchni lotniskowych, Warszawa, marzec 1985.
- [19] *Twardowski K., Traple J.*: Uwagi dotyczące wątpliwych wyników pomiarów. *Wiernictwo, Nafta, Gaz*, tom 23/2, 2006, s. 699-714.

## Serwis GDDKiA • Aktualności

### Złożono wnioski o wydanie ZRID na budowę S17 od Garwolina do granicy z województwem lubelskim

22 lipca do Delegatury Mazowieckiego Urzędu Wojewódzkiego w Siedlcach trafiły wnioski o wydanie zezwolenia na realizację inwestycji drogowej, czyli pozwolenia na budowę ekspresowej S17 na odcinku od Garwolina do granicy z województwem lubelskim. Wydanie decyzji przez wojewodę mazowieckiego umożliwi przejęcie nieruchomości pod budowaną drogę i rozpoczęcie procedury odszkodowawczej, a wykonawca będzie mógł rozpocząć prace.

Do wojewody złożono dwa wnioski o wydanie ZRID. Pierwszy dotyczy odcinka od końca obwodnicy Garwolina do Gończyc (o długości 12,2 km), a drugi od Gończyc do granicy województw (13 km). Decyzji spodziewamy się na przełomie roku. Oryginał doręczony zostanie wnioskodawcy, a pozostali zainteresowani dowiedzą się o niej w drodze obwieszczeń w Mazowieckim Urzędzie Wojewódzkim oraz w urzędach gmin, przez tereny których przebiega droga. Właściciele i użytkownicy wieczystości nieruchomości otrzymają zawiadomienia o wydaniu takiej decyzji na adres wskazany w ewidencji gruntów. Po 14 dniach od publicznego ogłoszenia informacji o wydaniu decyzji można będzie się od niej odwołać – tylko przez kolejnych 14 dni.

Decyzja ZRID wskaże nie tylko teren, na którym będą zrealizowane roboty związane z budową drogi ekspresowej. Stwierdzi również przejęcie z mocy prawa na rzecz Skarbu Państwa nieruchomości przeznaczonych na pas drogowy oraz zatwierdzi podział nieruchomości pod ten pas. Decyzja obejmie swoim zakresem nie tylko teren pod samą drogę, ale również skrzy-

żowania, rowy drogowe, skarpy, drogi gospodarcze, zbiorniki retencyjne czy też tzw. czasowe zajęcia pod przebudowę kolidujących instalacji. Wydanie decyzji ZRID umożliwi też rozpoczęcie procesu szacowania wartości nieruchomości przejętych pod budowę drogi, a później wypłatę odszkodowań. Przypomnijmy, że wysokość odszkodowania ustala wojewoda w postępowaniu administracyjnym, które kończy się wydaniem decyzji o ustaleniu odszkodowania. Generalna Dyrekcja Dróg Krajowych i Autostrad będzie wypłacać należne kwoty.

Umowę na zaprojektowanie i budowę S17 od końca obwodnicy Garwolina do granicy województw mazowieckiego i lubelskiego podpisano 22 września 2015 roku. Wykonawca opracował już niezbędną dokumentację do złożenia wniosku o wydanie decyzji ZRID, a teraz kontynuuje prace nad projektem wykonawczym. Po uzyskaniu decyzji ZRID rozpocznie roboty budowlane, które w okresach zimowych pomiędzy 15 grudnia i 15 marca będzie mógł zawiesić lub prowadzić, o ile nie będą przeszkadzały warunki atmosferyczne. Przewidywany termin zakończenia budowy to przełom kwietnia i maja 2019 roku. Wykonawcą jest firma Budimex, która budowę 25-kilometrowej trasy wyceniła na ok. 602 mln zł. Nawierzchnia drogi ekspresowej wykonana będzie w technologii betonu cementowego.

22-07-2016