



Seminarium
Głównego Urzędu Miar

Zastosowanie statystycznych
metod odpornościowych
przy opracowaniu wyników
porównań
międzylaboratoryjnych

referat wygłosi:
doc. dr inż. Zygmunt Warsza

3 marca 2017 r.

B. Zastosowanie metod odpornościowych w porównaniach międzylaboratoryjnych

- 1. Wprowadzenie**
 - 2. Metody statystyki odpornościowej**
 - 3. Istota badań międzylaboratoryjnych.**
 - 4. Zastosowanie dwu metod odpornych w porównaniach międzylaboratoryjnych do oceny precyzji metody pomiarowej. Przykład**
 - 4.1 Ocena metodą przeskalowanego odchylenia medianowego MAD_s**
 - 4.2 Ocena metodą odporną iteracyjną Hubera**
 - 4.3 Wnioski dotyczące przykładu**
- Podsumowanie**
- Literatura**
- Podziękowanie**

1. Wprowadzenie

W praktyce często wartości kilku obserwacji próbki pomiarowej istotnie odstają od pozostałych danych.

Nawet niewielka ich liczba może znacząco zmienić parametry statystyczne próbek o małej liczbie elementów.

W wielu przypadkach nie ma możliwości powtórzenia, lub uzupełnienia pomiarów.

Wartość i niepewność wyniku takich pomiarów, obliczone wg zasad przewodnika GUM mogą okazać się niewiarygodne.

Stosuje się różne testy statystyczne, np. test Grubbsa, do zidentyfikowania i usunięcia danych odstających (ang. *outliers*).

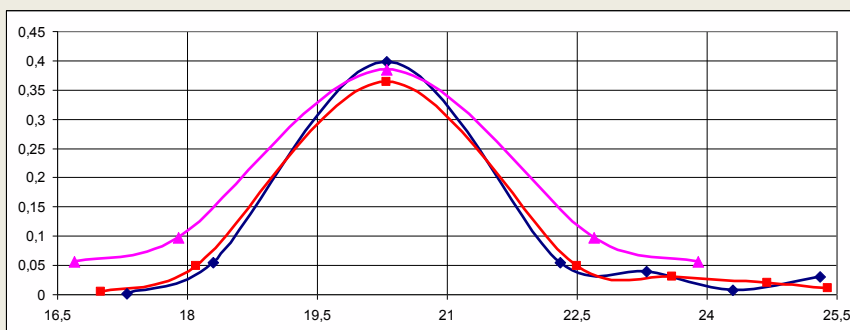
Jest to właściwe i skuteczne tylko dla dużych próbek.

Dla małych próbek - usunięcie kilku danych (1 - 3-ech) zmniejsza istotnie dokładność estymacji niepewności u_A

[wg GUM dod. E.1 $SD(u_A)/u_A \approx \{\sqrt{2(n-1)}\}^{-1}$ np. dla $n=4$ - 42%, $n=3$ - 52%]

TEŻA: Przy małej liczbie danych w próbce bardziej prawidłowe jest stosowanie do oszacowań metod odpornych na dane odstające.
Omówi się zastosowanie dwu metod odpornych

Rozkłady o tzw. „ciężkich ogonach”



- W praktyce często otrzymuje się rozkłady prawdopodobieństwa (jako przybliżenia histogramów) o obu bokach opadających wolniej niż w rozkładzie normalnym (ang. *heavy tails distributions* – rozkłady o „ciężkich ogonach”).
- Zwykle nie są to funkcje o charakterze wykładniczym.
- Właściwość ta występuje z reguły dla próbek o małej liczebności danych.
- W małych próbkach pojawia się też asymetria.
- Współczynnikiem charakteryzującym asymetrię jest skośność g (ang. *skewness*).

Metody przetwarzania danych próbek z wartościami odstającymi, czyli outlierami

Pogląd J. Tukey'a [5]:

„normalność rozkładu danych eksperymentalnych jest mitem i takiego rozkładu nie było w praktyce i nigdy nie będzie”.

Odnosi się to szczególnie do próbek o małej liczbie danych.

Powody powstania odstających wartości obserwacji:

- błędy wynikłe z nieprzestrzegania zasad prowadzenia doświadczenia,
- różnice dokładności dla różnych wartości danych,
- błędy i omyłki w odczycie przetwarzaniu wyników obserwacji,
- niesprawność użytych przyrządów pomiarowych.
- chwilowe wpływy zewnętrzne o dużej intensywności i wiele innych.

• 1. Metoda kontaminacji

• Założenia:

- parametry statystyczne warunków pomiarów, systemu pomiarowego i obiektu są niezmiennie.
- wartości danych w próbce są pobrane z populacji o rozkładzie normalnym, ale dane odstające, czyli outliery, są z innego rozkładu.

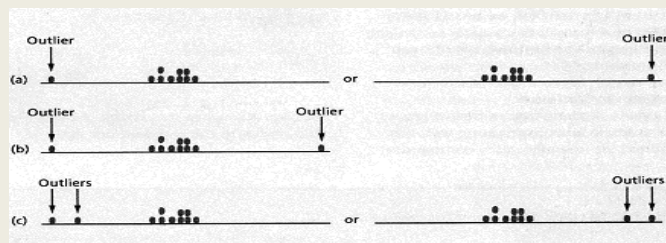
Outliery należy zidentyfikować przez testy statystyczne i usunąć.

• 2. Metody odporne na występowanie danych odstających, czyli outlierów

• Założenia:

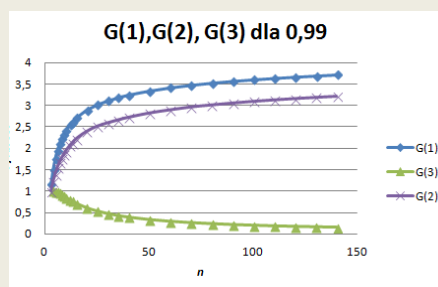
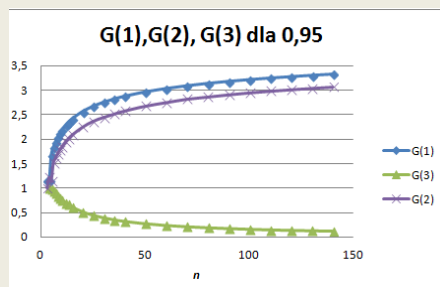
- w rzeczywistości występuje przestrzenno-czasowa zmienność warunków pomiaru oraz parametrów systemu i obiektu badań.
- dane próbki pochodzą z populacji o rozkładzie niegaussowskim.

1. Metoda klasyczna. Kryteria Grubbsa dla próbek o liczebności n



Trzy rodzaje punktów odstających (outlierów) wg testu Grubbsa

$$G_n \max = |x_n - \bar{x}| / s$$



Współczynnik skośności próbki g

Dla próbki o n elementach x_i z populacji generalnej X szacuje się współczynnik skośności g za pomocą jej momentów centralnych $m_2=s^2$ i m_3 , jako estymatorów momentów μ_2, μ_3 tej populacji.

W [3] i [4] dla skośności g próbki podano wzór

$$g = \frac{m_3}{s^3}$$

$$g = \frac{n \sum_{i=1}^n (x_i - \bar{x})^3}{(n-1)(n-2)s^3}$$

W programach statystycznych używa się też standaryzowanego współczynnika skośności próbki $SSKE = g/\sqrt{6/n}$ [4].

Dla $n > 150$ i populacji symetrycznych rozkład praktycznie jest normalny.

Spotyka się w literaturze nieco inne postacie wzoru dla współczynnika skośności g , ale różnice uwidaczniają się tylko dla bardzo małych próbek.

W statystyce stosuje się też różne nieklasyczne miary skośności, np. współczynnik Lorentza [4].

Wariancja współczynnika skośności

Wariancja współczynnika skośności g dla próbki o n elementach z populacji o rozkładzie normalnym

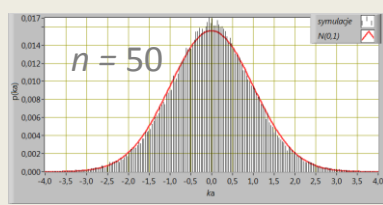
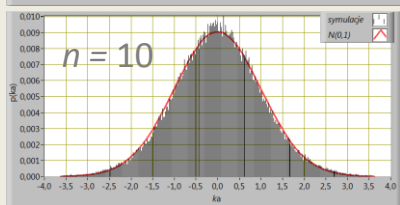
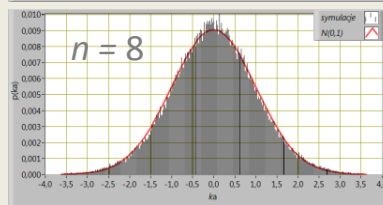
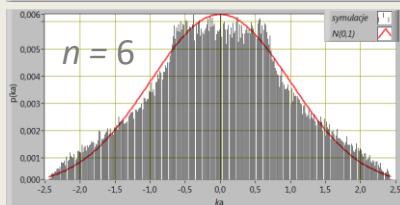
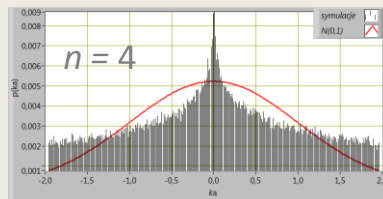
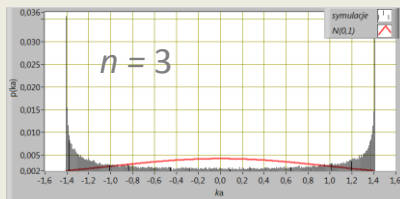
$$D(g) = \frac{6n(n-1)}{(n-2)(n+1)(n+3)}$$

Wariancja $D(g)$ próbek wg Smirnova [2]

$$D(g) = \frac{6(n-2)}{(n+1)(n+3)} = \frac{6}{n} \left[1 - \frac{12}{2n+7} + O\left(\frac{1}{n^3}\right) \right]$$

$O(\cdot)$ - człon resztkowy $1/n^3$

Histogramy skośności małych próbek o n elementach



Zależność parametrów współczynnika skośności g próbki (ang. skewness) od jej liczebności n

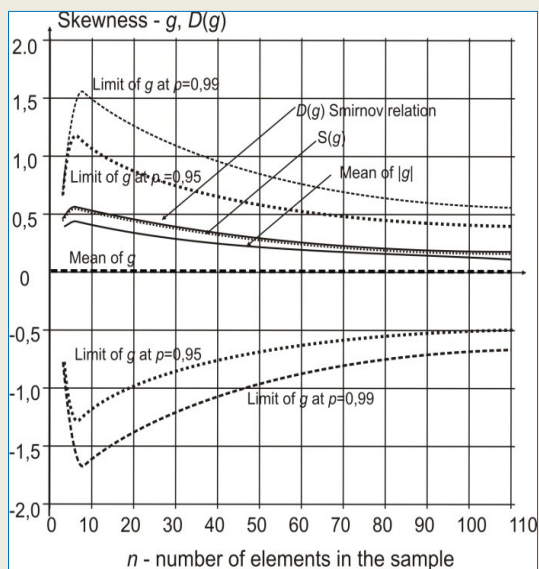
$$g = \frac{m_3}{s^3}$$

m_3 - trzeci moment centralny,
 s - odchylenie standardowe

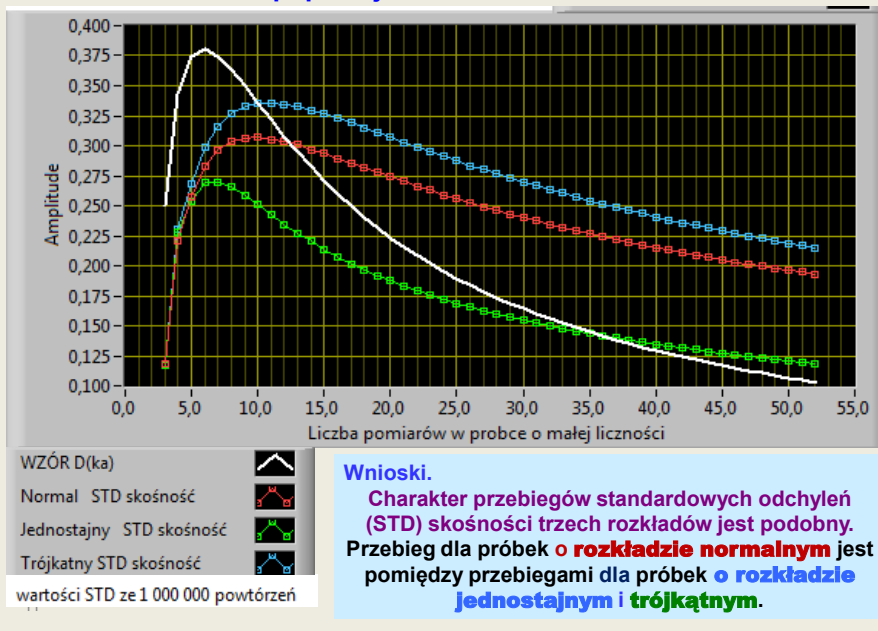
Średni moduł skośności $|g|$ i odchylenie standardowe $S(g)$ dla próbek z rozkładu normalnego o osiągają maksima dla bardzo małych n i maleją do zera dla $n \rightarrow \infty$.

Również ich rozkłady dla małych n znacznie odbiegają od normalnych.

Wstępują też pseudo-outliery.

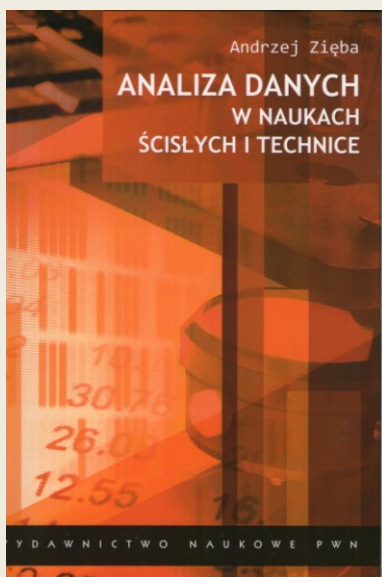


Porównanie zależności SD skośności od liczby elementów n próbek z populacji o kilku rozkładach



2. Metody statystyki odpornościowej

- Metody statystyki odpornościowej powstały w latach 1960 i rozwijają się nadal.
 - Metody te stosuje się już w różnych zadaniach przetwarzania danych, do kalibracji pomiarów wieloparametrowych, w specjalizowanych technikach pomiarowych, np. w chemometrii oraz w porównaniach międzylaboratoryjnych.
- Źródła wiedzy o metodach odpornych:**
- Istnieje wiele publikacji i pozycji książkowych w języku angielskim.
 - Niektóre metody odpornościowe są oprogramowane, m. in. w MatLab,
 - Międzynarodowe normy dotyczące badań laboratoryjnych ISO 5725-5, ISO 13528: 2009-01 (metody te nie trafiły jeszcze do przewodnika wyznaczania niepewności GUM i jego Suplementów !)
 - Książki autorów polskich zawierające rozdziały omawiające metody odporne:
 - Domański Cz., Pruska A. „Nieklasyczne metody statystyczne” PWE 2000 [6];
 - skrypt śp. J. Piotrowskiego z Politechniki Śląskiej i jego monografia z K. Kostyrko „Wzorcowanie aparatury pomiarowej” (2000 i wyd. nowe 2012) ,
 - A. Zięba „Analiza danych w naukach ścisłych i technice” PWN 2013
 - Publikacje autorstwa metrologów polskich:
 - D. Buchczik (Pol. Śl.) PAK 5/2008,
 - od 2012 r. artykuły E. Volodarskiego, L. Koshevoj i Z. Warsz (4 tematy, 15 art., w tym 8 po polsku) i rozdz. 11 w monografii Z. Warsz.
 - wiele międzynarodowych publ. angielskich Zespołu Chemometrii z Uniwersytetu Śląskiego
 - W języku rosyjskim:
 - przetłumaczona podstawowa pozycja J. Tukey’a z ubiegłego wieku,
 - kilka opracowań autorów rosyjskich.
 - E. Volodarski i L. Kosheva: „Techniczne aspekty akredytacji laboratoriów badawczych” Wydawca: Uniwersytet Techniczny w Winnicy, Ukraina 2013, rozdz. 7.5



Dodatek F. Statystyka odpornościowa
s. 242-256



Rozdz. 9. Procedury estymacji odporne na błędy grube
Rozdz. 10. Procedury estymacji charakterystyk wielowymiarowych s. 312-327 i 328-349

ISO 13528: 2009-01. Statistical method for use in proficiency testing by interlaboratory comparisons. Annex C (normative) Robust Analysis (p. 104-105)

**C.1 Robust analysis:
Algorithm A**

This algorithm yields robust values of the average and standard deviation of the data to which it is applied.

NOTE 1 Algorithms A and S given in this annex are reproduced from ISO 5725-5.

NOTE 2 Robustness is a property of the estimation algorithm, not of the estimates it produces, so it is not strictly correct to call the averages and standard deviations calculated by such an algorithm robust. However, to avoid the use of excessively cumbersome terminology, the terms "robust average" and "robust standard deviation" should be understood in this International Standard to mean estimates of the population mean or of the population standard deviation calculated using a robust algorithm.

The robust estimates x^* and s^* may be derived by an iterative calculation, i.e. by updating the values of x^* and s^* several times using the modified data, until the process converges. Convergence may be assumed when there is no change from one iteration to the next in the third significant figure of the robust standard deviation and of the equivalent figure in the robust average. This is a simple method to program on a computer.

Denote the p items of data, sorted into increasing order, by:

$$x_1, x_2, \dots, x_p, \dots, x_p$$

Denote the robust average and robust standard deviation of these data by x^* and s^* .

Calculate initial values for x^* and s^* as:

$$x^* = \text{median of } x_j \quad (i = 1, 2, \dots, p) \quad (\text{C.1})$$

$$s^* = 1,483 \text{ median of } |x_j - x^*| \quad (i = 1, 2, \dots, p) \quad (\text{C.2})$$

Update the values of x^* and s^* as follows. Calculate:

$$\delta = 1,5s^* \quad (\text{C.3})$$

For each x_j ($i = 1, 2, \dots, p$), calculate:

$$x_j^* = \begin{cases} x^* - \delta, & \text{if } x_j < x^* - \delta \\ x^* + \delta, & \text{if } x_j > x^* + \delta \\ x_j, & \text{otherwise} \end{cases} \quad (\text{C.4})$$

Calculate the new values of x^* and s^* from:

$$x^* = \sum x_j^* / p \quad (\text{C.5})$$

$$s^* = 1,134 \sqrt{\sum (x_j^* - x^*)^2 / (p - 1)} \quad (\text{C.6})$$

where the summation is over i .

ISO 13528: 2009-01. Statistical method for use in proficiency testing by interlaboratory comparisons. Annex C (normative) Robust Analysis (p. 105 -106)

C.2 Algorithm S

This algorithm is applied to standard deviations (or ranges). It yields a robust pooled value of the standard deviations or ranges to which it is applied.

Denote the p items of data, sorted into increasing order, by:

$$w_1, w_2, \dots, w_i, \dots, w_p$$

(These may be ranges or standard deviations.)

Denote the robust pooled value by w^* , and the degrees of freedom associated with each w_i by ν . (When w_i is a range, $\nu = 1$. When w_i is the standard deviation of n test results, $\nu = n - 1$.) Obtain the values of ξ and η required by the algorithm from Table C.1.

Calculate an initial value for w^* as:

$$w^* = \text{median of } w_i \quad (i = 1, 2, \dots, p) \quad (C.7)$$

Update the value of w^* as follows. Calculate:

$$\psi = \eta \times w^* \quad (C.8)$$

For each w_i ($i = 1, 2, \dots, p$), calculate:

$$w_i^* = \begin{cases} \psi, & \text{if } w_i > \psi \\ w_i, & \text{otherwise} \end{cases} \quad (C.9)$$

Calculate the new value of w^* from:

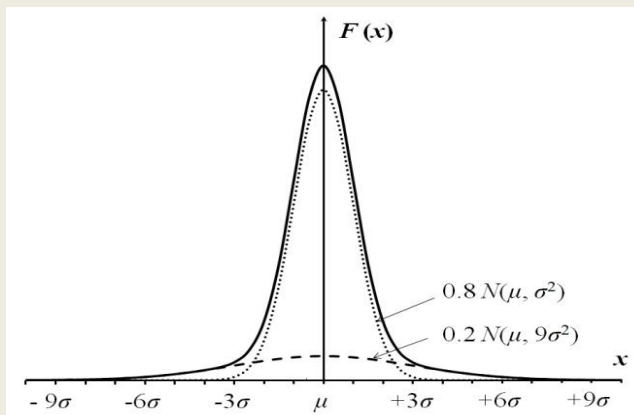
$$w^* = \xi \sqrt{\sum (w_i^*)^2 / p} \quad (C.10)$$

The robust estimate w^* may be derived by an iterative calculation by updating the value of w^* several times, until the process converges. Convergence may be assumed when there is no change from one iteration to the next in the third significant figure of the robust range. This is a simple method to program on a computer.

Degrees of freedom	Limit factor	Adjustment factor
ν	η	ξ
1	1,645	1,097
2	1,517	1,054
3	1,444	1,039
4	1,395	1,032
5	1,359	1,027
6	1,332	1,024
7	1,310	1,021
8	1,292	1,019
9	1,277	1,018
10	1,264	1,017

ANMERKUNG Die Werte von ξ und η sind aus ISO 5725-5:1998, Anhang B, abgeleitet.
NOTE: The values of ξ and η are derived in ISO 5725-5:1998, Annex B.

Rozkład mieszany wg Tukey'a uwzględniający dane odstające



$$F(x) = (1 - \varepsilon) N(\mu, \sigma^2) + \varepsilon N(\mu, 9\sigma^2) \quad \text{for } \varepsilon = 0.2.$$

$$F(x) = (1 - \varepsilon) \cdot p\left(\frac{x - \mu}{\sigma}\right) + \varepsilon \cdot p\left(\frac{x - \mu}{3\sigma}\right)$$

2. 1/1 Procedury odpornościowe

Metody statystyki odpornościowej zwiększają wiarygodność statystyczną identyfikacji parametrów próbek o małej liczności n .

I. Metoda absolutnego medianowego odchylenia MAD_n (Median Absolute Deviation).

$$\text{Dla próbki o } n \text{ elementach: } MAD_n = \text{med}\{|x_i - M_n|\} \quad (2)$$

gdzie: x_i - i-ty element próbki $M_n = \text{med}\{x_i\}$ - jej mediana.

Do dokładniejszego oszacowania odchylenia standardowego próbki w zależności od liczby jej elementów n stosuje się współczynnik korekcyjny $k(n)$ o wartościach podanych w [6], np. $k(4)=2,019$. Stąd otrzymuje się odchylenie standardowe

$$\hat{s} = k(n)MAD_n$$

Jest ono odporne na dane odstające.

Metoda ta jest bardzo prosta i nie jest wrażliwa na maksymalną wartość bezwzględną odchylenia elementu próbki od centrum jej rozkładu.

Odchylenie standardowe \hat{s} ma inną wartość, niż s obliczone z rzeczywistych danych próbki wg GUM.

17

2. 1/2 Zmodyfikowana ocena medianowa

[6] Randa Proposal for KCRV and Degree of Equivalence for GTRF Key Comparisons. NIST, (2000) and update (2005)

Tabela 1. Zmodyfikowana odporna procedura medianowa

$$s(x_n) \equiv k(n) \cdot MAD$$

n	$k(n)$	n	$k(n)$	n	$k(n)$
2	1,773	10	1,626	50	1,507
3	2.206	11	1,601	100	1,494
4	2,019	12	1,596	1000	1,484
5	1,800	13	1,581	2000	1,483
6	1,764	14	1,577	.	.
7	1,686	15	1,566	.	.
8	1,671	20	1,544	∞	1,483
9	1,633	25	1,530		

2. II./1 Metoda odporna Hubera (Algorytm A wg ISO -5)

Metoda Hubera, jest jednym z podstawowych narzędzi matematycznych statystyki odpornościowej (*robust statistics*).

W metodzie tej zakłada się, że tylko część środkowa rozkładu gęstości prawdopodobieństwa (*probability density function* PDF), tj. dla małych odchyłeń od estymatora wartości mierzand, nie odbiega zbytnio od przyjętego rozkładu normalnego jako modelu teoretycznego.

Dzięki małej wrażliwości na dane odstające (outliery) umożliwia ona poprawę wiarygodności wyników przy próbkach o małej liczności.

Metodę można zaliczyć do grupy metod odpornych iteracyjnych.

Metodę zilustruje się przykładem jej zastosowania do atestacji dokładności procedur badawczych i kontrolnych dokonywanej poprzez porównanie wyników w badaniach między-laboratoryjnych.

Inicjatorem przedstawienia tej metody polskiemu środowisku metrologicznemu na sympozjach PPM'12, Automation i innych był prof. Eugenij Wołodarski z Politechniki Kijowskiej (z pomocą doc. Zygmunta Warsz PIAP)

2. II./2 Podstawy matematyczne metody odpornej Hubera

- przyjmuje się model z outlierami, czyli danymi odstającymi, o bardziej rozciągniętych "ogonach" niż dla PDF rozkładu normalnego .
- Pozwala to zachować wygodne do analizy założenie o hipotetycznej jednorodności populacji do statystycznego oszacowania i równocześnie uwzględnić informację zawartą w odchyleniach odstających.
- Tylko dla centralnej części wartości danych stosuje metodę najmniejszych kwadratów MNK, a poza granicami tego przedziału, wykorzystuje kryterium mini-max MNM by zmniejszyć wpływ danych odstających (outlierów), ale ich nie usuwać.
- Dla zmniejszenia czułości na odstające wartości danych w [5] stosuje się funkcjonal

$$\sum_{i=1}^n \rho(x_i - \mu) \quad (1)$$

gdzie : - postać funkcji $\rho(x_i - \mu)$ zależy od pewnego parametru c

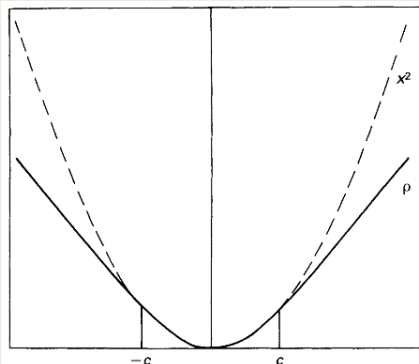
Dla małych odchyłeń $|\varepsilon| = |x_i - \mu| < c\sigma$ (gdzie: σ - odchylenie standardowe)

jest to funkcja kwadratowa,

Dla dużych odchyłeń $|\varepsilon| = |x_i - \mu| > c\sigma$ minimalizuje się moduł odchyłeń.

Te obszary czułości metody ilustruje rys 1.

II./3 Podstawy metody odpornej Hubera cd.



Rys 1. Względna funkcja czułości metody

$$|\varepsilon| = |x_i - \mu| \quad \rho(\varepsilon) = \varepsilon^2 \text{ dla } |\varepsilon| \leq c, \\ \rho(\varepsilon) < \varepsilon^2 \text{ dla } |\varepsilon| > c$$

Wartość stałej c zależy od stopnia "zanieczyszczenia" rozkładu.

- dla "zanieczyszczenia" 1% $c=2$,
- dla "zanieczyszczenia" 5% $c=1.4$.

Zwykle wybiera się $c=1,5$.

$$x_i^* = \begin{cases} x_i & \text{przy } |x_i - \hat{\mu}| \leq c\sigma, \\ \hat{\mu} - c\sigma & \text{przy } x_i < \hat{\mu} - c\sigma, \\ \hat{\mu} + c\sigma & \text{przy } x_i > \hat{\mu} + c\sigma \end{cases} \quad (2)$$

gdzie $x_i^* = \text{med}$ danych x_i uszeregowanych rosnąco

Jako ocenę środka grupowania rozkładu próbki odporną na dane odstające preferuje się medianę.

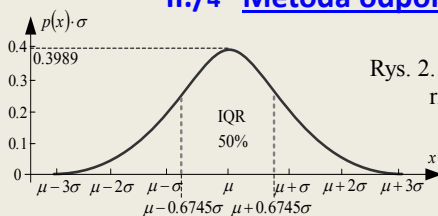
Z badań wynika [6], że jako środkowy przedział najlepszy jest rozstęp między trzecim ($P = 3/4$) i pierwszym ($P = 1/4$) kwartylami (*inter-quartile mid-range*).

Kwartyle te odległe są od środka tego przedziału o odchylenie ćwiartkowe.

Dla tego przedziału powierzchnia pod krzywą rozkładu Gaussa wynosi 50%.

W metodzie tej stosuje tzw. winsoryzację danych odstających, tj. sprowadzanie ich na granice środkowego przedziału międzykwartylowego.

II./4 Metoda odporna Hubera cd.



Rys. 2. Definicja rozstępu między-kwartylowego: rzędne pierwszego i trzeciego kwartyła

$$a = \mu - 0.6745\sigma, \quad b = \mu + 0.6745\sigma$$

Dla ustalenia zależności pomiędzy parametrami rozkładu PDF "obciążonego" i dla całej populacji, tj. aby spełnić warunek skalowania, konieczne jest przeliczenie odchylenia standardowego przy użyciu współczynnika korygującego,

który wyznacza się z początkowego rozkładu pdf

$$p(x) = \frac{c}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}$$

W rozkładzie "nieobciążonym" $c=1$

w "obciążonym" w kwartylach: $c = \frac{0,5}{p(\frac{b-m}{\sigma}) - p(\frac{a-m}{\sigma})}$

$$\text{Dla rozkładu normalnego } c = \frac{1}{0,6745} \approx 1,483 \quad \text{ i SD odporne } s^* = 1,483 \cdot MAD_n \quad (3)$$

Jako ocenę środka grupowania przyjmuje się medianę oryginalnej serii danych, bardziej odporną na dane odstające (outliery) niż średnia. Następnie z (3) oblicza się s^* i po wyborze $c=1,5$, granicę obszarów stosowania metody MNK i metody najmniejszego modułu MNM.

$$\varphi = c s^*$$

II./5 Metoda Hubera cd

Stosując warunki (2), otrzymuje się zmodyfikowany szereg danych x_i^* którego wartość średnia

$$\bar{x}^* = \frac{1}{n} \sum_{i=1}^n x_i^*,$$

Po podstawieniu x_i^* , \bar{x}^* , wyznacza się z (3) - odchylenie standardowe $s(x^*)$

Obliczoną wartość $s(x^*)$ wykorzystuje się z kolei do obliczenia "nowego" stabilnego odchylenia standardowego

$$s^* = 1,134 \cdot s(x^*)$$

(Współczynnik 1,134 odpowiada $c = 1,5$ dla rozkładu normalnego)

Otrzymaną wartość s^* wykorzystuje się do obliczenia nowej granicy φ

Kontynuuje się procedurę w sposób jak powyżej.

Na podstawie wartości \bar{x}^* obliczonych w bieżącym i poprzednim kroku iteracji określa się zbieżność algorytmu.

Procedurę powtarza się, aż zmiany \bar{x} i s^* pomiędzy kolejnymi krokami staną się minimalne. Zastosowania metody przedstawi się w przykładzie 1.



Sympozja: PPM' 2012, PPM' 2014,
Podstawowe Problemy metrologii
Automation 2014, Automation 2015
Kongres Metrologii 2013
Konferencje Międzynarodowe
Publikacje w języku polskim i angielskim

3. Istota badań międzylaboratoryjnych. Wprowadzenie



Levgen Volodarsky
Politechnika Kijowska



Larysa Kosheva
Instytut Lotnictwa Kijów



Zygmunt Warsza
PIAP Warszawa

Монография

Е.Т. Володарский, Л.О. Кошечкина

Технические аспекты аккредитации испытательных лабораторий



Аннотация

Рассмотрены технические вопросы аккредитации испытательных лабораторий, пути достижения сопоставимости и взаимного признания результатов испытаний. Приведено обоснование выбора унифицированных показателей точности результатов испытаний и методов их оценивания. Рассмотрены примеры практического применения предложенных методов оценивания результатов.

Для научных работников, преподавателей, аспирантов, разработчиков методик испытаний, аудиторов (ассессоров), экспертов органов по аккредитации и других специалистов в сфере испытаний.

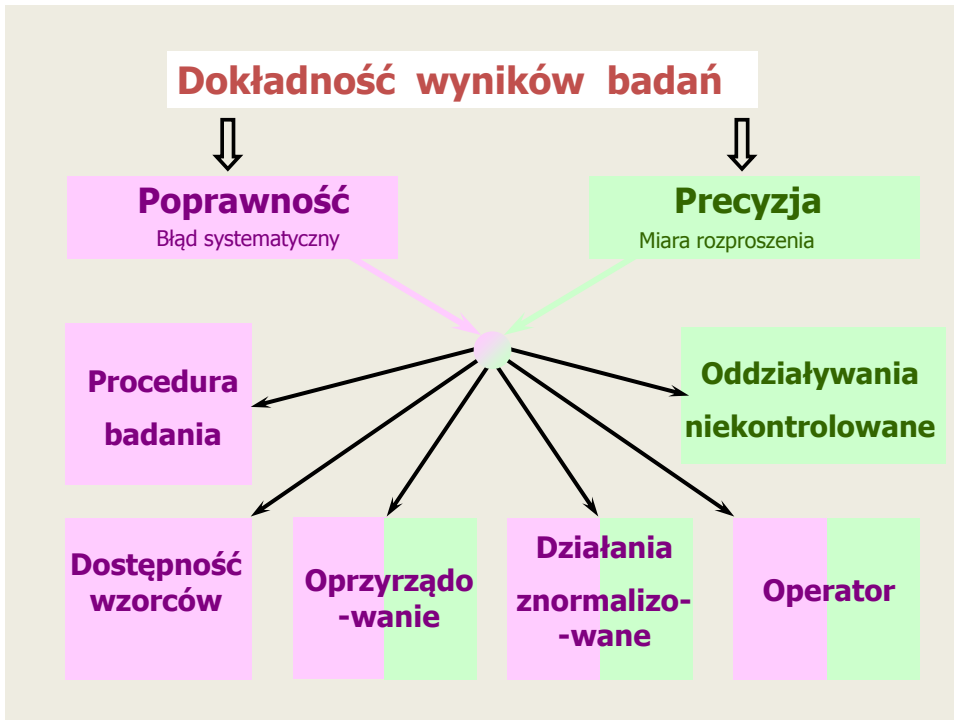
СОДЕРЖАНИЕ

Глава 1. Испытания как этап обеспечения качества продукции.....	18
Глава 2. Обеспечение метрологической сопоставимости и совместимости результатов лабораторных испытаний.....	41
Глава 3. Статистический подход к оцениванию количественных результатов испытаний.....	56
Глава 4. Установление характеристик точности методики испытаний.....	76
Глава 5. Внутривлабораторный эксперимент в системе обеспечения качества лабораторных испытаний	127
Глава 6. Внешний контроль качества результатов лабораторных испытаний.....	156
Глава 7. Неопределенность количественных результатов испытаний.....	176
Глава 8. Оценивание неопределенности принятия решения при контрольных испытаниях.....	234

Istota badań międzylaboratoryjnych WPROWADZENIE

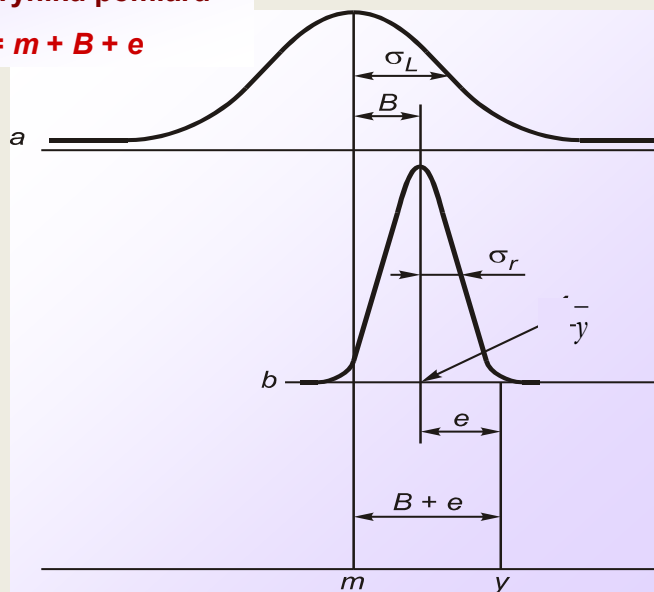
- **R**ozpatruje się zagadnienie wzajemnego uznawania (przez klienta i dostawcę) wyników oceny jakości produktów
- **W**zajemne uznawanie wyników badań może zapewnić laboratorium badawcze o odpowiedniej biegłości, tj. o technicznych kompetencjach potwierdzonych przez akredytację.
- **P**orównywalność wyników uzyskuje się dzięki ich dokładności i niezawodności.
- **P**orównywalne wyniki można osiągnąć przez kompleksowe stosowanie współczesnych metod statystycznych.



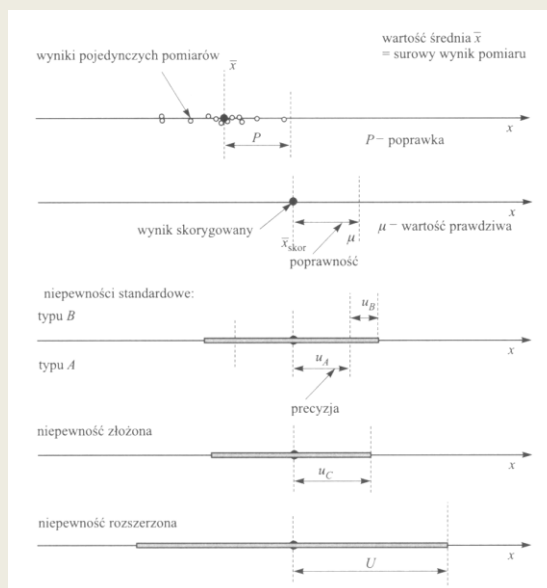


Podstawowy statystyczny model wyniku pomiaru

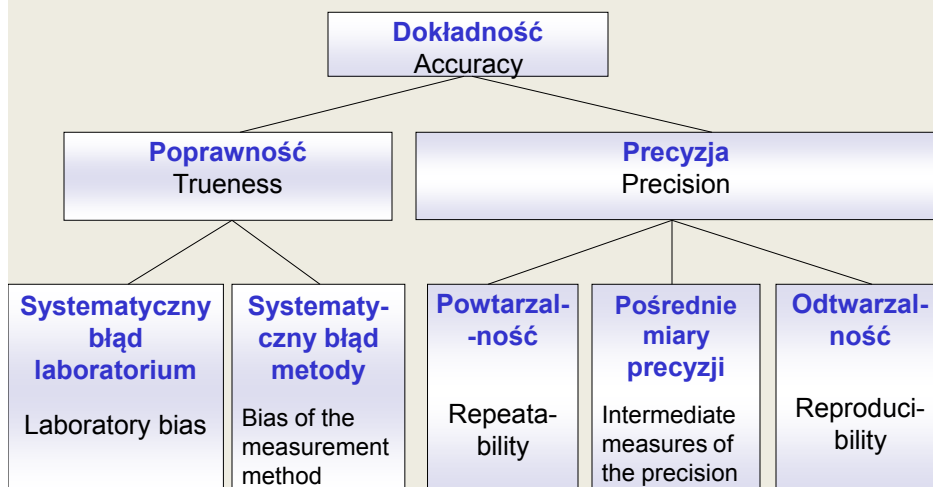
$$Y = m + B + e$$



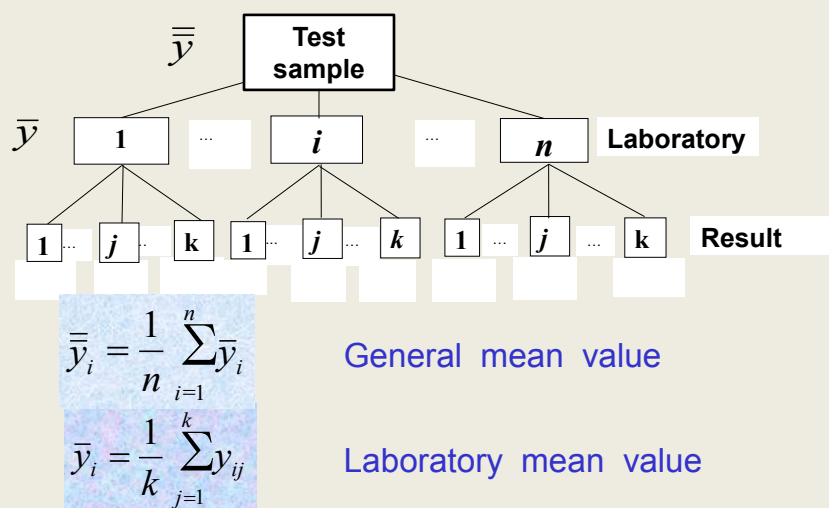
Ilustracja estymacji składowych niepewności wyniku pomiaru wg zaleceń GUM



Miary oceny dokładności metod pomiarowych



Organization of inter-laboratory experiment



Metoda (procedura) prowadzenia badań



- **W**spółczesne przepisy o certyfikacji metody badań wymagają sprawdzania dokładności przez porównania międzylaboratoryjne
- **P**orównania międzylaboratoryjne stanowią eksperymentalną realizację modelu **procedury badawczej**.
Uwzględnia się **możliwe warianty warunków przeprowadzania badań w laboratoriach**.
Pozwala to na ustalenie parametrów charakteryzujących daną procedurę badań.

Organization of inter-laboratory experiment main statistical parameters

$$s_{w_i}^2 = \frac{1}{k-1} \sum_{j=1}^k (y_{ij} - \bar{y}_i)^2$$

Laboratory variance

$$s_r^2 = \frac{1}{n} \sum_{i=1}^n s_{w_i}^2$$

Repeatability variance

$$s_L^2 = \frac{1}{n-1} \sum_{i=1}^n (\bar{y}_i - \bar{\bar{y}})^2$$

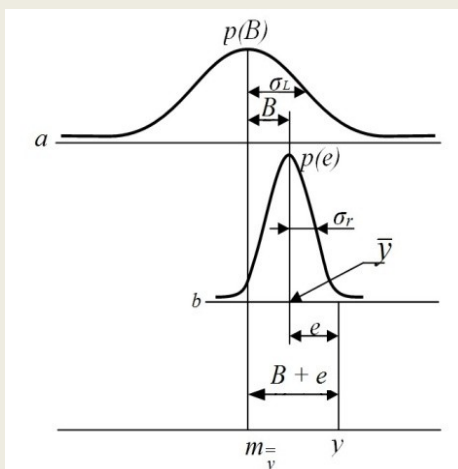
Between laboratory
variance

for $n \rightarrow \infty$

$$\sigma_R^2 = \sigma_r^2 + \sigma_L^2$$

Variance of reproducibility

Basic statistical model of the measurement result

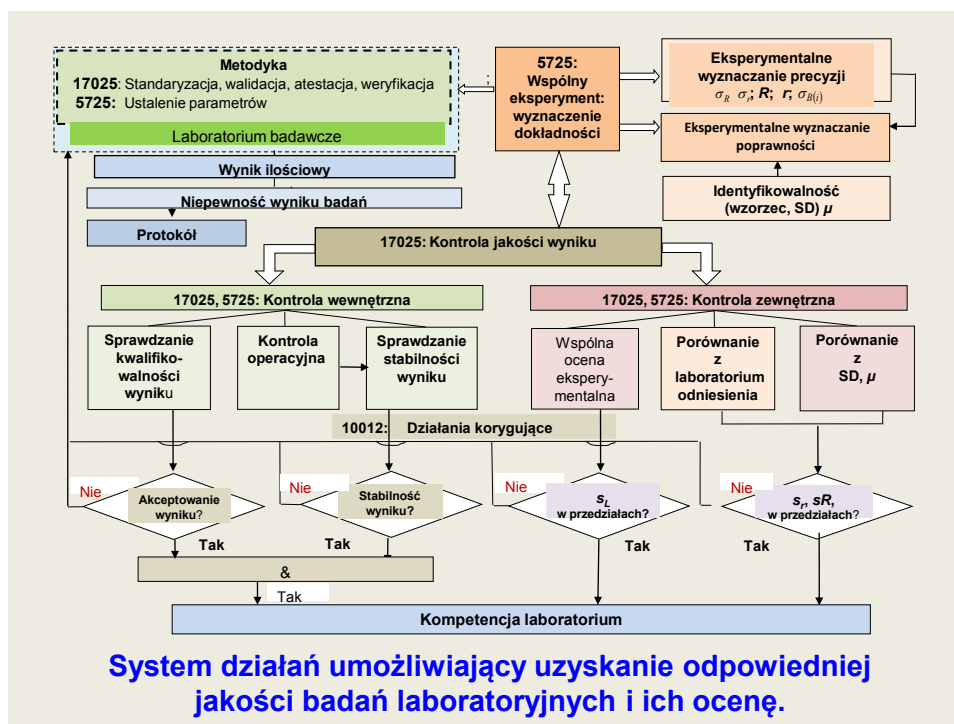


$$y = m_y + B + e$$

m_y - mean of test result

B - laboratory component
of bias

e - random component
in test result



Ustalenie wskaźników (miar) dokładności metody

Porównania międzylaboratoryjne umożliwiają ustalenie miar dokładności do użycia w dokumentacji normującej wymagania dla procedury danej metody badań.

- **O**dchylenia standardowe powtarzalności σ_r i odtwarzalności σ_R oraz ich przedziały r, R ;
- **O**dchylenia standardowe pośredniej precyzji
- **P**zesunięcie wyniku δ (poprawność, błąd $\sigma_{I(.)}$ systematyczny) spowodowane niedoskonałością procedury.

4.1/1 Zastosowanie metod odpornych. Przykład 1

Dane liczbowe zaczerpnięto z [3]. W dziewięciu laboratoriach przeprowadzono wspólny eksperyment polegający na pomiarach porównawczych pewnej metody badawczej w celu oszacowania jej dokładności. Założono wstępnie, że wiarygodność pomiarów wszystkich laboratoriów jest jednakowa. Z pomiarów w laboratoriach otrzymano wartości średnie:

24.140, 20.155, 19.500, 20.300, 20.705, 17.570, 20.100, 20.940, 21.185

Dwa podkreślone wyniki (pierwszy i siódmy) istotnie różnią się od pozostałych. Wynik wspólny dla eksperymentu zostanie znaleziony dwoma odpornymi metodami.

Metoda I: Obliczenie przeskalowanego odchylenia medianowego MAD_s

Jeśli założy się, że wszystkie obserwacje pochodzą z *pojedynczego rozkładu niegaussowskiego*, to logiczną konsekwencją jest zastosowanie algorytmu, który nie wyróżnia żadnej z wartości otrzymanych z pomiarów.

Sposób postępowania :

- **obliczenie mediany $med = 20,3$ i uznanie jej za wynik pomiaru,**
- **obliczenie odchylenia medianowego (*median absolute deviation*) $MAD = 0,64$** (również ze wszystkich punktów).

Za standardową niepewność pomiaru można uznać **przeskalowane odchylenie medianowe MAD_s**

$$s(x) \equiv MAD_s = \kappa \cdot MAD \quad (6)$$

$\kappa_\infty = 1,483$ jest teoretycznym stosunkiem s/MAD rozkładu normalnego gdy $n \rightarrow \infty$.

4.1/2 Zastosowanie metody I medianowej. Przykład 1. cd.

Dla analizowanych danych otrzymuje się

$$s(x) \equiv 1,483 \cdot 0,64 = 0,95.$$

Wg Zięby [16] taką odporną procedurę proponowali Burke (2001), Randa (2000), Müller (2001) i nieco inną M. Cox (2002) z ang. NPL.

Randa w 2005 r. zaproponował w internecie [17], by zamiast wartości asymptotycznej współczynnika $\kappa_\infty = 1,483$, która zaniża ocenę wyniku pomiarów, **dla próbek o skończonej liczebności n stosować wartości $\kappa(n)$ z tabeli 1.**

Tabela 1. Wartości mnożnika κ dla prób losowych o liczebności n wg [17].

n	$\kappa(n)$	n	$\kappa(n)$	n	$\kappa(n)$
2	1,773	10	1,626	50	1,507
3	2,206	11	1,602	100	1,494
4	2,019	12	1,596	1000	1,484
5	1,800	13	1,581	2000	1,483
6	1,764	14	1,577	.	.
7	1,686	15	1,566	.	.
8	1,671	20	1,544	∞	1,483
9	1,633	25	1,530		

Dla analizowanych 9-ciu danych: $s(x) \equiv \kappa(n) \cdot MAD = 1,633 \cdot 0,64 = \underline{1,045}$

4./3 Metoda II odporna Hubera (ISO 5725-5 Algorytm A)

Metoda stosuje tzw. **winsoryzację danych odstających**. Pozwala w obliczaniu odchylenia standardowego próbki uwzględnić wartości odstające o odchyleniu bezwzględnym maksymalnym [5], [7] sprowadzając je na granice $\pm \varphi$ środkowego międzykwartylowego przedziału dla normalnego rozkładu prawdopodobieństwa.

Po uporządkowaniu elementów próbki wg ich wartości x_1, x_2, \dots, x_n wyznacza się jako środek grupowania danych x^* i jako odchylenie standardowe s^* :

$$x^* = \text{med} \{x_i\}, \quad i = 1, \dots, n \quad (4) \quad \text{i} \quad s^* = 1,483 \text{MAD}_n \quad (5)$$

Oblicza się wartość $\varphi = c s^*$ i granice w (6) do porównywania danych odstających, (np. $c=1,5$ – zanieczyszczenie ok. 5% [7]).

$$x_i^* = \begin{cases} x^* - \varphi, & \text{przy } x_i < x^* - \varphi; \\ x^* + \varphi, & \text{przy } x_i > x^* + \varphi; \\ x_i - w & \text{w innych przypadkach.} \end{cases} \quad (6)$$

W kroku (j) procedury iteracyjnej, po zmodyfikowaniu wg (6) wartości z kroku ($j-1$) znajduje się kolejno nowe wartości średnie (7) i odchylenia standardowe (8). Procedurę zatrzymuje się dla kroku $j=m$, gdy różnica dwu kolejnych odchyleń standardowych jest dopuszczalnie mała.

$$\bar{x}_{(j)} = \sum_{i=1}^n x_{i(j)}^* / n; \quad (7)$$

$$s_{(j)}^* = 1,134 \sqrt{\sum_{i=1}^n (x_{i(j)}^* - \bar{x}_{(j)})^2 / (n-1)} \quad (8)$$

Oszacowane po $j=m$ krokach odchylenie standardowe $s_{\text{val}} = s_{(m)}^*$ jest obliczone na podstawie wszystkich otrzymanych danych doświadczalnych.

4. Przykład 1. Metoda II

Metoda II. Obliczenia oceny wyniku metodą odporną Hubera

Wyniki pomiarów poszczególnych laboratoriów ustawiono w porządku rosnącym (początek iteracji - kolumna „0” w tabeli 2) jako: $x_{i(0)}^*$, gdzie $i = (1, \dots, 9)$
Z tych danych oblicza się 2 początkowe wartości dla $n = 9$ laboratoriów:

$$\text{średnia} \quad x_0^* = \frac{1}{n} \sum_{i=1}^n x_{i(0)}^* = 20,511$$

$$\text{odchylenie standardowe SD} \quad s_0^* = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_{i(0)}^* - x_0^*)^2} = 1,727$$

gdzie: x_i^* pochodzi z kolumny "0" w tabeli 2.

Jako początkowe (tj. dla surowych danych) oszacowanie położenia centrum grupowania populacji przyjmuje się

$$\text{medianę} \quad \text{med } x_{i(0)}^* = x_5^* = 20,300$$

i oblicza się **odchylenie MAD** $s_1^* = 1,483 \cdot \text{med} \left\{ |x_{i(0)}^* - x_5^*| \right\}$

gdzie: seria różnic $|x_{i(0)}^* - x_5^*|$ ma wartości: **2,73, 0,8, 0,2, 0,145, 0, 0,405, 0,640, 0,889, 3,84**

Oblicza się $s_1^* = 1,483 \cdot 0,640 = 0,949$ i następnie **wartość prognozy**: $\varphi_1 = 1,5 \cdot s_1^* = 1,5 \cdot 0,949 = 1,424$

Granice przejścia od metody MNK (o najmniejszej sumie kwadratów odchyleń) do MNM (o najmniejszym module): $x_{(0)}^* - \varphi_1 = 20,300 - 1,424 = 18,876$ $x_{(0)}^* + \varphi_1 = 20,300 + 1,424 = 21,724$

Wyniki te umieszczono w kolumnie 1 tabeli 1 jako pierwszy krok iteracji.

Obliczenia kolejnych iteracji podano w kolumnach 1- 4 tabeli 2.

4./4 Przykład 1, metoda II cd.

Tabela 2 Przykład obliczeń odpornego wyniku pomiarów

Nr iteracji j	0	1	2	3	4
φ_j		1,424	1,478	1,514	1,539
$x_j^* - \varphi_j$		18,876	18,909	18,893	18,872
$x_j^* + \varphi_j$		21,724	21,865	21,921	21,950
$x_{1(j)}^*$	17,570	18,876	18,909	18,893	18,872
$x_{2(j)}^*$		19,500			
$x_{3(j)}^*$		20,100			
$x_{4(j)}^*$		20,155			
$x_{5(j)}^*$		20,300			
$x_{6(j)}^*$		20,705			
$x_{7(j)}^*$		20,940			
$x_{8(j)}^*$		21,185			
$x_{9(j)}^*$	24,140	21,724	21,865	21,921	21,950
Średnia \bar{x}_j^*	20,511	20,387	20,407	20,411	20,412
SD s_j^*	1,727	0,869	0,890	0,905	0,916
nowe \bar{x}_{j+1}	20,300	20,387	20,407	20,411	20,412
nowe s_{j+1}^*	0,949	0,985	1,009	1,026	1,039

Wartości średnie kroku czwartego i piątego $x_4^* = 20,411$ $x_5^* = 20,412$

praktycznie nie różnią.

Odporna ocena standardowego odchylenia wartości średniej też jest

zbliżona, tj.: $s_4^* = 1,026$ $s_5^* = 1,039$.

UWAGI

Przy podejściu tradycyjnym, tj. zidentyfikowaniu wartości odstających (outlierów) wg kryterium Grabbsa i ich wyeliminowaniu z obliczeń otrzyma się: zbliżoną wartość średnią $m = 20,4$ i mniejsze SD: $s = 0,501$.

Ucieka się od rzeczywistej sytuacji w eksperymencie, idealizując, że dane eksperymentalne są z populacji generalnej o rozkładzie normalnym.

Niepewność standardowa w wspólnych badaniach dla $n=7$ ma o 15,5 % większe względne odchylenie standardowe niż dla $n=9$ [wg. E1 GUM: $s(u_i) \approx 1/(\sqrt{2(n-1)})$].

4./5 Przykład 1. Porównanie wyników

Tabela 2. Porównanie wyników czterech metod (2 klasyczne i 2 odporne)

Method	For all data	Rejected x_1, x_9 by Grabbs crit.	Robust MAD_s	Robust iterative
Result value	$\bar{x}_0 = 20.511$	$m = 20.4$	$med = 20.3$	$\bar{x}_s^* = 20.412$
Std. uncertainty	$s_0 = 1.727$	$s = 0.501$	$s(x_9) = 1.045$	$s^* = 1.039$

Dla początkowych danych $x_i = x_{i(0)}$ 9 lab.: **wartość średnia** $\bar{x}_0 = 20.51$

odchylenie standardowe $s_0 = 1.73$

W modelu tradycyjnym (*cross-contamination*) zakłada się, że dobre są tylko obserwacje pobrane z rozkładu Normalnego. Konsekwencją jest użycie testów do wykrycia outlierów, np. testów Grubbsa, tj.

$$G_n \max = |x_n - \bar{x}| / s \quad (n=1, n=9)$$

Po usunięciu outlierów $x_{1(0)}, x_{9(0)}$, dla pozostałych danych otrzymuje się:

wartość średnią $m = 20.41$ **bardzo zbliżoną do poprzedniej** oraz

odchylenie standardowe $s = 0.50$ **znacznie mniejsze (3.5 raza)** niż poprzednio.

Te rezultaty dotyczą całego eksperymentu, ale obliczono je z pomiarów w tylko w 7-miu laboratoriach.

Są one mniej wiarygodne statystycznie.

4./5

Przykład 1. Wnioski

- Dla obu metod odpornych wartości centrum grupowania danych 20,30 i 20, 41 są bardzo zbliżone.
- Ich niepewności standardowe różnią się między sobą tylko o ok. 9% i są pomiędzy niepewnościami metod klasycznych.
- Iteracyjną metodą odporną Hubera otrzymano

$$s^* = 1.039 > s$$

- Wyznaczając estymatory wartości średniej i standardowego odchylenia tą metodą wykorzystano dane wszystkich laboratoriów. Można więc przewidywać, że są one najbliższe do rezultatów, które by otrzymano dla dużej liczby niezależnych pomiarów i traktować te odporne estymatory parametrów populacji generalnej jako najlepsze z możliwych dla danych uzyskanych w eksperymencie.

4./6 Wnioski cd.

- tak więc:
 - **zbieżność obu metod odpornych jest niemal całkowita.**
 - **Metoda odporna Hubera ma algorytm bardziej skomplikowany, lecz łatwy do zautomatyzowania. Służy ona głównie wyznaczeniu odpornej oceny niepewności, gdyż wprowadzony próg $\pm c\sigma$ zmniejsza czułość na dane odstające.**
 - **Przeprowadzona analiza wykazała przydatność zastosowania metody Hubera do wyznaczania parametrów statystycznych próbek danych o małej liczebności, pobranych z populacji generalnej o założonym rozkładzie normalnym, ale zawierających też wyniki odstające od pozostałych.**
 - **Metoda Hubera pozwala ocenić bardziej obiektywnie wartość wyniku i dokładność badanej metody pomiarowej.**
 - **Różnice pomiędzy wyznaczonymi statystycznie ocenami całego międzylaboratoryjnego eksperymentu i niepewnościami wyniku poszczególnych laboratoriów można wykorzystać do obiektywnej analizy rzeczywistego stanu organizacji procesu badań i sposobu prowadzenia działalności w laboratoriach.**
- W szczególności dotyczy to tych dwu, które miały wyniki odstające.**

PODSUMOWANIE

- **Metody odporne:**
 - nie odrzucają informacji zawartej w danych odstających, a jedynie zmniejszają jej znaczenie,
 - nie idealizują warunków pomiaru i modelu rozkładu rzeczywistego tylko jako gaussowskiego,
 - są bardziej właściwym sposobem obliczania niepewności próbek o małej liczebności.
- Przetwarzanie danych pomiarowych i wyrażanie niepewności pomiarów metodami odpornymi powinno pojawiać się nie tylko na konferencjach pomiarowych, ale i w praktyce metrologicznej, poczynając od GUM.



Literatura 1

- [1] Guide to the Expression of Uncertainty in Measurement (GUM), revised and corrected version of GUM 1995, BIPM_JCGM 100:2008.
- [2] Wyrażanie Niepewności Pomiaru. Przewodnik. Polskie tłumaczenie wersji GUM z 1995 r. (J. Jaworski), Wyd. Głównego Urzędu Miar Alfavero Warszawa 1999, 2002
- [3] ISO 5725:1994. Accuracy (trueness and precision) of measurement methods and results.
- [4] Sarhan Ahmed E., Greenberg Bernard G., editors: Contributions to order statistics, John Wiley & Sons, Inc., New York, 1962, pp. 482
- [5] Tukey John W.: Exploratory Data Analysis. Addison-Wesley. 1978
- [6] Domański Cz., Pruska K.: Nielasyczne metody statystyczne. Polskie Wydawnictwo Ekonomiczne SA. 2000
- [7] Frosch Moller S., von Frese J. Bro R.: *Robust methods for multivariate data analysis*. J. of Chemometrics 2005; 19: 549–563
- [8] Doksum, Kjell A. Mathematical Statistics: Basic and Selected Topics. 1. Second updated edition, Pearson Prentice-Hall 2007
- [9] Olive David J.: Applied Robust Statistics - Southern Illinois University Department of Mathematics. June 23, 2008
- [10] Huber P.J., Ronchetti E. M.: Robust Statistics. 2nd edition. Wiley 2011 pp. 380

Literatura 2 cd.

- [11] Daszykowski M. i inni: *Robust statistics in data analysis - A review basic concepts*. Chemometrics & Intelligent Laboratory Systems 85 (2007) 203–219
- [12] Stanimirova I. i inni: How to construct a multiple regression model for data with missing elements and outlying objects. Analytica Chimica Acta 581 (2007) 324–332
- [13] Stanimirova I., Walczak B.: *Classification of data with missing elements and outliers*, Talanta 76 (2008) 602–609
- [14] Piotrowski J., Kostyrko K.: Wzorcowanie Aparatury Pomiarowej. Wydawnictwo Naukowe PWN Warszawa Wydanie nowe 2012
- [15] Buchcik D.: *Procedura kalibracji przy wykorzystaniu metody najmniejszej mediany w przypadku modeli wielowymiarowych*. PAK nr 5 2008, str.294 – 297
- [16] Zięba A.: Analiza danych w naukach ścisłych i technice. PWN 2013
- [17] Randa J.: *Proposal for KCRV and Degree of Equivalence for GTRF Key Comparisons*. NIST, (2000) and update (2005). Preprint dostępny w internecie.
- [18] http://pl.wikipedia.org/wiki/Statystyka_odpornościowa
- [19] Piotrowski J. Procedury pomiarowe i estymacja sygnałów. Skrypt Politechniki Śląskiej nr 1889 Gliwice 1994

Literatura 3 : Z. Warsza i inni

W języku polskim

1. Volodarsky E. T., Warsza Z. L., Koshevaya L. A.: *Odporna ocena dokładności metod pomiarowych* **Pomiary Automatyka Kontrola (PAK)** vol. 58, nr 4/2012 s. 396 -401
2. Volodarsky E.T., Warsza Z. L., Koshevaya L.A., Palanychenko D.: *Zastosowanie estymacji odpornej w badaniach bieglności laboratorium przy niewielkiej liczbie pomiarów.* **Pomiary Automatyka Kontrola (PAK)** vol. 59 nr 6 2013 s. 554 -557
3. Volodarski E., Warsza Z., *Zastosowanie statystyki odpornościowej na przykładzie badań międzylaboratoryjnych.* **Przegląd Elektrotechniczny** - Electrical Review **11'** 2013 s. 260 - 267.
4. Volodarsky E.T., Warsza Z. L., Koshevaya L.A., *System oceny i zapewnienia kompetencji technicznej laboratoriów badawczych.* (Materiały X konferencji Naukowo – Technicznej) Podstawowe Problemy Metrologii **PPM'14**, Kościelisko 14-18 czerwca. Prace Oddz. PAN w Katowicach, seria: Konferencje nr 19 (2014), s.212 -216
5. Volodarsky E.T., Warsza Z. L., Koshevaya L.A., *System oceny i zapewnienia jakości badań bieglności laboratoriów przy ich akredytacji.* **Przemysł Chemiczny** vol. 93 n.8 (2014) s.1252 -54
6. Volodarsky E.T., Warsza Z. L., Koshevaya L.A., *System oceny statystycznej w badaniu bieglności laboratoriów badawczych.* **PAK (Pomiary Automatyka Kontrola)** 2014, 10 s. 816 -821
7. Eugenij Volodarsky, Zygmunt Warsza, Larysa Kosheva: *Ocena precyzji procedury pomiarowej w badaniach międzylaboratoryjnych metodą odporną z wykorzystaniem „algorytmu S”*, **Przemysł Chemiczny**, tom 94, nr 6 2015 s. 1008-1011
8. Eugenij T. Volodarsky, Zygmunt L. Warsza: *Ocena precyzji badań międzylaboratoryjnych metodą odporną "S-algorytm"* **Przegląd Elektrotechniczny** - Electrical Review **2015**, 10 s. 192 -196
9. Zygmunt Lech Warsza: **Metody rozszerzenia analizy niepewności pomiarów. Monografia.** Wydawnictwo Przemysłowego Instytutu Automatyki i Pomiarów (PIAP) Warszawa, 2016

Literatura 3 : Z. Warsza et al

In English

10. Volodarsky E. T., Warsza Z. L., *Application of two robust methods on the example of interlaboratory comparison.* in Monograph "Advanced Mathematical and Computational Tools in Metrology and Testing X" (**AMCTM X**) (F. Pavese, W. Bremsler, A.G. Chunovkina, N. Fischer, A.B. Forbes, Eds.), vol.10, Series on Advances in Mathematics for Applied Sciences vol. 86, World Scientific, N. Jersey • London • Singapore... (2015) p. 385 -391 ISBN:978-981-4678-61-2 + abstract 4 p., Int. Symposium X AMCTM Petersburg 9-12. Sept. 2014).
11. Volodarsky E. T., Warsza Z. L., *Examples of robust estimation with small number of measurements.* Progress in Automation, Robotics and Measuring Techniques. (editors: R. Szewczyk, C. Zieliński, M. Kaliczyńska), part 3 "Measuring Techniques and Systems". (ISBN 978-3-319-15834-1), **vol. 352** of series: "Advances in Intelligent Systems and Computing" (ISSN2194-5357) **Springer 2015** pp.285-291
12. Volodarsky E., Warsza Z., *Robust estimation in interlaboratory measurements with small number of measurements.* Measurement Automation Monitoring **MAM 2015** vol. 61 **no 4**, pp. 104 -110
13. Eugenij Volodarsky, Zygmunt Warsza, Larysa Kosheva, Adam Idzkowski: *Evaluation of Precision of the Inter-Laboratory Measurements Using Robust "S-Algorithm"* Problems and Progress in Metrology. Proceedings of Katowice Branch of Polish Academy of Sciences. Series: Conferences no 20 **PPM'15** pp. 07-10 June 2015 p. 53-59
14. Eugenij Volodarsky, Zygmunt Warsza, Larysa Kosheva, Adam Idzkowski: *Robust Algorithm S to assess precision of interlaboratory measurements* **MAM 2015** vol. 61 **no 4** pp. 111 -114
15. Eugenij Volodarsky, Zygmunt Warsza, Larysa Kosheva, Adam Idzkowski: *Evaluation of the precision of interlaboratory measurements by robust algorithm S.* **Proceedings of Symposium AMSA'15 Applied Methods of Statistical Analysis. Nonparametric Approach**, Novosibirsk & Bialokuriha 14-19 th September, **2015**) NGTU pp. 113 -123
16. Eugenij Volodarsky, Zygmunt L. Warsza, Larysa Kosheva, Adam Idzkowski: *Assessment of precision of the interlaboratory test data by using robust "Algorithm S"* in monograph: R. Jabłoński, T. Brezina (Editors) "Advanced Mechatronics Solutions" **no 393** of series **Advances in Intelligent Systems and Computing.** Springer **2016**, pp. 87-96, ISSN -2194-5357; DOI: 10.1007/978-3 319-23923-1_13

Podziękowanie

**Autor pragnie serdecznie podziękować
kierownictwu GUM i organizatorom
za zaproszenie do przedstawienia tej
problematyki na seminarium**