Przedmowa

Technika pomiarowa jest podstawą rozwoju działalności ludzkiej w obszarze wiedzy i kultury materialnej, wymianie dóbr i usług, gospodarce, ochronie zarówno życia i zdrowia, jak i środowiska. Pomiary wykonywane w różnych krajach, miejscach i czasie powinny dawać porównywalne ze sobą wyniki. Analiza danych i szacowanie dokładności wielkości mierzonej – czyli menzurandu – muszą być ujednolicone w skali międzynarodowej. Funkcję tę w ograniczonym zakresie pełni obecnie dokument o nazwie *Evaluation of measurement data – Guide to the Expression of Uncertainty in Measurement* oraz angielskim akronimie GUM. Pierwszą jego wersję wydała Międzynarodowa Organizacja Standaryzacji (ISO) w 1993 r. Dwa lata później ukazała się wersja poprawiona, a następnie – Suplementy. Ostatnia wersja Przewodnika GUM z 2008 r. jest już tylko w wersji elektronicznej. W 2012 r. ukończono kolejny Suplement 2 o wyznaczaniu niepewności wielowymiarowych pomiarów pośrednich. Opracowaniem dalszych Suplementów i nowej wersji GUM opartej na podejściu Bayesa zajmuje się obecnie utworzony w 1997 r. wspólny komitet ośmiu organizacji międzynarodowych ds. przewodników metrologicznych o skrócie JCGM, pod przewodnictwem dyrektora Międzynarodowego Biura Miar i Wag (BIPM) w Sèvres na obrzeżach Paryża. W języku polskim istnieje tylko tłumaczenie wersji GUM z 1995 r. *Wyrażanie Niepewności Pomiaru. Przewodnik* wraz z komentarzami prof. Janusza Jaworskiego, wydane przez Główny Urząd Miar w 1999 r. W praktyce pomiarowej wykorzystuje się aktualne angielskie wersje przewodnika GUM i jego Suplementów oraz międzynarodowy słownik terminów metrologicznych VIM (fr. *Vocabulaire international de métrologie – Concepts fondamentaux et généraux et termes associés*) i inne wydawnictwa BIPM.

Przewodnik GUM zawiera reguły szacowania i wyrażania niepewności pomiarów rekomendowane do stosowania przy wszystkich poziomach dokładności pomiarów w laboratoriach pomiarowych i innych placówkach wykonujących pomiary – od handlu do badań technicznych i podstawowych. Służby Miar i akredytowane laboratoria w większości krajów stosują te zalecenia już od wielu lat. Rozpowszechniły się też one w praktyce pomiarowej, w tym w kontroli procesów przemysłowych, analizie chemicznej, badaniu materiałów i urządzeń podczas ich wytwarzania i eksploatacji oraz w monitoringu środowiska i w nauce.

W Polsce zalecenia przewodnika GUM zostały bardzo szybko wdrożone do praktyki pomiarowej i dydaktyki akademickiej. Przyczyniło się do tego kilkanaście sympozjów o tematyce wyznaczania niepewności pomiarów, które w latach 1993–2013 prowadziła początkowo prof. Danuta Turzeniecka z Politechniki Poznańskiej, a następnie prof. Stefan Kubisa z Politechniki Szczecińskiej. Również na kongresach metrologii i corocznych konferencjach Podstawowych Problemów Metrologii były sekcje obejmujące te zagadnienia. Dotychczasowy polski wkład w kształtowanie

przepisów międzynarodowych dotyczących analizy i niepewności danych pomiarowych nie odzwierciedla jednak ani potencjału wiedzy metrologicznej, ani poziomu teorii pomiarów w kraju. Jednym z powodów była zmiana statusu Głównego Urzędu Miar w 1993 r. na urzędniczy (służba cywilna), który nie obejmuje prowadzenia prac badawczych. Ponadto wiele z oryginalnych publikacji polskich autorów o niepewności pomiarów ukazywało się głównie w języku rodzimym, co też ma znaczenie.

Wraz z rozwojem elektroniki i informatyki powstają nowe narzędzia pomiarowe oraz rozwijają się metody i procedury przeprowadzania pomiarów. Przepisy przewodnika GUM nie nadążają za bieżącymi potrzebami szybko rozwijającej się techniki pomiarowej. Stosowanie obecnej wersji tych zaleceń natrafia na ograniczenia, gdyż nie obejmują one wielu rodzajów pomiarów. Propozycje interpretacji, udoskonaleń i rozszerzania zakresu stosowania przewodnika prezentowane są na wielu konferencjach i w licznych publikacjach krajowych i zagranicznych. Wymaga to jednak wielu badań i ich weryfikacji w praktyce pomiarowej. Niezbędne jest też doprowadzanie do uwzględnienia ich w treści przepisów metrologicznych. Jest to jednak proces długotrwały. Europejskie i światowe środowiska metrologiczne – głównie Narodowe Instytuty Metrologiczne (NMI) – prowadzą prace badawczo-rozwojowe (w Unii Europejskiej w ramach EURAMET) niezbędne do opracowania nowych przepisów na miarę potrzeb współczesnej techniki pomiarowej. Polska metrologia od lat nie uczestniczyła w nich twórczo. W Głównym Urzędzie Miar zagadnieniami tymi zajmuje się jedynie dr Paweł Fotowicz. Aby taki stan zmienić polscy metrolodzy starszego pokolenia utworzyli już w 1991 r. Polskie Towarzystwo Metrologiczne (PTM), które zaczęło wydawać wkładkę *Metrologia i Technika Pomiarowa* w miesięczniku *Pomiary Automatyka Kontrola*. Tej inicjatywy społecznej nie wsparły jednak poprzednie kolejne kierownictwa Głównego Urzędu Miar. Autor, po przejęciu obowiązków prezesa PTM, dalszą działalność oparł zatem głównie na bezpośrednich kontaktach osobistych z metrologami zainteresowanymi rozwojem podstaw naukowych metrologii i techniki pomiarowej, w tym metod szacowania niepewności pomiarów oraz zajął się wspomaganiem publikowania prac z tej dziedziny autorów polskich i z innych krajów, głównie z Ukrainy.

Wykorzystując długoletnie doświadczenie, zarówno w konstruowaniu przyrządów pomiarowych, jak i w prowadzeniu metrologicznych prac badawczych autor formułował kolejno tematy z dziedziny podstaw metrologii, które jego zdaniem są istotne dla rozwoju techniki pomiarowej w przemyśle i szerzej – w wielu obszarach gospodarki krajowej. Ich rozwiązanie przekraczało możliwości jednej osoby. Doskonaleniem problematyki wyznaczania niepewności pomiarów i jej rozszerzeniem na obszary nieobjęte dotychczas zaleceniami przewodnika GUM – obok nielicznych metrologów polskich – byli też zainteresowani metrolodzy z innych krajów środkowoeuropejskich, w tym z Ukrainy i Białorusi oraz Rosji. Autor wraz z byłym długoletnim redaktorem naczelnym miesięcznika *Pomiary Automatyka Kontrola* – mgr. Tadeuszem Ustaborowiczem zainicjował i uczestniczył w powołaniu w 2008 r. nowego periodyku – kwartalnika *Pomiary Automatyka Komputery w Gospodarce i Ochronie Środowiska* wydawanego w Rzeszowie. Obecnie pismo wydawane jest w Lublinie pod nazwą *Informatyka Automatyka Pomiary w Gospodarce i Ochronie Środowiska*. Jest ono ukierunkowane na prezentację środkowo-europejskich dokonań, między in-

nymi w dziedzinie metrologii i techniki pomiarowej. Ponadto autor zainicjował i nawiązał nieformalną wymianę naukową i wspólne opracowywanie publikacji – realizowane głównie drogą internetową – zarówno z metrologami z polskich uczelni, jak i krajów ościennych (przede wszystkim z Ukrainy). Współpracę z autorem podjęli profesorzy z Polski: Tadeusz Skubis (Politechnika Śląska), Andrzej Zięba (Akademia Górniczo-Hutnicza) i Stefan Kubisa (Zachodniopomorski Uniwersytet w Szczecinie) oraz doktorzy: Jerzy Korczyński (Politechnika Łódzka), Jarosław Makal i Adam Idźkowski (Politechnika Białostocka). Z Ukrainy współpracowali profesorzy: Eugenij Volodarski (prezes stowarzyszenia Akademia Metrologii Ukrainy, Politechniki Kijowska) wraz z aspirantami i Larysa Kosheva (Akademia Lotnicza w Kijowie), Mikhailo Dorozhovets (Politechnika Lwowska i Politechnika Rzeszowska), Igor Zacharow (Instytut Radioelektroniki w Charkowie), Volodimir Kucheruk (Uniwersytet Techniczny w Winnicy) oraz Orest Serediuk (Instytut Nafty i Gazu w Iwano-Frankiwsku). Ponadto współpracowali: doc. Aleksandr Mikhal (Instytut Elektrodynamiki Ukraińskiej Akademii Nauk w Kijowie), dr Marina Galovska (jako aspirantka Politechniki Kijowskiej i następnie autorka wyróżnionej rozprawy doktorskiej w Politechnice w Brunszwiku), dr Sergiej Zabolotny (obecnie prorektor Uniwersytetu Technicznego w Czerkasach) oraz dr Volodimir Ezhela (główny specjalista ds. przetwarzania danych z Instytutu Fizyki Wielkich Energii IHEP „Rosatom" w Protvino w obwodzie moskiewskim). Rezultaty wielu wspólnych prac były prezentowane na kongresach, konferencjach i sympozjach krajowych i zagranicznych oraz publikowane w polskich i zagranicznych czasopismach naukowych.

Ostatnio zmieniła się sytuacja w Głównym Urzędzie Miar. Nowy prezes, dr Włodzimierz Lewandowski, podjął starania o przekształcenie tego urzędu w Narodowy Instytut Metrologiczny (NMI), w którym obok zadań metrologicznych związanych z utrzymaniem i doskonaleniem wzorców będą prowadzone prace badawczo-rozwojowe ukierunkowane na rozwój gospodarki. Prezes zapowiada też rychłe utworzenie zespołu rozwijającego podstawowe zagadnienia metrologii i techniki pomiarowej.

Zalecenia przewodnika GUM dotyczące analizy i wyznaczania niepewności pomiarów stosuje się obecnie w światowej praktyce pomiarowej, choć nie obejmuje on jeszcze wielu potrzeb badań użytkowych w nauce i gospodarce, w tym dotyczących jakości w przemyśle i innych dziedzinach. W szczególności brakuje zaleceń, jak w statystycznym opisie niepewności pomiarowych wykorzystywać optymalne estymatory dla danych o niegaussowskich rozkładach prawdopodobieństwa w pomiarach menzurandów jedno- i wielowymiarowych. Nie ma też zaleceń, jak jednolicie wyznaczać niepewność w pomiarach różnych procesów dynamicznych zmiennych w czasie i w przestrzeni. W modelowaniu pomiarów – poza metodą Monte Carlo opisaną w Suplemencie 1 do GUM – nie ma informacji, jak używać innych współczesnych zaawansowanych narzędzi matematycznych, w tym statystykę odpornościową. Przygotowanie nowego bayesowskiego ujęcia przepisów metrologicznych i dalsze rozszerzenie ich treści wymaga wielu szczegółowych badań i weryfikacji w praktyce o zasięgu międzynarodowym. Część z tych zagadnień została ujęta w treści niniejszej monografii. Autor zamierza też opracować następną monografię poświęconą tematyce metrologicznej w akredytacji i kontroli wiarygodności pomiarów w laboratoriach badawczych i kontrolnych.

Streszczenie

Niniejsza monografia prezentuje wyniki prac autora i współpracujących z nim metrologów z Polski, Ukrainy i innych krajów nad możliwościami rozszerzenia analizy niepewności pomiarów objętej międzynarodowym Przewodnikiem Wyznaczania Niepewności Pomiarów (GUM). Poświęcona jest ona omówieniu wybranych badań nad adaptacją współczesnych podstaw teorii prawdopodobieństwa i procesów stochastycznych do praktyki metrologicznej. Treść monografii obejmuje przedmowę, 12 autonomicznych rozdziałów i dwa samodzielne dodatki, łącznie 191 stron tekstu.

W rozdziale 1 przedstawiono syntezę stosowanych obecnie metod opisu dokładności przyrządów pomiarowych za pomocą błędów oraz dokładności pomiarów przez niepewności wg zaleceń przewodnika GUM. W następnych rozdziałach proponuje się metody rozszerzające te zalecenia. Są one ukierunkowane na zastosowania nie tylko w pomiarach Służby Miar, ale we wszystkich badaniach użytkowych.

W rozdziale 2, dla obserwacji pomiarowych uzyskiwanych przez próbkowanie ciągłego sygnału wielkości mierzonej, omówiono czyszczenie ich surowych danych z nieznanych regularnie zmiennych składowych systematycznych. Szacowanie niepewności typu A, gdy te oczyszczone dane powiązane są ze sobą statystycznie, czyli auto skorelowane wraz z wyznaczaniem ich funkcji autokorelacji otrzymaną z wartości tych danych eksperymentalnych, omówiono w rozdziale 3.

Przedstawiono wyniki badań określających możliwość zastosowania rozkładu prawdopodobieństwa w postaci jednego okresu funkcji $\cos^2$ jako alternatywy rozkładu normalnego, ale o ograniczonym zakresie dla prawdopodobieństwa równego 1 (rozdz. 4). Podano wyniki modelowania metodą symulacji Monte Carlo statystyk skośności i kurtozy próbek danych o małej liczebności z populacji o rozkładzie normalnym oraz równomiernym i trójkątnym (rozdz. 5).

Następnie omówiono inne niż średnia estymatory wartości i niepewności menzurandu dla danych wyników powtarzanych pomiarów, które modeluje się różnymi rozkładami niegaussowskimi, w tym równomiernym, arc sin, trapezowymi Trap oraz CTrap i ich splotami (rozdz. 6–9). Najlepszym estymatorem dla równomiernego rozkładu jest środek rozstępu. Dla rozkładów trapezowych zaproponowano nowy estymator dwuelementowy: 0,5(średnia + środek rozstępu). Jest on bardziej efektywny niż jednoelementowe estymatory, takie jak średnia, mediana i środek rozstępu. Niekonwencjonalne sposoby wyznaczania najefektywniejszego z rozpatrywanych estymatorów dla próbek o nieznanym *a priori* rozkładzie, realizowane są metodami symulacji Monte Carlo i resamplingu, które nie wymagają identyfikacji rodzaju rozkładu. Przedstawiono je w rozdziale 9.

W rozdziale 10 przedstawiono wyniki badań nad zastosowaniem do analizy niepewności metody maksymalizacji wielomianu (PMM) opartej na kumulantach. Natomiast w rozdziale 11 rozpatrzono dwie statystyczne metody odporne wyznaczania niepewności pomiarów: metodę o przeskalowanym odchyleniu medianowym MAD_S i metodą iteracyjną dwukryterialną podaną przez Hubera z zastosowaniem winsoryzacji danych odstających oraz ich zastosowanie w badaniach międzylaboratoryjnych.

Ostatni 12 rozdział zawiera omówienie kilku podstawowych zagadnień dotyczących wyznaczania niepewności i zaokrąglania wyników pomiarów pośrednich wielowymiarowych. Jest to wyjaśnienie i pewne rozszerzenie treści Suplementu 2 do przewodnika GUM.

Treść rozdziałów monografii uzupełniają dwa dodatki. Pierwszy zawiera przykłady obliczania bardziej efektywnych estymatorów menzurandu niż średnia dla rozkładów równomiernego i Laplace'a. W drugim dodatku przedstawiono kilka propozycji wyznaczania niepewności rozszerzonej, które można zastosować w nowym przewodniku GUM 2 opartym na podejściu Bayesa, w tym również dla środka rozstępu jako estymatora.

Monografia ta przeznaczona jest zarówno dla służby miar, jak i dla tych wszystkich, którzy dokonują i interpretują wyniki pomiarów w badaniach stosowanych w praktyce przemysłowej i we wszystkich innych dziedzinach gospodarki oraz w medycynie, wojskowości i w nauce.

Podziękowania

Serdecznie dziękuję autorom i współautorom wszystkich publikacji, których wyniki wykorzystałem w treści tej monografii, a w szczególności (w porządku alfabetycznym) profesorom Mykhailowi Drozhovetsowi, Stefanowi Kubisie, Eugenowi Volodarskiemu i Andrzejowi Ziębie oraz doktorom Marinie Galovskiej, Volodimirovi Ezheli, Jerzemu Korczyńskiemu i Sergiejowi Zabolotnemu. Współpracowali oni ze mną od 2007 r. nad rozwiązywaniem wielu problemów tu przedstawionych i znacznie wzbogacili moją wiedzę. Postanowiłem więc podzielić się nią z Czytelnikami.

Pragnę podkreślić, że nawiązanie treści tej monografii do poziomu międzynarodowych prac badawczych w dziedzinie metod analizy i szacowania niepewności pomiarów stało się możliwe dzięki ogromnej życzliwości i poparciu dyrekcji Przemysłowego Instytutu Automatyki i Pomiarów PIAP. Dzięki temu w latach 2010–2015 mogłem nie tylko opracować i opublikować kilkadziesiąt własnych i wspólnych prac, lecz także aktywnie uczestniczyć w wielu polskich i kilku najważniejszych europejskich konferencjach, sympozjach i warsztatach naukowych z podstaw teoretycznych metrologii i techniki pomiarowej.

Szczególne podziękowania należą się dr. Janowi Jabłkowskiemu, naczelnemu dyrektorowi Instytutu PIAP w powyższych latach, który ponadto skutecznie pokonał mój opór oraz zachęcił i przekonał mnie do opracowania niniejszej monografii.

Dziękuję też dr Małgorzacie Kaliczyńskiej, która dołożyła wiele starań jako redaktor naukowy tej monografii i cierpliwie akceptowała wnoszone do ostatniej chwili poprawki autorskie.

Dziękuje również recenzentom: prof. Andrzejowi Ziębie z Instytutu Fizyki Akademii Górniczo-Hutniczej w Krakowie oraz dr. Pawłowi Fotowiczowi – sekretarzowi naukowemu Głównego Urzędu Miar za wniesionych wiele bardzo cennych uwag merytorycznych i redakcyjnych.

Wreszcie, ogromne i nieocenione zasługi ma tu nade wszystko moja żona Danuta Nilsson, która przez wiele lat sama realizowała ogrom zadań życia codziennego, by umożliwić mi nieustanną, zwykle wielogodzinną pracę przy komputerze w ciszy domu w zdrowym leśnym klimacie Zalesia Górnego. Dziękuję Jej za to zarówno z całego serca, jak i z głębi duszy.

Zygmunt Lech Warsza Zalesie Górne, grudzień 2016 r.

Rozdział 1

Metody wyznaczania błędów i niepewności pomiarów oraz potrzeba ich rozszerzenia

Streszczenie. Wprowadzono w tematykę monografii, omówiono proces pomiarowy, podano definicje i podstawowe wzory do wyznaczania błędów pomiarowych, w tym granicznych oraz niepewności pomiaru wg międzynarodowego przewodnika GUM. Scharakteryzowano zakres oraz ograniczenia w stosowaniu niepewności. Podano potrzeby rozszerzenia sposobu szacowania niepewności metodą A na dane pozyskiwane przez próbkowanie sygnałów. Powinno być ono poprzedzone czyszczeniem danych ze składowych regularnych i uwzględniać autokorelację między obserwacjami próbki. Dla danych z rozkładów innych niż normalny należy stosować właściwe estymatory, a w małych próbkach przy występowaniu danych odstających statystykę odporną. Podano też podstawową literaturę polską i obcojęzyczną.

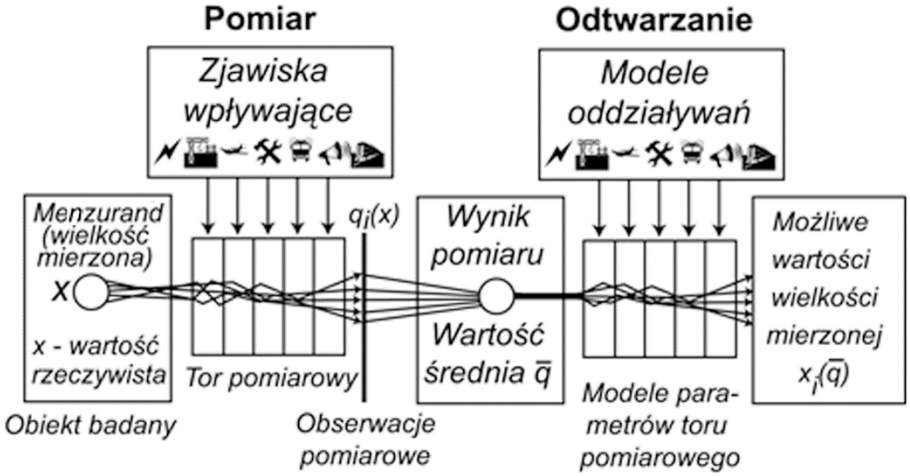
Pomiary towarzyszą człowiekowi od zarania dziejów. Jakościowe obserwacje, takie jak: mały – większy – największy lub blisko – dalej – najdalej, ciężki – cięższy – najcięższy itp., stopniowo przerodziły się w ujęcia ilościowe. Najstarsze przykłady – to mierzenie wielkości opadu w starożytnych Indiach na podstawie poziomu wody w określonym naczyniu, mierzenie odległości poprzez czas potrzebny na dotarcie karawany, mierzenie czasu klepsydrą lub wg okresowych zjawisk astronomicznych oraz kierunku wg położenia słońca i gwiazd. Stopniowo powstawały jednostki różnych wielkości, systemy ich podziału, urządzenia do pomiaru wartości różnych wielkości, metody ich użytkowania i sprawdzania oraz nazewnictwo. Szereg opisów można znaleźć już w Biblii. Stopniowo następowało ujednolicanie jednostek i instrumentarium stosowanego na coraz to większych obszarach i opracowywano przepisy dotyczące kalibracji i zasad użytkowania urządzeń pomiarowych. Tworzono też różne układy jednostek. Układ o symbolu SI jest przyjęty w skali globalnej.

W 1875 r. na mocy Konwencji Metrycznej utworzono Międzynarodowe Biuro Miar i Wag BIPM (fr. *Bureau International des Poids et Mesures*) w Sèvres pod Paryżem, które organizuje międzynarodowe porównania krajowych standardów pomiaru i przeprowadza kalibracje jednostek w państwach członkowskich. Bliski zakończenia jest też proces tworzenia wzorców wartości podstawowych jednostek miar opartych na zjawiskach kwantowych. Powstały też krajowe instytucje metrologiczne dbające o rozwój i utrzymanie wzorców państwowych, zapewnienie jednolitości i poprawności pomiarów, metod sprawdzania przyrządów w oparciu o te wzorce oraz o tworzenie i przestrzeganie przepisów prawnych dotyczących pomiarów (metrologia prawna). Jak jest to istotne dla gospodarki niech świadczy fakt, że szereg krajów Unii Europejskiej zrezygnowało z własnego pieniądza, ale utrzymuje Narodowe Instytuty Metrologiczne NMI (ang. *National Measurement Institute*). Wyłoniła się też samodzielna interdyscyplinarna dziedzina wiedzy stosowanej

o pomiarach i urządzeniach techniki pomiarowej o nazwie **metrologia**[1]. Zostały opracowane i nadal rozwijane są oryginalne podstawy teoretyczne i filozoficzne [20]. Technika pomiarów oraz przetwarzania i przekazywania sygnałów pomiarowych stanowi obecnie podstawę niemal każdego rodzaju działalności ludzkiej, od wytwarzania poprzez handel, użytkowanie, medycynę, edukację i naukę, aż do ochrony życia, mienia i środowiska. W poniższej monografii przedstawi się zasady opisu i oceny miar dokładności pomiarów oraz omówi propozycje ich rozszerzenia powstałe głównie z inicjatywy autora i opracowane z udziałem innych osób.

Proces pomiarowy i sposoby opisu wyniku pomiarów

Przebieg procesu pomiarowego ilustruje rysunku 1.



Rys. 1. Proces pomiarów i odtwarzania wartości wielkości mierzonej (menzurandu)
Fig. 1. The process of measurements and recovery of the measured value (measurand)

Na początku procesu pomiarowego następuje przetwarzanie wartości x wielkości mierzonej X w zależny od niej sygnał pomiarowy, który kolejno jest przetwarzany dalej w torze pomiarowym, aż do otrzymania obserwacji $q_i(x)$ w postaci sygnału o pożądanej formie lub jako odczyt wartości (wyników obserwacji). Proces ten nie jest idealny. Na obiekt mierzony i sygnał w torze pomiarowym oprócz wielkości mierzonej oddziałują inne wielkości zależnie od warunków pomiaru oraz zakłócenia zewnętrzne i wewnętrzne. Oddziaływania różnego pochodzenia oraz zmiany parametrów toru pomiarowego powodują, że przy powtarzaniu pomiaru poszczególne

[1] Alternatywną nazwę angielską ***measurement science*** o szerszym zakresie zaproponował w latach 70. XX wieku prof. Ludwik Finkelstein z City University w Londynie (urodzony we Lwowie).

obserwacje pomiarowe $q_i(x)$ przyjmują inne wartości niż mierzona wartość x badanej wielkości X, nazywanej obecnie *menzurandem*[2].

Przetwarzanie sygnału i propagację zakłóceń w torze pomiarowym rozpatruje się w kierunku **na wprost**, tj. od wielkości mierzonej do zbioru tzw. „surowych" wartości wyników obserwacji pomiarowych. Po wprowadzeniu poprawek uwzględniających znane zdeterminowane (czyli systematyczne) oddziaływania otrzymuje się skorygowane wartości obserwacji q_i. Ze zbioru n wartości q_i nazywanych próbką wyznacza się (estymuje) najbardziej prawdopodobny statystycznie wynik pomiaru, np. jako ich wartość średnia $\overline{q}$ oraz parametry charakteryzujące przedział rozproszenia obserwacji z przyczyn losowych i wskutek braku wiedzy o położeniu wartości rzeczywistej menzurandu. Na tej podstawie dalej w procesie pomiarowym odtwarza się położenie i szerokość przedziału dla zbioru wartości wielkości mierzonej x, które wg zadanego kryterium odpowiadają temu wynikowi pomiaru przy zadanym modelach oddziaływań, strukturze i parametrach toru pomiarowego. Procedura **odtwarzania** tego zbioru przebiega w kolejności **odwrotnej** niż przetwarzanie sygnału w początkowej części procesu pomiarowego z rys 1.

Realizacja procedury odtwarzania może odbywać się:

- poza systemem pomiarowym – obliczenia z wartości wyników obserwacji,
- w dalszej części toru pomiarowego, zwykle skomputeryzowanej, w sposób niejawny i praktycznie na bieżąco, poprzez operacje na zbiorze wartości sygnałów otrzymywanych dla poszczególnych obserwacji.

Z surowych wyników obserwacji pomiarowych, stosując poprawki, należy wyeliminować wpływy znanych oddziaływań systematycznych, które występują jako zdeterminowane w danym cyklu pomiarowym. Nie można jednak wyeliminować skutku bieżących wpływów stałych oddziaływań, gdyż wartość prawdziwa x nie jest znana. Mogą też być one inne w każdej z serii obserwacji oraz w ciągu całego okresu eksploatacji przyrządu lub badania obiektu, w tym wskutek zmian zarówno warunków pomiaru, jak i stabilności parametrów wewnętrznych. Natomiast przy regularnym lub innym znanym sposobie próbkowania ciągłej wielkości mierzonej możliwa jest identyfikacja i eliminacja wpływu nieznanych *a priori* zmiennych oddziaływań aperiodycznych i periodycznych.

Poprawny metrologicznie wynik pomiaru powinien zawierać zmierzoną najbardziej prawdopodobną wartość x badanej wielkości X wraz z oceną miary jej niedokładności wyznaczoną w określony, znany i powszechnie zaakceptowany sposób. Zadania te mogą spełniać różnie zdefiniowane parametry metrologiczne i dla zapewnienia jednolitości opisu niezbędne są ustalenia międzynarodowe.

Miary niedokładności pomiarów: błędy i niepewności pomiarowe

Do opisu dokładności pomiarów i przyrządów pomiarowych od dawna stosowano pojęcie *błąd pomiaru.* W najprostszym przypadku, tj. w pomiarze pojedynczej war-

[2] *Menzurand* definiowany jest obecnie jako pojęcie szersze niż pojedyncza wielkość mierzona. Obejmuje ono też wieloparametrowe modele matematyczne obiektów pomiaru, takie jak zależność funkcyjna $y = f(x)$, współrzędne wektora $\boldsymbol{X}$, charakterystyka częstotliwościowa $F(\omega)$, wielkości opisywane liczbami zespolonymi, np. w obwodach prądu przemiennego itp.

tości wielkości mierzonej, występuje błąd bezwzględny jako różnica pomiędzy wartością sygnału przyrządu lub jego wskazania, a wartością rzeczywistą wielkości mierzonej w chwili pomiaru. Wartość rzeczywista wielkości mierzonej jest zwykle nieznana i wykonując odpowiednie pomiary można ją poznać tylko z określoną dokładnością. Przy szacowaniu błędów pomiaru wartość rzeczywistą zastępuje się wartością poprawną, np. parametrem wzorca lub wskazaniem dokładniejszego przyrządu kontrolnego, albo też wartościami znamionowymi parametrów przyrządów i urządzeń pomiarowych.

Wielkości mierzone, zależne od nich sygnały przyrządów i systemów pomiarowych oraz błędy pomiarowe podlegają zwykle zmianom, zarówno w czasie wykonywania pojedynczej serii (próbki) powtarzanych pomiarów, jak i w okresie wielokrotnych badań obiektu oraz w okresie użytkowania przyrządu [13, 14]. Jest to skutek różnych wewnętrznych jak i zewnętrznych oddziaływań. Błędy stałe podczas wykonywania pomiarów, lub zmieniające się w sposób zdeterminowany, nazywa się systematycznymi, zaś zmieniające się losowo – błędami przypadkowymi. Zarówno w błędzie chwilowym, jak i w jego wartościach opisujących serię pomiarów można wyróżnić składowe o wartościach wyznaczalnych w danym eksperymencie pomiarowym, jak i składowe niewyznaczalne. Pierwsze z nich dają się łatwo usuwać przez poprawki. Łączny maksymalny możliwy wpływ pozostałych szacuje się jako błąd graniczny dla najgorszego przypadku znaków i wartości błędów składowych.

Matematyczna teoria opisu błędów pomiaru została zapoczątkowana przez Gaussa i Laplace'a w XVIII wieku. Jest oparta na teorii prawdopodobieństwa, zakłada przypadkowy rozrzut wyników obserwacji pomiarowych i operuje *błędami przypadkowymi* oraz prawdopodobieństwem ich występowania w określonym przedziale wartości.

W opisie właściwości metrologicznych przyrządów i systemów pomiarowych oraz ich podzespołów producenci podają błędy graniczne jako tolerancje dla warunków znamionowych i dla zakresów dopuszczalnych zmian warunków pracy lub dla innych określonych ich wartości. Definicje błędów pomiarowych i wzory dla ich propagacji przy ogólnej i kilku podstawowych zależnościach funkcyjnych zestawiono w tabeli 1.

Na podstawie wzorów (tab. 1) wyznaczany jest pożądany rodzaj błędu wypadkowego (zagadnienie analizy) i błędu granicznego (najgorszy przypadek wartości i znaków jego składowych), np. dla rezystancji wypadkowej zestawu różnie połączonych rezystorów o znanych tolerancjach, dla pomiarów mocy za pomocą kilku przetworników o znanych parametrach metrologicznych lub pomiarów rezystancji mostkiem dekadowym lub cyfrowym. Dla zadanego dopuszczalnego błędu wypadkowego istnieje też zwykle niejednoznaczne zagadnienie odwrotne – znalezienie dopuszczalnych przedziałów błędów składowych i tolerancji elementów przyrządu lub zestawu pomiarowego (zagadnienie syntezy).

Wyrażanie niedokładności urządzeń i systemów pomiarowych oraz wykonywanych nimi pomiarów przez różne rodzaje błędów pomiarowych (systematyczne oraz przypadkowe, chwilowe i graniczne) trwało do lat 90. ubiegłego wieku [9–11].

Tab. 1. Rodzaje błędów pomiarowych i klasyczny sposób wyznaczania wyniku pomiaru
Tab. 1. Types of measurement errors and the classic method of determining the measurement result

Pojedynczy pomiar wartości *x* wielkości mierzonej	
Wartości chwilowe błędów	
Błąd bezwzględny $\Delta x = x - x_{rz}$	**Błąd względny** $\delta_x = \frac{x - x_{rz}}{x_{rz}} \approx \frac{\Delta x}{x}$
gdzie: x – wartość wskazana, x_{rz} – wartość poprawna	
Błędy graniczne (moduły)	
$\lvert\Delta x\rvert = \lvert x - x_{rz}\rvert_{\max}$	$\lvert\delta_x\rvert = \left\lvert\frac{x - x_{rz}}{x_{rz}}\right\rvert_{\max} \approx \left\lvert\frac{\Delta x}{x}\right\rvert_{\max}$
Zapis wyniku pomiaru $x \pm \lvert\Delta x\rvert$	
Wyznaczanie błędów wartości *y* z zależności funkcyjnej pomiarów pośrednich x_i $y = f(x_1, x_2, \ldots x_n)$	
Błędy chwilowe	
$\Delta y \approx c_1\Delta x_1 + c_2\Delta x_2 + \ldots$ gdzie: $c_1 = \frac{\partial f}{\partial x_1}$, $c_2 = \frac{\partial f}{\partial x_2}, \ldots$	$\delta_y \approx a_1\delta_{x1} + a_2\delta_{x2} + \ldots$ gdzie: $a_1 = \frac{\partial f}{\partial x_1}\frac{x_1}{y}$, $a_2 = \frac{\partial f}{\partial x_1}\frac{x_2}{y}, \ldots$
Błędy graniczne	
$\lvert\Delta y\rvert = \lvert c_1\rvert\lvert\Delta x_1\rvert + \lvert c_2\rvert\lvert\Delta x_2\rvert + \ldots$	$\lvert\delta_y\rvert = \lvert a_1\rvert\lvert\delta_{x1}\rvert + \lvert a_2\rvert\lvert\delta_{x2}\rvert + \ldots$
Błędy przypadkowe nieskorelowane	
$\bar{\Delta} y = \sqrt{c_1^2\,\bar{\Delta} x_1^2 + c_2^2\,\bar{\Delta} x_2^2 + \ldots}$ (ρ_{ij}=0)	$\bar{\delta}_y \approx \sqrt{a_1^2\,\bar{\delta}_{x1}^2 + a_2^2\,\bar{\delta}_{x2}^2 + \ldots}$ (ρ_{ij}=0)
Błędy przypadkowe skorelowane	
$\bar{\Delta} y = \sqrt{c_1^2\,\bar{\Delta} x_1^2 + c_2^2\,\bar{\Delta} x_2^2 + \ldots 2c_1c_2\rho_{12}\bar{\Delta} x_1\,\bar{\Delta} x_2 + \ldots}$	$\bar{\delta}_y \approx \sqrt{a_1^2\,\bar{\delta}_{x1}^2 + a_2^2\,\bar{\delta}_{x2}^2 + \ldots 2a_1a_2\rho_{12}\bar{\delta}_{x1}\bar{\delta}_{x2} + \ldots}$
gdzie: ρ_{ij} współczynnik korelacji błędów wielkości x_i, x_j ($-1 \le \rho_{ij} \le 1$)	
Podstawowe zależności funkcyjne i ich błędy	
$y = k_1x_1 \pm k_2x_2, \ldots \pm k_nx_n$ (przy k_i=*const* ≥0)	
$\Delta y = k_1\Delta_{x1} \pm k_2\Delta_{x2}, \ldots \pm k_n\Delta_{xn}$	$\delta_y = k_1\frac{x_1}{y}\delta_{x1} \pm k_2\frac{x_2}{y}\delta_{x2} \pm \ldots$
$\lvert\Delta y\rvert = k_1\lvert\Delta_{x1}\rvert + k_2\lvert\Delta_{x2}\rvert, \ldots + k_n\lvert\Delta_{xn}\rvert$	$\lvert\delta_y\rvert = \left\lvert k_1\frac{x_1}{y}\right\rvert\lvert\delta_{x1}\rvert + \left\lvert k_2\frac{x_2}{y}\right\rvert\lvert\delta_{x2}\rvert + \ldots$
$y = k_M x_1^{\alpha} x_2^{\beta} \cdot \ldots \cdot x_n^{\vartheta}$ (przy k_M = *const*)	
$\Delta y = y\left(\frac{\alpha}{x_1}\Delta_{x1} + \frac{\beta}{x_2}\Delta_{x2} + \ldots + \frac{\vartheta}{x_n}\Delta_{xn}\right)$	$\delta_y = \alpha\,\delta_{x1} + \beta\,\delta_{x2} + \ldots + \vartheta\delta_{xn}$
Zapis wyniku pomiaru $y = \bar{y} \pm (\lvert\Delta y\rvert + \bar{\Delta} y)$	

Prawdopodobieństwo wystąpienia błędu granicznego (dla najgorszego przepadku) jest bardzo małe. Opis dokładności przyrządów i pomiarów poprzez ten błąd nie w pełni odpowiadał potrzebom zarówno służb metrologicznych, jak i pomiarów

użytkowych. Inny ujednolicony sposób opisu zaproponowano w publikacji „*Guide to the Expression of Uncertainty in Measurement*" (GUM) opracowanej wspólnie przez siedem organizacji międzynarodowych i wydanej po raz pierwszy przez Międzynarodową Organizację Standaryzacji ISO w 1993 r. [1].

Obecnie kolejne wersje tego Przewodnika i jego uzupełnienia, w tym Suplementy publikuje Międzynarodowe Biuro Miar i Wag BIPM [1–6]. Przewodnik ten zawiera reguły szacowania i wyrażania wyniku pomiarów oraz jego niedokładności o wszystkich istotnych w praktyce poziomach. W Przewodniku GUM jako miarę oceny niedokładności przy sprawdzaniu przyrządów, jak i dla wyników wszelkich pomiarów, wprowadzono jako pojęcie podstawowe – ***uncertainty***[3]. Ma ono charakter losowy i zdefiniowane jest inaczej niż błędy, tj. względem oszacowanej wartości wyniku pomiaru, czyli bez wykorzystywania pojęcia – wartość rzeczywista, którą w praktyce daje się określić jedynie teoretycznie. Jako odpowiednik tego nowego terminu, w polskiej literaturze metrologicznej zaczęto używać ***niepewność*** i to nawet kilka lat wcześniejszej niż ukazało się tłumaczenie Przewodnika GUM [2][4]. Niemal wszystkie kraje przyjęły zalecenia GUM do stosowania w Służbie Miar, akredytowanych laboratoriach oraz w innych placówkach – od pomiarów handlowych do pomiarów w badaniach technicznych i podstawowych.

Podstawowe definicje niepewności pomiarów

W Przewodniku GUM podano następującą definicję ogólną:

> **Niepewność** – parametr związany z wynikiem pomiaru, charakteryzujący rozrzut wartości mierzonej, który można w uzasadniony sposób przypisać wielkości mierzonej.

Tego terminu używa się też w nazwach pojęć pochodnych, ujętych tu opisowo wraz z używanymi dalej w tekście oznaczeniami:

- **niepewność standardowa** $u_A(x)$ **–** odchylenie standardowe, wyznaczane statystycznie, tj. metodą typu A z wartości obserwacji pomiarowych wg wzorów takich, jak dla rozkładu normalnego,
- **niepewność standardowa** $u_B(x)$ **–** szacowana metodą typu B z modelowania wpływów różnych oddziaływań,
- **niepewność standardowa złożona** $u(x)$ **–** dla wypadkowej kompozycji rozkładów przy opisie niedokładności wyniku pomiaru,
- **niepewność rozszerzona** $U_p(x)$ – wielokrotność $k_p u(x)$ jako połowa szerokości przedziału $\boldsymbol{x}$ o ustalonym poziomie ufności p.

[3] Pojęcie statystyczne o angielskiej nazwie ***uncertainty***, użył po raz pierwszy George Airy w końcu XIX wieku. Jako jego odpowiednik w literaturze polskiej utrwalił się termin **niepewność**. W GUM wprowadzono też szereg pojęć pochodnych, np. ***standard uncertainty*** – niepewność standardowa i ***expanded uncertainty*** **–** niepewność rozszerzona**.**

[4] Tłumaczenie polskie przewodnika GUM ukazało się dopiero w 1999 r. pod auspicjami Głównego Urzędu Miar jako: *Wyrażanie Niepewności Pomiaru. Przewodnik* [2]. Opracował je i bogatym komentarzem opatrzył prof. Janusz Jaworski z Politechniki Warszawskiej.

Samo pojęcie *niepewność* i jego nazwa, pomimo że utrwaliła się już w piśmiennictwie polskim, wzbudzała zastrzeżenia niektórych metrologów[5]. Według [2] komentarza prof. Janusza Jaworskiego do polskiego tłumaczenia przewodnika GUM termin ten jest niejednoznaczny. Używa się go zarówno jako pojęcie ogólne – cechę pomiaru, jak i wg powyższej definicji podstawowej – również dla parametrów rozkładu prawdopodobieństwa wyników pomiaru jako pojęć szczegółowych. Zmarły niedawno wybitny polski specjalista z dziedziny podstaw teoretycznych techniki pomiarowej prof. Janusz Piotrowski z Politechniki Śląskiej [13, 14] uważał, że należało przyjąć inne polskie słowo, np. *niedokładność.* Zdaniem autora i w nawiązaniu do dyskusji między profesorami: mechanikiem Janem Obalskim i elektrykiem Arturem Metalem prowadzonej w latach 60. Ubiegłego wieku na łamach dziewięciu numerów czasopisma „Pomiary Automatyka Kontrola", można nawet też było rozpatrzyć reaktywowanie terminu używanego wówczas przez pomiarowców elektryków – *uchyb pomiaru*[6], gdyż np. *wystrzelona kula ani nie błądzi, ani nie jest niepewna, lecz chybia celu*. Ponadto termin *niepewność* ma też pewien odcień pejoratywny. Jednakże stosuje się ją już teraz szeroko w polskim piśmiennictwie i w praktyce pomiarowej. Zapewne jest za późno na zmianę.

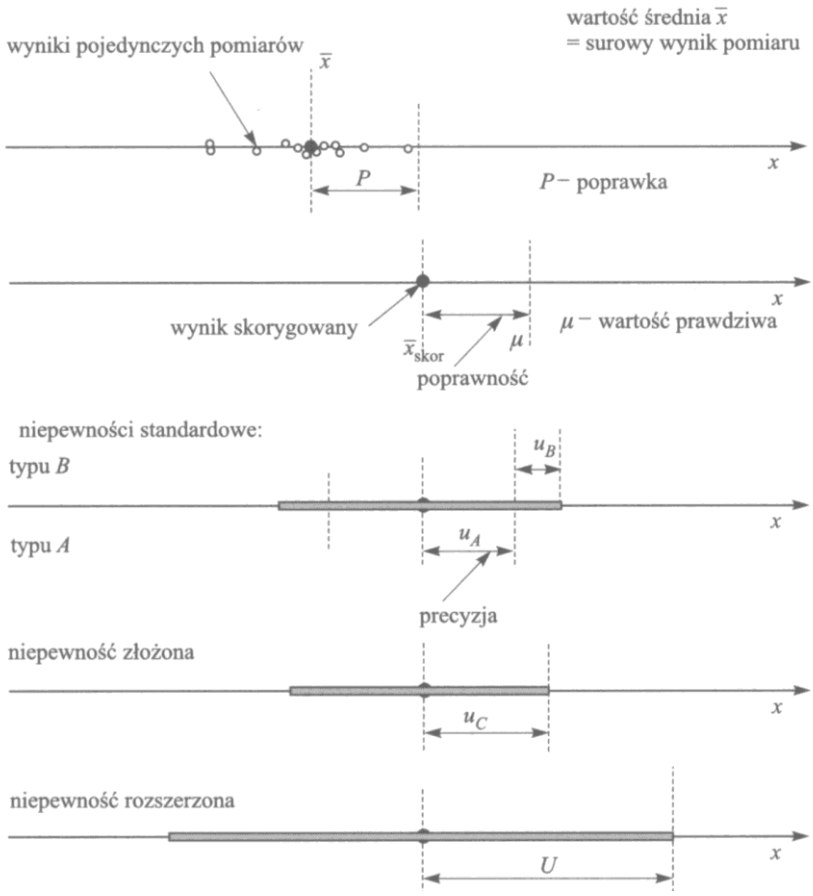
Na rysunku 2 przedstawiono poglądowo występowanie poszczególnych składowych niepewności wyniku pomiarów wartości x wielkości X wraz z definicjami poprawności i precyzji stosowanymi w opisach dokładności badań laboratoryjnych w chemii i innych dziedzinach eksperymentalnych [13].

Zalecenia Przewodnika GUM [1] dotyczą postępowania ze zbiorem pozyskanych obserwacji pomiarowych q_i wielkości mierzonej, czyli menzurandu (wg ostatniej polskiej wersji słownika VIM [8]). Wartości obserwacji koryguje się przez poprawki niwelujące błędy systematyczne znane w danym eksperymencie pomiarowym. Następnie skorygowane obserwacje traktuje się tak, jakby stanowiły próbę statystyczną z populacji o normalnym rozkładzie prawdopodobieństwa. Wyznacza się jej wartość średnią $\overline{q}$ i odchylenie standardowe. Do oceny niedokładności wynikającej z rozproszenia wyników obserwacji pomiarowych podstawę stanowi standardowe odchylenie wartości średniej nazwane niepewnością standardową, wyznaczane metodą statystyczną typu A i oznaczone jako $u_A(x)$.

Miarę niedokładności wyniku pomiaru wywołaną łącznym wpływem wszystkich oddziaływań, których na bieżąco nie można zidentyfikować ilościowo by skorygować wyniki, szacuje się poprzez niepewność wyznaczaną metodą typu B. Rozkład prawdopodobieństwa i wartość standardową $u_B(x)$ tej niepewności ocenia się jako występujące przez dłuższy okres, niż trwało uzyskanie danej serii obserwacji. Wyznacza się je na podstawie struktury systemu pomiarowego, właściwości użytych przyrządów, przewidywanych zakresów i rozkładów wartości czynników oddziałujących na wartość wyniku pomiarów oraz ich współczynników wpływu.

[5] Niepewność pomiarów zdefiniowano w Przewodniku bez stosowania pojęcia – wartość rzeczywista, wykorzystywanej dla błędów pomiarowych. Stoi to w sprzeczności z definicją podawaną w innym dokumencie Międzynarodowego Biura Miar i Wag BIPM, tj. w Międzynarodowym Słowniku Metrologii o akronimie VIM [7, 8].

[6] W innych językach słowiańskich, np. w ukraińskim, dla błędu pomiarowego stosowany jest termin zbliżony fonetycznie do uchybu: *pochibka*, a w rosyjskim – *pogreshnost* lub *oshibka*.



Rys. 2. Składowe niepewności wyniku pomiaru wg przewodnika GUM [1]
Fig. 2. Components of the uncertainty of the measurement result by GUM [1]

Obie niepewności składowe w zaleceniach GUM wyznaczane są zgodnie z zasadami teorii prawdopodobieństwa, ale inaczej, tj.:

1. niepewność u_A – na podstawie obliczeń statystycznych dla serii (próbki) wyników obserwacji,
2. niepewność u_B – przez heurystyczne szacowanie łącznego oddziaływania wewnętrznych i zewnętrznych czynników wpływających przy znanych lub zakładanych *a priori* ich rozkładach prawdopodobieństwa oraz współczynnikach wpływu[7].

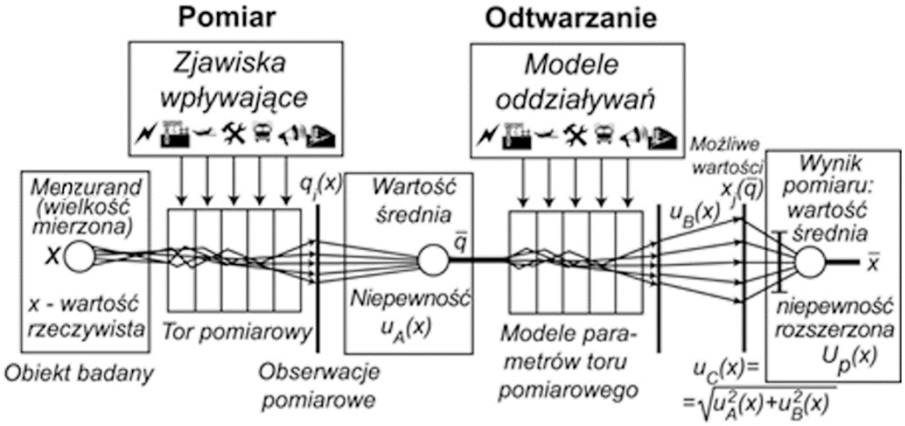
[7] Podejście statystyczne, niemal identyczne jak zastosowane w Przewodniku GUM, podał w Polsce około 40 lat wcześniej, tj. już w 1951 r., w swojej pracy doktorskiej Stanisław Trzetrzewiński, późniejszy profesor Politechniki Gdańskiej. Proponował on, by jako najbardziej prawdopodobny liczyć średni błąd graniczny. Ta pionierska praca była napisana po polsku i pewnie dlatego jej wyniki nie zostały wówczas spopularyzowane międzynarodowo.

Standardową niepewność u_B danego eksperymentu pomiarowego szacuje się według GUM metodą typu B randomizując spodziewane, lecz o wartościach nieznanych w danym eksperymencie pomiarowym, jej składowe jako skutki oddziaływań różnych wielkości wpływających. Dokonuje się tego zamiast stosowanej poprzednio analizy „najgorszego przypadku" dla sumy podstawowych i dodatkowych błędów granicznych użytych przyrządów oraz uwzględnia się przybliżenia operacji matematycznych opisujących dany eksperyment pomiarowy.

Niepewności standardowe wyznaczane metodami typu A i B są traktowane w przewodniku GUM jako statystycznie od siebie niezależne. Dla oceny dokładności wyniku pomiarów wartości menzurandu tworzy się rozkład wypadkowy jako kompozycję (splot) rozkładów niepewności składowych, w tym metodą Monte Carlo [5]. W postępowaniu uproszczonym, podanym we wszystkich wersjach przewodnika GUM, odchylenie standardowe $u_A(x)$ wartości średniej próbki (oszacowane na podstawie zbioru obserwacji pomiarowych) i heurystycznie oszacowaną niepewność standardową $u_B(x)$ składa się geometrycznie, tj. wg wzoru

$$u_C(x) = \sqrt{u_A^2(x) + u_B^2(x)} \,.$$

Wynik pomiarów charakteryzuje się przez skorygowaną wartość średnią $\bar{x}$ obserwacji w próbce, wyznaczoną po wprowadzeniu znanych poprawek i przez jej wypadkową niepewność standardową $u(x)$. Z niej wyznaczana jest niepewność rozszerzona $U_P(x)$ przy przyjętym arbitralnie współczynniku rozszerzenia k_P, np. jak dla rozkładu normalnego $k_P = 2$ lub $k_P = 3$, odpowiednio dla prawdopodobieństwa $P = 0{,}95$ lub $P = 0{,}99$ lub metodą Monte Carlo dla splotu rozkładów prawdopodobieństwa obu niepewności u_A i u_B. Jest to, symetryczny względem wartości średniej, przedział występowania wartości wielkości mierzonej o szerokości $2U_P(x)$, odpowiadający tym prawdopodobieństwom.



Rys. 3. Proces pomiarów i tworzenia wyniku menzurandu wraz z oceną jego niepewności
Fig. 3. The process of measurement and evaluation of the result value and uncertainty of the measurand

Na rys. 3 przedstawiono proces pomiarowy z rys. 1 z zaznaczonymi miejscami, w których wyznaczane są poszczególne niepewności wg zaleceń GUM. Wzory, które przy opracowywaniu opisu wyniku pomiarów służą do wyznaczania poszczególnych parametrów, w tym składowych niepewności, zestawiono w tabeli 2.

Więcej informacji, z przykładami obliczeń, można znaleźć w kolejnych opublikowanych wersjach przewodnika GUM, wraz z dotychczas wydanymi czteroma jego Suplementami [3–6] oraz w podstawowej literaturze o niepewnościach, polskiej [12–18, 37] i bardzo bogatej, głównie anglojęzycznej, literaturze zagranicznej [21–36].

Tab. 2. Parametry wyniku pomiaru (wartość średnia i niepewności) obliczane wg GUM [1]
Tab. 2. Parameters of the measurement result (the mean value and uncertainty) calculated according to GUM [1]

<table>
<tr><td colspan="2">Pomiary pojedynczej wartości x wielkości mierzonej</td></tr>
<tr><td colspan="2">Skorygowane wartości n obserwacji (próbka):
$q_1, q_2, ..q_i, ..., q_n$</td></tr>
<tr><td colspan="2">Wartość wyniku pomiarów (estymator menzurandu)
Jest to parametr charakterystyczny dla określonego rodzaju rozkładu obserwacji q_i
np. dla rozkładu normalnego najbardziej prawdopodobna jest
wartość średnia obserwacji pomiarowych
$\bar{x} = \bar{q} = \frac{1}{n}\sum_{i=1}^{n} q_i$</td></tr>
<tr><td>Niepewność standardowa typu A
$u_A(x)$
obliczana statystycznie ze zbioru wartości obserwacji jako:
odchylenie standardowe średniej
$u_A(x) = s(\bar{q}) = \frac{s(q_i)}{\sqrt{n}} = \sqrt{\frac{1}{n(n-1)}\sum_{i=1}^{n}(q_i - \bar{q})^2}$
gdzie: $s(q_i)$, $s(\bar{q})$ – odchylenia standardowe próbki i średniej.</td><td>Niepewność standardowa typu B
$u_B(x)$
szacowana z wiedzy o oddziaływaniach, ich parametrach i rozkładach jako:
odchylenie standardowe
$u_x(x) = \sqrt{\sum_{i=1}^{m} u_B^2\left(x|_{qi}\right) + 2\sum_{i=1}^{m-1}\sum_{j+i=1}^{m} r_{ij} u_B\left(x|_{qi}\right) u_B\left(x|_{qj}\right)}$
gdzie: r_{ij} – współczynniki korelacji pomiędzy składowymi niepewności od różnych wielkości wpływających.</td></tr>
<tr><td colspan="2">Złożona niepewność standardowa $u_C(x)$
wyniku pomiaru wartości x
$u_C(x) = \sqrt{u_A^2(x) + u_B^2(x)}$</td></tr>
<tr><td colspan="2">Niepewność rozszerzona $U_p(x)$ o poziomie ufności p
$U_p(x) = k_p u_C(y)$
gdzie: $k_p = t_p\left(v_{eff}\right)$ – współczynnik rozszerzenia, t_p – współczynnik z rozkładu Studenta
dla założonego poziomu ufności p i efektywnej liczby stopni swobody v_{eff}</td></tr>
<tr><td colspan="2">Zapis wyniku pomiaru
$x = \bar{x} \pm U_p(x)$</td></tr>
</table>

<table>
<tr><th colspan="2">Pomiary pośrednie wartości y wielkości mierzonej</th></tr>
<tr><td colspan="2">Równanie pomiaru
$y = f(x_1, x_2, \ldots x_m)$</td></tr>
<tr><td colspan="2">Wartość wyniku pomiaru
$\bar{y} = f(\bar{x}_1, \bar{x}_2, \ldots \bar{x}_m)$</td></tr>
<tr><td colspan="2">Niepewność standardowa złożona wielkości y
$u(y)$</td></tr>
<tr><td>nieskorelowane argumenty
$u(y) = \sqrt{\sum_{i=1}^{m} c_i^2 u_C^2(x_i)}$</td><td>skorelowane argumenty o wsp. korelacji r_{ij}
$u(y) = \sqrt{\sum_{i=1}^{m} u_C^2(x_i) + 2\sum_{i=1}^{m-1}\sum_{\substack{j=1 \\ j\neq i}}^{m} r_{ij}\, c_i u_C(x_i) c_j u_C(x_j)}$</td></tr>
<tr><td colspan="2">Niepewność rozszerzona $U(y)$ o poziomie ufności p
$U(x) = k\,u(y)$
gdzie: k – przyjęty arbitralnie współczynnik rozszerzenia (2 lub 3) lub wyznaczony dla rozkładu $u(y)$ jako splotu rozkładów $u_C(x_i)$ otrzymanego metodą Monte Carlo [2]</td></tr>
<tr><td colspan="2">Zapis wyniku pomiaru
$y = \bar{y} \pm U(y)$</td></tr>
</table>

Krótka charakterystyka literatury o niepewności pomiarów

Rozdział ten wprowadza ogólnie we współczesne zagadnienia wyznaczania wartości i miar niedokładności pomiarów. W jego bibliografii podano międzynarodowe przepisy zawierające zalecenia GUM oraz zestaw najważniejszych pozycji literaturowych, w tym podstawowe książki polskie oraz angielskie i inne. Poruszane w tej literaturze zagadnienia są opracowane na różnym poziomie, jednak tak, by były też w dużym stopniu dostępne dla średniej i wyższej kadry technicznej w przemyśle, energetyce i eksploatacji oraz dla szeregowych pracowników izb pomiarowych, akredytowanych i innych laboratoriów badawczych oraz wszystkich zainteresowanych opracowywaniem wyników pomiarów. Ogromu wszystkich publikacji na ten temat w czasopismach polskich i zagranicznych nie udałoby się tu zamieścić. Ograniczono się tu do wyboru pozycji najbardziej istotnych.

W treści Przewodnika GUM [1, 2] i w jego Suplementach [5–8] omówione są wszechstronnie obecnie przyjęte sposoby szacowania niepewności standardowej i rozszerzonej przy ograniczonej liczbie statystycznie niezależnych obserwacji pomiarowych z uwzględnieniem rozkładu Studenta i dla splotu ich rozkładów przy pomiarach pośrednich jedno- i wieloparametrowych.

Wyznaczanie niepewności komputerowo realizowaną metodą statystyczną Monte Carlo omawia Suplement 1 do GUM [3]. Opracowano też wiele procedur obliczeniowych i programów komputerowych, w tym dostępnych w Internecie, np. na stronach amerykańskiego Narodowego Instytutu Wzorców i Technologii NIST.

Kierunki dalszych prac międzynarodowych dotyczących zmian treści obecnej wersji przepisów GUM są zarysowane w publikacji W. Bicha i innych [19].

Ograniczenia w stosowaniu zaleceń Przewodnika GUM i jego Suplementów oraz zadania do opracowania stąd wynikłe

Przewodnik GUM uporządkował ujęcie metod wyznaczania i opisu niedokładności wyniku pomiarów, aby ujednolicić je w skali międzynarodowej do celów metrologicznych. Zawarte w nim metody szacowania dokładności wzorców, przyrządów pomiarowych i pomiarów opisane poprzez ich niepewności przyjęły do stosowania niemal wszystkie narodowe służby miar. Szersze ich wykorzystanie w rutynowej praktyce pomiarowej i badaniach natrafia jeszcze na szereg ograniczeń.

Z surowych wyników obserwacji pomiarowych nie można wyeliminować bieżących stałych wpływów oddziaływań, gdyż prawdziwa wartość x nie jest znana. Mogą też być one inne w każdej z serii obserwacji w ciągu całego okresu eksploatacji przyrządu lub badania obiektu, w tym wskutek zmian warunków pomiaru. Potrzebny jest tu bądź dodatkowy pomiar kontrolny, bądź bardziej doskonałe przyrządy i systemy o wewnętrznym zerowaniu i wzorcowaniu. W zbiorze obserwacji można natomiast zidentyfikować i usunąć wpływy oddziaływań zdeterminowanych progresywnych (aperiodycznych) i periodycznych, jeśli sposób próbkowania wielkości mierzonej jest znany, np. jest to zwykle stosowane regularne próbkowanie. Wartość średnia wyniku pomiaru nie ulega zmianie jedynie przy występowaniu składowej progresującej liniowej i dla wszystkich składowych oscylacyjnych o okresach będących parzystą wielokrotnością odstępu czasu ΔT próbkowania równomiernego. Powiększają one jednak odchylenie średnie standardowe i dlatego powinny też zostać usunięte z surowych wyników obserwacji przed dalszym ich przetwarzaniem. W zaleceniach GUM i jego Suplementach nie podano jak to uczynić. Najbardziej zaawansowane technicznie przyrządy i systemy pomiarowe są wyposażone w odpowiednie podzespoły, przetworzenia sygnału i oprogramowanie obejmujące procedury do identyfikacji i eliminacji wpływów składowych systematycznych o przebiegach zdeterminowanych: stałych oraz zmiennych periodycznie i aperiodycznie. Bardziej rozbudowane zestawy i systemy pomiarowe zawierają też automatyczne wzorcowanie i korekcję zera, filtrację cyfrową i przekształcenie Fouriera wartości obserwowanych sygnałów w zadanym oknie czasowym itp.). Nie ma jeszcze przepisów o jednolitym sposobie ujęcia niepewności przy takich przekształceniach.

W wielu pomiarach w praktyce laboratoryjnej i przemysłowej trzeba na bieżąco radzić sobie możliwie najprostszymi sposobami, wspomaganymi jedynie poprzez zewnętrzne, obecnie na ogół realizowane komputerowo, przetwarzanie zebranych wyników obserwacji pomiarowych.

Metody wyrażania niepewności zawarte w Przewodniku GUM i jego zaleceniach nie ujmują opisu niedokładności wielu pomiarów występujących w praktyce laboratoryjnej i przemysłowej. Na przykład jako najkorzystniejsze parametry (wartość średnia, odchylenie średnie standardowe) i wzory do obliczeń niepewności u_A wybrano takie, jak dla opisu danych pomiarowych rozkładem normalnym. Tymczasem dla innych rozkładów lepsze są inne estymatory menzurandu i ich niepewności. Aby to zagadnienie przybliżyć, w Dodatku 1 porównano oceny liczbowe wartości i dokładności menzurandu dla próbki z populacji o rozkładzie równomiernym oraz o rozkładzie Laplace'a. Obliczone je wg GUM dla średniej oraz dla bardziej efektywnego estymatora, tj. o mniejszym odchyleniu standardowym, każdego z rozkładów.

Zalecenia GUM nie ujmują też przypadków, w których wielkość mierzoną, zmiany parametrów obiektu mierzonego i toru pomiarowego oraz wielkości wpływające należy modelować w funkcji czasu (lub współrzędnych przestrzennych, np. w pomiarach optycznych i badaniu rozkładu pól), a więc procesami stochastycznymi stacjonarnymi i niestacjonarnymi, jako opisem najbliższym rzeczywistości [13, 14].

Dotychczas przy wyznaczaniu niepewności wg GUM nie bada się i nie uwzględnia wpływu autokorelacji między obserwacjami niedaleko odległymi w kolejności zbierania, czyli opisywanych szeregami losowymi, np. w dziedzinie czasu. Autokorelacja występuje przy szybkim próbkowaniu wskutek ograniczonego widma szumów i dryftów niskoczęstotliwościowych. Dotyczy to wielu pomiarów stosowanych w badaniach, naukowych i technicznych, w przemyśle i eksploatacji przy zbieraniu danych pomiarowych w różnych przedziałach czasowych.

Klasyczne metody statystyczne przetwarzania danych opierają się na założeniu modelowania prawdopodobieństwa rozrzutu ich wartości rozkładem normalnym. Jednak dla szeregu przypadków w praktyce liczba danych próbki nie jest wystarczająca do budowy w pełni wiarygodnych ich modeli parametrycznych. Ponadto w latach 60. ubiegłego wieku statystycy, na podstawie wyników szczegółowych badań ustalili, że rzeczywiste dane eksperymentalne, które przetwarzano przy założeniu o rozkładzie normalnym, zawierają zwykle średnio około 10% (o zakresie od 1% do 20%) danych odstających od tego rozkładu. Poprzednio, ze względu na żmudne obliczenia, opracowywano tylko dane o wysokiej jakości, a teraz, przy powszechnym stosowaniu komputerów można przetwarzać prawie każde dane i o innych rozkładach niż normalny.

W praktyce pomiarowej zdarza się też, że wartości jednej lub kilku obserwacji w próbce istotnie odstają od pozostałych danych (ang. nazwa *outliers*). Dla próbki o małej liczbie elementów nawet niewielka ich liczba może znacząco zmienić wyznaczane klasycznie jej parametry statystyczne. W wielu przypadkach nie ma możliwości powtórzenia, lub uzupełnienia pomiarów. Wartość wyniku pomiarów i niepewność statystyczna u_A, obliczone w dotychczas stosowany sposób wg zasad międzynarodowego przewodnika GUM, mogą okazać się niewiarygodne, a nawet sprzeczne ze zdrowym rozsądkiem. Aby uniknąć takiej sytuacji, tradycyjnie stosuje się testy statystyczne. Pozwalają one zidentyfikować, a następnie usunąć nietypowe, tj. odstające wartości danych. Postępowanie takie jest skuteczne tylko dla dużych próbek, gdyż dla małych testy te tracą wrażliwość na wartości odstające. Dotychczas takie testy opracowano tylko dla próbek z populacji o rozkładzie normalnym. Jeżeli w pomiarach uzyskać można próbki tylko o małej liczbie danych, np. przy dużym koszcie pojedynczego pomiaru, w badaniu metodą niszczącą nielicznych dostępnych obiektów fizycznych, lub gdy pomiarów nie można powtórzyć, to trzeba starać się wykorzystać wszystkie otrzymane dane. Usunięcie jednej lub kilku obserwacji z małej próbki zmniejsza wiarygodność jej oceny statystycznej. Do analizy tych przypadków opracowano tzw. metody odporne. Wykorzystują one również dane odstające traktując je w inny sposób niż pozostałe dane. Metody te również nie są objęte rekomendacjami GUM, ale występują już w przepisach dotyczących chemicznych i innych badań laboratoryjnych.

Do oceny miar dokładności wieloparametrowych pomiarów pośrednich, opisywanych liniowymi i nieliniowymi zależnościami, dopiero ostatnio ukazały się re-

komendacje BIPM – jak liczyć niepewności, w postaci Suplementu 2 [4]. Obejmują jednak one głównie przypadki, gdy rozrzuty obserwacji wielkości bezpośrednio mierzonych opisuje się rozkładami normalnymi.

Metody wyznaczania niepewności wg Przewodnika GUM i jego Suplementów nie obejmują też jeszcze analizy dokładności pomiarów dynamicznych oraz śledzenia niedokładności w cyfrowym przetwarzaniu sygnałów pomiarowych o różnych algorytmach. Od kilku lat zapowiadany jest jednak kolejny Suplement GUM o szacowaniu dokładności wielkości zmiennych w czasie.

W opisie właściwości metrologicznych przyrządów i ich podzespołów nadal stosuje się błędy graniczne. Podaje się je dla znamionowych parametrów urządzeń i do opisu dopuszczalnych przedziałów zmian wartości parametrów w czasie pracy (innych w laboratorium, przemyśle oraz dla eksploatacji urządzeń w środowisku otwartym, np. w transporcie i dla techniki wojskowej). Dla części pomiarów zarówno rutynowych, jak i stosowanych w badaniach naukowych i technicznych sposoby estymacji niepewności są znane w niewystarczającym stopniu, by je stosować szerzej w praktyce lub nawet jeszcze nie są opracowane.

Istniejące przepisy metrologiczne nie nadążają za bieżącymi potrzebami praktyki pomiarowej. Przedstawione w 2006 r. [19] zamierzenia dotyczące rozszerzenia treści przewodnika GUM zrealizowano dotychczas jako dwa Suplementy [3, 4]. Tylko częściowo zaspakaja to istniejące potrzeby. Skuteczne rozszerzenie zakresu stosowania w praktyce pojęcia niepewności jako miary dokładności pomiarów wymaga rozwinięcia metod zawartych w Przewodniku GUM i jego Suplementach ukierunkowanego na pomiary stosowane w praktyce różnych dziedzin. Zaawansowane są prace nad nową wersję GUM opartą na teorii prawdopodobieństwa warunkowego Thomasa Bayesa. Jednak z prac tych jeszcze nie wynika, czy podejście to wiele wniesie w przyszłości do szacowania dokładności pomiarów ciągłych procesów przemysłowych oraz do tych badań technicznych materiałów i urządzeń, w opisie których trzeba stosować rozkłady niegaussowskie i statystykę odporną. Dlatego też w kolejnych rozdziałach tej monografii przedstawione są niektóre wybrane propozycje w sposób niezbyt zmatematyzowany. Zdaniem autora mogą one okazać się już obecnie istotne w wielu badaniach użytkowych, gdyż z nich wynikły i dla nich są przeznaczone.

Literatura

1. Evaluation of measurement data – Guide to the expression of uncertainty in measurement (GUM), BIPM JCGM 100: 2008 (last corrected version)
2. *Wyrażanie Niepewności Pomiaru*. Przewodnik, tłumaczenie wydania GUM z 1995 r. z komentarzami J. Jaworskiego, Wydawnictwo Głównego Urzędu Miar Warszawa 1999.
3. Supplement 1 to GUM [1]. *Evaluation of measurement data – Propagation of Distribution using a Monte Carlo Method*. BIPM JCGM 101: 2008
4. Supplement 2 to GUM [1], *Evaluation of measurement data – Extension to any number of output quantities*. BIPM JCGM 102: 2011
5. *Evaluation of measurement data – An introduction to GUM and related documents*. BIPM JCGM 104: 2009

6. *Evaluation of measurement data – The role of measurement uncertainty in conformity assessment*. BIPM JCGM 106: 2012.
7. *International Vocabulary of Metrology – Basic and general concepts and associated terms* (VIM). Edition 3 BIPM JCGM 200: 2008.
8. *Międzynarodowy słownik metrologii. Pojęcia podstawowe i ogólne oraz terminy z nimi związane* (VIM). Polski Komitet Normalizacyjny Przewodnik. PKN-ISO/IEC Guide 99, 2009-07-28.
9. Jaworski J.: *Matematyczne Podstawy Metrologii*. Wydawnictwo Naukowo Techniczne Warszawa 1979.
10. Tylor J.R.: *Wstęp do analizy błędu pomiarowego.* Wydawnictwo Naukowe PWN, Warszawa 1995 (tłumaczenie oryginału ang.: *An Introduction to error analysis. The study of uncertainties in physical measurements*. Oxford University Press, California, 1982).
11. Rozenberg W.J.: *Wstęp do teorii błędów systemów pomiarowych*. PWN Warszawa 1982 (tłum. z ros. *Vviedenie v teoriu tochnosti izmeritelnykh sistem*. Moskva, Sovietskie Radio 1975).
12. Turzeniecka D.: *Analiza dokładności wybranych przybliżonych metod oceny niepewności*. Wydawnictwo Politechniki Poznańskiej 1999.
13. Piotrowski J., Kostyrko K.: *Wzorcowanie Aparatury Pomiarowej*. Wydawnictwo Naukowe PWN Warszawa wydanie 1. 2000, wydanie 2. zmienione i uaktualnione 2012.
14. Piotrowski J.: *Podstawy Miernictwa*. Wydawnictwa Naukowo-Techniczne Warszawa 2002.
15. Arendarski J.: *Niepewność Pomiarów.* Skrypt. Oficyna Wydawnicza Politechniki Warszawskiej, 2003.
16. Skubis T.: *Opracowanie wyników pomiarów. Przykłady.* Wyd. Politechniki Śląskiej Gliwice 2003.
17. Skubis T.: *Podstawy metrologicznej oceny wyników pomiaru.* Wyd. Politechniki Śląskiej Gliwice 2004.
18. Zięba A.: *Analiza danych w naukach ścisłych i technice*. Wydawnictwo Naukowe PWN Warszawa 2013.
19. Bich W., Cox M.G., Harris P.M.: *Evolution of the 'Guide to the Expression of Uncertainty in Measurement*. "Metrologia" 43(2006) 161–166.
20. Olejnik R.: *O pomiarze*. Wydawnictwo Politechniki Częstochowskiej.

Książki anglojęzyczne i inne o niepewności pomiarów

21. Dietrich C.F.: *Uncertainty, Calibration and Probability. The Statistics of Scientific and Industrial Measurement*. Second Edition 1991. Series in Measurement Science and Technology.
22. Hugh W. Coleman, W. Glenn Steele: *Experimentation and Uncertainty Analysis for Engineers*. Second Edition. John Wiley & Sons 1999, str. 275.
23. Roald H. Dieck: *Measurement Uncertainty: Methods and Applications*. ISA-The Instrumentation, Systems, and Automation Society, 3rd ed., 2002, str. 252.
24. Ignacio Lira: *Evaluating the Measurement Uncertainty. Fundamental and Practical Guidance*. Series in Measurement and Technology. Institute of Physics Publishing 2002, str. 243.

25. Shri Krishna Kimothi: *The Uncertainty of Measurements*. Physical and Chemical Metrology: Impact and Analysis. ASQ Quality Press 2002, str. 391.
26. Hanns L. Harney: *Bayesian Inference. Parameter Estimation and Decisions*. Springer-Verlag Berlin Heidelberg 2003, str. 263.
27. Stephen M. Stigler: T*he History of Statistics. The Measurement of Uncertainty before 1900*. Harvard University Press. Ninth edition, 2003, str. 410
28. Gertsbakh Ilya: *Measurement Theory for Engineers*. Springer-Verlag Berlin Heidelberg 2003, str. 150.
29. Rabinovich Semyon G.: *Measurement Errors and Uncertainties. Theory and Practice*. Springer Science and Media, 3rd edition, 2005, str. 308.
30. Wolfgang Fellin, Heimo Lessmann, Michael Oberguggenberger, Robert Vieider (Eds.): *Analyzing uncertaity in civil engineering*, Springer-Verlag Berlin Heidelberg New York 2005, str. 240.
31. Michael Grabe: *Measurement Uncertainties in Science and Technology*. Springer-Verlag Berlin Heidelberg 2005, str. 269.
32. Les Kirkup. Bob Frenkel: *An introduction to uncertainty in measurement using the GUM* (Guide to the expression of uncertainty in measurement). Cambridge University Press 2006, str. 233.
33. Bilal M. Ayyub, George J. Klir: *Uncertainty Modeling and Analysis in Engineering and the Sciences.* Chapman and Hall/CRC, Taylor and Francis Group 2006, str. 378.
34. Dennis V. Lindley: *Understanding Uncertainty*. A John Wiley & Sons, Inc. Publication 2006, str. 250.
35. Novitski P.V., Zograf I.A.: *Ocenka pogreshnostiej rezultatom izmierenii*, Energoatomizdat, Leningrad, 1985. str.248 (w języku rosyjskim).
36. Dorozhovets M.: *Opracovania rezultatov vimirovan*. Wydawnictwo Uniwersytetu „Politechnika Lwowska" Lviv 2007 str. 622 (w języku ukraińskim).
37. *Niepewność pomiarów w teorii i praktyce*. Praca zbiorowa pod redakcją P. Fotowicza, S. Kubisy S. Moskowicza i D. Suchockiej. Główny Urząd Miar. Warszawa 2011, str. 260.

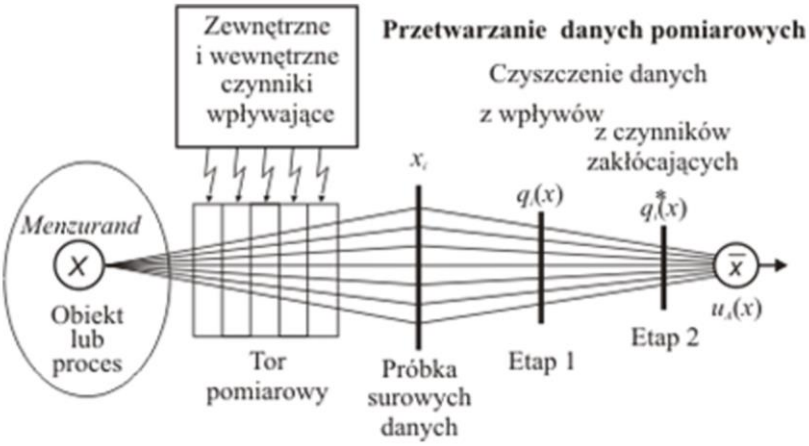
Rozdział 2

Wykrywanie i eliminacja wpływu dryftu i oscylacji w danych na niepewność pomiarów typu A

Streszczenie. Zaproponowano udoskonalenie metody typu A szacowania niepewności zalecanej przez przewodnik ISO GUM. Polega ono na wykryciu nieznanego *a priori* trendu i podstawowej składowej oscylacyjnej oraz ich wyeliminowaniu z surowych wyników obserwacji pomiarowych otrzymanych przy znanym, np. równomiernym próbkowaniu. Dla dwu przykładów liczbowych przedstawiono i omówiono wpływ takiego „czyszczenia" wyników surowych ze składowych systematycznych na wartość odchylenia standardowego próbki.

W wielu pomiarach przemysłowych i podczas eksploatacji, np. w badaniach diagnostycznych, często występują silne i nieznane *a priori* oddziaływania czynników zakłócających lub sygnał pomiarowy uzyskuje się łącznie z innymi zmiennymi sygnałami, zaś użyteczną informację zawierają tylko niektóre jego parametry. Wskutek tego, przy wielokrotnych pomiarach parametru o stałej wartości pojawiają się duże rozrzuty kolejnych wyników obserwacji o pochodzeniu nie tylko losowym, ale również zdeterminowanym, czyli systematycznym. Poprawny metrologicznie wynik pomiaru menzurandu powinien zawierać najbardziej prawdopodobną wartość badanej wielkości wraz z oceną jej dokładności wyznaczoną w określony, znany i powszechnie zaakceptowany jednolity sposób, np. według postępowania opisanego w międzynarodowym Przewodniku GUM [1] (patrz rozdział 1). Szacowanie składowej losowej u_A *niepewności pomiarów* metodą statystyczną typu A zalecaną przez GUM daje w powyższych przypadkach wartości zawyżone przy każdej liczbie obserwacji, gdyż na odchylenie standardowe rozrzutu wyników pomiaru wpływają też nieznane, a więc nieusunięte zakłócenia systematyczne o charakterze regularnym, takie jak aperiodyczny trend i oscylacje. Zakłócenia te należy wykryć i wyeliminować z surowych wyników pomiarowych, czyli „oczyścić" próbkę pomiarową.

Metoda tu proponowana dotyczy przypadków, gdy wartość x wielkości mierzonej występuje na tle zakłóceń i znana jest kolejność oraz wzajemne względne odległości poszczególnych obserwacji pomiarowych, np. pozyskano je przy regularnym próbkowaniu badanej wielkości. Proces pomiarowy należy analizować dla całego kanału pomiarowego, tj. od wartości wielkości mierzonej do wyjścia w postaci wyniku pomiaru menzurandu wraz z określoną miarą jego dokładności. Nie wyodrębnia się wpływów cząstkowych pochodzących od poszczególnych oddziaływań oraz od zmian parametrów wewnętrznych, ani nie lokalizuje miejsc ich występowania w obiekcie i torze pomiarowym. W opracowaniu tej tematyki współuczestniczył dr J. M. Korczyński [10–13].



Rys. 1. Model procesu pomiarowego o udoskonalonej ocenie szacowania niepewności typu A przez dodanie funkcji wykrywania i „czyszczenia" (filtracji) wyników surowych z nieznanych *a priori* zakłóceń – etap 2

Fig. 1. Model of the measurement process with improved assessment of estimating uncertainty of type A by adding detection and "clean" (filtering) the raw results from the unknown *a priori* interferences – stage 2

Proces pomiarowy był już omówiony szczegółowo w rozdziale 1. Wartość bieżąca x wielkości mierzonej X, czyli wartość menzurandu, dociera poprzez obiekt na wejście toru pomiarowego (rys. 1) i jest przetwarzana za pomocą czujników lub obwodów wejściowych w pierwotny sygnał pomiarowy. Podlega on następnie odpowiedniemu kształtowaniu – kondycjonowaniu, dalszemu przetwarzaniu i próbkowaniu, aż do otrzymania zbioru pojedynczych wartości sygnału w pożądanej formie lub zbioru ich odczytów, czyli tzw. surowych danych pomiarowych x_i. Proces ten nie jest idealny. Na obiekt badany i sygnał w torze pomiarowym oprócz wielkości mierzonej oddziałują warunki pomiaru oraz zakłócenia zewnętrzne i wewnętrzne, w tym zmiany parametrów toru. Jak już wspomniano oddziaływania te są różnego rodzaju, deterministyczne i przypadkowe. Zakłócenia pojawiają się zarówno podczas propagacji sygnału przez obiekt, jak i w początkowej części toru pomiarowego oraz następnie przy przetwarzaniu. Powodują, że poszczególne obserwacje x_i mają inne wartości niż mierzona wartość x badanej wielkości X. W pozostałej swojej części proces pomiarowy zawiera procedurę odtwarzania wartości wielkości mierzonej, czyli estymacji najbardziej prawdopodobnego wyniku pomiaru wraz z oceną jego dokładności wg założonego kryterium i przy przyjętych modelach oddziaływań, strukturze i parametrach toru pomiarowego. Zwykle estymatorem menzurandu jest wartość średnia $\bar{x} \equiv \bar{q}$ zbioru obserwacji skorygowanych przez znane poprawki. Jako miarę dokładności wyznacza się granice przedziału wartości mierzonej x, w którym wielkość ta będzie się znajdować z zadanym prawdopodobieństwem. W etapie 1 (rys. 1) wprowadzane są odpowiednie poprawki wtedy, gdy oddziaływania na sygnał pomiarowy są systematyczne (czyli zdeterminowane) i znane *a priori*. Otrzymuje się zbiór skorygowanych wartości sygnału lub odczytów $q_i(x_i)$.

W szeregu przypadkach, przez odpowiednie ukształtowanie procesu pomiarowego, można też wykryć i wyeliminować wpływy przejściowych zakłóceń przypadkowych i tych wszystkich oddziaływań systematycznych, które *a priori* nie są znane – etap 2 (rys. 1). Nie dotyczy to jedynie stałych co do wartości skutków oddziaływań w samym obiekcie. Są one nieznane w danym cyklu pomiarowym, gdyż brak jest informacji o wartości prawdziwej x menzurandu. Zakłócenia mogą być inne w każdej serii obserwacji w ciągu całego okresu eksploatacji przyrządu lub badania obiektu, w tym również wskutek zmian warunków pomiaru [3] i operacje takiego czyszczenia próbki muszą być każdorazowo powtarzane.

W rozdziale 1 podano, że szacowanie niepewności pomiarów wg Przewodnika GUM [1] odbywa się bez posługiwania się pojęciem: wartość rzeczywista lub prawdziwa, gdyż nie jest ona nigdy znana. Za najlepszy statystycznie wynik pomiaru menzurandu X (estymatę jego wartości) przyjęto w GUM arbitralnie wartość średnią $\bar{x} \equiv \bar{q}$ wyznaczoną ze zbioru n skorygowanych obserwacji q_i jako próby statystycznej z populacji wszystkich możliwych wartości sygnału. Ocena dokładności polega na wyznaczeniu dwu nieskorelowanych ze sobą składowych niepewności: niepewność standardową $u_A(x)$ wyznaczanej metodą statystyczną typu A jako odchylenie standardowe $s(\bar{q})$ średniej ze skorygowanych obserwacji próby i niepewność $u_B(x)$ szacowaną w inny zróżnicowany, sposób metodą B, w tym subiektywnie heurystycznie w oparciu o wcześniejszą wiedzę o tym procesie.

Sposób obliczania niepewności typu A i B wg GUM zestawiono w tabeli 2 (rozdz. 1). Jest on omówiony szczegółowo w licznej zagranicznej i polskiej metrologicznej literaturze książkowej.. Wyznaczanie niepewności $u_A(x)$ tylko nieznacznie zmodyfikowane w stosunku do stosowanego wcześniej dla pozornego błędu przypadkowego próby losowej z populacji o rozkładzie normalnym (Gaussa).

Teoretyczne założenia poprawności obliczeń są następujące:

1. Po wyeliminowaniu z obserwacji znanych składowych systematycznych próbkę traktuje się jako losową.
2. Wyniki obserwacji są niezależne statystycznie (nie są ze sobą skorelowane) i mają jednakowe wagi.
3. Parametry statystyczne, takie jak wartość średnia $\bar{x} \equiv \bar{q}$ oraz jej odchylenie standardowe $s(\bar{q})$, czyli niepewności typu A, dają się wyznaczyć i są właściwe do oceny wyniku i jego niepewności dla rozkładu prawdopodobieństwa, który z zadanym poziomem istotności odwzorowuje populację wyników obserwacji.

Zbiór surowych obserwacji, nawet po korekcji przez poprawki, nie jest jednak próbą z czysto losowej populacji, gdyż obserwacje te mogą zawierać zakłócenia systematyczne spowodowane różnymi niezidentyfikowanymi oddziaływaniami. Jednakże można istotnie zmniejszyć ich wpływ. Omówię to bardziej szczegółowo.

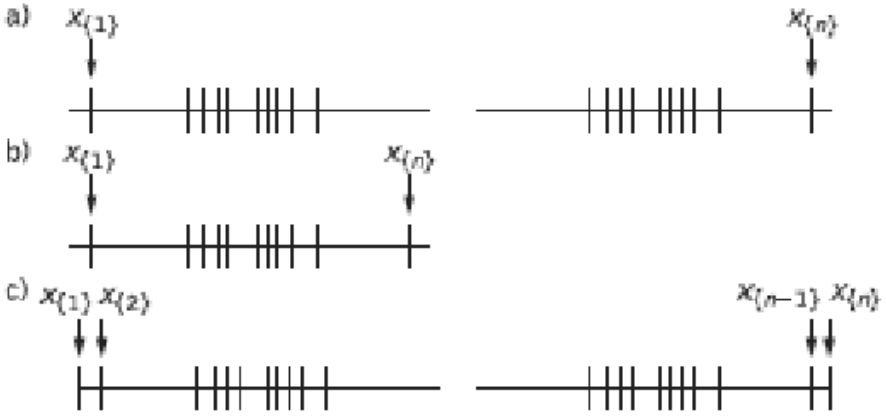
W stosunku do czasu zbierania obserwacji próbki (lub do ich liczby) zakłócenia występujące w obiekcie badanym i torze pomiarowym dzielą się na wolno- i szybkozmienne [3] oraz na ciągłe i impulsowe. Gdy dysponuje się dodatkową informacją o przebiegu procedury mierzenia w aparaturze pomiarowej, np. o powszechnie stosowanym równomiernym próbkowaniu oraz o jego częstotliwości i długości okresu zbierania danych próbki, to staje się możliwa identyfikacja i eliminacja wpływu niektórych z tych zakłóceń. Wymaga to bądź zastosowania on-line

odpowiednich procedur filtracji i przetwarzania wartości sygnałów surowych danych, bądź obliczeń *off-line* z odczytów ich wartości. Aby można było skutecznie oszacować statystyczną składową błędów – przez niepewność u_A, lub dla niektórych rozkładów innych niż normalny – przez jej odpowiednik w postaci standardowego odchylenia od wartości estymatora wyniku pomiaru korzystniejszego niż wartość średnia, np. środek rozpięcia dla rozkładu równomiernego, mediana dla rozkładu podwójnie wykładniczego, to obserwacje w próbce powinny być punktami stacjonarnego procesu losowego. Rzeczywista próbka może zawierać nieznane *a priori* zakłócenia różnego rodzaju, w tym:

- dane odstające, czyli tzw. outliery (od ang. *outliers*), zwykle krótkotrwałe impulsowe zakłócenia przypadkowe o wartościach istotnie różniących się od pozostałych oraz
- zakłócenia regularne (systematyczne), np. trend aperiodyczny i oscylacje.

Zakłócenia te należy wykryć i próbkę z nich oczyścić zanim przystąpi się do szacowania niepewności u_A według zaleceń Przewodnika GUM. Cel ten realizuje się w etapie 2 procesu pomiarowego z rys. 1. Zalecenia, jak takie zadania realizować, nie występują dotychczas w jawnej formie ani w podstawowych podręcznikach omawiających analizę i szacowanie niepewności pomiarów, zarówno polskich [3–5], jak i zagranicznych podanych w literaturze rozdziału 1, ani w GUM.

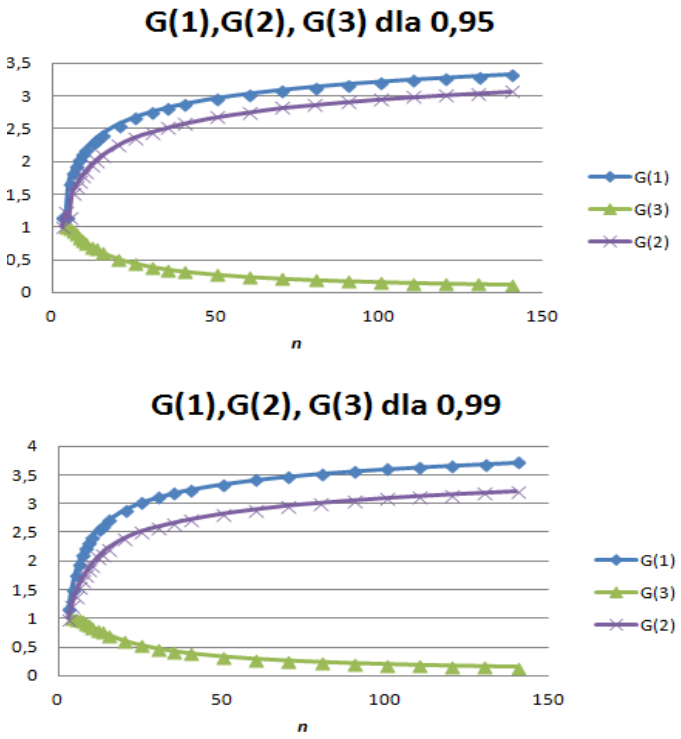
Do wykrywania outlierów w danych surowych próbek opracowano szereg testów. Jeśli dane te uzyskano z populacji o gaussowskiej składowej losowej[1], np. przez próbkowanie sygnału to zwykle używa się testów Grubbsa [5].



Rys. 2. Trzy przypadki wykrywania outlierów wg testów Grubssa: a) G(1), b) G(2), c) (3) [5]
Fig. 2. Three types of outlier identification by Grubbs tests: a) G(1), b) G(2), c) G(3) [5]

[1] Autor kilka lat temu na jednej z konferencji Podstawowych Problemów Metrologii przedstawił pogląd, że w pomiarach są też potrzebne testy dla wykrywania outlierów w próbkach z rozkładu równomiernego. Temat ten podjął prof. M. Dorozhovets i na ostatnim Kongresie Metrologii 2016 w Lublinie zaprezentował oryginalną propozycję takiego testu [16]. Wymaga ona jeszcze weryfikacji w praktyce pomiarowej.

Testy te polegają na kontroli, czy krańcowe dane nie przekraczają wartości uznanej za odstającą, czyli za outlier. Odległość granic zależy od liczby potencjalnych outlierów: pojedynczy G(1), dwa z jednej strony G(2), dwa po różnych stronach G(3) – rys. 2 [5] i od liczby elementów *n* w próbce. Wartości granic są stabelaryzowane. Przedstawiają je wykresy na rys. 3, dla prawdopodobieństw P = 0,95 i 0,99.



Rys. 3. Granice dla danych odstających, odniesione do odchylenia standardowego, w funkcji liczby elementów *n* próbki

Fig. 3. Relations (in SD values) between borders of outliers and numbers *n* of sample elements

Metoda wykrywania składowych systematycznych w surowych danych i ich eliminacja

Po usunięciu danych odstających, czyli *outlierów* wartości $y(i)$ sygnału wyjściowego uzyskiwane w wyniku próbkowania wynoszą

$$y(i) = f(i) + f_R(i) + N(0,\sigma) \tag{1}$$

gdzie: $i \in (1,n)$ – numer kolejny obserwacji w próbce, t_i – czas uzyskania obserwacji $y(i)$, $f(i)$ – wartości w chwilach t_i funkcji menzurandu znanej *a priori* lub jej estyma-

ty wyznaczonej z pomiarów (w szczególnym przypadku $f(i)$ = const), $f_R(i)$ – wartości innych niepożądanych składowych regularnych w chwili t_i, $N(0, \sigma)$ – składowa losowa o rozkładzie normalnym (lub innym znanym).

Występowanie w wynikach obserwacji pomiarowych próbki zakłóceń w postaci zmiennych składowych zdeterminowanych o nieznanym z góry pochodzeniu i przebiegu, a więc niemożliwych do natychmiastowego wyeliminowania, spowoduje, że średnia wyników obserwacji będzie obciążona przesunięciem o nieznanej wartości, a jej odchylenie standardowe będzie większe niż wynikające z przyczyn losowych. Statystycznego ciągu opisującego obserwacji próbki nie można wówczas traktować jako stacjonarnego (tj. o stałych parametrach podczas procesów pozyskiwania i przetwarzania danych próbki). Składowa stała dryftu może być też inna w każdej z próbek pozyskiwanych w różnych chwilach okresu badań obiektu oraz zmieniać się w trakcie pełnego okresu użytkowania przyrządu. Jest to skutek zmian warunków pomiaru i parametrów wewnętrznych powstających w całym torze pomiarowym, łącznie z obiektem badanym. Zmiany te można wykryć tylko po wykonaniu dodatkowych pomiarów kalibrujących. Na ich podstawie można dokonać korekty wartości wyniku pomiarów. Jeśli korekta ta nie jest możliwa, to wpływ trendu należy oszacować jako składową niepewności $u_B(x)$.

Wpływy regularnie zmienne aperiodycznie i okresowo można jednakże wykryć i częściowo usunąć poprzez analizę zbioru surowych wartości obserwacji pomiarowych przy znanym, np. równomiernym sposobie próbkowania. Trend pojawiający się w części elektrycznej toru pomiarowego, np. dryft wzmacniacza, można ponadto wyznaczyć i jego wpływ wyeliminować poprzez wzorcowanie wewnętrzne tej części toru. Jego wpływ usuwa się poprzez zerowanie w procesie ręcznej lub automatycznej adiustacji.

Jeśli obok trendu w próbce występują też zakłócające składowe oscylacyjne, to nie zmieniają one wartości średniej zbioru jej obserwacji q_i tylko wówczas, gdy ich okresy są wielokrotnością okresu ΔT próbkowania równomiernego. Największy wpływ występuje, gdy okres składowej oscylacyjnej równa się podwojonemu czasowi zbierania wyników próbki, a faza początkowa równa się zero. Natomiast takie niewyeliminowane składowe zawsze powiększają niepewność $u_A(x)$.

Zidentyfikowanie występowania składowych systematycznych, tj. trendu i składowych periodycznych występujących równocześnie w serii obserwacji pomiarowych wymaga zastosowania odpowiednich procedur filtracji i przetwarzania ich sygnałów lub obliczeń *off-line* z ich odczytów. W części przypadków można stosować metody uproszczone, np. kolejno wyznaczając i eliminując te składowe. Znajduje tu zastosowanie metoda najmniejszych kwadratów MNK. Przy większej liczbie składowych oscylacyjnych o porównywalnych amplitudach zagadnienie komplikuje się, gdyż dla każdej sinusoidy trzeba by wyznaczyć trzy parametry: częstotliwość, amplitudę i fazę. Natomiast stosowanie przekształcenia szybkiej transformaty Fouriera FFT przy krótkich próbkach natrafia na istotne ograniczenia. Trend można usuwać kilkoma metodami [2] – p. 9.1.3. Należy w tym celu wyznaczyć jego równanie, np. powszechnie znaną metodą najmniejszych kwadratów, w skrócie MNK. Proces czyszczenia próbki z trendu i pojedynczej składowej oscylacyjnej, omówiony w publikacjach [7–12] zostanie przedstawiony w tekście poniżej. Oba przykłady liczbowe stworzono za pomocą symulacji komputerowej.

Przykład 1

Próbka zawiera serię n = 121 surowych wyników obserwacji o 4½ znakach odczytu, które otrzymano w regularnych odstępach próbkowania badanego procesu. Ich wartości v_i podano w tabeli 1. Wpływy czynników oddziałujących na obiekt i tor pomiarowy nie są znane *a priori*. Należy wyznaczyć najlepszą wartość wyniku pomiaru średniego napięcia tej serii obserwacji oraz dokonać oceny dokładności tego wyniku w postaci odchylenia standardowego otrzymanego po wyeliminowaniu z surowych obserwacji wpływu głównych składowych systematycznych, w tym trendu (składowa aperiodyczna odchyleń) oraz przebiegów periodycznych.

Tab. 1. Surowe wyniki obserwacji w przykładzie 1
Tab. 1. Rough results of observations of Example 1

1,2200	1,2080	1,2186	1,2263	1,2497	1,2725	1,2981	1,2731
1,2500	1,2286	1,2181	1,2183	1,2162	1,2247	1,2253	1,2108
1,2409	1,2529	1,2696	1,2577	1,2397	1,2300	1,2341	1,2562
1,2449	1,2378	1,2203	1,1920	1,2056	1,2092	1,2198	1,2227
1,2210	1,2134	1,2064	1,2138	1,2154	1,2220	1,2352	1,2479
1,2385	1,2277	1,2206	1,2320	1,2466	1,2679	1,2412	1,2279
1,1897	1,2123	1,2291	1,2498	1,2450	1,2343	1,2356	1,2420
1,2239	1,2101	1,2057	1,2044	1,2011	1,1940	1,1941	1,1836
1,1956	1,2002	1,2159	1,2142	1,1963	1,1840	1,1726	1,1657
1,1553	1,1726	1,1932	1,2146	1,1983	1,1904	1,1736	1,1874
1,2003	1,1950	1,1911	1,1754	1,1594	1,1748	1,1799	1,1817
1,1816	1,1907	1,1937	1,1982	1,1956	1,1977	1,1868	1,1684
1,1455	1,1648	1,2019	1,2126	1,2086	1,1885	1,1760	1,1729
1,1706	1,1692	1,1921	1,2036	1,2229	1,1996	1,1810	1,1609
1,1314	1,0975	1,0704	1,0845	1,0954	1,1146	1,1172	1,1148
1,1263							

Rozwiązanie

Odchylenia Δv_i surowych wartości obserwacji od ich wartości średniej $\bar{v} = 1{,}2029$ w kolejności ich pozyskania przy równomiernym próbkowaniu przedstawiono na rys. 4a. Można zauważyć, że mają one tendencję zmniejszania się wraz ze wzrostem kolejnego numeru i obserwacji w próbce. Wyznaczono histogram dla 8 jednakowych przedziałów, na które podzielono pełny zakres zmian odchyleń $\Delta v_{i\,max} - \Delta v_{i\,min}$ = 0,2277. Łącząc punkty histogramu otrzymano wypadkowy rozkład surowych danych przedstawiony jako krzywa 1 (rys. 5). Różni się on od rozkładu normalnego (4) na tym rysunku. Jest wyraźnie niesymetryczny i przy założonym poziomie istotności 5% nie spełnia kryterium $\chi 2$ (tab. 2).

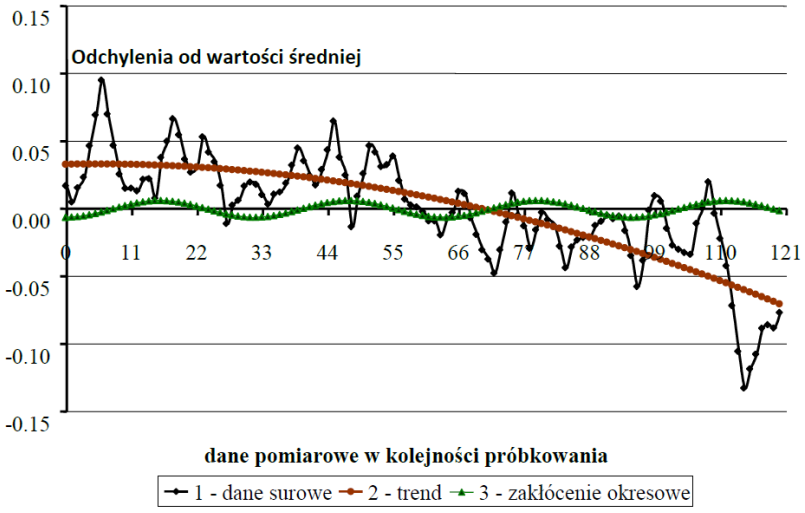
Żadne wpływy nie były tu znane *a priori*, więc nie ma jak wprowadzić poprawek. Natomiast przy równomiernym (lub innym znanym) próbkowaniu można zdobyć dodatkową informację. Stosując metodę najmniejszych kwadratów wyznaczono równanie trendu w odchyleniach obserwacji próby jako wielomian drugiego stopnia:

$$y(\nu) = -7{,}82 \cdot 10^{-6}\nu^2 + 7{,}82 \cdot 10^{-5}\nu + 0{,}033$$

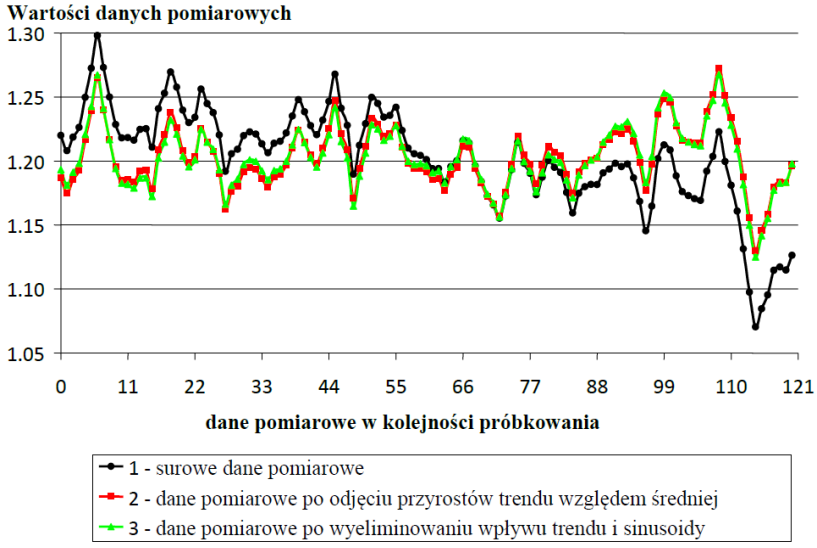
Należy założyć, że przechodzi on przez zero dla wartości średniej próbki, gdyż zmian tej wartości nie można wykryć bez kalibracji zewnętrznej z wykorzystaniem

innego dokładniejszego przyrządu lub wzorca. Po odjęciu zmian trendu od surowych odchyleń, czyli jego przyrostów liczonych względem wartości średniej, otrzymuje się wartości skorygowane $\Delta q_i^{'}$ o mniejszym zakresie zmian, tylko 0,143.

a)

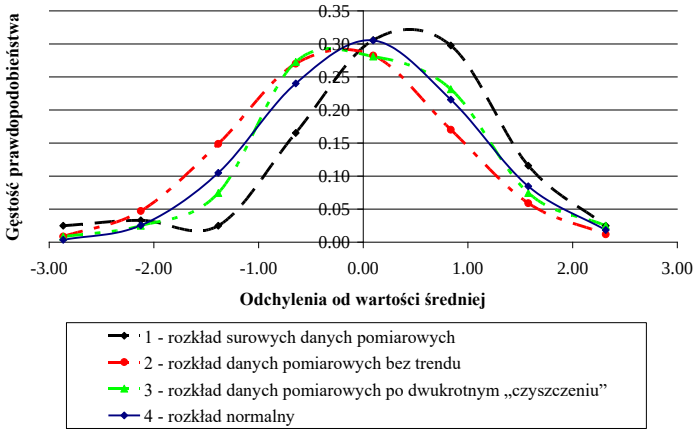


b)



Rys. 4. Przebiegi danych pomiarowych próbki wg kolejności próbkowania: a) obwiednia odchyleń danych surowych od wartości średniej i jej składowe systematyczne, b) dane surowe po usunięciu trendu, a następnie i sinusoidy

Fig. 4. Sets of measurement data in order as they are regularly sampled: a) deviations of raw measurement data from mean value and systematic components discovered in them, b) sets of measurement data after cleaning from trend and from trend and oscillation



Rys. 5. Histogramy odchyleń danych pomiarowych od ich wartości średniej
Fig. 5. Histograms of the dispersion from mean value of the measurement data: 1 – for raw data; 2 – after elimination of the trend; 3 – after elimination of the trend and the oscillation

Podstawowe parametry statystyczne surowych wyników o równomiernym próbkowaniu i po eliminacji z nich wpływu składowej pochodzącej od samego trendu liniowego oraz od trendu i składowej oscylacyjnej, podano łącznie w tabeli 2.

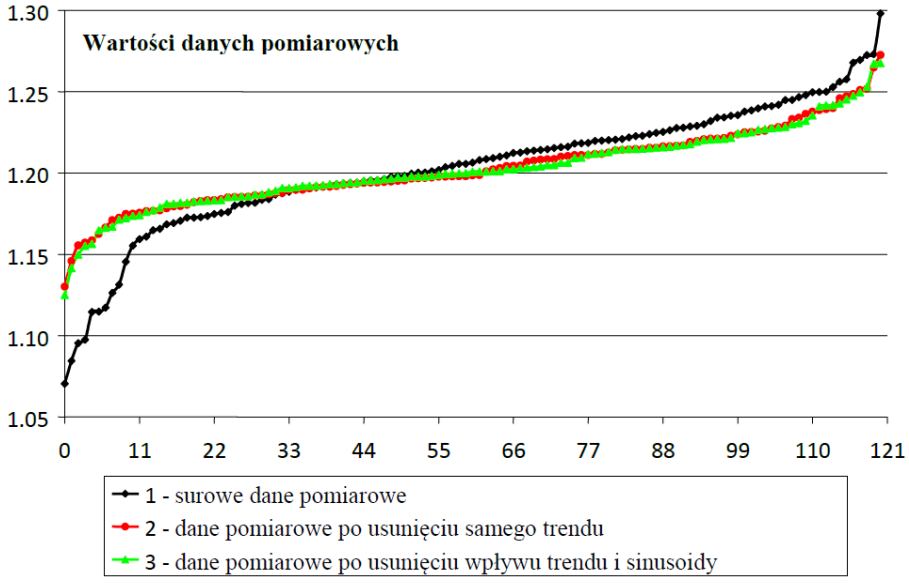
Tab. 2. Parametry statystyczne próbki surowych obserwacji i po kolejnych „czyszczeniach” z trendu i z przebiegu oscylacyjnego
Tab. 2. The statistical parameters of collected measurement data sample of raw observations and after two steps of cleaning: firstly from trend and then from oscillations

Wyniki pomiarów	Dane surowe v_i	Dane oczyszczone z:	
		trendu q_i'	trendu i sinusoidy q_i
Kryterium $\chi 2$ dla rozkładu normalnego	$\chi^2 = 52{,}28 > \chi^2_{5;\,0{,}05} = 11{,}1$	$\chi^2 = 3{,}83 < \chi^2_{5;\,0{,}05} = 11{,}1$	$\chi^2 = 3{,}25 < \chi^2_{5;\,0{,}05} = 11{,}1$
	rezultat negatywny	**rezultat pozytywny**	
Wartość średnia	$\bar{v} = 1{,}2029$	$\bar{q}' = 1{,}2028\,*$	$\bar{q} = 1{,}20262\,*$
Standardowe odchylenie próbki	$s(v_i) = 0{,}03953 \approx 0{,}040$	$s(q_i') = 0{,}02409\,*$	$s(q_i) = 0{,}02407\,*$
Niepewność u_A	$0{,}003593 \approx 0{,}0036$	$0{,}002190 \approx 0{,}0022$	$0{,}002188 \approx 0{,}0022$
Stosunek odchyleń standardowych	$\frac{s(v_i)}{s(q_i')} = \frac{0{,}03953}{0{,}02409} \approx 1{,}6407\,*$		$\frac{s(v_i)}{s(q)} \approx 1{,}6420\,*$
Wynik końcowy pomiaru dla $U=2u_A$	bez uwzględnienia autokorelacji		$x = 1{,}2026 \pm 0{,}0044$
	z uwzględnieniem autokorelacji [15] (rozdz. 3)		$x = 1{,}2026 \pm 0{,}0095$

* Wartości podano bez zaokrąglenia do 2 cyfr, by lepiej uwidocznić małe różnice.

Na rys. 5 podano też funkcję gęstości rozkładu próbki opartą na nowym histogramie 2. Z analizy tabeli 2 wynika, że spełnia już ona założone kryterium $\chi 2$. Wartość średnia jest taka sama jak poprzednio dla surowych wartości obserwacji.

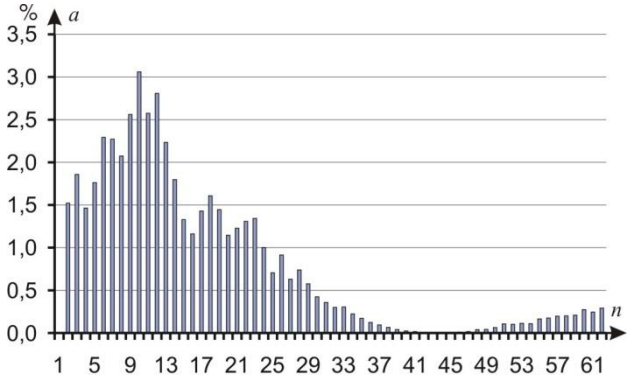
W podobny sposób wykryto też i wyznaczono metodą MNK oraz usunięto zawarty w danych pomiarowych niewielki sinusoidalny przebieg oscylacyjny(rys. 4a). Histogram 3 po dwukrotnym czyszczeniu i odpowiadający jego punktom rozkład normalny 4 też podano na rys. 5. Przedstawione w kolejności próbkowania na rys. 4b łącznie dane surowe i ich wartości otrzymane po każdym z obu etapów „czyszczenia", uszeregowano wg wartości i podano na rys.6.



Rys. 6. Dane surowe (1) i po kolejnych czyszczeniach (2, 3) uporządkowane wg ich wartości
Fig. 6. The raw measurement data (1) and after steps of their cleaning (2, 3) ordered by values

Na podstawie rysunków 5 i 6 wstępnie można ocenić, jaki model rozkładu przyjąć dla oczyszczonych danych eksperymentalnych. Z porównania ich obu wynika, że jest on zbliżony do rozkładu normalnego, gdyż dla tego rozkładu wykres na rysunku 6 ma kształt obróconej o 90° jego dystrybuanty. Jeśli na osi *y* rys. 6 zastosowano by nieliniową podziałkę dostosowaną do tego rozkładu, to jego wykres otrzymano by w postaci linii prostej. Ułatwia to ocenę wizualną jak dalece rozkład rzeczywisty odbiega od normalnego. Dawniej do tego celu stosowano specjalny papier milimetrowy o takiej podziałce, tzw. probabilistyczny, dziś do wizualizacji danych można wykorzystać odpowiedni program komputerowy.

Postanowiono też dodatkowo sprawdzić, czy w odchyleniach próbki po wyeliminowaniu obu składowych występują jeszcze jakieś inne harmoniczne o dużej amplitudzie. W tym celu wyznaczono widmo odchyleń o częstotliwościach liczonych względem okresu próbkowania i amplitudach *a* w % wartości skutecznej RMS wszystkich składowych. Przedstawiono je na rys. 7.



Rys. 7. Widmo częstotliwościowe skorygowanych podwójnie odchyleń próbki z przykładu 1 jako funkcja okresu harmonicznej odniesionej do czasu próbkowania ΔT

Fig. 7. Frequency spectrum of data cleaned from non-periodical and periodical component measurement data deviations of Example 1 as function of period of harmonics related to the sampling time ΔT

Zasadnicza, tj. początkowa część tego widma wykazuje, że proces losowy jest niskoczęstotliwościowy i nie występują już w nim wyraźnie widoczne składowe oscylacyjne. Potwierdza to przyjęte założenie, że dalsze poszukiwanie i eliminacja składowych oscylacyjnych wyższego rzędu w wynikach surowych już jest nieuzasadniona. Wyznaczono nawet metodą MNK kolejną z nich, ale jej amplituda okazała się pomijalnie mała.

Rozkład wyników obserwacji pomiarowych oczyszczonych z obu tych składowych systematycznych jest już wystarczająco bliski rozkładowi normalnemu.

Wniosek

Dla surowych wyników obserwacji otrzymano eksperymentalne odchylenie standardowe próbki o ok. 64% większe niż dla tych danych po wyeliminowaniu trendu. Jest też one jeszcze nieznacznie, tj. o ok. 64,2% większe od standardowego odchylenia otrzymanego po oczyszczeniu wyników surowych z obu składowych systematycznych, tj. z trendu i z przebiegu periodycznego. Proces czyszczenia wpłynął więc istotnie na zmniejszenie niepewności u_A.

Sprawdzenie dokładności procesu czyszczenia surowych obserwacji

W celu sprawdzenia skuteczności metody sukcesywnego czyszczenia zostanie dokonana następująca symulacja. Początkowe wartości obserwacji próbki pobranej z pewnej populacji losowej zostaną zanieczyszczone znaną składową oscylacyjną. Tak zanieczyszczona próbka będzie stanowiła dane surowe. Następnie dokona się procesu czyszczenia takiej próbki. W tym celu wyznaczy się równania podstawowych składowych systematycznych, w tym składową oscylacyjną, i porówna się ją z funkcją pierwotną. Sposób uproszonego postępowania przy wyznaczaniu skorygowanych wartości q_i będzie podobny jak w przykładzie 1, tj. zastosuje się kolejno metodę MNK. Zilustruje to przykład 2.

Przykład 2

Próbka pomiarowa wraz z dodanym przebiegiem periodycznym o nieznanych dla eksperymentatora parametrach, stanowi zbiór surowych wyników 144 obserwacji pomiarowych otrzymanych kolejno przy regularnym próbkowaniu. Podano je w tabeli 3. Należy wykryć występujące w próbce podstawowe regularne zakłócenia, wyznaczyć ich równania, oczyścić z nich próbkę i na podstawie dodatkowej informacji o wprowadzonych zanieczyszczeniach ocenić skuteczność tej operacji. Ponadto należy określić właściwy przybliżony model rozkładu oczyszczonych danych i jego podstawowe parametry.

Tab. 3. Surowe wyniki pomiarów x_i przykładu 2
Tab. 3. Rough results of observations x_i of Example 2

6,822	2,699	5,044	1,816	-0,161	-0,546	1,935	4,881
2,419	-0,939	-0,298	6,267	5,456	2,478	1,651	-1,473
-1,642	-0,459	-0,121	5,459	2,909	5,266	-1,158	-0,123
-0,041	6,291	1,052	-0,016	0,094	8,007	6,843	6,622
7,135	0,491	2,224	6,554	4,375	7,798	4,114	7,142
6,845	4,101	7,056	7,158	6,413	8,447	4,648	7,645
8,899	11,732	7,704	10,874	7,318	4,332	7,215	7,382
12,499	9,169	12,006	11,745	11,177	7,582	13,859	8,480
7,433	9,193	10,012	7,291	8,923	5,934	13,750	8,464
9,873	9,430	9,783	8,307	5,442	7,183	10,296	10,020
11,525	10,785	12,501	5,314	12,002	5,604	4,885	4,755
6,672	9,822	4,861	4,237	3,979	8,964	6,566	8,831
1,829	8,438	3,358	1,417	8,454	4,869	7,054	3,330
8,075	6,166	3,312	1,908	1,258	5,028	7,681	4,551
7,377	1,164	5,820	3,133	3,222	2,954	3,066	4,290
5,562	8,535	1,452	4,571	3,260	9,625	8,141	5,622
7,375	9,423	5,955	8,816	11,407	11,274	7,272	6,626
8,935	4,091	9,406	5,393	7,987	5,159	7,365	10,211

Rozwiązanie

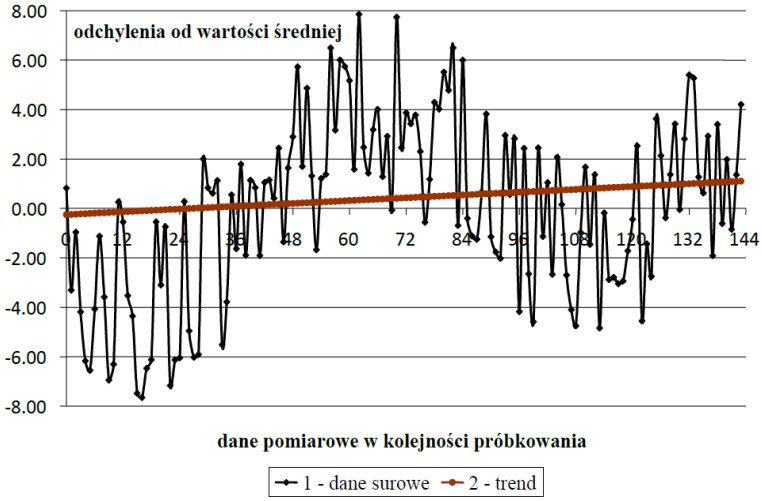
Odchylenia wyników surowych od wartości średniej są dosyć duże. Wraz z występującym w nich trendem przedstawiono je na rys. 8, a ich histogram znajduje się na rys. 10 – (czarne słupki). Wyznaczono równanie tego trendu – prosta 2, przyjmując, że jest on w przybliżeniu liniowy i podobnie jak w przykładzie 1, z danych surowych usunięto liniowo narastające przyrosty względem średniej próbki wynikające z tego równania. Otrzymano punkty podane na rys. 9 – obwiednia 1. Z oceny wizualnej wynika, że odchylenia mogą zawierać pojedynczą sinusoidę – krzywa 2. Metodą MNK wyznaczono jej równanie o przebiegu podanym na tym rysunku. Optymalne wartości jej amplitudy U_m, częstotliwości $f{\cdot}T$ (unormowanej w stosunku do czasu zbierania T próbki) i jej fazy φ wynoszą:

$$U_m = -3{,}063; \qquad f{\cdot}T = 1{,}465; \qquad \varphi = 0{,}678$$

W symulacji sprawdzającej skuteczność wykrywania składowych systematycznych w wartościach obserwacji próbki dodano do nich sinusoidę o znanych parametrach

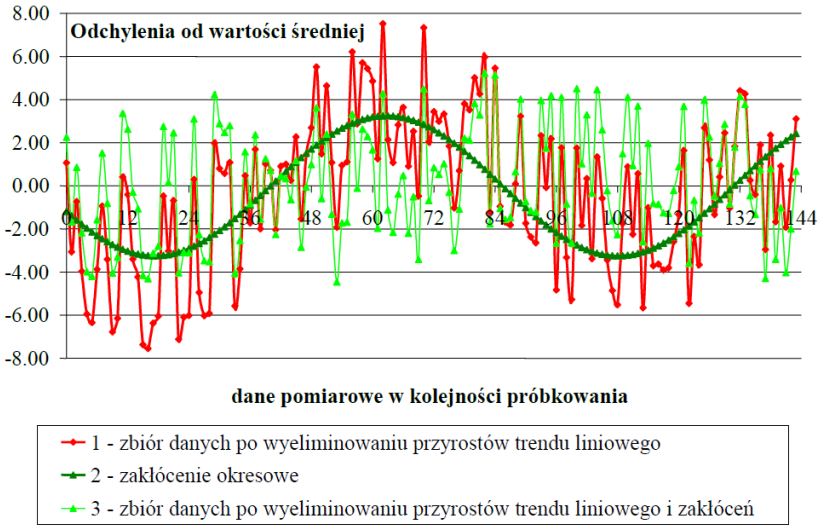
$$U_m = -3{,}123; \qquad f{\cdot}T = 1{,}481; \qquad \varphi = 0{,}561$$

Parametry tej sinusoidy i powyżej zidentyfikowanej metodą MNK różnią się niewiele. Różnice wynikają zapewne stąd, iż w wartościach początkowych próbki przy skończonej liczbie obserwacji przypadkowo pojawiła się mała składowa okresowa.



Rys. 8. Odchylenia Δx_i surowych wyników pomiarów od ich średniej zestawione w kolejności próbkowania

Fig. 8. Deviations Δx_i of raw measurement results from their mean presented in the order of sampling

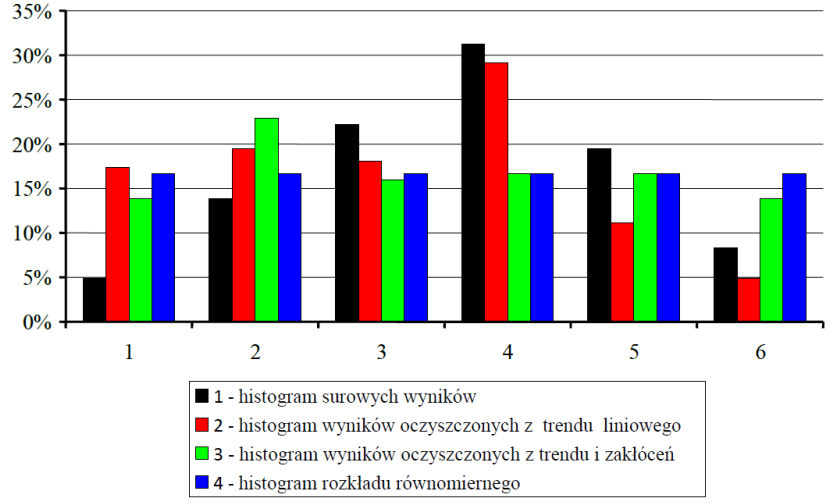


Rys. 9. Zbiór odchyleń Δq_i' dla przykładu 2 po wyeliminowaniu z surowych przyrostów Δx_i trendu liniowego oraz odchylenia Δq_i po oczyszczeniu też i z sinusoidy

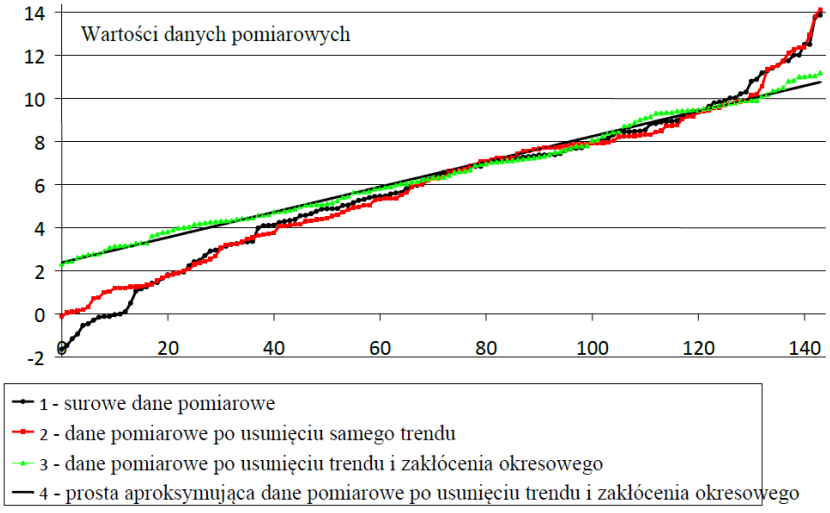
Fig. 9. Deviations of raw measurement data of example 2 after steps of their cleaning: 1 – Δq_i' from the linear trend, 3 – Δq_i from trend and then from sinusoid 2

Na rys. 9 podano również wartości obserwacji próbki otrzymane po pełnej korekcji – obwiednia 3. Rozrzut wyników jest jeszcze mniejszy i nie widać by występowała w nich jeszcze jakaś regularność.

Na rys. 10 zawarto trzy histogramy: dla wyników surowych, po usunięciu z nich trendu liniowego, a następnie przebiegu sinusoidalnego. Natomiast na rys. 11 dla kolejnych etapów procesu czyszczenia podano dane uporządkowane wg wartości.



Rys. 10. Histogramy odchyleń obserwacji od ich wartości średniej dla przykładu 2
Fig. 10. The histograms of deviations of observations from the mean value for Example 2



Rys. 11. Rozkłady danych z przykładu 2 uporządkowane wg ich wartości
Fig. 11. Distributions of data from Example 2 ordered according to their values

Z obu ostatnich rysunków wynika, iż rozkład końcowy po oczyszczeniu danych jest zbliżony do równomiernego. W jego histogramie na rys. 10 jedynie wartość drugiego przedziału jest zauważalnie wyższa od pozostałych. Również na rys. 11, wartości obserwacji próbki po dwukrotnym czyszczeniu – krzywa 3 układają się najbliżej linii prostej jako wykresu rozkładu równomiernego.

Można zauważyć pewien paradoks, który wystąpił się w tym przykładzie. Otóż wzrokowo rozkład wyników surowych był bliższy normalnego niż równomiernego, chociaż nie spełniałby zwykle zakładanego poziomu zgodności wg kryterium $\chi 2$. Jako hipotezę modelu rozkładu dla populacji z której pochodzi próbka powinno się więc ostatecznie przyjąć rozkład równomierny. W sposób podobny do opisanego w przykładzie 1 sprawdzono tę hipotezę wg kryterium $\chi 2$ przy założonym poziomie istotności $\alpha = 0{,}05$ i dla histogramu o 8 przedziałach uzyskano wynik pozytywny. Wartość średniej 6,007 nieco wzrosła po usunięciu składowej sinusoidalnej do 6,601. Powodem jest niecałkowita liczbę jej okresów w długości próbki. Natomiast wartości standardowego odchylenia średniej, czyli niepewności u_A szacowanej wg GUM, są dla kolejnych etapów następujące: 0,2950, 0,2798, 0,2065. Wartość u_A malała więc po każdej operacji czyszczenia i łącznie zmniejszyła się 1,43 razy.

Dla próbki surowej oraz po każdym z etapów jej czyszczenia wyznaczono też współrzędną środka rozpięcia jako estymatora wyniku pomiarów statystycznie lepszego niż średnia dla próbki z populacji o rozkładzie równomiernym. Wyniosła ona kolejno 6,108, 7,009, 6,186. Odpowiednie odchylenia standardowe tych środków są następujące: 0,382, 0,0351, 0,0218. Po przeprowadzeniu czyszczenia próbki ocena dokładności środka rozpięcia jest lepsza i to o przeszło 40% (tj. odchylenie standardowe bez korekcji jest aż ok. 1,6 razy większe). Dzięki czyszczeniu ocena dokładności obu estymatorów wyniku pomiaru uległa istotnej poprawie.

Podsumowanie

Celem obu przedstawionych przykładów liczbowych było pokazanie jak usunięcie trendu i pojedynczego przebiegu oscylacyjnego z surowych wyników obserwacji wpływa na ich wartość średnią i niepewność u_A dla próbek przybliżanych rozkładem normalnym lub równomiernym. W przykładzie 2 sprawdzono też, czy dokładność wyznaczenia składowej oscylacyjnej próbki jest wystarczająca. W obu przypadkach uzyskano istotnie mniejsze odchylenie standardowe wyniku pomiaru.

W praktyce pomiarowej jest bardzo istotne, aby upewnić się, czy wykryte składowe regularne istnieją rzeczywiście i jaka jest przyczyna ich się pojawienia, a nie, że są li tylko skutkiem przypadkowego rozłożenia się punktów w pojedynczej próbce o skończonej długości. W tym celu należy zbadać wg przyjętego do porównań kryterium, np. $\chi 2$, czy przebiegi te powtarzają się w kilku kolejnych próbkach o zbliżonych warunkach pomiaru, lub zarejestrować jedną długą próbkę i analizować odcinkami jej stacjonarność.

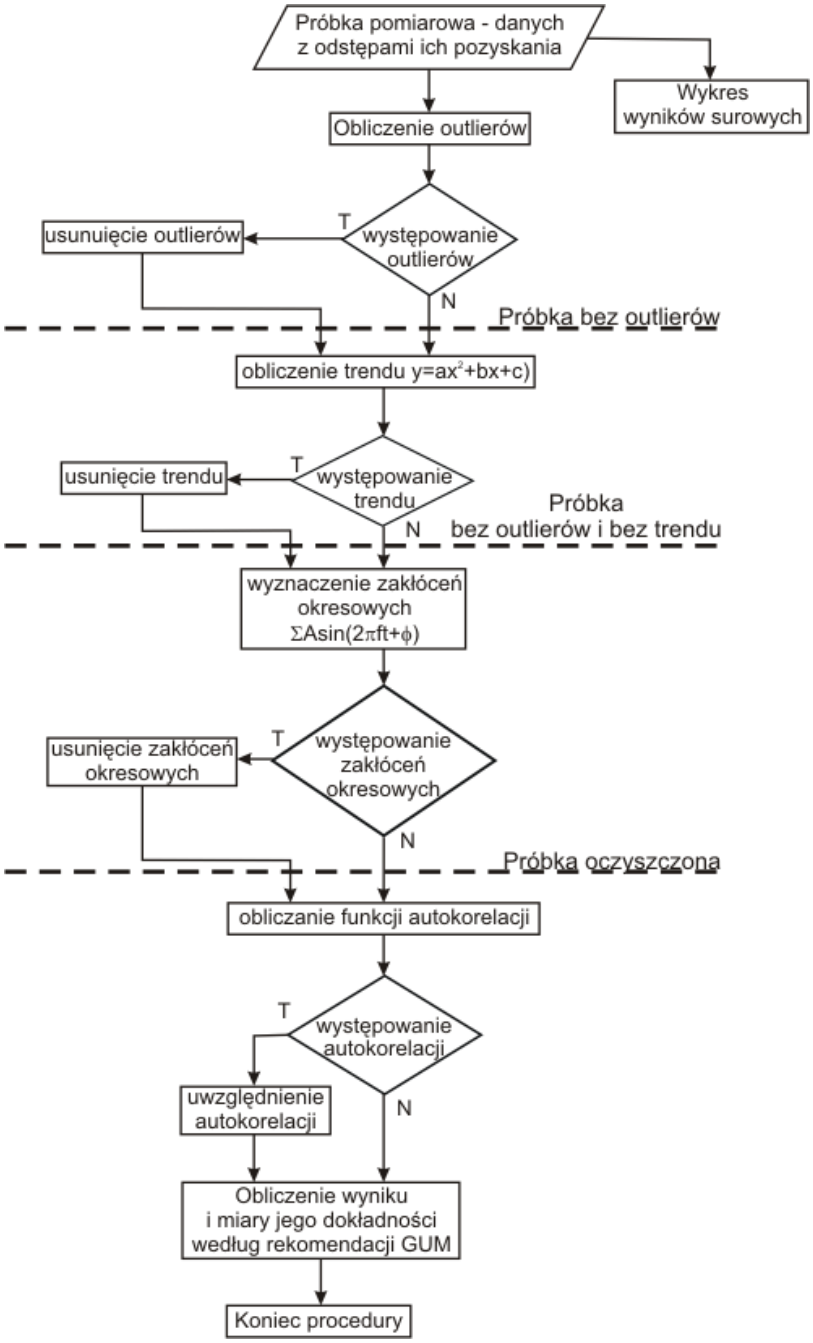
Z przedstawionych rozważań wynika, że przy stosowaniu w pomiarach regularnego próbkowania zalecaną przez Przewodnik GUM procedurę wyznaczenia niepewności typu A jako miary losowego rozrzutu wyników można by w dość prosty sposób rozszerzyć na opis pomiarów o dużych i nieznanych zakłóceniach. W tym celu, po wprowadzeniu poprawek – etap 1 z rys 1, należy następnie wykryć i oczyścić surowe wyniki obserwacji z krótkich zakłóceń impulsowych outlierów, a na-

stępnie – z nieznanych z góry zakłóceń systematycznych w postaci trendu i oscylacji. Jest to możliwe wtedy, gdy dla menzurandu (mierzonej wielkości) w postaci wartości stałej lub wartości średniej znane są względne położenia poszczególnych obserwacji, np. z ich kolejności przy równomiernym próbkowaniu lub prowadzonym wg innej procedury określającej względną pozycję każdej obserwacji. Wówczas wg omówionych powyżej zasad przetwarzania wartości takich obserwacji z wykorzystaniem metody najmniejszych kwadratów MNK można wykryć i wyeliminować wpływ zakłóceń w postaci trendu (dryftu) oraz oscylacji występujących w odchyleniach tych wyników od ich wartości średniej. Uzyskuje się wynik pomiarów o mniejszej niepewności typu A. Schemat takiej ulepszonej procedury do szacowania niepewności u_A podano na rysunku 12.

Można by włączyć procedurę czyszczenia danych surowych do rekomendacji międzynarodowych jako metodę standardową poprzedzającą wyznaczanie niepewności u_A wg zaleceń GUM oraz opracować jej standaryzację dla różnych przypadków występujących w praktyce pomiarowej.

Potrzebę stosowania przedstawionej tu metody ***czyszczenia surowej próbki*** wynikłej z propozycji udoskonalenia metody wyznaczania niepewności typu A przedstawiono wstępnie w [7–10], oraz rozwinięto w wielu następnych publikacjach. Propozycje te obejmują m. in. uwzględnianie wpływu autokorelacji obserwacji na szacowanie niepewności typu A oraz dla modeli rozkładów wyników obserwacji pomiarowych odbiegających istotnie od normalnego stosowanie innych, lepszych estymatorów statystycznych wyniku pomiaru, tj. o mniejszym odchyleniu standardowym niż dla wartości średniej. Czyszczenie danych próbki [9, 11–13] stanowi uzupełnienie umożliwiające poprawę tych ocen i powinno je poprzedzać.

Wybrane propozycje dotyczące wyznaczania niepewności typu A oraz uściślenia w szacowania niepewności typu B i niepewności złożonej u_C i rozszerzonej u zostaną omówione szczegółowo w następnych rozdziałach. Umożliwiają one istotnie zwiększenie wiarygodności oceny niepewności pomiarów i to nie tylko najdokładniejszych laboratoryjnych w państwowej służbie miar i nauce, ale przede wszystkim powinno się je wykorzystywać w wielu badaniach użytkowych w przemyśle i eksploatacji.



Rys. 12. Schemat algorytmu ulepszonej procedury szacowania niepewności u_A
Fig. 12. Scheme of the algorithm of improved procedure for estimating the uncertainty u_A

Literatura

1. *Wyrażanie Niepewności Pomiaru. Przewodnik*, tłumaczenie z komentarzem J. Jaworskiego: Guide to the Expression of Uncertainty in Measurement, ISO 1992, revised and corrected 1995, Główny Urząd Miar Warszawa 1999.
2. Bendat J.S., Piersol A.G.: *Metody analizy i pomiaru sygnałów losowych*, WNT Warszawa 1976.
3. Piotrowski J., Kostyrko K.: *Wzorcowanie Aparatury Pomiarowej*, Wydawnictwo Naukowe PWN Warszawa 2000, 2012.
4. Skubis T.: *Podstawy metrologicznej oceny wyników pomiaru*. Wyd. Politechniki Śląskiej Gliwice 2004.
5. Zięba A.: *Analiza danych w naukach ścisłych i technice*. Wydawnictwo Naukowe PWN Warszawa 2013.
6. Pavese F., Ichim D.: *SAODR: Sequence analysis for outlier data rejection*, "Measurement Science and Technology", 15 (2004), 2047–2052.
7. Dorozhovets M., Warsza Z. L.: *Udoskonalenie metod wyznaczania niepewności wyników pomiaru w praktyce*. „Przegląd Elektrotechniczny", 1/2007, 1–13.
8. Dorozhovets M., Warsza Z. L.: *Propozycje rozszerzenia metod wyznaczania niepewności wyniku pomiarów wg Przewodnika GUM*. „Pomiary Automatyka Robotyka", część (1) 1/2007, 6–15, część (2) 2/2007, 45–52.
9. Warsza Z. L., Dorozhovets M., Korczynski M.J., *Methods of upgrading the uncertainty of type A evaluation*, Part 1. Elimination the influence of unknown drift and harmonic components, Proceedings of 15th IMEKO TC4 Symposium, Sept. 2007, Iasi Romania, 193–198.
10. Dorozhovets M., Warsza Z., *Methods of upgrading the uncertainty of type A evaluation*, Part 2. Elimination of the influence of autocorrelation of observations and choosing the adequate distribution, Proceedings of 15th IMEKO TC4 Symposium, Iasi Romania, 2007, 199–204.
11. Dorozhovets M., Warsza Z., Korczyński J., *Udoskonalenie metody typu A wyznaczania niepewności pomiarów. Wykrywanie i eliminacja wpływu dryftu i oscylacji przy równomiernym próbkowaniu*. „Pomiary Automatyka Kontrola", Nr 12, 2007, 8–11.
12. Warsza Z.L., Korczyński M.J.: *Eliminacja wpływu nieznanych a priori składowych systematycznych na niepewność typu A pomiarów o równomiernym próbkowaniu*. „Pomiary Automatyka Robotyka", Nr 2, 2008, 5–13.
13. Warsza Z.L., Korczyński M.J.: *Eliminacja nieznanych a priori składowych systematycznych z niepewności typu A pomiarów o równomiernym próbkowaniu*. „Przegląd Elektrotechniczny", Nr 05/2008, 109–114.
14. Warsza Z.L., Korczynski M.J.: *Improving of the type A uncertainty evaluation by refining the measurement data from a priori unknown systematic influences*. Advances in Intelligent Systems and Computing 267, 2014, Springer, 727–732.
15. Warsza Z.L.: *Szacowanie niepewności w pomiarach z autokorelacją obserwacji*. Elektronika nr 8/2012 (wkładka Technika Sensorowa), 108–113
16. Dorozhovets M., *Testowanie obserwacji istotnie odstających w próbkach o rozkładzie jednostajnym*. „Mechanik" Nr 11/2016, 1596–1599

Rozdział 3

Ocena niepewności pomiarów z autokorelacją danych

Streszczenie. Opisano rozszerzenie podanej w Przewodniku GUM metody wyznaczania niepewności typu A na pomiary o równomiernym próbkowaniu menzurandu w funkcji czasu z uwzględnieniem wpływu autokorelacji obserwacji. Obejmuje ona wyodrębnienie składowej losowej w surowych danych pomiarowych przez identyfikację i usunięcie z nich estymatora sygnału menzurandu oraz innych składowych regularnie zmiennych i outlierów. Podano wzory do obliczania odchylenia standardowego pojedynczej obserwacji i odchylenia wartości średniej próbki przy istnieniu autokorelacji między obserwacjami. Jej wpływ uwzględnia się przez zastosowanie bądź współczynników zwiększających oba odchylenia standardowe, bądź użycie mniejszej od rzeczywistej tzw. „efektywnej liczby" równoważnych nieskorelowanych obserwacji. Dzięki temu można oszacować niepewność rozszerzoną wg dotychczasowej procedury GUM. Podano też metodę wyznaczania estymaty funkcji autokorelacji z danych próbki. Rozważania zilustrowano przykładami.

Opracowanie międzynarodowych przepisów metrologicznych, zalecanych do stosowania w skali globalnej służyć ma temu, aby pomiary w różnych krajach i miejscach były wykonywane rzetelnie, a ich wyniki wiarygodne i porównywalne. Tworzenie tych przepisów jest procesem długotrwałym i wymaga opracowania ich podstaw teoretycznych, zbioru zaleceń ujętych w sposób prosty i jednoznaczny oraz poprzedzającej weryfikacji w praktyce w różnych dziedzinach. Zwykle przepisy te nie nadążają za bieżącymi potrzebami szybko rozwijającej się techniki pomiarowej oraz w szeregu przypadkach trzeba zastosować podejście jeszcze w nich nieujęte. Dotyczy to w szczególności pomiarów przeprowadzanych przyrządami i systemami o zautomatyzowanym działaniu z próbkowaniem sygnału wejściowego i z jego przetwarzaniem a/c. Nie ma jeszcze ustalonych międzynarodowo zasad jak wyznaczać niepewność pomiaru wielkości, oraz przebiegu sygnałów i procesów zmiennych w funkcji czasu, chociaż rozpoczęto już opracowywanie podstaw naukowych do ich tworzenia [19] i istnieje szereg wcześniejszych opracowań obejmujących częściowo tę problematykę [M5, M6].

Poniżej zostanie omówiona możliwość stosunkowo łatwego rozszerzenia zaleceń obecnej wersji międzynarodowego Przewodnika GUM [1] na pomiary wielkości zmiennych w czasie, o obserwacjach uzyskiwanych w trakcie regularnego próbkowania sygnału zależnego od wartości menzurandu. Przedstawiona zostanie metoda przetwarzania danych próbki umożliwiająca poprawne oszacowanie niepewności, gdy obserwacje są powiązane statystycznie (autoskorelowane). Podany jest też sposób wyznaczania estymaty funkcji autokorelacji z tych danych.

Wstępne przetwarzanie surowych danych

Zmiany surowych wartości powtarzanych obserwacji pomiarowych wywołane są przez przyczyny zarówno zdeterminowane, jak i losowe. W ogólnym przypadku sygnał wyjściowy systemu pomiarowego może zawierać zarówno składowe regularnie zmienne w czasie trwania procesu pozyskiwania danych pomiarowych wraz ze składową stałą, jak i składową losową. W pomiarach użytkowych, w tym przemysłowych obu rodzajom zmian mogą podlegać zarówno parametry obiektu badanego, jak i parametry wewnętrzne toru pomiarowego oraz warunki pomiaru. Wartości parametrów statystycznych serii obserwacji jako próbki, otrzymanej np. przez równomierne próbkowanie menzurandu, zwykle będą się różnić dla każdej z próbek pobranych z tej samej populacji, przy różnych częstotliwościach próbkowania, przy innym odcinku czasu zbierania danych, przy mierzeniu zestawem różnych przyrządów, a nawet tymi samymi przyrządami, gdy pomiary zostaną wykonane w innym momencie okresu użytkowania tych przyrządów [4]. W ogólnym przypadku losowe zmiany wartości obserwacji są niestacjonarne, ale można traktować je jako stacjonarne, gdy pomiary zgromadzi się w odpowiednio krótkim odcinku czasu.

Przy wyznaczaniu niepewności menzurandu o wartości stałej lub zmiennej w znany zdeterminowany sposób, należy przyjąć następujące założenia:

- czas T_0 zbierania n obserwacji pomiarowych tworzących próbkę, czyli szerokość okna pomiarowego została wybrana właściwie, tj. tak, że można założyć stacjonarność składowej losowej obserwacji w próbce,
- próbkowanie jest równomierne o częstości $i/t_i = n/T_0$ oraz
- składowa losowa jest addytywna, tj. niezależna od wartości menzurandu, a jej prawdopodobieństwo opisuje się rozkładem normalnym.

Jeśli pojawiają się krótkotrwałe zakłócenia o odstających danych, to przed wyznaczaniem wartości i niepewności wyniku pomiarów należy je wykryć i wyeliminować z surowych danych [5]. Wówczas wartości $y(i)$ sygnału wyjściowego uzyskiwane w wyniku próbkowania wynoszą

$$y(i) = f(i) + f_R(i) + N(0,\sigma) \tag{1}$$

gdzie: $i \in (1,n)$ – kolejny numer obserwacji w próbce, t_i – czas do uzyskania obserwacji $y(i)$, $f(i)$ – wartości w chwilach t_i funkcji menzurandu znanej *a priori* lub jej estymaty wyznaczonej z pomiarów, $f_R(i)$ – wartości innych niepożądanych składowych regularnych w chwili t_i, $N(0, \sigma)$ – składowa losowa o rozkładzie normalnym.

Zadanie pomiarowe może polegać na badaniu różnych regularnych składowych próbki, w tym najczęściej wartości stałej procesu. Jeśli z surowych danych wyznaczy się i odejmie się wartości znanej funkcji $f(i)$ estymującej określony menzurand, to pozostanie jeszcze sygnał resztkowy $v(i) = f_R(i) + N(0, \sigma)$. Zawiera on niepożądane i nieznane *a priori* składowe o charakterze regularnym $f_R(i)$ oraz składową losową $N(0, \sigma)$. Z wartości tej składowej szacuje się wyznaczaną metodą statystyczną składową u_A niepewności pomiaru. Niezależnie od tego, czy niepożądane składowe regularne sygnału są znane *a priori*, czy też zidentyfikowano je przez analizę danych próbki pomiarowej, powinno się je usunąć z danych pomiarowych, gdyż mogą wpływać na parametry estymowanej z próbki wartości stałej lub funkcji mierzonej,

w tym na jej wartość średnią oraz na kształt histogramu próbki, a więc na ocenę wartości i niepewności wyniku pomiaru.

Przeanalizujemy wpływ różnych rodzajów składowych regularnych i sposoby ich eliminacji. Skutków niepożądanych oddziaływań w postaci wartości stałej nie można wyznaczyć z samych surowych danych pomiarowych i ich usunąć, gdyż do wyznaczenia poprawki dla niej trzeba znać wartość rzeczywistą lub poprawną. Jedynym sposobem jest tu dodatkowa kalibracja z wykorzystaniem dokładniejszego przyrządu kontrolnego lub wzorca.

Składowa regularna aperiodyczna, czyli trend, nawet tylko liniowy, jeśli nie zostanie rozpoznany i nie skoryguje się jego wpływu, to spowoduje, że wartość średnia zbioru obserwacji będzie obciążona przesunięciem (błędem systematycznym) o nieznanej wartości. W najnowszej aparaturze elektronicznej poprzez odpowiednio częstą automatyczną adiustację wewnętrzną usuwa się na bieżąco dryft, ale tylko powstający w części instrumentalnej toru pomiarowego. Polega ona na powtarzalnym okresowo automatycznym zerowaniu sygnału wyjściowego przy zwieraniu wejścia. Sposób ten nie zapewnia eliminacji trendu występującego w sygnale przed wejściem układu pomiarowego, tj. w samym menzurandzie i na drodze sygnału do wejścia układu. Wówczas identyfikacji dokonuje się *a posteriori* w procesie przetwarzania danych wyjściowych z użyciem odpowiednich metod obliczeniowych.

Wartość średnia danych pomiarowych pozyskanych w równomiernym próbkowaniu sygnału będzie niezależna od występujących w nich składowych okresowych tylko wtedy, gdy próbka obejmuje pełną liczbę ich okresów oraz, gdy okresy te są wielokrotnością odstępu próbkowania ΔT. Warunki te na ogół nie są spełnione. Niepożądane składowe oscylacyjne mogą nieco wpływać na wartość średnią wyniku oraz istotnie powiększają ocenę składowej statystycznej jego niepewności.

W wielu pomiarach wielkości procesowych zmiennych w czasie (lub w przestrzeni) konieczne jest stosowanie odpowiednich procedur filtracji w czasie rzeczywistym (on-line) w trakcie analogowego i cyfrowego przetwarzaniu sygnału lub podczas obliczeń off-line zebranych odczytów. Należą do nich: metoda regresji, filtracja cyfrowa, szybkie przekształcenie Fouriera FFT, falki itp.

Proces identyfikacji i usuwania niepożądanych składowych zdeterminowanych z surowych danych pomiarowych nazywane są czyszczeniem danych [5, 6]. Proste metody czyszczenia zostały omówione w rozdziale 2 i zilustrowane przykładami.

Metoda oceny niepewności próbki o skorelowanych obserwacjach

Wprowadzeniem w zagadnienie powinno być wcześniejsze zapoznanie się z podstawowymi zasadami wyznaczania wartości wyniku i jego niepewności, podanymi w Przewodniku GUM [1, 2]. Opisano je w skrócie w rozdziale 1. Zalecenia GUM dotyczą menzurandu jako wielkości stałej $x = const$, mierzonej bezpośrednio lub pośrednio. Wynik pomiaru należy obliczać z odpowiednio skorygowanych surowych danych pomiarowych. Czyszczenie tych danych (rozdział 2) obejmuje usunięcie za pomocą poprawek znanych *a priori* błędów systematycznych oraz wykluczenie obserwacji odstających, czyli outlierów, zidentyfikowanych np. na podstawie testów Grubbsa. W przypadku próbkowania wielkości ciągłych w znany, np. regularny sposób, istnieje dodatkowa możliwość zidentyfikowania w danych surowych nieizna-

nych *a priori* regularnych zakłóceń aperiodycznych i periodycznych (rozdział 2). Należy je również usunąć z tych danych w procesie czyszczenia.

Wynik pomiaru $x = const$ opisują dwa składniki:

- średnia, $\bar{x} = \bar{q}$ jako najbardziej prawdopodobna wartość dla zbioru skorygowanych obserwacji q_i próbki pomiarowej,
- niepewność rozszerzona $u(x)$ wyniku pomiaru jako miara niedokładności o określonym prawdopodobieństwie.

Wyznaczenie niepewności rozszerzonej $u(x) = k_P\, u_C(x)$ wymaga oszacowania niepewności składowych $u_A(x)$ i $u_B(x)$, odpowiednio metodami typu A i B, a z nich niepewności standardowej $u_C(x) = \sqrt{u_A^2(x) + u_B^2(x)}$. Niepewność $u(x)$ można też wyznaczyć bezpośrednio dla splotu rozkładów możliwych wartości $u_A(x)$ i $u_B(x)$.

Niepewność $u_A(x)$ stanowi estymatę odchylenia standardowego średniej $\bar{x}$ dla całej populacji możliwych wartości q_i. Wyznaczanie $u_A(x)$ wg zasad GUM ogranicza się do przypadków, gdy obserwacje te można traktować jako przypadkowo zmienne, niezależne od siebie, czasu lub przestrzeni, czyli jako realizacje zmiennej losowej, a nie procesu losowego. Zakłada się też, że rozrzut danych z zadowalającym przybliżeniem można opisać rozkładem normalnym, a dla małej liczby pomiarów – rozkładem Studenta. Oba przebiegają w nieograniczonym przedziale wartości x. Rozkłady te są stabelaryzowane i zachowują kształt przy splocie kilku z nich. Rozkład normalny od czasów Gaussa i Laplace'a służył do opisu błędów przypadkowych. W rzeczywistości liczba źródeł i zakresy zmian parametrów oddziałujących przypadkowo na niepewność pomiarów są ograniczone i rozrzut danych pomiarowych występuje w przedziale skończonym. Inne rozkłady mogą więc lepiej modelować rzeczywiste dane eksperymentalne i obecnie, gdy dysponuje się doskonałymi skomputeryzowanymi możliwościami obliczeniowymi, nie ma poprzednich ograniczeń w ich stosowaniu. Ponadto przyjęta wg Przewodnika GUM średnia $\bar{x}$ jako estymator wartości menzurandu jest najkorzystniejszym, tj. o najmniejszym odchyleniu standardowym parametrem tylko dla próbek zawierających dane z populacji o rozkładzie normalnym. Z innych badań (omawianych w rozdziałach 5–7) wynika, że np. dla próbek o rozkładzie równomiernym i dla części rozkładów trapezowych bardziej efektywny jest środek rozpięcia będący średnią z krańcowych wartości obserwacji po usunięciu outlierów.

Niepewność $u_B(x)$ służy do oszacowania wpływów pozostających po korekcji danych surowych o rozpoznanym pochodzeniu, ale o nieznanej stałej wartości w danym eksperymencie, więc nieusuniętych przez poprawki. Zastępuje ona stosowany poprzednio w wyznaczaniu dokładności pomiarów maksymalnie możliwy (ang. *the worth case*) tzw. graniczny błąd systematyczny. Błąd ten dotychczas stosowany jest w zdefiniowanych wymaganiach dla przyrządów podlegających prawnej kontroli. W Przewodniku GUM przyjęto, że składowe $u_B(x_{Qi})$ niepewności $u_B(x)$, pochodzące od wielkości wpływających Q_i są niezależne od przyczyn ich powstawania i charakteru (instrumentalne, związane z metodą pomiaru, od oddziaływań otoczenia i z innych źródeł). Traktuje się je jednolicie jako losowe, szacuje ich wariancje i jako wielkości drugiego rzędu w stosunku do wartości mierzonej sumuje się liniowo również dla funkcji nieliniowych opisujących pomiar. Dokładność oszacowania

$u_B(x)$ zależy od wiedzy eksperymentatora o oddziaływaniach i parametrach przyrządów zastosowanych w danym eksperymencie.

W rozdziale 1, w drugiej kolumnie tabeli 2 podano wzory wg GUM do wyznaczania niepewności pomiaru. Wynika z nich, że uzyskanie większej dokładności pomiarów, np. przez zadane proporcjonalne zmniejszenie ich niepewności typu A, wymaga kwadratowego wzrostu liczby obserwacji pomiarowych n podlegających losowemu rozrzutowi. Współczesna technika umożliwia wykonywanie pomiarów cyfrowych z niemal dowolnie nastawianą częstością w dziedzinie czasu, lub zmian innej wielkości. Aby zwiększyć liczbę obserwacji n w ograniczonym czasie badań lub skrócić ten czas często ustawia się możliwie gęste próbkowanie. Ograniczenie czasu pozyskiwania obserwacji może również wynikać z wymagań dotyczących śledzenia szybkich zmian badanego procesu. Przy zbyt gęstym próbkowaniu rezultaty obserwacji mogą stać się powiązane ze sobą statystycznie, czyli występuje autokorelacja, szczególnie, gdy zjawiska losowe w sygnale są intensywne przy małych częstotliwościach, np. szum 1/*f*. W takich przypadkach należy znać funkcję autokorelacji lub wyznaczyć jej estymatę z danych pomiarowych i uwzględnić jej wpływ w oszacowaniu niepewności typu A. Przy dużej autokorelacji danych stosując rutynowo wzór dla $u_A(x)$ z kolumny 2 tabeli 2 (rozdział 1) otrzyma się zbyt korzystną, tj. znacznie zaniżoną wartość, gdyż zalecenia GUM takiego przypadku dotychczas nie obejmują. W Przewodniku GUM [1–3] oraz innych krajowych i europejskich instrukcjach metrologicznych, podręcznikach, a nawet w publikacjach o wyznaczaniu niedokładności pomiarów, np. [M3] zakłada się *a priori*, że rozrzut wartości obserwacji próbki opisuje się rozkładem normalnym i że są one statystycznie niezależne.

Od lat, na wielu seminariach metrologicznych autor sygnalizował, że z teorii sygnałów wynika, iż przy próbkowaniu obserwacje pomiarowe procesu można traktować dopiero jako niezależne, gdy odległość między nimi jest szersza niż zastępczy promień funkcji autokorelacji oraz, że informację o statystycznym ich powiązaniu można pozyskać z analizy wartości kilku kolejnych obserwacji. Pozostawało to bez odzewu w polskim środowisku metrologów. Przełom rozpoczął się dopiero, gdy w 2006 r. M. Dorozhovets na seminarium w Politechnice Rzeszowskiej przedstawił prosty sposób ujęcia wpływu autokorelacji przez zastępczą liczbę pomiarów nieskorelowanych n_{eff} [9]. W [10] i kilku następnych wspólnych publikacjach [11–14], omówiono zagadnienie – jak uwzględniać autokorelację przy wyznaczaniu niepewności pomiarów z zachowaniem zasad obliczania niepewności wg GUM. Niemal równolegle z publikacjami [9–13] ukazała się praca o wyznaczaniu niepewności danych autoskorelowanych N. F. Zhanga [8] z *National Institute of Standards and Technology* NIST – głównej amerykańskiej placówki metrologicznej, a nieco później praca T. J. Witta z BIPM [14]. Obaj autorzy nie stosowali jednak w rozważaniach zastępczej liczby pomiarów n_{eff}.

Mykhaylo Dorozhovets zbadał też wpływ nieznajomości funkcji autokorelacji na dokładność niepewności u_A [16]. Tematykę uwzględniania autokorelacji analizował następnie A. Zięba z AGH. Natrafił on na znacznie wcześniejsze publikacje dotyczące tego zagadnienia, uściślił wzory określające niepewność podane w [9–12] i omówił, jak wyznaczać estymatę funkcji autokorelacji z pomiarów [16–19, M7]. Dorozhovets natomiast opracował jeszcze metodę pośredniego wykrywania autokorelacji w obserwacjach losowych z zastosowaniem testu F i sprawdził jej skutecz-

ność przez symulację metodą Monte Carlo [21, 23, 25]. Poniżej omówiono stan zagadnienia wynikający z prac [10–13, 19–25] podanych w literaturze tego rozdziału.

Jeśli składowe deterministyczne w danych próbki pomiarowej są pomijalne lub wyeliminowane zostały różnymi metodami – przez poprawki, filtrację i czyszczenie danych wyjściowych, to w odpowiednio krótkim czasie zbierania danych można je z wystarczającym przybliżeniem opisać stacjonarnym szeregiem losowym uzyskanym z próbkowania badanego procesu.

Związek statystyczny między realizacjami X_i, X_{i+k} takiego szeregu charakteryzuje się funkcją autokorelacji

$$\rho_k = \frac{\operatorname{cov}(X_i, X_{i+k})}{\sigma^2}. \tag{2}$$

Funkcja ρ_k zależy od widma częstotliwościowego i jest bądź znana *a priori* dla badanego procesu, bądź z danych pomiarowych próbki należy znaleźć jej estymatę r_k. Praktycznie we wszystkich pomiarach fizycznych autokorelacja jest dodatnia.

Przy występowaniu autokorelacji z wariancji sumy zmiennych losowych wynika następujący związek między odchyleniem standardowym średniej i pojedynczej obserwacji [9–13].

$$\frac{\sigma(\bar{x})}{\sigma} = \frac{1}{\sqrt{n_{eff}}}, \tag{3}$$

$$\text{gdzie} \qquad n_{eff} = \frac{n}{1 + 2\sum_{k=1}^{n-1}(1 - k/n)\rho_k} \equiv \frac{n}{1 + D_\rho}. \tag{4}$$

Dla korelacji dodatnich zachodzi $0 < \rho_k < 1$ i $n_{\text{eff}} < n$.

Dla obserwacji statystycznie niezależnych, tj. gdy $\rho_k \to 0$ (dla $k \geq 1$), człon $D_\rho = 0$, $n_{eff} = n$ – zależność (3) przechodzi w znane wyrażenie $\sigma(\bar{x})/\sigma = 1/\sqrt{n}$.

Gdy $\rho_k \to 1$, to obserwacje stają się całkowicie skorelowane (tj. ściśle uzależnione) i z zależności (4) wynika

$$D_\rho \to \frac{2}{n}\sum_{k=1}^{n-1}(n-k)\cdot 1 = n - 1.$$

Wówczas zgodnie z (4) $n_{eff} = 1$, tj. zanika składowa losowa. Rozrzut obserwacji nie będzie już losowy i będzie zależał tylko od występowania regularnie zmiennych błędów systematycznych. Przy całkowitym ich wyeliminowaniu, dla menzurandu o wartości stałej kolejno powtarzane obserwacje będą jednakowe, a w pomiarach zmiennego menzurandu wartości obserwacji będą leżeć na krzywej odległej od opisującej go funkcji o występujący w danym eksperymencie nieznany błąd systematyczny. Znajomość wartości efektywnej liczby obserwacji n_{eff} umożliwia prawidłowe estymowanie niepewności pomiarów u_A próbki o składowej losowej z autokorelacją.

W tabeli 1 zestawiono dla porównania wzory do wyznaczenia niepewności dla próbki pomiarowej o danych nieskorelowanych i z autokorelacją [20, 22–24].

Tab. 1. Wzory do wyznaczenia niepewności próbki pomiarowej o danych nieskorelowanych i z autokorelacją

Tab. 1. Formulas for evaluation of the uncertainty of the measurement sample of non-correlated and autocorrelated data

Parametr	bez autokorelacji (wg GUM)	z autokorelacją
Efektywna liczba pomiarów	n	$n_{eff} = \dfrac{n}{1+2\sum_{k=1}^{n-1}(1-k/n)\rho_k} \equiv \dfrac{n}{1+D_\rho}$ (5) gdzie: ρ_k - funkcja autokorelacji w punkcie k
Odchylenie standardowe pojedynczej obserwacji	$s(x_i)=\sqrt{\dfrac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2}$ gdzie $\bar{x}$ - wartość średnia	$s_a(x_i)=k_a s(x_i)$ (6) gdzie: $k_a=\sqrt{\dfrac{n_{eff}(n-1)}{n(n_{eff}-1)}}\approx 1$ (6a)
Odchylenie standardowe średniej	$s(\bar{x})=\dfrac{s(x_i)}{\sqrt{n}}=\sqrt{\dfrac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n(n-1)}}$ $\equiv u_A$	$s_a(\bar{x})=\dfrac{s_a(x_i)}{\sqrt{n_{eff}}}=k_b s(\bar{x})=\sqrt{\dfrac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n(n_{eff}-1)}}$ (7) gdzie: $k_b=\sqrt{\dfrac{n-1}{n_{eff}-1}}$ (7a)
Liczba stopni swobody	$\nu=n-1$	$\nu_{eff} \cong \dfrac{n}{1+2\sum_{k=1}^{n-1}\rho_k^2}-1$ (8)
Względna dyspersja odchylenia standardowego	$\dfrac{u(s)}{s}=\dfrac{u(s(\bar{x}))}{s(\bar{x})}\cong\dfrac{1}{\sqrt{2\nu}}$	$\dfrac{u(s_a)}{s_a}=\dfrac{u(s_a(\bar{x}))}{s_a(\bar{x})}\cong\dfrac{1}{\sqrt{2\nu_{eff}}}$ (9)
Niepewność rozszerzona	$U=k_p u=k_p\sqrt{u_A^2+u_B^2}$	$U=k_p u=k_p\sqrt{s_a^2+u_B^2}$ (10)

Ze wzorów (3), (6) i (7) w tabeli 1 wynika, że do oszacowania niepewności $s_a(q_i)$ i $s_a(\bar{x})$, można bądź posługiwać się n_{eff}, bądź używać tylko samych współczynników korekcyjnych k_a i k_b dla $s(q_i)$ i $s(\bar{x})$ wyznaczonych wg GUM. Współczynniki te otrzymuje się po wyznaczeniu członu D_ρ z (4).

Dla dużych n i n_{eff} współczynnik $k_a \to 1$, a nawet i dla małych n, niewiele odbiega od 1. Na przykład dla $n = 10$ i $D_\rho = 1$ otrzymuje się $n_{eff} = 5$ i $k_a = 1{,}06$. Wyniki obliczeń $s_a(q_i)$ wg (6) i $s(q_i)$ wg GUM różnią się więc tu zaledwie o 6%, co jest w pełni dopuszczalne. Autokorelacja jest natomiast istotna dla odchylenia standardowego wartości średniej. Z zależności (7a) otrzymuje się $k_b = 1{,}5$. Stąd wynika, że dla tego przypadku niepewność średniej u_{Aeff} danych skorelowanych wg (7) jest w rzeczywistości aż o 50%(!) większa niż u_A obliczona wg GUM.

Dla danych z autokorelacją można też wprowadzić zastępczą, tzw. efektywną liczbę stopni swobody, określoną w przybliżeniu zależnością (8) podaną w [16]. Wynika z niej, że $\nu_{eff} \neq n_{eff} - 1$. Za pomocą tej liczby, podobnie jak przez ν dla danych nieskorelowanych, opisuje się względną dyspersję obu odchyleń standardowych (9), która również wzrasta wraz ze zmniejszaniem się stosunku n_{eff}/n.

Zaletą powyższego ujęcia jest to, że dla znanej funkcji autokorelacji ρ_k lub jej estymaty r_k oszacowanej z danych próbki, wyznacza się bądź współczynniki k_a i k_b oraz skorygowane oszacowania niepewności $s_a(q_i)$ i $s_a(\bar{x})$, bądź zastępczą liczbę pomiarów n_{eff} i ν_{eff}, które tak samo jak n i ν można stosować dalej w algorytmie obliczania niepewności rozszerzonej $U(x)$ wg GUM. Jeśli wiadomo, że mają one inny znany rozkład, to należy stosować wzory dla najlepszego estymatora tego rozkładu i jego odchylenia standardowego. Uwzględnianie autokorelacji przy wyznaczaniu niepewności z próbek danych pochodzących z rozkładów niegaussowskich nie zostało jeszcze opracowane.

Dla znanej funkcji autokorelacji można też określić maksymalną częstość równomiernego próbkowania, przy której obserwacje można traktować jako nieskorelowane. Minimalny odstęp $\Delta T = T_0/n$ między obserwacjami, czyli okres próbkowania powinien być większy od połowy zastępczej szerokości funkcji autokorelacji ρ_k (szerokości prostokąta o wysokości 1 i polu pod tą krzywą). Dla krótszych odstępów ΔT należy stosować wzory (4)–(8) uwzględniające autokorelację.

Ponadto z warunku Nyquista wynika, że do wyznaczenia przebiegu sygnału o paśmie częstotliwościowym 0–B konieczne są co najmniej dwa pomiary dla każdego okresu przebiegu harmonicznego o najwyższej pożądanej częstotliwości. Stąd przy równomiernym próbkowaniu minimalna liczba obserwacji w próbce $n_{\min}$ zależy od górnej częstotliwości B sygnału i czasu zbierania obserwacji T w następujący sposób

$$n_{\min} = 2BT. \tag{11}$$

Estymator funkcji autokorelacji próbki

Funkcja autokorelacji zwykle nie jest znana i trzeba ją estymować z danych pomiarowych. Najczęściej stosowaną i implementowaną w programach komputerowych jest jej postać

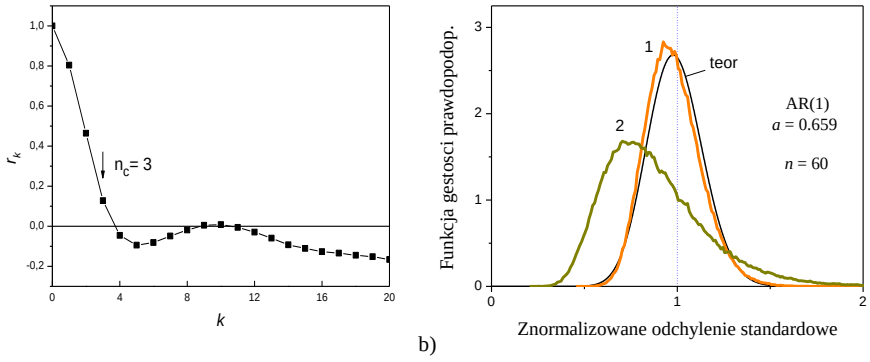
$$r_k = \frac{\sum_{i=1}^{n-k} (x_i - \bar{x})(x_{i+k} - \bar{x})}{s^2(q_i)}. \tag{12}$$

Wykres estymaty r_k (rys. 2a) ma zwykle dwie, różne jakościowo części. Dla małej liczby odstępów k występuje *zbocze* opadające, w którym jest zawarta realna informacja o funkcji autokorelacji. Pozostała część w postaci *ogona* jest obrazem dość dużych fluktuacji skorelowanego szumu.

Według Zięby [15] zastąpienie we wzorze (4) funkcji ρ_k przez jej estymatę r_k wg (12), tak jak to podano w [5], daje estymator efektywnej liczby obserwacji o niekorzystnych właściwościach. Przyczyną jest wpływ ogona funkcji autokorelacji. Sumowanie w (12) ogranicza się do kilku początkowych elementów estymaty r_k, tj.

$$\hat{n}_{eff} = \frac{n}{1+2\sum_{k=1}^{n_c}\left(1-\frac{k}{n}\right)r_k} \tag{13}$$

Graniczny odstęp n_c określany jest przez ostatni niezerowy element estymaty r_k przed jej pierwszym przejściem przez zero (metoda FTZ – ang. *first transit through zero*). Na przykład dla krzywej z rys. 1a wartość $n_c = 3$. Z symulacyjnych badań metodą Monte Carlo [16, 17] wynika, że powoduje to zmniejszenie ujemnego obciążenia estymatora $\hat{n}_{eff}$.



Rys. 1. a) Początkowa część estymaty funkcji autokorelacji $\{r_k\}$ obliczona z danych z rys. 1a (wyznaczonych z danych surowych po odjęciu trendu), b) Funkcje gęstości prawdopodobieństwa dla estymatorów znormalizowanego odchylenia standardowego wg modelu AR(1)[1] i wielkości próby losowej $n = 60$ [3]. Krzywe 1 i 2 pochodzą z symulacji MC i dotyczą odpowiednio s_a/σ i $s_a(\bar{x})/\sigma(\bar{x})$. Zależność teoretyczną 'teor' obliczono za pomocą wzoru (16) dla $v_{eff} = 22{,}7$ [16–24]

Fig. 1. a) The initial part of the estimate $\{r_k\}$ of autocorrelation function ρ_k computed from data of Figure 1a (after removing the trend from the raw data); b). Probability density functions for normalized estimators of the standard deviation by the first order autoregressive model AR(1)[2] and the random sample size $n = 60$ [5, 6]. Curves 1 and 2 – derived from MC simulations and relate, respectively, s_a/σ and $s_a(\bar{x})/\sigma(\bar{x})$. Function 'teor' is calculated theoretically by the equation (16) for $v_{eff} = 22.7$ [16–25]

Metoda FTZ obowiązuje tylko dla korelacji dodatnich. Uzyskaną wartość $\hat{n}_{eff}$ stosuje się do obliczenia $s_a(x_i)$ i $s_a(\bar{x})$ wg (5b) i (5). Na rysunku 1b pokazano otrzymane metodą MC przykłady rozkładów prawdopodobieństwa dwu estymatorów. Przebieg oznaczony jako *teor* na rys. 1b obliczono ze wzoru dla obserwacji nieskorelowanych o odchyleniu standardowym $z = s/\sigma$, wynikającego z rozkładu χ^2, w którym za ν podstawiono efektywną liczbę stopni swobody v_{eff}.

[1] Model AR(1) – autoregresyjny szereg czasowy pierwszego rzędu, dla którego $r_k = a^k$ [13, 14, M4].

[2] Model AR(1) – autoregressive time series of the first order, for which $r_k = a^k$ [4].

Przykład

W przykładzie z rozdziału 2 dla 120 danych oczyszczonych z trendu i niewielkiej składowej sinusoidalnej, ale bez uwzględniania wpływu autokorelacji, otrzymano odchylenie standardowe pojedynczego pomiaru i średniej odpowiednio $s(q_i) = 0{,}024$ i $s(\bar{q}) = 0{,}0022$ (tabela 2).

Tab. 2. Parametry statystyczne próbki surowej i po czyszczeniach z trendu i sinusoidy oraz po uwzględnieniu wpływu autokorelacji

Tab. 2. Statistical parameters of the raw sample, then after cleaning from trend and sinusoid and with influence of autocorrelation

Wyniki pomiarów	Dane surowe v_i	Dane oczyszczone z:	
		trendu q_i'	trendu i sinusoidy q_i
Kryterium $\chi 2$ dla rozkładu normalnego	$\chi^2 = 52{,}28 > \chi^2_{5,\,0{,}05} = 11{,}1$	$\chi^2 = 3{,}83 < \chi^2_{5,\,0{,}05} = 11{,}1$	$\chi^2 = 3{,}25 < \chi^2_{5,\,0{,}05} = 11{,}1$
	rezultat negatywny	**rezultat pozytywny**	
Wartość średnia	$\bar{v} = 1{,}2029$	$\bar{q}' = 1{,}2028$	$\bar{q} = 1{,}2026$
Standardowe odchylenie próbki	$s(v_i) = 0{,}040$	$s(q_i') = 0{,}02409 \approx 0{,}024$	$s(q_i) = 0{,}02407$
Niepewność u_A	$0{,}003593 \approx 0{,}0036$	$0{,}002190 \approx 0{,}0022$	$0{,}002188 \approx 0{,}0022$
Stosunek odchyleń standardowych	$\frac{s(v_i)}{s(q_i')} = \frac{0{,}03953}{0{,}02409} \approx 1{,}641$		$\frac{s(v_i)}{s(q)} \approx 1{,}6420$
Wynik pomiaru	bez uwzględnienia autokorelacji dla $u = 2u_A$		$x = 1{,}2026 \pm 0{,}0044$
	z uwzględnieniem autokorelacji dla: $s_a(q_i) = 0{,}0244$, $s_a(\bar{q}) = 0{,}0047$, $U = 2s_a$		$x_a = 1{,}2026 \pm 0{,}0095$

Dalej zostaną obliczone odchylenia $s_a(q_i)$ i $s_a(\bar{q})$ dla estymaty r_k funkcji autokorelacji danych q_i wyznaczonej zgodnie z rys. 1a dla $n_c = 3$. Otrzymana z (11) wartość $r_k = 0{,}81$ jest dość duża, co potwierdza przewidywaną autokorelację tych danych. Dla tej wartości r_k z (9) wyliczana jest estymata efektywnej liczby obserwacji $n_{eff} = 32{,}1$ wynosząca tylko około 25% liczby rzeczywistych obserwacji $n = 120$. Na podstawie zależności (6a) i (7a) – tabela 1 – otrzymuje się współczynniki $k_a = 1{,}012$ i $k_b = 1{,}96$. Po ich uwzględnieniu uzyskano zaokrąglone do 3 cyfr znaczących wartości $s_a(q_i) = 0{,}00244$ oraz $s_a(\bar{q}) = 0{,}00472$. Autokorelacja nie ma tu istotnego wpływu na odchylenie $s_a(q_i)$ pojedynczego pomiaru, zaś znaczny na standardowe odchylenie średniej $s_a(\bar{q})$, czyli na odpowiednik niepewności u_A dla skorelowanych obserwacji. Wzrasta ona około dwukrotnie. Wyliczone wartości zapisano w tabeli 2. Zapis końcowy wyniku pomiarów bez i z uwzględnieniem autokorelacji znajdują się w ostatnim jej wierszu.

Z wzorów (6) i (7) w tabeli 1 wynika też, że względna standardowa niepewność wyznaczenia obu odchyleń wynosi $u(s_a)/s_a \approx 11\%$. Jest więc nieco większa niż 6,48% dla przypadku nieskorelowanych 120 obserwacji, ale nieco mniejsza niż 12,5% dla rzadziej pobieranych 33 nieskorelowanych obserwacji. Tak więc przy

niemal czterokrotnie częstszym próbkowaniu zysk na dokładności wyznaczenia niepewności u jest tu praktycznie pomijalny.

Propozycję realizacji przyrządu wirtualnego o strukturze zawierającej funkcje czyszczenia danych i wyznaczania na bieżąco niepewności typu A mierzonego sygnału opisano w [7].

Wnioski

Podjęcie omówionej tematyki było wynikiem potrzeb praktyki pomiarowej, wymagającej by przy próbkowaniu sygnału poprawnie wyznaczać niepewność statystyczną otrzymywanych danych pomiarowych. Zaprezentowane ujęcie przeznaczone jest do powszechnego stosowania. Najważniejsze wnioski podano poniżej.

1. Przy stosowaniu w praktyce metody zmniejszenia niepewności pomiarów, polegającej na zwiększaniu liczby obserwacji w ograniczonym czasie zbierania obserwacji pomiarowych, może pojawić się negatywny wpływ autokorelacji.
2. Aby uniknąć pomyłek przy wyznaczeniu niepewności $u_A(x)$, należy w otrzymanej próbce obserwacji pomiarowych zidentyfikować odpowiednimi metodami obliczeniowymi nie tylko mierzone, ale i niepożądane składowe systematyczne, aperiodyczne i periodyczne. Należy wyeliminować je przed wyznaczaniem niepewności pomiarów typu A, gdyż pozostawione w danych pomiarowych regularne składowe zwiększą jej wartość.
3. Dopiero dla tak oczyszczonych danych należy oszacować funkcję autokorelacji. Autokorelacja powoduje zwiększenie składowej niepewności typu A wyniku pomiaru, gdyż liczba niezależnych obserwacji pomiarowych jest wtedy mniejsza niż zebranych. Trzeba wówczas wyznaczyć bądź współczynniki korygujące wartości odchylenia standardowego próbki oraz niepewności u_A wartości średniej, bądź też zastępczą, tzw. efektywną liczbę pomiarów, którą należy uwzględniać przy szacowaniu niepewności typu A. Wzory dla takiego przypadku podano w tabeli 1 i uwzględniono je w obliczeniu niepewności wyniku w tabeli 2 dla przykładu, gdy występuje autokorelacja obserwacji próbki.
4. Programy komputerowe stosowane do obliczania niepewności typu A i niepewności rozszerzonej U powinno być uzupełnione poprzedzającymi algorytmami do identyfikacji i eliminacji z surowych danych ich składowych aperiodycznych i periodycznych, a także algorytmami do wyznaczania estymatorów funkcji autokorelacji oczyszczonych danych.
   Należy też zmodyfikować wzory stosowane w tych programach, aby uwzględniały wpływ autokorelacji – przez użycie w nich efektywnej liczby obserwacji n_{eff} lub współczynników k_a oraz k_b z tabeli 1.
5. Omówiony tu sposób wyznaczania niepewności u_A dla danych z autokorelacją dotyczy, podobnie jak w Przewodniku GUM, pomiarów modelowanych rozkładem normalnym. Da innych rozkładów wpływ będzie jakościowo podobny, ale wzory mogą być inne. Wymaga to jeszcze dalszych badań.

Literatura

1. *Guide to the Expression of Uncertainty in Measurement*, revised and corrected version GUM of 1995, BIPM_JCGM 100:2008.

2. *Wyrażanie Niepewności Pomiaru. Przewodnik,* tłumaczenie i komentarz J. Jaworskiego, Wydawnictwo Głównego Urzędu Miar Alfavero, Warszawa 1999, 2002.
3. Evaluation of measurement data – Supplement 1 to the "Guide to the expression of uncertainty in measurement", Propagation of distributions using a Monte Carlo method. BIPM_JCGM:101, 2008.
4. Pavese F., Ichim D., *SAODR: Sequence analysis for outlier data rejection*, "Measurement Science and Technology", Vol. 15, 2004, 2047–2052.
5. Warsza Z.L., Dorozhovets M., Korczyński M.J., *Methods of upgrading the uncertainty of type A evaluation* (1). *Elimination the influence of unknown drift and harmonic components*, Proceedings of 15th IMEKO TC4 Symposium, Iasi Romania, 2007, 193–198.
6. Warsza Z.L., Korczyński J., *Eliminacja wpływu nieznanych a priori składowych systematycznych na niepewność u_A pomiarów o równomiernym próbkowaniu*, „Pomiary Automatyka Robotyka", Nr 2, 2008, 5–13.
7. Warsza Z.L., Korczyński J.M., *A New Instrument Enriched by Type A Uncertainty Evaluation*; Proceedings of 16th IMEKO TC4 Symposium "Exploring New Frontiers of Instrumentation and Methods for Electrical and Electronic Measurements", September 22–24, 2008, Florence Italy (CD).
8. Nien Fan Zhang, *Calculation of the uncertainty of the mean of autocorrelated measurements*. "Metrologia", 43, 2006, 276–281.
9. Dorozhovetz M., *Wybrane problemy praktycznej oceny błędów i niepewności wyników pomiaru*. Zeszyty Naukowe Politechniki Rzeszowskiej, Nr 122 (*Elektrotechnika* z. 29), 2006, 9–44.
10. Dorozhovets M., Warsza Z.L., *Udoskonalenie metod wyznaczania niepewności wyników pomiaru w praktyce*. „Przegląd Elektrotechniki", Nr 1, 2007, 1–13.
11. Dorozhovets M., Warsza Z.L, *Propozycje rozszerzenia metod wyznaczania niepewności wyniku pomiarów wg Przewodnika GUM* (1). „Pomiary Automatyka Robotyka", Nr 1, 2007, 16–25.
12. Dorozhovets M., Warsza Z., *Wyznaczanie niepewności typu A pomiarów o skorelowanych rezultatach obserwacji*. „Pomiary Automatyka Kontrola", Nr 2, 2007, 20–25.
13. Dorozhovets M., Warsza Z., *Methods of upgrading the uncertainty of type A evaluation, Part 2. Elimination of the influence of autocorrelation of observations and choosing the adequate distribution*. Proceedings of 15th IMEKO TC4 Symposium, Iasi Romania, 2007, 199–204.
14. Witt T.J., *Using the autocorrelation function to characterize time series of voltage measurements*. "Metrologia", No. 44, 2007, 201–209.
15. Dorozhovets M., *Wpływ braku znajomości a priori funkcji autokorelacji obserwacji na ocenę standardowej niepewności ich wartości średniej*. „Pomiary Automatyka Kontrola", Nr 12, 2009, 89–92.
16. Zięba A., *Effective number of observations and unbiased estimators of variance for autocorrelated data – an overview*, "Metrology and Measurement Systems", No. 17, 2010, 3–16.

17. Zięba A., *Niepewność pomiaru dla ciągu obserwacji samoskorelowanych*. Monografia: Niepewność pomiarów w teorii i praktyce. Praca zbiorowa. GUM Warszawa 2011, 109–118.
18. Zięba A., Ramza P., *Standard deviation of the mean of autocorrelated observations estimated with the use of the autocorrelation function estimated from the data*. "Metrology and Measurement Systems", No. 18, 2011, 529–542.
19. Warsza Z.L., Zięba A., *Niepewność typu A pomiaru o obserwacjach samoskorelowanych*. „Pomiary Automatyka Kontrola", Nr 2, 2012, 157–162.
20. Warsza Z.L., *Szacowanie niepewności w pomiarach z autokorelacją obserwacji.* „Elektronika", Nr 8, 2012 (wkładka *Technika Sensorowa*), 108–113.
21. Dorozhovets M., *Using F-Test for the Indirect Detecting of the Autocorrelation of Random Observations*. Proceedings of The 7th IEEE Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications. 2013, Berlin, Germany.
22. Warsza Z.L., *Evaluation of the type A uncertainty in measurements with autocorrelated observations.* Journal of Physics. Conference series 459(2013) 012035. Joint MEKO TC1+TC7+TC13 Symposium: Measurement Across Physical and Behavioral Sciences Genova 4–6 Sept. 2013, DOI: 10.1088/1742-6596/459/1/0120356.
23. Warsza Z.L., Dorozhovets M., *Evaluation of the uncertainty type A of the random stationary signal component from its autocorrelated observations*. Proceedings of 20th IMEKO TC4 International Symposium. Benevento, Italy, Sept. 15–17, 2014, 367–372.
24. Warsza Z.L., *Evaluation of the standard deviation of the random component of the measured signal from its autocorrelated observations*. Advances in Intelligent Systems and Computing, Vol. 267, Recent Advances in Automation, Robotics and Measuring Techniques (R. Szewczyk, C. Zieliński, M. Kaliczyńska, eds.), Springer 2014, 733–741.
25. Warsza Z.L., Dorozhovets M., *Evaluation of the uncertainty type A of the random stationary signal component from its autocorrelated observations.* "Measurement Automation Monitoring", Vol. 61, No. 8, 2015, 399–402.

Literatura uzupełniająca książkowa

M1 Bendat J.S., Piersol A.G., *Metody analizy i pomiaru sygnałów losowych*. WNT, Warszawa 1976 (tłumaczenie wydania z 1971 r. oryginału: *Random Data. Analysis and measurement procedures* John Wiley & Sons N.York, Chichester, Brisbane).
M2 Korn G.A., Korn T.M., *Matematyka dla pracowników naukowych i inżynierów*. PWN, Warszawa 1983 (tłum. polskie oryginału: *Mathematical Handbook for Scientists and Engineers* McGraw-Hill, Co., New York, San Francisco, 1968).
M3 Tylor J.R., *Wstęp do analizy błędu pomiarowego*. Wydawnictwo Naukowe PWN, Warszawa 1995 (tłumaczenie z ang.: *An Introduction to error analysis. The study of uncertainties in physical measurements*. Oxford University Press California 1982).
M4 Box G.E.P., Jenkins G.M, Reinsel G.C., *Time Series Analysis: Forecasting and Control*. Prentice Hall, New Jersey 1994. (polskie tłumaczenie poprzedniego

wydania: *Analiza szeregów czasowych. Prognozowanie i sterowanie.* PWN, Warszawa 1983).

M5 Piotrowski J., *Procesy pomiarowe i estymacja sygnałów*. Skrypt nr 1889 Politechniki Śląskiej, Gliwice 1994.

M6 Piotrowski J., Kostyrko K., *Wzorcowanie aparatury pomiarowej*. PWN, Warszawa 2000, 2012.

M7 Zięba A., *Analiza danych w naukach ścisłych i technice*. Wydawnictwo Naukowe PWN, Warszawa 2013.

Rozdział 4

Funkcja cos^2 jako model rozkładu gęstości prawdopodobieństwa – właściwości i przykłady stosowania

Streszczenie. Omówiono niekonwencjonalny rozkład gęstości prawdopodobieństwa (PDF) w postaci jednego okresu cosinusoidy przesuniętej w górę i o polu równym 1 jako alternatywę rozkładu normalnego (funkcji Gaussa). Wykazano, że w przedziale ±2,5 odchylenia standardowego jeden okres cosinusoidy przesuniętej w górę o jej amplitudę *A*, czyli funkcja $2A\cos^2(\cdot)$ nie odbiega od funkcji rozkładu normalnego więcej niż o ±2,1%. Taki rozkład oznaczono symbolem COS^2. Dla rzeczywistych wyników pomiarów o ograniczonym zakresie rozrzutu ma on mniejszą niepewność typu A niż wyznaczana wg międzynarodowego Przewodnika GUM. Rozważania teoretyczne i symulacje poparto przykładem z praktyki metrologicznej. Rozkład COS^2 można z powodzeniem stosować do szybkiego szacowania niepewności pomiarów, w tym do automatyzacji tych obliczeń oraz w statystyce.

Do opisu rozkładu gęstości prawdopodobieństwa (PDF) zbioru wyników obserwacji pomiarowych zwykle stosuje się dwuparametrową wykładniczą funkcję Gaussa nazywaną rozkładem normalnym. Rozkład ten przebiega w nieograniczonym zakresie wartości wielkości mierzonej ±∞ i jest modelem teoretycznym, który występuje tylko przy nieskończenie dużej liczbie czynników losowych niezależnie oddziałujących na cały tor pomiarowy wraz z obiektem badanym. Nie zachodzi to jednak w praktyce ani w eksperymentach pomiarowych, ani w opisach statystycznych rozrzutu mierzonych właściwości obiektów i procesów technicznych. Przy losowym oddziaływaniu skończonej liczby czynników rozstęp, czyli przedział rozrzutu wartości wyników obserwacji pomiarowych, jest zawsze ograniczony. Pojawiają się więc niedogodności w stosowaniu rozkładu normalnego, gdyż:

- dla wyników pomiarów o ograniczonym rozrzucie, standardowa niepewność $u_A(x)$ wielkości mierzonej, liczona jako standardowe odchylenie średniej $s(\bar{x})$ jak dla rozkładu normalnego o nieograniczonym przebiegu funkcji Gaussa, jest zawsze nieco zawyżona. W nieco mniejszym stopniu dotyczy to też stosowania rozkładu Studenta przy małej liczbie danych i zależności $u(x) = \sqrt{\frac{n-1}{n-3}}\, s(\bar{x})$,
- nie można oszacować ograniczonej szerokości przedziału dla rozrzutu rzeczywistych danych, w którym wynik pomiaru miałby prawdopodobieństwo 1,
- rozkład normalny jest funkcją wykładniczą Gaussa. Jej całka nie jest opisana przez funkcje elementarne i wartości tej całki trzeba wyznaczać numerycznie korzystając z tabel, kalkulatora lub komputera.

W wielu pomiarach użytkowych niedogodności te nie są istotne, gdyż wystarcza ocena niepewności z dokładnością do 20%. Natomiast dokładność wyznaczania

niepewności powinna być większa przy kalibracji wzorców oraz w badaniach surowców, urządzeń i wyrobów przy kontroli ich parametrów i jakości w handlu i diagnostyce, tj. gdy opis parametrów rozrzutu właściwości służy decyzjom o kwalifikacji i cenie oraz w badaniach stanu środowiska. Tych trudności nie rozwiązuje w wystarczająco prosty sposób rozkład pseudo-normalny o obciętych końcach i polu pod krzywą znormalizowanym do 1 ze względu na skomplikowaną postać jego opisu. Dane eksperymentalne bywają też opisywane innymi modelami rozkładów o skończonym rozstępie.

W międzynarodowym Przewodniku GUM [1] składową statystyczną niepewności pomiarów u_A zaleca się szacować metodą typu A tak, jak dla rozkładu normalnego. Wynikiem prac nad udoskonaleniem i rozszerzeniem zakresu stosowania metod wyznaczania niepewności objętych Przewodnikiem GUM są kolejne Suplementy, np. [2] (patrz rozdz. 1).

Prawidłowy proces wyznaczania wypadkowej, tzw. rozszerzonej niepewności wyniku pomiarów wymaga splatania rozkładów jej składników u_A i u_B. Obok dość uciążliwej metody analitycznej polegającej na sumowaniu wariancji obu składowych, stosuje się metodę numeryczną Monte Carlo opisaną w Suplemencie 1 [2]. Wymaga ona wiele setek tysięcy operacji i w szeregu przypadkach, nawet przy szybkich obliczeniach komputerowych, może być zbyt powolna do zastosowania w automatycznych systemach pomiarowych dla obliczeń w czasie rzeczywistym (on-line). Hetman i Korczyński w pracy [4] zaproponowali zastosowanie metody szybkiej transformaty Fouriera FFT. Obejmuje ona bądź generację wartości funkcji rozkładów składowych, bądź zapamiętywanie tabel ich wartości. Są jednak pewne trudności w splataniu rozkładów dyskretnych zawierających funkcję delta Diraca.

W badaniach procesów przemysłowych w trakcie ich eksploatacji zwykle wymagane jest zautomatyzowane przetwarzanie sygnałów pomiarowych, wbudowane w systemy przyrządów i czujniki, szczególnie przeznaczone do pracy ciągłej. Tak realizowany pomiar umożliwia na bieżąco wyznaczanie składowych regularnych i statystycznych sygnału wartości mierzonej (menzurandu), bądź wyznaczanie niepewności pomiarów wartości stałej lub o określonym przebiegu [3]. Trudno jest tu zastosować metodę Monte Carlo ze względu na zbyt długi czas obliczeń w stosunku do wymagań bieżącej kontroli i sterowania procesów, a implementacja metody FFT w prostszych urządzeniach pomiarowych zwykle nie jest możliwa.

Na podstawie wyników tej analizy sformułowano dwie tezy robocze jako zadania dla opisywanych tu badań.

- Informacja o szerokości przedziału, w którym występują wyniki pomiarów pozwoli na lepsze oszacowanie niepewności pomiarowych typu A i B niż według Przewodnika GUM [1, 2]. Uzyska się większą dokładność estymacji wartości menzurandu (szerokość przedziału niepewności rozszerzonej) dla tych samych eksperymentalnych danych.
- Należy zbadać zastosowanie pojedynczego okresu cosinusoidy podniesionej o amplitudę lub o inną stałą wartość, jako prostego modelu rozkładu PDF o ograniczonym zakresie i kształcie bliskim środkowej części funkcji Gaussa.

Dla sprawdzenia tych tez konieczne było opracowanie podstaw teoretycznych wraz ze szczegółowym zbadaniem właściwości takiego niekonwencjonalnego modelu PDF. Wstępne wyniki badań opisano w [5]. Już po ich ukończeniu natrafiono na

publikacje sprzed pół wieku (1961 r.), gdzie Green i Rabb [10] proponowali jedną z postaci funkcji $\cos^2(\cdot)$ do opisu rozkładu wyników badań w psychometrii. Nie użyli oni jednak żadnego kryterium do doboru optymalnych parametrów tej funkcji i jedynie zalecali wówczas stosowanie cosinusoidy do wizualnego dopasowywania jej na oscyloskopie do otrzymanego eksperymentalnie rozkładu obserwacji. Przyjęta przez nich arbitralnie postać funkcji $\cos^2(\cdot)$ odbiegała dość znacznie (nawet o około 14%) od funkcji Gaussa. Nie spotkano też w późniejszej literaturze śladów stosowania tej funkcji w technice pomiarowej, chociaż podstawowe wzory dla parametrów rozkładu COS^2 podawane są w niektórych poradnikach i w Internecie.

Podstawowe właściwości rozkładu +COS

Rozpatrzmy model funkcji gęstości rozkładu prawdopodobieństwa PDF w postaci jednego okresu funkcji cosinus o amplitudzie A, okresie $2X$ i przesuniętej o odcinek B w dodatnim kierunku osi y:

$$f(x) = B + A\cos\pi\frac{x}{X} \qquad \text{dla} \quad \left(-1 \le \frac{x}{X} \le 1\right) \tag{1}$$

gdzie: x – wartość bieżąca wyniku obserwacji, X – pół zakresu rozpięcia równe 0,5 okresu cosinusoidy, A, B – amplituda cosinusoidy i jej podniesienie.

Powyższa funkcja $f(x)$ ma sens fizyczny jako rozkład PDF dla $f(x) > 0$ oraz w przedziale ograniczonym do jednego okresu $\pm(X \le 0{,}5X_T)$. Zaproponowano, by dla przypadku ogólnego $B \ge 0$ rozkład ten nazywać podniesioną cosinusoidą i oznaczać symbolem +COS [5].

Funkcja $f(x)$ określona wzorem (1) ma trzy niezależne parametry (A, B, X), wyznaczalne z następujących warunków:

- pole pod krzywą między krańcami $\pm X$ wynosi $S = 1$, lub ma inną wartość $S < 1$, np. jak dla rozkładu Gaussa w tym przedziale,
- zadany jest przedział X lub punkty, przez które ma przechodzić funkcja +COS, bądź ma ona być styczna na krańcach do osi x,
- wyznacza się minimum różnicy między danymi a krzywą według określonego kryterium statystycznego, np. średnio-kwadratowego, minimum sumy modułów, Kołmogorowa-Smirnowa, Czebyszewa, χ^2, lub innego kryterium.

W pracach [5–7] rozpatrzono tylko takie funkcje $B + A\cos(\pi x/X)$, które na końcach zakresu $\pm X$ są styczne do osi x (tj. $B = A$) lub do prostej równoległej do tej osi powyżej niej. Wstępna analiza i symulacje wykazały, że funkcje $B + A\cos(\cdot)$ o zakresie mniejszym niż okres, będą bardziej odbiegały od krzywej Gaussa niż funkcje styczne. Dla podniesionej cosinusoidy o okresie $2X$, pole S pod krzywą określa zależność $S = B \cdot 2X$. Jeżeli wartość pola $S = 1$, to $X = \frac{1}{2}\, B$ i tylko dwa parametry A i B (lub A i X) są wówczas niezależne, tj.:

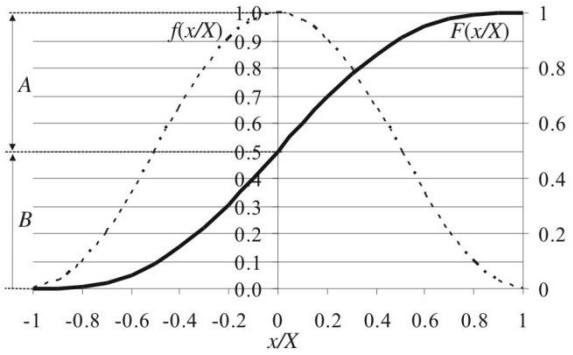
$$f(x) = B + A\cos(2\pi x B) \qquad \text{dla} \quad -1 \le 2xB \le 1 \tag{2}$$

W publikacjach [6, 7] wyznaczono dystrybuantę oraz odchylenie standardowe rozkładu +COS, które wynosi

$$\sigma_{+\cos} = X\sqrt{\frac{1}{3} - \frac{4}{\pi^2}AX} = \frac{1}{2B}\sqrt{\frac{1}{3} - \frac{2}{\pi^2}\frac{A}{B}} \tag{2a}$$

Rozwiązanie jest możliwe dla stosunku $A/B \leq (1/6)\,\pi^2 \approx 1{,}78$.

W ogólnym przypadku $B > A$ kształt funkcji jest zbliżony do dwustronnie obciętego rozkładu normalnego i też jest kłopotliwy w stosowaniu. Dlatego preferuje się szczególny przypadek rozkładu +COS dla $A = B$ i o styczności do osi x w punktach krańcowych $x = \pm X$ (rys. 1).



Rys. 1. Przebieg gęstości prawdopodobieństwa $f(x/X)$ i dystrybuanty $F(x/X)$ dla rozkładu +COS o $A = B$ i wartości średniej zero

Fig. 1. Probability density distribution function $f(x/X)$ and distribution $F(x/X)$ of +COS for $A = B$ and data mean value equal zero

Rozkład COS² czyli cosinusoida podniesiona o amplitudę

Rozpatrzmy bliżej szczególną postać rozkładu +COS z rys. 1 o podniesieniu B równym amplitudzie A. Jego funkcja $f(x)$ jest styczna do osi x dla $x = \pm X$. Wówczas

$$f(x) = A\left(1 + \cos\pi\frac{x}{X}\right) \qquad \text{dla} \quad A = B \tag{3}$$

Jeśli przedział całkowania rozciąga się między punktami styczności $\pm X$ funkcji z osią x, to dla pola $S = 1$ otrzymuje się następujące wartości całek jej składników z (3) $A \cdot 2X = 1$ i $2\sin(x = \pm X) = 0$. Tak więc tylko jeden parametr A lub X jest wtedy niezależny. Stąd

$$A = 0{,}5/X \qquad \text{lub} \qquad X = 1/(2A) \tag{4}$$

Po elementarnych przekształceniach zależności (3) i (4) otrzymuje się

$$f(x) = A\left(1 + \cos 2\pi xA\right) \equiv 2A\cos^2\left(\pi xA\right) \tag{5}$$

Z zależności (5) wynika, że dla $A = B$ funkcja (3) jest $\cos^2(\cdot)$ o amplitudzie $2A$ i mniejszym o połowę argumencie.

Zaproponowano, by ten szczególny przypadek rozkładu +COS nazywać rozkładem COS^2 [5]. Jego parametry A i X są powiązane zgodnie z (3) i wystarczy znaleźć tylko jeden z nich. Podobnie jak rozkład Gaussa jest to rozkład dwuparametrowy. Określa go wartość średnia i parametr charakteryzujący rozrzut, np. X (połowa rozstępu) lub amplituda A.

Rozkład skumulowany CPDF, czyli dystrybuantę rozkładu otrzymuje się z całki $f(x)$ jako pole pod krzywą w funkcji x. Dla rozkładu COS^2, tj. gdy $A = B$, $X = 1/(2A)$, opisuje ją przedstawiona na rys. 1 prosta funkcja

$$F(x) = \frac{1}{2}\left(\frac{x}{X} + 1\right) + \frac{1}{2\pi}\sin\pi\frac{x}{X}\Bigg|_{-X \le x \le X} \tag{6}$$

W tabeli 1 podano wartości dystrybuanty $F(x)$ o przebiegu jak nas rys 1, wyznaczone z (6) dla 21 wartości argumentu x/X.

Tab. 1. Wartości dystrybuanty rozkładu COS^2
Tab. 1. Values of COS^2 distribution

x/X	–1	–0,9	–0,8	–0,7	–0,6	–0,5	–0,4	–0,3	–0,2	–0,1	0
$F(x)$	0	0,001	0,006	0,021	0,049	0,091	0,149	0,221	0,307	0,401	0,5
x/X	1	0,9	0,8	0,7	0,6	0,5	0,4	0,3	0,2	0,1	0
$F(x)$	1	0,999	0,994	0,979	0,951	0,909	0,851	0,780	0,694	0,599	0,5

Odchylenie standardowe dla rozkładów COS^2 (rys. 1 i 2) określa zależność

$$\sigma = \sqrt{\int_{-\infty}^{+\infty} x^2 f(x)\,dx} = \sqrt{\int_{-X}^{+X} x^2 A\cos^2(\pi x A)\,dx} \tag{7}$$

Po rozwiązaniu (7) lub z zależności (2a) dla $A = B$ otrzymuje się prosty wzór

$$\sigma_{+\cos} = X\sqrt{\frac{1}{3} - \frac{2}{\pi^2}} = \frac{1}{2A}0{,}3615 \tag{8}$$

Jeżeli znana jest wartość X lub A, to z (8) łatwo wyznaczyć powiązaną z nimi wartość $\sigma_{+\cos}$ i na odwrót. Jeżeli $\sigma_{+\cos} = 1$, to $A = 0{,}1808$ i $X = 2{,}766$ lub też, gdy dany jest zakres X, np. $X = 2{,}5$ to z (8) $\sigma_{+\cos} = 0{,}9060$.

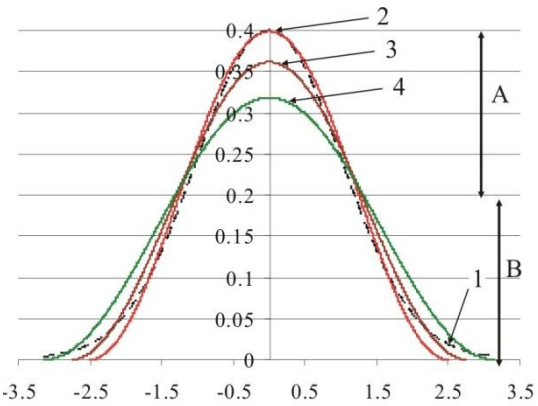
Współczynniki k_{PC} rozkładu COS^2 dla **niepewności rozszerzonej** U o różnym prawdopodobieństwie P poziomu ufności i liczbie obserwacji $n > 10$, podano w tabeli 2. Dla $n \le 10$ należy stosować nieco wyższe wartości tych współczynników, zbliżone do wartości jak przy rozkładzie Studenta-Gosseta.

Tab. 2. Współczynniki k_{PC} rozkładu COS^2 dla oceny niepewności rozszerzonej
Tab. 2. Coefficients k_{PC} of COS^2 distribution for estimation of expanded uncertainty

P (%)	50	68,3	90	95	99	99,7	100
$k_{PC}(x/X)$	0,265	0,385	0,596	0,683	0,816	0,878	1

Porównanie rozkładu COS² z rozkładem normalnym

Rozkład +COS i jego szczególna postać COS² jest rozkładem symetrycznym niegaussowskim o ograniczonym zakresie rozrzutu wyników pomiarowych, czyli rozstępie. Kilka przypadków rozkładu COS² podano na rys. 2 wraz ze znormalizowanym rozkładem normalnym $N(0, 1)$ – krzywa 1, czyli o odchyleniu standardowym $\sigma = 1$.



Rys. 2. Funkcje gęstości prawdopodobieństwa (PDF): 1 – znormalizowany rozkład Gaussa, $N(0, 1)$ oraz trzy rozkłady COS² o równaniu $f(x) = A\cos^2\pi Ax$, tj.: 2 – krzywa przechodząca przez wierzchołek funkcji Gaussa, krzywa 3 – o standardowym odchyleniu $\sigma_{+COS}=1$, krzywa 4 – propozycja Greena [10] $f_{GR}(x)$ o $A = 1/2\pi$

Fig. 2. Probability density functions: PDF: 1 – $N(0, 1)$ Normal distribution (Gauss) and three distributions COS² given by equation $f(x) = A\cos^2\pi Ax$, i.e.: 2 – the curve crossing across top of Gauss function, 3 –curve of which standard deviation equals 1, $\sigma_{+COS} = 1$ and 4 – the Green suggested curve $f_{GR}(x)$ [10] of which $A = 1/2\pi$

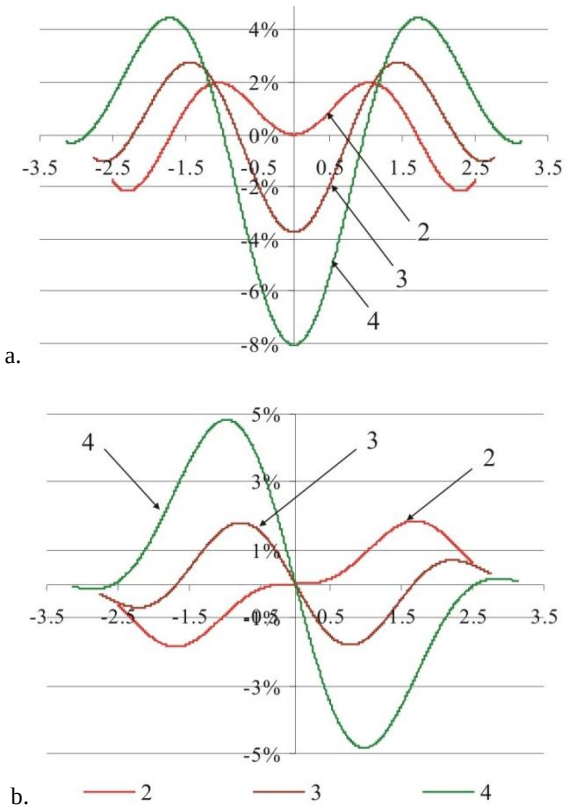
Dla krzywej nr 2 stycznej do wierzchołka rozkładu normalnego $N(0, 1)$

$$A = \frac{1}{2}\frac{1}{\sqrt{2\pi}} = 0{,}19947 \qquad \text{i} \qquad X = \sqrt{2\pi} \approx 2{,}5$$

Wartości A można wyznaczyć też dla założonych zakresów $2X_{opt}$, równych lub mniejszych od okresu $2X$ funkcji (4) przy dopasowaniu funkcji cos² do histogramu danych eksperymentalnych w różny sposób. Można zastosować np. kryterium minimum sumy kwadratów różnic lub modułów różnic (metody MNK i MNM), przybliżać do krzywej otrzymanej po uporządkowaniu obserwacji wg ich wartości, bądź też do ich PDF lub CPDF. Z symulacji wynikło, że wszystkie tak otrzymane funkcje optymalne rozkładu COS² przecinają oś y na wysokości mniejszej niż wierzchołek rozkładu Gaussa dobranego dla tych samych danych [5–9].

Aby porównać przebieg rozkładu COS² z funkcją Gaussa dokonano ich wzajemnej aproksymacji opisanej szczegółowo w publikacjach [5–9]. Znaleziono parametry kilku wariantów rozkładu +COS minimalizując różnicę między obu krzywymi w zadanym arbitralnie przedziale spełniającym warunek $X_{opt} \le X_T$ na osi x (lub odpowiadającym mu przedziale na osi y), np. dla odchyleń standardowych funkcji

Gaussa ±2σ, ±2,5σ. Dla znormalizowanego rozkładu Gaussa o $\sigma = 1$ były to przedziały ±2, ±2,5. Wykorzystano metodę dwuparametrową MKM, znajdując wartości A i B dla minimum kwadratu różnicy między krzywymi, lub też jeden z tych parametrów, np. A wybrano tak, by krzywa przechodziła przez zadany punkt, np. wierzchołek rozkładu Gaussa, lub by były jednakowe odchylenia standardowe σ obu funkcji. **Niedokładność przybliżenia** określono na podstawie różnic między PDF lub dystrybuantami CPDF obu rozkładów przez odchylenie średnie standardowe oraz maksimum i minimum (±) lokalnych odchyleń. Różnice Δ_{PDF} i Δ_{CPDF} trzech funkcji COS^2 (rys. 2) i znormalizowanego rozkładu Gaussa $N(0, 1)$ pokazano na rys. 3.



Rys. 3. a. Różnice Δ_{PDF} rozkładów COS^2 nr 2, 3, 4 (rys. 2) i standardowego normalnego N (0, 1); b. różnice Δ_{CPDF} skumulowanych rozkładów COS^2 nr 2, 3, 4 i std. normalnego N (0, 1)

Fig. 3. a. Differences Δ_{PDF} of distributions COS^2 no 2, 3, 4 of Fig 2 vs. Normal N (0, 1); b. Differences of Δ_{CPDF} cumulative distributions COS^2 no 2, 3, 4 vs. N (0, 1)

Różnice Δ_{PDF} rozkładu COS^2 w postaci krzywej 2, jak również Δ_{CPDF} jej dystrybuanty nie przekraczają ±2,1%. Natomiast Δ_{PDF} krzywej 3 rozkładu COS^2 o tym samym odchyleniu standardowym $\sigma = 1$, zmienia się w nieco większym zakresie (–3,7%, +2,8%), ale nawet tę niedokładność można zaakceptować w większości przypadków przy szacowaniu niepewności typu A.

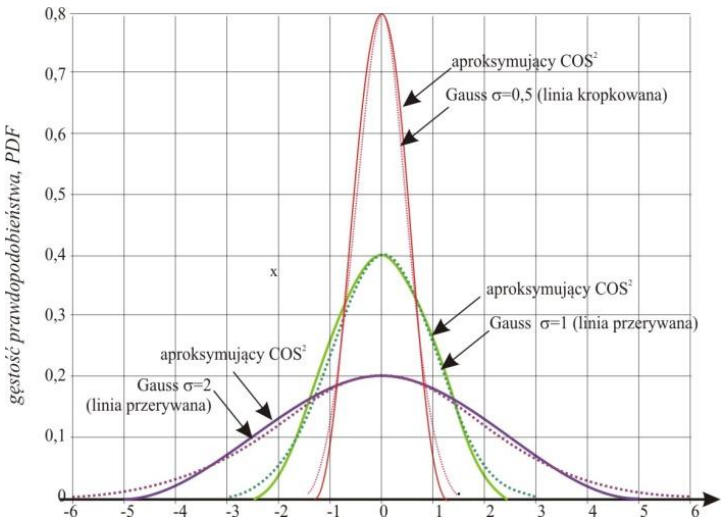
Na rysunku 3 podano też dla porównania parametry charakteryzujące funkcję $f_{GR}(x)$ zaproponowaną przez Greena [8]. Krzywa ta ma $A = 1/(2\pi)$ i odchylenie standardowe $\sigma_{+COS} \approx 1{,}14$. Różnice Δ_{PDF} i Δ_{CPDF} – krzywe 4 na rys. 3 między $f_{GR}(x)$ i znormalizowaną funkcją Gaussa są znacznie większe niż dla funkcji COS^2 nr 3 przechodzącej przez wierzchołek rozkładu normalnego i dochodzą do +4,5% oraz do –8,1%. Są to wartości nie do zaakceptowania dla większości pomiarów w pracach naukowych i badaniach technicznych.

Najlepszą zgodność $\approx \pm 0{,}01$ z rozkładem Gaussa uzyskuje się dla rozkładu +COS o podniesieniu $B > A$ i o optymalnie dobranych parametrach A i B [5–7]. Różnice są dwukrotnie mniejsze niż dla najlepszego COS^2. Przebieg takiego rozkładu jest podobny do obciętej krzywej Gaussa i też nie jest wygodny w użyciu.

Rozkład normalny o odchyleniu standardowym σ_i, może być aproksymowany przez odpowiedni rozkład COS^2 o $A_i = B_i$ i $X_i = 1/(2A_i)$ w takich przypadkach, jak w kolumnach 2–6 tabeli 3. Ze wzoru (8) wynika:

$$A_i \sigma_i = A \quad i \quad X_i \sigma = X \sigma_i \tag{9}$$

Trzy rozkłady PDF funkcji COS^2 aproksymujących funkcje Gaussa o odchyleniach standardowych $\sigma_i = (0{,}5;\ 1;\ 2)$ podano na rys. 4.



Rys. 4. Rozkłady Normalne o zerowej wartości średniej, $\sigma = (1/2,\ 1,\ 2)$ i odpowiadające im rozkłady COS^2 wyznaczone metodą najmniejszych kwadratów

Fig. 4. Normal distribution functions of mean value equal zero and $\sigma = (1/2,\ 1,\ 2)$ and relevant the best fitting COS^2 distributions obtained by MSM method

Do wyznaczenia rozszerzonej niepewności pomiarów konieczne jest wykonanie splotu rozkładów opisujących poszczególne składowe niepewności wypadkowej. Tylko splot samych funkcji Gaussa pozostaje funkcją Gaussa. Jego niepewność rozszerzoną można wyznaczyć na podstawie wypadkowego odchylenia standardowego. Odchylenie standardowe σ_* splotu dwu nieskorelowanych rozkładów wynosi

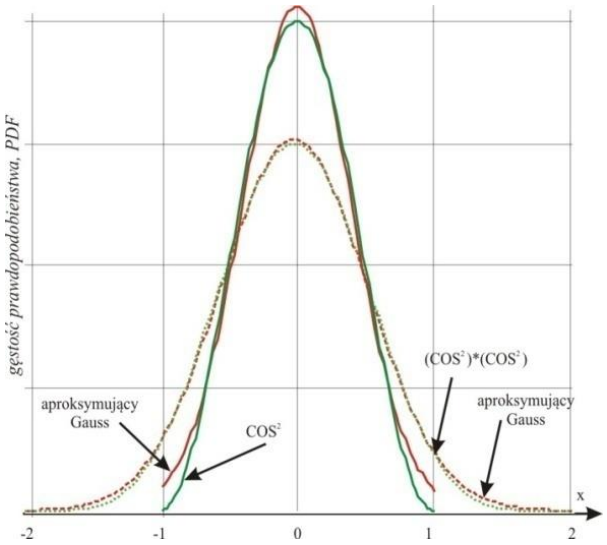
$$\sigma_* = \sqrt{\sigma_1^2 + \sigma_2^2} \tag{10}$$

gdzie: σ_*, σ_1, σ_2 – odchylenia standardowe obydwu składników

Dla splotu dwu rozkładów COS^2 otrzymuje się

$$\sigma_{2(+COS)*} = \sqrt{\sigma^2_{COS^2 1} + \sigma^2_{COS^2 2}} = 0.3615\sqrt{X^2_{COS^2 1} + X^2_{COS^2 2}} \tag{11}$$

Jako przykład, na rys. 5 przedstawiono splot dwu jednakowych rozkładów COS^2 o funkcji $f(x) = A(1+\cos 2\pi xA)$ i parametrach: $-X \le x \le +X$, $A = B = 0{,}2$, X = 2,5, $\sigma_{COS^2} = 0{,}904$, $\sigma = 0{,}997$, kurtosis[1] E = 2,4 (jak dla PDF rozkładu trójkątnego).



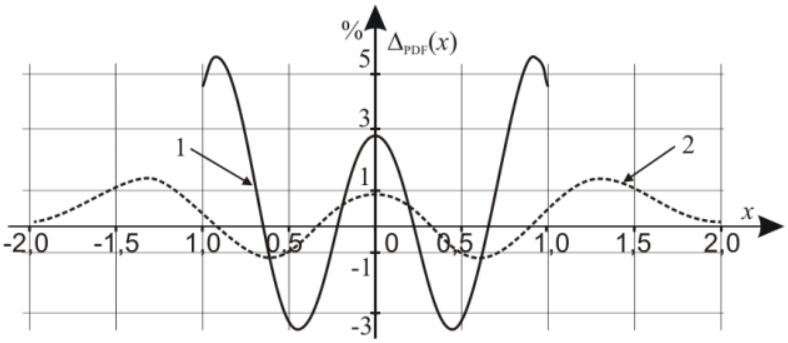
Rys. 5. Rozkłady prawdopodobieństwa funkcji: COS^2 i splotu dwu jednakowych funkcji COS^2 oraz najbardziej zbliżonych do nich rozkładów Gaussa (linie przerywane)

Fig. 5. PDF-s of: COS^2 and convolution of two the same COS^2; and curves of their best fitting Gauss distributions (dotted line)

Taki splot ma zakres $4X$ = 10, SD: $\sigma_{2(COS^2)*} = 1{,}28$, kurtozę E = 2,71 i odchylenie standardowe $\sigma_{BestGauss} = 1{,}33$. Splot dwu funkcji COS^2 bardziej zbliża się do rozkładu normalnego niż pojedyncza funkcja COS^2, jako przejaw działania centralnego twierdzenia granicznego.

Z rysunku 6 wynika, że odchylenie standardowe różnicy tego splotu i odpowiadającego mu rozkładu normalnego jest dwukrotnie mniejsze niż dla pojedynczego COS^2, tj. otrzymuje się odpowiednio $\overline{\Delta}_{PDF} = (0{,}11;\ 0{,}22)\%$.

[1] Kurtosis E (eksces) jest stosunkiem momentu centralnego 4. rzędu i standardowego odchylenia rozkładu i ocenia smukłość względem rozkładu normalnego o E = 3.



Rys. 6. Różnice Δ_{PDF} rozkładów Gaussa i COS^2 – 1 oraz ich splotów – 2
Fig. 6. Δ_{PDF} differences of Gauss and COS^2 – 1, and their convolutions – 2

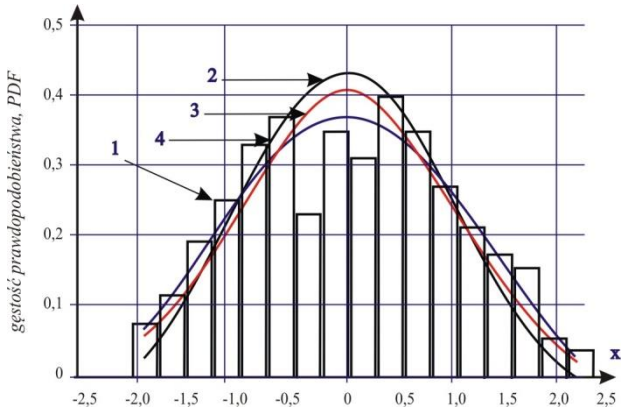
Przykład 1.
Oszacowanie niepewności próbki według rozkładów normalnego i COS²

Z wygenerowanej cyfrowo populacji pobrano losowo próbkę o n = 200 nieskorelowanych wartościach x_i symulujących niezależne od siebie statystycznie obserwacje pomiarowe. Przy zastosowaniu dwu rozkładów – normalnego i COS^2, wyznaczono wynik pomiaru w postaci estymatora wartość menzurandu i jej niepewności rozszerzonej $U(k_P)$ o określonym prawdopodobieństwie objęcia P dla przypadku, gdy niepewność u_B jest pomijalna.

Splot rozkładu COS^2 z różnymi innymi PDF daje wyniki bardzo zbliżone do splotu tych samych rozkładów z funkcją Gaussa. Najlepszym estymatorem dla obu modeli rozkładów jest wartość średnia $\bar{x}$ próbki. Jej niepewności – standardowa i rozszerzone, będą nieco różniły się dla obu rozkładów. Przy pomijalnej niepewności u_B wyznacza się je statystycznie i tylko na podstawie danych eksperymentalnych. Dla rozkładu Gaussa stosuje się zalecenia Przewodnika GUM o szacowaniu niepewności u_A, zaś dla rozkładu COS^2 – opisane powyżej zależności (4), (7)–(9) i wykorzystuje się wartości funkcji z tabel 1 i 2. Oba te modele można zastosować do wyznaczania niepewności pomiaru, gdyż na poziomie istotności α = 0,05 spełniają zarówno test Kołmogorowa, jak i χ^2.

Na rysunku 7 przedstawiono: 1 – histogram o j = 17 przedziałach o wysokościach h_j jako stosunkach liczb zawartych w nich odchyleń od wartości średniej próbki do liczby n wszystkich danych n_i; 2 – funkcję Gaussa o odchyleniu standardowym próbki $S(x_i)$ = 0,978; 3 – funkcję COS^2 o pół-zakresie X_0 = 2,31 jako odległości między wartością średnią $\bar{x}$ = 0 i najdalszym od niej wynikiem x_{iMAX}; 4 – funkcję COS^2 o X_2 = 2,71 wyznaczonym wg (9) z odchylenia standardowego próbki $S(x_i)$. Odległość między krańcowymi odchyleniami $x_{i\,MAX} - x_{i\,MIN}$ próbki wynosi 4,22.

Dla wartości średniej $\bar{x}$ próbki w tabeli 3 podano jej niepewności rozszerzone U (przy $u_B << u_A$) o kilku poziomach objęcia k_P (dla prawdopodobieństwa P) wyznaczone dla rozkładów Gaussa i COS^2. Stopień przybliżenia każdego z rozkładów do danych próbki oceniono popularnym testem zgodności χ^2.



Rys. 7. Rozkłady dla analizowanych symulowanych danych pomiarowych: 1 – histogram o 17 przedziałach wysokości h_i, 2 – jej PDF Gaussa i dwa PDF-y rozkładu COS^2: 3 – o pół-zakresie X_0 do najdalszej obserwacji x_n, 4 – o zakresie $2X_2$ dla SD próbki

Fig. 7. Distributions of analyzed simulated data: 1 – histogram of 17 bins of h_i, 2 – their Gauss PDF and two PDFs of distribution COS^2: 3 – of half range X_0 to the most distant observation x_n, 4 – of half range $2X_2$ for SD of sample

Dla próbki o n = 200 pomiarach i histogramu o 17 przedziałach otrzymuje się:

$$\chi^2 = 200\sum_{j=1}^{17}\frac{\left(h_j - p_j\right)^2}{p_j} \tag{12}$$

gdzie wysokości h_j przedziałów histogramu są ułamkami n_j/n zawartych w tych przedziałach danych eksperymentalnych i danych próbki.

Tab. 3. Niepewności rozszerzone U o poziomach objęcia k_P dla wartości średniej próbki
Tab. 3. Expanded uncertainties U of confidence levels k_P for mean value of the sample

Poziom ufności P	COS^2(0, X_i)			$N(0,\overline{\sigma})$
	Współczynnik rozszerzenia k_{PC}	X_1 = 2,31	X_2 = 2,71	$S(x_i)$ = 0,978
		Niepewność rozszerzona ±U		
0,500	0,265	0,043	0,051	0,047
0,683	0,385	0,063	0,074	0,069 (u_A)
0,900	0,596	0,097	0,114	0,114
0,950	0,683	0,112	0,131	0,136
0,990	0,816	0,133	0,156	0,178
0,997	0,878	0,143	0,168	0,205
1	1	0,163	0,192	∞
Kryterium χ2	$\chi^2_{14,0,05} = 23,7$	19,5	6,05	7,98
$u_A(U) = \sqrt{\sum\left(f(x_j) - h_j\right)^2} / \sqrt{n-1}$		0,071	0,012	0,016

Przy połowie rozstępu X_0 = 2,11, tj. odległości między wartością średnią i dalszym z krańców histogramu, test χ^2 dla COS^2 był negatywny. Aby go spełnić rozsze-

rzono X_1 do 2,31. Jeszcze lepszy wynik tego testu dla COS² niż dla Gaussa uzyskano dla większego X_2 = 2,71. Odpowiada on odchyleniu standardowemu próbki $S(x_i)$ = 0,978.

Dla powyższych trzech rozkładów wyznaczono też różnice wysokości h_{xj} słupków histogramu i prawdopodobieństw dla środkowych ich punktów x_j oraz odchylenia standardowe tych różnic. Podano je jako $u_A(U)$ w ostatnim wierszu tabeli 3.

Wynikiem pomiaru jest wartość średnia próbki wraz z oceną jej dokładności przez niepewność rozszerzoną U o wymaganym prawdopodobieństwie P, tj.:

$$x = \bar{x} \pm U$$

gdzie: U – niepewność rozszerzona o prawdopodobieństwie objęcia P.

Przy założeniu, że wpływ niepewności typu B w tych rozważaniach jest pomijalny, otrzymano następujące wyniki:

dla modelu Gaussa dla $P = 0{,}997$ **:**

$$x = \bar{x} \pm k_{PG} \frac{S(x_i)}{\sqrt{n}} = \bar{x} \pm 0{,}205$$

dla modelu COS² dla $P = 0{,}997$ **:**

$$x = \bar{x} \pm k_{PC} \frac{X_1}{\sqrt{n}} = \bar{x} \pm 0{,}143 \qquad \text{lub} \qquad x = \bar{x} \pm k_{PC} \frac{X_2}{\sqrt{n}} = \bar{x} \pm 0{,}192$$

Dla rozkładu danych eksperymentalnych rozpatrywanej próbki oba warianty rozkładu COS² są lepszymi modelami niż rozkład Gaussa. Np. niepewność rozszerzona o poziomie prawdopodobieństwa 0,997 dla modelu Gaussa jest o 43,4% większa niż dla COS² o połowie rozstępu X_1 = 2,31, lub o 22% – przy X_2 = 2,71. Ponadto dla modeli COS² można podać szerokość przedziału, wewnątrz którego wartość średnia $\bar{x}$ znajduje się z prawdopodobieństwem 1! Jest on równy zakresowi $2X_i$ tego rozkładu. Dla tego przykładu wynosi on ±(2,36 lub 2,78) u_A wg modelu Gaussa.

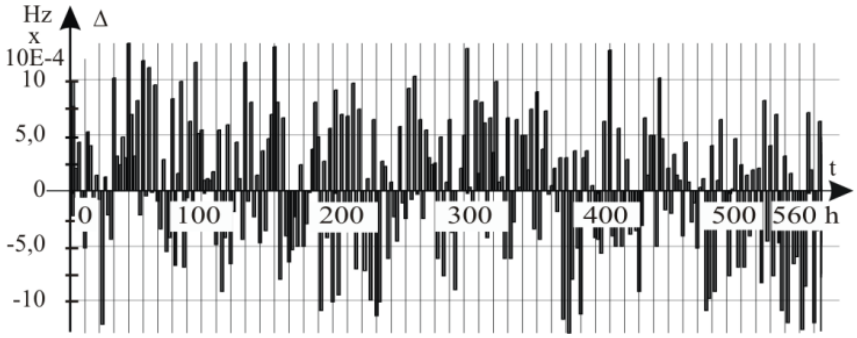
Test χ^2 jest właściwy do stosowania dla rozkładów o ograniczonym zakresie PDF i stopniowo narastających brzegach, tj. takich jak trapez i COS². Gęstości prawdopodobieństwa p_i punktów bliskich ich krańcom są małe, a ich składniki w teście χ^2 wg (12), wskutek dużych współczynników wagi $1/p_i$, zbyt dominujące. Lepiej jest stosować inne kryteria do oceny rozbieżności modelu rozkładu i danych eksperymentalnych, np. test Kołmogorowa-Smirnowa, Lilljeforsa i Shapiro-Wilka.

Przykład 2.
Ocena dokładności synchronizacji częstotliwości w telekomunikacji

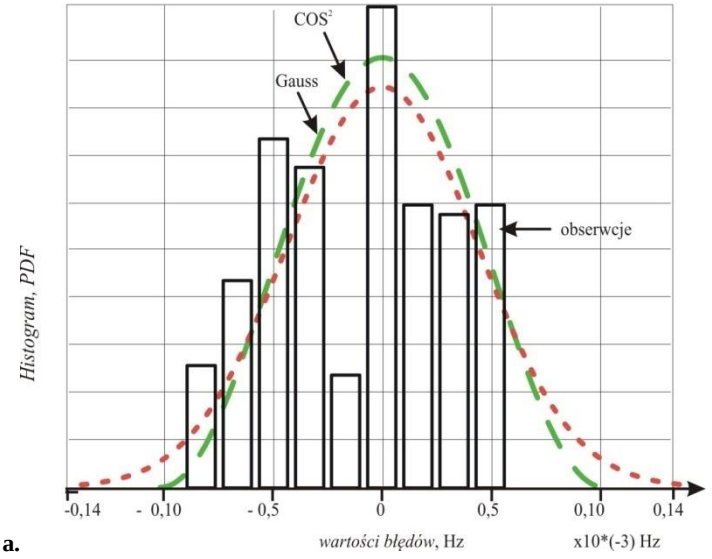
Przykład ten pokazuje przypadek już spotkany w praktyce metrologicznej, gdy rozkład COS² lepiej odwzorowywał rzeczywiste obserwacje pomiarowe niż rozkład Normalny. Są to wyniki badań stabilności częstotliwości generatora wzorcowego dużej częstotliwości stosowanego w telekomunikacji, uzyskane z Ukrainy [7].

Skuteczna transmisja sygnałów cyfrowych w sieci telekomunikacyjnej wymaga bardzo dokładnych zegarów. Przykład ten dotyczy pomiarów kalibracji wzorca

czasu w urządzeniu telekomunikacyjnym zapewniającym ich synchronizację. Badaniom stabilności podlegały zegary kwarcowe. Poziom odniesienia zapewniał pierwotny wzorzec czasu z oscylatorem cezowym o dokładności $5 \cdot 10^{-12}$. Pomiary wykonywano w odstępach co dwie godziny. Odchyłki od wartości średniej 2,048 MHz przedstawiono na rys. 8. Z 281 obserwacji wykluczono dwie jako outliery, gdyż miały wartości nadmiernie odstające. Pozostałe 279 poddano dalszemu przetwarzaniu. W zebranych wynikach obserwacji nie wykryto żadnych wpływów periodycznych lub aperiodycznych i nie stwierdzono też autokorelacji wyników.

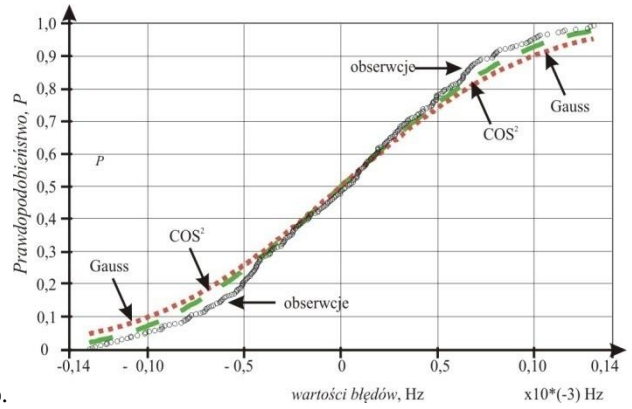


Rys. 8. Błędy pozorne synchronizacji częstotliwości generatorów w funkcji czasu
Fig. 8. Apparent errors of generators' frequency vs. time



a.

Rys. 9. a. Histogram błędów pozornych oraz odpowiadający im model Gaussa i COS^2 wyznaczone metodą najmniejszych kwadratów
Fig. 9. a. Histogram of apparent errors and relevant Gaussian and COS^2 PDF models calculated based on minimal square error method



b.

Rys. 9. b. Skumulowany histogram błędów pozornych i odpowiadający im model Gaussa oraz COS^2

Fig. 9. b. Cumulative histogram of apparent errors and its Gauss and COS^2 models

Na rys. 9a przedstawiono histogram oraz aproksymujące go dwa modele funkcji gęstości prawdopodobieństwa PDF – konwencjonalny (funkcja Gaussa) i COS^2. Na rys. 9b podano ich skumulowane funkcje gęstości prawdopodobieństwa. Oba te modele można zastosować do wyznaczania niepewności pomiaru, gdyż na poziomie istotności $\alpha = 0{,}05$ spełniają zarówno test Kołmogorowa jak i χ^2.

W tabeli 4 zestawiono wartości niepewności U_{COS^2} i U_{Gauss} obliczone na podstawie modelu COS^2 i Gaussa, wraz z ich współczynnikami rozszerzenia dla poziomów ufności p w pierwszej kolumnie. Procentową różnicę między niepewnościami modelu COS^2 w stosunku do modelu Gaussa, odniesioną do U_{COS^2} podano w kolumnie 6. Stosując moduł rozpięcia jako zakres cosinusoidy otrzymano węższy przedział niepewności. Różnica jest wyraźnie większa dla wyższych poziomów ufności. Jej wartość została obliczona z zależności:

$$\Delta = \frac{U_{Gauss} - U_{COS^2}}{U_{COS^2}} 100 \quad [\%]$$

Tab. 4. Porównanie niepewności rozszerzonej częstotliwości generatora wzorcowego przy stosowaniu modelu Gaussa i COS^2

Tab. 4. Uncertainty comparison of standard generator frequency obtained for Gaussian and COS^2 models

p	k_{pcos}^2	U_{COS^2} [Hz]	k_{pGauss}	U_{Gauss} [Hz]	Δ [%]
0,5	0,265	2,08E-06	0,675	2,38E-06	15
0,683	0,385	3,02E-06	1	3,53E-06	17
0,9	0,596	4,67E-06	1,65	5,80E-06	24
0,95	0,683	5,35E-06	1,96	6,92E-06	29
0,99	0,816	6,40E-06	2,58	9,09E-06	41
0,997	0,878	6,88E-06	2,97	1,05E-05	52
1	1	7,84E-06			

Porównanie estymatorów rozkładu COS²

Wyznaczono odchylenia standardowe SD kilku estymatorów jednoelementowych rozkładu $\cos^2$ oraz efektywnego estymatora dwuelementowego (o najmniejszym SD) opisanego równaniem:

$$X_{eff}^{\cos} = 0{,}85 \cdot \bar{X} + 0{,}15 \cdot q_{V/2}.$$

Obliczenia symulacyjne wykonano metodą Monte Carlo dla następujących parametrów: amplituda rozkładu $A = 0{,}2$, liczebność próbki $n = 500$ i liczba eksperymentów MC 5000. Rezultaty obliczeń zestawiono w tabeli 5. Najlepszy jest estymator dwuelementowy, ale SD jego i średniej arytmetycznej różnią się nieznacznie.

Tab. 5. Porównanie estymatorów rozkładu $\cos^2$
Tab. 5. Standard deviations SD of few estimators of distribution $\cos^2$

Estymator	Odchylenie standardowe SD
Średnia arytmetyczna	0,101
Środek rozstępu	0,200
Mediana	0,142
$X_{eff}^{\cos}$	0,096

Wnioski i podsumowanie

Ograniczony zakres rozrzutu wartości danych otrzymywanych w badaniach eksperymentalnych i statystyce stosowanej uzasadnia opisywanie ich modelami odpowiednio dobranych rozkładów niegaussowskich. Można wówczas dokładniej oszacować składową statystyczną niepewności wyniku wartości mierzonej niż metodą statystyczną typu A wg Przewodnika GUM [1]. W pracy podano wyniki badań zastosowania niekonwencjonalnego modelu rozkładu gęstości prawdopodobieństwa (PDF) w postaci jednego okresu podniesionej cosinusoidy, oznaczonej jako rozkład +COS. Oto wnioski:

- Najprostszą postać rozkładu +COS uzyska się w przypadku szczególnym, gdy podniesienie cosinusoidy równa się jej amplitudzie, czyli dla funkcji $\cos^2(\cdot)$. Granicznymi wartościami rozstępu rozkładu są wówczas punkty styczności jej sąsiednich minimów z osią x. Ta postać rozkładu, oznaczona symbolem $\cos^2$, pozwala uniknąć niedogodności opisu danych rzeczywistych przez wykładniczą funkcję Gaussa przebiegającą w nieskończonym przedziale wartości $\pm\infty$.
- Z badań symulacyjnych wynika, że dla rozkładu $\cos^2$ można tak zoptymalizować jego zakres $2X$ (czyli rozstęp rozkładu) i związaną z nim jednoznacznie amplitudę A cosinusoidy, że różnica w stosunku do rozkładu normalnego w przedziale $\pm 2{,}5\sigma$ nie przekroczy $\pm 2{,}1\%$. Dla mniejszych przedziałów dokładność przybliżania obu rozkładów jest nawet nieco większa [5–7].
- Dla cosinusoidy podniesionej o nieco więcej niż jej amplituda, tj. o $B > A$, przybliżenie jest jeszcze dwukrotnie dokładniejsze, ale taka funkcja, podobnie

jak obcięty na krańcach rozkład normalny, ma pionowe odcinki ograniczające i jej opis analityczny jest zbyt skomplikowany.
- Sploty rozkładów COS^2 i normalnego z innymi rozkładami są podobne, a zależności – dość proste.
- Model COS^2 łatwo wykorzystać w obliczeniach niepewności. Dwa parametry: wartość średnią i rozstęp, dobiera się odpowiednio do danych eksperymentalnych. Niewymagane jest tworzenie histogramu oraz wyznaczanie odchylenia standardowego ze wszystkich pomiarów wzorem o postaci pierwiastka, jak dla rozkładu Gaussa. Można też nie stosować komputera i rozbudowanych tablic.
- Proces szacowania niepewności rozszerzonej przebiega następująco: należy określić wstępnie zakres $2X_i$ funkcji COS^2, nieco szerszy niż wg krańcowych wartości rozrzutu pomiarów i sprawdzić czy przy tej wartości X_i spełnione jest wybrane kryterium dla różnic rozkładu COS^2 i danych eksperymentalnych. Jeśli tak, to mnożąc wartość X_i przez odpowiedni współczynnik z tabel 2 lub 4 uzyskuje się wartość niepewności rozszerzonej o zadanym prawdopodobieństwie, w tym też równym 1 dla $2X_i$, gdyż przedziały o określonym prawdopodobieństwie są powiązane stałymi współczynnikami z okresem funkcji $\cos^2(\cdot)$.
- Przydatność rozkładu COS^2 w praktyce zilustrowano Przykładem 1 wyznaczenia niepewności dla symulowanej próbki danych zaczerpniętych z populacji losowej. Uzyskano dokładniejsze oszacowanie niepewności typu A niż według zaleceń przewodnika GUM opartych na funkcji Gaussa.
- W Przykładzie 2 zaczerpniętym z praktyki metrologicznej otrzymano, że rozkład COS^2 lepiej odwzorowuje rzeczywiste obserwacje pomiarowe niż rozkład normalny. Były to wyniki badań stabilności częstotliwości generatora wzorcowego o wysokiej częstotliwości, stosowanego w telekomunikacji.
- Przyjęty w przykładzie 2 model COS^2 spełniał zarówno test Kołmogorowa-Smirnowa jak i χ^2 lepiej niż rozkład normalny. Można więc było stosować go do szacowania niepewności wyniku pomiarów. Otrzymano dla P = 0,95 o kilkanaście % mniejszą niepewność rozszerzoną niż typu A wg GUM opartą na funkcji Gaussa. Uzyskany węższy przedział niepewności, tym bardziej różni się od przedziału wg modelu Gaussa, im wyższy jest poziom ufności. Istnieje też przedział o prawdopodobieństwie 1.
- Rozkład COS^2 jest łatwiejszy do wyznaczania niż normalny. Można go zastosować w programowalnych inteligentnych przyrządach i przetwornikach do pomiaru chwilowych wartości. Umożliwia rozszerzenie ich funkcji o automatyzację wyznaczania na bieżąco wartości średniej lub innego estymatora menzurandu, jego trendu, i parametrów rozrzutu wyników wraz z ciągłą oceną niepewności oraz odpowiednią formą prezentacji [3].

Literatura

1. *Guide to the expression of uncertainty in measurement* (GUM) OIML 2008; Polskie tłumaczenie J. Jaworskiego: Wyrażanie Niepewności Pomiaru. Przewodnik, Główny Urząd Miar, Alfavero 1999.
2. Supplement 1 Propagation of distributions using a Monte Carlo method. OIML G1-101 Ed. 2007.

3. Korczyński M.J., Warsza Z.L., *A New Instrument Enriched by Type A Uncertainty Evaluation*. Proc. of 16th IMEKO TC4 International Symposium in Florence, 22–24 Sept. 2008, paper 1181.
4. Warsza Z.L., Korczynski M.J., *Shifted up cosine function as model of probability distribution* – Sbornik dokładov Miezhdunarodnoj nauczno-prakticzeskoj konferencji "Metrologia 2009", 14–15.04.2009 BelGIM Minsk Bielarus 179–184.
5. Warsza Z.L., Korczyński M.J., *Podniesiona cosinusoida jako model rozkładu gęstości prawdopodobieństwa*. Materiały konferencji: Podstawowe Problemy Metrologii PPM '09, Sucha Beskidzka, Maj 2009, Polska Akademia Nauk, Oddz. w Katowicach, Prace Komisji Metrologii, seria Konferencje nr 14, 116–121.
6. Warsza Z.L., Korczyński M.J., Galovska M., *Shifted up cosine function as model of probability distribution*. Proceedings of IMEKO World Congress Fundamental and Applied Metrology. September 6−11, 2009, Lisbon, Portugal, paper no 530, 2417–2422.
7. Warsza Z.L., Korczynski M.J., *Shifted up cosine function as model of probability distribution.* Proceedings of 19th Symposium Metrology and Metrology Assurance Sept. 9–14, 2009, Sozopol Bulgaria, 42–49.
8. Warsza Z.L., Korczyński M.J., *Shifted up cosine function as unconventional model of probability distribution*. Journal of Automation, Mobile Robotics & Intelligent Systems (JAMRIS), Vol. 4, No. 1, 2010, 49–55.
9. Korczyński M.J., Warsza Z.L., *Podniesiona cosinusoida jako rozkład gęstości prawdopodobieństwa – właściwości i przykłady zastosowań.* „Pomiary Automatyka Kontrola", Vol. 56, Nr 9/2010, 1006–1011.
10. Raab D.H., Green E.H., *A cosine approximation to the normal distribution*, "Psychometrika", Vol. 26, No. 4, 1961, 447–450.

Rozdział 5

Statystyki skośności i kurtozy małych próbek pomiarowych z populacji o rozkładzie normalnym i dwu innych

Streszczenie. Przedstawiono podstawowe właściwości statystyczne rozkładów skośności i kurtozy dla zbioru próbek o określonych małych liczbach elementów. Rozkłady te wyznaczono metodą Monte Carlo. Próbki wielokrotnie pobierano losowo z populacji o rozkładzie normalnym. Znajomość statystyk skośności i kurtozy umożliwia bardziej wiarygodne oszacowanie odchylenia standardowego i niepewności estymatora wartości menzurandu z próbek pomiarowych o małej liczbie obserwacji, gdy znany jest rozstęp populacji rozrzutu ich wartości. Porównano średni moduł i odchylenie standardowe skośności próbek z populacji o rozkładzie normalnym, równomiernym i trójkątnym.

Dane eksperymentalne, w tym wartości obserwacji pomiarowych, są zwykle niesymetrycznie rozproszone losowo i różnie skoncentrowane. Dla przyrządu lub systemu pomiarowego z czujnikami różnych wielkości jest to zbiór odczytów lub wartości sygnału wyjściowego otrzymywanych przez próbkowanie. Nierównomierność rozrzutu dotyczy nie tylko próbek z populacji o niesymetrycznych rozkładach prawdopodobieństwa, ale też i niezbyt dużych próbek z populacji o rozkładzie normalnym i o innych rozkładach symetrycznych. Asymetria takich próbek wzrasta przy zmniejszaniu się liczby ich elementów. Jednak w wielu przypadkach występujących w praktyce dysponuje się tylko małą liczbą obserwacji pomiarowych i z różnych przyczyn nie ma możliwości jej zwiększenia. Ograniczenie to spowodowane jest:

- brakiem większej liczby obiektów badanych, (np. przy walidacji metody stosowanej tylko w kilku akredytowanych laboratoriach),
- dużym kosztem badań lub ograniczonym czasem ich wykonania – brakiem możliwości ponownego przeprowadzenia pomiarów, np. w badaniach odległych terenowo i w medycynie, w badaniach procesów jednorazowych i badaniach niszczących obiekt lub zmieniających nieodwracalnie jego właściwości.

We wszystkich powyższych przypadkach można jedynie wykorzystać otrzymaną małą próbkę danych i dołożyć starań, by przy ich przetwarzaniu jak najlepiej oszacować wynik pomiarów. Dotychczas przy wyznaczaniu tego wyniku, zgodnie z rekomendacją Przewodnika GUM [2], traktuje się próbkę danych tak, jakby pochodziła z populacji o rozkładzie normalnym. Estymatorem wartości menzurandu jest jej wartość średnia, a niepewność typu A, powstająca wskutek rozrzutu danych, opiera się na wariancji próbki. Należałoby zbadać: na ile istotna może być też znajomość innych parametrów statystycznych małych i bardzo małych próbek, w tym jej skośności i kurtozy w przypadkach, gdy rodzaj rozkładu populacji danych nie jest znany. Dalej przedstawiono pierwszy etap tych badań, w którym metodą symulacji Monte Carlo (MC) wyznaczono statystyki skośności i kurtozy małych i bardzo ma-

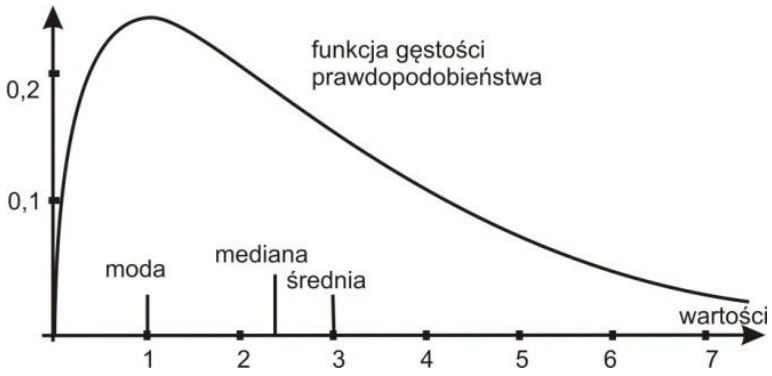
łych próbek pobranych losowo z populacji o rozkładzie normalnym i dla porównania z dwu innych rozkładów – równomiernego i trójkątnego [9, 10].

Współczynniki skośności populacji i próbki

Na rysunku 1 podano przykład rozkładu gęstości prawdopodobieństwa PDF (ang. *probability density function*) o asymetrii prawostronnej (wydłużone prawe ramię). Zaznaczono położenie podstawowych parametrów statystycznych tego rozkładu: mody, mediany i średniej. Asymetrię rozkładu opisują się różnymi współczynnikami skośności. Proste miary skośności (ang. *skewness*) są następujące [1]:

$$A_d = \frac{\mu - d}{\sigma} \quad (1a) \qquad A_m = 3\frac{\mu - m}{\sigma} \quad (1b) \qquad A_Q = \frac{Q_1 + Q_3 - 2m}{2Q} \quad (1c)$$

gdzie: μ – średnia arytmetyczna, m – mediana, d – dominanta (moda), σ – odchylenie standardowe, Q_1, Q_3 – pierwszy i trzeci kwartyl, Q – odchylenie ćwiartkowe.



Rys. 1. Parametry rozkładu o dystrybuancie (PDF) asymetrycznej prawostronnie
Fig. 1. Parameters of the right-asymmetric distribution function (PDF)

Miary (1a–1c) nadają się jedynie do porównywania asymetrii jednego rodzaju rozkładów. Jednolity opis asymetrii różnych rozkładów umożliwia podany przez Pearsona współczynnik skośności γ_1 [1, 3]. Dla populacji X jest on stosunkiem momentu centralnego μ_3 i odchylenia standardowego w trzeciej potędze σ^3 – wzór (2) (tab. 1). Współczynnik γ_1 jest zerem dla rozkładów symetrycznych, jest dodatni dla rozkładów o asymetrii prawostronnej i ujemny dla rozkładów o asymetrii lewostronnej.

Tab. 1. Współczynniki skośności wg Pearsona dla rozkładu i próbki
Tab. 1. Pearson skewness coefficients of distribution and sample

Dla populacji	$\gamma_1 = \mathrm{E}\left[\left(\frac{X-\mu}{\sigma}\right)^3\right] = \frac{\mu_3}{\sigma^3} = \frac{\mathrm{E}\left[(X-\mu)^3\right]}{(\mathrm{E}\left[(X-\mu)^2\right])^{3/2}}$ (2)
Dla próbki	$g_1 = \frac{m_3}{m_2^{3/2}} = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^3}{\left(\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2\right)^{3/2}}$, (3)

Dla próbki o n elementach x_i z populacji generalnej X współczynnik skośności Pearsona g_1 określa się przez jej momenty centralne $\bar{x}$, $m_2 = s^2$ i m_3 –wzór (3) (tab. 1). Są one estymatorami momentów centralnych μ, μ_2, μ_3 populacji generalnej.

Skośność próbek z populacji o rozkładzie normalnym

Współczynnik skośności próbki g_1 wg wzoru (3) jest obciążony [1]. Wprowadzono więc zmodyfikowaną jego postać:

$$g = \frac{\sqrt{n(n-1)}}{n-2} g_1 \tag{4}$$

W statystycznych programach obliczeniowych stosowany jest też współczynnik skośności próbki o innej postaci [3], tj.:

$$g = \frac{n\sum_{i=1}^{n}(x_i - \bar{x})^3}{(n-1)(n-2)s^3} \tag{5}$$

oraz standaryzowany współczynnik skośności [3]

$$SSKE = g\sqrt{\frac{n}{6}} \, . \tag{6}$$

Dla $n > 150$ oraz populacji symetrycznych ma on rozkład normalny.

Różnice wartości między różnymi postaciami współczynnika skośności nie są duże i występują dla bardzo małych próbek.

Wariancja współczynnika skośności

Wariancja współczynnika skośności g dla próbki o n elementach z populacji o rozkładzie normalnym wynosi [1]:

$$D(g) = \frac{6n(n-1)}{(n-2)(n+1)(n+3)} \tag{7}$$

Znana jest nieco inna zależność podana przez Smirnova [4]

$$D(g) = \frac{6(n-2)}{(n+1)(n+3)} = \frac{6}{n}\left[1 - \frac{12}{2n+7} + O\left(\frac{1}{n^3}\right)\right] \tag{8}$$

gdzie: $O(\cdot)$ – człon resztkowy $1/n^3$.

Zależność (8) dotyczy próbek o średniej liczebności, począwszy od $n > 25$ elementów. Przy dużej liczbie n człon resztkowy $O(\cdot)$ staje się pomijalny i dla próbek o $n>150$ wariancja $D(g) \rightarrow 6/n$.

Kurtoza małych próbek

Kurtoza jest miarą spłaszczenia (smukłości) rozkładu. Kurtoza K próbki informuje o stopniu koncentracji jej danych. Oblicza się ją na podstawie stosunku wartości momentów m_4 i $m_2 = s^2$ próbki, tj.:

$$Kurtoza = \frac{m_4}{m_2^2} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^4}{(n-1)\,s^4}$$

Do porównywania innych rozkładów z rozkładem normalnym (Gaussa) o kurtozie 3, stosuje się współczynnik ekscesu kurtozy $K = Kurtoza - 3$. Dla próbek, gdzie $n \geq 4$ w [3] podano wzór dla K

$$K = \frac{n(n+1)\sum_{i=1}^{n}(x_i - \bar{x})^4}{(n-1)(n-2)(n-3)s^4} - \frac{3(n-1)^2}{(n-2)(n-3)s^4}$$

oraz zdefiniowano standaryzowany współczynnik kurtozy:

$$SK(n) = K\left(\frac{24}{n}\right)^{-\frac{1}{2}}$$

Modelowanie skośności i kurtozy małych próbek metodą Monte Carlo

W literaturze nie znaleziono szczegółowych informacji o rozkładach skośności g oraz ekscesu kurtozy K dla małych próbek o liczbie elementów: $3 < n < 25$. Rozkłady te wyznaczono więc metodą symulacji Monte Carlo i przeanalizowano ich właściwości statystyczne dla różnych liczb elementów n próbki [9–11].

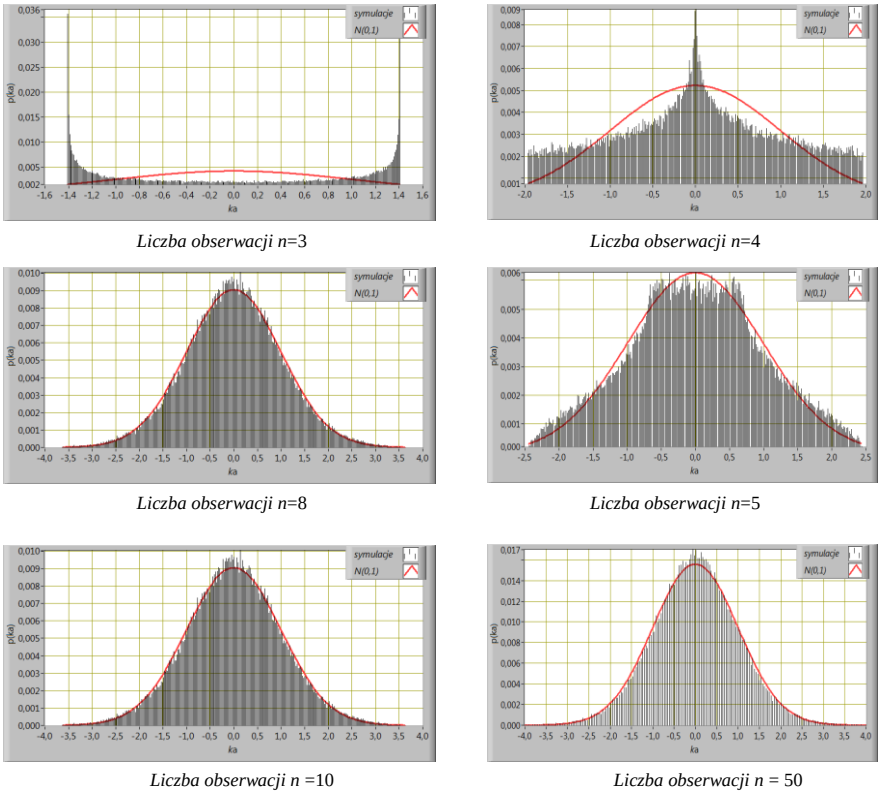
Dalej zostanie dokładnie omówiony eksperyment symulacyjny przeprowadzony dla współczynnika skośności Pearsona. Z numerycznego modelu badanej populacji X pobierano wielokrotnie losowo próbki o określonej liczbie elementów n i dla każdej z tych próbek obliczono współczynnik skośności g wg wzoru wynikającego z (3)

$$g = \frac{m_3}{s^3} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^3}{n\left(\sqrt{\sum_{i=1}^{n}\frac{(x_i - \bar{x})^2}{n-1}}\right)^3} \tag{9}$$

gdzie: m_3 – moment centralny trzeciego rzędu dla próbki, s – jej nieobciążone odchylenie standardowe.

W pierwszej kolejności badania dotyczyły próbek z populacji generalnej o rozkładzie normalnym. Wybrano losowo serie po N = 100 000 próbek o liczbie

elementów n = (3, 4, 5, ..., 50). Dla każdego n wyznaczono histogramy współczynnika skośności g ze wzoru (9). Niektóre z histogramów przedstawiono na rys. 2.



Liczba obserwacji n=3 *Liczba obserwacji n*=4

Liczba obserwacji n=8 *Liczba obserwacji n*=5

Liczba obserwacji n =10 *Liczba obserwacji n* = 50

Rys. 2. Przykłady rozkładów współczynnika skośności g dla próbek o małej liczbie elementów n

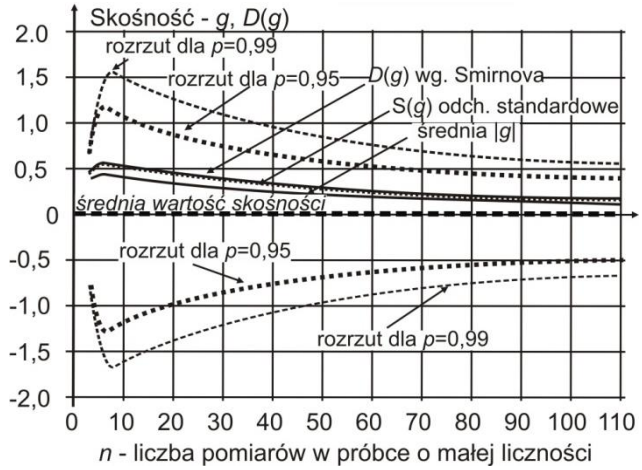
Fig. 2. Examples of the skewness coefficient g distributions of samples with low number of elements n

Wraz ze wzrostem liczby elementów n próbki, kształt histogramu współczynnika skośności próbki g stopniowo zbliża się do funkcji Gaussa. Na histogramach przedstawiono też rozkłady gęstości prawdopodobieństwa ich najlepszych rozkładów Gaussa.

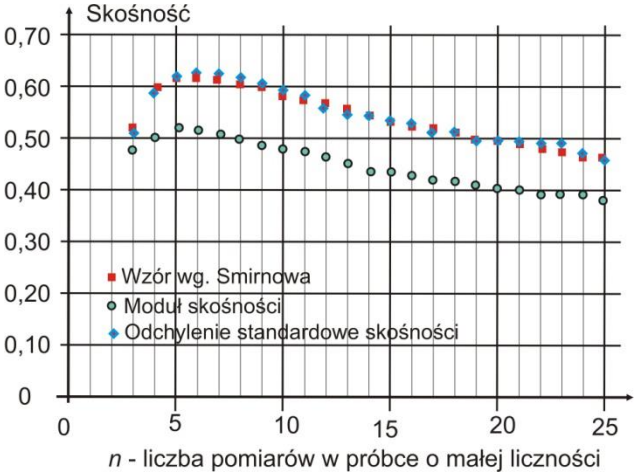
Ponadto dla zbioru próbek o liczbach elementów $n \geq 3$ obliczono odchylenia standardowe współczynnika skośności $s(g)$, średnią wartość jego modułu $|\overline{g}|$, wariancję $D|\overline{g}|$ i standardowe odchylenie $\Delta|g|$ od tej wartości oraz współczynnik kurtozy K. Wyniki tych obliczeń, otrzymane przy stosowaniu różnych wzorów, podano w postaci wykresów na kolejnych rysunkach.

Na rysunku 3 przedstawiono przebiegi obliczonych parametrów skośności w funkcji liczby elementów n próbki po wygładzeniu metodą najmniejszych kwa-

dratów oraz odchylenie standardowe $s(g) = \sqrt{D(g)}$, tj. wg wzoru (8) Smirnowa. Zaś rys 4 podaje w powiększeniu zależność średniego modułu skośności $|\overline{g}|$ i jego odchylenia standardowego $\sqrt{D|\overline{g}|}$ dla początkowych liczb elementów *n*. Osiągają one maksimum około *n* = 6 i potem maleją ze wzrostem liczby *n* elementów próbki.

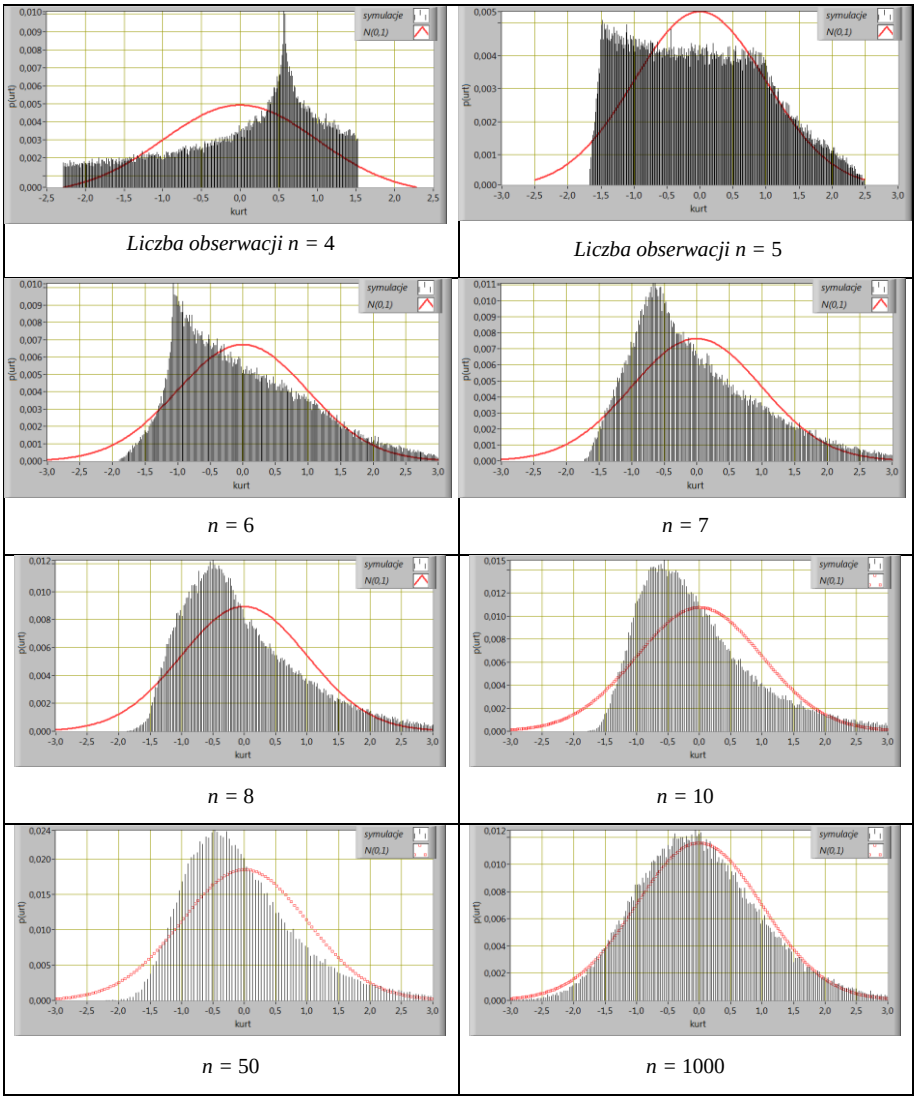


Rys. 3. Parametry skośności *g* w funkcji liczby elementów *n* próbek z populacji Gaussa
Fig. 3. Parameters of skewness *g* of *n*-element samples from Gauss population



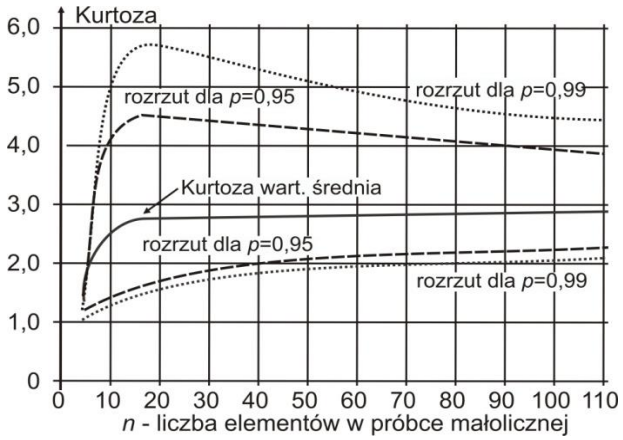
Rys. 4. Parametry skośności *g* dla próbek o bardzo małych *n* < 25
Fig. 4. Parameters of skewness *g* of the very small samples (*n* < 25)

Dla zbiorów próbek pobranych 100 000 razy z populacji o rozkładzie normalnym, korzystając z symulacji metodą Monte Carlo, wyznaczono też rozkłady kurtozy (rys. 5) oraz przebiegi ich podstawowych parametrów statystycznych w funkcji liczby elementów *n* próbki (rys. 6 i rys. 7).

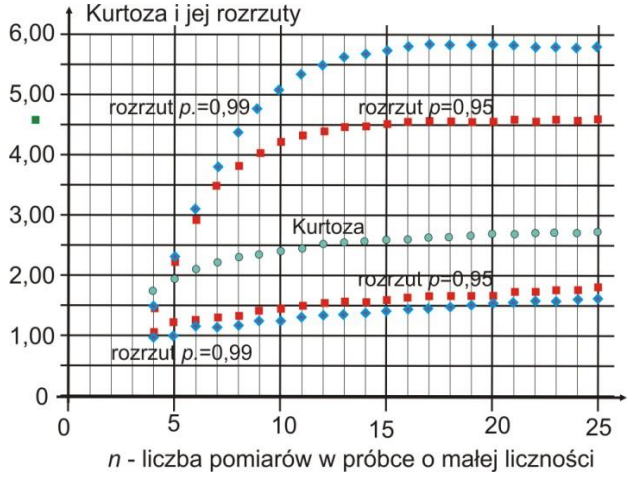


Rys. 5. Histogramy ekscesu kurtozy próbek o kilku licznościach *n* z populacji o rozkładzie normalnym (dla serii $N = 100\ 000$)

Fig. 5. Histograms of kurtosis excess for samples of observations *n* is taken from population of normal distribution ($N = 100\ 000$ trials)



Rys. 6. Parametry statystyczne kurtozy K w funkcji liczby elementów n próbek
Fig. 6. Statistics of kurtosis K as function of number n of sample elements

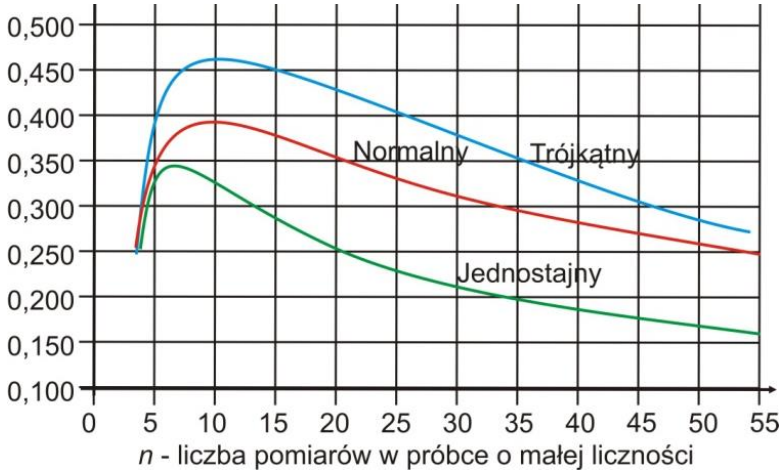


Rys. 7. Parametry statystyczne histogramów kurtozy K dla bardzo małych próbek $N(0, 1)$
Fig. 7. Statistical parameters of the kurtozis K histograms of very small samples of $N(0, 1)$

Porównanie skośności próbek z rozkładu normalnego, jednostajnego i trójkątnego

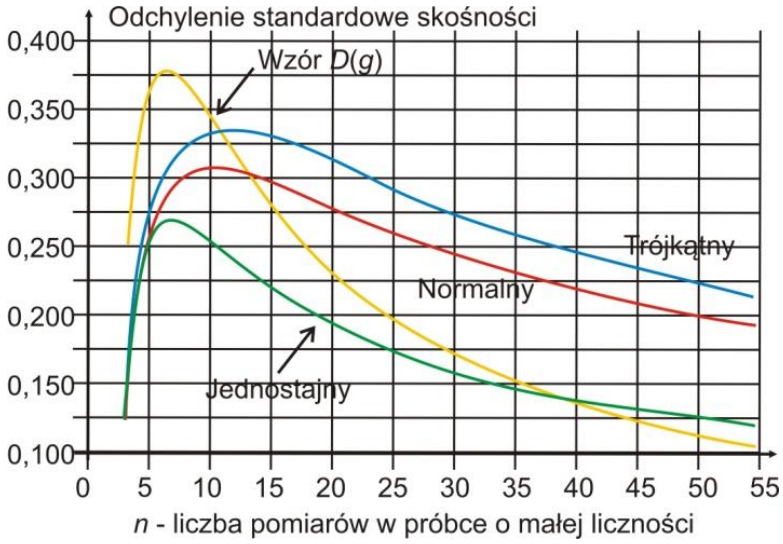
Na rysunku 8, dla $N = 10^6$ próbek z trzech populacji o rozkładach: normalnym, jednostajnym i trójkątnym, podano zależności średniego modułu skośności $|\overline{g}|$ od liczby elementów n próbki, a na rys. 9 – ich odchylenia standardowe SD. Z wykresów wynika, że próbki z populacji o rozkładzie normalnym mają wartości $|\overline{g}|$ pośrednie pomiędzy próbkami z populacji o rozkładzie jednostajnym i trójkątnym. Liczba ob-

serwacji n_{max}, przy której występuje maksimum średniego modułu skośności $|\overline{g}|$, jest największa dla rozkładu trójkątnego.



Rys. 8. Średni moduł skośności $|\overline{g}|$ próbek o n elementach z 3 różnych populacji ($N = 10^6$)

Fig. 8. Mean skewness module of n element samples of 3 different populations ($N = 10^6$)



Rys. 9. Wariancja $D(g)$ i odchylenie standardowe SD skośności g próbek o n elementach z populacji o trzech różnych rozkładach (N = 1 000 000 pobrań)

Fig. 9. Variance and standard deviation STD of the skewness g of n element samples from population of three different PDF (values from N = 1 000 000 trials)

Wnioski i podsumowanie

Dla zbioru małych próbek z populacji o rozkładzie normalnym wartość średnia modułu współczynnika skośności odbiega znacznie od zera, a średnia kurtoza – od wartości 3 dla całej populacji. Moduł skośności osiąga maksimum dla liczby *n* elementów próbki około 6, a kurtoza – dla *n* około 20. Następnie, wraz ze wzrostem *n* oba parametry bardzo powoli maleją, odpowiednio do wartości 0 i 3 tych parametrów dla populacji. Kształty rozkładów obu parametrów dla małych *n* również zdecydowanie odbiegają od krzywej Gaussa.

Dla próbek z populacji o rozkładach niegaussowskich, np. równomiernego, trapezowego i trójkątnego, wartość średnia nie jest najlepszym estymatorem menzurandu [5–7], a przebiegi skośności i kurtozy w funkcji *n* są inne niż rozkładu normalnego.

Skośność i kurtoza próbki nie są dotychczas uwzględniane przy wyznaczaniu zarówno średniej jako estymatora wartości menzurandu, jak i niepewności typu A jako składnika w ocenie dokładności pomiaru. Pewną próbę, aby dla znanych symetrycznych rozkładów niegaussowskich wykorzystywać kurtozę podjęli I. Zakharov i E. Klimova [12]. Do tych samych celów autor z S. Zabolotnym zbadali wstępnie możliwość wykorzystania szeregów stochastycznych i kumulantów [13].

Planowane jest przeprowadzenie badań, których celem jest stwierdzenie, czy rozmnażając dane małych próbek z kilku różnych populacji metodą resamplingu i uwzględniając ich skośność i kurtozę można wyznaczyć lepsze wartości estymatorów dla średniej i jej niepewności rozszerzonej niż wg GUM, w tym stosując metodę Monte Carlo wg Suplementu 1 [8].

Literatura

1. Klonecki W., *Statystyka dla inżynierów*. PWN Warszawa (1999).
2. Guide to the Expression of Uncertainty in Measurement GUM. ISO/IEC/OIML/BIPM, last ed. BIPM JCGM 100 (2008).
3. Dobosz M., *Statystyczna analiza wyników badań*. Wydawnictwo Exit, Warszawa 2004.
4. Bolshev L.N., Smirnov N.V., *Tablicy matematicheskoj statistiki*. Nauka (1983) s. 416 (po rosyjsku).
5. Warsza Z.L., *Jednoelementowe estymatory wartości menzurandu próbek o kilku niegaussowskich rozkładach prawdopodobieństwa – przegląd*. „Pomiary Automatyka Kontrola", nr 1 (2011), 101–104.
6. Warsza Z.L., *Dwuelementowe estymatory wartości menzurandu próbek pomiarowych o trapezowych rozkładach prawdopodobieństwa – przegląd prac*. „Pomiary Automatyka Kontrola", nr 1 (2011), 105–108.
7. Kubisa S., Warsza Z., *Środek rozstępu jako estymator menzurandu dla próbek z populacji o rozkładzie jednostajnym i płasko-normalnym*. „Pomiary Automatyka Kontrola", Vol. 60, nr 6, 2014, 398–401.
8. *Guide to the Expression of Uncertainty in Measurement* (GUM), OIML ed. 2008. Supplement 1: Propagation of distributions using a Monte Carlo method, G 1 – BIPM JCGM 101 (2007).

9. Warsza Z.L., Korczyński J., *Statystyki skośności i kurtozy małych próbek pomiarowych z populacji o rozkładzie normalnym i kilku innych*. „Pomiary Automatyka Kontrola", Vol. 60, nr 12 (2014), 1119–1123.
10. Warsza Z.L., Korczyński M.J., *Statistical properties of skewness and kurtosis of small samples from normal and two other populations*. [In:] Szewczyk R., Zieliński C., Kaliczyńska M. (eds) Progress in Automation, Robotics and Measuring Techniques. Advances in Intelligent Systems and Computing, Vol. 352. Springer, 293–301.
11. Warsza Z.L., Korczyński J., *Statistical properties of skewness and kurtosis of small samples from Normal population.* Systemy Obrobotki Informacji – Information Processing Systems, vipusk 2(127) (2015) Ukraina, 24–28.
12. Zakharov I.P., Klimova E.A., *Application of excess method to obtain reliable estimate of coverage factor*. Systemy Obrobotki Informacji – Information Processing Systems, vipusk 3 (119) (2014) Kharkiv Ukraina, 24–28 (w języku rosyjskim) УДК 006.91
13. Warsza Z.L., Zabolotnii Serhii W., *A polynomial estimation of measurand parameters for samples of non-Gaussian symmetrically distributed data.* [In:] Szewczyk R., Zieliński C., Kaliczyńska M. (eds) Automation 2017. ICA 2017. Advances in Intelligent Systems and Computing, Vol. 550. Springer, 468–480.

Rozdział 6

Środek rozstępu jako estymator menzurandu dla danych z populacji o rozkładzie równomiernym i płasko-normalnym

Streszczenie. Dla próbek o różnej liczności z populacji o rozkładzie równomiernym zbadano metodą symulacji Monte Carlo właściwości statystyczne środka rozpięcia próbki jako estymatora wartości mierzonej i porównano go z wartością średnią. Ma on mniejsze odchylenie standardowe niż średnia arytmetyczna zalecana w Przewodniku GUM. Obliczono dla takich próbek rozkład podobny do rozkładu Studenta i niepewność rozszerzoną. Otrzymano, że dla populacji o rozkładzie płasko-normalnym (splot rozkładu równomiernego i normalnego) przewaga środka rozpięcia szybko maleje przy wzroście udziału rozkładu normalnego.

Poprawny metrologicznie wynik pomiaru powinien zawierać najbardziej prawdopodobną wartość menzurandu (np. mierzonej wielkości) wraz z oceną jej dokładności, wyznaczoną w określony i powszechnie zaakceptowany jednolity sposób. Służby miar stosują postępowanie zalecane w międzynarodowym Przewodniku Wyrażania Niepewności Pomiarów o angielskim akronimie GUM [1, 1a]. W opisie dokładności przyrządów nadal używa się dopuszczalnych błędów granicznych (patrz rozdział 1).

Wraz z powszechną elektronizacją i komputerowym przetwarzaniem danych w nauce, przemyśle i innych dziedzinach wykonuje się wiele pomiarów, z których część nie spełnia założeń GUM, np. gdy rozrzut wartości wielkości mierzonej lub składowe losowe sygnału modeluje się lepiej rozkładami niegaussowskimi. W danych pomiarowych mogą występować zakłócenia losowe ciągłe lub przejściowe – tzw. outliery. Dla większych próbek eliminuje się je z danych za pomocą specjalnych testów. Zaś dla małych próbek w połowie XX wieku powstało nowe narzędzia matematyczne – statystyka odpornościową (ang. *robust statistics*).

Poniżej przedstawione zostaną wyniki badań właściwości statystycznych próbek o rozkładzie równomiernym oraz w różnym stopniu płasko-normalnym. Wykonano je metodą Monte Carlo i opisano we wspólnych pracach z S. Kubisą [12–14].

Zależności podstawowe

Metodę wyznaczania niepewności pomiaru wg Przewodnika GUM [1, 2] nazwano tu **klasyczną.** Rozproszenie wartości x_i danych próbki o n obserwacjach otrzymanych w powtarzanych pomiarach traktuje się w niej tak, jakby pochodziły z populacji generalnej o rozkładzie normalnym i były statystycznie niezależne. Z danych pomiarowych próbki eliminuje się skutki znanych oddziaływań systematycznych i jako estymator wartości menzurandu wyznacza się ich średnią arytmetyczną

$$\bar{x} = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i \tag{1}$$

W szacowaniu dokładności menzurandu wg GUM [1] wykorzystuje się odchylenie standardowe eksperymentalne średniej $\bar{x}$ próbki, oznaczone jako s_{cl} i opisane

$$s_{\mathrm{cl}} = \sqrt{\frac{\sum_{i=1}^{n} (\bar{x} - x_i)^2}{n \cdot (n-1)}} \tag{2}$$

To otrzymywane metodą statystyczną odchylenie s_{cl} nazwano w GUM niepewnością standardową wyznaczaną metodą typu A i oznacza się je zwykle symbolem $u_{\mathrm{A}}(x)$. Ponadto, na podstawie wiedzy obserwatora szacuje się też niepewność standardową $u_{\mathrm{B}}(x)$. Jest to wypadkowe odchylenie standardowe oddziaływań mogących wystąpić w tym typie eksperymentu, ale o wartościach nieznanych, a więc i niemożliwych do usunięcia z danych próbki. Następnie oblicza się złożoną niepewność standardową

$$u_c(x) = \sqrt{u_{\mathrm{A}}^2(x) + u_{\mathrm{B}}^2(x)}$$

i z niej, przy współczynniku rozszerzenia k_p dla poziomu ufności p, lub ze splotu rozkładów u_{A} i u_{B} metodą Monte Carlo [2], znajduje się niepewność rozszerzoną

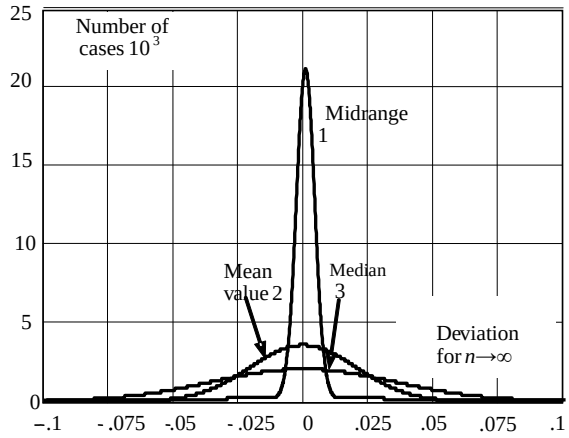
$$U(x) = k_p \cdot u_{\mathrm{c}}(x).$$

Suplement 1 [2] do GUM zaleca metodę MC jako najbardziej uniwersalną, bazującą na elementarnych zależnościach matematycznych, możliwą do stosowania nawet dla silnie nieliniowych funkcji pomiaru, a także w przypadkach nietypowych, np. dla rozkładów asymetrycznych [3]. Jeżeli jednak wiadomo, że obserwacje pochodzą z innej populacji generalnej o znanym rozkładzie prawdopodobieństwa, np. równomiernym, to lepsza może być alternatywna metoda nazywana tu **specjalną**.

Cramer, w znakomitej ponadczasowej monografii [4], wykazał analitycznie, że dla próbek z populacji o rozkładzie równomiernym najlepszym estymatorem menzurandu jest środek rozstępu (ang. *midrange*), który ma najmniejsze odchylenie standardowe. Potwierdzają to przykłady liczbowe w pracach [6, 7] i w Dodatku 1 oraz otrzymane metodą Monte Carlo rozkłady odchyleń trzech estymatorów takich próbek o dużej liczebności $n \to \infty$ – rys. 1 [8, 12]. Wzory dla środka rozstępu x_{v} i jego odchylenia standardowego s_{v}, zaczerpnięte z [5], podano w tabeli 1.

Tab. 1. Parametry statystyczne próbki o równomiernym rozkładzie prawdopodobieństwa
Tab. 1. Statistical parameters of the sample of uniform PDF

Rozstęp próbki (*range of sample*):	$V = x_{i\max} - x_{i\min}$	
Środek rozstępu próbki (*midrange*):	$x_{\mathrm{V}} = \frac{x_{i\max} + x_{i\min}}{2}$	(3)
Odchylenie standardowe środka rozstępu x_{V} (*Standard Deviation of midrange* x_{V}):	$s_{\mathrm{V}} = \frac{V}{\sqrt{2}(n-1)} \cdot \sqrt{\frac{n+1}{n+2}}$	(4)



Rys. 1. Histogramy odchyleń estymatorów menzurandu dla 10^3 próbek losowych z populacji 200×2^{20} o rozkładzie równomiernym [7]: 1 – środek rozpięcia, 2 – wartość średnia; 3 – mediana

Fig. 1. Histograms of deviations of estimators of measurand value for samples from rectangular distribution simulated by 200×2^{20} random numbers: 1 – midrange; 2 – mean value; 3 – median

Dla próbek z populacji o rozkładzie równomiernym najlepszym estymatorem wartości mierzonej jest więc środek rozstępu próbki x_V (3). Estymatorem jego odchylenia standardowego jest wyznaczone eksperymentalnie odchylenie s_V (4).

Poniżej przedstawiono wyniki badań właściwości obu estymatorów w funkcji stopni swobody danych próbki $\nu = n - 1$, otrzymane metodą symulacji Monte Carlo. Obserwacje symulowano liczbami pseudolosowymi z populacji o odchyleniu standardowym $\sigma = 1$ przy liczbie symulacji $M = 2 \times 10^5$ i liczbach stopni swobody:

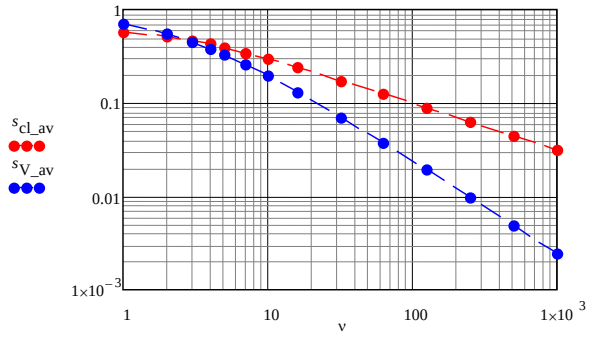
$$\nu = 1, 2, 3, 4, 5, 7, 10, 16, 32, 63, 125, 250, 500, 1000.$$

Badania przeprowadzono dla przypadku, gdy populacja liczb pseudolosowych, z której pochodzą obserwacje, ma czysty rozkład równomierny i dla kilku przypadków, gdy ta populacja ma rozkład równomierny zanieczyszczony normalnym.

Właściwości próbek z populacji o rozkładzie równomiernym

Estymatory menzurandu $\bar{x}$ i x_V obliczone ze wzorów (1) i (3) oraz ich odchylenia standardowe s_{cl} *i* s_V ze wzorów (3) i (4) mają zwykle parami różne wartości dla tych samych danych próbki. W eksperymencie MC, dla każdej z symulacji (ang. *trial*) o numerze m ($m = 1, \ldots, M$) ze wzorów (2) i (4) obliczano wartości $s_{cl\,m}$ i $s_{V\,m}$ dla próbek o n elementach pobranych losowo z *czystego* rozkładu równomiernego o odchyleniu standardowym $\sigma = 1$. Następnie z M symulacji MC obliczono średnie $s_{cl\,av}$ i $s_{V\,av}$. Ich wartości dla różnych liczb stopni swobody $\nu = n - 1$ podano na rys. 2.

Porównanie przebiegu wykresów z rys. 2 sugeruje wyższość estymatorów (3), (4) nad estymatorami klasycznymi (1) i (2), gdyż np. dla $\nu = 10^3$ wartość s_{V_av} jest około 13 razy mniejsza od wartości s_{cl_av}.



Rys. 2. Wartości średnie odchyleń standardowych s_{cl} i s_V w funkcji liczby stopni swobody ν
Fig. 2. Mean values of std. deviations s_{cl}, s_V as functions of number ν of degree of freedom

Powyższy wniosek nie jest w pełni miarodajny metrologicznie, bo bardziej istotne są wartości niepewności rozszerzonych. Przy metodzie klasycznej wg GUM niepewność ta, oznaczona tu jako $U_{cl\,m}$, jest iloczynem odchylenia empirycznego $s_{cl\,m}$ i współczynnika rozszerzenia k_{cl} obliczonego wg rozkładu t-Studenta dla ν stopni swobody i poziomu ufności p:

$$U_{cl\_m} = k_{c1} \cdot s_{cl\_m} \tag{5}$$

Przy niepewności $u_B = 0$ podobnym iloczynem wyraża się niepewność rozszerzona dla środka rozstępu x_V:

$$U_{V\_m} = k_V \cdot s_{V\_m} \tag{6}$$

gdzie k_V jest specjalnym współczynnikiem rozszerzenia dla estymatorów (3) i (4).

Niepewność rozszerzoną U_{Vp} o prawdopodobieństwie p dla środka rozstępu x_v rozkładu równomiernego można wyznaczyć analitycznie. W czasie przeprowadzania badań symulacyjnych MC nie dysponowano jednak odpowiednią zależnością, gdyż Dorozhovets dopiero w pracy [15] z 2016 r. podał wzór:

$$U_{Vp} = k_V \cdot s_V = \left[(1-p)^{-\frac{1}{n-1}} - 1\right] \cdot (n-1) \cdot \sqrt{\frac{n+2}{2 \cdot (n+1)}} \tag{6a}$$

gdzie: $k_V = (1-p)^{-\frac{1}{n-1}} - 1$, $s_V = \frac{V \cdot \sqrt{2}}{n-1} \cdot \sqrt{\frac{n+1}{n+2}}$ – patrz wzór (3)

Dla małych próbek uwzględnia się też rozszerzanie rozkładu wraz ze spadkiem liczebności n. Zmienna t-Studenta dla rozkładu normalnego zdefiniowana jest jako:

$$t \overset{\text{def}}{=} \frac{\bar{x} - E(x)}{s_{cl}} = \frac{\Delta \bar{x}}{s_{cl}} \tag{7}$$

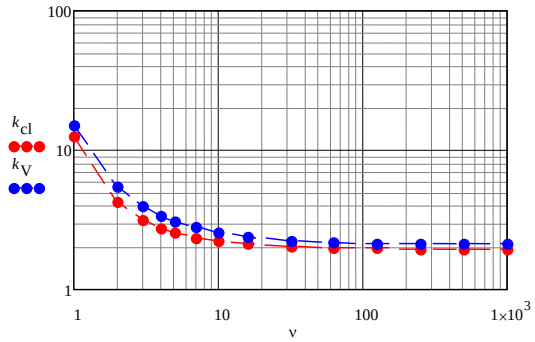
przy czym: $E(x)$ jest oczekiwaną wartością mierzoną o wartości znanej z eksperymentów symulacyjnych, a $\Delta \bar{x}$ – błędem estymaty $\bar{x}$.

Podobnie można zdefiniować zmienną t_V dla rozkładu równomiernego

$$t_V \overset{\text{def}}{=} \frac{x_V - E(x)}{s_V} = \frac{\Delta x_V}{s_V} \tag{8}$$

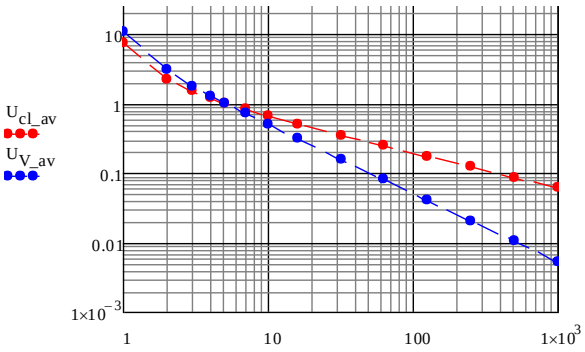
Rozkład prawdopodobieństwa zmiennej t_V nazwiemy rozkładem (statystyką) quasi-Studenta. Rozkład ten można wyznaczyć za pomocą symulacji Monte Carlo i obliczyć odpowiedni współczynnik rozszerzenia k_V.

Wykresy współczynników k_{cl} i k_V dla poziomu ufności p = 95% pokazuje rys. 3.



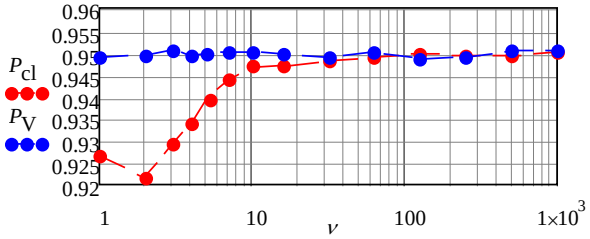
Rys. 3. Współczynniki rozszerzenia w funkcji liczby ν stopni swobody: p = 95%
Fig. 3. Coverage factors as function of number ν degree of freedom: p = 95%

Symulowano niepewności rozszerzone U, ale związane tylko z losowym rozrzutem obserwacji (tj. dla $u_B = 0$) dla prawdopodobieństwa p = 0,95. Dla kolejnych wartości $\nu = n - 1$ obliczono M razy niepewności $U_{cl\,m}$ (5) i $U_{V\,m}$ (6) oraz ich wartości średnie $U_{cl\,av}$ i $U_{V\,av}$ dla zbiorów o liczebności M. Podano je w funkcji liczby stopni swobody ν (rys. 4). Wykresy te potwierdzają wyższość estymatorów (3) oraz (4) nad estymatorami klasycznymi (1), (2) przy liczbach ν większych od około 5. Dla $\nu = 10^3$ wartość U_{V_av} jest około 11 razy mniejsza od U_{cl_av} (rys. 4).



Rys. 4. Średnie wartości niepewności rozszerzonych w funkcji liczby stopni swobody ν
Fig. 4. Average values of expanded uncertainties as functions of number ν degree of freedom

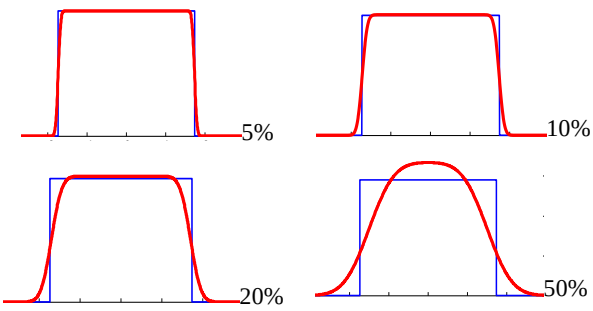
Przedstawione wyniki symulacji metodą Monte Carlo można zweryfikować przez empiryczne sprawdzenie prawdopodobieństwa zdarzenia – czy błędy estymat wartości mierzonej zawierają się w granicach obliczonej niepewności. Dla każdej z n obserwacji należy obliczyć liczbę sukcesów i podzielić ją przez liczbę M symulacji. Takie ilorazy P_{cl}, P_V powinny być bliskie postulowanemu poziomowi ufności p = 95%. Wyniki weryfikacji przedstawiają wykresy na rys. 5. P_V jest zadowalające dla metody o estymatach (3), (4). Natomiast P_{cl} jest zbyt małe dla próbek o małych liczbach ν o estymatach (1), (2) wg GUM.



Rys. 5. Empiryczne sprawdzenie prawdopodobieństw w funkcji liczby ν stopni swobody
Fig. 5. Empirical verification of probability as function of number ν of degree of freedom

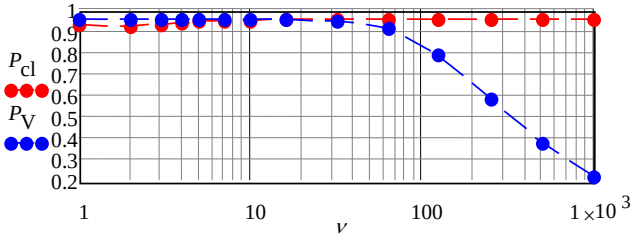
Właściwości statystyczne próbek z populacji o rozkładach płasko-normalnych

We współczesnych zelektronizowanych pomiarach często otrzymuje się próbki wartości sygnału z populacji, której rozkład jest splotem rozkładu równomiernego i normalnego, czyli z populacji o rozkładzie płasko-normalnym [11]. Taką populację można scharakteryzować stopniem udziału λ rozkładu normalnego. Oznacza to, że przy odchyleniem standardowym populacji σ = 1, odchylenie standardowe składnika o rozkładzie normalnym wynosi $\sigma_N = \lambda$, a dla składnika głównego o rozkładzie równomiernym $\sigma_J = \sqrt{1-\sigma_N^2}$. Na rysunku 6 przedstawiono wykresy rozkładu płasko-normalnego o czterech różnych stopniach λ udziału rozkładu normalnego.



Rys. 6. Rozkłady płasko-normalne o różnym λ na tle rozkładu jednostajnego
Fig. 6. Flatten-gaussian distributions of different λ and uniform PDF

Dla $\lambda = 0{,}05 = 5\%$ składnik o rozkładzie równomiernym ma odchylenie standardowe $\sigma_J \approx 0{,}9987 = 99{,}87\%$. Wykresy odchyleń eksperymentalnych i niepewności rozszerzonych o zawartości rozkładu normalnego $\lambda = 5\%$ nie różnią się istotnie od wykresów dla samego rozkładu równomiernego z rys. 2 i rys. 4. Natomiast wykresy prawdopodobieństw z rys. 7 i rys. 4 różnią się znacznie. Zbyt małe wartości prawdopodobieństwa P_V dla większej liczby danych $n = \nu+1$ wskazują na konieczność korekty współczynników rozszerzenia k_V dla rozkładu płasko-normalnego.



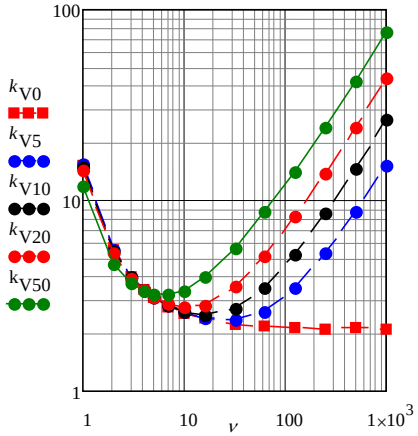
Rys. 7. Empiryczne sprawdzenie prawdopodobieństw Pcl i P_V w funkcji liczby ν
Fig. 7. Empirical verification of probabilities P_{cl} and P_V for degree of freedom numbers ν

Stosując metodę Monte Carlo obliczono współczynniki rozszerzenia k_V dla poziomu ufności $p = 95\%$ i przy wartościach stopnia udziału rozkładu normalnego $\lambda = 0\%, 5\%, 10\%, 20\%, 50\%$. Wyniki obliczeń przedstawiono w tabeli 2 i na rys. 8.

Tab. 2. Współczynnik rozszerzenia k_V ($p = 95\%$) w funkcji liczby stopni swobody ν dla wybranych wartości stopnia udziału rozkładu normalnego λ
Tab. 2. Coverage factor k_V ($p = 95\%$) as function of number ν of degree of freedom for some values of λ

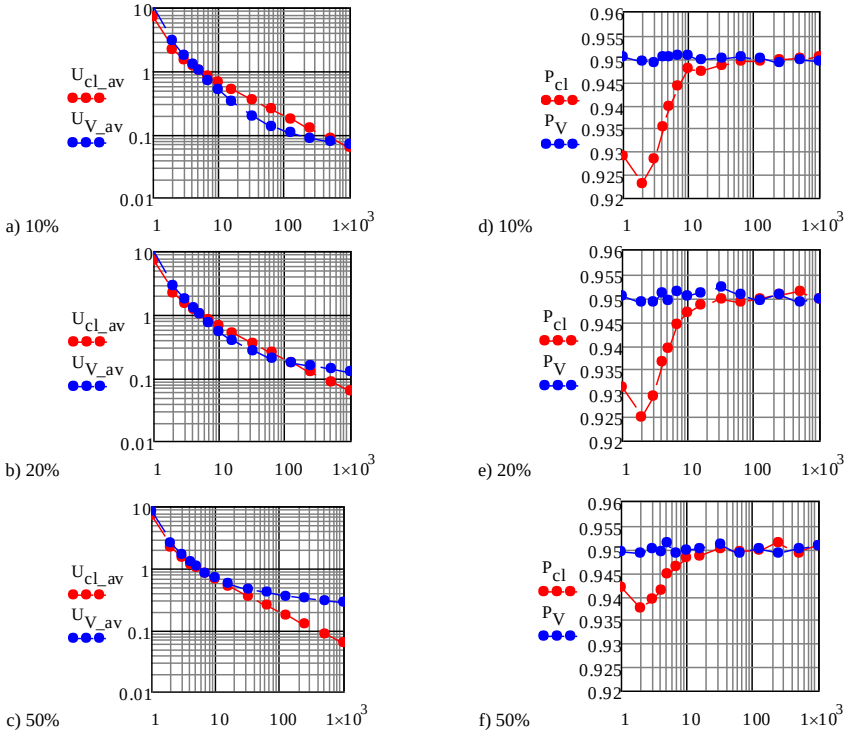
Liczba stopni swobody ν	Stopień λ zwartości rozkładu normalnego w %				
	$\lambda = 0\%$	$\lambda = 5\%$	$\lambda = 10\%$	$\lambda = 20\%$	$\lambda = 50\%$
1	15,31	15,54	15,01	14,40	11,94
2	5,51	5,47	5,42	5,31	4,66
3	3,98	3,95	3,96	3,91	3,67
4	3,42	3,41	3,41	3,40	3,34
5	3,09	3,09	3,10	3,12	3,23
7	2,79	2,79	2,81	2,88	3,19
10	2,57	2,59	2,62	2,74	3,36
16	2,41	2,42	2,52	2,83	3,96
32	2,24	2,39	2,71	3,55	5,63
63	2,18	2,61	3,48	5,13	8,68
125	2,14	3,45	5,23	8,20	14,23
250	2,13	5,31	8,60	13,93	24,33
500	2,14	8,79	14,79	24,30	42,57
1×10^3	2,13	15,39	26,44	43,61	76,29

Wraz ze wzrostem stopnia udziału rozkładu normalnego λ w populacji płasko-normalnej maleje efektywność środka rozpięcia w stosunku do wartości średniej. Ilustrują to wykresy średniej niepewności rozszerzonej i prawdopodobieństwa weryfikującego poprawność jej obliczeń podane na rys. 9.



Rys. 8. Współczynniki rozszerzenia (p = 95%) rozkładu płasko-normalnego o stopniu zawartości rozkładu normalnego λ = 0%, 5%, 10%, 20%, 50% w funkcji liczby stopni swobody ν

Fig. 8. Coverage factors (p = 0.95) of flatten-Gaussian PDF of normal PDF level λ = 0%, 5%, 10%, 20%, 30%, 50% as function of number ν of degree of freedom



Rys. 9. Niepewności $U_{\mathrm{cl\ av}}$, $U_{\mathrm{V\ av}}$ – a, b, c i prawdopodobieństwa weryfikujące P_{cl}, P_{V} – d, e, f dla stopni udziału rozkładu normalnego λ = 10%, 20% i 50%

Fig. 9. Uncertainties $U_{\mathrm{cl\ av}}$, $U_{\mathrm{V\ av}}$ – a, b, c and control probabilities P_{cl}, P_{V} – d, e, f for levels of standard deviation of normal distribution λ = 10%, 20% and 50%.

Uwagi i wnioski

Wyniki obliczeń symulacyjnych MC obarczone są błędami odwrotnie proporcjonalnymi do pierwiastka z liczby symulacji *M*. W szczególności błąd obliczeń prawdopodobieństwa ma rozkład dwumianowy (Bernoulliego) o odchyleniu standardowym:

$$\sigma_P = \sqrt{\frac{P \cdot (1-P)}{M}} \tag{9}$$

Dla $p = 0{,}95$ i $M = 2 \cdot 10^5$ otrzymuje się $\sigma_P \approx 5 \cdot 10^{-4}$. Dla dużych wartości *M* rozkład błędu obliczeń prawdopodobieństwa zbliża się do rozkładu normalnego oraz błąd graniczny można oceniać przedziałem 3-sigmowym, czyli ok. $15 \cdot 10^{-4}$ (0,15%). Uzasadnia to nieregularności w przebiegach wykresów (rys. 5 i 9).

Dla samego rozkładu równomiernego użycie estymatorów (3) i (4) powinno się preferować dla próbek o liczebności $n > 6$, tj. dla liczby ν jest nieco większego niż 5 (rys. 4). Wtedy średnia niepewność rozszerzona U_{V_av} środka rozpięcia, obliczona za pomocą współczynnika rozszerzenia k_V, jest mniejsza od średniej niepewności U_{cl_av} obliczonej klasycznie wg GUM i to tym bardziej, im większe jest *n*.

Dla danych z populacji o rozkładzie równomiernym splecionym nawet z niewielką zawartością składnika o rozkładzie normalnym, np. dla $\lambda = 5\%$, pojawia się konieczność zwiększenia współczynnika rozszerzenia (rys. 8). W obliczeniach założono, że ten dodatkowy składnik ma rozkład normalny. Można się tego samego spodziewać, gdyby miał on np. rozkład równomierny. Jego złożenie ze składnikiem głównym daje rozkład trapezowy, zbliżony do płasko-normalnego.

Wzrastający stopień udziału rozkładu normalnego λ zmniejsza efektywność środka rozpięcia x_V jako estymatora menzurandu (rys. 9 a, b, c). Dla $\lambda = 20\%$ (rys. 9a) efektywność ta występuje tylko dla liczb stopni swobody ν od ok. 5 do ok. 100, a dla $\lambda = 50\%$ (rys. 9c) podejście specjalne jest gorsze od klasycznego w całym zakresie liczb *v*. Stopień udziału 50% nie oznacza, że odchylenie standardowe składnika dodatkowego wynosi 50% całkowitego odchylenia standardowego populacji obserwacji. Odchylenie standardowe składnika głównego ma wtedy wartość $\sqrt{1-0{,}5^2} \approx 0{,}87$ (87%) odchylenia standardowego.

Przy znacznym zmniejszeniu udziału rozkładu równomiernego w rozkładzie płasko-normalnym bardziej efektywna jest metoda klasyczna wg GUM (1), (2). Jednak przy małych liczbach stopni swobody $\nu < 20$ daje ona nieco za niski poziom prawdopodobieństwa weryfikującego P_{cl} (rys. 9 d, e, f).

Warto zaznaczyć, że inne proste rozkłady też mają estymatory efektywniejsze niż średnia (rozdz. 7), np. dla rozkładu typu U (arc sin) – środek rozstępu [9], rozkładu Laplace'a, czyli dwuwykładniczego – mediana [7, 9]. Stosując metodę Monte Carlo wykryto, że dla rodziny rozkładów trapezów liniowych Trap najkorzystniejszy jest estymator dwuelementowy 0,5·(środek rozpięcia + średnia) [10] – rozdz. 8.

Literatura

1. *Evaluation of measurement data – Guide to the expression of uncertainty in measurement* (GUM). ISO first edition 1992, BIPM JCGM 100: 2008 (last corrected version)

1a. *Wyrażanie Niepewności Pomiaru. Przewodnik,* tłumaczenie wyd. 2 GUM z 1995 r. z komentarzami J. Jaworskiego, Wydawnictwo Głównego Urzędu Miar, Warszawa 1999.
2. Supplement 1 to GUM [1]. *Evaluation of measurement data – Propagation of Distribution using a Monte Carlo Method*. BIPM JCGM 101: 2008.
3. Kubisa S., Moskowicz S., *Studium przechodniości oceny przedziału ufności wyniku pomiaru, określanej za pomocą symulacji Monte Carlo*. „Pomiary Automatyka Kontrola", Nr 6, 2007, 7–10.
4. Cramer H., *Mathematical Methods of Statistics,* Stockholm Univ. (Chapter 19.1) 1946.
5. Novitski P.V., Zograf I.A., *Evaluation of measurement result errors*. Energoatomizdat, Leningrad 1985 – 248 p. (in Russian)
6. Dorozhovets M., Warsza Z.L., *Udoskonalenie metod wyznaczania niepewności wyników pomiaru w praktyce.* „Przegląd Elektrotechniczny", Nr 1, 2007, 1–13.
7. Dorozhovets M., Warsza Z.L., *Wpływ nieadekwatności wyboru parametrów rozkładu prawdopodobieństwa na niepewność typu A*. „Pomiary Automatyka Kontrola", Nr 9 bis, 2007, 25–28.
8. Warsza Z.L., Dorozhovets M., *Type A uncertainty evaluation of autocorrelated observations and choosing the best estimators of data distribution*. Proceedings of 18th National Symposium Metrology and Metrology Assurance. Sept. 2008, Sozopol Bulgaria, 70–78.
9. Warsza Z.L., *Jednoelementowe estymatory wartości menzurandu próbek o kilku niegaussowskich rozkładach prawdopodobieństwa – przegląd*. „Pomiary Automatyka Kontrola", Nr 1, 2011, 101–104.
10. Warsza Z.L., *Dwuelementowe estymatory wartości menzurandu próbek pomiarowych o trapezowych rozkładach prawdopodobieństwa – przegląd prac*. „Pomiary Automatyka Kontrola", Nr 1 2011, 105–108.
11. Fotowicz P., *Wykorzystanie rozkładu płasko-normalnego przy obliczaniu niepewności pomiaru*, „Pomiary Automatyka Kontrola", Vol. 57, Nr 6, 2011, 101–104.
12. Kubisa S., Warsza Z.L., *Środek rozstępu jako estymator menzurandu dla próbek z populacji o rozkładzie równomiernym i płasko-normalnym*. „Pomiary Automatyka Kontrola", Vol. 60, Nr 6, 2014, 398–401.
13. Kubisa S., Warsza Z.L., *Midrange as estimator of measured value for samples from population of uniform and Flatten-Gaussian distributions. "*Proceedings of 20th IMEKO TC4 International Symposium..." Benevento, Italy, September 15–17, 2014, 752–757.
14. Warsza Z., Kubisa S., *Comparisons by MC Method the Mean and Mid-range as Estimators of Measurand for Samples from Uniform and Flatten-Gaussian Populations*. "Proceedings of Symposium AMSA '15: Applied Methods of Statistical Analysis. Nonparametric Approach", Novosibirsk & Bialokuriha 14–19 September 2015, NGTU, 101–112.
15. Dorozhovets M., *Testowanie obserwacji istotnie odstających w próbach o rozkładzie jednostajnym*, „Mechanik", Nr 11/2016, 1596–1599.

Rozdział 7

Jednoelementowe estymatory menzurandu dla danych o kilku niegaussowskich rozkładach prawdopodobieństwa

Streszczenie. Przedstawiono wyniki badań efektywności różnych jednoelementowych estymatorów wartości menzurandu (wielkości mierzonej) dla danych o kilku niegaussowskich rozkładach prawdopodobieństwa wykonane metodą symulacji Monte Carlo. Dla rozkładu trapezowego Trap o bokach liniowych oraz krzywoliniowego wklęsłego CTrap wyznaczono odchylenia standardowe wartości średniej, środka rozpięcia i mediany w funkcji liczby n danych i kurtozy rozkładu próbki oraz stosunku długości podstaw β trapezu. Dla trapezu liniowego o β od 1 do 0,35 środek rozstępu próbki ma mniejsze odchylenie standardowe SD niż jej wartość średnia. Podano też wyniki symulacji stosunków standardowych odchyleń średniej i środka rozpięcia dla rodzin próbek danych modelowanych splotami rozkładów arcsin, normalnego i równomiernego.

Przewodnik GUM [1] zaleca, aby po usunięciu znanych błędów systematycznych, wyznaczać metodą statystyczną typu A niepewność standardową jako standardowe odchylenie (w skrócie SD) wartości średniej. Postępuje się więc tak, jak dla rozkładu normalnego przy nieskorelowanych (statystycznie niezależnych) danych eksperymentalnych. Tymczasem rozkład ten jest tylko modelem teoretycznym o funkcji Gaussa przebiegającej w nieograniczonym zakresie $\pm\infty$. Nie występuje on w rzeczywistości, gdyż wymagałoby to nieskończonej liczby oddziaływań losowych. Zwykle liczba ta jest skończona i może dominować kilka z nich, a rozrzuty wartości danych eksperymentalnych występują w ograniczonym przedziale. Próbkę rzeczywistą lepiej modelują niegaussowskie rozkłady prawdopodobieństwa, z estymatorami środka grupowania o mniejszym SD niż dla wartości średniej. Poprawny wybór najefektywniejszego estymatora próbki i prawidłowe oszacowanie jego SD rzutuje na dokładność oceny wyniku pomiarów. Amerykański National Institute of Standards and Technology NIST dopuszcza wyznaczanie niepewności typu A wg innych niż Przewodnik GUM zależności statystycznych [3], a w Przewodniku NASA [4] omawiane są jeszcze inne rozkłady prawdopodobieństwa danych pomiarowych.

W USA i kilku krajach europejskich prowadzone są badania o skali międzynarodowej nad udoskonaleniem zaleceń przewodnika GUM i rozszerzeniem ich na pomiary dynamiczne oraz pomiary procesów. Tematyką tą zajmował się też autor [5a–c, 6a–c, 10a–d]. W rozdziale 2 dla pomiarów o równomiernym próbkowaniu opisano sposób czyszczenia surowych danych pomiarowych z nieznanych *a priori* regularnych zakłóceń oraz możliwość automatyzacji szacowania niepewności procesów na bieżąco. W rozdziale 3 omówiono metodę wyznaczania standardowego odchylenia estymatorów dla próbki o skorelowanych danych i wprowadzono zastępczą liczbę pomiarów $n_{eff} < n$, którą można wykorzystywać w procedurach wg GUM.

W rozdziale 4 przedstawiono analizę właściwości i zastosowanie niekonwencjonalnego rozkładu w postaci jednego okresu funkcji $\cos^2$, przy stosowaniu którego dla danych próbki o ograniczonym zakresie rozrzutu uzyskano dokładniejsze oszacowanie niepewności niż dla rozkładu normalnego. W rozdziale 6 wykazano, że najlepszym jednoelementowym estymatorem menzurandu próbki o rozkładzie równomiernym jest środek rozpięcia, a dla dwuwykładniczego – mediana. W pracach [10a–d] zbadano efektywność estymatorów kilku rozkładów definiowanych jako sploty, w tym rozkładów trapezowych. Poniżej przedstawiono przegląd wyników uzyskanych dla estymatorów jednoelementowych (1C).

Przykłady efektywności estymatorów jednoelementowych

Istotę zagadnienia wyboru estymatorów przybliżą wstępnie przykłady liczbowe dla kilku próbek pobranych z populacji losowych o różnych rozkładach (tabela 1).

Tab. 1. Odchylenia standardowe trzech estymatorów menzurandu dla próbek o kilku rozkładach danych

Tab. 1. Standard deviations of three measurand estimators of few sample data PDF

Rozkład	Odchylenia standardowe dla trzech estymatorów			Najbardziej efektywny estymator	Standardowa niepewność estymatora
	S_x	$S_{qv/2}$	S_{med}		
Normalny	**0,010**	0,220	0,013	wartość średnia	$u_A = S_x / \sqrt{n}$
Równomierny	0,006	**1,4·10^{-4}**	0,010	środek rozpięcia	$\frac{V}{\sqrt{2}\ (n-1)}\sqrt{\frac{n+1}{n+2}}$
Dwuwykładniczy	0,007	0,870	**7·10^{-5}**	mediana	$S_x / \sqrt{2n}$
Trójkątny	**0,004**	0,0045	0,0049	wartość średnia	$S_x / \sqrt{n}$
Arc-sinus	0,067	**5·10^{-5}**	0,146	środek rozpięcia	$S_x \cdot \sqrt{5}\pi^4 / n^2$
$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} x_i$, $S_x = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \overline{X})^2}{n-1}}$,	$q_{V/2} = \frac{x_{\max} - x_{\min}}{2}$, $S_{qV/2} = \sqrt{\frac{\sum_{i=1}^{n}(x_i - q_{V/2})^2}{n-1}}$			$x_{med} = x_{s,(n+1)/2}$ lub $x_{med} = \frac{x_{s,n/2} + x_{s,n/2+1}}{2}$, $S_{med} = \sqrt{\frac{\sum_{i=1}^{n}(x_i - x_{med})^2}{n-1}}$	

Na dole tabeli 1 podano wzory dla trzech podstawowych estymatorów jednoelementowych tych próbek, tj. dla wartości średniej $\overline{X}$, środka rozstępu $q_{V/2}$ i mediany X_{med} (dla parzystej i nieparzystej liczby elementów) oraz dla standardowych odchyleń danych próbki liczonych względem każdego z nich. Środek rozstępu i medianę wyznaczono po uporządkowaniu elementów próbki wg ich wartości. Wyniki obliczeń odchyleń standardowych próbki dla tych estymatorów zestawiono w kolumnach 2–4.

Z tabeli 1 wynika, że podane w niej wartości standardowych odchyleń danych każdej z próbek od jej estymatorów różnią się między sobą dość istotnie, mimo że wartości samych estymatorów, np. środka rozstępu i średniej, zwykle są bardzo do siebie zbliżone. Dla każdego z rozkładów jeden z estymatorów najlepiej spełnia

warunek efektywności, tj. ma najmniejsze odchylenie standardowe (SD). Wartość średnia jest najlepszym estymatorem tylko dla rozkładu normalnego i trójkątnego.

Podane przykłady próbek modelowanych (tab. 1) rozkładami niegaussowskimi uzasadniają sens poszukiwania dla nich innego niż wartość średnia, bardziej efektywnego estymatora, tj. o jak najmniejszym odchyleniu standardowym i zastosowanie tego odchylenia do szacowania przedziału niepewności wyniku pomiaru o zadanym prawdopodobieństwie. W pierwszej kolejności omówione zostaną dwa rozkłady trapezowe: liniowy Trap i krzywoliniowy CTrap.

Najlepsze estymatory jednoelementowe próbek danych o rozkładach trapezowych

Rozkład o postaci trapezu liniowego Trap

Rozkład gęstości prawdopodobieństwa (PDF) w postaci symetrycznego trapezu o prostoliniowych bokach i polu powierzchni równym 1 w Suplemencie 1 do Przewodnika GUM [2] (tab. 1) oznaczony jest symbolem Trap(a, b). Można nim opisywać zarówno rozkłady populacji parametrów badanego obiektu [4, 10], jak i niepewności wyznaczane metodami: typu B [8] i typu A. Sygnały losowe o takich rozkładach pojawiają się w systemach pomiarowych jako wynik splotu składowych o rozkładach równomiernych występujących przy kwantowaniu sygnałów analogowych i zliczaniu impulsów. Na przykład wynikiem sumowania dwu różnych rozkładów równomiernych jest rozkład o PDF w postaci symetrycznego liniowego trapezu (rys. 1a). Rodzinę trapezów dla różnych stosunków szerokości splatanych rozkładów równomiernych przedstawia rys. 1b [9]. Mają one górną podstawę $2a$ i dłuższą dolną $2b$, a ich stosunek $\beta = a/b$ dla $b \geq a$ zmienia się w zakresie $\beta \in (0, 1)$. Według p. 4.3.9 GUM środek $x_i = 0{,}5(b_+ - b_-)$ populacji o rozkładzie Trap ma wariancję

$$u^2(x_i) = b^2(1+\beta)/6$$

Przypadkami granicznymi rodziny rozkładów Trap są:

- trójkąt ($\beta = 0$) – dla jednakowych obu splatanych rozkładów,
- prostokąt ($\beta = 1$) – gdy jeden ze splatanych rozkładów równomiernych ma pomijalną szerokość.

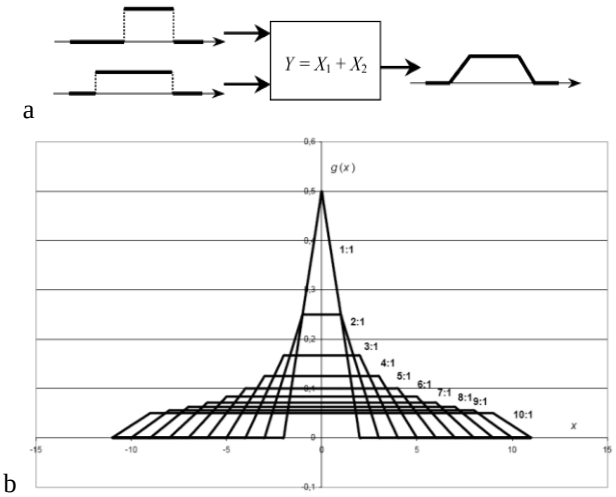
Miarą efektywności estymatorów są ich odchylenia standardowe SD. Najbardziej efektywnym jednoelementowym estymatorem wartości menzurandu dla próbek o rozkładzie trójkątnym jest wartość średnia, zaś dla próbek o rozkładzie prostokątnym – środek rozpięcia [5]. Aby wybrać właściwy estymator próbki modelowanej dowolnym z trapezów liniowych, trzeba podzielić je na dwa zbiory, w których lepszą efektywność ma jeden z tych estymatorów. Granicę podziału $\beta \equiv \alpha$ określono metodą symulacji Monte Carlo.

Model rozkładu trapezowego w metodzie MC uzyskano jedną z dwu technik:

- przez generację liczb losowych jako sumy dwu liczb o rozkładach równomiernych (rys. 1a),
- metodą funkcji odwrotnej (wyznaczonej dla trapezu [8]).

Podczas symulacji kolejne próbki o liczbie elementów n pobierano wielokrotnie z populacji losowych o rozkładzie Trap, które tworzono dla różnych stosunków

$\beta = a/b$ w zakresie $\beta \in (0, 1)$. Przy liczności każdej z próbek od 20 do 10 000 otrzymano stabilne wartości ich parametrów.



Rys. 1. Rozkład o postaci symetrycznego trapezu liniowego: a) wynik sumowania dwu rozkładów równomiernych: b) rodzina splotów dwu rozkładów równomiernych o różnych stosunkach szerokości [9]

Fig. 1. Linear symmetrical trapeze distribution: a) result of sum of two uniform distributions; b) set of convolutions of two uniform distributions of different ratios of width

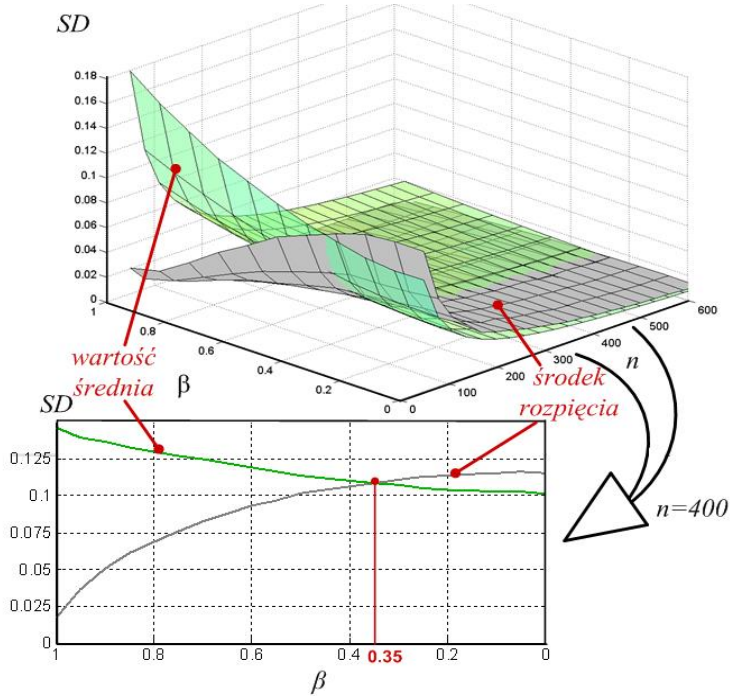
Standardowe odchylenia SD wartości średniej i środka rozpięcia próbek, otrzymane z modelowania MC, pokazano u góry rysunku 2. Są to funkcje parametrów n i β przedstawione jako powierzchnie w przestrzeni trójwymiarowej. U dołu podano ich przekrój dla $n = 400$. Odchylenia standardowe mediany są znacznie większe – na tym rysunku nie są przedstawione.

Z badań tych wynika ważny dla praktyki metrologicznej i statystyki stosowanej wniosek (niespotkany w literaturze wcześniejszej niż w pracach [10a–d]), że wybór najdokładniejszego estymatora 1C dla trapezowego rozkładu PDF zależy od jego kształtu. Dla trapezu liniowego jest to:

- środek rozpięcia w zakresie stosunku podstaw $1 \geq \beta > 0{,}35$,
- wartość średnia dla przedziału $0{,}35 \geq \beta \geq 0$, czyli jak dla rozkładu normalnego.

Taki sam rezultat otrzymano dla splotu dwóch rozkładów równomiernych o stosunku zakresów około 2,05, który odpowiada powyższej granicznej wartości $\beta \approx 0{,}35$. Wyznaczona analitycznie [11] graniczna wartość stosunku $\beta \equiv \alpha$ pokryła się z wynikami wcześniejszych symulacji MC

$$\alpha = \frac{-10 + 3\pi + \sqrt{9\pi^2 - 72\pi + 140}}{14 - 3\pi - \sqrt{9\pi^2 - 72\pi + 140}} = 0{,}3545779$$

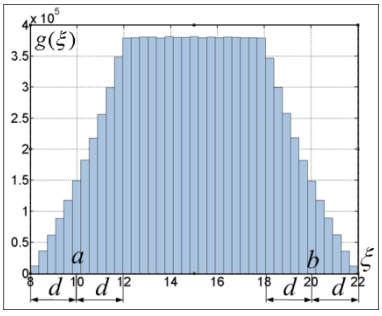


Rys. 2. Standardowe odchylenia SD wartości średniej i środka rozpięcia próbek o PDF w postaci trapezu liniowego Trap(*a, b*) [2] w funkcji stosunku podstaw trapezu β i liczności próbki *n*; u dołu – dla *n* = 400 w funkcji β [9d]

Fig. 2. Dependences of sample mean and midrange standard deviations *SD* on size *n* and on ratio β of linear trapeze bases as PDF Trap(*a, b*) [2] – upper figure; for *n* = 400 – lower figure [9d]

Rozkład o postaci trapezu krzywoliniowego CTrap

Rozpatrzono jeszcze rozkład trapezowy o symbolu CTrap. Histogram danych tego rozkładu zasymulowanych w wąskich przedziałach przedstawiono na rys. 3.



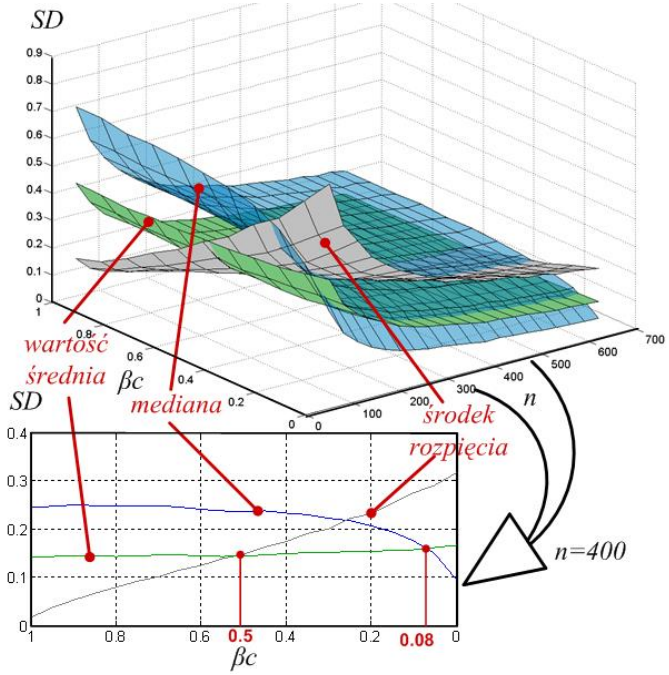
Rys. 3. Rozkład PDF CTrap(*a,b,d*) o kształcie trapezu krzywoliniowego

Fig. 3. Curvilinear trapeze PDF CTrap(*a,b,d*)

Rozkład CTrap jest uwzględniony w tabeli 1 Suplementu 1 do GUM [2]. Jego PDF jest krzywoliniowym wklęsłym trapezem. Stosuje się go wtedy, gdy długości górnej i dolnej podstawy nie są dokładnie znane i mogą zmieniać się w przedziałach $\pm d$, czyli gdy wynoszą $a \pm d$ i $b \pm d$, przy czym $(a, b, d) > 0$ oraz $a+d < b-d$.

Na rysunku 4 pokazano powierzchnie odchyleń standardowych SD trzech podstawowych estymatorów jednoelementowych (1C) tego trapezu w funkcji liczby obserwacji n w próbce i stosunku β_c największej górnej i najmniejszej dolnej podstawy jako

$$\beta_c=(a_2-a_1-2d)/(b_2-b_1+2d)$$



Rys. 4. Odchylenia standardowe SD wartości średniej, środka rozpięcia $q_{V/2}$ i mediany X_{med} próbek z populacji o PDF w kształcie trapezu krzywoliniowego CTrap w funkcji liczby elementów n próbki i stosunku podstaw trapezu β_c, poniżej – przekrój dla n = 400 [10d]

Fig. 4. SD of single component estimators as function of bases ratio $\beta \equiv \beta_c$ and of sample size n of curvilinear trapeze PDF – upper figure; cross-section for n = 400 – lower figure [10d]

Występują tu aż trzy charakterystyczne obszary. W pierwszym wąskim obszarze o małych β_c, tj. dla zakresu $0 < \beta_c < 0{,}08$, najlepsza jest mediana, w następnym $0{,}08 < \beta_c < 0{,}5$ – wartość średnia, zaś dla $0{,}5 < \beta_c < 1$ dominuje środek rozpięcia. Przy wyznaczaniu niepewności pomiarów należy wziąć pod uwagę, że jest ona wielkością drugiego rzędu i w części praktycznych przypadków, np. przy sprawdzaniu przyrządów wystarcza, gdy szacuje się ją z niepewnością nawet rzędu 20%. Wówczas:

$$d = \frac{(b-a)}{2} \cdot \frac{(1-\beta_c)}{(1+\beta_c)} = 0{,}3\frac{(b-a)}{2} \Rightarrow \beta = 0{,}54$$

Można przyjąć, że wartość średnia jest najbardziej efektywnym estymatorem aż do granicy $\beta = 0$ na rys. 4. Wówczas rozkłady o postaci trapezu krzywoliniowego będą miały, jak poprzednio, tylko dwa obszary o różnych najefektywniejszych estymatorach 1C, ale o granicy podziału nieco przesuniętej w górę do około $\beta_c = 0{,}5$.

Kurtoza rozkładów trapezowych

Novitzky i Zograph [7] podając topograficzną klasyfikację rozkładów, omawiają efektywność środka rozpięcia dla kilku typów rozkładu i efektywność estymatorów w funkcji odwrotności współczynnika kurtozy, czyli kontrkurtozy $\varepsilon^{-1} \in (0,1)$.

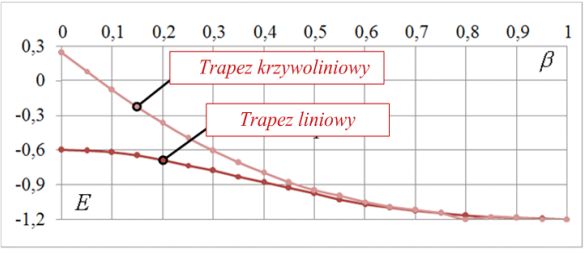
Należy tu przytoczyć definicje kurtozy E i współczynnika kurtozy ε, gdyż często te pojęcia statystyczne są ze sobą mylone. Wyznacza się je zarówno dla populacji, jaki dla próbek o liczbie elementów n (patrz rozdział 5). Współczynnik kurtozy ε opisany jest wzorem

$$\varepsilon \equiv \frac{\mu_4}{S_x^4},$$

gdzie: μ_4 – moment centralny rzędu czwartego, S_x – odchylenie standardowe.

Natomiast kurtoza $E = \varepsilon - 3$. Dla rozkładu normalnego $\varepsilon = 3$ oraz $E = 0$.

Na rysunku 5, dla dwu rodzajów rozkładów trapezowych, liniowego Trap oraz krzywoliniowego CTrap, przedstawiono kurtozę E (zwaną też ekscesem) w funkcji stosunku podstaw obu trapezów $\beta = \beta_C$. Odchylenia standardowe estymatorów środka rozpięcia i wartości średniej trapezu liniowego są jednakowe dla trapezu o kurtozie $E = -0{,}805$ odpowiadającej jednoznacznie stosunkowi długości podstaw $\beta = 0{,}35$.



Rys. 5. Współczynnik kurtozy E rozkładów trapezowych Trap i Ctrap w funkcji stosunku β długości podstaw ich trapezów

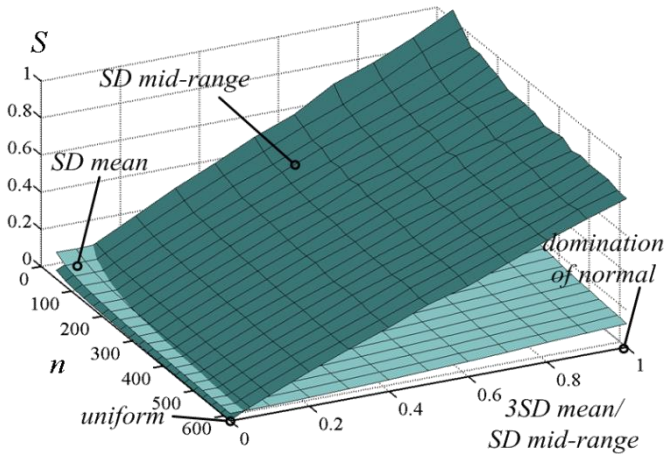
Fig. 5. Coefficient of kurtosis E of Trap and Ctrap distributions as function of ratio β of their trapeze bases lengths

Estymatory jednoelementowe (1C) splotów innych rozkładów

Splot rozkładu równomiernego z normalnym

Standardowe odchylenia S estymatorów 1C tego splotu przedstawiono na rys. 6. Wynika z niego, że środek rozpięcia jest lepszym estymatorem (o mniejszym SD) od średniej tylko w wąskim obszarze bliskim $\beta = 1$ dla rozkładu równomiernego. Obszar ten ponadto zwęża się wraz ze wzrostem liczby elementów n próbki. Szcze-

gółowemu porównaniu obu powyższych estymatorów menzurandu był poświęcony rozdział 6.

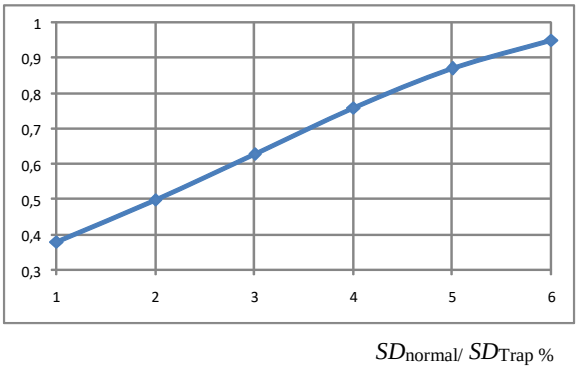


Rys. 6. Standardowe odchylenia (SD) średniej i środka rozpięcia rodziny splotów rozkładu normalnego i równomiernego

Fig. 6. Standard deviations (SD) of mean and midrange of the family of uniform and normal PDF convolution

Splot rozkładu trapezowego Trap z normalnym

Dla tego splotu zmniejsza się graniczny stosunek podstaw trapezów $\beta = \alpha$ o lepszym środku rozpięcia lub średniej jako estymatorze 1C próbki. Na rys. 7 przedstawiono to dla rozkładu Trap przy różnych % SD dodanego rozkładu normalnego (np. szum). Przy 5% średnia dominuje już aż dla 95% zakresu β.

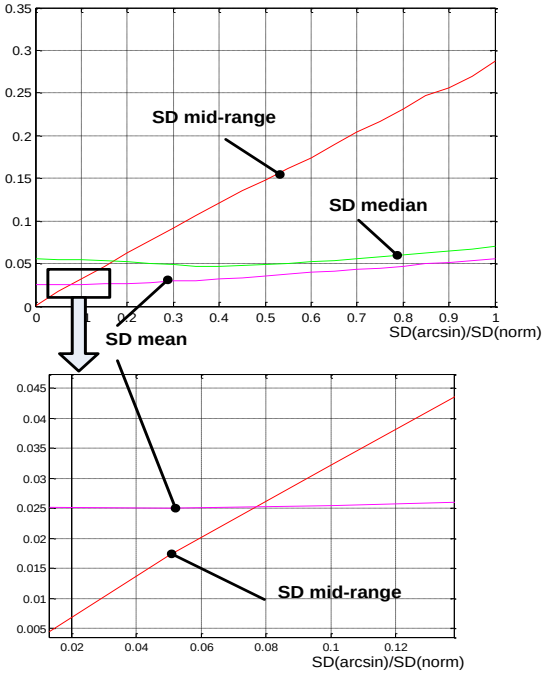


Rys. 7. Zmiany granicznej wartości stosunku podstaw β rozkładu Trap przy wzrastającym udziale rozkładu normalnego w %

Fig. 7. Changes of border β of the ratio of Trap PDF bases (linear trapeze) for increasing % of the added normal PDF

Splot rozkładu arc sin z normalnym

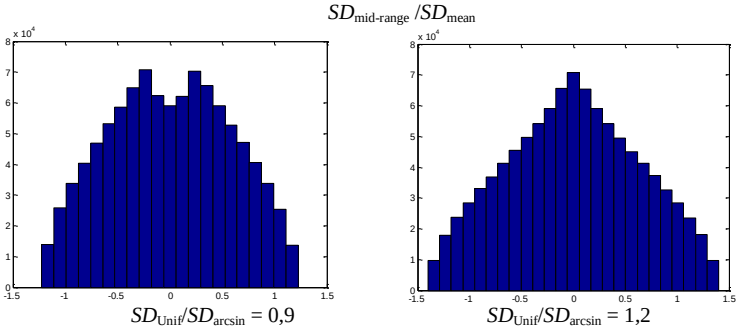
Standardowe odchylenia środka rozpięcia i średniej dla tego splotu podano na rys. 8. W większości zakresu stosunku SD obu składników dominuje średnia arytmetyczna.



Rys. 8. Standardowe odchylenia estymatorów 1C splotu rozkładów arc sin i normalnego
Fig. 8. SD of 1C (one-component) estimators of arc sin and normal PDF convolution

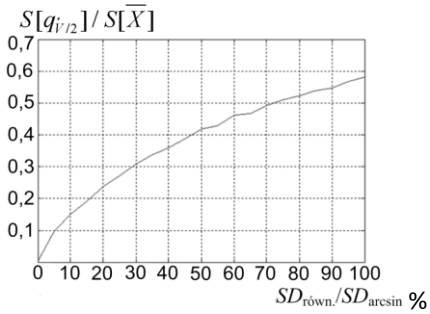
Splot rozkładów: równomiernego i arc sin

Rozkłady PDF dla dwu różnych stosunków odchyleń standardowych splatanych rozkładów przedstawiono na rys. 9.



Rys. 9. Dwa przykłady PDF splotu rozkładów równomiernego i arc sin
Fig. 9. Two examples of PDF-s of the Uniform and arc sin convolution

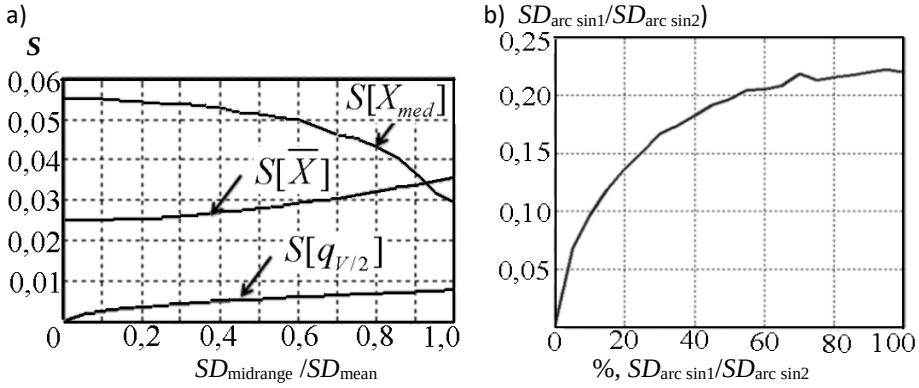
Zależność stosunku odchyleń standardowych środka rozpięcia i średniej i od stosunku standardowych odchyleń obu splatanych rozkładów w % podano na rys. 10. W całym zakresie $SD_{równ.}/SD_{arc\ sin}$ najlepszym estymatorem 1C jest środek rozpięcia.



Rys. 10. Stosunek odchyleń standardowych SD środka rozpięcia i średniej splotu rozkładów równomiernego i arc sin
Fig. 10. Ratio of mid-range and mean SD of uniform and arcsin PDF convolution

Splot dwu rozkładów arc sin

Zależności SD estymatorów splotu arcsin*arcsin przedstawiono na rys 11a, b.



Rys. 11. Odchylenia standardowe SD estymatorów 1C splotu dwu rozkładów arc sin
Fig. 11. Standard Deviations SD of 1C estimators for convolution of two arc sin PDF

W całym zakresie (0, 1) stosunku $SD_{arc\ sin1}/SD_{ar\ csin2}$ splatanych dwu rozkładów arc sin odchylenie standardowe środka rozpięcia $SD_{midrange}$ splotu jest najmniejsze.

Podsumowanie

- Prawidłowe oszacowanie wyznaczanej statystycznie niepewności (odpowiednik u_A z Przewodnika GUM) zależy od wyboru efektywnego, tj. najbardziej dokładnego estymatora PDF próbki, stosowanego w przetwarzaniu jej danych.
- Najlepszy jednoelementowy (1C) estymator wartości menzurandu dla próbki o rozkładzie PDF modelowanym symetrycznym trapezem zależy od jego

kształtu opisanego przez stosunek podstaw β. Jeśli trapez jest bliski prostokąta ($1 \geq \beta \geq 0{,}35$) to lepszy jest środek rozstępu próbki, a poniżej ($0{,}35 \geq \beta \geq 0$) aż do rozkładu trójkątnego ($\beta = 0$) – wartość średnia. Inne wnioski są w tekście.

Literatura

1. *Wyrażanie Niepewności Pomiaru. Przewodnik*, tłumaczenie z angielskiego Guide to the Expression of Uncertainty in Measurement, BIPM, revised and corrected 2nd ed. 1995 Wydawnictwo Głównego Urzędu Miar, Warszawa 2002.
2. Supplement 1 to the Guide to the expression of uncertainty in measurement (GUM) – *Propagation of distributions using a Monte Carlo method*, Guide OIML G 1-101 Ed. 2007 (E).
3. Taylor B.N., Kuyatt Ch.E., *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*, NIST Technical Note 1297, National Institute of Standards and Technology, USA1994.
4. *Measurement Uncertainty Analysis Principles and Methods*, NASA Measurement Quality Assurance Handbook – ANNEX 3, July 2010.

5a. Dorozhovets M., Warsza Z.L., *Udoskonalenie metod wyznaczania niepewności wyników pomiaru w praktyce*. „Przegląd Elektrotechniczny", Nr 1, 2007, 1–13.

b. Dorozhovets M., Warsza Z.L., *Methods of upgrading the uncertainty of type A evaluation (2). Elimination of the influence of autocorrelation of observations and choosing the adequate distribution*, Proceedings of 15th IMEKO TC4 Symposium, Sept. 2007 Iasi Romania, 199–204.

c. Warsza Z., Dorozhovets M., *Uncertainty type A evaluation of autocorrelated measurement observations*. Biuletyn WAT (Wojskowej Akademii Technicznej), Vol. 57, Nr 2, 2008, 143–152.

6a. Warsza Z.L., Dorozhovets M., Korczynski M.J., *Methods of upgrading the uncertainty of type A evaluation* (1). *Elimination the influence of unknown drift and harmonic components*, Proceedings of 15th IMEKO TC4 Symposium, Sept. 2007, Iasi Romania, 193–198.

b. Dorozhovets M., Warsza Z., Korczyński M.J., *Udoskonalenie metody typu A wyznaczania niepewności pomiarów…* „Pomiary Automatyka Kontrola", Nr 12, 2007, 8–11.

c. Warsza Z.L., Korczyński M.J., *Eliminacja nieznanych a priori składowych systematycznych z niepewności typu A pomiarów o równomiernym próbkowaniu*. „Przegląd Elektrotechniczny", Nr 05, 2008, 109–114.

7. Novickij P.V., Zograf I.A., *Ocenka pogreshnostiej resultatov izmierenii*, Energoatomizdat, Leningrad 1985
8. Kacker R.N., Lawrence J.F., *Trapezoidal and triangular distributions for Type B evaluation of standard uncertainty*. "Metrologia", Vol. 44, 2007, 117–127.
9. Fotowicz P., *Ocena dokładności przybliżenia splotu rozkładu normalnego i prostokątnego rozkładem trapezowym*, „Pomiary Automatyka Robotyka", Nr 5, 2001, 9–11.

10a. Warsza Z.L., Galovska M., *About the best measurand estimators of trapezoidal probability distributions*. "Przegląd Elektrotechniczny – Electrical Review", Nr 5, 2009, 86–91.

b. Warsza Z.L., Galovska M., *The best measurand estimators of trapezoidal PDF*. Proceedings of IMEKO World Congress ”Fundamental and Applied Metrology”, September 2009, Lisbon Portugal, 2405–2410.
c. Warsza Z.L., Galovskaja M.V., *Vybor najlutshej ocenki izmierajemoj velichiny na primiere trapecievidnych raspredelenij*, “Sistemy Obrobotki Informacii”, Vol. 4, No. 78, 2009, Charków, 28–31.
d. Warsza Z.L., Galovska M., *Estymatory wartości menzurandu dla trapezowych rozkładów prawdopodobieństwa danych pomiarowych*, „Pomiary Automatyka Komputery w Gospodarce i Ochronie Środowiska”, Nr 1, 2010, 12–16.
11. Endovitskyi P., Ukraiński Narodowy Uniwersytet Techniczny „Kijowski Instytut Politechniczny”, praca niepublikowana
12. Galovska M., Warsza Z.L., *Estymatory wartości menzurandu próbek danych o rozkładach niegaussowskich*. Metrologia dziś i jutro. Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 2010, 59–72.
13. Warsza Z.L., Galovska M., *Estymatory wartości menzurandu danych o niegaussowskich rozkładach prawdopodobieństwa.* „Przegląd Elektrotechniczny – Electrical Review”, Nr 12, 2010, 253–258.
14. Warsza Z.L., *Jednoelementowe estymatory wartości menzurandu próbek o kilku niegaussowskich rozkładach prawdopodobieństwa – przegląd*, „Pomiary Automatyka Kontrola”, Nr 1, 2011, 101–104.

Rozdział 8

Dwuelementowe estymatory menzurandu dla danych o trapezowych rozkładach prawdopodobieństwa

Streszczenie. Dla próbek modelowanych symetrycznymi trapezowymi rozkładami prawdopodobieństwa omówiono estymatory dwuelementowe (2C) wartości menzurandu jako kombinację liniową wartości średniej i środka rozstępu próbki. Wyznaczono metodą Monte Carlo ich współczynniki dla minimum odchylenia standardowego (SD) jako funkcje kurtozy E lub stosunku podstaw trapezu β. Wykazano, iż estymator 2C trapezowego rozkładu liniowego Trap w zakresie $0 < \beta < 0{,}75$ ma mniejsze SD niż środek rozstępu i średnia danych próbki. Określono też przedziały β dla najlepszych estymatorów rozkładu CTrap – trapez krzywoliniowy. Zaproponowano uproszczony estymator 2C o obu współczynnikach 0,5, który jest zbliżony do optymalnego. Poprzez symulację i przykład liczbowy sprawdzono jego skuteczność dla różnych β. Ten estymator nie jest objęty zaleceniami GUM. W praktyce umożliwia uzyskać dokładniejszą wartość niepewności typu A niż wyznaczana wg GUM.

W wyniku występowania zjawisk losowych w badanych obiektach i w systemach pomiarowych powstaje rozrzut wartości wyników pomiaru wielkości mierzonej (w ogólnym przypadku parametrów menzurandu). Do jego modelowania, obok normalnego rozkładu prawdopodobieństwa opisanego funkcją Gaussa, stosuje się też inne elementarne rozkłady nazywane niegaussowskimi, np. rozkład równomierny oraz różne ich sploty, w tym rozkład trapezowy. W publikacjach [7–10] omówiono otrzymane metodą symulacji Monte Carlo (MC) wyniki badań efektywności wartości średniej $\overline{X}$, środka rozstępu $q_{V/2} = 0{,}5(x_{max} - x_{min})$ i mediany X_{med} jako estymatorów jednoelementowych (1C) menzurandu dla próbek danych o rozkładzie gęstości prawdopodobieństwa (PDF) w postaci symetrycznego trapezu o bokach liniowych, oraz trapezu krzywoliniowego wklęsłego, oznaczonych w Suplemencie 1 do Przewodnika GUM [1] odpowiednio symbolami Trap oraz CTrap. Podano odchylenia standardowe (SD) tych estymatorów w funkcji liczebności n próbki danych i stosunku długości podstaw trapezu β lub kurtozy E. Dla rozkładów Trap wartość średnia próbki ma mniejsze SD niż środek rozstępu tylko w przedziale $0 < \beta < 0{,}35$, który dominuje powyżej, tj. dla $0{,}35 \leq \beta \leq 1$ (rozdz.7, rys. 2). Dla trapezów krzywoliniowych CTrap przedział dominacji środka rozstępu jest węższy, tj. $1 > \beta > 0{,}5$ (rozdz. 7, rys. 4). W przedziale $0{,}5 > \beta > 0{,}08$ najmniejsze SD ma wartość średnia, następnie pojawia się wąski przedział o najmniejszym SD dla mediany. Jako kolejne zagadnienie analizowano [7–10], czy i jaką dalszą poprawę oszacowania dokładności wyniku pomiaru z próbek danych o rozkładach trapezowych uzyska się stosując kilkuelementowe estymatory liniowe. Synteza tych badań i porównanie ich wyników z estymatorami 1C z rozdziału 7 jest przedstawiona na kolejnych stronach.

Estymatory dwuelementowe (2C) jako funkcje współczynnika kurtozy *E*

Zakharov i Stephen [4] badali metodą Monte Carlo dla rozkładów symetrycznych efektywność liniowego estymatora trójelementowego (3C) o postaci

$$\hat{X} = k_1 \overline{X} + k_2 q_{V/2} + k_3 X_{med} \tag{1}$$

gdzie: $\overline{X}$ – wartość średnia, $q_{V/2}$ – środek rozstępu, X_{med} – mediana.

Z przeprowadzonych symulacji [4] wynikało, że dla rozkładu Trap(*a, b*) o zakresie kurtozy $E \in (-1{,}15; -0{,}2)$ wystarczą tylko dwa współczynniki k_1 oraz k_2, gdyż $k_3 = 0$ i estymator (1) staje się dwuelementowy. Estymatory takie są oznaczane dalej symbolem 2C. Przy wartościach k_1, k_2 podanych w [4] estymator ten jest obciążony, gdyż nie spełnia jednego z podstawowych wymagań efektywności, tj. warunku $k_1 + k_2 = 1$ [3]. Wspólnie z Mariną Galowską sprawdziliśmy i zweryfikowaliśmy te współczynniki metodą Monte Carlo [7–10] otrzymując wartości:

$$k_1 = -1{,}05E + 1{,}22, \quad k_2 = 1{,}05E - 0{,}22 \tag{2}$$

Natomiast dla próbek o rozkładzie CTrap(*a, b, d*), tj. o PDF w kształcie wklęsłego trapezu krzywoliniowego, z symulacji metodą Monte Carlo w zakresie kurtozy $E \in (-1{,}2; 0{,}2)$ otrzymano dla nieobciążonego estymatora (1) współczynniki:

$$\begin{aligned}
k_1 &= \begin{cases} 0, \textit{if } E \le -1.15; \\ 1{,}05 \cdot E + 1{,}22, \textit{if } -1{,}15 < E \le -0{,}2; \\ 1, \textit{if } E \ge -0{,}2; \end{cases} \\
k_2 &= 1 - k_1 = \begin{cases} 1, \textit{if } E \le -1.15; \\ -1{,}05 \cdot E - 0{,}22, \textit{if } -1{,}15 < E \le -0{,}2. \\ 0, \textit{if } E \ge -0{,}2; \end{cases} \\
k_3 &= 0.
\end{aligned} \tag{3}$$

Taką samą wartość *E* mogą mieć rozkłady o różnych kształtach [3]. Wyznaczanie współczynników k_1, k_2 estymatora (1) przy $k_3 = 0$, poprzedzić trzeba oszacowaniem kurtozy *E* rozkładu eksperymentalnego próbki, lub przyjętego jej modelu. Stosując metodę Monte Carlo w [9] należy brać pod uwagę, że obliczenia kurtozy *E* na podstawie danych próbki są mało dokładne, szczególnie dla małej liczebności *n*.

Estymatory 2C jako funkcje stosunku podstaw trapezu *β*

Standardowe odchylenia estymatorów dwuelementowych dla rozkładów Trap oraz CTrap o trapezowych PDF dogodniej jest przedstawiać bezpośrednio w funkcji ich parametrów geometrycznych, najlepiej jednolicie, tak jak dla estymatorów jednoelementowych, tj. w funkcji stosunku β lub β_c podstaw trapezu [7, 10]. Zależności współczynników k_1, k_2 od β lub β_c otrzymuje się po przeliczeniu ze współczynnika kurtozy *E* wg krzywych podanych w rozdziale 7 na rys. 5 [7,10].

Zbadano też metodą MC efektywność nieobciążonego estymatora 2C dla całej rodziny rozkładów o PDF w postaci symetrycznych trapezów liniowych [7]. Estymator taki zapisuje się ogólnie jako

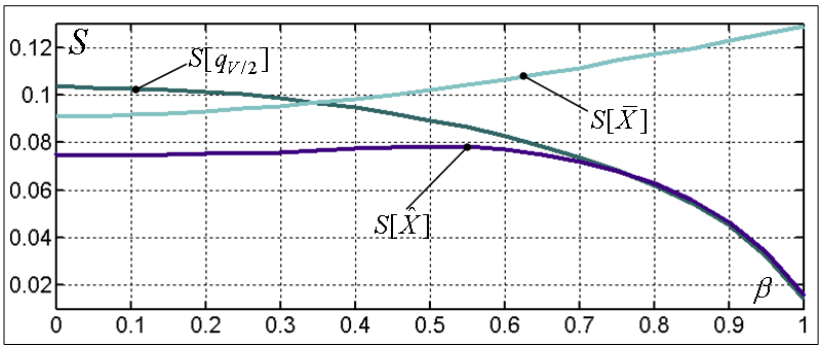
$$\hat{X} = k_1 \overline{X} + (1 - k_1)\, q_{V/2} \tag{4}$$

Estymator $\hat{X}$ symulowano metodą MC dla różnych kształtów rozkładów trapezowych począwszy od prostokąta, aż do trójkąta i próbek o różnej liczebności *n*. Wyznaczono wartości k_1 dla warunku $\min S[\hat{X}]$ z niepewnością < 10% i $k_2 = 1 - k_1$

$$k_1 = \begin{cases} -0{,}12\beta + 0{,}56, \; if \;\; 0 < \beta < 0.5 \\ 1 - \beta, \; if \;\; 0.5 < \beta < 1 \end{cases}, \qquad k_2 = 1 - k_1 = \begin{cases} 0{,}12\beta + 0{,}44, \; if \;\; 0 < \beta < 0.5 \\ \beta, \; if \;\; 0.5 < \beta < 1 \end{cases}. \tag{5}$$

Dla trapezów o różnych stosunkach podstaw β, wartości tych współczynników pozostają stałe dla liczebności próbki *n* już od 10, aż do 10 000.

Na rysunku 1 porównano odchylenia standardowe $S\left[\hat{X}\right]$ estymatora 2C, średniej $S\left[\overline{X}\right]$ i środka rozstępu $S[q_{V/2}]$ w funkcji stosunku podstaw β trapezu liniowego jako modelu dla PDF próbki. Z porównania wynika, że estymator dwuelementowy $\hat{X}$ jest lepszy niż estymatory jednoelementowe w szerokim przedziale $0 \le \beta \le 0{,}75$.

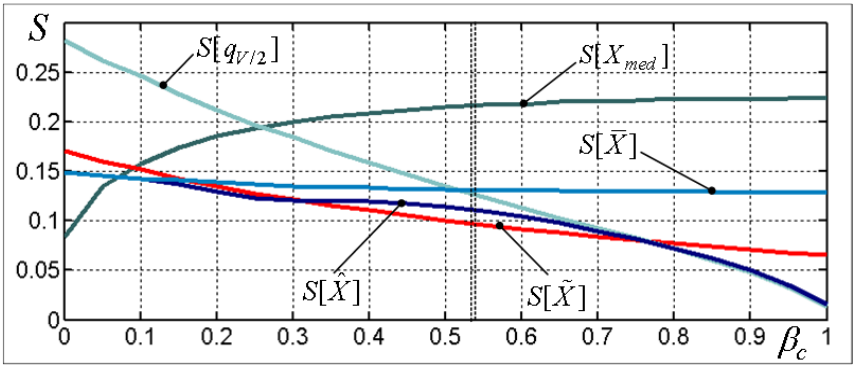


Rys. 1. Odchylenia standardowe estymatorów menzurandu: $S[\hat{X}]$ – estymatora 2C wg (2) oraz $S[\overline{X}]$ – średniej i $S[q_{V/2}]$ – środka rozstępu dla próbki o rozkładzie Trap(*a*, *b*) w funkcji stosunku podstaw β trapezu. Liczebność próbki *n* = 400

Fig. 1. Standard deviations $S[\hat{X}]$ of measurand estimator 2C from eq. (2), $S[\overline{X}]$ and $S[q_{V/2}]$ of the sample of linear trapeze PDF Trap(*a*, *b*) as function of a ratio of bases β. Number of sample elements *n* = 400

Dla próbki o PDF modelowanym trapezem krzywoliniowym CTrap(*a*, *b*, *d*) przebiegi standardowych odchyleń estymatorów w funkcji jego β_C są podobne [7–10] (rys. 4 w rozdz. 7). W wąskim zakresie $\beta_C \in [0; 0{,}8]$ najlepszym estymatorem jest tu mediana, czego też nie zauważano we wcześniejszej literaturze. Estymator

dwuelementowy $\hat{X}$ jest lepszy od średniej i od środka rozpięcia $q_{V/2}$ w całym zakresie β_C, gdyż jego odchylenie standardowe $S[\hat{X}]$ jest najmniejsze, szczególnie gdy $\beta_C \to 1$, tj. dla trapezów bliskich prostokąta.



Rys. 2. Odchylenia standardowe estymatorów menzurandu w funkcji stosunku podstaw β_C trapezu krzywoliniowego jako PDF próbki

Fig. 2. Standard deviations of different measurand estimators as function of a ratio of bases β_C of curvilinear trapeze as model of the sample PDF

Uproszczony estymator dwuelementowy

Do stosowania w praktyce zaproponowano [7] uproszczony liniowy estymator 2C o jednakowych współczynnikach $k_1 = k_2 = 0{,}5$

$$\tilde{X} = 0{,}5\,\bar{X} + 0{,}5\,q_{V/2} \tag{6}$$

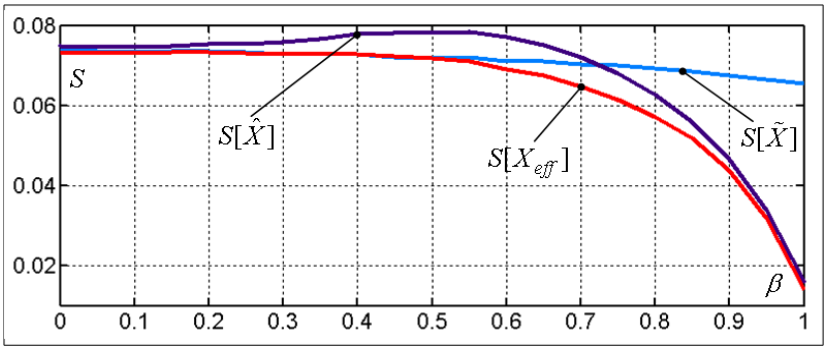
Właściwości tego dwuelementowego estymatora przeanalizowano metodą Monte Carlo. Na rysunku 2 zamieszczono przebieg jego odchylenia standardowego $S[\tilde{X}]$ w funkcji β dla próbki o PDF jako trapezie liniowym. Z porównania rezultatów wynika, że uproszczony estymator 2C wg (6) niewiele odbiega od najlepszych k_1, k_2 wg (5) w szerokim zakresie stosunku podstaw trapezów $0 < \beta < 0{,}75$. Jeśli rozkład trapezowy traktować jako splot dwu rozkładów równomiernych, to stosunek szerokości prostokątów wynosi $2(1 + 2\beta)$. Zakresowi $1 > \beta > 0{,}75$ odpowiada stosunek szerokości powyżej 8, czyli dominuje jeden z rozkładów równomiernych, wówczas najlepszym jest środek rozpięcia $q_{V/2}$ (rys. 1 i 2).

Uproszczony estymator 2C wg (6) można stosować też dla rozkładu o postaci trapezu krzywoliniowego CTrap(a, b, d). Wynika to z analizy wykresów (rys. 2).

Estymator (6) nie jest najbardziej efektywny w całym zakresie β. Sprawdzenie kilku przypadków dla $\beta > 0{,}54$ uzasadnia, że dla rozkładów trapezowych można rekomendować stosowanie różnych postaci estymatora X_{eff} w dwu zakresach β, tj.:

$$X_{eff} = \begin{cases} \dfrac{1}{2}\cdot q_{V/2} + \dfrac{1}{2}\bar{X}, & if \quad 0{,}54 < \beta < 0{,}8; \\ q_{V/2} & if \quad \beta > 0{,}8. \end{cases} \tag{7}$$

Po uwzględnieniu (5) otrzymano zależności od β podane na rys 3.



Rys. 3. Odchylenia standardowe S dwu estymatorów 2C wg (5) i (6) oraz estymatora X_{eff} (7) w funkcji stosunku podstaw β trapezu

Fig. 3. Standard deviations S of two 2C estimators of (5) and (6) and estimator X_{eff} of (7) as function of ratio β of trapeze bases

Podstawy teoretyczne oceny niepewności rozkładów trapezowych

Odchylenie standardowe $S[\hat{X}]$ dwuelementowego estymatora (4)

$$S(\hat{X}) = \sqrt{k_1^2 S^2[\bar{X}] + (1-k_1)^2 S^2[q_{V/2}] + 2\rho k_1(1-k_1) S[\bar{X}] \cdot S[q_{V/2}]} \tag{8}$$

gdzie: $S[\bar{X}] = \dfrac{S_x}{\sqrt{n}}$; $S^2[q_{V/2}] = \dfrac{V^2(1-\beta^2)(4-\pi)}{8n}$ [13]; $V = X_{max} - X_{min}$

W wyniku modelowania MC [7, 10] otrzymano wartości współczynnika korelacji ρ rekomendowane dla różnych zakresów n. Podano je w tabeli 1.

Tab. 1. Zalecane wartości współczynnika korelacji ρ

Tab. 1. Recommended values of correlation coefficient ρ

n	[100; 200)	[200; 300)	[300; 500)	$n \to \infty$
ρ	0,25	0,2	0,15	$\rho \to 0$

Dla dużych n wpływ korelacji jest nieduży i wzór (8) upraszcza się do postaci

$$S(\hat{X}) = \sqrt{k_1^2 S^2[\bar{X}] + (1-k_1)^2 S^2[q_{V/2}]} \tag{9}$$

Różnice między wartościami otrzymanymi analitycznie i z modelowania były mniejsze niż 5%.

Najlepszy estymator powinien mieć minimalną wariancję. Jego współczynnik k_1 otrzymuje się dla $\dfrac{\partial S^2[X_{eff}]}{\partial k_1} = 0$. Optymalną wartość k_1 wyznaczono analitycznie na podstawie zależności (9). Dla dużych n wynosi ona

$$k_1 = \frac{S^2[q_{V/2}]}{S^2[\bar{x}] + S^2[q_{V/2}]} \tag{10}$$

Standardowe odchylenie (odpowiednik niepewności u_A) dla estymatora 2C wg (6) o $k_1 = k_2 = 0{,}5$ oraz estymatora X_{eff} wg (7) w zakresie $0{,}54 < \beta < 0{,}8$ wyznacza się jako

$$u_A(\tilde{X}) = \tfrac{1}{2}\sqrt{S^2[\bar{X}] + S^2[q_{V/2}] + 2\rho \cdot S[\bar{X}] \cdot S[q_{V/2}]} =$$

$$= \tfrac{1}{2}\sqrt{\frac{S^2[\bar{X}]}{n} + \frac{V^2(1-\beta^2)(4-\pi)}{8n} + \rho\frac{S^2[\bar{X}]\cdot V}{n}\sqrt{\frac{(1-\beta^2)(4-\pi)}{2}}} \tag{11}$$

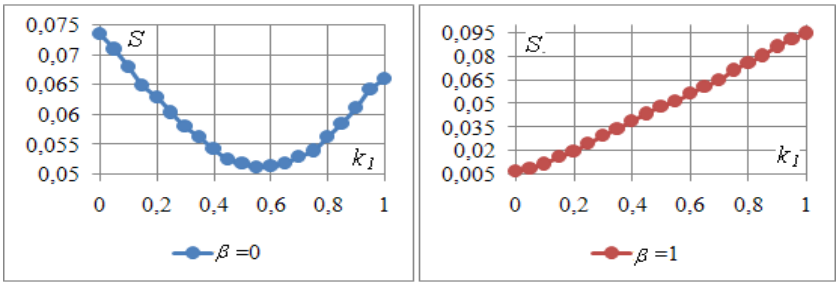
Dla próbek o dużym n składnik zależny od ρ staje się pomijalny i wynosi

$$u_A(\tilde{X}) = \tfrac{1}{2}\sqrt{S^2[\bar{X}] + S^2[q_{V/2}]} \tag{11a}$$

Różnice między wartościami otrzymanymi na drodze analitycznej i z modelowania były mniejsze od 5%.

Estymator 2C w przypadkach szczególnych

Zależności standardowych odchyleń S w funkcji k_1 dla granicznych przypadków trapezowego PDF (trójkąt, prostokąt) dla dużych n, tj. przy pomijalnym składniku od współczynnika korelacji, podano na rys. 4.



Rys. 4. Zależności odchyleń standardowych S od współczynnika k_1 dla granicznych przypadków PDF trapezu liniowego
Fig. 4. Standard deviation ***S*** as function of coefficient k_1 for both border cases of trapeze PDF

Wynika z nich wniosek intuicyjnie oczywisty, że trójkątny rozkład PDF ma właściwości pomiędzy rozkładami normalnym, a równomiernym. Można dla niego stosować jako najlepsze estymatory dwuelementowe 2C, natomiast dla rozkładu normalnego najlepsza jest wartość średnia. Analitycznie dla dużych n otrzymuje się:

Dla rozkładu trójkątnego ($\beta = 0$)

$$S^2\left[q_{V/2}\right] = \frac{V^2(4-\pi)}{8n}, \qquad \text{oraz} \qquad k_1 = \frac{3(4-\pi)}{2+3(4-\pi)} \approx 0{,}56\,.$$

Wyniki pokrywają się z rezultatami symulacji metodą Monte Carlo.

Dla rozkładu trapezowego o stosunku długości podstaw $\beta = 0{,}35$ otrzyma się $S[\bar{X}] = S[q_{V/2}]$ i z (10):

$$k_1 = \frac{\sigma_X^2}{n} \Big/ \frac{2\sigma_X^2}{n} = 0{,}5\,,$$ a więc zgodnie z (5).

Dla rozkładu prostokątnego ($\beta = 1$):

$$k_1 = \frac{\dfrac{3\sigma_X^2}{2(n+1)(n+2)}}{\dfrac{\sigma_X^2}{n} + \dfrac{3\sigma_X^2}{2(n+1)(n+2)}} = \frac{3n}{2n^2+9n+2}\,,$$

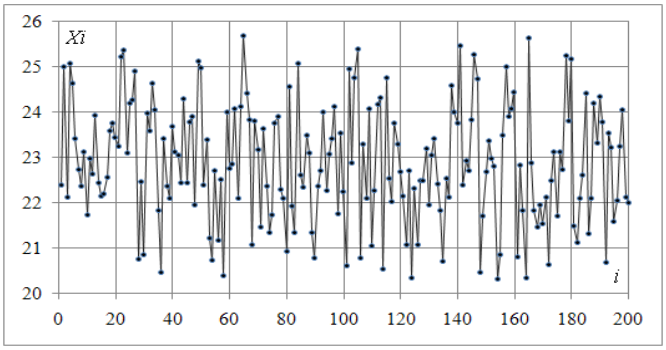
Jeżeli $n \to \infty$, to $k_1 \to 0$. Dla $n = 30$, $k_1 = 0{,}04$, zaś dla $n = 10$, $k_1 = 0{,}1$.

Otrzymane wartości k_1 są bardzo bliskie zera dla $n \to \infty$.

Przykład

Rozważania zilustrowano przykładem liczbowym symulującym eksperyment pomiarowy [7]. Z populacji generalnej o trapezowym PDF pobrano próbkę o $n = 200$ danych o wartościach losowych jak na rys. 5. Próbka była czysto losowa, gdyż nie modelowano w niej zmieniających się wraz z próbkowaniem błędów systematycznych. Sprawdzenie funkcji autokorelacji dla kilku sąsiednich obserwacji dało wynik negatywny. Przyjęto więc, że wpływ autokorelacji jest pomijalny. Przy tej liczbie $n = 200$ można stosować do obliczeń wzory podane wcześniej w tekście rozdziału.

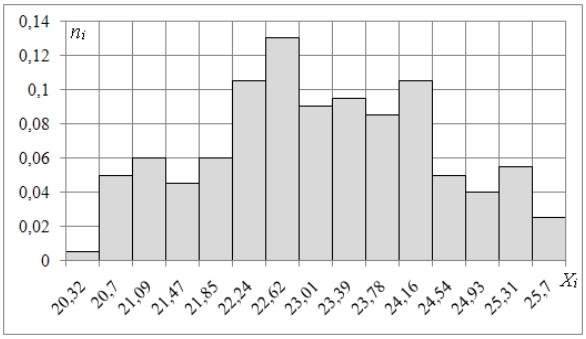
Wyznaczono dla próbki najlepszy estymator wartości mierzonej i jego niepewność standardową typu A oraz rozszerzoną przy pomijalnej niepewności typu B.



Rys. 5. Wartości obserwacji w próbce
Fig. 5. Values of sample observations

Dane próbki uszeregowano w histogram o 15 przedziałach (rys. 6) – przyjęto tu nieco większą liczbę przedziałów niż 9 wg wzoru Sturgessa, ale liczba pomiarów w przedziale jest wystarczająca – średnio około 13,3. W pierwszej kolejności wybie-

ra się najwłaściwszy model PDF dla próbki. Według kryterium χ^2 sprawdzono zgodność histogramu z trzema rozkładami teoretycznymi. Przy poziomie istotności 0,05 wyniki dla rozkładu normalnego i równomiernego były ujemne [7]. Metodą najmniejszych kwadratów wybrano trapezowy rozkład PDF o długości podstaw 1,79 i 5,38. Ich stosunek wynosi β = 1/3. Dla takiego trapezu, przy 11 stopniach swobody, wynik testu $\chi 2$ był pozytywny. Jeszcze korzystniejszy wynik zgodności dał test Kołmogorowa-Smirnowa dla dystrybuanty. Przy negatywnym wyniku trzeba zmniejszyć wymagania lub uzyskać próbkę o większej liczbie obserwacji n, a przy nieznanej populacji dobierać inny model rozkładu. Przedstawione tu postępowanie dotyczy przypadków, gdy liczba obserwacji $n > 10$ i znany jest, nawet tylko orientacyjnie, stosunek podstaw β.



Rys. 6. Histogram względnych częstości danych próbki
Fig. 6. Histogram of relative frequencies of the data sample

Dla symulowanego eksperymentalnego rozkładu próbki otrzymano następujące wartości estymatorów:

$$\bar{X} = 22{,}873\,, \quad q_{V/2} = 23{,}010, \quad \tilde{X} = 22{,}942$$

Odchylenie standardowe próbki: S_X = 1,309.

Odchylenie standardowe średniej: $S[\bar{X}] = S_x / \sqrt{n} = 0{,}0926$.

Odchylenie standardowe środka rozstępu próbki:

$$S\left[q_{V/2}\right] = \sqrt{\frac{V^2(1-\beta^2)(4-\pi)}{8n}} = 0{,}089$$

Odchylenie standardowe estymatora 2-elementowego

$$u_A(\tilde{X}) = \tfrac{1}{2}\sqrt{S^2[\bar{X}] + S^2[q_{V/2}] + 2 \cdot 0{,}2 \cdot S[\bar{X}] \cdot S[q_{V/2}]} = 0{,}0703$$

Niepewność standardowa dwuelementowego estymatora o współczynnikach według wzorów (5) niewiele różni się od wyznaczonej powyżej.

Współczynniki rozszerzenia estymatorów $q_{V/2}$ i $\tilde{X}$ nie są znane, ale ich rozkłady są nieco smuklejsze niż normalny, więc nie będą one większe niż dla średniej.

Dla wszystkich trzech estymatorów przyjęto prawdopodobieństwo $P = 0{,}95$ i jednakowe w przybliżeniu współczynniki rozszerzenia $k_P = k_{0,95} = 1{,}96$.

Rozszerzona niepewność estymatora $\tilde{X}$ wynosi $U_p(\tilde{X}) = k_p \cdot u_A(\tilde{X})$. Wyniki dla trzech estymatorów menzurandu zawiera tabela 2.

Tab. 2. Trzy estymatory wartości i niepewności wyniku pomiaru dla próbki o rozkładzie Trap
Tab. 2. Three estimators of measurement result and uncertainty for the Trap PDF sample

	Wartość wyniku i jego niepewność standardowa S	Wartość wyniku z niepewnością rozszerzoną oraz granice jej przedziału
$\bar{X}$	$X = 22{,}87;\ u_A = 0{,}09$	$X = (22{,}87 \pm 0{,}19),\ P = 0{,}95$ $X \in (22{,}68;\ 23{,}06),\ P = 0{,}95$
$q_{V/2}$	$X = (23{,}01;\ S[q_{V/2}] = 0{,}09$	$X = (23{,}01 \pm 0{,}18),\ P = 0{,}95$ $X \in (22{,}83;\ 23{,}19),\ P = 0{,}95$
$\tilde{x}$	$X = 22{,}94:\ S[\tilde{X}] = 0{,}07$	$X = (22{,}94 \pm 0{,}14),\ P = 0{,}95$ $X \in (22{,}87;\ 23{,}15),\ P = 0{,}95$

Zgodnie z przewidywaniami, najdokładniejszy jest ostatni wynik dla uproszczonego estymatora 2C o symbolu $\tilde{X}$, gdyż ma najmniejsze $u_A = S[\tilde{X}]$.

Można też zauważyć, że w tym przykładzie wartości menzurandu otrzymane dla różnych estymatorów różnią się niewiele i każdy mieści się w przedziale rozszerzonej niepewności dwu pozostałych.

Podsumowanie i wnioski

- Prawidłowe oszacowanie wyznaczanej statystycznie niepewności dla próbki danych o trapezowym rozkładzie PDF (odpowiednik u_A wg Przewodnika GUM) istotnie zależy od wyboru efektywnego, tj. możliwie najbardziej dokładnego jej estymatora.
- W pracach [7–10] wykazano, że wybór najlepszego jednoelementowego estymatora wartości menzurandu dla próbki o rozkładzie PDF modelowanym symetrycznym trapezem zależy od jego kształtu opisanego przez stosunek długości podstaw β. Jeśli kształt trapezu zbliżony jest do prostokąta, tj. w przedziale $1 \geq \beta \geq 0{,}35$ najlepszym estymatorem jest środek rozstępu próbki, zaś poniżej $\beta = 0{,}35$, aż do rozkładu trójkątnego o $\beta = 0$ lepsza jest wartość średnia.
- Jeszcze lepszy dla rozkładów trapezowych jest estymator dwuelementowy 2C jako forma liniowa średniej i środka rozpięcia.
- W szerokim zakresie kształtów trapezu ($0{,}75 \geq \beta \geq 0$) można stosować uproszczony estymator dwuelementowy wg wzoru (6) – o jednakowych współczynnikach $k_1 = k_2 = 0{,}5$. Jest on wystarczająco dokładny w praktyce, a jego odchylenie standardowe mało zależy od liczby pomiarów $n > 10$ i od stosunku podstaw $\beta \leq 0{,}55$ (rys. 3).
- Estymatory te można stosować też w metodzie Bootstrap [11, 12], która nie wymaga by wyznaczenie wyniku było poprzedzone wyborem modelu rozkładu dla próbki.

- Dla próbki o rozkładzie trapezowym przy liczbie danych $n \geq 10$ można przyjąć, że współczynniki jej estymatorów nie zależą od *n*. Próbki o mniejszym $n < 10$, trzeba modelować indywidualnie.
- Wszystkie wnioski zostały sprawdzone pozytywnie za pomocą symulacji Monte Carlo i na kilku przykładach numerycznych.
- Podane powyżej, inne niż wartość średnia, jedno- i dwuelementowe estymatory optymalne dla rozkładów trapezowych nie były omawiane w publikacjach [5, 6] o rozkładach trapezowych. Nie występują też w zaleceniach Przewodnika GUM o wyznaczaniu niepewności pomiarów. Jedynie zalecenia amerykańskiego instytutu metrologicznego NIST [2] dopuszczają inne formuły uznane w teorii prawdopodobieństwa. Wyniki prac przedstawiano częściowo na kilku metrologicznych konferencjach krajowych i międzynarodowych [8, 9, 12].
- Proponowany dla rozkładów trapezowych liniowy estymator 2C wg (4) i (5) oraz jego uproszczona wersja (6) o jednakowych współczynnikach 0,5 dla wartości średniej i środka rozpięcia można stosować nie tylko w praktyce pomiarowej dla przypadków nie objętych dotąd zaleceniami Przewodnika GUM i jego Suplementami, ale też i w innych użytkowych badaniach statystycznych, w których trapezowe modele rozkładów są często używane [5], np. dla liczby sprawnych urządzeń w przedsiębiorstwie, ludzi w miejscowości wg wieku.

Literatura

1. *Evaluation of measurement data – Guide to the expression of uncertainty in measurement* (GUM), BIPM, JCGM 100, Ed. 2008, and Supplement 1 – *Propagation of distributions using a Monte Carlo method* (Ed. 2007).
2. Taylor B.N., Kuyatt Ch.E., *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*, NIST Technical Note 1297, 1994.
3. Novickij P.V., Zograf I.A., *Ocenka pogreshnostiej resultatov izmierenii* (*Ocena błędów wyników pomiarów*), Energoatomizdat, Leningrad 1985.
4. Zakharov I.P., Shtefan N.V., *Algorithms for reliable and effective estimation of type A uncertainty*. "Measurement Techniques", Vol. 48, No. 5, 2005, 427–437, DOI: 10.1007/s11018-005-0160-7.
5. Van Dorp J.R., Kotz S., *Generalized Trapezoidal Distributions*. "Metrika", Vol. 58, Issue 1, July 2003.
6. Kacker R.N., Lawrence J.F., *Trapezoidal and triangular distributions for Type B evaluation of standard uncertainty*. "Metrologia", Vol. 44, 2007, 117–127.
7. Warsza Z.L., Galovska M., *About the best measurand estimators of trapezoidal probability distributions.* "Przegląd Elektrotechniczny – Electrical Review", Nr 5, 2009, 86–91.
8. Warsza Z.L., Galovska M., *The best measurand estimators of trapezoidal PDF*. Proceedings of IMEKO World Congress "Fundamental and Applied Metrology", September 2009, Lisbon Portugal, 2405–2410.
9. Warsza Z.L., Galovskaja M.V., *Vybor najlutshej ocenki izmierajemoj velichiny na primiere trapecievidnych raspredelenij*, "Sistemy Obrobotki Informacii" 4(78), Kharkow 2009, 28–31.

10. Warsza Z.L., Galovska M., *Estymatory wartości menzurandu dla trapezowych rozkładów prawdopodobieństwa danych pomiarowych*. „Pomiary Automatyka Komputery w Gospodarce i Ochronie Środowiska", Nr 1, 2010, 12–17.
11. Galovska M., Warsza Z.L., *The ways of effective estimation of measurand.* „Pomiary Automatyka Komputery w Gospodarce i Ochronie Środowiska", Nr 1, 2010, 18–20.
12. Galovska M., Warsza Z.L., *Estymatory wartości menzurandu próbek danych o rozkładach niegaussowskich*, Metrologia dziś i jutro, Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 2010, 59–72.
13. Endovitskyi P., Ukraiński Narodowy Uniwersytet Techniczny „Kijowski Instytut Politechniczny", praca nieopublikowana.
14. Warsza Z.L., *Dwuelementowe estymatory wartości mezurandu próbek danych pomiarowych o trapezowych rozkładach prawdopodobieństwa – przegląd prac*, „Pomiary Automatyka Kontrola", R. 57, Nr 1, 2011, 105–108.
15. Warsza Z.L., *Effective Measurand Estimators for Samples of Trapezoidal PDFs.* "Journal of Automation, Mobile Robotics and Intelligent Systems", Vol. 6, No. 1, 2012, 35–41.
16. Warsza Z.L., *Wyznaczanie niepewności próbek pomiarowych o kilku niegaussowskich rozkładach prawdopodobieństwa*. „Elektronika: konstrukcje, technologie, zastosowania" (wkładka: *Technika Sensorowa*), Vol. 53, Nr 9, 2012, 134–139.

Rozdział 9

Wybór estymatora menzurandu metodami Monte Carlo i resamplingu

Streszczenie. Podstawowym celem rozdziału jest rozszerzenie sposobów wyboru estymatora danych empirycznych, gdy dane te są jedyną dostępną informacją, tj. gdy nieznany jest rozkład ich populacji generalnej. Przedstawiono metodę opartą na symulacji Monte Carlo i dwie metody powtórnego próbkowania, czyli resamplingu o angielskich nazwach *jackknife* (wycinania) i *bootstrap*. Ich przydatność do uzyskiwania lepszego niż wg GUM estymatora menzurandu, tj. o możliwie najmniejszym odchyleniu standardowym, zilustrowano przykładami liczbowymi.

Prawidłowe statystyczne szacowanie jest niezmiernie istotne w przetwarzaniu danych pomiarowych. Przeprowadzone niewłaściwie, spowoduje błędną ocenę wartości i niepewności pomiaru menzurandu i innych jego parametrów. Głównym celem niniejszego rozdziału jest przedstawienie możliwości wyboru najlepszego, tj. o najmniejszym odchyleniu standardowym, estymatora wartości mierzonej dla próbki danych empirycznych w przypadkach, gdy stanowi ona jedyną dostępną informację.

Estymatorem menzurandu (wartości mierzonej) o najmniejszym odchyleniu standardowym, wyznaczanym z próbki o obserwacjach pomiarowych pobranych z populacji o rozkładzie normalnym, jest wartość średnia. Jej odchylenie standardowe (SD) nazwano w Przewodniku GUM [1] niepewnością typu A. Klasyczny sposób sprawdzenia, czy dla próbki o n danych hipoteza o normalności rozkładu jej populacji generalnej jest do przyjęcia, polega na zastosowaniu dla jej histogramu jednego z testów zgodności, np. testu $\chi 2$ dla $v = n - 1$ stopni swobody, przy zadanym dopuszczalnym poziomie istotności α. Przy wyniku pozytywnym, jako estymatora menzurandu używa się następnie średniej arytmetycznej oraz jej odchylenia standardowego, czyli niepewności typu A jako składnika w procedurze oceny niepewności całkowitej wyniku pomiaru (patrz rozdział 1).

Rozkład normalny jest jednakże modelem teoretycznym, który realizuje się w rzeczywistości z przybliżeniem, gdyż liczba źródeł oddziaływania losowego jest skończona. Rozrzut wartości obserwacji pomiarowych zachodzi w ograniczonym przedziale i zwykle lepiej odwzorowują to inne, tj. niegaussowskie teoretyczne modele rozkładu prawdopodobieństwa [1].

W rozdziałach 5–8 tej monografii przedstawiono wyniki szczegółowej analizy estymatorów parametrów menzurandu dla próbek danych o różnych modelach rozkładu gęstości prawdopodobieństwa (PDF). Na podstawie prac [1, 2] omówiono m.in. rozkłady trapezowe, które stanowią splot dwu rozkładów równomiernych często występujących w sygnałach elektronicznych systemów pomiarowych. Wykazano, że wybór najdokładniejszego wśród estymatorów jednoskładnikowych (1C)

menzurandu zależy od kształtu PDF populacji możliwych wyników obserwacji. Dla rozkładu o symbolu Trap [4], czyli o PDF w postaci symetrycznego trapezu o bokach liniowych, gdy jego kształt jest bliski prostokąta, tj. dla zakresu $1 \geq \beta \geq 0{,}35$, (gdzie: $\beta = a/b$ – stosunek krótszej górnej i dolnej podstawy), najlepszym estymatorem 1C jest środek rozpięcia próbki. Poniżej $\beta = 0{,}35$ aż do $\beta = 0$ (dla rozkładu trójkątnego) lepsza jest wartość średnia. Wartości obu tych estymatorów dla $n > 10$ można już traktować w praktyce jako niezależne od n.

W pracach [2, 3] wykazano też, że dla próbek o trapezowych rozkładach PDF jeszcze lepszym niż powyższe estymatory 1C jest estymator dwuskładnikowy (2C) w postaci liniowej kombinacji dwu poprzednich o optymalnych wartościach współczynników zależnych od stosunku podstaw β trapezu. Dla szerokiego zakresu β zaproponowano nowy estymator 2C o formie uproszczonej, tj. o jednakowych współczynnikach proporcjonalności 0,5 dla obu składników (rozdział 8). Ten estymator 2C różni się od najlepszych 2C bardzo niewiele i może być stosowany w praktyce z wystarczającą do przyjęcia dokładnością.

Przy stosowaniu w pomiarach rozkładów niegaussowskich występuje też inny ważny w praktyce problem. Jest to trudność w identyfikacji właściwej funkcji rozkładu prawdopodobieństwa (PDF) dla próbki w przypadkach, gdy test zgodności wypada pozytywnie dla kilku różnych modeli rozkładów. Na podstawie tego testu brak wtedy podstaw do podjęcia decyzji, który z nich wybrać. Należy też dla najlepszych estymatorów środka grupowania wartości rozpatrywanych rozkładów znać równania odchyleń standardowych (niepewności standardowej), uwzględniające liczbę obserwacji n w próbce.

Niniejszy rozdział poświęcony jest podejściu numerycznemu, które nie wymaga wyboru z góry określonego teoretycznego modelu rozkładu. Zostaną rozpatrzone przykład, w których rozważy się i przetestuje algorytmy umożliwiające taką realizację. Dzielą się one na dwa grupy: symulacja metodą Monte Carlo oraz metodą resamplingu, czyli powtórnego próbkowania danych próbki.

Modelowanie metodą Monte Carlo

Dane pierwotne występują zwykle w postaci próbki liczącej n wartości obserwacji, lub mogą być zgrupowane w histogram. Stosowana tu metoda modelowania Monte Carlo polega na tworzeniu z tych danych innych próbek oraz pseudo-populacji o cechach statystycznych korespondujących z danymi pierwotnymi.

Działania należy zacząć od utworzenia listy rozpatrywanych estymatorów. Na przykład mogą to być to estymatory jedno-składnikowe: średnia arytmetyczna $\overline{X}$, środek rozstępu $q_{V/2}$ i mediana X_{med}. Następnie, na podstawie próbek o powielonych danych oblicza się wartości tych estymatorów i ich odchylenia standardowe (SD) oraz wybiera się z nich estymator o najmniejszym SD.

Można też rozpatrywać estymatory liniowe trzyskładnikowe o postaci:

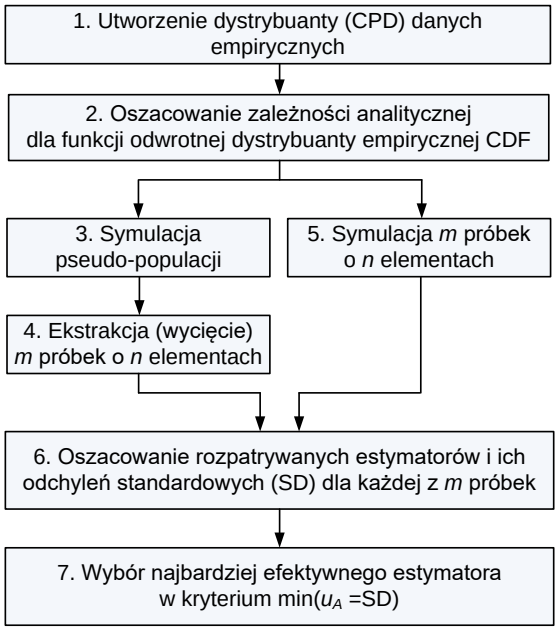
$$\hat{X} = k_1 \overline{x} + k_2 X_{med} + k_3 q_{V/2} \tag{1}$$

oraz ich przypadki szczególne dwuskładnikowe. Współczynniki wyznacza się z warunku optymalizacji $\min\left\{SD\left(\hat{X}\right)\right\}$.

Istnieje kilka sposobów symulacji rozkładu dla danych pierwotnych.

Symulacja oparta na danych niezgrupowanych i aproksymacji funkcji odwrotnej

Do oceny niepewności można stosować bezpośrednio metodę funkcji odwrotnej, jak to proponuje się w [4]. Jej algorytm podano na rys. 1.



Rys. 1. Metoda oparta na symulacji Monte Carlo
Fig. 1. Method based on Monte Carlo simulation

A oto zalecenia dalszego postępowania.

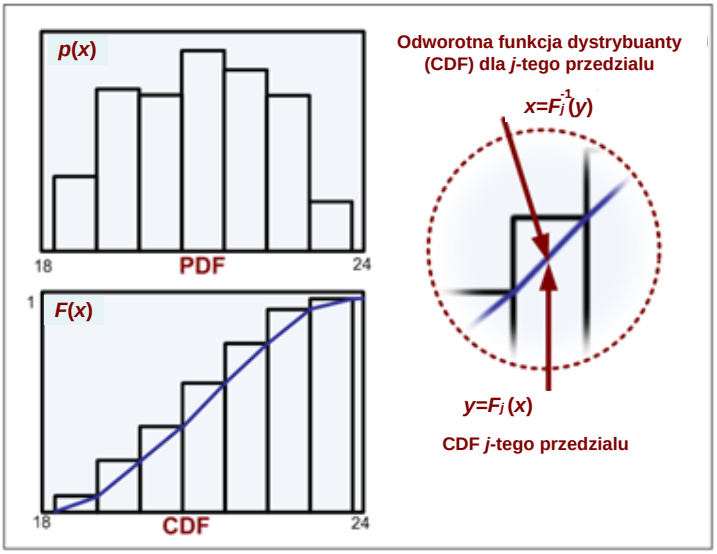
- W większości przypadków analityczne znalezienie funkcji odwrotnej nie jest możliwe i lepiej dokonywać jej aproksymacji dla dystrybuanty (CDF).
- Model aproksymowanej krzywej nie musi być wielomianem. Powinno się go wybierać wg kryterium najmniejszego błędu lub wg dokładności aproksymowanej wartości. Preferuje się aproksymację wielomianową tylko dlatego, że występuje ona w wielu pakietach oprogramowania.
- Dla skomplikowanych estymatorów należy stosować procedurę szacowania współczynników (wag) składników.

Jakość rozważanej procedury określa się za pomocą:

- liczby próbek *m*,
- dokładności aproksymacji funkcji odwrotnej CDF,
- wykazu wybranych estymatorów lub procedury szacowania ich współczynników (wag).

Symulacja oparta na zgrupowanych danych i odcinkami linearyzowanej funkcji odwrotnej

W tym przypadku dystrybuanta CDF jest reprezentowana przez odcinkowo-liniową aproksymację (rys. 2) i nie ma konieczności stosowania metody najmniejszych kwadratów. Symulacja wykonywana jest metodą funkcji odwrotnej dla każdego przedziału. Funkcję odwrotną do liniowej wyznaczają dwa punkty skumulowanych częstości względnych. W kolejnych postępowaniach powtarza się tę operację.



Rys. 2. Zgrupowane dane i linearyzowana odcinkami funkcja odwrotna
Fig. 2. Grouped data and inverse piecewise linear function

Symulacja na podstawie parametrów kształtu i uogólnienia rozkładu

Większość modeli PDF używanych do oceny niepewności daje się uogólnić. Stosuje się przede wszystkim system krzywych Pearsona [4] oraz inne systemy: Johnsona, Burra i innych. W zależności od wartości parametrów próbki odpowiadających czterem pierwszym centralnym momentom, możemy dopasować formę krzywej Pearsona i zasymulować rozkład. Istnienie wiele uogólnionych rozkładów, np. rozkłady *gh* i rozkłady λ Tukeya [4], które umożliwiają uproszczenie tego podejścia.

Metody resamplingu

Metody resamplingu są jednym z najnowszych sposobów szacowania statystycznego, opartego na realizacji komputerowej [5]. Istnieją dwa sposoby jego realizacji.

Metoda „jackknife" (z wycinaniem)

Polega na usuwaniu elementów z otrzymanego zbioru początkowych danych i ponownym obliczaniu estymatora.

Dla *i*-tej replikacji „jackknife” (tj. po usunięciu *i*-tego elementu):

$$\tilde{\theta}_i = f(x_1, x_2 ..., x_{i-1}, x_{i+1}, ..., x_n) \tag{2}$$

Odchylenie standardowe estymatora „jackknife” z replikacji *m* wynosi:

$$SD(\tilde{\theta}) = \sqrt{\frac{1}{m-1}\sum_{i=1}^{m}\left(\tilde{\theta}_i - \bar{\tilde{\theta}}\right)^2} \qquad \text{gdzie } \bar{\tilde{\theta}} = \frac{1}{m}\sum_{i=1}^{m}\tilde{\theta}_i \,. \tag{3}$$

W metodzie „jackknife” liczba *m* tworzonych próbek jest ograniczona przez liczbę obserwacji *n*, tj. $m \le n-1$.

Metoda „bootstrap”

W metodzie „bootstrap” tworzy się zbiór próbek przetworzonych z próbki danych pierwotnych. Każdą z nich otrzymuje się przez losowe wybieranie elementów z próbki oryginalnej (ekstrakcja z jednakowym prawdopodobieństwem $1/n$). Te nowe próbki – „resample” uzupełniają próbkę oryginalną. Algorytm tej metody resamplingu podano na rysunku 3.

Rys. 3. Algorytm metody resamplingu
Fig. 3. Algorithm of the resampling method

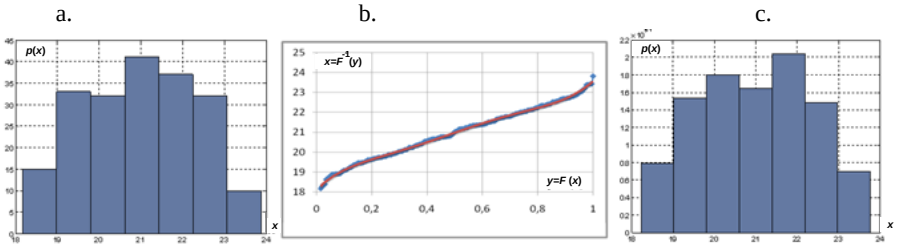
Metodę można stosować, gdy obserwacje są niezależne i jednakowo rozłożone Jej dokładność zależy od błędów statystycznych i błędu modelowania. Przetwarzanie oryginalnej próbki można wykonać tyle razy, na ile pozwoli moc obliczeniowa komputera.

Przykład

Histogram zasymulowanego rozkładu o liczbie elementów n=200 podano na rys. 4a.

Symulacja metodą Monte Carlo

Dokonano symulacji Monte Carlo dla danych niezgrupowanych. Aproksymację funkcji odwrotnej CDF pokazano na rys. 4b. Próbkę zasymulowano M = 1 000 000 razy w oparciu o znalezioną funkcję odwrotną. Jest to tak zwana pseudo-populacja. Jej histogram jest na rysunku 4c. Dla ustalenia zgodności z pierwotnymi danymi sprawdzono i porównano wartości niektórych parametrów (tabela 1).



Rys. 4. Symulacja metodą Monte Carlo rozkładu empirycznego: a. histogram danych pierwotnych, b. aproksymacja funkcji odwrotnej CDF, c. histogram „pseudo-populacji" (z symulacji metodą Monte Carlo)

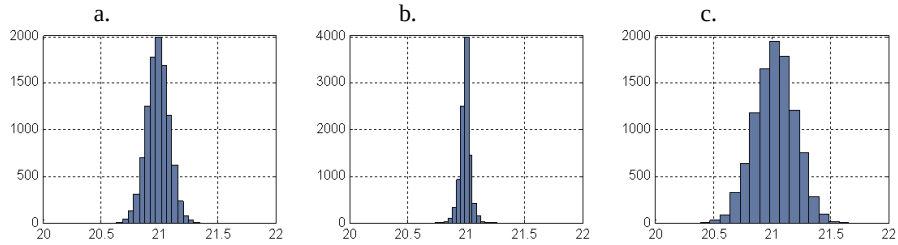
Fig. 4. Monte Carlo simulation of empirical distribution: a. Histogram of initial data, b. Inverse CDF approximation, c. Histogram of pseudo-population (by Monte Carlo simulation)

Tab. 1. Porównanie wybranych parametrów próbki otrzymanych metodami MC

Tab. 1. Comparison of some parameters of sample obtained by MC method

Parametr	Dane początkowe	"pseudo-populacja"	"pseudo-populacja" (z grupowaniem)
Dyspersja	1,82	1,82	1,91
Asymetria	–0,02	–0,03	–0,042
Kurtoza	2,08	2,05	2,07

Wyniki zastosowania tej metody MC podano na rysunku 5 i w tab. 2.



Rys. 5. Rozkłady: a. wartości średniej, b. środka rozpięcia i c. mediany próbki danych

Fig. 5. Distributions of sample mean, midrange and median

Z przedstawionych na rysunku 5 histogramów estymatorów jednoelementowych najbardziej efektywny jest środek rozstępu, gdyż ma najmniejsze odchylenie standardowe (jest najwęższy).

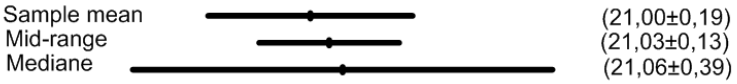
Porównanie wyników symulacji metodą MC i bootstrap

Parametry próbki o znanych danych początkowych (kolumna 3) następnie dla sprawdzenia wyznaczano metodami Monte Carlo oraz metodą bootstrap. Rezultaty symulacji podane są w tabeli 2.

Tab. 2. Porównanie kilku estymatorów jednoelementowych
Tab. 2. Comparison of few one-component estimators

Estymator	Parametr	Dane początkowe	Symulacja bez grupowania m = 10 000	Symulacja z grupowaniem m = 1000 000	Metoda bootstrap (m = 200)
średnia	średnia	21,000	20,984	20,978	21,001
	SD	0,0957	0,095	0,097	0,096
środek rozstępu		21,0275	20,994	21,014	21,035
	SD	–	**0,046**	**0,048**	**0,066**
mediana		21,1190	21,023	21,016	21,059
	SD	–	0,170	0,140	0,200

Wartości SD odchyleń standardowych trzech estymatorów jednoelementowych uzyskane każdą z metod symulacji różnią się. Jednak wykazują, że środek rozstępu jest dla tej próbki najbardziej efektywny, gdyż ma najmniejsze SD. Na rysunku 6 przedstawiono otrzymane metodą bootstrap istotne różnice między rozszerzoną niepewnością wartości średniej i długością równoważnych jej przedziałów dla środka rozstępu i mediany dla p = 0,95.

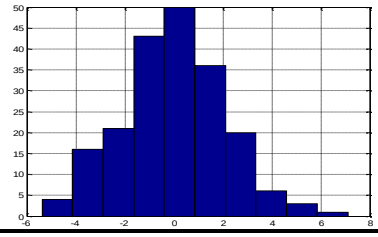
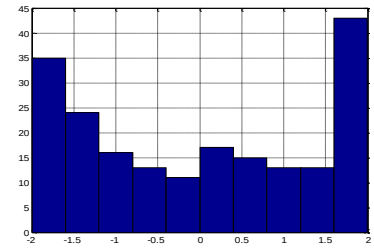
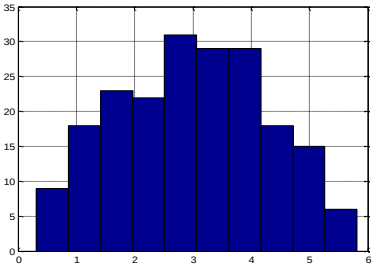


Rys. 6. Niepewność rozszerzona średniej i równoważne jej przedziały środka rozstępu i mediany dla p = 0,95
Fig. 6. Expanded uncertainty of mean and coverage intervals of midrange and median for p = 0.95

Wyznaczenie najlepszych estymatorów próbek bez konieczności poszukiwania ich rozkładów

W tabeli 3 zamieszczono wyniki wyznaczenia metodami MC i bootstrap parametrów kilku estymatorów trzech różnych próbek o histogramach podanych w kolumnie pierwszej. Dla każdej próbki pogrubiono najlepszy z rozpatrywanych estymatorów, tj. o najmniejszym odchyleniu standardowym.

Tab. 3. Porównanie rezultatów dla kilku próbek
Tab. 3. Comparison of results for few samples

Dane pierwotne (n = 200)	u_A wg GUM	Estymator	SD estymatora	
			symulacja MC wg Pearsona	resampling (bootstrap)
Próbka z populacji o rozkładzie normalnym 	0,14	**średnia**	**0,15**	**0,15**
		środek rozstępu	0,70	0,60
		mediana	0,18	0,19
Próbka z populacji o rozkładzie typu U 0.0994 	0,10	średnia	0,10	0,10
		środek rozstępu	**0,0014**	**0,0010**
		mediana	0,21	0,24
Próbka z populacji o rozkładzie trapezoidalnym (stosunek podstaw β = 0,35) 	0,09	średnia	0,09	0,09
		środek roz-stępu	0,09	0,10
		mediana	0,13	0,13
		Estymator 2C $\left(\frac{\bar{X} + q_{V/2}}{2}\right)$	**0,070**	**0,070**

Wyniki symulacji z tabeli 3 potwierdzają znane wcześniej właściwości estymatorów dla rozkładu normalnego i typu U oraz estymatorów zbadanych po raz pierwszy w pracach [2, 3] dla rozkładów trapezowych. Najlepszymi z nich estymatorami, tj. o najmniejszym odchyleniu standardowym SD dla próbek z populacji o rozkładach rozpatrzonych w tym przykładzie, są:

– dla rozkładu normalnego – średnia arytmetyczna,

– dla rozkładu typu U – środek rozstępu,
– dla rozkładu trapezowego Trap – estymator 2C zaproponowany przez Warszę i Galovską w [2, 3] (rozdział 8).

Istnieje ścisły dowód analityczny dla rozkładu normalnego [8].

Wnioski

Rozpatrzono metody poszukiwania najlepszego estymatora menzurandu i jego niepewności dla próbki danych pomiarowych. Procedury Monte Carlo dają dobre wyniki, ale wymagają próbki danych pierwotnych o odpowiednio dużej liczebności do uzyskania dobrego przybliżenia przy tworzeniu funkcji odwrotnej do jej dystrybuanty (CDF) i histogramu.

Nieklasyczne podejście – metody resamplingu, nie wymagają próbek o dużych rozmiarach. Istnieją dwie różne ich procedury: jackknife i bootstrap. Procedury te wymagają dodatkowych operacji algorytmicznych dla aproksymacji funkcji odwrotnej, w tym przez odcinki linii prostej. Metodą bootstrap można wyprodukować więcej próbek wtórnych – resampli, ale wymaga ona dodatkowego oprogramowania do realizacji pseudo-losowej populacji danych.

Powyższe algorytmy można wykorzystać w badaniach różnych obiektów dla tworzenia wielu różnych estymatorów wartości menzurandu, w tym estymatorów w postaci formy liniowej o kilku elementach oraz w analizie niepewności pomiaru próbek z populacji o nieznanym rozkładzie.

Literatura

1. Dorozhovets M., Warsza Z.L., *Udoskonalenie metod wyznaczania niepewności wyników pomiaru w praktyce*. „Przegląd Elektrotechniczny”, Nr 1, 2007, 1–13.
2. Warsza Z., Galovska M., *The best measurand estimators of trapezoidal distributions*, Proceedings of XIX IMEKO Congress, Lisbon Sept. 2009, 2405–2410.
3. Warsza Z.L., Galovska M., *Estymatory wartości menzurandu dla trapezowych rozkładów prawdopodobieństwa danych pomiarowych.* „Pomiary Automatyka Komputery w Gospodarce i Ochronie Środowiska”, Nr 1, 2010, 12–16.
4. Supplement 1 to the Guide to the expression of uncertainty in measurement (GUM) – Propagation of distributions using a Monte Carlo method, Guide OIML G 1-101 Edition 2007 (E).
5. Sim C. H., M.H.Lim: Evaluating expanded uncertainty in measurement with a fitted distribution, *Metrologia* 45 (2008), 178–184.
6. Michael R. Chernick. *Bootstrap methods: a guide for practitioners and researchers*, 2nd ed., Wiley, 2007, p. 369.
7. Galovska M., Warsza Z.L., *The ways of effective estimation of measurand*. „Pomiary Automatyka Komputery w Gospodarce i Ochronie Środowiska”, Nr 1 2010, 18–20.
8. Mańczak K., *Technika planowania eksperymentu*. Wydawnictwa Naukowo-Techniczne, Warszawa 1976.

Rozdział 10

Wielomianowy estymator parametrów menzurandu o symetrycznych rozkładach niegaussowskich oparty na kumulantach danych

Streszczenie. Przedstawiono zastosowanie metody maksymalizacji wielomianu o angielskim akronimie PPM do oceny wartości i niepewności menzurandu dla symetrycznych niegaussowskich próbek danych pomiarowych o nieznanym *a priori* rozkładzie PDF ich populacji. Metoda ta opiera się na zastosowaniu wielomianów stochastycznych i wykorzystuje statystykę wyższego rzędu - opis parametrów losowych przez momenty i kumulanty. Podano wyrażenia analityczne umożliwiające estymację parametrów menzurandu i analizę ich dokładności dla wielomianu stopnia $s = 3$. Wykazano, że niepewność menzurandu oszacowana z danych próbki metodą PPM jest zwykle mniejsza niż wyznaczona klasycznie wg Przewodnika GUM. Stosunek obu estymat niepewności zależy od wartości współczynników kumulantów wyższego rzędu i charakteryzuje stopień odstawania rozkładu danych próbki od modelu Gaussa. Wyniki modelowania statystycznego metodą Monte Carlo potwierdziły skuteczność proponowanego podejścia.

Wynik pomiarów menzurandu uzyskuje się przez analizę statystyczną zbioru wartości wielu powtarzanych obserwacji pomiarowych, czyli próbki pomiarowej. Właściwy wybór metod i algorytmów przetwarzania wartości obserwacji wymaga przyjęcia model matematycznego, który adekwatnie opisuje ich rozproszenie. Dla menzurandu o stałej wartości zwykle używa się modelu, który charakteryzuje się estymatorem tej wartości o najmniejszym odchyleniu standardowym, i rozkładem o symetrycznej funkcji gęstości prawdopodobieństwa (PDF) wokół tej wartości. Wybór estymatora zależy od rodzaju tego rozkładu. Powinno się wybrać model rozkładu, który jak najlepiej przybliża dane doświadczalne. W przewodniku wyznaczania niepewności pomiarów GUM estymatorem menzurandu jest wartość średnia próbki, która jest najlepszym estymatorem tylko dla rozkładu normalnego (Gaussa) [1].

W praktyce najczęściej rodzaj rozkładu danych pomiarowych nie jest znany *a priori*. Badania eksperymentalne dotyczące identyfikacji i analizy rozkładów dla rzeczywistych danych wykazały, że funkcji Gaussa nie można uznać za uniwersalne narzędzie, nadające się bez ograniczeń do opisu wszystkich rodzajów błędów w systemach pomiarowych. W praktyce, warunki zastosowania centralnego twierdzenia granicznego zwykle nie są spełnione, gdyż liczba składników wpływów losowych jest ograniczona. Ponadto w wielu przyrządach elektronicznych może wystąpić jakiś rodzaj nieliniowości charakterystyki, taki jak ograniczenie liniowego zakresu amplitudy sygnału, nieliniowość kwantyzacji itp. Dlatego do opisania danych pomiarowych używa się też inne rozkłady, np. rozkład równomierny, trójkątny, arc sin, różne rozkłady trapezowe itp. [1b, 2].

Metody klasyczne bazują na pojęciu gęstości prawdopodobieństwa i nazywa się je metodami parametrycznymi. Główne trudności występujące przy stosowaniu tych metod dotyczą konieczności posiadania informacji a priori o formie rozkładu, a także potencjalnie wysokiego stopnia złożoności przy ich realizacji i analizie właściwości. Dlatego opracowywane są inne metody statystyczne, które pozwalałyby zminimalizować wymaganą *a priori* ilość informacji. Jednym z alternatywnych sposobów przetwarzania danych niegaussowskich jest wykorzystanie statystyki wyższego rzędu – opis przez momenty lub kumulanty [3]. W tym rozdziale omawia się wyniki wspólnej pracy z S. Zabolotniim [15].

Poniżej zostaną przedstawione podstawowe równania opisujące ogólne kumulanty i ich związki ze statystycznymi momentami rozkładów prawdopodobieństwa.

Funkcję charakterystyczną zmiennej losowej ξ otrzymuje się przez przekształcenie Fouriera o postaci

$$f_\xi(u) = \frac{1}{2\pi} \int_{-\infty}^{\infty} p_\xi(x) e^{ju\,x} dx$$

gdzie: p – gęstość prawdopodobieństwa; u, x – zmienne pomocnicze, $j = \sqrt{-1}$.

Momenty m_i wyznacza się ze znanego wzoru

$$m_i = \int_{-\infty}^{\infty} x^i \, p_\xi(x) dx = j^{-i} \left[\frac{d^i}{du^i} f_\xi(u) \right]_{u=0}$$

Kumulanty χ_i otrzymuje się po rozwinięciu funkcji $f_\xi(u)$ w szereg Taylora-Maclaurina. Są one opisane następującymi wzorami:

$$\chi_i = j^{-i} \left[\frac{d^i \ln f_\xi(u)}{du^i} \right]_{u=0}.$$

Między kumulantami i momentami istnieje jednoznaczny związek, na przykład

$$\chi_1 = m_1,\ \chi_2 = \mu_2,\ \chi_3 = \mu_3,\ \chi_4 = \mu_4 - 3\mu_2^2,\ \chi_5 = \mu_5 - 10\mu_2\mu_3,$$
$$\chi_6 = \mu_6 - 15\mu_2\mu_4 - 10\mu_3^2 + 30\mu_2^3 \ \ldots$$

gdzie: m_1 – moment pierwszego rzędu, μ_i – momenty centralne rzędu i.

W analizie właściwości zmiennych losowych używa się też współczynników kumulant jako parametrów bezwymiarowych następująco znormalizowanych względem wariancji, czyli kumulanty rzędu drugiego χ_2:

$$\gamma_i = \chi_i \Big/ \sqrt{\chi_2^i}$$

Najbardziej użyteczne są współczynniki asymetrii γ_3 i kurtozy γ_4. Momenty i kumulanty kilku podstawowych rozkładów podano w tabeli 5 na końcu rozdziału.

Opis z użyciem kumulant ma szereg właściwości, które pozwalają skutecznie rozwiązywać różne zadania statystyczne. Główne zalety to:

- odrębny i niezależny sens statystyczny kumulant, które do pewnego stopnia można rozpatrywać jako niezależne od siebie;
- dla rozkładu Gaussa wszystkie kumulanty rzędu $r \geq 3$ są równe zero. Używając pewnej liczby kumulant (lub ich współczynników) rzędu wyższego niż 3, można opisywać wymagany stopień niegaussowości zmiennych losowych;
- w przeciwieństwie do innych statystyk wyższego rzędu, na przykład momentów, skończonemu zbiorowi kumulant zawsze odpowiada jednoznacznie funkcja rzeczywista, która aproksymuje rozkład prawdopodobieństwa;
- w praktyce bardzo ważną właściwością jest inwariantność kumulanty względem zmian początku i zakresu skali zmiennych losowych;
- ponadto występuje właściwość kompozycji, która polega na tym, że kumulanta *i*-tego rzędu sumy statystycznie niezależnych zmiennych losowych jest sumą kumulant tego samego rzędu wszystkich składowych [4].

Przykładem wykorzystania zalet opisu z użyciem kumulantów w rozwiązywaniu zadań metrologicznych jest zastosowanie podanej w [5] metody kurtozy do obliczenia błędu pomiaru sumy składowych. Jednak zastosowanie tego algorytmu ogranicza się do danych opisanych przez niektóre rozkłady o symetrycznych PDF. Ponadto jest to metoda graficzno-analityczna, trudna do automatyzacji.

Udoskonaloną, łatwą do zautomatyzowania metodę kurtozy podano w pracy [6]. Jest ona uniwersalnym narzędziem do szacowania niepewności rozszerzonej dla szerokiej klasy niepewności typu B wraz z udziałem niepewności typu A przy dowolnym stopniu swobody. Modele addytywnych i multiplikatywnych błędów dodatkowych w torach pomiarowych z użyciem kumulant analizowano w pracy [7]. Rolę współczynnika kumulanty 4. rzędu (kurtozy) jako jednej z najbardziej istotnych cech rozkładów symetrycznych badano szczegółowo [8]. Poniżej, do opisu właściwości symetrycznych rozkładów danych pomiarowych, poza kurtozą, zostanie zastosowana kumulanta 6. rzędu. Przy niewielkiej korelacji zakłada się współczynniki $r_{ij} = 0$.

Cel badań

Poniżej przedstawi się rezultaty wspólnych z S. Zabolotniim badań dotyczących możliwości zastosowania niekonwencjonalnej semiparametrycznej metody statystycznej o akronimie PPM do szacowania estymat opisujących wartość i niepewność wyniku pomiaru menzurandu na podstawie danych próbki pomiarowej [18]. To narzędzie matematyczne o angielskiej nazwie *Polynomial Maximization Method* i jej akronim zaproponował Yurij Kunchenko w monografii [9]. Wykorzystuje ono wielomiany stochastyczne. W języku polskim proponuje się stosować nazwę: Metoda Maksymalizacji Wielomianu, lub w skrócie: metoda wielomianowa PMM. Metoda ta pozwala na ocenę właściwości probabilistycznych rozkładów jednomodalnych odbiegających od funkcji Gaussa i określonych przez wartości współczynników kumulant 3. rzędu lub wyższych. Tego rodzaju właściwości probabilistyczne wraz z dodatkową nieliniową transformacją danych mogą poprawić ocenę niepewności wyniku pomiarów, gdyż estymata wariancji wyznaczana metodą PPM jest mniejsza niż wg GUM. Zachętę do podjęcia tematu stanowiły wyniki wcześniejszych prac S. Zabolotniiego dotyczących zastosowania metody PMM. Opracował on adapta-

cyjne algorytmy do oszacowania *a posteriori* (off-line) położenia punktu skokowej zmiany (*Change Point*) wartości średniej lub wariancji w nie-gaussowskim szeregu losowym otrzymanym przez próbkowanie badanego procesu lub sygnału [10, 15].

Niniejszy rozdział ma następujące cele szczegółowe:

- zastosowanie metody maksymalizacji wielomianu PPM w syntezie algorytmów do estymacji niepewności menzurandu na podstawie serii danych pomiarowych z populacji o symetrycznym jednomodalnym rozkładzie niegaussowskim,
- analizę teoretyczną dokładności estymat wielomianu,
- badanie skuteczności tych algorytmów przez modelowanie metodą Monte Carlo.

Matematyczne sformułowanie problemu

Mierzoną wartość a wyznacza się jako wynik z serii powtarzanych obserwacji pomiarowych, których dane zawierają błędy przypadkowe. Zbiór wyników pomiarów należy interpretować jako próbkę $\vec{x} = \{x_1, x_2, ...x_n\}$ składającą się z n niezależnych i jednolicie losowo rozłożonych wartości danych, opisanych przez model $\xi = a + \xi_0$. W modelu tym występuje stały składnik a = const (wartość mierzona) oraz symetrycznie rozłożony składnik ξ_0 jako zmienna losowa o wartości średniej równej zeru i o rozkładzie prawdopodobieństwa różniącym się od normalnego (błędy przypadkowe). Właściwości probabilistyczne składnika ξ_0 zostaną opisane przez kumulantę (wariancję) χ_2 i współczynniki kumulant γ_4 i γ_6. Należy na podstawie zbioru wartości obserwacji w próbce $\vec{x}$ znaleźć estymator $\hat{a}$ parametru a jako wartość mierzoną i estymator jej odchylenia standardowego $\sigma_{(a)s}$, czyli niepewność standardową u_A.

Metoda maksymalizacji wielomianu (PMM)

Według metody PMM [9], oszacowanie parametru ϑ jest rozwiązaniem następującego wielomianowego równania stochastycznego tego parametru:

$$\sum_{i=1}^{s} h_i(\vartheta)\left[\frac{1}{n}\sum_{v=1}^{n} x_v^i - \alpha_i(\vartheta)\right]\Bigg|_{\vartheta=\hat{\vartheta}} = 0 \tag{1}$$

gdzie: s – stopień równania użytego do estymacji, $\alpha_i(\vartheta)$ – teoretyczne wartości momentów początkowych i-tego rzędu, $h_i(\vartheta)$ – współczynniki dla minimum wariancji.

Współczynniki $h_i(\vartheta)$ (dla $i = \overrightarrow{1,s}$) znajduje się rozwiązując układ liniowych równań algebraicznych odpowiedniego stopnia s, otrzymany dla warunków minimalizacji wariancji parametru ϑ, a mianowicie:

$$\sum_{i=1}^{s} h_i(\vartheta)\, F_{i,j}(\vartheta) = \frac{d}{d\vartheta}\alpha_j(\vartheta), \quad j = \overrightarrow{1,s}, \tag{2}$$

gdzie: $F_{i,j}(\vartheta) = \alpha_{i+j}(\vartheta) - \alpha_i(\vartheta)\,\alpha_j(\vartheta)$ – parametr nazwany przez Kunchenkę korrelantem centralnym o wymiarach i, j [9].

Kunczenko wykazał też, że wielomianowe estymaty $\hat{\vartheta}$, będące rozwiązaniami równań stochastycznych (1), są zgodne i asymptotycznie nieobciążone.

Układ równań (1) rozwiązuje się analitycznie metodą Kramera, to znaczy

$$h_i(\vartheta) = \frac{\Delta_{is}}{\Delta_s}, \quad i = \overrightarrow{1,s} \tag{3}$$

gdzie: $\Delta_s = \det\left\|F_{i,j}\right\|$; $(i, j = \overrightarrow{1,s})$ – objętość wielomianu stochastycznego stopnia *s*, Δ_{is} – wyznacznik otrzymany z Δ_s po zastąpieniu *i*-tej kolumny przez kolumnę wyrazów wolnych układu równań (2).

W pracy [9] wykazano, że wielomianowe oceny $\hat{\vartheta}$, będące rozwiązaniami stochastycznych równań o postaci (1), są spójne i asymptotycznie nieobciążone. W celu obliczenia oceny niepewności należy znaleźć objętość pozyskiwanej informacji o szacowanych parametrach ϑ, opisaną ogólnie przez równanie:

$$J_{sn(\vartheta)} = n\sum_{i=1}^{s} h_i(\vartheta)\frac{d}{d\vartheta}\alpha_i(\vartheta) \tag{3}$$

Sens statystyczny funkcji $J_{sn(\vartheta)}$ jest taki, jak w klasycznej koncepcji Fischera o ilości informacji, gdyż jeżeli $n \to \infty$, to jej odwrotność prowadzi do wariancji estymat:

$$\sigma^2_{(\vartheta)s} = \lim_{n\to\infty} J^{-1}_{sn(\vartheta)}\Big|_{\vartheta=\vartheta_0} \tag{4}$$

gdzie: ϑ_0 – rzeczywista wartość szacowanego parametru.

Wielomianowe oszacowanie wartości mierzonej

W praktyce pomiarowej, w wielu przypadkach do znalezienia estymaty $\hat{a}$ parametru *a* używa się statystyki liniowej, czyli średniej arytmetycznej [1, 2]:

$$\hat{a} \equiv \bar{x} = \frac{1}{n}\sum_{v=1}^{n} x_v \tag{5}$$

Parametr $\hat{a}$ ze wzoru (5) jest oszacowaniem matematycznym wartości oczekiwanej zmiennej losowej. Wyznacza się go z zastosowaniem metody momentów (MM). Według [2], estymata o postaci (5) ma minimalne odchylenie standardowe (SD), gdy zmienna losowa ma rozkład Gaussa i jej wartości $\vec{x} = \{x_1, x_2, ...x_n\}$ pobierane są losowo. W dalszych rozważaniach zakłada się, że wartości te nie są ze sobą skorelowane (funkcja autokorelacji równa zeru). Jeśli rozkład wartości danych otrzymanych w powtarzanych pomiarach różni się znacząco od funkcji Gaussa, to inne estymaty, wyznaczane metodami alternatywnymi, w tym nieliniowymi, mogą być lepsze. Na przykład środek rozstępu dla rozkładu równomiernego ma wartość SD mniejszą niż estymata $\hat{a}$, czyli niż niepewność standardowa wyznaczana wg GUM metodą typu A.

Oszacowanie wartości średniej, uzyskiwane metodą PMM przy użyciu wielomianu stopnia $s = 1$, jest identyczne z postacią (5) wg metody liniowej MM [9] . Kunchenko wykazał też, że symetria rozkładu, podobnie jak dla momentów, charakteryzuje się brakiem nieparzystych współczynników kumulant oraz estymaty otrzymane przy użyciu wielomianów stopnia $s = 2$ degenerują się do estymat według metody liniowej MM. Zatem najpierw rozpatrzy się utworzenie algorytmu do statystycznej estymacji parametru a dla stopnia wielomianu $s = 3$. W metodzie PMM przy użyciu takiego wielomianu ocena $\hat{a}$ jest rozwiązaniem równania:

$$h_1(a)\sum_{v=1}^{n}(x_v - a) + h_2(a)\sum_{v=1}^{n}\left[x_v^2 - \left(a^2 + \chi_2\right)\right] + h_3(a)\sum_{v=1}^{n}\left[x_v^3 - \left(a^3 + 3a\chi_2\right)\right]\Big|_{a=\hat{a}} = 0\,, \tag{6}$$

gdzie: $h_1(a)$, …, $h_3(a)$ – optymalne współczynniki, minimalizujące wariancję.

Dla $s = 3$ współczynniki te znajduje się rozwiązując układ trzech równań liniowych (6) zapisany w następującej postaci macierzowej

$$\chi_2\begin{bmatrix} 1 & 2a & 3a^2 + \chi_2(\gamma_4 + 3) \\ 2a & 4a^2 + \chi_2(\gamma_4 + 2) & 6a^3 + a\chi_2(5\gamma_4 + 12) \\ 3a^2 + \chi_2(\gamma_4 + 3) & 6a^3 + a\chi_2(5\gamma_4 + 12) & 9a^4 + 3a^2\chi_2(5\gamma_4 + 12) + \\ & & + \chi_2^2(15\gamma_4 + \gamma_6 + 15) \end{bmatrix}\begin{bmatrix} h_1(a) \\ h_2(a) \\ h_3(a) \end{bmatrix} = \begin{bmatrix} 1 \\ 2a \\ 3a^2 + 3\chi_2 \end{bmatrix} \tag{6a}$$

Współczynniki $h_1(a)$, …, $h_3(a)$ są dane następującymi wyrażeniami

$$h_1(\alpha) = \frac{-\chi_2^4}{\Delta_3}\left(3\alpha^2\Lambda_1 - \chi_2^2\Lambda_2\right), \qquad h_2(\alpha) = \frac{3\chi_2^4}{\Delta_3}\alpha A\,, \qquad h_3(\alpha) = \frac{-\chi_2^4}{\Delta_3}A \tag{7}$$

gdzie: $\Delta_3 = \chi_2^6\left(12 + 24\gamma_4 + 7\gamma_4^2 - \gamma_4^3 + 2\gamma_6 + \gamma_4\gamma_6\right)$, $\Lambda_1 = 2\gamma_4 + \gamma_4^2$,

$\Lambda_2 = 12 + 30\gamma_4 + 12\gamma_4^2 + 2\gamma_6 + \gamma_4\gamma_6$, $\gamma_i = \chi_i \Big/ \sqrt{\chi_2^i}$ – współczynniki i-tej kumulanty.

Po wstawieniu współczynników (7) do zależności (1) otrzymuje się następujące równanie dla a:

$$Aa^3 + Ba^2 + Ca + D\Big|_{a=\hat{a}} = 0\,, \tag{8}$$

gdzie: $A = \Lambda_1$, $B = -3\Lambda_1\frac{1}{n}\sum_{v=1}^{n}x_v$, $C = \Lambda_1\frac{1}{n}\sum_{v=1}^{n}x_v^2 - \chi_2\Lambda_2$, $D = \Lambda_2\chi_2\frac{1}{n}\sum_{v=1}^{n}x_v - \Lambda_1\frac{1}{n}\sum_{v=1}^{n}x_v^3$

Po zamianie zmiennej a na $y = a + \frac{B}{3A}$ otrzymuje się równanie (8) w postaci kanonicznej:

$$y^3 + 3Py + 2Q = 0 \tag{9}$$

gdzie: $2Q = \frac{2B^3}{27A^3} - \frac{BC}{3A^2} + \frac{D}{C}$, $3P = \frac{3AC - B^2}{3A^2}$.

Liczba pierwiastków rzeczywistych równania (9) zależy od znaku wyrażenia $G = Q^2 + P^3$. Jeśli $G > 0$, to równanie (9) ma jeden pierwiastek rzeczywisty i dwa zespolone, jeżeli $G < 0$ – trzy pierwiastki rzeczywiste, a gdy $G = 0$ to dla $P = Q$ jest tylko jedno rozwiązanie zerowe, lub dla $P^3 = -Q^2$ są dwa pierwiastki rzeczywiste.

Pierwiastki równania (8) znajduje się analitycznie ze wzorów Cardana:

$$a_{(1)} = U + Y - \frac{B}{3A}, \quad a_{(2)} = \varepsilon_1 U + \varepsilon_2 Y - \frac{B}{3A}, \quad a_{(3)} = \varepsilon_2 U - \varepsilon_1 Y - \frac{B}{3A}, \tag{10}$$

gdzie: $U = \sqrt[3]{-Q + \sqrt{Q^2 + P^3}}$, $Y = \sqrt[3]{-Q - \sqrt{Q^2 + P^3}}$, $\varepsilon_{1,2} = -0{,}5 \pm i\,0{,}5\sqrt{3}$.

Gdy według metody wielomianowej PMM poszukiwana ocena ma wiele rozwiązań, wybiera się ten pierwiastek rzeczywisty równania ogólnego (4), przy którym maksymalizuje się ilość uzyskiwanej informacji [9]. Z badań symulacyjnych, które omawia się poniżej, wynika, że przy stopniu wielomianu $s = 3$, z otrzymywanych trzech pierwiastków (10) takie właściwości jako estymata $\hat{a}$ ma estymata $a_{(1)}$.

Analiza dokładności oszacowania

Do ilościowej oceny zmian niepewności pomiarów proponuje się zastosować współczynnik redukcji wariancji estymat:

$$g_{(a)s} = \frac{\sigma^2_{(a)s}}{\sigma^2_{(a)1}} \tag{11}$$

Współczynnik ten jest stosunkiem wariancji $\sigma^2_{(a)s}$ estymaty parametru a, którą można znaleźć metodą PMM za pomocą wielomianu s-tego rzędu, i wariancji $\sigma^2_{(a)1}$ dla estymaty liniowej tego parametru (4) uzyskanej metodą momentów MM (tak jak w GUM). Jest to równoważne zastosowaniu wielomianów stopnia $s = 1$ w PMM.

Liniowa estymata $\hat{a}$ o postaci (4) jest nieobciążona (nie ma przesunięcia) i jest zgodna [2]. Jej wariancja σ^2_α nie zależy od wartości estymowanego parametru, lecz jest określona przez wariancję losowej składowej wartości obserwacji w próbce (drugiego rzędu kumulanta χ_2 zmiennej losowej ξ_0) oraz przez liczebność próbki n:

$$\sigma^2_{(\alpha)1} = \frac{\chi_2}{n} \tag{12}$$

Za pomocą wyrażenia (7) opisującego optymalne współczynniki w PMM oraz w oparciu o wzór ogólny (3) otrzymuje się wyrażenie opisujące ilość informacji szacowanych parametrów a. Uzyskuje się ją z liczebności próbki n za pomocą wielomianów stochastycznych stopnia $s = 3$:

$$J_{3n(a)} = \frac{n}{\Delta_3}\chi_2^5\left(12 + 24\gamma_4 + 9\gamma_4^2 + 2\gamma_6 + \gamma_4\gamma_6\right),$$

gdzie Δ_3 ma taką samą postać jak w (7).

Wariancja asymptotyczna $\sigma^2_{(\alpha)3}$ jest odwrotnością wartości $J_{3n(\alpha)}$ o postaci:

$$\sigma^2_{(\alpha)1} = \frac{\chi_2}{n}\left[1 - \frac{\gamma_4^2}{6 + 9\gamma_4 + \gamma_6}\right] \tag{13}$$

Tak więc, współczynnik zmniejszenia wariancji wynosi

$$g_{(\alpha)3} = 1 - \frac{\gamma_4^2}{6 + 9\gamma_4 + \gamma_6} \tag{14}$$

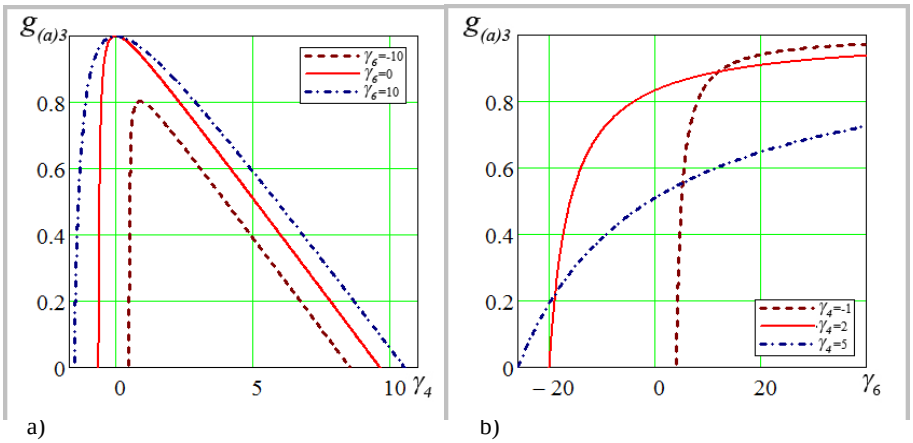
Jest on funkcją wartości współczynników kumulant wyższych rzędów (γ_4 i γ_6) i nie zależy od wartości kumulanty $\chi_2 = m_2$ oraz od liczebności n próbki.

Współczynniki kumulant wyższego rzędu dla zmiennych losowych nie mogą przyjmować dowolnych wartości, gdyż ich kombinacja ma pewien obszar wartości dopuszczalnych [9]. Na przykład, dla symetrycznych rozkładów prawdopodobieństwa zmiennych losowych o właściwościach określonych przez współczynniki kumulant stopnia 4 i 6, obszar dopuszczalnych wartości tych parametrów jest ograniczony przez dwie nierówności

$$\gamma_4 > -2 \qquad \text{i} \qquad \gamma_6 + 9\gamma_4 + 6 > \gamma_4^2 .$$

Z ostatniej nierówności i analizy wzoru (14) wynika, że współczynnik redukcji wariancji współczynników kumulant $g_{(\alpha)3}$ jest liczbą z przedziału (0; 1].

Na rysunku 1 podano wykresy uwzględniające te ograniczenia.



Rys. 1. Zależności współczynnika redukcji wariancji $g_{(\alpha)3}$ od współczynników kumulant γ_4, γ_6
Fig. 1. Dependence of variance reduction coefficient $g_{(\alpha)3}$ on cumulant coefficients γ_4, γ_6

Wykresy te ilustrują zależności współczynnika $g_{(\alpha)3}$ i są wykonane jako przekroje funkcji dwu zmiennych opisanej wzorem (14). Przedstawiają zależność dla jednego z parametrów przy trzech określonych wartościach drugiego. Można również zauważyć, że wariancja estymaty wielomianowej $\hat{a}$ znacznie wzrasta i dąży

asymptotycznie do zera, gdy wartości współczynników kumulant zbliża się do granicy obszaru ograniczonego z lewej przez parabolę. Tylko gdy współczynnik kurtozy wynosi zero, nie obserwuje się redukcji estymaty wariancji. W innych przypadkach (gdy ($\gamma_4 \neq 0$) ocena $\hat{a}$ metodą PMM ma mniejszą niepewność niż liniowa estymata tego parametru, np. niepewność typu A wg GUM. Zysk zależy od wartości współczynników kumulant wyższego rzędu.

Modelowanie statystyczne estymacji średniej metodą wielomianową

Na podstawie powyższych rozważań opracowano pakiet oprogramowania w środowisku programowym MATLAB. Pozwala on na modelowanie statystyczne proponowanego wielomianowego estymatora menzurandu o symetrycznym niegaussowskim rozkładzie rozproszenia jego wartości, w oparciu o statystykę wyższego rzędu. Pakiet ten wykorzystuje metodę Monte Carlo. Otrzymuje się analizę porównawczą dokładności różnych algorytmów estymacji statystycznej. Pozwala on także na wykorzystanie probabilistycznych właściwości estymat wielomianowych.

Eksperymentalnie otrzymane wartości współczynnika zmniejszenia wariancji $g_{(a)3}$ (11) stanowią kryterium dla porównań skuteczności metody PMM. Oblicza się go dla *k*-krotnych eksperymentów o tej samej wartości początkowej parametrów modelu. Estymowany współczynnik $\hat{g}_{(a)3}$ jest stosunkiem wariancji $\hat{\sigma}^2_{(a)3}$ oszacowań parametrów metodą PMM dla wielomianu 3. rzędu i wariancji $\hat{\sigma}^2_{(a)1}$ liniowych oszacowań tego parametru wg wzoru (5).

Wiarygodność wyników symulacji za pomocą statystycznych algorytmów estymacji zależy od dwóch czynników: całkowitej liczebności *n* wektora wejściowego $\bar{x}$ (z którego wyznacza się estymatę wartości poszukiwanego parametru) i liczby eksperymentów *M*, które przeprowadza się metodą Monte Carlo przy tych samych warunkach początkowych (wartościach współczynników kumulant (γ_4 i γ_6 opisujących probabilistyczny charakter modelu). Wyniki uzyskane przy zastosowaniu metody Monte Carlo podano w tabeli 1.

Tab. 1. Wyniki symulacji Monte Carlo estymowanych parametrów (część 1)
Tab. 1. The results from the Monte Carlo simulation of estimated parameters (Part 1)

Rozkład		Wartości teoretyczne			Symulacja Monte Carlo		
		γ_4	γ_6	$g_{(a)3}$	$\hat{g}_{(a)3}$		
					$n = 20$	$n = 50$	$n = 200$
Arc sin		–1,3	8,2	0,2	0,27	0,22	0,21
Równomierny		–1,2	6,9	0,3	0,41	0,33	0,31
Trapezowy	$\beta = 0{,}75$	–1,1	6,4	0,36	0,47	0,39	0,37
	$\beta = 0{,}5$	–1	5	0,55	0,63	0,58	0,55
	$\beta = 0{,}25$	–0,7	2,9	0,76	0,82	0,78	0,77
Trójkątny		–0,6	1,7	0,84	0,89	0,86	0,85
Laplace'a		3	30	0,86	0,91	0,87	0,86

Dane doświadczalne wartości współczynnika redukcji wariancji $g_{(a)3}$ uzyskano z $M = 10^4$ prób MC (*trials*) dla różnych symetrycznych rozkładów obserwacji pomiarowych. W obliczeniach estymat parametru a wielomianu (10) nie brano pod uwagę informacji *a priori* o rodzaju rozkładu, ale tylko trzy wartości parametrów modelu χ_2 oraz γ_4 i γ_6. Wartości te można otrzymać z analitycznych zależności pomiędzy parametrami rozkładu gęstości prawdopodobieństwa (PDF) i jego momentami, na podstawie których oblicza się odpowiednie kumulanty.

W praktyce w sytuacji, gdy informacja o PDF rozkładu nie jest dostępna *a priori*, estymaty poszukiwanych parametrów łatwo uzyskuje się na drodze algorytmicznej za pomocą procedur wykorzystujących relacje asymptotyczne (15):

$$\hat{\chi}_2 = \hat{\mu}_2, \qquad \hat{\gamma}_4 = \hat{\mu}_4 / \hat{\mu}_2^2 - 3, \qquad \hat{\gamma}_6 = \hat{\mu}_6 / \hat{\mu}_2^3 - 15\hat{\mu}_4 / \hat{\mu}_2^2 + 30. \tag{15}$$

gdzie: $\hat{\mu}_i$ – i-ty moment centralny próbki:

$$\hat{\mu}_i = \frac{1}{k}\sum_{v=1}^{k}(x_v - \bar{x})^i \tag{16}$$

Estymatory o postaci (16) są zgodne i dla $n \to \infty$ asymptotycznie nieobciążone [11]. Metodę obliczania niezbędnej wartości k próbki, która zapewni uprzednio wymagany błąd względny oszacowania współczynników kumulant do 6. rzędu włącznie, przedstawiono w [12].

Analiza danych z tabeli 1 wykazuje znaczną korelację pomiędzy obliczeniami analitycznymi i wynikami modelowania statystycznego. Wraz ze wzrostem liczebności n próbki, różnica pomiędzy teoretycznymi i eksperymentalnymi wartościami współczynnika redukcji wariancji $g_{(a)3}$ maleje. Na przykład, dla $n = 50$ różnica ta nie przekracza 10%, a gdy $n = 200$, to jest mniejsza niż 5%. Wyniki te potwierdzają asymptotyczną właściwość (4) charakteryzującą ocenę wielkości wydobytej informacji $J_{sn(a)}$ o estymowanych parametrach, opisaną ogólnie wzorem (3). Służy ona do obliczania dyspersji estymat wielomianowych jako rozwiązań układu równań o ogólnej postaci (1), gdyż dla wartości rzeczywistej ϑ_0: $\sigma^2_{(\vartheta)s} = \lim\limits_{n\to\infty} J^{-1}_{sn(\vartheta)}\Big|_{\vartheta=\vartheta_0}$.

Z pracy [2] wynika, że przy symetrycznym rozkładzie danych, jako estymator wartości menzurandu, poza średnią arytmetyczną, stosuje się też inne parametry statystyczne, w tym medianę i środek rozstępu. Skuteczne wykorzystanie odpowiednich statystyk zależy od typu rozkładu, w tym udziału („ciężaru”) ogonów, liczby modów i ograniczonego rozstępu rozproszenia danych. Wartość współczynnika kurtozy γ_4 stosuje się też jako kryterium do wyboru odpowiedniej statystyki. Mediana jest estymatorem lepszym niż średnia dla rozkładów jednomodalnych Laplace'a, Cauchy'ego o „cięższych ogonach” niż rozkład normalny. Środek rozstępu (*midrange*) jest skuteczniejszym estymatorem dla rozkładów o ograniczonym rozstępie (arc sin, równomierny). Estymuje się też parametry liniowych kombinacji różnych statystyk. Na przykład w pracy [13] wykazano, że dla rozkładu trapezowego jako kombinacji dwu równomiernych najlepszy jest estymator dwuelementowy o optymalnych współczynnikach wagowych średniej i środka rozstępu, zależnych od stosunku podstaw trapezu β.

Wyniki badań uzupełniających przedstawiono w tabeli 2. Parametry modelowania (liczba doświadczeń, objętość i wybór parametrów rozkładu obserwacji pomiarowych) tej tabeli są określone podobnie, jak dla tabeli 1.

Tab. 2. Wyniki estymacji parametrów wg symulacji metodą Monte Carlo (część 2)
Tab. 2. The results from Monte-Carlo simulation of parameters' estimation (Part 2)

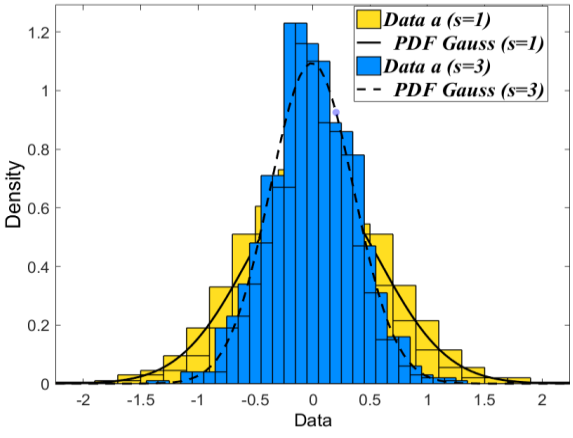
Rozkład		Symulacja Monte Carlo					
		$\hat{p}_{(a)3}$			$\hat{q}_{(a)3}$		
		$n=20$	$n=50$	$n=200$	$n=20$	$n=50$	$n=200$
Arc sin		0,09	0,06	0,05	1,9	4,2	17
Równomierny		0,16	0,12	0,1	1,6	3	10
Trapezowy	$\beta=0{,}75$	0,18	0,14	0,13	1,17	1,1	1,03
	$\beta=0{,}5$	0,29	0,23	0,21	0,86	0,75	0,72
	$\beta=0{,}25$	0,42	0,38	0,36	0,76	0,7	0,67
Trójkątny		0,57	0,54	0,53	0,73	0,68	0,66
Laplace'a		1,38	1,42	1,56	0,11	0,04	0,01

W tabeli 2 porównano dokładności estymat wielomianowych jako stosunki:

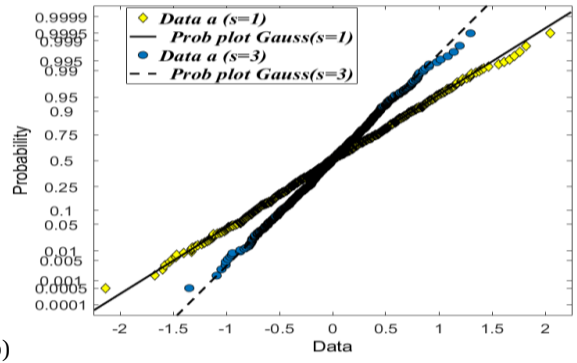
$$\hat{p}_{(a)s}=\frac{\hat{\sigma}^2_{(a)s}}{\hat{\sigma}^2_{(a)med}},\quad \hat{q}_{(a)s}=\frac{\hat{\sigma}^2_{(a)s}}{\hat{\sigma}^2_{(a)mr}} \tag{17}$$

gdzie $\hat{\sigma}^2_{(a)med}$ i $\hat{\sigma}^2_{(a)mr}$ są eksperymentalnymi odchyleniami estymat mediany i środka rozstępu uzyskanymi przez statystyczne modelowanie metodą Monte Carlo.

Wyniki eksperymentu przedstawiono też graficzne na rysunku 2. Rysunek ten pokazuje, że istnieje duże zbliżenie rozkładu do modelu gaussowskiego.



a)



b)

Fig. 2. Przykład eksperymentalnych oszacowań mierzonego parametru o trapezowym rozkładzie (β = 0,5) błędów przypadkowych: a) histogramy i aproksymacja PDF funkcją Gaussa; b) funkcja rozkładu empirycznego i wykresy prawdopodobieństwa Gaussa

Fig. 2. Example of experimental estimates of the measured parameter with a trapezoidal distribution (β = 0.5) of random errors: a) histograms and Gaussian approximation of PDF; b) empirical distribution function and Gaussian probability plots

Hipotezę o ważności rozkładu Gaussa dla estymat wielomianu zbadano też za pomocą testu Lillieforsa w oparciu o statystyki Kołmogorowa-Smirnowa [14]. Testem tym bada się normalność rozkładu w oprogramowaniu MATLAB. W tabeli 3 przedstawiono wyniki tych badań w postaci szeregu parametrów wyjściowych testu.

Tab. 3. Wyniki badania adekwatności modelu Gaussa dla estymat wielomianu (s = 3) na podstawie testu Lillieforsa

Tab. 3. The results of testing the adequacy of the Gaussian distribution model of polynomial estimates (s = 3) on the basis of Lilliefors test

Rozkład		Parametry wyjściowe testu Lillieforsa						
		P			*LSTAT*			*CV*
		$n = 20$	$n = 50$	$n = 200$	$n = 20$	$n = 50$	$n = 200$	
Arcsin		0,001	0,003	0,25	0,03	0,01	0,007	0,009
Równomierny		0,001	0,006	0,31	0,02	0,01	0,007	
Trapezowy	$\beta = 0{,}75$	0,001	0,12	0,5	0,02	0,008	0,006	
	$\beta = 0{,}5$	0,07	0,18	0,5	0,008	0,007	0,006	
	$\beta = 0{,}25$	0,24	0,26	0,5	0,007	0,007	0.006	
Trójkątny		0,31	0,5	0,5	0,007	0,006	0,005	
Laplaca		0,001	0,09	0,5	0,016	0,008	0,006	

P jest poziomem istotności, który odpowiada wartości próbki testowej statystyki; *LSTAT* – wartość próbki testu statystycznego; *CV* – krytyczna wartość statystyki.

Jeżeli *LSTAT* < *CV* to hipoteza zerowa przy danym poziomie krytycznym jest prawdziwa. Wyniki w tabeli 3 uzyskano dla różnych rodzajów błędów pomiaru oraz liczebności n próbki $\bar{x}$ oraz stałego poziomu istotności α = 0,05 hipotezy zerowej (Gaussa) i $M = 10^4$ eksperymentów MC. Z analizy tych i innych wyników badań

eksperymentalnych wynika, że dla modeli błędów pomiaru o rozkładzie, który znacząco różni się od modelu Gaussa (np. arc sin i równomierny mają duże wartości bezwzględne współczynników kumulant wyższego rzędu), normalne rozkłady estymat wielomianu obserwuje się tylko przy odpowiednio dużych liczbach elementów próbki, tj. $n > 100$.

Normalizacja trapezowego rozkładu danych pomiarowych o dużej wartości $\beta = 0{,}75$ pojawia się dla $n \geq 50$. Dla średnich i małych wartości $\beta \leq 0{,}5$ i dla rozkładu trójkątnego ($\beta = 0$) normalizacja zachodzi już dla małych próbek $n = 20$.

Dla określonych warunków normalizacji otrzymano wyrażenia analityczne opisujące dokładność oszacowania wielomianu. Pozwalają one obliczyć standardową niepewność wyników pomiarowych. Nie jest wymagana informacja *a priori* o rodzaju rozkładu, a tylko o wartościach ograniczonej liczby kumulant opisujących właściwości probabilistyczne próbki pomiarowej. Ponadto, korzystanie z addytywności funkcji kumulant pozwala w prosty sposób uwzględniać niepewności generowane przez wiele źródeł i o różnych właściwościach probabilistycznych.

Wyznaczanie niepewności rozszerzonej

Wyznaczenie niepewności rozszerzonej $U = k_p u(Y)$ o zadanym prawdopodobieństwie p wymaga znajomości współczynnika rozszerzenia k_p. Dla danych z populacji o rozkładzie jednomodalnym symetrycznym w [17] podano dla k_p wzór:

$$k_p \leq \frac{2}{3 \cdot \sqrt{1-p}} \tag{18}$$

Dla $p = 0{,}95$ dla Gaussa otrzymuje się $k_p = 1{,}96$, a z wzoru (18) $k_p = 2{,}98$.

Inne zależności do wyznaczania niepewności rozszerzonej omówiono w Dodatku 2.

W tabeli 4 zawarto wartości rozszerzonej niepewności pomiaru U o poziomie ufności $p = 0{,}95$ i współczynnik $g_{(a)3}$ dla znormalizowanej próbki, tj. o wariancji równej 1.

Tab. 4. Niepewności rozszerzone wg GUM i metodą wielomianową PMM
Tab. 4. Comparison of expanded uncertainty estimation by GUM and PMM methods

Rozkład		Niepewność rozszerzona U dla p = 0,95															
		wartości teoretyczne									symulacja Monte Carlo						
		$n = 20$			$n = 50$			$n = 200$			$n = 20$		$n = 50$		$n = 200$		
		$s = 1$	$s = 3$	$g_{(a)3}$	$s = 1$	$s = 3$	$g_{(a)3}$	$s = 1$	$s = 3$	$g_{(a)3}$	$s = 1$	$s = 3$	$s = 1$	$s = 3$	$s = 1$	$s = 3$	
Arc sin		0,44	0,23	0,52	0,28	0,12	0,43	0,14	0,06	0,43	0,44	0,2	0,28	0,13	0,14	0,06	
Równomierny			0,24	0,55		0,15	0,54		0,08	0,57	0,44	0,28	0,28	0,16	0,14	0,08	
Trapezowy	$\beta = 0{,}75$		0,26	0,59		0,17	0,61		0,09	0,64	0,44	0,3	0,28	0,18	0,14	0,09	
	$\beta = 0{,}5$		0,33	0,75		0,21	0,75		0,10	0,71	0,44	0,35	0,28	0,21	0,14	0,1	
	$\beta = 0{,}25$		0,38	0,86		0,24	0,86		0,14	1	0,44	0,4	0,28	0,24	0,14	0,12	
Trójkątny			0,39	0,89		0,25	0,89		0,13	0,93	0,44	0,42	0,28	0,26	0,14	0,13	
Laplace			0,40	0,91		0,26	0,93		0,13	0,93	0,44	0,41	0,28	0,26	0,14	0,13	

Przedstawione w tabeli 4 wyniki uzyskano na podstawie obliczeń teoretycznych, przy założeniu estymatorów niepewności standardowej próbek takich, jak dla rozkładu normalnego wg Przewodnika GUM oraz dla metody wielomianowej PMM z $M = 10^5$ wyników (trials) symulacji Monte Carlo o wielomianach stopnia $s = 1$ oraz $s = 3$. Przy wyznaczaniu niepewności rozszerzonej U założono, że wartości estymatorów odchylenia standardowego dla rozpatrywanych liczebności n danych próbki mają rozkład Gaussa.

Z porównania wartości współczynnika $g_{(a)3}$ wynika, że dla stopnia wielomianu $s = 3$ najwięcej uzyskuje się dla większych liczebności próbki n i rozkładów danych znacznie odbiegających od rozkładu normalnego, tj. arc sin, równomierny i trapezowy o $\beta = 0{,}75$. Niepewność U dla $p = 0{,}95$ wg metody PMM jest mniejsza nawet poniżej 50%. Wyniki z symulacji metodą Monte Carlo oraz z obliczeń różnią się nieznacznie.

Wnioski

Łączna analiza uzyskanych wyników teoretycznych i eksperymentalnych umożliwia sformułowanie podanych poniżej ogólnych wniosków dotyczących możliwości stosowania w metrologii wielomianów stochastycznych jako narzędzia matematycznego, które zaproponował Kunchenko [9].

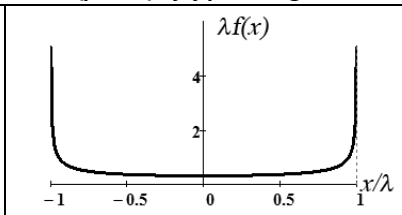
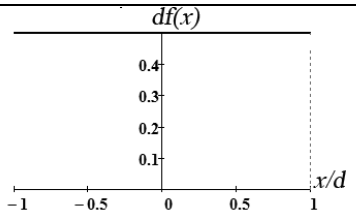
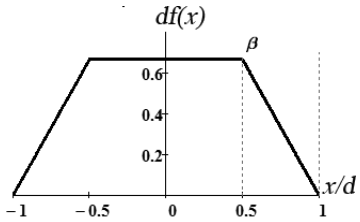
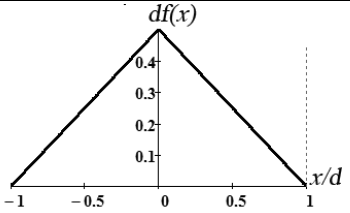
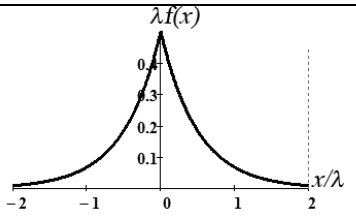
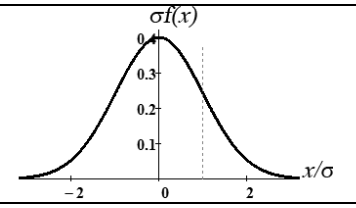
- Wielomiany te umożliwiają konstruowanie algorytmów służących do znajdowania nieliniowych estymat mierzonego parametru w obecności symetrycznych, ale niegaussowskich rozkładów addytywnych błędów losowych opisanych przez statystyki wyższego rzędu.
- Badania wykazały, że szacowanie wielomianowe powinno się stosować dla tych probabilistycznych modeli błędów pomiarowych (za wyjątkiem Gaussa), dla których efektywność estymaty średniej arytmetycznej jest wyższa niż innych estymat statystycznych (mediany, środka rozstępu).
- Ogólnie, oceny wielomianowe, syntetyzowane przy $s = 3$, charakteryzują się większą dokładnością w porównaniu z liniowymi ocenami średniej arytmetycznej. Stopień poprawy zależy od poziomu niegaussowości danych próbki wyrażającej się liczbowo przez bezwzględne wartości współczynników kumulant wyższych rzędów.
- Zmniejsza się wariancja i niepewność standardowa estymat parametrów niegaussowskich rozkładów prawdopodobieństwa składowej losowej występującej w obserwacjach pomiarowych. Zależy to od uwzględnienia określonej liczby kumulant. Takie oszacowanie danych próbki pomiarowej jest prostszym zadaniem niż dobór odpowiedniego rozkładu dla tych danych i sprawdzenie jego adekwatności dla oceny niepewności pomiarów.
- Dokładność wyników uzyskanych metodą PPM maleje wraz ze zmniejszaniem się liczebności próbki pomiarowej.

Literatura

1. Evaluation of measurement data – Guide to the expression of uncertainty in measurement (GUM), BIPM JCGM 100: 2008 (last corrected version)

1a. *Wyrażanie Niepewności Pomiaru*. Przewodnik, tłumaczenie wyd. 2 z 1995r. i komentarz J. Jaworskiego, Wydawnictwo Głównego Urzędu Miar Warszawa 1999.
1b. Supplement 1 to GUM – Propagation of distributions using a Monte Carlo method, Guide OIML G 1-101 Ed. 2008.
2. Novickij P.V., Zograf I.A., *Ocenka pogreshnostiej resultatov izmierenii,* Energoatomizdat, Leningrad 1991 (in Russ).
3. Mendel J.M., *Tutorial on higher-order statistics* (*spectra*) *in signal processing and system theory: theoretical results and some applications*. Proc. IEEE, 79(3), 1991, 278–305.
4. Kendall M.G., Stuart A., *The Advanced Theory of Statistics*, Vol. 1: Distribution Theory, 3rd Edition, Griffin 1969.
5. Toybert P., *Otsenka tochnosti rezultatov izmereniy* (Estimation of accuracy of measurement results). Energoatomizdat, Leningrad 1988 (in Russ).
6. Zaharov I.P., Klimova E.A., *Application of excess method to obtain reliable estimate of expanded uncertainty*. "Systemy obrobky informatsii", No. 3(119), 2014, 24–28 (in Russ).
7. Kuznetsov B.F., Borodkin, D.K., Lebedeva L.V., *Cumulant models of additional errors*. "Sovremennye tekhnologii. Sistemnyi analiz. Modelirovanie", No. 1(37), 2013, 134–138.
8. De Carlo L.T., *On the meaning and use of kurtosis*. "Psychological methods", *2*(3), 1997, 292–307. doi:10.1037/1082-989X.2.3.292.
9. Kunchenko Y., *Polynomial Parameter Estimations of Close to Gaussian Random Variables*. Germany, Aachen: Shaker Verlag 2002.
10. Zabolotnii S.W., Warsza Z.L., *Semi-parametryczna estymacja punktu zmiany parametrów w szeregach niegaussowskich metodą maksymalizacji wielomianu.* „Przegląd Elektrotechniczny", ISSN 0033-2097, R. 91, Nr 1, 2015, 102–107.
11. Cramér H., *Mathematical methods of statistics*, Vol. 9, Princeton University Press, 1999.
12. Beregun V.S., Garmash O.V., Krasilnikov A.I., *Mean square error of estimates of cumulative coefficients of the fifth and sixth order*. "Electronic Modelling", Vol. 36, N1, 2014, 17–28.
13. Warsza Z., Galovska M., *About the best measurand estimators of trapezoidal probability distributions*. "Przegląd Elektrotechniczny", 85(5), 2009, 86–91.
14. Lilliefors H.W., *On the Kolmogorov-Smirnov test for normality with mean and variance unknown*. "Journal of the American Statistical Association", 62(318), 1967, 399–402.
15. Zabolotnii S.W., Warsza Z.L., *Semi-parametric estimation of the change-point of parameters of non-gaussian sequences by polynomial maximization method*. In: R. Szewczyk et all (eds.), Challenges in Automation, Robotics and Measurement Techniques, *Proceedings of Automation 2016.* Series: Advances in Intelligent Systems and Computing, Vol. 440, Springer International Publishing Switzerland 2016, 903–919, DOI 10.1007/978-3-319-29357-8_80.
16. Kubisa S., Warsza Z.L., *Środek rozstępu jako estymator menzurandu dla próbek z populacji o rozkładzie równomiernym i płasko-normalnym*. „Pomiary Automatyka Kontrola", Vol. 60, No. 6, 2014, 398–401.
17. Bich W., Cox M., Michotte C., Towards a new GUM – an update. "Metrologia", Vol. 53 2016, S149– S159 IOP publishing | BIPM.
18. Warsza Z.L., Zabolotnii S.W., *A polynomial estimation of measurand parameters for samples of non-Gaussian symmetrically distributed data*. In: R. Szewczyk et all **(**eds.): Innovations in Automation, Robotics and Measurement Techniques. *Proccedings of Automation 2017*. Series: Advances in Intelligent Systems and Computing, Vol. 550. Springer International Publishing AG 2017, 468–480, DOI 10.1007/978-3-319-54042-9_45.

Tab. 5. Centralne momenty i kumulanty kilku podstawowych rozkładów symetrycznych
Tab. 5. Central moments and cumulants of few symmetrical distributions

Scentrowana funkcja PDF rozkładu ($\mu_1 = 0$) i jej przebieg	
Arc sin $f(x) = \dfrac{1}{\pi\sqrt{\lambda^2 - x^2}}$ $-\lambda < x < \lambda$ $\mu_r = \dfrac{(r-1)!!}{r!!}\lambda^r$	
Równomierny (szczególny przypadek trapezu $\beta = 1$) $f(x) = \dfrac{\mathbf{1}(x+d) - \mathbf{1}(x-d)}{2d}$ $\mu_r = \dfrac{(2d)^r}{2^r(r+1)}$	
Trapezowe $f(x) = \begin{cases} \dfrac{x+d}{d^2-(\beta\cdot d)^2}, & d \le x < -\beta\cdot d \\ \dfrac{1}{\beta\cdot d + d}, & -\beta\cdot d \le x < \beta\cdot d \\ \dfrac{d-x}{d^2-(\beta\cdot d)^2}, & \beta\cdot d \le x < d \end{cases}$	
Trójkątny (szczególny przypadek trapezu β=0) $f(x) = \begin{cases} \dfrac{x+d}{d^2}, & d \le x \le 0 \\ \dfrac{d-x}{d^2}, & \le x \le d \end{cases}$	
Laplace'a $f(x) = \dfrac{1}{2}\lambda e^{-\lambda\lvert x\rvert}$ $\mu_r = \dfrac{r!}{\lambda^r}$	
Normalny (Gaussa) $\mu_r = \sigma^r(r-1)!!$, r - parzyste	

Momenty centralne μ_r	**Kumulanty χ_r**
$\mu_2 = \frac{\lambda^2}{2}$ $\mu_4 = \frac{3}{8}\lambda^4$ $\mu_6 = \frac{5}{16}\lambda^6$	$\chi_2 = \frac{\lambda^2}{2}$ $\chi_4 = -\frac{3}{8}\lambda^4$ $\chi_6 = \frac{5}{4}\lambda^6$
$\mu_2 = \frac{d^2}{3}$ $\mu_4 = \frac{d^4}{5}$ $\mu_6 = \frac{d^6}{7}$	$\chi_2 = \frac{d^2}{3}$ $\chi_4 = -\frac{2}{15}d^4$ $\chi_6 = \frac{16}{63}d^6$
$\mu_2 = \frac{d^2 \cdot (\beta^2+1)}{6}$ $\mu_4 = \frac{d^4 \cdot (\beta^4+\beta^2+1)}{15}$ $\mu_6 = \frac{d^6 \cdot (\beta^2+1) \cdot (\beta^4+1)}{28}$	$\chi_2 = \frac{d^2 \cdot (\beta^2+1)}{6}$ $\chi_4 = -\frac{d^4 \cdot (\beta^4+6\beta^2+1)}{60}$ $\chi_6 = \frac{d^6 \cdot (\beta^2+1) \cdot (\beta^4+14\beta^2+1)}{126}$
$\mu_2 = \frac{d^2}{6}$ $\mu_4 = \frac{d^4}{15}$ $\mu_6 = \frac{d^6}{28}$	$\chi_2 = \frac{d^2}{6}$ $\chi_4 = -\frac{d^4}{60}$ $\chi_6 = \frac{d^6}{126}$
$\mu_2 = \frac{2}{\lambda^2}$ $\mu_4 = \frac{24}{\lambda^4}$ $\mu_6 = \frac{720}{\lambda^6}$	$\chi_2 = \frac{2}{\lambda^2}$ $\chi_4 = \frac{12}{\lambda^4}$ $\chi_6 = \frac{960}{\lambda^6}$
$\mu_2 = \sigma^2$ $\mu_4 = 3\sigma^4$ $\mu_6 = 15\sigma^6$	$\chi_2 = \sigma^2$ $\chi_4 = 0$ $\chi_6 = 0$

Rozdział 11

Zastosowanie statystyki odpornościowej do oceny dokładności pomiarów

Streszczenie. Przedstawiono dwie metody oceny niepewności próbki danych doświadczalnych o niewielkiej liczebności, odporne na zawarte w niej odstające obserwacje pomiarowe o wartościach znacznie różniących się od pozostałych, czyli tzw. outliery. Metodami tymi wyznacza się parametry statystyczne wyniku pomiaru na podstawie wszystkich danych eksperymentalnych, ale dane odstające traktuje się odmiennie jako mniej wiarygodne. Rozważania ilustruje przykład liczbowy wykorzystujący dane pomiarowe z badań międzylaboratoryjnych (*key comparision*). Porównano otrzymane w nim wyniki obliczone metodą o przeskalowanym odchyleniu medianowym MAD_S i metodą iteracyjną dwukryterialną podaną przez Hubera z zastosowaniem winsoryzacji danych odstających. Podano wnioski.

W praktyce pomiarowej, szczególnie w badaniach technicznych i przemysłowych, zdarza się iż wartości jednej lub kilku obserwacji w próbce istotnie odstają od pozostałych danych. Nawet niewielka ich liczba może znacząco zmienić klasyczne parametry statystyczne próbki o małej liczbie elementów, a w wielu przypadkach nie ma możliwości powtórzenia lub uzupełnienia pomiarów. Wartość wyniku pomiarów i niepewność statystyczna u_A, obliczone w dotychczas stosowany sposób zgodnie z zasadami międzynarodowego Przewodnika GUM [1], mogą okazać się niewiarygodne, a nawet sprzeczne ze zdrowym rozsądkiem. Aby uniknąć takiej sytuacji, tradycyjnie stosuje się testy statystyczne. Pozwalają one zidentyfikować, a następnie usunąć nietypowe, tj. odstające wartości danych (ang. *outliers*). Postępowanie takie jest skuteczne tylko dla dużych próbek, gdyż dla małych próbek, służące do tego celu testy, np. test Grubbsa, tracą wrażliwość na wartości odstające. W przypadku uzyskania podczas badań eksperymentalnych próbek o małej liczbie danych (np. o dużym koszcie, wykonywanych metodami niszczącymi nieliczne posiadane obiekty lub o pomiarach niemożliwych do powtórzenia), należy starać się wykorzystać wszystkie otrzymane dane. Usunięcie dowolnej obserwacji z takiej próbki zmniejsza wiarygodność jej oceny statystycznej. Stosunek odchylenia standardowego $s(u_A)$ do wartości oczekiwanej niepewności u_A oszacowanej w eksperymencie dla próbki o n nieskorelowanych obserwacjach wynosi około $[2(n–1)]^{–1}$ [1, 2] (dodatek E.1). Na przykład dla $n = 10$ stosunek ten równa się 24%, dla $n = 4$ – wynosi 42%, a dla $n = 3$ – rośnie aż do 52%. Usunięcie tylko jednej obserwacji z tak mało licznej próbki powoduje wzrost odchylenia standardowego $s(u_A)$ aż o około 24%. Z drugiej strony rozrzut odchylenia medianowego MAD, bardziej odpornego niż $s = u_A$, jest około 1,5 razy większy.

Klasyczne metody statystyczne przetwarzania danych opierają się na założeniu modelowania prawdopodobieństwa ich rozrzutu rozkładem normalnym. Jednak dla

szeregu przypadków w praktyce liczba danych nie jest wystarczająca do budowy w pełni odpowiadających im modeli parametrycznych. W latach 60. ubiegłego wieku statystycy, bazując na wynikach szczegółowych badań ustalili, że eksperymentalne dane przetwarzane przy założeniu o występowaniu rozkładu normalnego zawierają zwykle średnio około 10% (od 1% do 20%) danych odstających od tego rozkładu, czyli zanieczyszczeń. Dawniej, ze względu na żmudne obliczenia, opracowywano tylko dane o wysokiej jakości. Natomiast obecnie, mając do dyspozycji wydajne komputery – można przetwarzać prawie każde dane.

Już Poincare wskazywał na nieuzasadnioną, ale panującą powszechnie głęboką wiarę w uniwersalność rozkładu normalnego. Matematycy sądzą, że rozkład normalny można zaobserwować w każdym eksperymencie. Eksperymentatorzy zaś uważają, że matematycy mogą udowodnić teoretycznie, na podstawie centralnego twierdzenia granicznego, iż musi zachodzić rozkład normalny. W nielicznych tylko przypadkach, np. gdy wynik pomiaru wyznacza się z dużej liczby zliczeń, normalność rozkładu dla samego menzurandu można traktować jako ścisłe prawo fizyki, a przy małej ich liczbie, że spełniony jest rozkład Poissona. Wszystkie założenia teoretyczne opierają się na możliwości prowadzenia eksperymentu (uzyskiwania obserwacji) w tych samych stałych warunkach. W makro rzeczywistości zwykle występuje przestrzenno-czasowa zmienność obiektu badań, systemu pomiarowego oraz warunków otoczenia. Dlatego to John Tukey [5] wyznawał nawet pogląd, że normalność rozkładu danych uzyskanych w eksperymencie jest mitem i takiego rozkładu nie ma i nigdy się nie uzyska. Jest to, jak widać, zagadnienie dla filozofów. Natomiast w praktyce pomiarowej, szczególnie przy małej liczbie danych powodem powstania odstających wartości niektórych obserwacji pomiarowych mogą być błędy wynikające z nieprzestrzegania zasad prowadzenia doświadczenia, niewykryte uszkodzenia przyrządów pomiarowych, błędy i omyłki w przetwarzaniu danych, wpływy zewnętrzne o dużej intensywności i wiele innych.

Przyjęty kształt rozkładu danych doświadczalnych jest tylko proponowanym umownym modelem, któremu one powinny podlegać. Sam wybór może nie być idealny i część wartości obserwacji podlega innemu rozkładowi niż przyjęty. Parametryczne podejście przy założeniu, że rozkład jest znany (i powinien być normalny) tak głęboko zakorzeniło się w praktyce statystycznego przetwarzania danych, że rezygnacja z niego zwykle nie wydaje się zasadna.

Celem jest przedstawienie oceny dokładności próbki metodą iteracyjną opartą na statystyce odpornościowej (ang. *robust statistics*). Dzięki małej wrażliwości na dane odstające (ang. *outliers*) umożliwia ona poprawę wiarygodności wyników przy próbkach również o małej liczności. Podstawy tej statystyki opracowano kilkadziesiąt lat temu [4–10]. Rozwijana jest dalej dla różnych zadań przetwarzania danych, w tym do kalibracji pomiarów wieloparametrowych [6–10, 12, 14, 15], w wyspecjalizowanych technikach pomiarowych, np. w Polsce w chemometrii na Uniwersytecie Śląskim [11–13]. Niektóre metody odpornościowe są już oprogramowane, m.in. w środowisku MATLAB, ale nie trafiły jeszcze do podstawowych międzynarodowych przepisów metrologicznych, takich jak GUM. Podstawowe informacje o statystyce odpornościowej przedstawił jako pierwszy polskiemu środowisku metrologicznemu Andrzej Zięba z Akademii Górniczo-Hutniczej na sympozjum Podstawowych Problemów Pomiarów w 2007 r. oraz ich zasady ujął w swojej książce [17].

Ocena rozproszenia danych metodami odpornymi

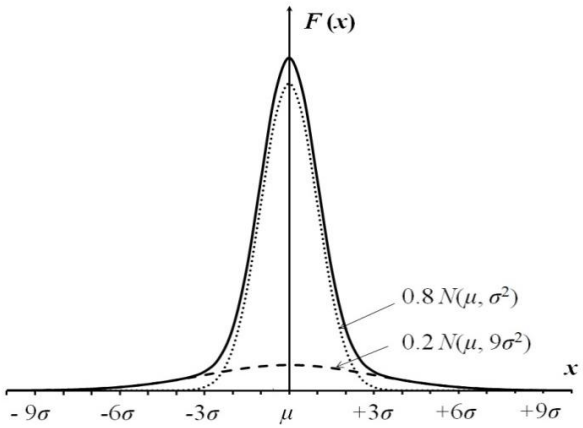
W praktyce najbardziej rozpowszechnione są takie procedury analizy statystycznej, które są optymalne przy założeniu normalności rozkładu danych. Są one jednak dość wrażliwe na niewielkie odchylenia rzeczywistego rozrzutu danych od tego założenia. Występuje to szczególnie przy estymacji wariancji. W rzeczywistości często mamy do czynienia z rozkładami różniącymi się od idealnego rozkładu normalnego. Dlatego też ostatnio w przetwarzaniu danych doświadczalnych coraz szerzej stosuje się metody odporne [2–4, 13, 14]. Pod pojęciem odporność rozumie się w statystyce stabilność wartości wyznaczanych parametrów próbek pomiarowych, czyli ich niewrażliwość na różne odchylenia i niejednorodności rozrzutu elementów. W ogólnym przypadku powstają one z przyczyn nieznanych.

Podstawowym modelem w metodach odpornych nie jest zmienna losowa o pojedynczym rozkładzie, ale model mieszany. Dla różnych próbek danych tylko środkową część ich uporządkowanych wartości można dość stabilnie modelować rozkładem normalnym. Zaś dolne fragmenty zboczy modelu rzeczywistego rozkładu, czyli jego ogony, są bardziej rozciągnięte niż dla rozkładu normalnego części środkowej i mało stabilne. Obserwacje pomiarowe w tych krańcowych obszarach występują rzadziej. Względem rozkładu normalnego dla części środkowej pojedyncze z nich, szczególnie dla małych próbek, mogą być rozpoznawane jako outliery pomimo, że w rzeczywistości są pseudo-outlierami. Przyjęcie takiego podejścia pozwala zachować tradycyjny pogląd o hipotetycznej jednorodności populacji generalnej, stanowiący dotychczas podstawę wszystkich ocen statystycznych. Jedynie w ogonach rozkładu rzeczywistego dopuszcza się możliwość występowania odchyleń odstających od rozkładu normalnego części środkowej. Równocześnie jednak na ogony nakłada się pewne ograniczenia. Modeluje się je rozkładem normalnym o innych parametrach lub inną statystyką. Do opisu takiej sytuacji można stosować podejście zaproponowane przez Johna Tukeya [5]. Założył on, że istnieje duża liczba n danych pomiarowych, wśród których są przypadkowo zmieszane „dobre" i „złe" obserwacje x_i z populacji o średniej wartości μ. Dobre i złe obserwacje pobrane z tej populacji występują odpowiednio z prawdopodobieństwem $(1 - \varepsilon)$ lub ε, gdzie ε – mała liczba. Oba rodzaje obserwacji x_i mają różne rozkłady normalne, tj.: pierwszy – $N(\mu, \sigma^2)$, a drugi – $N(\mu, 9\sigma^2)$, ale o tej samej wartości średniej μ (rys. 1). Odchylenie standardowe „złych" jest 3 razy większe niż „dobrych". Przy założeniu, że wszystkie wartości x_i są niezależne, rozkład wypadkowy $F(x)$ jest opisany wzorem:

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

gdzie $\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{y^2}{2}}\,dy$.

Otrzymuje się rozkład o dolnych częściach zboczy przebiegających nieco powyżej zboczy rozkładu normalnego sporządzonego dla środkowych obserwacji próbki (rys. 1).



Rys. 1. Rozkład mieszany wg Tukey'a uwzględniający dane odstające
$F(x) = (1 - \varepsilon)\, N(\mu, \sigma^2) + \varepsilon\, N(\mu, 9\sigma^2)$ dla $\varepsilon = 0{,}2$

Fig. 1. Tukey joint distribution includings outliers
$F(x) = (1 - \varepsilon)\, N(\mu, \sigma^2) + \varepsilon\, N(\mu, 9\sigma^2)$ for $\varepsilon = 0.2$

Wśród opracowanych dotąd metod odpornościowych i ich algorytmów wdrożonych do stosowania szeroko rozprzestrzeniło się inne podejście, zaproponowane przez Hubera [3] i również uznane już za klasyczne. Wprowadził on pewną stałą wartość *c* zależną od stopnia „zanieczyszczenia" populacji generalnej. Wartość ta służy do określenia granic części środkowej histogramu danych pomiarowych, którą modeluje się rozkładem normalnym jako podstawowym [6, 9–11, 17]. Obserwacje o wartościach w obszarach bocznych – ogonach histogramu danych, występują rzadziej i wg jednego z kryteriów stosowanych dla rozkładu podstawowego, mogą być odstające. Położone skrajnie obserwacje ściąga się na granice obszaru środkowego. Dla nowego ich zbioru następuje zmniejszenie odchylenia standardowego i zwężenie tego obszaru. Dostosowywanie należy powtarzać i przebiega ono dalej iteracyjnie [9] tak długo, aż różnice między wartościami parametrów obliczanymi w kolejnych krokach będą poniżej wymaganego poziomu dokładności obliczeń. Zastosowanie metody Hubera do oceny wartości średniej wyników w badaniu procedur laboratoryjnych i biegłości laboratorium zostało już przedstawione z udziałem autora w kilku czasopismach polskich [9–11].

Dalej zostanie omówiona szczegółowo iteracyjna metoda Hubera przetwarzania danych, należąca do typu o angielskim symbolu IRLS (ang. *iteratively reweighted least squares*). Metodę stosuje się do atestacji dokładności procedur badawczych i kontrolnych dokonywanej poprzez porównanie wyników w badaniach międzylaboratoryjnych. W dwu przykładach liczbowych porówna się wyniki z otrzymywanymi klasycznie i metodą przeskalowanego odchylenia medianowego MAD_S.

Metoda odporna dwukryterialnaHubera z winsoryzacją danych odstających

Statystyka odpornościowa polega na zastosowaniu metod, które zapewniają mniejszy niż modele klasyczne wpływ danych odstających i innych anomalii na wynik

pomiaru. Znajdują one szerokie zastosowanie przy szacowaniu odchyleń standardowych σ. Pod pojęciem odporność rozumie się niewrażliwość na nieprawidłowości i niejednorodności w próbce wywołane przez różne, w ogólnym przypadku nieznane przyczyny. Iteracyjna dwukryterialna metoda przetwarzania danych została opracowana przez Hubera [4, 10]. Dzięki małej wrażliwości na dane odstające, czyli outliery, umożliwia ona poprawę wiarygodności wyników dla próbek o małej liczbie elementów. Zakłada się w niej, że część środkowa rozkładu gęstości prawdopodobieństwa PDF (ang. *probability density function*), obejmująca małe odchylenia od estymatora wartości menzurandu, nie odbiega zbytnio od przyjętego gaussowskiego modelu teoretycznego.

W stosowanej obecnie statystycznej metodzie obliczania niepewności opartej na zaleceniach międzynarodowego Przewodnika GUM [1] zakłada się, że przy powtarzaniu pomiarów uzyskuje się dane, które są próbką pobraną z populacji generalnej modelowanej rozkładem normalnym o wartości średniej μ i odchyleniu standardowym σ. Do obliczenia wyniku pomiarów z danych próbki wyznacza się oceny, czyli estymatory tych obu parametrów statystycznych, odpowiednio jako ich wartość średnią $\bar{x}$, oraz jej odchylenie standardowe jako niepewność pomiaru u_A. Przy występowaniu w pomiarach zmiennych losowo oddziaływań można jedynie ocenić prawdopodobieństwo, z jakim $\bar{x}$ jako estymator wartości menzurandu może znajdować się w danym przedziale, czyli określić jej niepewność u_A. Ocena ta osiąga minimum dla średniej wartości próbki $\bar{x}$, ale jest wrażliwa na dane odstające od pozostałych czyli outliery, które zwiększają wartość niepewności u_A z kwadratem ich odległości od średniej. W klasycznej procedurze statystycznej wykrywa się takie dane wg testów Grubbsa (rozdz. 2) i usuwa się je z próbki. Dla próbek o małej liczbie elementów wiarogodność wyznaczanych statystycznie jej parametrów jest mała i zmaleje jeszcze po usunięciu outlierów. Ponadto otrzymywane w praktyce wyniki nawet dla małych próbek z rozkładu normalnego zwykle rozkładają się niesymetrycznie względem ich wartości średniej. Im próbka jest mniejsza, tym jej asymetria wzrasta i średnia wartość modułu skośności osiąga maximum dla około $n = 6$ elementów. Obliczone jej parametry statystyczne mogą znacznie różnić się od parametrów populacji, czyli nie być poprawne. Badania rozkładów skośności i kurtozy takich małych próbek wykonane metodą Monte Carlo omówiono w rozdziale 5.

W statystyce odpornościowej rozważa się modele rozkładów wraz z danymi odstającymi, czyli outlierami. Cechą charakterystyczną takich rozkładów są bardziej rozciągnięte ogony niż dla normalnego rozkładu gęstości prawdopodobieństwa w postaci funkcji Gaussa. Skorzystanie z tej właściwości pozwala zachować wygodne do analizy założenie o hipotetycznej jednorodności populacji, na którym opierają się wszelkie statystyczne oszacowania i częściowo uwzględnić informację zawartą w odchyleniach odstających. Z danych próbki pomiarowej oblicza się estymator wartości menzurandu, zwykle jako ich wartość średnią. Wskutek występowania wpływów od losowo zmiennych oddziaływań można jedynie ocenić prawdopodobieństwo, z jakim ta wartość może znajdować się w danym przedziale, czyli określić jej niepewność u_A.

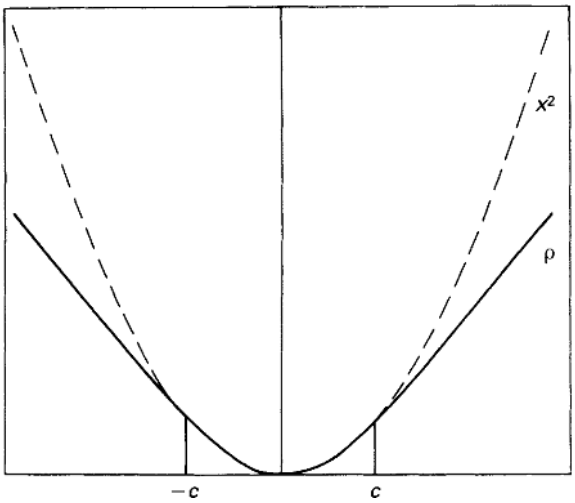
W omawianej tu metodzie odpornej zakłada się, że tylko dla centralnej części rozkładu danych doświadczalnych można przyjąć założenie o normalnym rozkładzie populacji. Dla danych z tej części należy przeprowadzać uśrednianie, tj. stosować

metodę najmniejszych kwadratów MNK. Bardziej odporne na duże odchylenia jest podane przez Laplace'a kryterium modułowe MNM. Dlatego też w tej metodzie odpornej dokonuje się symbiozy obu kryteriów. Dla centralnej części wartości danych stosuje się metodę najmniejszych kwadratów, a poza granicami tego przedziału, wykorzystuje kryterium modułowe w celu zmniejszenia wpływu danych odstających (outlierów), ale ich nie usuwa się. Dla zmniejszenia czułości na występowanie odstających wartości danych zaproponowano [5], by w obliczeniach stosować modyfikację położenia danych o modułach odchyleń ε_i opisaną wyrażeniem

$$|\varepsilon_i| = |x_i - \hat{\mu}| > \varphi \tag{2}$$

gdzie: $\varphi = c\sigma$ – odległość od środka grupowania danych $\hat{\mu}$, c – stała określająca przyjęty stopień odporności, σ – odchylenie standardowe populacji.

Dla obserwacji o odchyleniach mniejszych od $c\sigma$, funkcja czułości $\rho(\varepsilon)$ jest kwadratowa, a dla większych minimalizuje się moduł odchyleń $|\varepsilon| = |x_i - \mu|$. Te obszary czułości metody ilustruje rys. 2.



Rys. 2. Względna funkcja czułości ρ stosowana w metodzie odpornościowej
Uwaga: $\rho(\varepsilon) = \varepsilon^2$ dla $|\varepsilon| \le c$ oraz $\rho(\varepsilon) < \varepsilon^2$ dla dużych $|\varepsilon|$

Fig. 2. The relative sensitivity function ρ used in a robust method
Note: $\rho(\varepsilon) = \varepsilon^2$ for $|\varepsilon| \le c$ and $\rho(\varepsilon) < \varepsilon^2$ for large $|\varepsilon|$

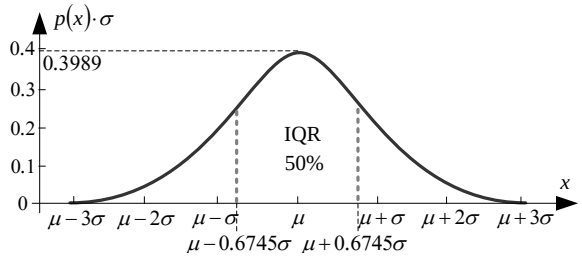
Wybrana w taki sposób funkcja $\rho(\varepsilon)$ pozwala na mniej ostre podejście do danych odstających na wartość $|\varepsilon| > c\sigma$ od środka grupowania rozkładu danych. Stała c określa stopień odporności. Wartość tej stałej zależy od stopnia zanieczyszczenia rozkładu. Tak więc dla zanieczyszczenia 1% przyjmuje się $c = 2$, a dla zanieczyszczenia 5% $c = 1{,}4$. Zwykle wybiera się $c = 1{,}5$. Zgodnie z tak wybranym kryterium modyfikuje się skrajne z pozyskanych danych, a mianowicie:

$$x_i^* = \begin{cases} x_i & \text{przy } |x_i - \hat{\mu}| \le c\sigma, \\ \hat{\mu} - c\sigma & \text{przy } x_i < \hat{\mu} - c\sigma, \\ \hat{\mu} + c\sigma & \text{przy } x_i > \hat{\mu} + c\sigma \end{cases} \tag{3}$$

gdzie $\hat{\mu} = med\{x_i\}$ danych x_i uszeregowanych rosnąco.

Opisane wzorami (3) przetworzenie danych nazywane jest *winsoryzacją* od nazwiska matematyka amerykańskiego Winsora.

Parametry statystyczne populacji generalnej μ i σ zwykle nie są znane i trzeba używać ich estymat wyznaczonych z danych próbki. Jako odporną na dane odstające ocenę środka grupowania rozkładu $\hat{\mu}$ próbki wstępnie preferowana jest mediana $med\{x_i\}$. Badania wykazały [9], że najlepszy jest środek rozstępu między trzecim (p = 3/4) i pierwszym (p = 1/4) kwartylami próbki (ang. *inter-quartile midrange*) (rys. 3). Ich odległość od tego środka nazywana jest *odchyleniem ćwiartkowym* [16].



Rys. 3. Definicja rozstępu międzykwartylowego (*a*, *b*) (o odchyleniu ćwiartkowym). Linie przerywane – rzędne: 1. kwartyla $a = \mu - 0{,}6745\sigma$ oraz 3. kwartyla $b = \mu + 0{,}6745\sigma$

Fig. 3. Definition of inter-quartile midrange (*a*, *b*): Dotted lines – ordinates of first quartile $a = \mu - 0.6745\sigma$ and third quartile $b = \mu + 0.6745\sigma$

Długość odcinka między kwartylami wynika jednoznacznie z rozkładu gęstości prawdopodobieństwa przyjętego dla rozrzutu danych próbki. Powierzchnia pod krzywą rozkładu dla tego przedziału wynosi 50%.

Odchylenie standardowe populacji σ występujące we wzorze (3) zwykle nie jest znane. Jako początkowe oszacowanie zakresu przejścia z rozkładu pełnego do obciętego wprowadza się bezwzględne odchylenie medianowe MAD (ang. *median absolute deviation*), które jest oszacowaniem zakresu przejścia z rozkładu pełnego do obciętego

$$MAD_n = med\left\{\left|x_i - M_n\right|\right\} \tag{4}$$

gdzie: $M_n = med\{x_i\}$, n – liczba elementów w próbce.

Dla ustalenia zależności między parametrami rozkładu PDF obciętego oraz przewidywanego dla całej populacji, tj. do spełnienia warunku skalowania, trzeba przeliczyć odchylenie standardowe σ za pomocą współczynnika korygującego, który wyznacza się z początkowego rozkładu gęstości prawdopodobieństwa:

$$f(x) = \frac{d}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{5}$$

W rozkładzie nieobciętym $d = 1$, zaś w obciętym w kwartylach

$$d = \frac{0{,}5}{\Phi(\frac{b-\mu}{\sigma}) - \Phi(\frac{a-\mu}{\sigma})} \tag{6}$$

gdzie: $\Phi(z)$ – znormalizowana skumulowana funkcja (CPDF) rozkładu normalnego.

Z tabeli rozkładu normalnego wynika

$$d = \frac{1}{0{,}6745} = 1{,}4826 \approx 1{,}483. \tag{7}$$

Tak więc zmiana skali dla przejścia od rozkładu nieobciętego do obciętego wynosi: 1,483. Ocena odchylenia standardowego s^* próbki o liczbie obserwacji n, wyznaczona na podstawie znormalizowanego rozstępu między kwartylami, wynosi

$$s^* = 1{,}483 \cdot MAD_n \tag{8}$$

Różni się ona od odchylenia standardowego próbki s wg Przewodnika GUM [1], tj.:

$$s = \sqrt{\frac{1}{n-1}\sum (x_i - \bar{x})^2} \tag{9}$$

Huber [10] wykazał, że odchylenie s^* stanowi nieobciążoną ocenę odchylenia standardowego σ i będzie bardzo stabilne nawet, jeśli próbka zawiera do 50% odstających wartości obserwacji (outlierów). Aby uzyskać stabilność przy występowaniu takich danych, należy dla wybranej funkcji ρ wyznaczyć medianę *med* oryginalnego zestawu danych. Wartość tę przyjmuje się jako początkową ocenę x_0^* środka grupowania danych jako bardziej odporną na wartości odstające niż ich średnia. Następnie oblicza się MAD z (4) i s_1^* z (8) oraz po wyborze $c = 1{,}5$ znajduje się granice przejścia $\pm\varphi_1 = c \cdot s_1^*$ między obszarami stosowania metody najmniejszych kwadratów MNK i metody najmniejszego modułu MNM.

Stosując warunki (3), otrzymuje się zmodyfikowany szereg danych pierwszego kroku iteracji $x_{i(1)}^*$ o wartości średniej

$$\bar{x}_1^* = \frac{1}{n}\sum_{i=1}^{n} x_{i(1)}^* \tag{10}$$

Po podstawieniu x_i^* i $\bar{x}_1^*$ wyznacza się z (8) – odchylenie standardowe $s_2(x^*)$. Obliczoną wartość wykorzystuje się z kolei do obliczenia nowego stabilnego odchylenia standardowego

$$s_2^* = 1{,}134 \cdot s_2(x^*) \tag{11}$$

Współczynnik 1,134 odpowiada wartości c = 1,5 dla rozkładu normalnego. Otrzymaną wartość s_2^* wykorzystuje się do obliczenia nowej granicy φ_2 i kontynuuje się procedurę w taki sam sposób, jak omówiony powyżej. Na podstawie wartości $\bar{x}_j^*$ oraz $\bar{x}_{j-1}^*$ obliczonych w bieżącym j i poprzednim j–1 kroku iteracji określa się zbieżność algorytmu. Procedurę powtarza się, aż zmiany $\bar{x}_j^*$ i s_j^* między kolejnymi krokami staną się minimalne.

Zastosowanie metody przedstawiono dalej w przykładzie 1. Dotyczy on wykorzystania tej metody odpornej do weryfikacji pewnej metody badawczej przez porównanie wyników pomiarów badania jednorodnej próbki, uzyskanych w wielu laboratoriach. Wymaga to omówienia istoty takiego porównania.

Istota porównań międzylaboratoryjnych

W celu sprawdzenia lub walidacji [14] danej metody pomiarowej prowadzi się wspólny eksperyment, w którym uczestniczą wybrane laboratoria o odpowiednim przygotowaniu merytorycznym. Rozkłady wyników poprawnych badań laboratoryjnych powinny z założenia asymptotycznie zbliżać się do rozkładu normalnego, ale tak nie jest i właśnie takie porównanie ujawnia nagminne występowanie wartości odstających. Wstępnie przyjmuje się, że pomiary wykonywane przez wszystkie uczestniczące w eksperymencie laboratoria mają jednakową powtarzalność. Jednak zdarza się w praktyce, że powtarzalność wyników w niektórych z nich jest z obiektywnych powodów gorsza niż w pozostałych. Przy stosowaniu klasycznych testów statystycznych oddziela się z danych wyniki odstające [10], a pozostałe dane służą do wyznaczenia wariancji jako oceny odtwarzalności. Bardzo często niektóre z tych wyników są bliskie wartości granicznej oddzielającej odchylenia quasi-odstające i odstające. Nieuwzględnienie wszystkich wyników wpływa znacząco na wiarygodność oszacowania wariancji. Właśnie aby tego uniknąć, preferuje się stosowanie metod odpornych.

Międzylaboratoryjne badania atestacyjne (ang. *key comparisons*) polegają na określaniu poprawności (ang. *trueness*) badanej metody [3] i jej powtarzalności (ang. *repeatability*) szacowanej na podstawie średniej dyspersji wyników otrzymanych przez laboratoria. Stanowi ona obiektywną ocenę rzeczywistych technicznych i organizacyjnych możliwości danego laboratorium. Według [3], dotychczas przy ocenie wskaźników precyzji – powtarzalności i odtwarzalności (ang. *reproducibility*) usuwa się odstające wartości wyników obserwacji za pomocą testu Grubbsa. Usunięcie najgorszych danych uzyskanych przez niektóre laboratoria idealizuje rzeczywiste warunki badań i zwiększa niepewność wspólnego wyniku. Nieuwzględnienie odstających wyników zmniejsza wiarygodność statystycznych ocen dokładności, tj. powtarzalności i odtwarzalności metody pomiarowej badanej w porównaniach międzylaboratoryjnych. Jest to szczególnie istotne przy kosztownych testach i próbach niszczących, gdy liczba badanych obiektów i danych w każdej z próbek jest

zwykle ograniczona. W przykładzie dwoma metodami odpornymi na wartości odstające wyznaczy się i porówna wyniki i niepewności uzyskane przez laboratoria.

Przykład 1

Dane liczbowe dla tego przykładu 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 $n = 9$ następujących wartości średnich x_i:

$$\underline{24{,}140} \quad 20{,}155 \quad 19{,}500 \quad 20{,}300 \quad 20{,}705$$
$$\underline{17{,}570} \quad 20{,}100 \quad 20{,}940 \quad 21{,}185.$$

Dwa podkreślone wyniki (x_1 oraz x_6) istotnie odstają od pozostałych. Przy modelu tradycyjnym (*kontaminacji*) zakłada się, że poprawne obserwacje pochodzą tylko z rozkładu normalnego. Konsekwencją jest zastosowanie testu, np. Grubbsa, do znalezienia wartości odstających i ich odrzucenie. Z pozostałych wyznacza się wynik wspólny dla całego eksperymentu jako wartość średnią $\bar{x} = 20{,}41$ i standardowe odchylenie $s = 0{,}50$. Tak obliczone s tylko na podstawie 7 wartości będą jako oceny mniej wiarygodne statystycznie niż dla 9 danych.

Jeśli natomiast założy się, że wszystkie 9 obserwacji pochodzi z *pojedynczego rozkładu niegaussowskiego*, to logiczną konsekwencją jest zastosowanie algorytmu, który przy obliczaniu wyniku nie usuwa żadnej z wartości danych otrzymanych w pomiarach. Stosuje się to w dwu rozpatrywanych tu metodach odpornych: o przeskalowanym odchyleniu medianowym MAD_S i iteracyjnej dwukryterialnej.

Obliczenie przeskalowanego odchylenia medianowego MAD_S

Sposób postępowania w tej metodzie jest bardzo prosty:

- wyznacza się ze wszystkich $n = 9$ danych x_i medianę *med* i uznaje ją za estymator wartości wyniku pomiaru,
- następnie oblicza się odchylenie medianowe (ang. *median absolute deviation*) *MAD* wg wzoru(4),
- za standardową niepewność pomiaru $s(x)$ uznaje się *przeskalowane odchylenie medianowe* MAD_S

$$s(x) \equiv MAD_s = \kappa \cdot MAD \tag{7}$$

Wartość asymptotyczna $\kappa(\infty) = 1{,}483$ (odpowiadająca $d = 1{,}483$ dla populacji generalnej) jest stosunkiem $s(x)/MAD$ dla rozkładu normalnego w granicy $n \to \infty$.

Według Zięby [17] taką odporną procedurę proponowali Randa (2000), Burke (2001), Müller (2001) i nieco inną Cox (2002). W rozdziale 5 swojej książki [17] Zięba, za opublikowanym w Internecie preprintem Randy z 2005 r. [16], podał modyfikację tej metody, tj. by zamiast wartości asymptotycznej $\kappa(\infty) = 1{,}483$, która dla próbek o skończonej liczbie elementów n zaniża wartość oceny wyniku pomiarów, stosować wartości $\kappa(n)$ z tabeli 1.

Tab. 1. Wartości mnożnika $\kappa(n)$ dla prób losowych o liczebności n wg [17]
Tab. 1. Values of coefficient $\kappa(n)$ for random samples of n elements [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 próbki o n danych spełniony jest warunek $s(x_n) > s(x_\infty)$. Stąd dla analizowanego zbioru 9-ciu danych otrzymuje się:

$$med = 20{,}3; \quad MAD = 0{,}64$$

oraz

$$s(x_\infty) \equiv 1{,}483 \cdot 0{,}64 = 0{,}95$$

$$s(x_n) \equiv \kappa(n) \cdot \mathrm{MAD} = 1{,}633 \cdot 0{,}64 = 1{,}045.$$

Według obliczeń tą metodą otrzymuje się estymatę wartości menzurandu med =20,3 i estymatę jej odchylenia standardowego $s(x_n)$ = 1,045.

Obliczenia oceny metodą iteracyjną Hubera

W tabeli 2 podano przykład obliczeń odpornego wyniku pomiarów wg metody Hubera, nazwanej „Algorytmem A" w aneksie C normy PN-ISO 5725-5 [3]. W kolumnie „0" tej tabeli zamieszczono w porządku rosnącym wyniki pomiarów otrzymane w i = 1, …, 9 laboratoriach jako dane $x^*_{i(0)}$ stanu początkowego iteracji. Natomiast kolumny 1–4 zawierają kolejno dane dla następnych kroków iteracji.

Dla danych początkowych oblicza się:

wartość średnią:

$$\bar{x}^*_0 = \frac{1}{n}\sum_{i=1}^{9} x_{i(0)} = 20{,}511$$

i odchylenie standardowe próbki:

$$s^*_0 = \sqrt{\frac{1}{n-1}\sum_{i=1}^{9}(x^*_{i\,(0)} - \bar{x}^*_0)^2} = 1{,}727$$

gdzie: $x^*_{i(0)}$ – dane z kolumny „0" tabeli 2.

Jako początkowe (dla surowych danych $x^*_{i(0)}$) oszacowanie położenia centrum grupowania populacji przyjmuje się medianę

$$x^*_0 \equiv med\,(x^*_{i(0)}) = x^*_{5(0)} = 20{,}300$$

i z (6) jej odchylenie standardowe

$$s^*_1 = 1{,}483 \cdot med\left\{\left|x^*_{i(0)} - x^*_0\right|\right\}$$

gdzie: seria różnic $\left|x^*_{i(0)} - x^*_0\right|$ przyjmuje wartości:

2,73 0,80 0,20 0,145 0 0,405 <u>0,640</u> 0,889 3,84.

Tab. 2. Przykład obliczeń odpornego wyniku pomiarów
Tab. 2. Example of calculation of the robust measurement result

Nr iteracji j	0	1	2	3	4
φ_j		1,424	1,478	1,514	1,539
$x^*_j - \varphi_j$ $x^*_j + \varphi_j$		18,876 21,724	18,909 21,865	18,893 21,921	18,872 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

Otrzymuje się $s^*_1 = 1{,}483 \cdot 0{,}640 = 0{,}949$ i następnie

$$\varphi_1 = 1{,}5 \cdot s^*_1 = 1{,}5 \cdot 0{,}949 = 1{,}424$$

Określa ona granice dla odchyleń przejście z metody o najmniejszej sumie kwadratów MNK do metody najmniejszego modułu MNM:

$$x^*_0 - \varphi_1 = 20{,}300 - 1{,}424 = 18{,}876$$

$$x^*_0 + \varphi_1 = 20{,}300 + 1{,}424 = 21{,}724$$

Ponieważ $x^*_{1(0)} < (x^*_0 - \varphi_1)$, $x^*_{9(0)} > (x^*_0 + \varphi_1)$, to wartości $x^*_{1(0)}$, $x^*_{9(0)}$ wykraczają poza te progi i w kolumnie 1 danym $x^*_{1(1)}, x^*_{9(1)}$ przypisuje się wartości progów, tj.:

$$x^*_{1(1)} = 18{,}876, \quad x^*_{9(1)} = 21{,}724,$$

a pozostałe dane $x^*_{2(1)} = x^*_{2(0)}$, …, $x^*_{8(1)} = x^*_{8(0)}$ nie zmieniają się.

Po podobnych jak poprzednio obliczeniach otrzymuje się:

wartość średnią $\bar{x}^*_1 = \frac{1}{n}\sum_{i=1}^{9} x_{i(1)} = 30{,}387;$

i odchylenie standardowe $s^*_1 = 0{,}869.$

Do ich obliczenia użyto wartości $x^*_{i(1)}$ pobrane z kolumny „1".
Centrum grupowania rozkładu będzie teraz

$$x^*_1 = \bar{x}^*_1 = 30{,}387.$$

Następnie wyznacza się nową wartość

$$s^*_2 = 1{,}134\sqrt{\frac{1}{n-1}\sum_{i=1}^{9}(x^*_{i\,(1)} - x^*_1)^2} = 0{,}985$$

gdzie: $med\left\{\left|x^*_{i(1)} - x^*_1\right|\right\} = 0{,}664.$

Górna i dolna granice wynoszą teraz odpowiednio

$$x^*_1 - \varphi_2 = 18{,}909 \quad \text{i} \quad x^*_1 + \varphi_2 = 21{,}865,$$

gdzie: $\varphi_2 = 1{,}5 \cdot s^*_2 = 1{,}5 \cdot 0{,}985 = 1{,}478.$

W drugim etapie iteracji otrzymuje się wartości zmodyfikowane:

$$x^*_{1(2)} = x^*_1 - \varphi_2 = 18{,}909 \text{ i } x^*_{9(2)} = x^*_1 + \varphi_2 = 21{,}865.$$

Wprowadza się zmienione dane $x^*_{1(2)}$ i $x^*_{9(2)}$ do kolumny 2 i wraz z niezmodyfikowanymi danymi $x^*_{i(0)}$ (i = 2, ..., 8) stosuje się je w tym etapie iteracji. Wyniki obliczeń umieszcza się w kolejnej kolumnie 3 i dalej realizuje się ten sam algorytm jak poprzednio. Wartości średnie dla kroku czwartego i piątego, tj. $\bar{x}^*_4 = 20{,}411$ oraz $\bar{x}^*_5 = 20{,}412$, praktycznie nie różnią się. Odporne oceny odchylenia standardowego s^*_i wartości średniej pomiarów uzyskanych w laboratoriach są też zbliżone, tj. $s^*_4 = 1{,}026$ i $s^*_5 = 1{,}039.$

Analiza wyników obliczeń

Tradycyjne podejście polega na zidentyfikowaniu wartości odstających (outlierów) za pomocą kryterium Grubbsa i wyeliminowaniu z dalszego przetworzenia wyników krańcowych $x^*_{1(0)} = 17{,}570$ i $x^*_{9(0)} = 24{,}140$. W tym przypadku bierze się pod uwagę wyniki tylko z $n = 7$ laboratoriów. Otrzymuje się wówczas średnią wartość $m = 20{,}4$ i oszacowane dla nich odchylenie standardowe próbki $s = 0{,}501$. Niepewność wspólnych badań będzie mniej wiarygodna, gdyż wg dodatku E.1 do GUM [1, 2] ma ona większe własne odchylenie standardowe niż odchylenie s^* wyznaczone metodami odpornymi dla $n = 9$.

Wartości centrum grupowania danych otrzymane dla obu metod odpornych są zbieżne dlatego, że oceniano dane i dokładność wyników tego samego badania. Ich niepewności różnią się tylko o 9%, ale ponieważ dla próbki o 9 danych dokładność oceny jest ok. $1/\sqrt{2(n-1)} = 25\%$, to zbieżność obu metod odpornych jest w praktyce niemal całkowita. Korzystając z metody dwukryterialnej uzyskuje się $s^* = 1{,}039 > s$ dobrze odpowiadające rzeczywistej dyspersji wartości otrzymanych wyników badań. Po usunięciu dwu krańcowych wyników początkowych $x^*_{1(0)}$ i $x^*_{9(0)}$ otrzymano mniejszą wartość s. Odbiega to od rzeczywistości, gdyż idealizuje się pomiar zakładając, że dane eksperymentalne należą do populacji generalnej o rozkładzie normalnym.

Dla potwierdzenia tej tezy rozpatrzony zostanie prosty przykład o symulowanych danych pomiarowych.

Przykład 2

Dane pobrano z populacji o rozkładzie normalnym o wartości średniej $m = 75{,}238$ oraz odchyleniu standardowym $\sigma = 13{,}475$. Próbka zawierała tylko cztery wartości:

$$75{,}3 \quad 76{,}0 \quad 76{,}3 \quad 102{,}1$$

Stosując kryterium Grubbsa, największa wartość z próbki (102,1) zostanie uznana jako odstająca, chociaż w rzeczywistości $m + 2\sigma = 102{,}201$. Z odchylenia medianowego MAD_S otrzymuje się estymator $s(x) = 1{,}0095$ daleko odbiegający od rzeczywistej wartości odchylenia standardowego $\sigma = 13{,}475$.

Natomiast stosując omówiony algorytm iteracyjny otrzyma się wartość $s^*_{28} = 14{,}882$ dużo bliższą rzeczywistej wartości σ i bardziej wiarygodnie opisującą możliwy rozrzut wartości obserwacji pomiarowych. Otrzymano ją dopiero po 28 iteracjach. Nie stanowi to jednak problemu przy stosowaniu współczesnej techniki obliczeniowej.

Podsumowanie

Metoda przeskalowanego odchylenia medianowego jest bardzo prosta, ale nie daje poprawnych rezultatów, gdy wartość odstająca daleko odbiega od grupy pozostałych danych. Uwidacznia się to szczególnie dla mało licznych próbek.

Metoda Hubera (Algorytm A wg normy [3]) ma bardziej skomplikowany, lecz łatwy do zautomatyzowania algorytm. Dzięki wprowadzeniu progu $\pm c\sigma$ zmniejszającego czułość na dane odstające, metoda ukierunkowana jest na wyznaczenie odpornej oceny niepewności.

Przeprowadzona analiza wykazała przydatność zastosowania rozważanej odpornej metody dwukryterialnej do wyznaczania parametrów statystycznych próbek o małej liczbie danych, gdy pobrane są one z populacji generalnej o założonym rozkładzie normalnym, ale zawierają wyniki istotnie odstające od pozostałych. Pozwala ona ocenić obiektywnie wartość wyniku i dokładność metody badań.

Ponadto występujące w przykładzie różnice między statystycznymi ocenami niepewności metody pomiaru wyznaczonymi przez poszczególne laboratoria i niepewnością wyniku całego międzylaboratoryjnego eksperymentu można wykorzystać jako podstawę obiektywnej analizy rzeczywistego stanu organizacji procesu badań i sposobu prowadzania działalności w tych dwu laboratoriach, które miały wyniki odstające.

Zastosowanie metod odpornych w badaniach laboratoryjnych to zagadnienie obszerne i aktualne, które znacznie wykracza poza ramy tej monografii. Niektóre problemy z tej dziedziny zostały rozpatrzone szczegółowo w kilku polskich i anglojęzycznych publikacjach profesorów Eugenija Volodarskego i Larysy Koshevej z Kijowa opracowanych wspólnie z autorem i innymi współautorami [19–26]. Opublikowali też oni monografię w języku rosyjskim o technicznych aspektach akredytacji laboratoriów badawczych, wydaną przez Narodowy Uniwersytet Techniczny w Vinnicy na Ukrainie [27].

Literatura

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*. Tłumaczenie GUM z 1995 r. z komentarzem J. Jaworskiego, Wydawnictwo Głównego Urzędu Miar Alfavero Warszawa 1999.
3. ISO 5725:1994 – Accuracy (trueness and precision) of measurement methods and results.
4. Sarhan A.E., Greenberg B.G. (eds), *Contributions to order statistics*, John Wiley & Sons, 1962.
5. Tukey J.W., *Exploratory Data Analysis*. Addison-Wesley. 1978.
6. Domański Cz., Pruska K., *Nieklasyczne metody statystyczne*. Polskie Wydawnictwo Ekonomiczne, 2000.
7. Frosch Møller S., von Frese J., Bro R., *Robust methods for multivariate data analysis*. "Journal of Chemometrics", Vol. 19, 2005, 549–563.
8. Doksum Kjell A., *Mathematical Statistics: Basic and Selected Topics*. 1. Second updated edition, Pearson Prentice-Hall 2007.
9. Olive D.J., *Applied Robust Statistics*. Southern Illinois University Department of Mathematics. June 23, 2008.
10. Huber Peter J., Ronchetti Elvezio M., *Robust Statistics*. 2nd edition. Wiley 2011.

11. Daszykowski M., Kaczmarek K., Vander Heyden Y., Walczak B., *Robust statistics in data analysis – A review: Basic concepts*. "Chemometrics and Intelligent Laboratory Systems", Vol. 85, Issue 2, 2007, 203–219.
12. Stanimirova I., Serneels S., Van Espen P.J., Walczak B., *How to construct a multiple regression model for data with missing elements and outlying objects*. "Analytica Chimica Acta", Vol. 581, Issue 2, 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 2000, wydanie nowe 2012.
15. Buchcik D., *Procedura kalibracji przy wykorzystaniu metody najmniejszej mediany w przypadku modeli wielowymiarowych*. „Pomiary Automatyka Kontrola", Nr 5, 2008, 294–297.
16. Randa J., *Proposal for KCRV and Degree of Equivalence for GTRF Key Comparisons*. NIST, (2000) and update (2005). Preprint dostępny on-line.
17. Zięba A., *Analiza danych w naukach ścisłych i technice*. Wydawnictwo Naukowe PWN, Warszawa 2013.
18. http://pl.wikipedia.org/wiki/Statystyka_odpornościowa.
19. Volodarsky E., Warsza Z.L., Kosheva L., *Odporna ocena dokładności metod pomiarowych*. „Pomiary Automatyka Kontrola", Vol. 58, Nr 4, 2012, 396–401.
20. Volodarsky E., Warsza Z.L., Kosheva L., Palanyczenko D., *Zastosowanie estymacji odpornej w badaniach biegłości laboratorium przy niewielkiej liczbie pomiarów*. „Pomiary Automatyka Kontrola", Vol. 59, Nr 6, 2013, 554–557.
21. Volodarski E., Warsza Z.L., *Zastosowanie statystyki odpornościowej na przykładzie badań międzylaboratoryjnych*. „Przegląd Elektrotechniczny", 11/2013, 260–267.
22. Volodarsky E., Warsza Z.L., Koshevaya L.A., *System oceny i zapewnienia jakości badań biegłości laboratoriów przy ich akredytacji*. „Przemysł Chemiczny", Vol. 93, Nr 8, 2014, 1252–1254.
23. Volodarsky E., 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), Vol. 3 "Measuring Techniques and Systems". (ISBN 978-3-319-15834-1), Vol. 352 of series: *Advances in Intelligent Systems and Computing* (ISSN 2194-5357) Springer 2015, 285–291.
24. Volodarsky E., Warsza Z.L., *Robust estimation in interlaboratory measurements with small number of measurements*. "Measurement Automation Monitoring", Vol. 61, No 4, 2015, 104–110.
25. Volodarsky E., Warsza Z.L., Kosheva L., Idźkowski A., *Robust Algorithm S to assess precision of interlaboratory measurements*. "Measurement Automation Monitoring", Vol. 61, No 4, 2015, 111–114.
26. Volodarsky E., Warsza Z.L., Kosheva L., Idźkowski A., *Assessment of precision of the interlaboratory test data by using robust "Algorithm S"*. [in monograph:] R. Jabłoński T. Brezina (Editors) Advanced Mechatronics Solutions, Vol. 393 of series *Advances in Intelligent Systems and Computing*. Springer 2016, 87–96.
27. Volodarski E. T., Kosheva L. A., *Technicheskije Aspekty Akreditacii Ispytatelnykh Laboratorij* – monografia, Winnicki Narodowy Uniwersytet Techniczny Ukrainy VNTU Vinnica 2013.

Rozdział 12

Zasady wyznaczania wyników pomiarów pośrednich wielowymiarowych

Streszczenie. Omówiono zasady wyznaczania wyników w pomiarach pośrednich wielowymiarowych multimenzurandu, czyli estymatorów wartości, niepewności i współczynników korelacji wielu wielkości powiązanych statystycznie. Ujmuje je Suplement 2 do Przewodnika GUM opublikowany w 2012 r. Podano opis takich pomiarów za pomocą algebry wektorów losowych oraz przykład liniowej transformacji danych pomiarowych dwu wielkości. Omówiono skutki nieprawidłowego zaokrąglania przy dokładnym przetwarzaniu danych wejściowych oraz dla danych wyznaczanych z próbek o małej liczbie pomiarów. Omówiono propozycję uściślenia podanej w Przewodniku GUM poprawki do liniowej propagacji wariancji przy wyznaczaniu niepewności pomiarów o funkcjach nieliniowych. Przedstawiono propozycję opracowania międzynarodowych zasad publikowania danych pomiarowych z badań eksperymentalnych w dualnej postaci: tak jak dotychczas – na papierze i towarzyszącej dostępnej w Internecie publikacji z oryginalnymi wynikami pomiarów, aby użytkownik mógł sam pozyskiwać i weryfikować potrzebne mu dane eksperymentalne.

We wszystkich dyscyplinach naukowych opartych na eksperymencie, w przemyśle i w innych dziedzinach gospodarki wykonuje się coraz więcej pomiarów pośrednich wielu wielkości powiązanych ze sobą statystycznie, czyli wyznacza się parametry multimenzurandu wielowymiarowego (ang. *multivariate*). Na przykład pozyskuje się równocześnie wartości sygnałów wyjściowych czujników o parametrach zależnych od wielu wielkości, przetwarza się i je wyniki wykorzystuje do monitorowania i sterowania w czasie rzeczywistym (on-line). Wartości te podlegają rozrzutom losowym i mogą być ze sobą w różny sposób skojarzone, nie tylko deterministycznie, ale i statystycznie przez obiekt badany, system pomiarowy i oddziaływania zewnętrzne. Z danych pomiarowych wyznacza się estymatory wartości, niepewności i współczynników korelacji kilku wielkości analitycznie powiązanych z mierzonymi bezpośrednio, czyli mierzy się je pośrednio. Wyznaczania niepewności i zaokrąglania wyników pomiarów pośrednich parametrów multimenzurandu oraz prezentacji ich danych cyfrowych wykonuje się obecnie dość dowolnie, gdyż proces tworzenia międzynarodowych przepisów metrologicznych nie może nadążyć za zróżnicowanymi w różnych dziedzinach potrzebami techniki pomiarowej. Międzynarodowy przewodnik o wyrażaniu niepewności pomiarów GUM [1] oraz poradnik NASA [2] zawierają zalecenia dla szacowania wyników pomiarów pojedynczych wielkości. Niezbędna jest też standaryzacja procedur wyznaczania, przetwarzania, oceny dokładności i sposobu publikowania wyników pomiarów multimenzurandu, aby otrzymywane dane miały wysoką jakość i odpowiednią wiarygodność metrologiczną. Zagadnieniom tym jest poświęcony dopiero niedawno ukończony Suplement 2 do GUM, dostępny w Internecie [1a]. Suplement ten liczy ponad 70 stron. Poza

wstępem i odniesieniami do innych przepisów związanych z nim tematycznie obejmuje: pojęcia i definicje, notację i podstawy wektorowego opisu pomiarów wielowymiarowych, wielowymiarową funkcją gęstości prawdopodobieństwa, wyznaczanie niepewności multimenzurandu z uwzględnieniem dotychczasowych zasad GUM, zastosowanie metody Monte Carlo, 5 przykładów, 4 aneksy i bogatą bibliografię. Pomimo tak bogatej treści nie zawiera on omówienia szeregu zagadnień niezbędnych do wdrożenia w praktyce pomiarowej jednolitego sposobu szacowania wyników pomiarów multimenzurandu oraz zasad przetwarzania i rozpowszechniania jego danych.

W literaturze polskiej pomiary wielowymiarowe omawiano tylko ogólnie i zaledwie w kilku publikacjach [3, 12–16]. Zagadnienia wyznaczania estymatorów parametrów statystycznych dla tych pomiarów zaczęto rozpatrywać bardziej szczegółowo dopiero w latach 2010–2012 [6–9, 12]. Ich zaczynem była inicjatywa autora, by zaprosić do udziału w Kongresie Metrologii KM 2010 w Łodzi dr. Vladimira Ezhelę – głównego specjalistę w Ośrodku Przetwarzania Danych rosyjskiego Instytutu Fizyki Wielkich Energii (Rosatom). W pracach zgłoszonych na KM 2010 [4, 5], opartych na długoletnim doświadczeniu z przetwarzania wyników wielowymiarowych eksperymentów w badaniach podstawowych z fizyki, Ezhela przedstawił propozycję ujednolicenia wymagań dla zaokrąglania i przetwarzania danych takich pomiarów oraz konieczność zmiany formy prezentacji ich wyników na dualną – drukiem i elektronicznie. W szczególności dotyczyłoby to wielkości wyznaczanych pośrednio z tych samych wejściowych danych eksperymentalnych i z wartości podstawowych stałych fizycznych publikowanych przez CODATA [11]. Rozdział ten opiera się na treści kilku wspólnych prac z W. Ezhelą [6–9]. Autor starał się w nich połączyć wieloletnie jego doświadczenie w badaniach podstawowych i przetwarzaniu danych w fizyce i własne z techniki pomiarowej i metrologii. Najważniejsze zagadnienia i wnioski omawia się poniżej.

Opis danych pomiarów wielowymiarowych czyli multimenzurandu

Mierzoną pojedynczą wielkość skalarną opisuje się zmienną losową X. Minimalna struktura danych wyrażająca wynik pomiarów takiego menzurandu zawiera estymator współrzędnej punktu skupienia dla rozrzutu wartości obserwacji pomiarowych i towarzyszącą mu rozszerzoną niepewność pomiaru [1]. Jest to przedział wyrażany jako wielokrotność odchylenia standardowego σ, w którym z zadanym prawdopodobieństwem, czyli dla wymaganej ufności, występuje określony estymator. Jako estymator wartości mierzonej przyjęto w GUM wartość średnią $\overline{x}$, która ma najmniejsze odchylenie standardowe dla próbek danych z populacji o rozkładzie normalnym. Jednakże dla próbek z populacji o rozkładach niegaussowskich mniejsze odchylenia standardowe mają inne estymatory, np. środek rozstępu dla rozkładu równomiernego (rozdz. 6) i zbliżonych doń rozkładów trapezowych (rozdz. 7), dla których jeszcze dokładniejszy jest estymator dwuelementowy (rozdz. 8).

W ogólnym przypadku pośrednich pomiarach wielowymiarowych, po przetworzeniu zbioru bezpośrednio mierzonych m wielkości wejściowych X_i wg funkcji $F_i(X_1, \ldots, X_m)$ uzyskuje się zbiór n wielkości wyjściowych Y_i. Jako wyniki pomiarów wyznaczane są estymatory wartości poszczególnych wielkości składowych, ich

odchylenia standardowe σ_i lub niepewności rozszerzone $u(X_i)$ i wzajemne powiązania statystyczne opisane przez kowariancje σ_{ik} lub współczynniki korelacji ρ_{ik}.

Przy przetwarzaniu danych trzech lub więcej skojarzonych losowo wielkości mierzonych opis analityczny wielowymiarowego rozkładu normalnego komplikuje się. Opis operacji dokonywanych na wielowymiarowych danych pomiarowych upraszcza się przy zastosowaniu algebry wektorów losowych. Zbiór mierzonych wielkości losowych X_i przedstawia się wówczas przez wejściowy wektor losowy $\boldsymbol{X}$, zaś zbiór wyznaczanych pośrednio danych wielkości losowych Y_i (ang. *observables*) tworzy wyjściowy wektor $\boldsymbol{Y}$. Zależność obu wektorów [3, 6, 12–14] zapisuje się ogólnie jako

$$\boldsymbol{Y} = \boldsymbol{F}(\boldsymbol{X}) \tag{1}$$

gdzie: $\boldsymbol{X} = [X_1, X_2, \ldots X_m]^T$ – menzurand wejściowy jako wektor losowy o m elementach X_i; $\boldsymbol{Y} = [Y_1, Y_2, \ldots Y_n]^T$ o n elementach – menzurand wyjściowy jako wektor o n elementach Y_j; $\boldsymbol{F}(\cdot)$ – operator, ogólnie nieliniowy wiążący oba wektory.

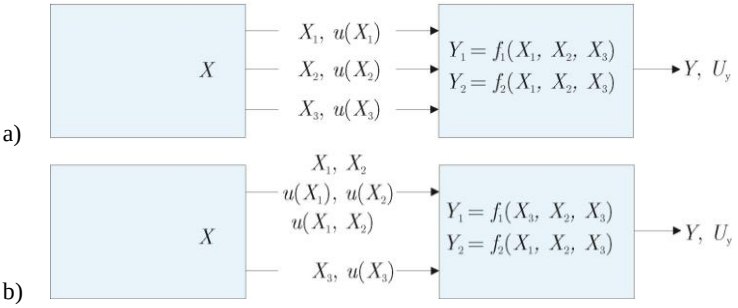
Dla liniowych zależności opisujących pomiar, operator $\boldsymbol{F}$ ze wzoru (1) staje się macierzą $\boldsymbol{S}$ o wymiarach n x m, tj.

$$\boldsymbol{Y} = \boldsymbol{S} \cdot \boldsymbol{X} \tag{1a}$$

W tym przypadku liczba wielkości wyjściowych n nie może przekroczyć *liczby m* wielkości wejściowych, gdyż układ niezależnych równań liniowych ma spełniać warunek $n \leq m$. Z pomiarów m wielkości wejściowych X_i można wtedy otrzymać co najwyżej $n = m$ wielkości wyjściowych Y_j.

Przy nieliniowym operatorze $\boldsymbol{F}$ liczba wielkości wyjściowych n nie jest ograniczona przez m i równa się liczbie niezależnych funkcji F_j wiążących składowe wektorów $\boldsymbol{X}$ i $\boldsymbol{Y}$. Może więc być nawet $n > m$.

Na rys. 1a i 1b podano dwa przykłady wyznaczania danych dwuwymiarowego ($n = 2$) menzurandu $\boldsymbol{Y}$ z przetwarzania danych pomiarowych menzurandu $\boldsymbol{X}$ o $m = 3$ składowych.



Rys. 1. Przykłady pośredniego wyznaczania danych dwu wielkości $\boldsymbol{Y} = [Y_1, Y_2]^T$ z danych trzech mierzonych wielkości wejściowych $\boldsymbol{X} = [X_1, X_2, X_3]^T$: a) nieskorelowanych, b) skorelowanych X_1, X_2

Fig. 1. Examples of indirect evaluation of measurement data of two jointed output variables $\boldsymbol{Y} = [Y_1, Y_2]^T$ from measurements of three input variables $\boldsymbol{X} = [X_1, X_2, X_3]^T$: a) no correlated, b) correlated X_1, X_2

W przypadku b) wielkości wejściowe X_1, X_2 są ze sobą skorelowane. Opisuje się to ich kowariancją σ_{12}, lub niepewnością $u(X_1, X_2)$. Skorelowanie wielkości wyjściowych zależy od skorelowania danych na wejściu i funkcji przetwarzania. Może się ono pojawić dla nieskorelowanych wielkości wejściowych.

W jednoczesnych pomiarach wielu wielkości poszczególne m-wymiarowe obserwacje są realizacjami wektora losowego. Gęstość prawdopodobieństwa ich występowania podlega rozkładowi wielowymiarowemu w przestrzeni współrzędnych X_i wektora $\boldsymbol{X}$. Kształt rozkładu może być różny, a jego geometria zależy nie tylko od rozkładów poszczególnych składowych wektora, ale i od ich statystycznego powiązania. Przy nieliniowym przetwarzaniu danych, rozkład i jego przekroje na wejściu i wyjściu mają inne kształty [1a]. Dla zadanego prawdopodobieństwa rozrzuty końca wektora losowego $\boldsymbol{X}$ o m elementach X_i występują w pewnym ograniczonym obszarze przestrzeni m-wymiarowej. Jest to np. hiperprostopadłościan o m bokach równoległych do osi współrzędnych i równych odchyleniom standardowym σ_i jego składowych X_i [5]. Otrzymuje się go bezpośrednio z jednowymiarowych przedziałów rozrzutu składowych X_i, gdy są one statystycznie niezależnie, np. gdy każdy mierzy się osobno i modeluje rozkładem równomiernym. Koniec średniego wektora $\overline{\mathbf{x}}$ znajduje się w centrum tej bryły.

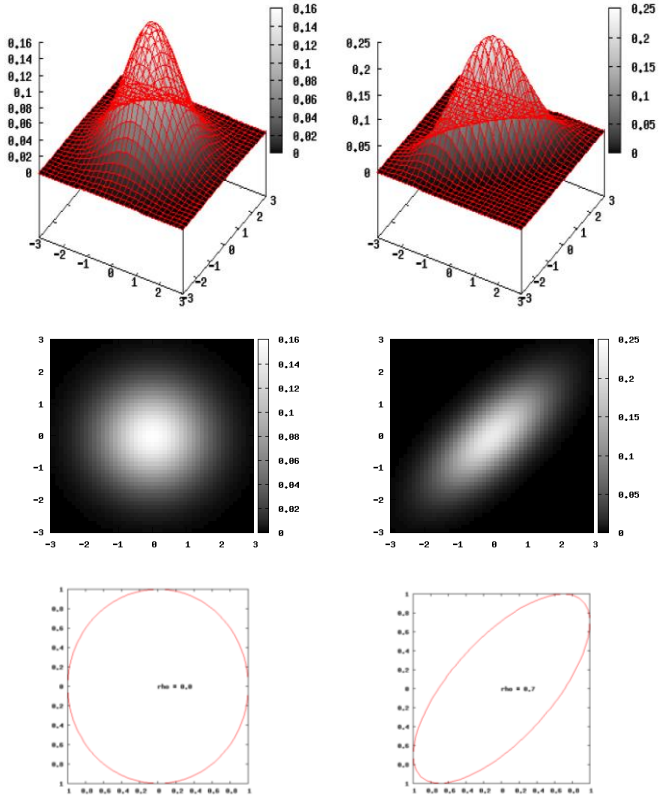
Jeśli rozrzuty każdej z m wielkości składowych X_i multimenzurandu są opisane rozkładem normalnym i mierzy się je równocześnie, to koniec wektora $\boldsymbol{X}$ podlega m-wymiarowemu rozkładowi normalnemu. W wielu przypadkach takim rozkładem z dopuszczalnym przybliżeniem modeluje się rzeczywiste rozrzuty obserwacji wielowymiarowych. Jego opis jest bardziej złożony niż pojedynczej zmiennej losowej – obok wartości średnich współrzędnych $\overline{x_i}$ końca wektora średniego $\overline{X}$ i ich odchyleń standardowych σ_i, występują też współczynniki korelacji wzajemnej ρ_{ik}. Podobnie jest opisany wektor wyjściowy $\boldsymbol{Y}$. Wyznaczenie skorelowania jest konieczne, gdy wszystkie, lub kilka składowych wektora $\boldsymbol{Y}$ są wspólnie przetwarzane numerycznie lub równocześnie uczestniczą w innych pomiarach.

Dla składowych menzurandu w postaci wektora, podobnie jak dla pojedynczej wielkości mierzonej, istnieją granice obszaru rozrzutu danych o zadanym prawdopodobieństwie, czyli o danym poziomie ufności. Takim obszarem dla rozkładu m-normalnego jest m-hiperelipsoida wpisana w m-hiperprostopadłościan. Bryły te dobrze opisują analitycznie dodatnio określone formy kwadratowe [5, 12–15]. Dla m-wymiarowego rozkładu normalnego macierz kowariancji o wymiarach $m \times m$ opisuje m-wymiarową elipsoidę. Kąty nachylenia jej osi i położenia punktów styczności zależą od współczynników korelacji pomiędzy składowymi wektora $\boldsymbol{X}$ [6]. Jeśli dane pomiarowe tych składowych nie są skorelowane, to osie hiperelipsoidy są równoległe do osi współrzędnych wektora.

Dla wektora $\boldsymbol{X}$ o liczbie elementów $m = 2$, dwa przypadki dwuwymiarowego rozkładu normalnego przedstawiono na rys. 2. W tym przypadku przekrojem dla prawdopodobieństwa $p(x,y) = 0{,}5$ jest elipsa wpisana w prostokąt o bokach $\pm\sigma_1$, $\pm\sigma_2$ i nachyleniu osi głównej zależnej od współczynnika korelacji ρ_{xy}. Wzory dla tego rozkładu podano w [5, 12–14]. Dwuwymiarowe prawdopodobieństwo $p(x,y)$ wynosi

$$p(x,y)=\frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho_{xy}^2}}\exp\left(-\frac{1}{2(1-\rho_{xy}^2)}\left(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}-\frac{2\rho_{xy}xy}{\sigma_x\sigma_y}\right)\right) \qquad (2)$$

gdzie: σ_x^2, σ_y^2 – wariancje, $\rho_{xy}=E(x-\bar{x})(y-\bar{y})$ współczynnik korelacji, $\bar{x}$, $\bar{y}$ – wartości średnie.



Rys. 2. Dwuwymiarowy normalny rozkład gęstości prawdopodobieństwa $p(x, y)$ dla dwu współczynników korelacji: $\rho_{xy} = 0$ oraz $\rho_{xy} = 0{,}7$. Poniżej: przykłady poziomicy dla $p(x, y) = const$ [14]

Fig. 2. Two-dimensional normal PDF $p(x, y)$ of correlation coefficients: $\rho_{xy}= 0$ and $\rho_{xy}= 0.7$. Below: border line of data dispersion space for $p(x, y) = const$ [15]

Równanie linii o stałej gęstości prawdopodobieństwa $p(x,y) = const$ wyznacza się z (2) przez przyrównanie wykładnika do wartości stałej. Dla $p = 0{,}5$ mamy

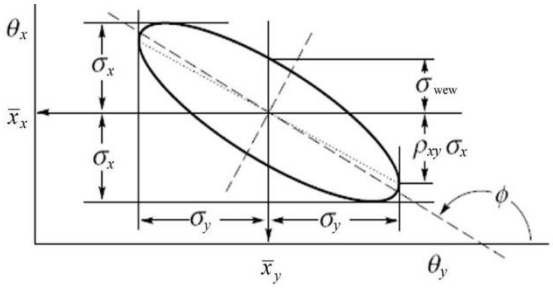
$$\frac{(x-\bar{x})^2}{\sigma_x^2}-2\rho_{xy}\frac{(x-\bar{x})}{\sigma_x}\frac{(y-\bar{y})}{\sigma_y}+\frac{(y-\bar{y})^2}{\sigma_y^2}=1-\rho_{xy}^2 \qquad (3)$$

Jest to równanie elipsy o środku położonym w końcu wektora średniego $[\bar{x}, \bar{y}]^{\mathrm{T}}$ i wpisanej w prostokąt $\pm\sigma_x$, $\pm\sigma_y$. Wartości średnie $\bar{x}$, $\bar{y}$ oraz obie wariancje σ_x^2, σ_y^2 i współczynnik korelacji ρ_{xy} oblicza się z wyników pomiarów w znany sposób.

Postacie rozkładu (2) dla dwu wartości współczynnika korelacji ρ_{xy} i przebiegi linii o stałej gęstości prawdopodobieństwa $p(x, y) = const$ podaje rys. 2. Przy $\sigma_x \neq \sigma_y$ i braku skorelowania ($\rho_{xy} = 0$) – lewa figura rys. 2, osie elipsy wpisanej w prostokąt są równoległe do jego boków $2\sigma_x$, $2\sigma_y$. Dla współczynnika korelacji $0 < |\rho_{xy}| \leq 1$, dłuższa oś elipsy jest nachylona do osi odciętych x pod kątem

$$\phi = \frac{1}{2} arc\ tg\ \ 2\rho_{xy}\frac{\sigma_x\sigma_y}{\sigma_x^2 - \sigma_y^2}\ .$$

Taką elipsę dla $\rho_{xy} = +0{,}7$ przedstawia dolna prawa figura z rys. 2 [12], a dla ujemnego współczynnika ρ_{xy}, otrzymywanego gdy $\sigma_y^2 > \sigma_x^2$ – rys. 3.



Rys. 3. Zależności dla elipsy stałej gęstości $p(x, y) = const$ iwspółczynnika korelacji $\rho_{xy} < 0$
Fig. 3. Formulas for ellipse of $p(x, y) = const$ and of the of correlation coefficient $\rho_{xy} < 0$

Punkty styczności elipsy z bokami prostokąta $\pm\sigma_y$, $\pm\sigma_x$ są odległe od osi układu x, y w środku elipsy odpowiednio o $x_0 = \pm\rho_{xy}\sigma_x$ i o $y_0 = \pm\rho_{xy}\sigma_y$. Przy pełnej korelacji $\rho_{xy} = \pm 1$ elipsa degeneruje się do jednej z przekątnych prostokąta.

Wynik pomiarów wektora wejściowego $\boldsymbol{X}$ wyznacza się w postaci dwuelementowej – jako wektor wartości średnich $\bar{\boldsymbol{x}}$ i macierz kowariancji $\boldsymbol{c}$, lub w postaci trójelementowej, gdy wraz z $\bar{\boldsymbol{x}}$ podawana jest macierz diagonalną odchyleń standardowych $\boldsymbol{\sigma}$ i macierz korelacji $\boldsymbol{r}$ w skrócie korelatorem. Macierz $\boldsymbol{c}$ jest powiązana z macierzami $\boldsymbol{r}$ i $\boldsymbol{\sigma}$ następującą zależnością [6, 12–15]:

$$\boldsymbol{c} = \boldsymbol{\sigma}\cdot\boldsymbol{r}\cdot\boldsymbol{\sigma}^{\mathrm{T}} \tag{4}$$

Elipsoidalny obszar rozrzutu danych pomiarowych występuje dla dodatnio określonej macierzy kowariancji $\boldsymbol{c}$ (i $\boldsymbol{r}$), tj. gdy jej wartości własne λ_i, czyli jednokrotne pierwiastki równania charakterystycznego, są dodatnie

$$\det[\boldsymbol{c} - \lambda\boldsymbol{1}] = 0 \tag{5}$$

Forma przedstawiania parametrów wyznaczanego pośrednio wektora wyjściowego $\boldsymbol{Y}$ jest identyczna jak wektora $\boldsymbol{X}$. Wartości średnie współrzędnych wektora $\mathbf{Y}$ wyznacza się z układu równań (1). Dla niepewności stosuje się liniową propagację wariancji, prawdziwą przy zależnościach liniowych (2) oraz jako przybliżenie dla małych przyrostów współrzędnych zależności nieliniowych, które występują przy wyznaczaniu niepewności. Macierze kowariancji $\boldsymbol{c_Y}$ i $\boldsymbol{c_X}$ wektorów wejściowego i wyjściowego są wówczas powiązane następującym równaniem liniowym:

$$\boldsymbol{c_Y} = \boldsymbol{S} \cdot \boldsymbol{c_X} \cdot \boldsymbol{S}^{\mathrm{T}} \tag{6}$$

gdzie: $\boldsymbol{S}$ – macierz współczynników wrażliwości

$$\boldsymbol{S} \equiv \frac{\partial(\boldsymbol{Y})}{\partial(\boldsymbol{X})} \equiv \begin{bmatrix} \frac{\partial y_1}{\partial x_1}, & \cdots & \frac{\partial y_1}{\partial x_m} \\ \cdots & \cdots & \cdots \\ \frac{\partial y_m}{\partial x_1}, & \cdots & \frac{\partial y_m}{\partial x_m} \end{bmatrix}$$

Z macierzy kowariancji $\boldsymbol{c_Y}$ można otrzymać macierz niepewności oznaczoną na rys. 1 jako $\boldsymbol{U_Y}$.

Macierz korelacji wielkości wyjściowych $\boldsymbol{r_Y}$ otrzymuje się bezpośrednio z macierzy kowariancji $\boldsymbol{c_Y}$, lub przekształcając korelator wejściowy $\boldsymbol{r_X}$ wg (4).

Jeśli któraś z wielkości ma rozkład niegaussowski, to obszar rozrzutu jest o innym kształcie i do jego opisu trzeba stosować inne funkcje, w tym kopuły (ang. *copulas*) [7]). W wielu przypadkach można w przybliżeniu zastąpić ten obszar hiperelipsoidą.

Wyniki pomiarów dwuwymiarowych o rozkładzie (2) można przedstawić jedną z dwu równoważnych struktur macierzowych: wektor średni i macierz kowariancji $\boldsymbol{c_Y} \equiv \boldsymbol{c}_{xy}$ lub wektor średni z niepewnościami i macierz korelacji $\boldsymbol{r}_{xy}$.

$$\left(\begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix}, \begin{bmatrix} \sigma_x^2 & \sigma_x\sigma_y \cdot \rho_{xy} \\ \sigma_x\sigma_y \cdot \rho_{xy} & \sigma_y^2 \end{bmatrix} \right) \Rightarrow \left(\begin{bmatrix} \bar{x} \pm \sigma_x \\ \bar{y} \pm \sigma_y \end{bmatrix}, \begin{bmatrix} 1 & \rho_{xy} \\ \rho_{xy} & 1 \end{bmatrix} \right) \tag{7}$$

Równanie charakterystyczne macierzy $\boldsymbol{c}_{xy}$ ma postać

$$\det|\boldsymbol{c}_x - \lambda \boldsymbol{1}| = \lambda^2 - (\sigma_x^2 + \sigma_y^2)\lambda + \sigma_x^2\sigma_y^2(1-\rho_{xy}^2) = 0 \tag{6a}$$

a jej wartości własne λ_1, λ_2, czyli pierwiastki równania (6a) wynoszą

$$\lambda_{1/2} = \tfrac{1}{2}(\sigma_x^2 + \sigma_y^2) \pm \tfrac{1}{2}\sqrt{(\sigma_x^2 + \sigma_y^2)^2 - 4\sigma_x^2\sigma_y^2(1-\rho_{xy}^2)} \tag{6b}$$

W przypadkach krańcowych: dla $\rho_{xy} = 0$ elipsa jest położona poziomo, a pierwiastki (6a) $\lambda_1 = \sigma_x^2$, $\lambda_2 = \sigma_y^2$ są kwadratami obu średnic; zaś dla $|\rho_{xy}| = 1$ elipsa staje się jedną z przekątnych prostokąta, gdyż $\lambda_2 = 0$, a $\lambda_1 = \sigma_x^2 + \sigma_y^2$.

Dla $0 \le \rho_{xy} \le 1$ oba pierwiastki spełniają warunek $\lambda_{1/2} \ge 0$, a długości średnic pochylonej elipsy wynoszą [4]:

$$a^2 = -\frac{1-\rho_{xy}^2}{\lambda_2}, \quad b^2 = -\frac{1-\rho_{xy}^2}{\lambda_1}$$

Dotyczy to też pierwiastków $\lambda'_{1/2} = 1 \pm \rho_{xy}$ równania $(1 - \lambda')^2 - \rho_{xy}^2 = 0$ dla macierzy korelacji $\boldsymbol{r}_{xy}$. Jej elipsa o półosiach $(1 \pm \rho_{xy})^{-1}$ wpisana jest w kwadrat ± 1 oraz styczna w $\pm\rho_{xy}$. Obie macierze $\boldsymbol{c}_{xy}$, $\boldsymbol{r}_{xy}$ są półokreślone dodatnio.

Transformację liniową dwu mierzonych bezpośrednio parametrów zilustruje analitycznie i liczbowo przykład 1 [6]. Zawiera on obliczenia dla liniowego przetwarzania danych multimenzurandu.

Przykład 1 [5]

Menzurand wyjściowy o dwu parametrach x, y wyznacza się z wejściowych danych pomiarowych ζ, η o dwuparametrowym rozkładzie normalnym jako ich sumę i różnicę. Ta liniowa operacja ma postać wektorową

$$\begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix} = \boldsymbol{S} \begin{bmatrix} \bar{\zeta} \\ \bar{\eta} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \bar{\zeta} \\ \bar{\eta} \end{bmatrix} = \begin{bmatrix} \bar{\zeta} + \bar{\eta} \\ \bar{\zeta} - \bar{\eta} \end{bmatrix}$$

Standardowe niepewności dla każdej ze składowych wektora wyjściowego $[x, y]^{\mathrm{T}}$ można otrzymać znaną metodą liniowej propagacji niepewności. Są one ze sobą skorelowane. Ich współczynnik korelacji ρ_{xy} można obliczyć metodą wektorową. Z równań (4)–(6) wynika

$$\boldsymbol{c}_{xy} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \sigma_\zeta^2 & \sigma_\zeta \sigma_\eta \rho_{\zeta\eta} \\ \sigma_\zeta \sigma_\eta \rho_{\zeta\eta} & \sigma_\eta^2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} \sigma_x^2 & \sigma_\zeta^2 - \sigma_\eta^2 \\ \sigma_\zeta^2 - \sigma_\eta^2 & \sigma_y^2 \end{bmatrix}$$

gdzie: $\sigma_x = \sqrt{\sigma_\zeta^2 + \sigma_\eta^2 + 2\sigma_\zeta \sigma_\eta \rho_{\zeta\eta}}$, $\sigma_y = \sqrt{\sigma_\zeta^2 + \sigma_\eta^2 - 2\sigma_\zeta \sigma_\eta \rho_{\zeta\eta}}$

Elementy niediagonalne macierzy kowariancji $\boldsymbol{c}_{xy}$ nie zależą od $\rho_{\zeta,\eta}$. Z $\boldsymbol{c}_{xy}$ wyznacza się macierz korelacji

$$\boldsymbol{r}_{xy} = \begin{bmatrix} 1 & \rho_{xy} \\ \rho_{xy} & 1 \end{bmatrix}$$

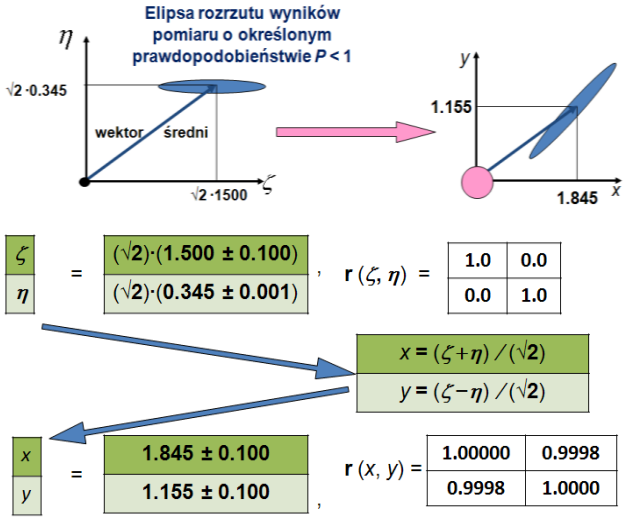
gdzie: $\rho_{xy} = \dfrac{\sigma_\zeta^2 - \sigma_\eta^2}{\sigma_x \sigma_y}$ – współczynnik korelacji wielkości wyjściowych x, y zależny od współczynnika korelacji $\rho_{\zeta,\eta}$ występującego we wzorach dla σ_x i σ_y.

Przy nieskorelowanych składowych wektora wejściowego $[\zeta, \eta]^{\mathrm{T}}$, tj. dla $\rho_{\zeta,\eta} = 0$, macierz $\boldsymbol{c}_{xy}$ upraszcza się

$$\boldsymbol{c}_{xy} = \begin{bmatrix} \sigma_\varsigma^2 + \sigma_\eta^2 & \sigma_\varsigma^2 - \sigma_\eta^2 \\ \sigma_\varsigma^2 - \sigma_\eta^2 & \sigma_\varsigma^2 + \sigma_\eta^2 \end{bmatrix}$$

i wówczas: $\sigma_x = \sigma_y = \sqrt{\sigma_\varsigma^2 + \sigma_\eta^2}$ oraz $\rho_{xy} = \dfrac{\sigma_\varsigma^2 - \sigma_\eta^2}{\sigma_\varsigma^2 + \sigma_\eta^2}$.

Na rys. 4 (u góry) przedstawiono schematycznie operację liniowego przekształcenia dwuelementowego wektora wejściowego $[\zeta, \eta]^T$ o nieskorelowanych składowych w wektor $[x, y]^T$, a poniżej dane liczbowe obu wektorów.



Rys. 4. Przykład liniowego przekształcenia wektora losowego $[\zeta, \eta]^T$ w wektor $[x, y]^T$ [6]
Fig. 4. Example of linear transformation of 2D random vector $[\zeta, \eta]^T$ to vector $[x, y]^T$ [6]

Z podanych na rys. 4 wartości elementów macierzy korelacji $\boldsymbol{r}(x, y)$ wynika, że niepewności składowych wektora wyjściowego $[x, y]^T$ są silnie ze sobą skorelowane. Jednakże zaokrąglenie wartości elementów niediagonalnych do 3 cyfr znaczących, jak to zaleca Przewodnik GUM w punkcie 7.2.6, dałoby $\rho_{xy} = 1$. Spowodowałoby to degenerację elipsy do przekątnej prostokąta, tj. do całkowicie deterministycznej współzależności niepewności obu składowych.

Przy przetwarzaniu wektora losowego $\boldsymbol{X}$ wg nieliniowego operatora $\boldsymbol{F}$ zgodnie z (1) kształt geometryczny obszaru rozrzutu końca wektora $\boldsymbol{Y}$ zmieni się zależnie od $\boldsymbol{F}$ [1a]. Dlatego też należy śledzić zmiany granic tego obszaru oraz położenie wierzchołka wektora średniego $\overline{\boldsymbol{Y}}$.

Przy funkcjach przetwarzania o małych nieliniowościach, lub dla małych wartości niepewności składowych wektora $\boldsymbol{X}$ zwykle można w przybliżeniu też stosować liniową propagację wariancji wg wzoru (6) [7, 13].

Przykład przetwarzania nieliniowego omówiono w [1, 4, 8]. Wykorzystano dane z Przykładu H.2 w GUM [1], powtórzonego też w Suplemencie 2, w którym z trzech próbek o 5-ciu równoczesnych pomiarach napięcia, prądu i przesunięcia fazowego w dwójniku impedancyjnym wyznacza się wartości i niepewności modułu jego impedancji $Z = \sqrt{R^2 + X^2}$ i jej składowe prostokątne *R, X*.

Ujęcie wektorowe można też stosować w analizie i prognozowaniu traktowanych losowo błędów granicznych pomiarów wieloparametrowych. Jeśli nie daje się a priori oszacować wartości ich współczynników korelacji r_{ij}, to dla dużych ich wartości przyjmuje się je jako równe 1 z odpowiednim znakiem, lub też najniekorzystniejszy ich wariant. Dla niewielkiej korelacji zakłada się współczynniki $r_{ij} = 0$.

O zniekształcaniu danych multimenzurandu przy przetwarzaniu

Przy wyznaczaniu z surowych danych pomiarowych estymatorów parametrów menzurandu stosuje się zaokrąglanie otrzymywanych wyrażeń liczbowych. Zalecenia dla tej prostej operacji nieliniowej opisano w Przewodniku GUM [1], ale tylko dla pojedynczej wielkości mierzonej. W Suplemencie 2 [1a] oraz innych przepisach metrologicznych i przewodnikach nie ma ścisłych zaleceń jak poprawnie zaokrąglać dane dla pomiarów wielu skojarzonych parametrów. Niezależne zaokrąglanie wartości i niepewności składowych oraz elementów macierzy kowariancji (lub korelacji) wektora losowego może zniekształcić wyniki zarówno wykonanych prawidłowo pomiarów multimenzurandu wejściowego ***X***, jak i wyznaczanych pośrednio z nich składowych multimenzurandu wyjściowego ***Y,*** w przypadku ogólnym powiązanego z ***X*** zależnością nieliniową (1).

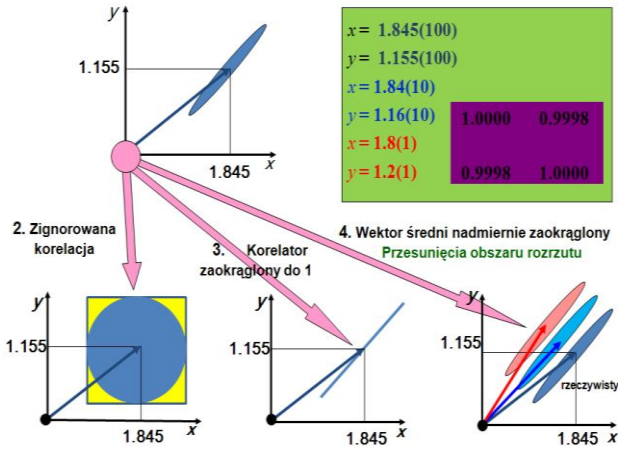
V. Ezhela wyznaje pogląd, podzielany też przez niektórych badaczy w naukach podstawowych, że surowe lub prawidłowo zaokrąglone dane wektora wejściowego ***X*** należy traktować jako absolutnie dokładne przy wszelkich dalszych przekształceniach tych danych. Ponadto uważa, że konsekwentnie należy:

- utrzymywać dodatnią półokreśloność macierzy kowariancji, która daje *n*-elipsoidalny obszar rozrzutu pomiarów o określonym prawdopodobieństwie dla zależności liniowych (jako zastępczy dla zależności nieliniowych);
- utrzymywać wierzchołek zaokrąglanego wektora wewnątrz tego obszaru.

W czołowych czasopismach naukowych Ezhela znalazł wiele przykładów oszacowań wyników pomiarów i ich procedur przetwarzania, które nie spełniają takich wymagań tzw. „dobrej praktyki” (uwagi o symbolu ■ na końcu jego pracy [4]). Przyczyny takiego stanu są następujące:

(i) Podaje się wartości średnie i odchylenia standardowe składowych multimenzurandu, a pomija macierz korelacji.
(ii) Estymatory są „nadmiernie” zaokrąglane, tj. tak, że macierz kowariancji nie jest dodatnio określona.
(iii) Końcowy wektor średni po nadmiernym zaokrągleniu wychodzi na wiele odchyleń standardowych poza granice obszaru rozrzutu bądź surowych wyników obserwacji pomiarowych, bądź ich wartości przetworzonych w pomiarach pośrednich wg operatora ***F*** ze wzoru (1).

Na rysunku 5 pokazano konsekwencje takich zniekształceń dla poprawnych początkowych danych wektora $[x, y]^T$ z rys. 4, zaś w Przykładzie 2 oszacowano liczbowo warianty a i b dla wartości danych z Przykładu 1 i przypadku 4 z rys.5.



Rys. 5. Przypadki nieprawidłowo zaokrąglanych skorelowanych dwuwymiarowych danych [6]
Fig. 5. Cases of presentation of the improper rounded correlated 2D data [6]

Przykład 2

Wyniki pomiarów pośrednich menzurandu $\boldsymbol{Y}$ wg rys. 5:

$$\boldsymbol{Y} = [1{,}845(100);\ 1{,}155(100)]$$

Przypadek 4 a. Zaokrąglenie pojedynczej ostatniej cyfry:

$$\boldsymbol{Y}_1 = [1{,}84(10);\ 1{,}16(10)]$$

Wektor różnicy: $\Delta\boldsymbol{Y}_1 = \boldsymbol{Y}_1 - \boldsymbol{Y} = [-0{,}005;\ 0{,}005]$

Przypadek 4 b. Zaokrąglenie dwu ostatnich cyfr:

$$\boldsymbol{Y}_2 = [1{,}8(10);\ 1{,}16(10)]$$

Wektor różnicy: $\Delta\mathbf{Y}_2 = \boldsymbol{Y}_2 - \boldsymbol{Y} = [-0{,}045;\ 0{,}045]$

W analizie macierzowej wektorów losowych położenie końca wektora zaokrąglonego względem centrum elipsoidalnego obszaru rozproszenia wyznacza się za pomocą miary odległości wg Mahalanobisa [10]. Długości odchyleń składowych odnosi się do ich odchyleń średnich standardowych wg formy kwadratowej

$$\chi^2 = \Delta\boldsymbol{Y}\,\frac{1}{\sigma_x\sigma_y}\cdot\boldsymbol{r}(x,y)^{-1}\Delta\boldsymbol{Y}^T \qquad (7)$$

Punkty elipsy stycznej wewnętrznie do prostokąta o bokach $\pm(\sigma_x, \sigma_y)$ mają $\chi^2 = 1$. Dla zaokrągleń 4a i 4b otrzyma się:

$$\chi_1^2 = [-0{,}005;\ 0{,}005] \cdot \frac{1}{0{,}01} \cdot \frac{1}{1-0{,}9998^2} \cdot \begin{bmatrix} 1 & -0{,}9998 \\ -0{,}9998 & 1 \end{bmatrix} \cdot \begin{bmatrix} -0{,}005 \\ 0{,}005 \end{bmatrix} = 25 >> 1$$

$$\chi_2^2 = [-0{,}045; 0{,}045] \frac{1}{0{,}01} \cdot \frac{1}{1-0{,}9998^2} \begin{bmatrix} 1 & -0{,}9998 \\ -0{,}9998 & 1 \end{bmatrix} \cdot \begin{bmatrix} -0{,}045 \\ 0{,}045 \end{bmatrix} = 2025 >> 1$$

Końce obu zaokrąglonych wektorów średnich wychodzą tu znacznie poza obszar rozproszenia danych pierwotnych, tj. niezaokrąglonych, na wiele odchyleń standardowych. Często w publikacjach obszar ten podaje się wokół końca zaokrąglonego wektora (prawy dolny róg rys. 5) i jego wymiary zaokrągla się do góry by objął obszar danych surowych.

Przykład 2 ilustruje, że niezależne zaokrąglanie składowych wektora średniego i ich niepewności, takie jak wg zaleceń GUM dla pojedynczej wielkości mierzonej, może dać niepoprawny rezultat przy przetwarzaniu skojarzonych statystycznie danych wielowymiarowych. Trudno jest to wykryć i skorygować bez dostępu do danych oryginalnych.

Rozważania te posłużyły V. Ezheli do uzasadnienia konieczności standaryzacji struktury oraz procedury zaokrąglania wartości zmierzonych i przetwarzanych danych w pomiarach wieloparametrowych. Proponuje on by zestaw danych wielowymiarowego menzurandu obejmował następujące parametry:

- wektor średni, tj. wartości średnie jego składowych,
- wektor odchyleń standardowych tych wartości,
- współczynniki korelacji,
- obliczoną minimalną wartość własną dodatnio określonej macierzy kowariancji lub korelatora,
- informację o zastosowanej cyfrowej precyzji obliczeń.

Przy tak określanych danych Ezhela wyznaczył powiązane ze sobą progi dopuszczalnego zaokrąglania przy dokładnym przetwarzaniu parametrów statystycznych wektora losowego [4, 5, 7]. Dla niepewności standardowych i współczynników korelacji otrzymał wielocyfrowe wartości liczbowe znacznie przekraczające liczby znaków stosowane w pomiarach i zalecane przez przepisy metrologiczne.

Zdaniem autora, takie wygórowane wymagania wynikają z przyjęcia nieprawidłowego założenia, że niepewności parametrów statystycznych wektora wejściowego są pomijalnie małe, czyli jego dane traktuje się jako zmierzone absolutnie dokładnie. Nie zachodzi to dla rzeczywistych danych pomiarowych. Progi podane przez Ezhelę dotyczą więc dokładności samego przetwarzania wektora losowego.

Przy szacowaniu wyników pomiarów wieloparametrowych należy uwzględnić też wpływ liczebności próbek na dokładność estymowanych parametrów i na zaokrąglanie ich wartości liczbowych. Omówi się to szczegółowo poniżej.

O zaokrąglaniu danych multimenzurandu przy ograniczonej liczebności próbek

W praktyce pomiarowej zwykle nie można pominąć niepewności składowych multimenzurandu wejściowego $\boldsymbol{X}$ zarówno o charakterze losowym, opisywanych staty-

styczną metodą typu A, jak i niepewności pochodzenia instrumentalnego szacowanych metodą typu B. W krańcowych przypadkach dominuje jeden ich rodzaj.

Przyczyny losowych rozrzutów obserwacji pomiarowych są dwojakie, tj.:

- wielkości mierzone w obiekcie badanym są już losowe,
- wartości tych wielkości są stałe, ale wskutek dużych zakłóceń w obiekcie i torach pomiarowych pojawiają się rozrzuty wartości kolejnych obserwacji pomiarowych.

Udziały obu powyższych źródeł losowości są zwykle różne.

W pierwszym z przypadków zadanie pomiarowe może polegać na dokładnym zbadaniu nie tylko samych estymat wartości średnich składowych menzurandu $\boldsymbol{X}$, ale i autokorelacji przy wyznaczaniu ich odchyleń standardowych (rozdz. 3) oraz współczynników korelacji i innych parametrów statystycznych opisujących model rozkładu i estymatorów niepewności tych wszystkich parametrów. Należy wtedy pobrać z obiektu badanego możliwie jak największą liczbę obserwacji, aby parametry statystyczne były bliskie parametrom populacji badanego wektora losowego $\boldsymbol{X}$**.** Tylko dla populacji o skończonej liczbie elementów i znanych wartościach, gdy pobierze się je wszystkie, to takie dane wejściowe można traktować jako pozbawione błędów losowych.

Przy pośrednim wyznaczaniu parametrów wektora wyjściowego $\boldsymbol{Y}$, utratę informacji zminimalizuje się, gdy po cyfrowym przetwarzaniu i zaokrągleniu koniec tego wektora pozostanie wewnątrz obszaru opisującego rozrzut przetworzonych surowych danych wejściowych. Powinno się więc śledzić, czy przy przetwarzaniu można przyjąć, że zachował się elipsoidalny obszar rozrzutu zmiany granic obszaru ich rozproszenia oraz położenie wierzchołka wektora średniego względem tego obszaru, np. wykorzystując miarę odległości Mahalanobisa wg (7). Właśnie dla spełnienia tych warunków V. Ezhela wyznaczył powiązane ze sobą wzory dla progów zaokrąglania wartości, niepewności i współczynników korelacji menzurandu wyjściowego $\boldsymbol{Y}$, tj. minimalne dopuszczalne liczby ich cyfr [7, 10]. Z wzorów tych wynikła konieczność wyznaczania bardzo dużej liczby znaków, nie do zaakceptowania przy szacowaniu niepewności dla większości pomiarów multimenzurandu. Progi te mogą posłużyć jedynie do oceny dokładności samego procesu przetwarzania wektora losowego, gdyż nie uwzględniają niepewności danych wejściowych.

Nawet gdy pomijalne są niepewności typu B, wynikające z nieusuniętych błędów instrumentalnych i błędów metody pomiarowej dla składowych wektora wejściowego $\boldsymbol{X}$, to wymagana dokładność cyfrowego przetwarzania danych istotnie zależy od liczb obserwacji N w próbkach pomiarowych. Liczby ta determinuje wielkość wyznaczanych statystyczne niepewności typu A nie tylko dla wartości składowych wektorów $\boldsymbol{X}$ i $\boldsymbol{Y}$, ale i dla wszystkich innych parametrów statystycznych wielowymiarowego rozkładu każdego z nich.

Dla jednowymiarowego rozkładu normalnego względna niepewność standardowa odchylenia standardowego jest równa w przybliżeniu od $[2(N - 1)]^{-1/2}$. W tabeli E.1 Przewodnika GUM [1] podano, że dla próbki liczącej np. tylko 5 pomiarów wynosi ona aż 36%. Dla 50 pomiarów niepewność ta spada, ale zaledwie do 10%. Niepewności statystyczne współczynników korelacji będą jeszcze większe, gdyż zależą od wartości i niepewności dwu składowych wektora. W takim przypadku precyzja przetwarzania danych wejściowych nie musi więc być nadmierna. Dla próbek o ograniczonej liczebności powinno wystarczyć by była ona o rząd większa niż naj-

mniejsza (lub następna) cyfra pewna dla niepewności. Niepewność, jako wielkość drugiego rzędu wystarczy zaokrąglać do dwu, lub przy małej liczbie pomiarów w próbce, nawet tylko do jednej cyfry, a estymator jej wartości – na tym samym miejscu dziesiętnym, czyli podobnie, jak według zaleceń Przewodnika GUM [1, 1a] dla pomiarów jednej wielkości: niepewność – do co najwyżej 2 cyfr, a tylko współczynniki korelacji bliskie 1 – do 3 cyfr.

Podane przez V. Ezhelę zależności dla progów bezpiecznego zaokrąglania przy precyzyjnym przetwarzaniu cyfrowym danych wektorów losowych [7] nie dotyczą próbek pomiarowych o małej lub średniej liczebności oraz nie obejmują też przypadków, gdy przypuszcza się, że dokładność odchyleń standardowych i współczynników korelacji wektora wejściowego $\boldsymbol{X}$ jest mała ze względu na nieusunięte nieznane błędy instrumentalne i oddziaływania zewnętrzne (niepomijalne niepewności typu B). Dla takich próbek powinno się wyznaczać wartości średnie składowych wektora $\boldsymbol{Y}$ z takimi niepewnościami, jakie wynikają z niepewności danych wejściowych. Wyniki końcowe należy tak zaokrąglać, by nie wprowadzić dużych dodatkowych błędów. Dwie takie propozycje omówiono w [9, 10]. Niezależne zaokrąglanie składowych wektora wg zasad dla skalarnych wielkości może spowodować pojawienie się wyników niespełniających i tych wymagań. Przedstawia to przykład 3.

Przykład 3

W tabeli 1 i na rys. 6 zilustrowano zaokrąglanie danych surowych wyników pomiaru o dwuwymiarowym rozkładzie normalnym.

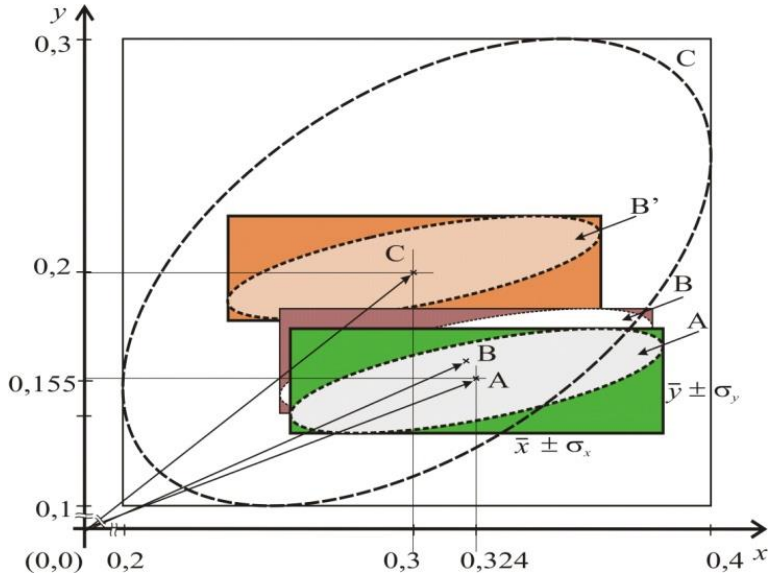
Tab. 1. Zaokrąglenia składowych i odchyleń standardowych składowych dwuwymiarowego wektora losowego [*x, y*]

Tab. 1. Rounding of components and standard deviations of the 2D random vector [*x, y*]

Zaokrąglenia	$\bar{x}_i$	$\sigma(\bar{x}_i)$	$\bar{y}_i$	$\sigma(\bar{y}_i)$	$2\sigma(\bar{x}_i)$	$2\sigma(\bar{y}_i)$
Dane surowe	0,3242	0,0664	0,1555	0,0256	0,1328	0,0512
A. do 3 cyfr	0,324	0,067	0,156	0,026	0,133	0,051
B. do 2 cyfr	0,32	0,07	0,16	0,03	0,14	0,05
C. do 1 cyfry	0,3	< 0,1	0,2	< 0,1	< 0,2	< 0,1

Po jednolitym zaokrągleniu surowych wartości średnich $\bar{x}$, $\bar{y}$ wg zasad GUM jak dla skalara, do trzech, dwu i jednej cyfry liczby najbliższej, otrzymano punkty położenia A, B i C końców wektora średniego oraz elipsy A, B i C o środkach w tych końcach. Na rys. 6 elipsy te wpisane są w prostokąty odpowiednio o bokach $\pm\sigma(\bar{x})$, $\pm\sigma(\bar{y})$ odchyleń standardowych tak samo zaokrąglonych, ale tylko w górę i przy niezmiennej wartości współczynnika korelacji r_{xy}. Elipsa A pokrywa się z elipsą dla surowych, tj. niezaokrąglonych danych. Punkty B i C są przesunięte względem środka elipsy A. Elipsa B' jest przeniesioną do punktu C elipsą B. Wszystkie elipsy znajdują się wewnątrz dużego prostokąta C. Elipsy B i C są większe od A i powinny obejmować ją i elipsę danych surowych, ale je przecinają, gdyż ich środki są przesunięte. Zmiana nachylenia osi elips B i C przez zmianę współczynnika korelacji r_{xy} nie rozwiąże tego problemu.

Opis elipsoidalnego obszaru rozrzutu przyjęty jako niesymetryczny względem końca zaokrąglonego wektora byłby niewygodny w praktyce. Sytuacja poprawia się dla elips odpowiadających rozszerzonym niepewnościom obu składowych (nie są podane na rysunku 2), np. o współczynniku rozszerzenia $k_{px} = k_{py} = 2$, czyli dla elips wpisanych w prostokąty o dwukrotne powiększonych i zaokrąglanych bokach $\pm 2\sigma_x$, $\pm 2\sigma_y$. Z danych tabeli 1 wynika np., że $2\sigma(\bar{x}_B) - (\bar{x} - \bar{x}_B) > 2\sigma(\bar{x}_B)$. Obie takie elipsy B_2, C_2 będą już obejmować podwójnie powiększoną elipsę A_2.



Rys. 6. Elipsy A, B, C rozrzutu wektora $[x, y]^T$ o dwuwymiarowym rozkładzie normalnym uzyskane przy jednolitym (jak wg GUM dla skalarów) zaokrąglaniu składowych i ich niepewności odpowiednio do 3, 2 i 1 cyfry oraz stałym współczynniku korelacji r_{xy} = const [9]

Fig. 6. Scattering ellipses A, B, C of vector $[x, y]^T$ of two variable Normal distribution obtained after recommended by GUM for scalars unified rounding of its components and uncertainty to 3, 2 and 1 digit and constant correlation coefficient r_{xy} = const [9]

Zaokrąglanie współczynników korelacji o wartościach bliskich ±1 wymaga bardzo ostrożnego traktowania.

Z Przykładu 3 wynika, że dla pomiarów multimenzurandu próbkami o ograniczonej liczebności brakuje jeszcze odpowiednich ustaleń o jednolitym zaokrąglaniu wartości, niepewności i współczynników korelacji mierzonych wielkości wektorowych i o precyzji wymaganej przy ich przetwarzaniu. W niepewności wyznaczanej dla tych parametrów powinno się uwzględnić wpływ liczby pomiarów w próbkach i ich niepewności typu B od nieusuwalnych błędów instrumentalnych. Dwie propozycje zaokrąglania dla danych przykładu H.2 z GUM podano w [7].

Wniosek. Struktura danych opisujących pomiary wieloparametrowe powinna koniecznie zawierać informację o liczbie pomiarów w próbce i o niepewności typu B

każdej z mierzonych wielkości. Jest to niezbędne do wyznaczenia niepewności statystycznych (typu A) uczestniczących w ocenie standardowych niepewności składowych multimenzurandu i niepewności współczynników korelacji. Niepewności te służą do ustalania precyzji obliczeń i zaokrąglania wyznaczanych pośrednio parametrów wektora $\boldsymbol{Y}$ [10].

O propagacji niepewności funkcji silnie nieliniowych wg GUM

W Przewodniku GUM [1], do obliczeń niepewności u_c pojedynczej wielkości mierzonej o funkcji przetwarzania $f(\cdot)$ podano wzór dla liniowej propagacji wariancji

$$u_{\mathrm{c}}^2(y) = \sum_{i=1}^{N} \left(\frac{\partial f}{\partial x_i} \right)^2 u_{\mathrm{c}}^2(x_i) \tag{8}$$

Dla funkcji nieliniowych jest to przybliżenie wystarczające dla małych niepewności. Dla dużych nieliniowości w Uwadze do punktu 5.1.2 dodano dodatkowy wyraz, tj:

$$\sum_{i=1}^{N} \sum_{j=1}^{N} \left[\frac{1}{2} \left(\frac{\partial^2 f}{\partial x_i \partial x_j} \right)^2 + \frac{\partial f}{\partial x_i} \cdot \frac{\partial^3 f}{\partial x_i \partial x_j \partial x_j} \right] u^2(x_i)\, u^2(x_j) \tag{9}$$

Wzór ten wielokrotnie powtórzono w wielu dokumentach metrologicznych i podręcznikach z fizyki i metrologii oraz w czołowych czasopismach światowych. Ezhela zauważył jednak [5], że składnik w nawiasie z trzecią pochodną może spowodować, iż z obliczeń otrzyma się ujemną wartość wariancji $u^2(F)$ – patrz Przykład 4, co jest sprzeczne z matematyczną definicją wariancji jako wielkości dodatniej.

Przykład 4 [6]

Należy wyznaczyć niepewność standardową wielomianu $F(x)=1 - x + 2x^2 + 3x^3 + 4x^4$ dla zmiennej losowej x o rozkładzie normalnym z wariancją σ^2 wokół $x = 0$. Dla liniowej propagacji z (8) otrzymuje się:

$$u'^2(F) = \sigma^2$$

Według wzoru (8) rozszerzonego o (9) wariancja $u^2(F)$ wynosi:

$$u^2(F) = \sigma^2\{F'(0)^2 + \sigma^2[(1/2)\, F''(0)^2 + F'(0)\, F'''(0)]\} = = \sigma^2 (1 - 10\, \sigma^2).$$

Drugi składnik w nawiasie powoduje, że przy $\sigma^2 > 0{,}1$ otrzyma się ujemną wartość wariancji $u^2(F)$. Dodatnią wariancję $u^2(F) > 0$ otrzyma się z (8) po uzupełnieniu o (9), ale z pominięciem w (9) drugiego składnika z $F'''(0)$, tj.

$$u^2(F) = \sigma^2\{F'(0)^2 + \sigma^2\, (1/2)\, F''(0)^2\} = \sigma^2\, (1 + 4{,}5\, \sigma^2)$$

Ezhela wykazał więc, że zalecane w Uwadze do punktu 5.1.2 GUM dodanie dodatkowego składnika wg wzoru (9) okazało się w tym przypadku niemożliwe do zastosowania dla zbyt dużych σ^2. Uwaga powinna być więc zmieniona lub uzupełniona ograniczeniami dotyczącymi stosowania tego składnika, wynikającymi z wartości kolejnych członów rozwinięcia w szereg funkcji opisującej pomiar.

Powinno się więc analizować dokładność wyznaczania niepewności przy stosowaniu liniowego przybliżenia wariancji dla funkcji nieliniowych, nawet w pomiarach pojedynczej wielkości.

W pośrednich pomiarach wielowymiarowych o nieliniowych funkcjach przetwarzania można stosować dwie metody wyznaczania niepewności i współczynników korelacji danych wyjściowych:

- wykorzystując liniowe przybliżenie propagacji wariancji wg zależności (6),
- przetwarzając wartości każdej z wielowymiarowych obserwacji wejściowych wg funkcji nieliniowej opisującej pomiary pośrednie i następnie wyznaczając parametry rozkładu dla tak otrzymanego losowego wektora wyjściowego. Przy tej metodzie i niezbyt dużych rozrzutach na wyjściu unika się konieczności wprowadzania poprawek dla niepewności uwzględniających nieliniowości.

W pomiarach wieloparametrowych opisywanych funkcjami nieliniowymi wymiary m i n wektorów wejściowego i wyjściowego oraz rząd T wyrazów rozwinięcia w szereg Taylora są następująco ze sobą powiązane [9]:

$$n \le n_{th} = \frac{(m+T)!}{m! \times T!} - 1 \tag{10}$$

gdzie: n i m – wymiary wektorów wyjściowego i wejściowego, T – rząd szeregu Taylora wielkości wyjściowej.

W opisie i obliczeniach pomiarów wielowymiarowych wykorzystuje się macierze wrażliwości zawierające pierwszych pochodne. Zwykle nie sprawdza się stosowalności aproksymacji liniowej. Może to też stanowić dodatkowe źródło niepoprawności danych końcowych.

Stosowanie nieliniowej propagacji wariancji może być istotne np. przy wyznaczaniu dużych wartości niepewności względnych dla pomiarów parametrów rozkładów losowych wektorów wielkości fizycznych powiązanych nieliniowymi, nawet dość prostymi, np. tylko iloczynowymi zależnościami algebraicznymi (jak w mostkach pomiarowych). Ujednolicone zalecenia, wraz z przykładami, powinna zawierać rozszerzona wersja Suplementu 2 do GUM o wyrażaniu wyników pomiarów wielowymiarowych, lub można by nawet opracować nowy dokument tylko o niepewnościach pomiarów pośrednich opisanych funkcjami nieliniowymi.

O propozycji dualnej formy publikowania danych pomiarowych

W przewodniku GUM [1] (w tym w załączniku H.2) i w jego Suplemencie 2 [1a] oraz w innych poradnikach metrologicznych, np. [2] i w literaturze podstawowej [3, 12–14] nie ma gruntownych zaleceń dla procedur numerycznego przetwarzania i wyrażania danych multimenzurandu oraz przekazywania tych wyników. Analiza zebranych przez V. Ezhelę przykładów w pracy [4] wykazała występowanie wielu nierzetelności w prezentowanych wyników eksperymentalnych prac naukowych nawet w czołowych czasopismach naukowych. Również w korektach danych kilku podstawowych stałych fizycznych FPC (ang. *Fundamental Physical Constancies*) [4, 8, 9, 11] opublikowanych w latach 1998–2010 przez międzynarodową organizację CODATA (ang. *Committee on Data for Science and Technology*) Ezhela wykrył, że ciągle występuje nieprawidłowa macierz korelacji, tj. o ujemnych wartościach

własnych. Wielowymiarowe rozkłady tych danych nie będą wówczas rozkładami normalnymi. Szybka adaptacja do stosowania w praktyce zaleceń Suplementu 2 do GUM [1a] pozwoliłaby zgromadzić doświadczenie z wyznaczania wyników pomiarów wielowymiarowych, czyli pomiarów multimenzurandu. W treści tego Suplementu nie ma sformułowanych wymagań co do poprawnego zaokrąglania i przekazywania takich danych pomiarowych przy ograniczonej liczebności próbek.

Forma komunikacji w nauce i technice jest obecnie w znacznym stopniu zorientowana na tradycyjną formę przekazu – na papierze. Zdaniem i to nie tyko Ezheli, nie wystarcza już ona do wymiany danych z doświadczeń wieloparametrowych i dla wyznaczania wartości i niepewności pochodnych stałych fizycznych oraz w metrologii o najwyższych dokładnościach. Jednakże dzięki rozwojowi elektronicznych środków publikowania, czyli e-publikatorów, otwarły się nowe możliwości, które pozwalają uniknąć występujących ograniczeń.

Jako rozwiązanie zapobiegające powyższym niedogodnościom Ezhela zaproponował przejście na dwuczęściową formę publikacji [4]. Tradycyjnemu tekstowi opisowemu, dobrze już redakcyjnie sformalizowanemu, towarzyszyłyby związane z nim pliki komputerowe przedstawiające dane źródłowe o odpowiedniej rozdzielczości cyfrowej, w pełni czytelne dla komputera, przeznaczone do długotrwałego przechowywania i dostępne w Internecie. Dzięki temu recenzenci i potencjalni użytkownicy raportowanych danych pomiarowych będą mogli sprawdzać ich kompletność i spójność. Ułatwiłoby to pobieranie i zapobiegnie degradacji tych danych przy przekształcaniu ich formy zapisu i zapewniło praktycznie wieczne ich przechowywanie. Propozycja ta nie trafiła dotąd do praktyki jako sposób na podwyższanie jakości danych i weryfikację wiedzy. Jednak podobną opinię wyrażono z kręgów OECD, by publikację i rozpowszechnianie danych oprzeć na e-publikacjach (cytat w [4]).

Zdaniem V. Ezheli [4, 8, 9], nie ma jeszcze w pełni przygotowanych podstaw merytorycznych i organizacyjnych do pełnego wykorzystywania istniejących już obecnie ogromnych możliwości e-publikatorów w zachowywaniu i przekazywaniu poprawnych wyników pomiarów o praktycznie nieograniczonej objętości z hyperlinkami i z szybką transmisją. Dla realizacji tego celu należy sformułować odpowiednie standardy dla plików odczytywanych i zrozumiałych dla komputera, związanych z publikacjami na papierze i zawierających pełne dane pomiarowe prawidłowo przedstawione liczbowo.

Przy stosowaniu plików komputerowych pojawi się np. kilka nowych zagadnień wymagających ustaleń międzynarodowych, np.:

- jaka powinna być minimalna struktury danych, aby prawidłowo zapisać wyniki badań?
- jak kontrolować precyzję numeryczną danych przy ich przetwarzaniu, aby nie utracić wyników i ich dokładności?

Przed metrologią i informatyką oraz innymi dziedzinami działalności eksperymentalnej stoi więc duże wyzwanie – opracowanie zaleceń do wyrażania i przekazywania wielowymiarowych danych pomiarowych za pomocą e-publikacji i wdrożenia do praktyki metod numerycznej weryfikacji tych danych. Podstawę do opracowania projektu takich przepisów mogą stanowić standardy robocze już stosowane w światowych centrach danych, np. w NIST (USA). Przykłady omówiono w [8, 9]. Opierając się na dotychczasowym trybie i tempie udoskonaleń i rozszerzania zakre-

su przepisów GUM dotyczących niepewności pomiarów autor nie jest w stanie przewidzieć kiedy tak duże zmiany proponowane przez Ezhelę mogą się pojawić.

Podsumowanie i wnioski

W tym ostatnim rozdziale przedstawiono w zarysie problemy występujących przy wyrażaniu i publikowaniu wyników pomiarów pośrednich wielowymiarowych, czyli pomiarów multimenzurandu. Omówiono w skrócie sposób opisu wyników tych pomiarów w ujęciu wektorowym. Podano przykłady zaokrąglania danych pomiarów dwuparametrowych przy ich przetwarzaniu i wyznaczaniu wyników oraz wskazano na konieczność poprawy zalecenia 5.1.2 w GUM o niepewności funkcji silnie nieliniowych.

Nie ma jeszcze powszechnie akceptowanych standardów numerycznego wyrażania danych pomiarów wieloparametrowych i możliwości ich numerycznej weryfikacji przy tradycyjnej formie publikacji „na papierze" oraz przy formie elektronicznej stosowanej obecnie w dość dowolny sposób. Można by tego uniknąć po ujednoliceniu sposobu wyrażania danych wielowymiarowych w formie „czytelnej i zrozumiałej" dla komputera. Przewiduje się też istotne zmiany w sposobie posługiwania się danymi pomiarowymi.

Nawiązano do propozycji podanej przez V. Ezhelę w [4] i opisanej we wspólnych publikacjach [8, 9], aby zakresem przepisów międzynarodowych objąć dualny sposób prezentacji i publikowania danych pomiarów wieloparametrowych – na papierze i w postaci towarzyszącej e-publikacji.

Należałoby też zbadać – jak stosowane w podstawowych badaniach fizycznych metody i procedury precyzyjnego przetwarzania ogromnych zbiorów danych wieloparametrowych można zaadaptować do wykorzystania w metrologii najwyższych dokładności i w wieloparametrowych pomiarach użytkowych o bardziej odpowiedzialnych zastosowaniach. Po niezbędnych uproszczeniach można by też je stosować i w pomiarach użytkowych z wieloparametrowymi próbkami o małej liczbie obserwacji. Dotyczy to w szczególności pomiarów o rozrzutach wyników opisywanych niepewnością typu A, porównywalną lub większą od niepewności typu B ujmującej wpływy nieusuwalnych lecz nieznanych instrumentalnych błędów systematycznych i błędów metody. Wyniki tych badań mogły by stanowić podstawę uzupełnień przyszłej udoskonalonej wersji Suplementu 2 do GUM [1a].

Zracjonalizowane sposoby wektorowego opisu wielowymiarowych danych pomiarowych i omówiony sposób wyznaczania ich niepewności, zracjonalizowany w stosunku do Suplementu 2 w GUM, mogą być już użyteczne w wielu dziedzinach nauki i techniki, m.in. w identyfikacji obiektów sterowania i identyfikacji powiązanych ze sobą zmian parametrów obiektów pod wpływem wielu równoczesnych oddziaływań, w badaniach materiałów elektronicznych, w wieloparametrowych pomiarach powierzchni, w nanotechnologii, w monitoringu i diagnostyce urządzeń, procesów przemysłowych i stanu środowiska oraz w badaniach medycznych.

Literatura

1. *Wyrażanie Niepewności Pomiaru Przewodnik*. Wydawnictwo Głównego Urzędu Miar, Warszawa, Alfavero 2002 – polskie tłumaczenie J. Jaworskiego: *Eva-*

luation of measurement data – Guide to the expression of uncertainty in measurement (GUM) wyd. 2 z 1995 r. Aktualna wersja GUM: www.bipm.org/utils/common/documents/jcgm/JCGM_100_2008_E.pdf.

1a. Supplement 2. *Extension to any number of output quantities*, Joint Committee for Guides in Metrology JCGM 102 2011

2. *Measurement Uncertainty Analysis Principles and Methods*, NASA Measurement Quality Assurance Handbook – ANNEX 3, NASA-HDBK-8739.19-3, 2010

3. Muciek A., *Matematyczny model propagacji niepewności w pomiarach pośrednich*. Podstawowe Problemy Metrologii, Materiały Sympozjum ppm'03, seria: Konferencje nr 5, Oddz. PAN w Katowicach, 2003, 593–604.

4. Ezhela V., *Physics and Metrology*. Materiały V Kongresu Metrologii KM 2010, Politechnika Łódzka, CD.

5. Ezhela V., *Comments on some clauses of GUM which provoking the incorrect presentation of measured data in scientific literature*. Materiały V Kongresu Metrologii KM 2010, Politechnika Łódzka, CD.

6. Warsza Z., Ezhela V., *Wyznaczanie parametrów multimenzurandu z pomiarów wieloparametrowych, Część 1. Podstawy teoretyczne w zarysie*. „Pomiary Automatyka Robotyka", Nr 2, 2011, 55–61.

7. Warsza Z., Ezhela V., *Wyznaczanie parametrów multimenzurandu z pomiarów wieloparametrowych, Część 2. Reguły zaokrąglanie, nieścisłości w przewodniku GUM*. „Pomiary Automatyka Robotyka", Nr 6, 2011, 64–70.

8. Ezhela V., Warsza Z., *Nieścisłości stałych podstawowych i propozycja standaryzacji dualnego sposobu publikowania wyników pomiaru multimenzurandu*. „Pomiary Automatyka Kontrola", Vol. 57, Nr 5, 2011, 486–490.

9. Warsza Z., Ezhela V., *O wyrażaniu i publikowaniu danych pomiarów wieloparametrowych – stan aktualny a potrzeby*. „Pomiary Automatyka Robotyka", Nr 10, 2011, 68–76.

10. Warsza Z.L., *Evaluation and Numerical Presentation of the Results of Indirect Multivariate Measurements. Outline of Some Problems to Be Solved*. Chapter in book: Advanced Mathematical & Computational Tools in Metrology and Testing AMCTM IX, ed. by Franco Pavese, Markus Bär et all, Series on Advances in Mathematics for Applied Sciences, Vol. 84, World Scientific Books Singapur 2012, 418–425

11. Mohr P.J., Taylor B.N., Newell D.B., *The 2010 CODATA Recommended Values of the Fundamental Physical Constants*, NIST (2011, ver. 6.2). [http://physics.nist.gov/cuu/Constants/index.html]

12. Warsza Z.L., *Zasady wyznaczania wyników pomiarów pośrednich wieloparametrowych – w zarysie*. „Elektronika", Nr 6, 2012 **(**wkładka *Technika Sensorowa*), 61–67.

13. Fotowicz P., *Propagacja niepewności i rozkładów przy wyznaczaniu obszaru rozszerzenia dwuwymiarowego menzurandu wektorowego*. Materiały X Konferencji Naukowo-Technicznej PPM'14, 2014, 76–79.

Uzupełniająca podstawowa literatura polska

14. Szydłowski H. i inni, *Teoria pomiarów*. PWN Warszawa 1981(354 –379)

15. Kukiełka L., *Podstawy badań inżynierskich*. PWN, Warszawa 2002

16. Pawłowski J., *Wprowadzenie do teorii kopuł*. Kraków, marzec 2009, Internet

Dodatek 1

Przykłady obliczenia efektywnych estymatorów próbek z rozkładu równomiernego i Laplace'a

W praktyce mogą pojawić się nieprawidłowości w ocenie dokładności wyniku pomiarów, gdy rzeczywisty rozkład prawdopodobieństwa danych pomiarowych istotnie różni się od rozkładu przyjętego jako ich model matematyczny. Zgodnie z zaleceniami Przewodnika GUM oblicza się wartość średnią arytmetyczna i wyznacza jej odchylenie średnie standardowe jako niepewność standardową u_A. Oba te parametry są najbardziej wiarygodnymi estymatorami tylko dla rozkładu normalnego opisanego funkcją Gaussa. Natomiast dla wielu innych rozkładów prawdopodobieństwa istnieją efektywniejsze, tj. o mniejszym odchyleniu standardowym parametry statystyczne do oceny wartości i niepewności wyniku pomiaru, które powinno się wykorzystywać (patrz rozdziały 6–9). Zilustrują to podane poniżej dwa przykłady oceny dokładności menzurandu dla próbek o danych z populacji o rozkładzie równomiernym oraz dwuwykładniczym czyli Laplace'a. Symulowane przykłady liczbowe zaczerpnięto ze wspólnych z M. Dorozovetsem prac [4, 5].

1. Efektywność estymatorów próbki z rozkładu równomiernego

Dla próbki z populacji o rozkładzie równomiernym (używa się też nazw: rozkład jednostajny lub prostokątny) wg zasad statystyki [M3, 125–127] najbardziej efektywną oceną wyniku pomiaru jest środek rozstępu (ang. *midrange*), gdyż ma on najmniejsze odchylenie standardowe (rozdz. 6, rys. 1). Oblicza się go jako wartość średnia q_{sr} z najmniejszego oraz największego z elementów próbki po ich uporządkowaniu według wartości i po usunięciu danych odstających, czyli outlierów, wg kryteriów innych dla tego rozkładu [6] niż testy Grubbsa dla rozkładu normalnego:

$$q_{sr} = \frac{q_{s,1} + q_{s,n}}{2} \tag{1}$$

Standardowe odchylenie (niepewność) środka rozstępu próbki oblicza się [M3] ze wzoru:

$$u_A(x_{r,s}) \equiv s_{s,r} = \frac{V}{\sqrt{2\,(n-1)\cdot(n-2)}} \tag{2}$$

gdzie $V = q_{s,n} - q_{s,1}$ jest rozstępem danych próbki.

Według zaleceń GUM próbkę danych z rozkładu równomiernego opisuje się przez wartość średnią i jej odchylenie standardowe, czyli niepewność u_A wyznaczana metodą statystyczną typu A. Jest to jednak estymator mniej efektywny niż środek

rozstępu $q_{r,s}$, gdyż jego standardowe odchylenie $u_A(q_{r,s})$ jest mniejsze i przy maleniu liczby obserwacji n w próbce wzrasta proporcjonalnie do $1/n$, a standardowa niepewność u_A wartości średniej – do $1/\sqrt{n}$. Środek stępu jest więc korzystniejszym parametrem do opisu dokładności próbek z populacji o rozkładzie równomiernym. Stwierdzenia te zostanie poparte poniżej przykładem liczbowym.

2. Przykład 1. Oszacowanie parametrów menzurandu dla próbki z populacji o rozkładzie równomiernym

Wykonano pomiary wartości natężenia prądu upływu izolacji I[nA] pozyskiwane kolejno przez równomierne próbkowanie. Warunki pomiaru i aparatura pomiarowa nie ulegała zmianie. Otrzymano n = 120 podanych poniżej danych liczbowych $q_j(I)$ skorygowanych przez poprawki:

1,6682	1,6773	1,6881	1,6719	1,6793	1,6833	1,6700	1,6635
1,6704	1,6705	1,6871	1,6733	1,6819	1,6711	1,6647	1,6637
1,6727	1,6832	1,6749	1,6635	1,6659	1,6888	1,6861	1,6758
1,6730	1,6621	1,6615	1,6654	1,6664	1,6856	1,6765	1,6843
1,6617	1,6649	1,6648	1,6863	1,6676	1,6634	1,6632	1,6900
1,6853	1,6839	1,6850	1,6613	1,6666	1,6808	1,6725	1,6836
1,6701	1,6798	1,6780	1,6677	1,6771	1,6767	1,6734	1,6797
1,6658	1,6754	1,6791	1,6882	1,6736	1,6840	1,6710	1,6601
1,6696	1,6697	1,6869	1,6750	1,6844	1,6832	1,6809	1,6682
1,6898	1,6710	1,6672	1,6732	1,6761	1,6667	1,6724	1,6621
1,6893	1,6712	1,6763	1,6751	1,6766	1,6718	1,6623	1,6687
1,6800	1,6795	1,6852	1,6832	1,6897	1,6653	1,6891	1,6675
1,6656	1,6658	1,6730	1,6845	1,6680	1,6664	1,6661	1,6840
1,6763	1,6847	1,6610	1,6844	1,6673	1,6610	1,6859	1,6738
1,6818	1,6693	1,6860	1,6797	1,6700	1,6653	1,6632	1,6760

Należy oszacować najlepszą wartość wyniku pomiaru oraz jej miarę niedokładności.

Rozwiązanie

Opisane zostaną kolejno wykonywane czynności.

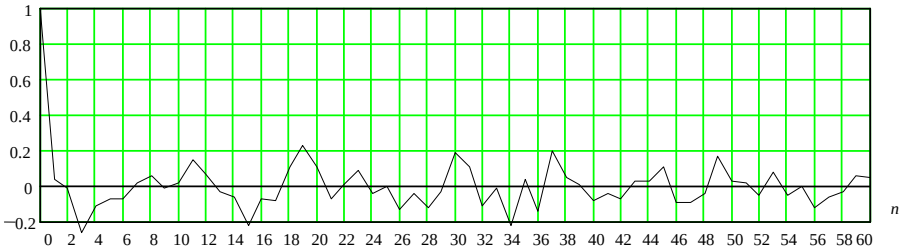
1. Zrezygnowano z utworzenia wykresu wartości otrzymywanych kolejno obserwacji, gdyż nie wnosi on istotnej informacji ilościowej.

2. Zakłada się, że wyniki obserwacji pochodzące z tej samej populacji wartości sygnału pomiarowego (lub ich odczytów). Należy najpierw sprawdzić, czy nie są one ze sobą skorelowane. Jest to istotne, gdyż jak wykazano w rozdziale 3, efektywna liczba obserwacji wykorzystywana w obliczeniach niedokładności jest wówczas mniejsza niż liczna określona w Przewodniku GUM. Do realizacji celu konieczna jest tylko informacja o wzajemnych stosunkach odstępów czasu między kolejnymi obserwacjami. Jeśli próbkowanie było równomierne, to wartości ρ_k funkcji autokorelacji należy wyznaczyć na podstawie ciągu kolejnych obserwacji wg wzoru:

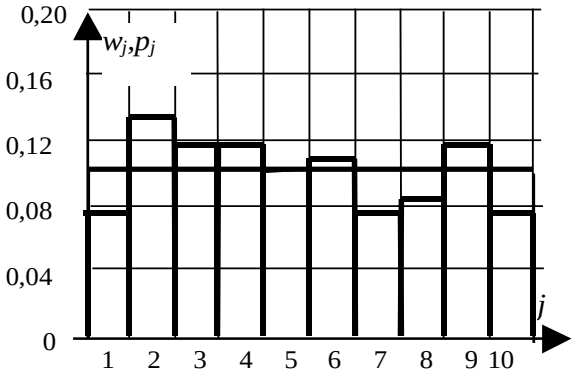
$$\rho_k = \frac{1}{n-k} \frac{\sum_{i=1}^{n-k} (q_i - \bar{q})(q_{i+k} - \bar{q})}{s^2(q_i)}$$

gdzie: k – liczba okresów próbkowania między równomiernie pobieranymi obserwacjami pomiarowymi.

Otrzymane kolejne wartości ρ_k przedstawiono jako punkty (rys. 1) i połączono linią łamaną. Funkcja ρ_k maleje do zera już dla drugiej obserwacji i następnie waha się wokół tej wartości. Oznacza to brak autoskorelowania wyników obserwacji.



Rys. 1. Eksperymentalna funkcja autokorelacji danych pomiarowych
Fig. 1. Experimental autocorrelation function of measurement data



Rys. 2. Histogram obserwacji z przykładu 2
Fig. 2. Histogram of observations of the example 2

3. W celu sprawdzenia rodzaju rozkładu sporządza się histogram próbki. Przyjęto $m = 10$ przedziałów – klas zgrupowania danych (rys. 2). Na osi x odłożono kolejne numery j przedziałów wartości obserwacji (j = 1, ..., m), a na osi y – częstość empiryczną $w_j = n_j/n$ (n_j – liczba pomiarów w przedziale j, n – liczba wszystkich pomiarów). Szerokość przedziałów wyznaczana jest z zależności:

$$h = \left[\max(q_i) - \min(q_i)\right] / m \tag{3}$$

Szerokość przedziałów wynosi 0,003. Z kształtu histogramu wywnioskowano, że jako model matematyczny dla wartości obserwacji można by przyjąć rozkład równomierny. Hipotezę tę należy sprawdzić ilościowo wykorzystując jedno z kryteriów, np. χ^2 [M1, M2]. Kryterium to oznacza, że przy liczbie stopni swobody histogramu $\nu = m-2-1$ i przyjętym poziomie istotności α rozbieżność między wynikami obserwacji, a założonym rozkładem nie powinna przekroczyć α [%]. Wartości χ^2 wyznacza się ze wzoru:

$$\chi^2 = m\sum_{j=1}^{m}\frac{\left(w_j - p_j\right)^2}{p_j} \tag{4}$$

gdzie: p_j – prawdopodobieństwo dla przedziału j wg założonego rozkładu.

Dla rozkładu równomiernego z danych tego przykładu z (4) wynika:

$$\chi^2 = 120\sum_{j=1}^{10}\frac{\left(w_j - p_j\right)^2}{p_j} = 5{,}0\,.$$

Z tabeli rozkładu χ^2, np. w [M1] uzyskuje się wartość $\chi^2_{\nu,\alpha}$. Dla histogramu z rys. 1 przy α = 0,025 i $\nu = 10-3 = 7$ otrzymuje się $\chi^2_{7,\,0{,}025} = 16$. Sprawdzenie hipotezy o rozkładzie równomiernym dało wynik pozytywny, gdyż

$$\chi^2 = 5{,}0 < \chi^2_{7,0{,}025} = 16\,.$$

4. Po tym sprawdzeniu przyjęto dla wyników obserwacji równomierny rozkład prawdopodobieństwa. Dane próbki po uporządkowaniu mają następujące wartości krańcowe

   1,6601 … **1,6900**

5. Środek rozpięcia próbki:

$$I_{\nu/2} = \left(q_{s,1} + q_{s,120}\right)/2 = \left(1{,}6601 + 1{,}6900\right)/2 = 1{,}67505$$

6. Rozpiętość czyli rozstęp próbki

$$V = q_{s,120} - q_{s,1} = 1{,}6900 - 1{,}6601 = 0{,}0299$$

7. Standardowe odchylenie wyniku pomiaru (niepewność standardowa środka rozpięcia)

$$s_I = \frac{V}{\sqrt{2(n-1)\cdot(n-2)}} = \frac{0{,}0299}{\sqrt{2\cdot 119\cdot 118}} \approx 0{,}00018$$

8. Jeżeli jako model zostanie przyjęty *a priori* normalny rozkład wyników obserwacji, to najlepszą oceną liczbową wyniku pomiaru będzie wartość średnia, podobnie jak jest to podane w Przewodniku GUM

$$\bar{I} = \frac{1}{120}\sum_{j=1}^{120} q_j(I) \approx 1{,}6745$$

9. Jej standardowa niepewność wyniosła by

$$u_A(\bar{I}) = \sqrt{\frac{\sum_{j=1}^{n}(q_j - \bar{I})^2}{n(n-1)}} \approx 0{,}00078$$

Jest ona aż około 4,4 razy większa od standardowej niepewności środka rozstępu, jeśli nie ma outlierów, lub zostały one z próbki usunięte! Ma to duże znaczenie dla takich pomiarów, w których niedokładność wynikająca z rozrzutu wartości obserwacji jest dominująca.

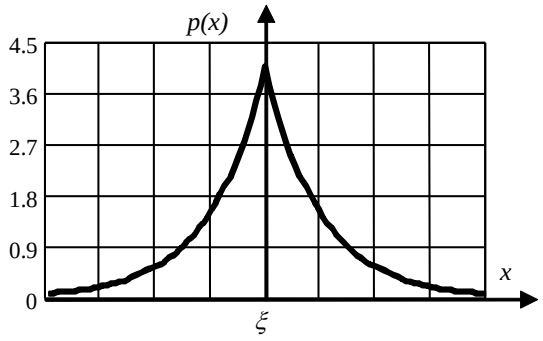
10. Jeśli niepewność typu B jest pomijalna, to wynik pomiaru można zapisać:

$$I = 1{,}6745 \pm 0{,}0008$$

Janusz Jaworski wykazał [3], że niepewności rozszerzone dla dużych próbek z rozkładu normalnego i równomiernego, wyznaczone w oparciu o ich niepewności standardowe dla współczynnika rozszerzenia k = 2 różnią się poniżej 5%. Na tej podstawie uważa on, że przy szacowaniu niepewności standardowej $u_A(x)$ wystarczy posługiwać się tylko wzorami podstawowymi z Przewodnika GUM, tj. tylko takimi jak dla rozkładu normalnego. Nie uzasadniają tego wniosku wyniki powyższego przykładu 1. Gdy ocenę niedokładności rozkładu równomiernego oparto na wyznaczeniu środka rozpięcia i jego odchylenia standardowego wg wzorów (1) i (2), to pojawiły się zbyt duże różnice.

3. Efektywność estymatorów próbki z populacji o rozkładzie Laplace'a

Rozkład Laplace'a (dwuwykładniczy) otrzymuje się jako różnicę dwu niezależnych zmiennych losowych o rozkładach wykładniczych [M3, tab. V]. Rozkład ten podano na rys. 3.



Rys. 3. Rozkład Laplace'a
Fig. 3. Laplace distribution

Opisuje go następująca zależność:

$$p(x) = \frac{1}{2\lambda} e^{-\frac{|x-\xi|}{\lambda}}, \quad -\infty < x < \infty \tag{5}$$

gdzie ξ jest punktem środkowym rozkładu, λ jest parametrem jego szerokości.

Dla tego rozkładu najbardziej efektywną oceną wyniku pomiaru nie jest średnia kwadratowa. Jest nią natomiast mediana (skrót ang. Med) jako wartość środkowa uporządkowanej próbki [M3].

Rozpatrzymy przypadek, gdy uzyskane wyniki obserwacji q_i nie są skorelowane. Należy uporządkować je kolejno wg wartości, tj.:

$$q_{s,1} \le q_{s,2} \le q_{s,3} \le \ldots \le q_{s,n}$$

W zapisie tym zastosowano nowe oznaczenia obserwacji $q_{s,i}$ wynikające z ich kolejności po uporządkowaniu. Wówczas medianę q_{med} oblicza się następująco:

– dla nieparzystej liczby obserwacji

$$q_{med} = q_{s,(n+1)/2} \tag{6a}$$

– dla parzystej liczby obserwacji

$$q_{med} = \frac{q_{s,n/2} + q_{s,n/2+1}}{2} \tag{6b}$$

Standardowe eksperymentalne odchylenie mediany próbki z populacji o rozkładzie Laplace'a jest √2 razy mniejsze od standardowego odchylenia wartości średniej próbki

$$s_{med} \approx \sqrt{\frac{s^2}{2n}} = \frac{1}{\sqrt{2}} \frac{s}{\sqrt{n}} = \frac{s}{\sqrt{2}} = \frac{u_A(\bar{q})}{\sqrt{2}} \tag{7}$$

4. Przykład 2. Obliczenia estymatorów menzurandu dla próbki z populacji o rozkładzie Laplace'a

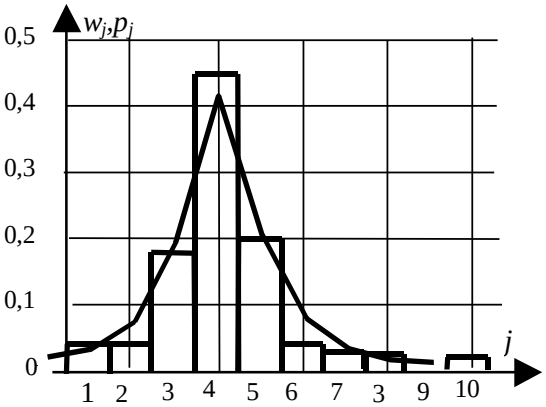
Dana jest próbka o n = 100 skorygowanych za pomocą poprawek wyników q_i obserwacji uzyskanych przy równomiernym próbkowaniu:

7,496	7,670	7,497	7,632	7,393	7,615	7,489	7,699	7,554
7,520	7,455	7,826	7,523	7,752	7,479	7,632	7,521	7,559
7,424	7,545	7,523	7,601	7,505	7,576	7,592	7,501	7,543
7,413	7,460	7,540	7,601	7,413	7,674	7,523	7,543	7,393
7,520	7,569	7,699	7,478	7,252	7,525	7,540	7,431	7,615
7,497	7,467	7,537	7,656	7,523	7,501	7,552	7,297	7,592
7,558	7,413	7,614	7,478	7,528	7,516	7,536	7,513	7,505
7,826	7,582	7,275	7,427	7,545	7,367	7,581	7,537	7,496
7,543	7,467	7,559	7,424	7,559	7,481	7,552	7,516	7,521
7,582	7,478	7,615	7,588	7,481	7,634	7,478	7,617	7,490
7,537	7,481	8,067	7,426	7,478	7,540	7,581	7,367	7,229
7,592								

Należy oszacować wartość najlepszego wyniku pomiaru i jego odchylenie standardowe.

Rozwiązanie

1. Sporządzono funkcję autokorelacji próbki. Otrzymano przebieg podobny jak na rys 1 dla rozkładu równomiernego. Wyniki obserwacji okazały się nieskorelowane.
2. Sporządzono histogram serii obserwacji zawartych w próbce wykorzystując m = 10 klas grupowania (rys. 4) o szerokości wynoszącej 0,0839. Szerokość przedziałów wyznaczono ze wzoru (3).



Rys. 4. Histogram obserwacji z przykładu 1
Fig. 4. Histogram of observations of the example 1

Z kształtu tego histogramu wywnioskowano o możliwości przyjęcia jako modelu matematycznego rozkładu Laplace'a. Sprawdzono hipotezę o tym rozkładzie wykorzystując kryterium χ^2 na poziomie istotności α = 0,05 i liczbie stopni swobody $v = n_p - 2 - 1 = 10 - 3 = 7$

$$\chi^2 = 100 \sum_{j=1}^{10} \frac{\left(w_j - p_j\right)^2}{p_j} = 9{,}102 \tag{5}$$

$$\chi^2 = 9{,}102 < \chi^2_{7,0,05} = 14{,}11$$

Uzyskano wynik pozytywny. Podobne sprawdzenie dla rozkładu normalnego daje większą różnicę.

- W wyniku sprawdzenia jako model matematyczny rozkładu prawdopodobieństwa wyników obserwacji przyjęto rozkład Laplace'a.
- Próbka jest parzysta i po uporządkowaniu ma następujące wartości środkowe:

... **7,523** **7,525** ...

3. Wartość mediany próbki wynosi

$$q_{med} = \left(q_{s,50} + q_{s,51}\right)/2 = \left(7{,}523 + 7{,}525\right)/2 = 7{,}524$$

4. Ocena wartości standardowego odchylenia wartości obserwacji próbki

$$s(q_i) = \sqrt{\frac{1}{100-1}\sum_{i=1}^{100}\left(q_i - \bar{q}\right)^2} = 0{,}113656 \approx 0{,}114$$

gdzie: $\bar{q} = \frac{1}{100}\sum_{i=1}^{100} q_i = 7{,}53113 \approx 7{,}531$ – wartość średnia próbki.

5. Standardowe odchylenie wyniku pomiaru (mediany) wynosi

$$s(q_{med}) = \frac{s(q_i)}{\sqrt{2n}} = \frac{0{,}113656}{\sqrt{2 \cdot 100}} \approx 0{,}008$$

6. Przy przyjęciu rozkładu normalnego jako modelu matematycznego wartości obserwacji, najlepszą oceną wyniku byłaby wartość średnia. Jej standardowa niepewność wyniosłaby nieco więcej, tj.:

$$u_A(\bar{q}) = \frac{s(q_i)}{\sqrt{n}} = \frac{0{,}113656}{\sqrt{100}} \approx 0{,}0114$$

5. Wnioski z przykładów liczbowych

1. Przy wyznaczaniu najlepszej oceny wyniku pomiaru oraz szacowaniu jego niedokładności, jeśli nie ma informacji, że wyniki obserwacji podlegają na pewno zadanemu *a priori* rozkładowi prawdopodobieństwa, to przy stosowaniu konwencjonalnego podejścia należy sporządzić histogram i sprawdzić hipotezę o przyjętym modelu rozkładu oraz przyjąć rodzaj rozkładu w jak największym stopniu odpowiadający wynikom obserwacji. Przyjęcie nieadekwatnego rozkładu może spowodować zmianę wartości wyniku pomiaru i znacznie zmienić ocenę jego niepewności, jak ilustrują to omówione powyżej przykłady 1 i 2.
2. Dla uniknięcia niejednoznaczności przy opisie niedokładności wyników obserwacji o innych rozkładach niż normalny standardowe odchylenia innych ich estymatorów niż wartość średnia powinno się oznaczać w odmienny sposób niż niepewność u_A wyznaczana metodą typu A według Przewodnika GUM. Odmienność rozkładu obserwacji trzeba uwzględnić też przy jego splocie z innymi rozkładami przy pomiarach pośrednich i z rozkładem wypadkowym oszacowanym dla niepewności typu B.
3. Przedstawione powyżej zasady wyboru właściwego rozkładu obserwacji pomiarowych wraz z eliminacją trendu i składowych okresowych (szczegółowo omawianych w rozdziale 2) oraz wyznaczaniem niepewności typu A dla obserwacji skorelowanych (rozdział 3) są propozycjami dotychczas nieobjętymi w treści Przewodnika GUM i jego Suplementów.

Literatura

1. *Guide to the Expression of Uncertainty in Measurement*, ISO 1992, revised and corrected 1995.
2. *Wyrażanie Niepewności Pomiaru. Przewodnik,* tłumaczenie i komentarz J. Jaworskiego, Wydawnictwo Głównego Urzędu Miar Alfavero, Warszawa 1999.

3. Jaworski J.M.: *Niepewność pomiaru, rozkłady zmiennych losowych modelujących błędy pomiaru.* Materiały XXXVI MKM Prace Komisji Metrologii Oddziału PAN w Katowicach, seria: Konferencje, Nr 6, 2004, 267–276.
4. Dorozhovets M., Warsza Z.L., *Propozycje rozszerzenia metod wyznaczania niepewności wyniku pomiarów wg Przewodnika GUM* (1), „Pomiary Automatyka Robotyka", R. 11, Nr 1, 2007, 6–15.
5. Dorozhovets M., Warsza Z.L., *Wpływ nieadekwatności wyboru parametrów rozkładu prawdopodobieństwa na niepewność typu A.* „Pomiary Automatyka Kontrola", Nr 9 bis, 2007, 25–28.
6. Dorozhovets M., *Testowanie obserwacji istotnie odstających w próbach rozkładzie jednostajnym*. „Mechanik", Nr 11, 2016, 1596–1599.

Uzupełniająca literatura matematyczna

[M1] Tylor J.R., *Wstęp do analizy błędu pomiarowego.* Wydawnictwo Naukowe PWN, Warszawa 1995 (tłumaczenie oryginału ang.: *An Introduction to error analysis. The study of uncertainties in physical measurements* Oxford University Press California 1982).
[M2] Bendat J.S., Piersol A.G., *Metody analizy i pomiaru sygnałów losowych*, WNT Warszawa 1976 (tłumaczenie polskie wcześniejszego wydania oryginału: *Random Data. Analysis and measurement procedure*, John Wiley & Sons N.York, Chichester, Brisbane Toronto… 1986).
[M3] Pacut A., *Prawdopodobieństwo Teoria Modelowanie Probabilistyczne w Technice*. WNT 1985.
[M4] Box G.E.P., Jenkins G.M., Reinsel G.C., *Analiza szeregów czasowych: prognozowanie i sterowanie*, PWN, Warszawa 1983 (polskie tłum. wydania wcześniejszego z 1971 r. oryginału *Time series Analysis: Forecasting and Control.* 3rd ed. Prentice Hall New Jersey 1994).

O propozycjach wyznaczania niepewności rozszerzonej dla nowego GUM

Stan prac prowadzonych przez Międzynarodowy Komitet d/s Przewodników w Metrologii JCGM nad nową wersją Przewodnika GUM 2 omawia publikacja [1]. Wersja ta jest oparta na podejściu według Bayesa. Proponuje się w niej zmodyfikowany wzór do wyznaczania standardowej niepewności pomiaru $u(y)$, tj.

$$u(y) = \sqrt{\frac{n-1}{n-3}u_{\mathrm{A}}^2 + u_{\mathrm{B}}^2} \tag{1}$$

Wzór ten uwzględnia zwiększanie się niepewności u_{A} obliczanej wg Przewodnika GUM 1 metodą statystyczną typu A, wraz ze zmniejszaniem się liczebności danych pomiarowych n w próbce. Jest ono opisane rozkładem Studenta tak, jak dla próbki z populacji o rozkładzie normalnym.

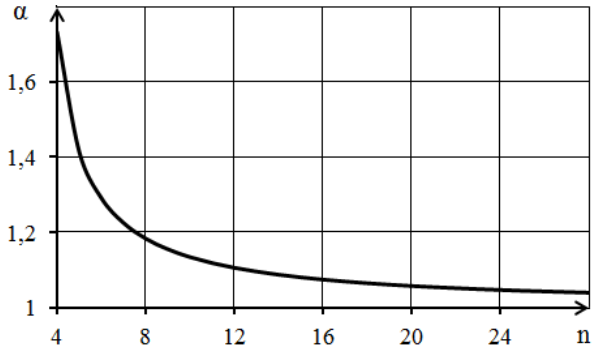
Standardowe odchylenie wartości średniej $\bar{x}$ próbki przy pomiarze pojedynczej wielkości X przy $u_{\mathrm{B}} = 0$ wynosi:

$$u(x) = \alpha\, s(\bar{q}) \tag{2}$$

gdzie: $\alpha \equiv \sqrt{\frac{n-1}{n-3}}$, $\bar{x} = \bar{q} = \frac{1}{n}\sum_{i=1}^{n} q_i$, $u_A(x) = s(\bar{q}) = \frac{s(q_i)}{\sqrt{n}} = \sqrt{\frac{1}{n(n-1)}\sum_{i=1}^{n}(q_i - \bar{q})^2}$

(patrz tab. 2 w rozdz.1)

Wykres zależności współczynnika α od n przedstawiono na rys. 1.



Rys. 1. Zależność współczynnika α od liczebności próbki n
Fig. 1. Coefficient α as function of the number n of sample elements

Przebieg charakterystyki współczynnika α w funkcji n jest taki, jak dla danych próbki z populacji o rozkładzie normalnym i wynika z rozkładu Studenta. Wartość współczynnika α wynikającą ze wzoru (1) można stosować tylko dla próbek o $n \geq 4$. Botsiura i Zakharov podają [7], że symulacja metodą Monte Carlo wartości średniej ze ścisłego wzoru rozkładu Studenta dla liczby $n = 3$ daje stabilną wartość równą 4.

Podejście Bayesa w zastosowaniu do zagadnień niepewności pomiarów i konsekwencje nowego wzoru (1) oraz sposób obliczania niepewności rozszerzonej U dla próbki z populacji o rozkładzie normalnym omówił w sposób bardzo przystępny Paweł Fotowicz w publikacjach [2–4].

Dla innych rozkładów i ich estymatorów rozszerzanie się rozkładu przebiega inaczej niż dla rozkładu normalnego i również zależy od liczebności próbki n. Przykładem są podane w pracy [4] i omówione w rozdziale 6 przebiegi estymatorów odchyleń standardowych wartości średniej oraz środka rozstępu próbek o liczebności n z populacji o rozkładzie równomiernym i płasko-normalnym. Są one wyznaczone metodą Monte Carlo. Wyznaczenie niepewności rozszerzonej $U_p(y)$ o zadanym prawdopodobieństwie p, realizowane na podstawie niepewności standardowej $u(y) \equiv s(y)$ dla rozpatrywanego estymatora y menzurandu danych próbki o dowolnym rozkładzie wymaga znajomości jego współczynników α_{py} i k_{py}. Zwykle przyjmuje się też, że wartości niepewności rozszerzonej U_p dowolnych estymatorów y dla próbek o większych liczebnościach n, pobieranych wielokrotnie z tej samej populacji mają w przybliżeniu rozkłady Gaussa.

Jeśli jako estymator menzurandu próbki przyjmuje się wartość średnią $\bar{y}$, to otrzyma się

$$U_{yp}(y) = \alpha_{yp} \cdot ky_p(\bar{y}) \cdot s(\bar{y}) \tag{3}$$

W pracy o nowym GUM [1] dla współczynnika rozszerzenia k_p wartości średniej populacji o rozkładzie dowolnym, w tym niesymetrycznym oraz o rozkładzie jednomodalnym symetrycznym podano następujące wzory przybliżone

$$k_P \leq \frac{1}{\sqrt{1-p}} \qquad k_P \leq \frac{2}{3 \cdot \sqrt{1-p}} \tag{4}$$

Dla $p = 0{,}95$ otrzymuje się odpowiednio $k_p = 4{,}47$ i $k_p = 2{,}98$, gdy dla rozkładu normalnego jest $k_p = 1{,}96$.

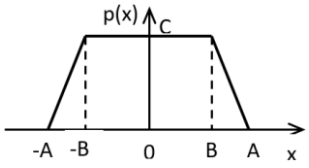
Wnikliwą analizę współczynnika k_p metodą Monte Carlo dla kilku rozkładów oraz różnych liczebności próbki n przeprowadził I. P. Zacharow z współpracownikami [6–8]. Współpracując z E. A. Klimovą podali oni dla wartości średniej kilku podstawowych rozkładów [6] wzory określające kurtozę K oraz współczynnik rozszerzenia k_p dla prawdopodobieństwa p. Zestawiono je w tabeli 1.

Tab. 1. Wzory współczynnika rozszerzenia k_p dla kilku rozkładów prawdopodobieństwa
Tab. 1. Formulas of the coverage factor k_p for few basic PDF distributions

Rozkład prawdopodobieństwa	Kurtoza K	Współczynnik rozszerzenia k_p dla prawdopodobieństwa p	k_p dla p = 0,95
arc sin	–1,5	$\sqrt{2}\sin\frac{\pi p}{2}$	1,40985
równomierny	–1,2	$p\sqrt{3}$	1,64545
trójkątny	–0,6	$\left(1-\sqrt{1-p}\right)\sqrt{6}$	1,90177
normalny	0	$t_p(\infty)$	1,95996
trapezowy Trap	$-1{,}2\frac{1+\lambda^4}{\left(1+\lambda^2\right)^2}$	$k_{0,95}=\sqrt{3}\frac{1+\lambda-\sqrt{0{,}2\lambda}}{\sqrt{1+\lambda^2}}$	
normalny + Studenta	$6/(\nu\text{-}4)$	$k_{0,95}=t_{0,95}(\nu)\sqrt{(\nu-2)/\nu}$	

gdzie: t_p – współczynnik wg rozkładu Studenta, $\nu=n\text{-}1$ – liczba stopni swobody

Rozkład Trap ma postać symetrycznego trapezu liniowego (rys. 2). Powstaje on jako splot dwu rozkładów równomiernych o różnej szerokości. Długości podstaw trapezu są sumą i różnicą długości rozstępów obu tych rozkładów. Krańcowymi przypadkami rozkładu Trap jest rozkład trójkątny (dla dwu jednakowych rozkładów równomiernych) i rozkład równomierny.



Rys. 2. Rozkład trapezowy Trap
Fig. 2. PDF of trapeze distribution Trap

Jeśli niepewności splatanych rozkładów równomiernych wynoszą u_1 i $u_2 = \lambda\, u_1$, (gdzie $0 \le \lambda \le 1$), to dla niepewności tworzonego przez nich rozkładu trapezowego otrzymuje się:

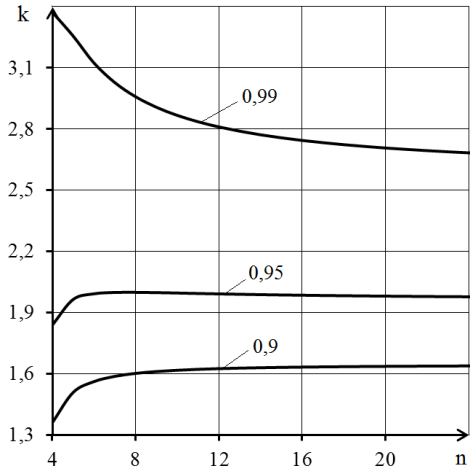
$$u = \sqrt{u_1^2 + u_2^2} = u_1\sqrt{1+\lambda^2} \tag{5}$$

Punkty charakterystyczne rozkładu Trap (rys. 2) opisane są w funkcji u_1 i λ wzorami:

$$A=\sqrt{3}(1+\lambda)u_1, \quad B=\sqrt{3}(1-\lambda)u_1, \quad C=1/\left[2\sqrt{3}(1+\lambda)u_1\right] \tag{6a, b, c}$$

Wartości kurtozy i współczynnika rozszerzenia $k_{0,95}$ rozkładu Trap podano w tabeli 1. Inne estymatory dla tego rozkładu omówiono w rozdziałach 7 i 8.

Zakharov i Botsiura w pracy [7] wyznaczyli zależności współczynnika rozszerzenia k_p w funkcji liczebności próbki n dla kilku wartości prawdopodobieństwa p z uwzględnieniem wzorów (1) i (2). Przedstawiono je na rysunku 3. Najbardziej płaski jest przebieg dla p = 0,95. Wyznaczyli też współczynnik k_p dla rozkładu płasko-normalnego jako splotu rozkładu równomiernego z rozkładem Gaussa.



Rys. 3. Współczynnik rozszerzenia k w funkcji liczebności próbki z populacji o rozkładzie normalnym

Fig. 3. Coverage coefficient k as function of number of elements in sample from set of normal PDF

Natomiast w pracy [8] ci sami autorzy zbadali metodą Monte Carlo dokładność różnych wzorów na współczynniki rozszerzenia niepewności wartości średniej, w tym również wg rosyjskiej normy GOST. Wzory (4) dają błędy sięgające nawet 100%. Do stosowania w praktyce preferują następujące wzory

$$k_{\mathrm{GUM}} = t_{0,95}(\nu_{\mathrm{eff}})\sqrt{\frac{\left(1+\gamma^2\right)}{\alpha^2+\gamma^2}} \tag{7}$$

$$k_3 = \sqrt{\frac{\left[t_{0,95}(n-1)\right]^2 + \left(k_{\mathrm{B}} u_{\mathrm{B}}\right)^2}{\alpha^2+\gamma^2}} \tag{8}$$

gdzie: $\nu_{\mathrm{eff}} = (n-1)\left(1+\gamma^2\right)^2$, $\alpha = \sqrt{\frac{n-1}{n-3}}$, $\gamma = \sqrt{n}\cdot u_{\mathrm{B}}(y)/s$, $t_{0,95}$ – dla rozkładu Studenta.

Z badań metodą MC dokładności obu współczynników rozszerzenia k_{GUM} i k_3 dla różnych kombinacji rozkładów normalnego i równomiernego oraz różnych wartości γ i n wynikło szereg interesujących wniosków. I tak wartości k_{GUM} nie zależną

od rozkładu niepewności u_B. Natomiast najmniejsze błędy, rzędu ±4,5%, uzyskuje się dla k_3 ze wzoru (8). Wzór ten należałoby uwzględnić w nowym GUM 2 opartym na podejściu wg Bayesa. Międzynarodowa akceptacja tej interesującej propozycji wymaga jeszcze szerszych badań symulacyjnych dokładności wzoru (8) z uwzględnieniem innych rozkładów niż normalny i równomierny, powtórzenia tych badań w kilku NMI i ośrodkach naukowych oraz publikacji w języku angielskim w głównych (*main stream*) czasopismach metrologicznych.

Dla rozkładów niegaussowskich inne estymatory są bardziej efektywne od średniej – patrz rozdziały 6–8 i publikacje [5, 9]. Pomimo to w literaturze jest niewiele informacji o wyznaczaniu ich niepewności rozszerzonej. M. Dorozhovets, z inspiracji autora, opracował kryteria dla wykrywania outlierów w próbkach o rozkładzie równomiernym [9]. Wyznaczył też analitycznie odpowiednik rozkładu Studenta dla niepewności s_V środka rozpięcia V tego rozkładu i podany poniżej wzór dla jego niepewności rozszerzonej U_{Vp}.

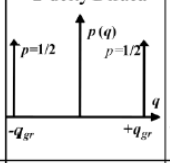
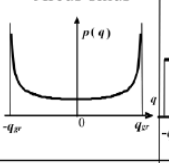
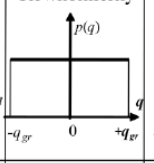
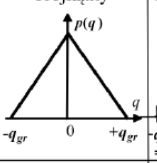
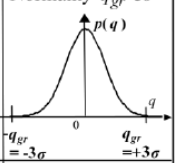
$$U_{Vp} = k_{VP} \cdot s_V = \left[(1-p)^{-\frac{1}{n-1}} - 1 \right] \cdot (n-1) \cdot \sqrt{\frac{n+2}{2 \cdot (n+1)}} \tag{9}$$

gdzie: $k_{Vp} = (1-p)^{-\frac{1}{n-1}} - 1$ – współczynnik rozszerzenia, $V = x_{i\max} - x_{i\min}$ – rozstęp, $s_V = \frac{V \cdot \sqrt{2}}{n-1} \cdot \sqrt{\frac{n+1}{n+2}}$ – odchylenie standardowe SD.

Ocena niedokładności przyrządów i nawet dopuszczalnych przedziałów zmian parametrów badanych obiektów nadal jest często dokonywana przez błąd graniczny jako przypadek najgorszy z możliwych. Jeżeli właściwości menzurandu lub dane sygnału w torze pomiarowym modelowane są jednym z rozkładów niegaussowskich, to aby oszacować niepewność rozszerzoną trzeba z błędu granicznego lub z rozstępu danych próbki wyznaczyć odchylenie standardowe σ. Zależności między tymi parametrami dla kilku podstawowych rozkładów podano w tabeli 2.

Tab. 2. Odchylenie standardowe kilku rozkładów gęstości prawdopodobieństwa o znanym rozstępie q_{gr}

Tab. 2. Standard deviation of few probability density distributions (PDF) of a known range q_{gr}

Rozkład $p(q)$	2-delty Diraca	Arcus sinus	Równomierny	Trójkątny	Normalny $q_{gr}=3\sigma$
					
Odchylenie standardowe $s(q)$	$\frac{q_{gr}}{1}$	$\frac{q_{gr}}{\sqrt{2}}$	$\frac{q_{gr}}{\sqrt{3}}$	$\frac{q_{gr}}{\sqrt{6}}$	$\frac{q_{gr}}{\sqrt{9}}$
w % od q_{gr}	100%	71%	58%	41%	33%

Literatura

1. Bich W., Cox M., Michotte C., *Towards a new GUM – An update*. "Metrologia", Vol. 53, 2016, S149–S159 IOP publishing | BIPM.
2. Fotowicz P., *Teoria prawdopodobieństwa warunkowego w metrologii*. Metrologia. Biuletyn Głównego Urzędu Miar nr1(16) 2010, 33–36.
3. Fotowicz P., *Alternatywne sposoby obliczania niepewności pomiaru*. „Pomiary Automatyka Kontrola", Vol. 56, Nr 11, 2010, 1305–1307.
4. Fotowicz P., *Modyfikacja sposobu obliczania niepewności pomiaru*. „Pomiary Automatyka Robotyka", Nr 3, 2016, 29–32.
5. Kubisa S., Warsza Z.L., *Środek rozstępu jako estymator menzurandu dla próbek z populacji o rozkładzie równomiernym i płasko-normalnym*. „Pomiary Automatyka Kontrola", Vol. 60, No. 6, 2014, 398–401.
6. Zakharov I.P., Klimova K.A., *Application of excess method to obtain reliable estimate of coverage factor*. "Sistemy Obrobotki Informacji" (SOI), 2014, Vypusk 3 (119), 24–28 (in Russ.)
7. Botsiura O.A., Zakharov I.P., *Peculiarities of evaluation of measurement uncertainty type a based on a Bayesian approach*. "Sistemy Obrobotki Informacji" (SOI) Kharkiv 2015, Vypusk 6 (131), 17–20 (in Russ.)
8. Botsiura O.A., Zakharov I.P., *Comparative analysis of various methods for calculating of coverage factor at implementation of Bayesian approach by the measurement uncertainty evaluation*. "Sistemy Obrobotki Informacji" (SOI) 2016, Vypusk 6 (143), 20–24 (in Russ.)
9. Dorozhovets M., *Testowanie obserwacji istotnie odstających w próbach o rozkładzie jednostajnym*. „Mechanik", Nr 11, 2016, 1596–1599.

Methods of Extension of the Measurement Uncertainty Analysis

Summary

This monograph presents the results of author and co-operating with him metrologists from Poland, Ukraine and other countries studies on the extension the analysis of measurement uncertainty covered by the international Guide to the Expression of Uncertainty in Measurement (GUM). It is devoted to the discussion of selected studies on the adaptation of modern basics of the probability theory and stochastic processes to metrological practice. The content of the monograph includes a preface, 12 autonomous chapters and two additives, together 191 pages of text.

Chapter 1 provides a synthesis of the currently used methods of description of the accuracy of measuring instruments through the errors and the uncertainty of measurement according to the GUM recommendations. In the following chapters are proposed methods to extend these recommendations. They are focused on applications not only in measurements of the Metrological Service, but in all utility studies.

In Chapter 2, for measurement observations obtained by sampling the continuous signal of the measured value discussed is cleaning of a row data from the unknown regularly variable systematic components. Estimating the uncertainty of type A, when these cleaned data linked with each other statistically, i.e. autocorrelated with autocorrelation function obtained experimentally from these data, is discussed in Chapter 3.

As a next presented are the results of research on the possibility of use the probability distribution in the form of the function $\cos^2$ as an alternative to the normal distribution (Ch. 4), but with a limited scope for the probability equal to 1. Lists are also the results of modeling using Monte Carlo method, for the statistics of skewness and kurtosis of small numerous data samples from normal, uniform and triangle distributions (Ch. 5).

Then other than the average the estimates of measurand and its uncertainty for the data of measurement results modeled by non-Gaussian distributions, including uniform, arc sin, trapezoidal Trap and CTrap distributions and their convolutions are considered in chapters 6–9. The midrange is the best for uniform distribution. A new two-element estimator 0.5(mean + mid range) is proposed for trapezoidal distributions (Ch. 8). It is more efficient then single-element estimators such as mean, median and mid range. The unconventional way of determining the most effective from the considered estimators for samples with unknown a priori distribution, is obtained by the use of Monte Carlo and resampling methods, without identifying the sample distribution (Ch. 9).

Chapter 10 gives the results of studies on the use in the uncertainty analysis the polynomial maximization method (PMM) – based on cumulants. Chapter 11 exam-

ines two the robust statistical methods for determining the measurement uncertainty and their implementation in interlaboratory studies. Compared are results obtained by method applied the rescaled median absolute deviation MAD_S and by the iterative Huber method with winsoryzation.

The last 12 chapter explains some of the basic issues concerning the determination of uncertainty and rounding of the indirect multidimensional measurement data. This is a clarification and some extension of the contents of Supplement 2 to the GUM.

The content of the monograph chapters is supplemented by two additions. The first contains examples of calculation the more efficient estimators of measurand than the average for uniform and Laplace distributions. The second presents a few suggestions for the determination of the expanded uncertainty, which are applicable in the new guide GUM 2 based on the Bayesian approach, and also for the midrange as the estimator.

This monograph is intended both for metrology services and for all those who make and interpret results of measurements in industry, as well as in all other areas of economy, medicine, the military and science.

Methods of Extension of the Measurement Uncertainty Analysis

Contents

Preface, Summary, Greetings

Chapters and their abstracts

Chapter 1

Methods of evaluation the measurement errors and uncertainty and need of their extension

Abstract. After the short introduction to the theme of monograph, the measurement process is discusses, the definitions and basic formulas for determining measurement errors, including the limited "worse case" errors and the uncertainty of measurement due to the international guide GUM are given. Characterized is scope and restrictions on the use of uncertainty. Given the need to expand the way of uncertainty estimation by method A data obtained by sampling the measurement signal. It should be preceded by cleaning the data from the regular unknown components and take into account the autocorrelation between the observations of a sample. For data distributions other than Normal appropriate estimators have to be applied, and in the case of small samples with the presence of outliers the robust statistics should be used. The list of a basic Polish and foreign language literature are given. In the attachment, two examples the accuracy evaluation of Uniform and Laplace pdf data samples by GUM or other method is compared.

Keywords: process of measurements, error, uncertainty, the GUM Guide

Chapter 2

Discrimination and elimination of the impact of drift and oscillation in the sample data on the uncertainty of type A

Abstract. A new approach to upgrading the type A uncertainty evaluation due GUM recommendations is presented. The set of raw data obtained by known, e.g. regular sampling can be "cleaned" by elimination the systematic components as trend and periodical ones. The positive influence of this cleanings on standard uncertainty type A is presented and disused in detail on two numerical examples.

Keywords: uncertainty type A, systematic components of rough data, trend, harmonic disturbances

Chapter 3
Evaluation of the uncertainty of auto-correlated measurement data

Abstract. Principles of the estimation of uncertainty calculation due international guide GUM [1] are presented and limitations of their type A method are discussed. Expanding of the application range of this method to regularly sampled statistically related observations is proposed by taking in consideration the influence of their autocorrelation function. Obvious is previous identification and cleaning of the raw sample data from regularly variable components. Equivalent number of effective not correlated observations is introduced for use. This upgraded method of uncertainty type A calculations is compatible with other GUM recommendations. It has been illustrated by the numerical example. Conclusions and literature is also included.

Keywords: measurement, uncertainty, autocorrelation, equivalent number of observations

Chapter 4
Function $\cos^2$ as model of probability density distribution – properties and examples of application

Abstract. Unconventional probability density distribution function (PDF) of the form of the one period of cosines function shifted up and of its area under curve equal of 1 is proposed as alternative to Normal distribution, i.e. Gauss function. It is presented that if a cosine function is shifted up by its amplitude *A*, such function ($2A\cos^2(\cdot)$), can be applied as approximation of the central part of Normal PDF (±2,5 of its standard deviation) with the accuracy better then 2,1%. In reality the distribution of measurements is always of limited range and uncertainty type A obtained by COS^2 PDF is smaller than recommended by international Guide GUM [1]. Analysed data from experiment and simulated data confirmed that new proposed unconventional distribution have an advantage of a lowering uncertainty interval by comparison to Gauss distribution at the same level of confidence and calculations are simpler and faster. Distribution COS^2 could be successfully applied for the fast uncertainty evaluation of measurements, for its automation and in statistics.

Keywords: COS^2 probability density function, replacement of Normal distribution

Chapter 5
Statistics of skewness and kurtosis of small measurement samples from populations of normal and two other distributions

Abstract. Statistics of skewness and kurtosis distributions and their basic parameters for a set of samples of certain small numbers of elements are finding. These distributions were determined using the Monte Carlo method. The samples were repeatedly taken at random from a normally distributed population and for comparison from the population of a few other simple distributions. Knowledge of skewness and kurtosis statistics may allows a more reliable estimate of the standard deviation and the uncertainty of the value of the measurand estimator from the

measurement samples of a small number of observations, when the range of their value distribution is known. Also compared are the average modules and standard deviation of skewness of small samples taken from three populations of normal, uniform and triangular distributions.

Keywords: distributions: normal, uniform, triangular; measurement sample, skewness, kurtosis

Chapter 6
Midrange as estimator of measurand for data from population of uniform and flatten-Gaussian distributions

Abstract. In this paper statistical properties of samples of varying number of observations, taken from a population of uniform distribution, have been examined by the Monte Carlo simulation. Their midrange has a smaller standard deviation than the mean value recommended by the Guide GUM as estimator of measurand. Calculated also is distribution similar as Student for Gauss, coverage factors and expanded uncertainty of such samples. Then also was found that for samples from the population of flatten-Gaussian distribution with increasing the level by the normal distribution the advantage of mid-range quickly decreases. Some final conclusions are enclosed.

Keywords: uniform PDF, flatten–Gaussian PDF, estimator, sample mid-range, mean value, uncertainty

Chapter 7
Single component estimators of measurand data of few non-Gaussian PDF-s

Abstract. The single-component estimators of the measurand value of data samples modelled by the few non-Gaussian probability distributions (PDF) are considered and their accuracy are evaluated. For symmetrical trapezoidal PDF of straight as well curved sides, using the Monte-Carlo method of simulation standard deviation (SD) of mean value, mid-range and mediana estimators are evaluated. It is established that in the ratio of upper and bottom bases of trapeze in the range from 1 to 0.35 the most accurate is the mid-range. Below this range smaller is standard deviation (SD) of the mean value. Investigated are also ratios of mean and midrange SD of arc sin convolutions with normal or uniform PDF models for simulated data samples.

Keywords: trapezoidal probability density functions – PDF, midrange, estimators, measurand uncertainty evaluation

Chapter 8
Two-component estimators of the measurand of data of trapezoidal probability distributions

Abstract. Two-component estimators (2C) of the measurand value of data samples modeled by symmetric trapezoidal probability distributions are considered and their accuracy are evaluated. For symmetrical trapezoidal PDF of straight as well curved

sides, using the Monte-Carlo simulation method standard deviations (SD) of above estimators are evaluated. Established are broad range $0 < \beta < 0{,}75$ of upper and bottom bases ratios β of the most accurate 2C estimator as the linear form of mid-range and mean values of the sample for linear trapezoidal PDF Trap(a, b), and shorter for CTrap(a, b, d**).** The new simplified 2C-estimator of equal coefficients is also proposed and positively tested. Above estimators successfully extend existing estimation of the measurand value and its accuracy by the uncertainty type A recommended in the international Guide GUM [1]. Problems solved above are not described before in literature. They could be effectively applied in practice of measurements and in applied statistics for estimation of more accurate results.

Keywords: trapezoidal probability distributions, mean, midrange, two-element estimators, evaluation of measurand uncertainty

Chapter 9
Choosing the measurand estimator using Monte Carlo and resampling methods

Abstract. The main purpose of this work is the expansion of the ways for choosing the best estimator of empirical data, when this data is only available information. Proposed methods are divided into two types: Monte Carlo simulation and resampling methods: "*jackknife"* and "*bootstrap".* Numerical examples are also given.

Keywords: effective estimator, Monte Carlo modeling, resampling methods

Chapter 10
Polynomial estimator of measurand based on cumulants of non-Gaussian symmetrically distributed data

Abstract. In this paper the non-standard method for evaluating of the average non-Gaussian-distributed symmetrically sampling with a priori partial description (unknown PDF) is propose. This method of statistical estimation is based on the apparatus of stochastic polynomials and uses the higher-order statistics (moment & cumulant description) of random variables. The analytical expressions for finding estimates and analyze their accuracy to the degree of the polynomial $s = 3$ are given. It is shown that the uncertainty estimates for received polynomial is generally less than the uncertainty estimates obtained based on the mean (arithmetic average). Reduction factor depends on the MSE values of higher order cumulant coefficients, characterizing the degree of the sampling distribution differences from the Gaussian model. The results of statistical modelling based on the Monte Carlo method, confirmed the effectiveness of the proposed approach.

Keywords: uncertainty**,** estimator, non-Gaussian model, stochastic polynomial, means value, variance, cumulant coefficients

Chapter 11
Application of robust statistic to evaluation the accuracy of measurements

Abstract. Presented are two methods of assessing the value and uncertainty of the measurand from the sample of experimental data that are resist to contained therein a small number of outliers, i.e. values of measurement data significantly different from the others. This allows setting credible statistical parameters of the measurement result based on all experimental data. The considerations illustrate two numerical examples of interlaboratory measurements. Compared are results obtained by method applied the rescaled median absolute deviation MAD_S and by the iterative Huber method with winsoryzation. Given are conclusions and bibliography.

Keywords: outliers, uncertainty of measurements, standard deviation, median, robust mean value, interquartile mid-range

Chapter 12
Backgrounds of the evaluation of indirect multivariate measurement results

Abstract. A short review of some problems arising in the correct numerical expression and evaluation of results of indirect multi-parameter measurements is given. As a theoretical basis for determining the estimates of values, uncertainties and correlation coefficients of the indirectly obtained multi-measurand, processed from data of the simultaneously measured variables, the algebra of random vectors is used. A numerical example illustrates the linear transformation of two variables and the types of improperly evaluated results of that may occur with over-rounding. Proposed is an upgrading of the GUM and its Supplement 2 clauses about proper formula of the uncertainty propagation for nonlinear functions. As necessary complement to it was the urgently needed standardization of the publishing measurement data in dual form: so far – on paper and on the accompanying e-publication containing the results of the original measurements. Then user can ease to take and review them himself. Given is also a brief set of conclusions and bibliography.

Keywords: multivariate measurements, correlated data uncertainty evaluation, fundamental constancies, e-publishing

Appendix 1
Examples of calculation of the effective estimators for data samples with uniform and Laplace distributions

Appendix 2
The proposals for the determination of the expanded uncertainty in the new GUM

English Summary, titles of Chapters and their abstracts