

BIBLIOTEKA  
Instytutu Łączności  
Nr  
INSTYTUT ŁĄCZNOŚCI  
WARSZAWA-MIEDZESZYN

**BIULETYN**

**INFORMACYJNY**

1-2(223-224)

1984



MINISTERSTWO ŁĄCZNOŚCI

---

BIBLIOTEKA  
Instytutu Łączności

Nr \_\_\_\_\_

# BIULETYN INFORMACYJNY

ROK 24

WARSZAWA 1984

NR 1-2/223-224/

---

INSTYTUT ŁĄCZNOŚCI

Branżowy Ośrodek  
Informacji Naukowej, Technicznej i Ekonomicznej

Redakcja Biuletynu Informacyjnego

---

Redaktor Naczelny - prof. mgr inż. Lesław Kędziński  
Z-ca Redaktora Naczelnego - doc. dr inż. Krystyn Plewko

Redaktorzy działów:

doc. mgr inż. Władysław Cetner, doc. mgr inż. Adam Moniuszko

Adres Redakcji:

Instytut Łączności

Branżowy Ośrodek

Informacji Naukowej, Technicznej i Ekonomicznej

Warszawa - Miedzeszyn, ul. Szachowa 1

ISSN 0209-1046

NA PRAWACH RĘKOPISU - DO UŻYTKU SŁUŻBOWEGO

Redaktor: mgr K. Juszkiewicz

Montaż tekstu: B. Skwara

---

Dział Wydawniczy Instytutu Łączności  
Format B5. Nakład 570. Wpłynęło do  
Działu Wydawniczego 4. IX. 1984 r.  
Druk ukończono w październiku 1984 r.

OKREŚLENIA I OCENY NIEPEWNOŚCI POMIARU W ŚWIETLE  
PRAKTYKI WIELKICH OŚRODKÓW METROLOGICZNYCH

/Przekład z jęz. francuskiego dokonany przez inż. Leszka  
Chodakowskiego i prof. dr Jerzego Dudziwicza  
z Instytutu Łączności w Warszawie/

SPIS TREŚCI

	Str.
1. Wprowadzenie	1
2. Kwestionariusz BIPM w sprawie niepewności pomiaru	3
2.1. Wstęp	3
2.2. Zagadnienia szczegółowe	4
2.2.1. Odchylenie standardowe	4
2.2.2. Granice przedziału ufności	5
2.2.3. Natura błędów systematycznych	6
2.2.4. Prezentacja błędów systematycznych	7
2.2.5. Propagacja błędów systematycznych	7
2.2.6. Łączenie błędów systematycznych i przypadkowych	8
2.2.7. Postać niepewności wynikowych	9
2.2.8. Pozostałe zagadnienia	10
2.3. Informacje uzupełniające	10
3. Raport w sprawie kwestionariusza BIPM dotyczącego wyrażania niepewności pomiarów	10
3.1. Wstęp	10
3.2. Odpowiedzi na pytania BIPM	14
3.3. Dane bibliograficzne	23
3.4. Wnioski	27
4. Sprawozdanie Grupy Roboczej CIPM ds. wyrażania niepewności pomiaru	30
4.1. Wstęp	30
4.2. Ankieta BIPM w sprawie wyrażania niepewności pomiarów	32
4.3. Określenie niepewności za pomocą metod statystycznych	34

	Str.
4.4. Niepewności, które nie mogą być określane metodami statystycznymi	36
4.5. Niepewność łączna i niepewność globalna	41
4.6. Uwagi końcowe	43
4.7. Zalecenie Grupy Roboczej ds. Wyrażania Niepewności Pomiaru przedstawione Międzynarodowemu Komitetowi Miar /CIPM/	45
4.8. Spis dokumentów przedstawionych na posiedzeniu	46
5. Zalecenie 1 /CI-1981/ CIPM	47
Wykaz literatury	48
Dodatek 1. Odpowiedź Narodowego Biura Wzorców /NBS-USA/	51
1. Przedmowa	51
2. Odpowiedzi na pytania BIPM	51
3. Informacje dodatkowe	59
3.1. Przykłady	59
3.2. Dane bibliograficzne	59
Dodatek 2. Odpowiedź Narodowego Biura Metrologii /BNM - Francja/	60
Dodatek 3. Odpowiedź Fizyczno-Technicznego Instytutu Federalnego /PTB-RFN/	65
1. Wstęp	
2. Projekt słownika dotyczącego niepewności pomiaru	67
3. Odpowiedzi na pytania BIPM	69
Dodatek 4. Odpowiedź Narodowego Laboratorium Fizycznego /NPL - Wielka Brytania/	71
1. Przedmowa	71
2. Wstęp	72
3. Rozróżnienie między błędem i niepewnością	72
4. Klasyfikacja i definicja niepewności	73
5. Terminologia dotycząca niepewności	74
6. Odpowiedzi na pytania BIPM	75
7. Informacje uzupełniające	81

## OKREŚLENIA I OCENY NIEPEWNOŚCI POMIARU W ŚWIETLE PRAKTYKI WIELKICH OŚRODKÓW METROLOGICZNYCH <sup>X/</sup>

### 1. WPROWADZENIE <sup>XX/</sup>

Biorąc udział w V Kongresie IMEKO w maju 1970 roku w Paryżu /Wersalu/ miałem możliwość zwiedzić słynne Międzynarodowe Biuro Miar BIPM mieszczące się w Pavillon de Breuteuil w Parc de Saint-Cloud /Sèvres/ koło Paryża. Wykorzystując tę okazję wpisałem się na listę prenumeratorów publikacji BIPM, które otrzymuję bezpłatnie od 13 lat <sup>xxx/</sup>; są to następujące rodzaje dokumentów:

1. Comptes rendus des séances de la Conference Générale des Poids et Mesures /Sprawozdania z kolejnych posiedzeń Generalnej Konferencji Miar/, wydawane co 4 lata.
2. Procès - Verbaux des séances du Comité International des Poids et Mesures /Protokoły z kolejnych posiedzeń Międzynarodowego Komitetu Miar/, wydawane co roku.
3. Sessions des Comités Consultatifs /Raporty z kolejnych posiedzeń Komitetów Doradczych <sup>xxxx/</sup>/, wydawane nieokresowo /co 1+4 lat/.
4. Recueil de Travaux du BIPM /Przegląd prac BIPM/, wydawany co kilka lat.

---

<sup>x/</sup> Przekład z jęz. francuskiego [2,3] dokonany przez inż. Leszka Chodakowskiego i prof. dr Jerzego Dudziewicza z Instytutu Łączności w Warszawie.

<sup>xx/</sup> Opracował prof. dr Jerzy Dudziewicz.

<sup>xxx/</sup> Wszystkie otrzymane publikacje są przechowywane w aktach Centralnej Izby Pomiarów Telekomunikacyjnych Instytutu Łączności, 04-894 Warszawa, ul. Szachowa 1, tel. 128-409.

<sup>xxxx/</sup> Obecnie istnieje 8 Komitetów Doradczych, a mianowicie: Elektryczności, Fotometrii i Radiometrii, Termometrii, Definicji Metra, Definicji Sekundy, Etalonów Jednostek Promieniowania Jonizującego, Jednostek Miar oraz Masy i Wielkości Pokrewnych.

5. Le Système International d'Unités /SI/ /Międzynarodowy Układ Jednostek Miar SI/, broszura wydawana co kilka lat /w 1981 roku - czwarte wydanie/.

Publikacji powyższych nie ma w wolnej sprzedaży, zatem czujemy się w obowiązku udostępnić czytelnikom polskim zawarte w nich najcenniejsze prace i informacje. Do nich można niewątpliwie zaliczyć sprawozdanie zawarte w 48 /1980 r./ i 49 /1981 r./ tomie Protokołu z 69 i 70 posiedzeń Międzynarodowego Komitetu Miar i dotyczące próby ujednolicenia terminologii i metod stosowanych przy określaniu i ocenie niepewności pomiaru. Zawarto w nim pełny tekst kwestionariusza opracowanego przez BIPM /por. pkt. 2/ i raportu w sprawie wyników ankietyzacji /por. pkt. 3/. W zakończeniu przytoczono tekst Zalecenia 1 /CI-1981/ przyjętego przez CIPM w 1981 roku /por. pkt. 5 / wraz z obszernym komentarzem specjalnej grupy roboczej, która opracowała projekt tego zalecenia /por. pkt. 4/. Artykuł zaopatrzone spisem dodatkowych pozycji bibliograficznych /por. s. 48/ wiążących się ściśle z daną tematyką; uwzględniono w niej głównie wydawnictwa krajowe. W dodatkach przytoczono pełny tekst najciekawszych odpowiedzi czterech przodujących laboratoriów metrologicznych na świecie /NBS - Dodatek 1, BNM - Dodatek 2, PTB - Dodatek 3 i NPL - Dodatek 4/.

Całe zagadnienie uznano za tak ciekawe i ważne, że postanowiono udostępnić czytelnikowi polskiemu, zamieszczając wolny przekład tych dokumentów w niniejszej publikacji. Sądzimy, że poniższym tekstem zainteresujemy wielu metrologów w kraju, niezależnie od branży, w jakiej prowadzą prace teoretyczne i praktyczne związane z techniką mierzenia. Należy przy tym podkreślić, że strona polska była również jednym z 21 respondentów ww. kwestionariusza. Rozpowszechnienie tego materiału wydaje się w związku z tym w pełni uzasadnione.



## 2. KWESTIONARIUSZ BIPM W SPRAWIE NIEPEWNOŚCI POMIARÓW

### 2.1. Wstęp

W czasie sesji CIPM, w sierpniu 1977 roku, została utworzona grupa robocza, która miała za zadanie studiowanie problemów dotyczących oceny i wyrażania niepewności wyników pomiarów.

Problemy te wywoływały już sporo kontrowersji - na ogół bez rezultatu. Zasadnicze trudności są związane, jak się wydaje, z nieokreśloną naturą tego, co zwie się zazwyczaj "błędami systematycznymi". Niejasność pozostaje w dalszym ciągu, gdy chce się połączyć ww. błędy z błędami pochodzenia "losowego" /"przypadkowego"<sup>x/</sup>, na przykład gdy chce się określić granice tolerancji.

Wydaje się, że można by uzyskać jednomysłność, w kwestii dotyczącej pomiarów o bardzo małej niedokładności /bardzo wielkiej precyzji/, dla których należy podawać możliwie wiele informacji o sposobie estymowania niepewności. Natomiast bardzo często zachodzi potrzeba w praktyce scharakteryzowanie niepewności lub tolerancji za pomocą pojedynczego parametru /lub w ostateczności za pomocą bardzo małej liczby parametrów/. Stanowi to właśnie ten problem, co do którego istnieje nadzieja, że można będzie uzyskać pewien postęp.

<sup>x/</sup> Określenia "losowy" i "przypadkowy" będą w niniejszym artykule uważane za synonimy. Wyrażenia te mają następujące odpowiedniki obcojęzyczne: "aléatoire" - francuski, "random" - angielski, "zufällig" - niemiecki, "случайный" - rosyjski. W odniesieniu do określenia rodzaju błędu przyjęto na ogół w polskiej literaturze technicznej stosować wyrażenie "błąd przypadkowy" [9], a nie "błąd losowy". W probabilistyce, statystyce matematycznej oraz w teorii procesów losowych stosuje się wyłącznie określenia "losowy" /np. zmienna losowa, funkcja losowa, proces losowy itd./. Tłumacze nie przestrzegają rygorystycznie którejkolwiek z ww. konwencji. /przyp. tłum./.

Niektóre organizacje branżowe, jak również laboratoria narodowe zaproponowały rozmaite reguły: są one często sprzeczne. Wydaje się, że trzeba jak najszybciej określić sytuację obecną, próbować uwypuklić znaczenie uzyskania zgodności w zakresie międzynarodowym, w przeciwnym bowiem razie sytuacja stanie się wkrótce nie do rozwikłania.

Biorąc pod uwagę dużą liczbę wypowiedzi, często bardzo rozbieżnych, trzeba starannie wybrać tematy dyskusji dla ww. grupy roboczej. Zasadniczym celem kwestionariusza jest uwypuklenie istotnych problemów i punktów, co do których dyskusja mogłaby przynieść praktyczne rezultaty. W szczególności trzeba unikać dyskusji czysto filozoficznych, czy matematycznych, które mają mały wpływ na praktyczne zastosowanie oraz trzeba znaleźć kompromis pomiędzy różnymi tendencjami: nadmiernego rygoru lub zawodnej tolerancji.

## 2.2. Zagadnienia szczegółowe

Zanim zostaną przytoczone same pytania wydaje się użyteczne przypomnieć krótko ich kontekst, w tym celu aby sprecyzować tematy i unikać nieporozumień.

### 2.2.1. Odchylenie standardowe

Wśród różnych parametrów, pozwalających charakteryzować niepewność zmiennej losowej  $X$ , odchylenie standardowe /standard deviation - ang., écart-type - franc./  $\sigma_x$  zajmuje miejsce uprzywilejowane. Jest ono określone następującym wyrażeniem:

$$\sigma_x^2 = E \left\{ \left[ X - E(X) \right]^2 \right\}$$

gdzie  $E(X)$  oznacza wartość oczekiwaną zmiennej losowej  $X$   
a  $\sigma_x^2$  - wariancję.

Przyjęto się estymować błąd standardowy - przyjmując jako punkt wyjścia próbkę o licznosci  $n$  - odchyleniem standardowym w próbce  $s_x$  wynikającym z następującego wyrażenia:

$$s_x^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2, \quad \text{gdzie } \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

Podobnie odchylenie standardowe w próbce dla wartości średniej  $\bar{x}$  jest estymowane za pomocą następującego wyrażenia:

$$s_{\bar{x}} = s_x / \sqrt{n}$$

Przypominamy, że przy ocenie tych wielkości nie potrzeba zakładać jakiegos konkretnego rozkładu; natomiast wartości  $X$  muszą być wzajemnie niezależne i wariancja musi istnieć.

Pytanie 1. Czy należy zalecać stosowanie odchylenia standardowego jako parametru charakteryzującego niepewność losową?

### 2.2.2. Granice przedziału ufności

Aby podjąć decyzję statystyczną co do określonego poziomu ufności, należy znać odpowiednie wartości graniczne danej wielkości wyznaczające granice przedziału ufności /na przykład dla prawdopodobieństwa - poziomu ufności -  $p = 95\%$ . Określając go należy założyć dla danej populacji pewien rozkład prawdopodobieństwa /przyjmuje się najczęściej rozkład normalny/; zależy on również od licznosci  $n$  próbki wykorzystywanej do obliczenia  $s_{\bar{x}}$ . Granice przedziału ufności określa się za pomocą wyrażenia typu  $\bar{x} \pm s_{\bar{x}} t(n-1, \frac{1-p}{2})$ . Hipoteza o normalności rozkładu

x/ Wyraz  $t(n-1, \frac{1-p}{2})$  oznacza kwantyl rzędu  $\frac{1+p}{2}$  / rozkładu t Studenta o  $(n-1)$  stopniach swobody; zakłada się, że zmienna losowa  $X$  ma rozkład normalny /przyp. tłum./.

staje się mniej krytyczna dla wystarczająco dużych wartości  $n$  i wtedy  $t$  zależy praktycznie tylko od  $p$ .

Przypominamy, że dla granic przedziału ufności bezpośrednio dodawanie kwadratowe jest wtedy tylko usprawiedliwione, gdy spełnionych jest więcej warunków takich, jak: normalność populacji, taki sam poziom ufności, takie same stopnie swobody dla wszystkich składników, podczas gdy sumowanie wariancji nie zależy od tych warunków.

Pytanie 2. Czy należy zalecać stosowanie umownego prawdopodobieństwa i odpowiednich granic przedziału ufności /zamiast odchylenia standardowego/?

### 2.2.3. Natura błędów systematycznych

Ocenianie błędów zwanych "systematycznymi", ich wyrażenie i ich praktyczne wykorzystywanie było już tematem wielu wypowiedzi. Prawdę mówiąc, samo pojęcie błędu systematycznego prowokuje do dyskusji i mało jest prób precyzyjnych definicji. Tak więc wydaje się trudne sformułowanie reguł praktycznych dla wielkości, która może obejmować, niekiedy jednocześnie, wiele pojęć różnych i źle zdefiniowanych. W szczególności można by zastanowić się, czy tradycyjne rozróżnienie między błędami przypadkowymi a błędami systematycznymi już z samej natury nie usprawiedliwia ich rozdziału i - w konkretnym przypadku - oddzielnego traktowania w odniesieniu do reguł propagacji błędów.

Zdarza się na przykład często, że można estymować wartość średnią i odchylenie standardowe pewnej wielkości /porównaj przypadek stałych fizycznych/, ale pojawiają się wątpliwości w przypisywaniu niepewności charakteru losowego lub systematycznego.

Pytanie 3. Czy istnieje istotna różnica między błędami przypadkowymi a błędami systematycznymi?

Pytanie 3b. Czy należy zalecać praktyczną regułę pozwalającą określać charakter błędu?

#### 2.2.4. Prezentacja błędów systematycznych

Realistyczna estymacja błędu o charakterze systematycznym, którego nie należy oczywiście mylić ze znaną poprawką /co do której zakłada się, że została już uwzględniona/ jest zawsze sprawą delikatną. Najczęściej opiera się ona na niezbyt pewnych danych i wymaga od eksperymentatora wiele rozważań. Niezbędna jest gruntowna znajomość szczegółów technicznych i teoretycznych podstaw metodyki pomiaru. Mimo wszelkich ostrożności nawet najbardziej doświadczonemu fizykowi mogą ujść uwadze istotne zjawiska systematyczne.

Najprostszy sposób ochrony przeciw takim ewentualnościom polega na oszacowaniu niepewności systematycznych w sposób "liberalny", na przykład na podstawie błędów "maksymalnych". Ale taka praktyka ma co najmniej dwie niedogodności: stwarza ona poważne problemy przy propagacji błędów /stosując normalne wzory zakłada się, że znane są wariancje/ oraz stwarza ryzyko ukrycia obecności nierozpoznanych wpływów, których obserwacja pozwala jedynie na rzeczywisty postęp w metodach pomiarowych.

Stosowanie niepewności systematycznych, które możliwie najlepiej przybliżają odchylenie standardowe, pozwala ominąć te niedogodności, ale ich normalne stosowanie napotyka jeszcze bariery psychologiczne.

Pytanie 4. Czy należy zalecać praktyczną regułę wyrażania błędów systematycznych?

#### 2.2.5. Propagacja błędów systematycznych

Jeżeli pewna wielkość jest funkcją wielu zmiennych podlegających wpływowi niepewności o charakterze losowym, to prawo propagacji błędów pozwala łączyć różne składniki i otrzymywać stąd błąd wynikowy /wykorzystuje się przy tym odpowiednie wariancje i kowariancje zmiennych, jak również odpowiednie pochodne cząstkowe; zakładamy w dalszych rozważaniach, że prawo to jest znane/.

O ile propagację błędów przypadkowych można uważać jako dobrze ugruntowaną, o tyle w odniesieniu do łączenia błędów systematycznych powstają wciąż różne propozycje.

Przy zastosowaniu błędów maksymalnych zwykło się niekiedy opisywać błędy systematyczne za pomocą rozkładu prostokątnego /jednostajnego/ i wiązać z tym specjalne reguły obliczeń. Metody tego rodzaju wydają się mało przekonywujące. Tym samym identyfikuje się właściwie brak dokładnego rozeznania z jednostajnym rozkładem gęstości prawdopodobieństwa w ograniczonym zakresie, co już samo przez się nie wydaje się realne. Logiczne wnioskowanie prowadziłoby do zastosowania dla granic maksymalnych reguł dodawania liniowego, aby zachować właściwość maksimum; wtedy zbyt szerokie granice, które się stąd uzyskuje, okazują się mało użyteczne w praktyce. Sposób bardziej tolerancyjny, polegający na tym, aby zamiast dodawania liniowego stosować dodawanie kwadratowe /w zależności od liczby składników/, wydaje się dyskusyjny.

Reguła bardziej przekonywująca, która bez wątpienia wzorowałaby się na relacjach znanych dla błędów przypadkowych, pozostaje w dalszym ciągu sprawą otwartą.

Pytanie 5. Czy należy zalecać praktyczną regułę łączenia błędów systematycznych?

#### 2.2.6. Łączenie błędów systematycznych i przypadkowych

Jeżeli traktuje się rozróżnianie między dwoma typami błędów raczej jako zwyczaj językowy często praktykowany, ale bez większych konsekwencji, to ich łączenie nie stanowi żadnego rzeczywistego problemu pod warunkiem, że wszystkie one są wyrażone w takiej postaci, która pozwala estymować ich odchylenia standardowe. Przeciwnie, trudności pojawiają się wtedy, gdy uważa się, że są to błędy dwóch zasadniczo różniących się typów i że ich łączenie nie ma sensu fizycznego. W tym przypadku traktowanie oddzielne wydaje się konieczne dla każdego typu. Pojawiają się rów-

niez pewne problemy wtedy, gdy wyraża się błędy systematyczne jako "wartości maksymalne", w tym przypadku bowiem zachodzi konieczność przypisania im /przynajmniej domyślnie/ pewnego prawdopodobieństwa, co sprowadza się do poziomu ufności i odchylenia standardowego. Estymowanie błędów maksymalnych, staje się więc wybiegiem, którego należy unikać.

Można sobie oczywiście wyobrazić inne łączenie /przy współczynnikach mniej lub bardziej dowolnych, mieszane sumy o składnikach liniowych i kwadratowych itp./, ale takie procedury mają małe szanse na rozpowszechnienie.

Pytanie 6. Czy należy zalecać praktyczną regułę łączenia błędów przypadkowych i systematycznych?

#### 2.2.7. Postać niepewności wynikowych

Przyjmuje się ogólnie, że na przykład dla sporządzenia certyfikatu wzorcowania niepewności odpowiadające odchyleniu standardowemu, a więc przy poziomie ufności ok. 70%, byłyby zbyt małe, ponieważ prawie w jednej trzeciej przypadków "wartość prawdziwa" byłaby na zewnątrz przedziału 1-sigmowego. W tej sytuacji stosuje się na ogół prawdopodobieństwo rzędu 95%. Przejście na inny poziom ufności nie przedstawia większego problemu, jeżeli przyjmuje się jako punkt wyjścia estymację odpowiedniego odchylenia standardowego. Jednak w większości przypadków spotykanych w badaniach naukowych takie przejście nie narzuca się samo przez się. A może należałoby zalecić powszechne stosowanie pewnego poziomu ufności?. Taki sam problem nasuwa się dla składników w tym przypadku, gdy nie można ani nie chce się wyznaczać niepewności "całkowitej".

Zresztą dla pomiarów o małej niedokładności prosty podział różnych składników niepewności wynikowej na dwie klasy byłby bez wątpienia zbyt uproszczony i prowadziłby do nadmiernych strat informacji. W tym przypadku lepiej już zestawić szczegółową listę wszystkich przyczyn błędów

i ich estymatorów. Lista powinna zawierać wszystkie informacje potrzebne przy ewentualnych nowych ocenach, np. w ramach kompilacji danych.

Pytanie 7. Czy należy zalecać praktyczną regułę wyrażania niepewności wynikowej?

#### 2.2.8. Pozostałe zagadnienia

BIPM próbowało wyjaśnić pewne kwestie na podstawie informacji, którymi obecnie dysponuje. Jest możliwe, że pominięto przy tym pewne sprawy, które wydają się Wam istotne, albo że wybrano wątpliwą kolejność prezentacji.

Pytanie 8. Jakie zagadnienia wydają się Wam istotne? Czy poprzednie zagadnienia powinny być według Was rozpatrywane w innym porządku?

#### 2.3. Informacje uzupełniające

- Opisy praktyk powszechnie stosowanych w Waszym laboratorium albo inne sugerowane metody; jeżeli to możliwe, podajcie przykłady typowych zastosowań /certyfikat, bibliografia itp./ - najlepiej przedstawić trudniejsze przypadki.
- Literatura, którą możecie polecić.

### 3. RAPORT W SPRAWIE KWESTIONARIUSZA BIPM DOTYCZĄCEGO WYRAŻANIA NIEPEWNOŚCI POMIARÓW

#### 3.1. Wstęp

Zgodnie z sugestią p. E. Amblera /NBS/, członka Międzynarodowego Komitetu Miar /CIPM/, Międzynarodowe Biuro Miar /BIPM/ przygotowało kwestionariusz na temat niepewności pomiaru; został on zaadresowany do dużej liczby laboratoriów i instytucji, które - jak sądzono - są szczególnie zainteresowane tymi zagadnieniami. Proszono je o udzielenie odpowiedzi, które w pewnej mierze mogłyby być traktowane jako odzwierciedlenie opinii ich krajów /por. pkt. 2/.



Przypominamy, że główny cel tego kwestionariusza polegał na umożliwieniu BIPM dokonania wyboru specyficznych zagadnień, które mogłyby być następnie dyskutowane w sposób efektywny przez odpowiednią grupę roboczą. Dalesze szczegóły na ten temat można znaleźć w pkt. 2.1.

Biorąc pod uwagę wielką różnorodność punktów widzenia na ten temat, uznano za konieczne wybranie najbardziej pilnych zagadnień; odpowiedzi potwierdziły w zupełności słuszność takiego sposobu postępowania. Mimo iż nie proszono respondentów o zajęcie stanowiska w konkretnej sprawie, lecz tylko o wyrażenie opinii, czy ten problem powinien stanowić przedmiot dyskusji czy też nie, to jednak nie zaskoczył nas fakt, że wiele odpowiedzi zawierało również pewną liczbę argumentów lub stwierdzeń, które odzwierciedlają, przynajmniej częściowo, stan poglądów panujących w obecnej chwili wśród personelu laboratoriów wobec pewnych istotnych problemów w tej dziedzinie.

Celem tego rozdziału jest zestawienie i podsumowanie informacji dotyczących wyrażenia niepewności pomiaru, jakie wynikają z otrzymanych odpowiedzi. Powinniśmy przy tym być przygotowani na to, że w odniesieniu do zagadnień najtrudniejszych duża część odpowiedzi nie wykazuje wyraźnego stanowiska. Często ma się wrażenie, że ta "milcząca większość" nie wyraziła swojego własnego zdania, a raczej poszukuje dopiero wyjaśnienia i rady. Można więc mieć nadzieję, że dzięki rozpowszechnieniu wielu różnych wypowiedzi, mimo iż niekiedy są one nie do pogodzenia, uzyska się wiele dodatkowego pożytecznego materiału informacyjnego. Możliwe, że szersze i głębsze zrozumienie różnych aspektów przyczyni się do odkrycia nowych sposobów podejścia i ewentualnie rozwiązania aktualnych problemów. Opracowanie wytycznych, które mogłyby być stosowane i akceptowane przez większość użytkowników, nie będzie rzeczą łatwą, lecz niewątpliwie warto dokonać wielu prób dla osiągnięcia tego celu.

Tablica 1 przedstawia listę krajów i organizacji, które nadeszły odpowiedzi /do końca 1978 roku/; w dalszym ciągu zreferujemy je uwzględniając odpowiednią numerację laboratoriów /zaznaczoną w nawiasach/.

Tablica 1.

Kraje i laboratoria uczestniczące w ankietyzacji  
w sprawie wyrażania niepewności pomiaru  
/w kolejności alfabetycznej w języku angielskim/

Lp.	Kraj	Nazwa laboratorium
1	2	3
1.	Republika Południowej Afryki	National Physical Research Laboratory /Narodowe Badawcze Laboratorium Fizyczne/, Pretoria.
2.	Stany Zjednoczone AP.	National Bureau of Standards /Narodowe Biuro Wzorców/, Washington
3.	Australia	National Measurement Laboratory /Narodowe Laboratorium Miar/, Lindfield
4.	Brazylia	Odpowiedzi zebrane przez L. Cintra do Prado, São Paulo /członek CIPM/
5.	Kanada	Conseil National de Recherches /Narodowa Rada ds. Badań/, Ottawa
6.	Chińska Republika Ludowa	Narodowy Instytut Metrologii, Beijing
7.	Czechosłowacka Republika Socjalistyczna	Czechosłowacki Instytut Metrologii, Bratysława
8.	Dania	Technical University, Dept. of Physics /Politechnika, Wydział Fizyki/, Lyngby
9.	Francja	Bureau National de Métrologie /Narodowe Biuro Metrologii/, Paryż

1	2	3
10.	Niemiecka Republika Demokratyczna	Amt für Standardisierung, Messwesen und Warenprüfung /Urząd ds. Normalizacji, Metrologii i Badań Towarów/, Berlin
11.	Republika Federalna Niemiec	Physikalisch - Technische Bundesanstalt /Fizyczno-Techniczny Urząd Federalny/, Braunschweig
12.	Wielka Brytania	National Physical Laboratory /Narodowe Laboratorium Fizyczne/, Teddington
13.	Węgierska Republika Ludowa	Narodowe Biuro Miar, Budapeszt
14.	Włochy	Istituto di Metrologia G. Colonnetti /Instytut Metrologii im. G. Colonnetti'ego/, Turyn
15.	Japonia	National Research Laboratory of Metrology /Narodowe Laboratorium Badawcze Metrologii/, Tokio
16.	Korea Południowa	Korea Standards Research Institute /Koreański Badawczy Instytut Wzorców/, Seul
17.	Holandia	Van Swinden Laboratorium /Laboratorium im. Van Swindena/, Delft
18.	Polska Rzeczpospolita Ludowa	Résumé odpowiedzi udzielonych przez cztery instytucje
19.	Rumuńska Republika Socjalistyczna	Institutul National de Metrologie /Narodowy Instytut Metrologii/, Bukareszt
20.	Szwecja	Statens Provningsanstalt /Państwowy Urząd Badań/, Borås
21.	Związek Socjalistycznych Republik Radzieckich	Odpowiedź pochodząca prawdopodobnie od Gosstandardu, Moskwa /uzyskana za pośrednictwem OIML, Paryż/.

### 3.2. Odpowiedzi na pytania BIPM

Ograniczymy się tu tylko do przytoczenia samych pytań. Dodatkowe wskazówki na temat zagadnień figurujących w kwestionariuszu można znaleźć w pkt. 2.

W odniesieniu do pewnych laboratoriów, których odpowiedzi były szczególnie interesujące, przytoczono w całości nadesłane dokumenty w Dodatku 1, 2, 3 i 4. Dotyczy to:

- /2/ NBS, Washington - por. Dodatek 1,
- /9/ BNM, Paryż - por. Dodatek 2,
- /11/ PTB, Braunschweig - por. Dodatek 3,
- /12/ NPL, Teddington - por. Dodatek 4.

Należy zaznaczyć, że nie zawsze można było jednoznacznie zaklasyfikować odpowiedzi jako "tak" lub "nie", w szczególności gdy zaopatrywano je w znaczną liczbę komentarzy. Opierając się na pełnych odpowiedziach podanych poniżej, czytelnik może nabierać pewnego pojęcia o stopniu dowolności klasyfikacji, która mogła mieć miejsce. W sumie jednak nie powinno stąd wynikać żadne istotne zniekształcenie wyników.

Przejdziemy teraz do konkretnych pytań

- a/ Pytanie 1. Czy należy zalecać stosowanie odchylenia standardowego jako parametru charakteryzującego niepewność losową?

To pytanie wstępne było stosunkowo proste i odpowiedzi na nie można było podzielić następująco:

- "tak": /1/, /5/, /6/, /7/, /18/, /19/, /21/;
- "tak", lecz z żądaniem, aby była dodana liczba stopni swobody: /2/, /4/, /8/, /9/, /16/;
- wiele laboratoriów wahało się zaznaczać wyraźną preferencję lub pozostawiały wolny wybór: /10/, /11/, /12/, /13/, /14/, /17/, /20/.

Wreszcie /15/ żąda, "aby było wyraźnie zaznaczone, czy odchylenie standardowe dotyczy powtarzalności, replikowal-

ności lub odtwarzalności"; wyróżniono przy tym siedem rodzajów niepewności o charakterze losowym. /3/ odpowiedziała "tak - w odniesieniu do sprawozdań w czasopiśmie naukowych i nie - w odniesieniu do raportów i certyfikatów wzorcownic".

W sprawie dodatkowych komentarzy - por. /2/ /Dodatek 1/ i /12/ /Dodatek 4/.

b/ Pytanie 2. Czy należy zalecać stosowanie umownego prawdopodobieństwa i odpowiednich granic przedziału ufności /zamiast odchylenia standardowego/?

To zagadnienie jest, oczywiście, związane z poprzednim; wobec tego poniższe komentarze będą dotyczyły obu.

Wśród odpowiedzi nie ma żadnego "tak" bez zastrzeżeń. Poza tym następujące laboratoria wyrażają zdecydowane "nie": /1/, /2/, /5/, /8/, /9/, /16/, /18/, /21/.

Większość odpowiedzi można zakwalifikować jako dosyć niejasne lub "niezdecydowane". Zaliczamy do nich: /3/, /4/, /6/, /7/, /10/, /12/, /13/, /14/, /15/, /17/, /19/, /20/. Jednakże wyliczając je należy wspomnieć, że przyczyny wahań są bardzo różne. W wielu odpowiedziach wskazuje się, że sugerowane rozwiązania nie stanowią rzeczywistych alternatyw, lecz każde z nich znajduje swoją użyteczną dziedzinę zastosowania. /15/ zaleca, aby "proces pomiarowy był prowadzony w taki sposób, aby zmiany losowe wyników charakteryzowały się rozkładem normalnym". Wreszcie /7/ sugeruje stosowanie "granicy prawdziwego rozstępu błędów", ponieważ ta wielkość "jest skończona w porównaniu z nieskończonym rozrzutem rozkładu normalnego". Należy uwzględnić trafną uwagę /9/ /por. Dodatek 2/.

Komentarze BIPM w sprawie pytania 1 i 2: Ze zbioru odpowiedzi wydaje się wynikać wyraźna preferencja w kierunku odchylenia standardowego pod warunkiem, że towarzyszy temu liczba stopni swobody /lub liczba wykonanych pomiarów/. Granice przedziału ufności znalazły mało zwolenników, a w wielu odpowiedziach podkreśla się nie bez racji, że takie

granice przedziału nie zawierają więcej informacji niż estymowane odchylenie standardowe, lecz ponadto wymagają przyjęcia rozkładu normalnego. Ich zastosowanie powinno być więc ograniczone do przypadków, gdy zachodzi potrzeba podjęcia decyzji statystycznej.

c/ Pytanie 3a: Czy istnieje istotna różnica między błędami przypadkowymi a błędami systematycznymi?

To pytanie na pozór "niewinne" wywołało całą lawinę różnych odpowiedzi, jak należało się tego spodziewać. Wyraźną odpowiedź "nie" dały /1/, /3/, /14/, /19/. Odpowiedzi "tak" są liczniejsze, lecz różnią się znacznie od siebie, gdy zawierają dalsze wyjaśnienia. Odpowiedzi "tak" udzieliły /5/, /8/, /13/, /17/, /18/. Podczas gdy dla /21/ błąd systematyczny jest równoważny poprawce<sup>x/</sup>, to /16/ zauważa, że "błąd przypadkowy na pewnym poziomie hierarchii metrologicznej mógłby być traktowany jako błąd systematyczny na następnym poziomie tej hierarchii".

Jednak najbardziej znaczącą grupę tworzą te ośrodki, które nie dają wyraźnej odpowiedzi /lub dostatecznie prostej/ i do których można również dołączyć niezdecydowanych. Do takiej grupy można zaliczyć /2/, /4/, /6/, /7/, /9/, /10/, /11/, /12/, /15/, /20/. Wśród nich /6/ i /15/ zauważają wyraźnie, że "mogą wymiennie traktować te dwa typy", a /9/ dodaje, że "istnieje również wiele przypadków pośrednich". /12/ sugeruje nowy podział z punktu widzenia genezy /por. Dodatek 2/, podczas gdy /2/ twierdzi, że "istotna różnica .... staje się oczywista, gdy informacja o niepewności... jest wykorzystana w sposób efektywny" /por. Dodatek 1/. Wreszcie /20/ uważa, że "istnieje mniejsza lub większa strefa, gdzie nie można z dużą ufnością twierdzić, że niepewność pomiaru należałoby traktować jako systematyczną lub losową".

<sup>x/</sup> Chodzi tu zapewne o poprawkę "wziętą" ze znakiem ujemnym /przypis tłumacza/.

d/ Pytanie 3b. Czy należy zalecać praktyczną regułę pozwalającą określić charakter błędu?

Ta kwestia jest również powiązana z poprzednią, jak wynika z odpowiedzi. Odpowiedzi "nie" pochodzą od tych samych laboratoriów jak poprzednio. Odpowiedzi "tak" pochodzą od /7/, /12/, /13/, /17/, /18/, /21/. Laboratoria /2/, /4/, /5/, /6/, /8/, /9/, /10/, /11/, /15/, /16/, /20/ nie zajęły wyraźnego stanowiska. Należy wziąć pod uwagę w szczególności komentarze opracowane przez /2/ i przez /9/ /por. Dodatki 1 i 2/.

Komentarze BIPM w sprawie pytań 3a i 3b. Chcąc ocenić to zagadnienie na podstawie różnorodności nadesłanych odpowiedzi, można by uznać te pytania za bardzo ważne. Ponieważ jednak opinie są jeszcze dość rozbieżne, byłoby rzeczą bardzo niebezpieczną uważać, że dyskusja doprowadzi w najbliższym czasie do praktycznych wniosków. Wydaje się, że zachodzi jeszcze potrzeba dalszej analizy, tak w aspekcie podstawowym jak i praktycznym, zanim będzie można mieć nadzieję na znalezienie zadawalającego rozwiązania tego trudnego problemu. Taką analizę prowadzi się prawdopodobnie w wielu różnych laboratoriach, lecz zabierze ona jeszcze wiele czasu, ponieważ ten problem jest dla wielu eksperymentatorów zupełnie nowy. Poza tym możliwe jest, że takie rozwiązanie będzie niezgodne z ich dotychczasową praktyką.

a/ Pytanie 4: Czy należy zalecać praktyczną regułę wyrażania błędów systematycznych?

Ponieważ nie żąda się wyraźnego sformułowania ewentualnej "reguły praktycznej", to byłoby całkiem logiczne udzielić po prostu odpowiedzi "tak" lub "nie". "Tak" bez żadnych innych komentarzy odpowiedziały /3/, /11/, /13/, /21/, podczas gdy /7/ dodało, że należałoby sformułować różne zalecenia w zależności od rodzaju pomiaru, sposobu na podstawie którego uzyskuje się poprawki, uwzględniające efekty systematyczne i cel pomiaru. Można również zrozumieć /20/, które mówi: "niezwykle trudno jest sformułować regułę, co

do której można by mieć nadzieję, że będzie stosowana i interpretowana w jednolity sposób".

Tylko laboratorium /8/ odpowiada "nie" dodając, że "lepiej jest nie mieć żadnej reguły niż taką, która byłaby niejednoznaczna, źle określona, a nawet fałszywa".

W kilku odpowiedziach zaleca się koncepcję "maksymalnej granicy" /górnego kresu/, na przykład w /10/, /17/, a także w /14/, chociaż z zachowaniem ostrożności. Wreszcie /18/ optuje za czymś w rodzaju odchylenia standardowego.

/4/, /5/, /6/, /9/, /19/ nie dały żadnej wyraźnej odpowiedzi na zadane pytanie. /2/ i /16/ sugerują podział na dwie różne kategorie /dane eksperymentalne lub ocena indywidualna/, podczas gdy /12/ woli opisywać niepewność za pomocą wyrażenia subiektywnego /por. Dodatek 4/.

Komentarze BIPM: Napotykamy tu wyraźnie jeden z podstawowych problemów. Większość laboratoriów wolałaby rozporządzać pewną "regułą praktyczną", lecz dotychczasowe propozycje w tej sprawie są zupełnie rozbieżne. Granice przedziału ufności /dla różnych poziomów ufności/ i granice maksymalne, jak również odchylenia standardowe mają - każde z nich - swoich zwolenników. Zagadnienie wyboru danej postaci wiąże się, oczywiście, z zagadnieniami 3a i 3b i zawiera ono istotne implikacje praktyczne. Jest rzeczą oczywistą, że każde proponowane rozwiązanie powinno być również oceniane z punktu widzenia jego ewentualnej użyteczności dla późniejszego przetwarzania /obróbki/ wyników; zrobiono również pewną aluzję do niektórych problemów zawartych w pytaniach od 5 do 7.

f/ Pytanie 5: Czy należy zalecać praktyczną regułę łączenia błędów systematycznych?

Ponieważ może występować więcej niż jeden błąd systematyczny jest rzeczą oczywistą, że potrzebna jest pewna reguła ich łączenia. Tak więc nikt nie odpowiedział po prostu "nie". Jednakże większość laboratoriów zorientowała się, że pełna odpowiedź zależy od odpowiedzi udzielonej na pytanie 4



oraz że ogólne i zarazem proste rozwiązanie tego zagadnienia nie jest takie oczywiste.

/3/, /4/, /10/, /11/, /13/, /17/, /21/, dały odpowiedź "tak", na ogół bez żadnego komentarza. Odpowiedzi pozytywne, lecz bardziej sceptyczne lub z zastrzeżeniami udzieliły /5/, /8/, /9/, /14/, /20/. Szczególnie /2/ przeprowadziło dyskusję na temat różnych możliwych sposobów łączenia /liniowego, kwadratowego/, lecz nie zaleciło konkretnej reguły /por. Dodatek 1/. /6/ sugeruje proste dodawanie wariancji. Jednak większości odpowiedzi nie dało się łatwo sklasyfikować. I tak na przykład /16/ twierdzi, że "można by proponować wiele reguł praktycznych zamiast jednej", co jest nieco mniej pesymistyczne niż wypowiedź /7/, że "można by uwzględnić kilkadziesiąt różnych przypadków". Możliwe, że wszystko to jest prawdziwe, przynajmniej w pewnej mierze, lecz na pewno nie stanowi to tego typu reguły, której wszyscy poszukują.

/12/ opisuje pięć stosowanych powszechnie metod i dodaje: "Potrzebna jest przekonująca teoria ..., która mogłaby być punktem wyjścia dla metody łączenia niepewności systematycznych". Być może użyteczną ideę przedstawia /18/ twierdząc, że "nie trzeba rozróżniać reguły składania /dodawania/ niepewności systematycznych lub losowych i systematycznych, lecz stosować prawo propagacji odchylenia standardowego, jak to ma miejsce w statystyce matematycznej. Można by więc wyrażać wszystkie niepewności za pomocą odchylenia standardowego wartości średniej; nie należałoby więc poszukiwać żadnej innej terminologii dla niepewności pomiaru".

Komentarze BIPM: Jest oczywiste, że ten problem powinien być istotnym przedmiotem dyskusji oraz że należałoby powitać z zadowoleniem ustalenie jednolitej reguły ogólnej. Wydaje się, że liczba zwolenników liniowego składania maleje, a w szczególności nikt nie proponuje stosowania go do granicznych wartości maksymalnych. Ten wynik można więc uznać za dość zachęcający, chociaż nie osiągnięto jeszcze ostatecznego celu.

g/ Pytanie 6. Czy należy zalecać praktyczną regułę łączenia błędów przypadkowych i systematycznych?

W tym przypadku opinie są wyraźnie rozbieżne. Podczas gdy /8/ jest bardzo sceptyczne w tym względzie i sądzi, że "taka reguła będzie częściej bardziej myląca niż wyjaśniająca", to /13/ i /17/ są zdania, że to zależy od "poziomu metrologii". Istnieje również wyraźna opozycja. Tak więc /20/ mówi, że "nie powinny być one składane", a /16/ - że "na ogół powinno się unikać łączenia tych dwóch typów błędów".

/4/, /9/, /10/, /11/, /21/ są zwolennikami jakiejś reguły praktycznej, lecz niekiedy z zastrzeżeniami, podczas gdy /19/ sądzi, że "byłoby korzystne mieć niezbyt dużą liczbę reguł praktycznych".

Składanie kwadratowe proponują /3/ i /6/, ewentualnie również /18/. /1/, /2/, /5/, /7/, /12/, /14/ dają odpowiedź niezbyt jasną /lub nawet nie dają żadnej/.

Komentarze PIPM: Bardzo popularny pogląd panujący w dawnych latach, że należy unikać jakiegokolwiek łączenia błędów różnego rodzaju, ulegnie prawdopodobnie zmianie. Obecny stan, bardziej realistyczny, jest niewątpliwie inspirowany potrzebami praktycznymi. Podczas gdy bez wątplenia istnieje teraz większość, która chce przyjąć pewną regułę składania niepewności losowych i systematycznych, to na razie nie zarysowuje się żaden jednolity pogląd na temat sposobu, w jaki można by to konkretnie wdrożyć do praktyki.

h/ Pytanie 7: Czy należy zalecać praktyczną regułę wyrażania niepewności wynikowej?

I tu znowu wydaje się oczywiste, że to zagadnienie jest bezpośrednio związane z poprzednim. /9/ i /20/ wyrażają co do tego sceptycyzm lub sprzeciw, a /20/ twierdzi, że "bardzo trudno ... sformułować jednolitą regułę". /21/ jest zdania, że "należy zalecać wiele reguł ... w zależności od rodzaju zagadnienia". Podczas gdy laboratoria /2/, /3/, /4/, /7/, /10/, /11/, /14/ wyrażają pogląd, że korzystne byłoby mieć jakąś regułę tego rodzaju, to cztery inne laboratoria są zda-

nia, że taka reguła powinna zależeć od "poziomu metrologii"; chodzi tu o /12/, /13/, /16/ i /17/. Natomiast /1/, /5/, /6/, /18/ nie zajęły wyraźnego stanowiska.

Komentarz BIPM: Użyteczność pojęcia "niepewność wynikowa" będzie zależała - niewątpliwie - w dużej mierze od konkretnego zastosowania; być może należałoby zaproponować pewien podział całego problemu. Tak więc elementy lub składniki, które "wchodziłyby" do tak pojętej wielkości, będą to z konieczności te, o których była mowa wyżej, co stanowiłoby pewne ograniczenie<sup>x/</sup>. W swojej najbardziej ogólnej postaci zagadnienie zaczyna coraz silniej "zatręcać" o dziedzinę, co do której BIPM może lub powinien starać się dać propozycję pewnego rozwiązania. Jednak odpowiedzialność za certyfikat wzorcowania powinna spoczywać całkowicie na laboratorium, które wydaje takie świadectwo i które z kolei jest związane w wielu przypadkach przepisami legalizacyjnymi. Tak więc wskutek tego nie należy oczekiwać czegoś więcej niż ogólnych sugestii dotyczącej "dobrej i rozsądnej praktyki".

a/ Pytanie 8: Jakże inne zagadnienia wydają się Wam istotne? Czy poprzednie zagadnienia powinny być wg Was rozpatrywane w innym porządku?

Większość laboratoriów /osiem/ wyraża swoją satysfakcję lub nie daje żadnej odpowiedzi /pięć/, co można również interpretować jako zgodę.

Poza tym trzy laboratoria /7/, /17/, /18/ nalegają usilnie, aby przywiązywać duże znaczenie do ściśle określonej terminologii, która zależy prawdopodobnie również od "poziomu metrologii". Obezwerny "glosariusz terminów odnoszących się do niepewności pomiaru" został przedstawiiony przez /11/ /por. Dodatek 3/ i w tej samej odpowiedzi proponuje się,

---

<sup>x/</sup> Zdanie oryginalnego tekstu francuskiego nie jest zbyt jasne /przyp. tłum./

aby "przedyskutować dane, które powinny być zamieszczane w certyfikatach metrologicznych w różnych dziedzinach nauki i techniki". Tym niemniej można sobie zadać pytanie, czy dyskusja na tak szczegółowe tematy jest możliwa w łonie grupy roboczej, która jest kompetentna raczej w problemach ogólnych związanych z niepewnościami. Jest rzeczą nieuniknioną, że pewne studia zostaną przekazane instancjom bardziej wyspecjalizowanym.

Pewna liczba innych propozycji dotyczyła następujących punktów:

- /20/ uważa, że byłoby korzystne przeprowadzić dyskusję na temat "ewentualnego zastosowania zakresu<sup>x/</sup> zamiast rozrzutu wyników indywidualnych, w szczególności gdy liczebność ich jest mała", jak również na temat pewnych innych pojęć, takich jak "powtarzalność, odtwarzalność lub niedokładność".
- /14/ zadaje następujące pytanie: "Czy należy zalecać jakieś kryterium praktyczne, wg którego można by ustalać, czy dwa lub więcej wyników pomiarowych zaopatrzonych poza tym w niedokładność wynikową nie będzie wykazywało niezgodności między sobą?".
- /21/ usiłuje przypomnieć, że "jest rzeczą bardzo istotną określać i uwzględnić prawa rozkładu błędów składowych w ostatecznym oszacowaniu błędu pomiaru".
- /17/ podkreśla, że "należałoby znacznie bardziej spopularyzować wyniki porównań międzynarodowych przeprowadzonych pod auspicjami BIPM".

---

x/ Zastosowanie odpowiednika polskiego "zakres" czy też "dziedzina" lub "domena" zamiast wyrażenia francuskiego "domaine" nie umożliwia usunięcia wątpliwości co do istoty sprawy, o którą chodziło respondentowi szwedzkiemu /przyp. tłum./.

Wreszcie znaleziono w /8/ zaskakującą uwagę, wg której "w szczególności BIPM udało się ukryć swój własny punkt widzenia - gdyż zapewne posiada jakieś". W braku lepszej informacji przyjmujemy to jako komplement.

Komentarze BIPM: Gdy przygotowywaliśmy ten kwestionariusz nie uszły naszej uwadze problemy związane z terminologią. Są one całkiem realne i są bardzo ważne pod względem praktycznym, tym niemniej jednak nie stanowią istoty poruszanych tu zagadnień. Wtedy gdy pojęcia podstawowe są jeszcze kwestionowane, lepiej jest odsunąć zagadnienia słownictwa. Zresztą doświadczenie wskazuje, że ponieważ wszyscy są językoznawcami - dyskusje na temat terminologii bywają na ogół długie i mało efektywne. Dodatkowe problemy wynikające z barier językowych - znając dobrze się przyciągającą, z jaką takie dyskusje oddziałują na purystów - doradzają nam unikanie tych spraw, o ile to jest tylko możliwe, przynajmniej w pierwszym okresie. Wydaje się, że specjaliści mogliby się tym zająć w dalszej przyszłości.

### 3.3. Dane bibliograficzne

W odpowiedzi na naszą prośbę w sprawie informacji o publikacjach dotyczących wyrażenia niepewności pomiaru otrzymaliśmy pewną liczbę dokumentów, których spis podano poniżej w porządku chronologicznym; dołączyliśmy do tego pozycje, które były nam znane ekadınad.

Biorąc pod uwagę wielką liczbę dokumentów i artykułów, które opublikowano na temat zagadnień odnoszących się do wyrażenia na niepewność pomiaru, oraz ich niejednokrotnie dość wątpliwą wartość naukową, nie usiłowano zestawiać wyczerpującej listy. W większości cytowanych pozycji można znaleźć dalsze pozycje bibliograficzne.

Kilka pozycji, które ujmują zagadnienia ogólne z podobnego punktu widzenia jak nasz i które w związku z tym mogą być szczególnie interesujące dla czytelnika, zostały zaznaczone gwiazdką.

a/ Reguły lub wytyczne /"recepty"/

- \* - P.J. Campion, J.E. Burns, A. Williams: "A code of practice for the detailed statement of accuracy". /HMSO, London, 1973/, 52 strony.
- "Methods for treating the results of measurements". Proc. Institutes of Metrology USSR, No. 172 /232/, /Energia, Leningrad, 1975/, 72 strony.
- "Glosario de terminos empleados en metrologia", /Comité de Metrologia de la AECC, Madrid, 1976/, 35 stron.
- "The expression of uncertainty in electrical measurements". /No. 3003 British Calibration Service, 1977/, 16 stron.
- "The Australian Standards for the Measurement of Physical Quantities", National Measurement Laboratory /CSIRO, Melbourne, 1977/, 22 strony.
- O. Mathiesen: "How should the result of a measurement be reported? ", Svensk Mätplatskalender, 1977, 14 stron /po szwedzku/.
- "Expresión de los resultados de un calibración o medida /Comité de Metrologia de la AECC, Madrid, 1978/, 16 stron.
- "Metrology. Expression of result of measurement", Svensk standard SS 014150 /projekt, 1978/, 12 stron.
- "Measurement of fluid flow - Estimation of uncertainty of a flow-rate measurement", ISO 5168-1978 /E/, 26 stron.
- "Assessment of uncertainty in calibration and use of flow measurement devices", 79/31646 /British Standards Institution, London/, 78 stron.

- "Grundbegriffe der Messtechnik; Begriffe für Unsicherheit beim Messen und Fehler, Korrektion, Fehlergrenzen bei Messgeräten", DIN 1319, Część 3 /projekt/, /Deutsches Institut für Normung, Berlin, ca, 1979/, 17 stron.
- "Guidelines for estimation and statement of overall uncertainty in measurement results", CSC /80/ MS-9, National Physical Laboratory of India, Standards and Industrial Research Institute of Malaysia /Commonwealth Science Council, London, 1980/, 13 stron.

b/ Prace szczegółowe na temat niepewności pomiaru

- C. Eisenhart: "Realistic evaluation of the precision and accuracy of instrument calibration systems", J. Res. NBS 67C, s. 161-187 /1963/.
- H.O.W. Diettrich: "Über die Konsistenz von Messergebnissen", Rapport PTB VI C/N-4 /PTB, Braunschweig, 1965/, 19 stron.
- H.H.Ku: "Notes on the use of propagation of error formulas", J. Res. NBS 70C, s.263-273 /1966/.
- C. Eisenhart: "Expression of the uncertainties of final results", Science, 160, s. 1201-1204 /1968/.
- K. Weichselberger: "Ueber die Addition zufälliger und systematischer Fehler", Operations Research Verfahren 5, s. 423-444 /1968/.
- C. F. Diettrich: "Uncertainty, Calibration and Probability" /Hilger, London, 1973/, 411 stron.
- F. E. Grubbs: "Errors of measurement, precision, accuracy and the statistical comparison of measuring instruments", Technometrics 15, s. 53-66 /1973/.
- H. H. Ku: "Statistical methods applicable to counting experiments and evaluation of experimental data", Nucl. Instr. and Meth. 112, s. 377-383 /1973/.

- A. Williams: "Statement of results of experiments and their accuracy", Nucl. Instr. and Meth. 112, s. 373-376 /1973/.
- L. Gonella: "Proposal for a revision of the measure theory and terminology", Alta Frequenza, 44, s. 622-628 /1975/.
- G.C. Martarelli, A. Zanini: "Expression in unitary form of normal distributions", Alta Frequenza 44, s. 629-633 /1975/.
- W. Masin: "Ueber die Informationsfähigkeit eines Messgerätes", Metrologia 11, s. 169-177 /1975/.
- H. Reich: "Die Unsicherheit von Messungen mit Ionisationskammer - Dosimetern", PTB-Mitteilungen 86, s. 421-426 /1976/.
- G. Becker: "Ueber die Unsicherheit von primären Cäsiumstrahlnormalen für Zeit und Frequenz", PTB-Mitteilungen 87, s. 131-137 /1977/.
- P. J. Campion: "Problems in the statement of uncertainties" /NPL, Teddington, 1977/, 4 strony.
- F. G. Perey: "Covariance matrices of experimental data", Proc. Int. Conf. on Neutron Physics and Nuclear Data; Harwell 1978, s. 104-115.
- L. M. Schwartz: "Statistical uncertainties of analyses by calibration of counting measurements", Analytical Chemistry 50, s. 980-985 /1978/.
- W. Wöger: "Remarks on the confidence limits resulting from two models for the randomization of systematic uncertainties" /PTB Braunschweig, 1978; projekt/, 16 stron.
- \* - J. W. Müller: "Some second thoughts on error statements", Nucl. Instr. and Meth. 163, s. 241-251 /1979/.



- \* - S. R. Wagner: "On the quantitative characterization of the uncertainty of experimental results in metrology", PTB-Mitteilungen 89, s. 83-89 /1979/.
- Liu Chih Min: "A method for the combination of measuring errors" /NIM Beijing, bez daty/, 20 stron /po chińsku/.
- Hsiao Ming Yao: "The calculating methods frequently used for estimating experimental errors" /NIM Beijing, bez daty/, 23 strony /po chińsku/.
- Hsiao Ming Yao: "The precise calculation of the confidence factor in combination of errors" /NIM Beijing, bez daty/, 34 strony /po chińsku/.

### 3.4. Wnioski

W odniesieniu do większości problemów ujętych tym kwestionariuszem trudno jest lub w ogóle nie można wyciągnąć jednoznacznych wniosków na podstawie otrzymanych odpowiedzi. Nie stanowiło to zresztą naszego motywu rozsyłania kwestionariusza. Tym niemniej warto by powiedzieć kilka słów o aktualnej sytuacji ogólnej i o przejawiających się tendencjach. Ponieważ takie opinie ulegają ciągłej ewolucji, zachodzi konieczność zachowania rozeądku i umiaru.

Nie ma raczej wątpliwości, że zagadnienie ogólne dotyczące sposobu wyrażania niepewności pomiaru /eksperymentu/ wysu-  
nęło się na pierwszy plan w ciągu ostatnich lat. Ten nawrót zainteresowań jest uderzający dla tych, którzy pamiętają za-  
pewne sytuację panującą w owym czasie. - powiedzmy około dzie-  
sięciu lat temu - gdy wydawało się, że wszystko jest już u-  
porządkowane.

Rosnąca potrzeba coraz lepszej dokładności pomiarów w nauce i technice, jak również skuteczne sposoby przetwarzania danych za pomocą komputerów, którymi obecnie może dysponować każdy, uwypukliły wagę zagadnienia praktycznego dotyczącego wyrażania niepewności pomiaru. Obecnie większość

sposobów łączenia danych "grzeszy ciężko" swoim niespójnym i niepełnym charakterem informacji o różnych niepewnościach pomiaru lub nawet ich całkowitym brakiem.

Tradycyjny podział niepewności na składowe losowe i systematyczne, mimo iż często pożyteczny i łatwy w praktyce eksperymentalnej, stanowi w większości przypadków tylko ominięcie problemu, który i tak stanie przed eksperymentatorem, który rzadko będzie umiał wykorzystywać dokładnie te dane we własnych pracach.

Prowadzi nas to do pierwszego zagadnienia, zapewne do jednego z zagadnień podstawowych, a mianowicie do problemu typów błędu. Jeżeli dojdzie się do wniosku, że konieczne jest ich rozróżnianie, to natychmiast stajemy przed problemem ich dalszego wykorzystania.

Zagadnienia sformułowane w ankiecie i otrzymane odpowiedzi wskazują wyraźnie, że różne poruszone problemy są ze sobą ściśle powiązane. Wynika stąd, że żaden z nich nie może być w istocie rozwiązany bez uwzględnienia konsekwencji, jakie dany wybór będzie miał na pozostałe problemy. Narzucony jest również warunek konieczny, aby propozycje i wytyczne były użyteczne i łatwe w użyciu. Tak więc prosta reguła, która może być łatwo zrozumiana lub udowodniona, będzie na pewno lepsza od jakiegokolwiek skomplikowanego systemu.

Zupełnie niezależnie od problemów terminologicznych i klasyfikacyjnych nasuwa się kwestia bardziej praktyczna: jak należy wyrażać składowe niepewności całkowitej, którą się ocenia na ogół metodami mającymi raczej charakter oceny subiektywnej niż teorii statystycznej? I tutaj nie można tracić z pola widzenia dalszych implikacji; w szczególności wystąpi to w zagadnieniu łączenia ich z innymi składowymi niepewnościami. Nie traćmy jednak zbyt odważnie, gdy zdarzy się, że to czego trzeba unikać wyda nam się bardziej oczywiste niż to co należy wybrać. Pozostaje nadzieja, że ogólne i zdrowe zasady pomogą nam znaleźć rozwiązanie i przyczynić się do zmniejszenia nieporządku, który panuje dzisiaj w tych zagadnieniach.

Różnorodność otrzymanych odpowiedzi wskazuje wyraźnie, że nie osiągnięto jeszcze jednomyślnej opinii. Poza tym, jest zupełnie możliwe, że postawione zagadnienia wywołają u pewnych uczestników dalsze refleksje i nakłonią ich do weryfikacji pewnych koncepcji, które były dotąd uważane za ugruntowane. Rzeczywiście, wydaje się, że proces polegający na przemyśleniu pewnej liczby podstawowych problemów będzie przebiegał pomyślnie. Jest także możliwe, że najważniejszym rezultatem kwestionariusza będzie to, że okaże się, iż większość uczestników nie ma jeszcze definitywnego poglądu na wiele poruszonych tu problemów; są oni świadomi trudności i wydaje się, że oczekują rozsądnych propozycji. Taka sytuacja powinna być korzystna dla osiągnięcia porozumienia, które - jak sądzimy - mogłoby być akceptowane przez znaczną większość.

Dla pewnych eksperymentatorów /a może i dla wielu/ będzie to oznaczało zmianę zwyczajów. Jesteśmy świadomi tego, że jest to proces nie łatwy. Jest to jednak warunek nieunikniony, aby osiągnąć cel, którym jest jednolitość.

Kilka implikowanych problemów jest bardzo trudnych i mogą mieć poważne reperkusje. Pełna prezentacja różnych opinii ma więc wielką wagę i sądzimy, że różnorodność odpowiedzi podanych w tym raporcie może przyczynić się do wzajemnego zrozumienia. Pogodzenie rozbieżnych poglądów nie będzie rzeczą łatwą i szybką. Przeciwnie, przedsięwzięcie będzie niewątpliwie trudne i będzie wymagało wiele czasu, lecz perspektywa osiągnięcia w końcu upragnionego wyniku, który będzie bardziej jaśny, bardziej przydatny i poważniejszy akceptowany, dodaje temu usiłowaniu pełnej wartości rzeczywistego sukcesu.

#### 4. SPRAWOZDANIE GRUPY ROBOCZEJ CIPM DS. WYRAŻANIA NIEPEWNOŚCI POMIARU

##### 4.1. Wstęp

Grupa Robocza zebrała się w Międzynarodowym Biurze Miar w Sévres w dniach od 21 do 23 października 1980 roku dla przedyskutowania oceny niepewności wyników eksperymentalnych i uzgodnienia zaleceń praktycznych.

W posiedzeniach brali udział delegaci następujących laboratoriów:

- Biuro Narodowe Metrologii /BNM/: Narodowy Instytut Metrologii /INM/ Narodowej Wyższej Szkoły Sztuk i Rzemiosł, Paryż, Francja /P. Riéty/.
- Narodowy Instytut Metrologii /NIM/, Beijing, Chińska Republika Ludowa /Y. Jiang/.
- Instytut Metrologii im. G. Colonnetti'ego /IMGC/, Turyn, Włochy /S. Sartori/.
- Narodowe Biuro Wzorców /NBS/ Waszyngton, USA /R. Collé/.
- Narodowe Laboratorium Fizyczne /NPL/, Teddington, Wielka Brytania /A. Williams/.
- Narodowe Laboratorium Fizyczne Indi /NPLI/, New Delhi, Indie, /K. Chandra/.
- Narodowe Badawcze Laboratorium Fizyczne /NPRL/, Pretoria, Republika Południowej Afryki /R. Lake/.
- Federalny Urząd Metrologii /OFM/, Berno, Szwajcaria /P. Koch/.
- Fizyczno-techniczny Urząd Federalny /PTB/, Braunschweig, Republika Federalna Niemiec /F. Melchert/.
- Państwowy Urząd Badań /SP/, Borås, Szwecja /O. Mathiesen/.
- Laboratorium im. van Swindena /VSL/, Delft, Holandia /R. Kaarle/.

- Dyrektor Międzynarodowego Biura Miar /P. Giacomo/.

W posiedzeniu asystowali również: T. J. Quinn, zastępca dyrektora BIPM, A. Alliey, P. Carre i J. W. Müller /BIPM/.

Zebrań przewodniczył P. Giacomo, a sprawozdawcą został mianowany R. Kaarle.

Przewodniczący powitał wszystkich uczestników i życzył im powodzenia w obradach. Następnie przytoczył przyczyny, które skłoniły BIPM do zajęcia się problematyką wyrażania niepewności<sup>x/</sup> wyników eksperymentów. W sierpniu 1977 roku E. Ambler /NBS/, członek Międzynarodowego Komitetu Miar /CIPM/, zwrócił uwagę CIPM na zagadnienie dotyczące sposobu, za pomocą którego należy wyrażać niepewność związaną z wynikiem pomiaru. Wydawało się, że słusznie wybrano BIPM, jako instytucję, która powinna zająć się takim problemem, który - jak wiadomo - jest dość złożony, lecz ma duże znaczenie praktyczne. Rzeczywiście, problem nie jest nowy i pojawiał się dosyć często w pracach BIPM, zwłaszcza w toku dyskusji w łonie wielu Komitetów Doradczych. Ostatecznie CIPM powierzył BIPM troskę o szczegóły dotyczące tej problematyki.

Brak jednolitości w prezentowaniu niepewności stał się rzeczywiście poważnym problemem. W szczególności analiza porównań międzynarodowych okazała się niekiedy wręcz niemożliwa, ponieważ nie można było przypisywać określonego znaczenia podawanej niedokładności. Ta trudność jest dobrze znana każdemu, kto zajmuje się analizą wyników eksperymentalnych różnego pochodzenia. Istnieje również konieczna potrzeba ogólnej koncepcji, na której można by oprzeć metody obliczeniowe, na przykład w celu wyznaczenia granic niepewności lub

<sup>x/</sup> Słowo "niepewność" jest tu stosowane w jego ogólnym znaczeniu, które może się różnić od znaczenia węższego, które mu się niekiedy przypisuje w sposób umowny. Słowo "niepewność" wydaje się być tu znacznie lepsze niż słowo "błąd"; pewne powody tego podano w odpowiedzi NPL na kwestionariusz BIPM /por. Dodatek 4/.

tolerancji w metrologii prawnej lub przemysłowej. Sytuacja stała się w ostatnich latach coraz bardziej złożona z tego powodu, że różne jednostki organizacyjne wydawały sprzeczne zalecenia. Tak więc zachodzi pilna potrzeba opracowania przez BIPM zaleceń, które byłyby do przyjęcia dla wszystkich zanim jeszcze sytuacja nie stanie się zbyt uciążliwa.

#### 4.2. Ankieta BIPM w sprawie wyrażania niepewności pomiarów

Chcąc poznać różne stanowiska w tej sprawie zajmowane przez laboratoria narodowe BIPM postanowiło przygotować kwestionariusz obejmujący podstawowe zagadnienia w tej dziedzinie. Sposób, w jaki zostały sformułowane pewne podstawowe problemy, był prawdopodobnie nowy i nieoczekiwany dla pewnych osób, które otrzymały kwestionariusz; można było oczekiwać, że ankieta wzbudzi owocne dyskusje w łonie różnych naukowych jednostek organizacyjnych. Sądzono jednak, że konieczne było wyjaśnienie problemów podstawowych, zanim będzie można osiągnąć ewentualne uzgodnienie w skali międzynarodowej.

Ankieta zaadresowano do 32 laboratoriów narodowych, o których wiadano, że interesują się szczególnie tym zagadnieniem. Kwestionariusz został również przesłany do wiadomości pięciu organizacjom międzynarodowym. Przewodniczący przypomina, że J.W. Müller grał główną rolę w zestawianiu tekstu kwestionariusza i w analizie odpowiedzi, w związku z czym podziękował mu za jego wkład pracy. Przy końcu 1978 roku 21 laboratoriów nadesłało odpowiedzi. Dależe częściowa odpowiedź nadeszła dopiero kilka tygodni przed posiedzeniem [10] x/.

J. W. Müller przedstawił krótkie résumé dotyczące wyniku ankiety. Przypomniał, że jeden z głównych celów polegał na wyborze przedmiotów dyskusji dla grupy roboczej.

x/ Liczby w nawiasach kwadratowych dotyczą numerów pozycji w piśmie dokumentów zestawionego w p. 4.8.

Wśród zagadnień wątpliwych stawiano pytanie, czy niepewność losowa powinna być wyrażana przez odchylenie standardowe, czy też należy odwoływać się w związku z tym do granic przedziału ufności. Większość odpowiedzi wyraźnie faworyzowała pierwszą propozycję pod warunkiem, że będzie się zaznaczać dodatkowo liczbę stopni swobody. Przejście do odpowiedniego przedziału ufności polega na wyborze dowolnego poziomu ufności /prawdopodobieństwa/ i w praktyce wymaga przyjęcia /najczęściej "milczącego"/ populacji o rozkładzie normalnym. Dlatego też korzystniej jest stosować koncepcję odchylenia standardowego, ponieważ nie zależy to od tych dodatkowych hipotez.

Dalsze zagadnienia dotyczyły konieczności rozróżniania wielu typów niepewności i najlepszego sposobu ich scharakteryzowania. Zagadnienia te były trudniejsze. W konsekwencji odpowiedzi na te pytania wykazywały dużą rozbieżność. Jeżeli chodzi o niepewności "systematyczne", to wybór między wartościami granicznymi /maksymalnymi/, przedziałem ufności lub odchyleniami standardowymi nastroczał wiele wątpliwości. Większość laboratoriów uczestniczących w ankietyzacji uważała za istotne, aby ustalić regułę uzgodnioną w skali międzynarodowej dla łączenia /składania/ niepewności zwanych "losowymi" i "systematycznymi" w jedną niepewność "globalną". Ogólnie mówiąc, w toku analizy odpowiedzi na różne pytania można było mieć wrażenie, że istniała bez wątpienia jak gdyby "milcząca większość", która oczekiwała wyraźnej dyrektywy w tych tak kontrowersyjnych sprawach dotyczących niepewności eksperymentalnych /pomiaru/. Raport BIPM-80/3 /por. pkt. 3/ zawiera résumé wszystkich odpowiedzi jak również szczegółowe odpowiedzi podane przez NBS /por. Dodatek 1/, BNM /por. Dodatek 2/, PTB /por. Dodatek 3/ i NPL /por. Dodatek 4/.

Przewodniczący zwrócił uwagę, że głównym celem posiedzenia jest określenie pewnych jasnych i prostych reguł postępowania dla wyrażania niepewności. Te reguły powinny być oparte na przekonujących argumentach. Powinny być one

stosowane w sposób ogólny i akceptowane przez znaczną większość użytkowników. Pewne reguły arbitralne, jeżeli są one rzeczywiście niezbędne, powinny być ograniczone do minimum. Byłoby również bardzo pożądane, aby znaleźć przepisy, które mogłyby być stosowane na wszystkich poziomach metrologii. Może się to okazać dość trudne, ponieważ rzeczywiste potrzeby są bardzo różne. Natomiast można przypuszczać, że w większości przypadków podstawowe zasady wyznaczania niepewności są w istocie takie same. Trzeba więc, aby dykusje naświetliły cechy wspólne, a nie dotyczyły szczegółów, mimo iż te ostatnie mogą mieć także duże znaczenie praktyczne. Jest oczywiste, że cały wysiłek polegający na tym, aby opisać szereg pomiarów za pomocą małej liczby parametrów, implikuje bezwzględnie pewną stratę informacji. Jeżeli uważa się to za niedopuszczalne, to należy zachować całą serię danych, lecz w większości przypadków nie jest to ani możliwe, ani pożądane. Statystyka skłania do przyjęcia pewnego rozwiązania, a mianowicie do stosowania momentów, a zwłaszcza momentów niższego rzędu. Stąd wynika właśnie tak rozpowszechniony zwyczaj polegający na charakteryzowaniu szeregu danych eksperymentalnych za pomocą takich parametrów, jak: wartości średnie, wariancje i kowariancje. Sprawa polega właściwie na tym, aby wiedzieć, czy taka procedura matematyczna przedstawia użyteczny model w normalnych sytuacjach.

#### 4.3. Określenie niepewności za pomocą metod statystycznych

W przypadku niepewności typu losowego, którą można wyznaczyć za pomocą wielu powtarzanych pomiarów, grupa robocza jednogłośnie stwierdza, że należy je wyrażać za pomocą estymatorów wariancji  $/s_i^2/$  lub estymatorów "odchyień standardowych"  $/s_i/$ , podając zawsze odpowiednią liczbę stopni swobody  $\Psi_1$ . Jeżeli sądzi się, że te pomiary są w pewien sposób "powiązane" /skorelowane/, to należy również określić i wy-



znaczać odpowiednie kowariancje, zakładając że waga problemu usprawiedliwia dodatkowy nakład pracy.

Powstaje wątpliwość co do roli przedziałów ufności i ich związków z odchyleniami standardowymi. Jeden z uczestników uważa, że w przypadku dokładności najwyższej klasy nie ma rzeczywistej konieczności precyzowania poziomu ufności. Przy znanej funkcji rozkładu można zawsze określić odpowiedni przedział. Jeżeli można uważać, że populacja ma rozkład normalny, co często się czyni, to odpowiedniość jest dobrze znana i nie potrzeba jej przytaczać. Jednakże informacja tego typu może się okazać konieczna w tym przypadku, gdy na przykład wyznaczona niepewność powinna być porównana z granicą tolerancji narzuconą przez przepisy albo z tolerancjami pomiarowymi zadanymi z góry lub też ze specyfikacjami produkcyjnymi.

F. Melchert zwrócił uwagę, że projekt nowej normy zachodnio-niemieckiej DIN 1319 /część trzecia/ jest oparty na wykorzystaniu rozkładu /tablic/ t Studenta, a więc zakłada, że można stosować bardzo powszechnie rozkład normalny. Nawet gdyby można było krytykować ten projekt z punktu widzenia ściśle naukowego, stanowi on pożyteczną wskazówkę w zakresie metrologii prawnej.

W rezultacie grupa robocza zaleca wyrażać ten typ składowej niepewności za pomocą odchylenia standardowego estymowanego za pomocą wyrażenia  $s_1$  /odchylenia standardowego w próbie/ i podawać odpowiednią liczbę stopni swobody  $V_1$ .

W punkcie 4.4 będzie przeprowadzona krótka analiza na temat ewentualnej konieczności "kojarzenia" wyrażenia na niepewność z danym poziomem ufności, co ma często miejsce w metrologii praktycznej lub prawnej, jak również na temat możliwych sposobów jego realizacji.

#### 4.4. Niepewności, które nie mogą być określane metodami statystycznymi

Jeżeli można wyznaczyć niepewności ekperymentalne /z próbki/ przyjmując jako punkt wyjścia szereg pomiarów, jak to pokazano w poprzednim punkcie, to nie ma żadnego problemu, ponieważ można stosować przy ich wyznaczaniu dobrze znane metody. Jednakże istnieje wiele ważnych przypadków, gdzie jest to niemożliwe. Skłania nas to - jak sugerował na przykład R. Collé - do rozróżnienia między dwiema następującymi dużymi kategoriami składników niepewności:

- kategoria A: wyznaczenie można oprzeć na metodach statystycznych /obiektywnych/;
- kategoria B: estymacja powinna nawiązywać do "innych metod" ; implikuje to nieuchronnie istnienie pewnego elementu oceny subiektywnej.

Tradycyjnego rozróżniania między niepewnościami /lub "błędami", jak to często nazywano jeszcze niedawno/ "losowymi" a "systematycznymi" unika się tutaj rozmyślnie, głównie z tego powodu, że składniki "systematyczne" były w przeszłości definiowane na bardzo różne sposoby /często nie dające się pogodzić ze sobą/. Grupa robocza zaleca nie stosować tego terminu, ponieważ jego sens jest często niejednoznaczny, niepewny i źle interpretowany. Jeżeli "kategoria A" odpowiada niewątpliwie w większości przypadków temu, co dotychczas kwalifikowano jako "losowe", to "kategoria B" może być bardzo daleka od tego, co rozumiało się ogólnie jako "systematyczne". Aby wyjaśnić różnicę podejścia można by chyba powiedzieć, że podczas gdy klasyfikacja tradycyjna była oparta na skutku, jaki wywiera dana składowa niepewność na wynik końcowy, to nowe kategorie opierają się na sposobie, w jaki można uzyskać niepewność. Ponieważ wykorzystanie wyniku ekperymentu może być różnorodne, może się więc zdarzyć /i zdarza się w istocie/, że niepewność, którą uważano na początku za losową, staje się systematyczna i odwrotnie; sprawia to

wiele trudności w praktyce. Takie przemiany nie tylko stanowią problemy przy budowie jednoznacznej klasyfikacji, ale poza tym, utrudniają określenie niepewności "globalnej" lub "wynikowej", gdy trzeba uwzględnić składniki obu typów dla danego wyniku i gdy zakłada się, że te dwa typy podlegają różnej obróbce matematycznej. A. Williams sygnalizuje poważne trudności, jakie napotyka się, chcąc postępować logicznie przy łączeniu niepewności "losowych" i "systematycznych", gdy składowe "systematyczne" zostały określone w sposób różniący się od sposobu stosowanego przy składowych "losowych". Nowa klasyfikacja jest niezależna od dalszego wykorzystywania wyników.

Po przeprowadzeniu dyskusji grupa robocza doszła do wniosku, że najlepszy sposób uwzględnienia naszej niedoskonałej znajomości wartości liczbowej, jaką należy przyporządkować składowej zwanej "systematyczną", polega na traktowaniu jej zgodnie z modelem probabilistycznym. Gdybyśmy rozporządzali dostateczną ilością czasu /lub pieniędzy/ i innymi metodami /lub różnymi przyrządami/, nasza znajomość mogłaby być pogłębiona i rozrzut zastosowany w modelu byłby na ogół mniejszy.

W toku dyskusji ustalono ściśle, że w żadnym przypadku pojęcie niepewności pomiaru /niezależnie od kategorii, do jakiej ona należy/ nie należy mieszać z pojęciem poprawki; przyjmowano w dyskusjach, że uwzględniono już wszystkie znane poprawki w takiej mierze, w jakiej to było tylko możliwe dla eksperymentatora. Tym niemniej jest możliwe, że te poprawki nie były dostatecznie dobrze znane. W tym przypadku należy rozważyć wynikające stąd niepewności dodatkowe. Przyjmuje się, że mają one charakter losowy i mogą należeć do kategorii A lub B.

Ogólnie biorąc, niepewności kategorii A będą lepiej znane niż niepewności kategorii B, lecz może się też zdarzyć, że tak nie będzie. Również w większości przypadków niepewności kategorii A nie będą zupełnie zależały od osoby, która będzie dokonywała analizy wyników, lecz nawet tutaj obiekty-

wność nie będzie absolutna, gdyż będziemy niejednokrotnie dysponowali wieloma możliwymi metodami statystycznymi przy analizie wyników, a wybór metody będzie oczywiście sprawą subiektywną. Wynika stąd, że rozróżnienie między kategoriami A i B nie musi być wyraźne i jest możliwe, że zajdzie przypadek, co do którego będzie istniała wątpliwość w sprawie klasyfikacji. Zobaczymy dalej, że nie ma to istotnej konsekwencji. Rzeczywiście, można uważać, że rozróżnienie między dwiema kategoriami A i B jest raczej natury praktycznej niż "podstawowej". Powinno ono pomóc osobie analizującej wyniki przy uświadomieniu sobie różnych problemów, z którymi spotyka się ona przy wyborze wartości liczbowej charakteryzującej niepewność. Ten podział może również mieć pewne uświadomienie natury pedagogicznej.

Przewodniczący stawia z kolei pytanie, czy można charakteryzować składniki kategorii A i B w analogiczny sposób. Wydaje się, że sytuacja jest następująca: skoro można słusznie traktować niepewności kategorii A, które będą dalej oznaczone przez  $s_1$ , jako estymatory odchylenia standardowego, w sensie statystycznym, to nie będzie to prawdziwe dla niepewności kategorii B, które będą dalej oznaczone symbolem  $u_j$ . Mimo to w trosce o prostotę - P. Koch sugeruje, aby stosować "bez dyskryminacji" terminy "wariancja" jednocześnie do  $s_1^2$  i  $u_j^2$ , jak również "odchylenie standardowe" do  $s_1$  i  $u_j$ . Jednakże R. Collé nalegał usilnie, aby nie nazywać odchyleniami standardowymi wielkości  $u_j$ , ponieważ statystycy nie zaakceptują tego. Po dyskusji grupa robocza przyjęła całkowicie ten punkt widzenia, chcąc uniknąć jakichkolwiek niezgodności z dobrze ugruntowanymi pojęciami. Można nawet podnieść zarzut wobec słowa "estymować" przy wyznaczaniu wartości  $u_j$ , ale stosuje się je tu tylko z braku bardziej odpowiedniego terminu. Bliższe temu, co by się chciało wyrazić, byłyby prawdopodobnie wyrażenia amerykańskie "estimate" lub francuskie "apprécier".

Z całą pewnością należy przede wszystkim udzielić odpowiedzi na następujące istotne pytanie: Co chcemy naprawdę wyrazić przez wielkości  $u_j^2$ ? W rzeczywistości jest oczywiste, że istnieje wiele możliwości, lecz pewne z nich mogą się okazać bardziej przydatne niż pozostałe. Warto przypomnieć tu, że ten sam problem nasunął się w przeszłości, gdy należało scharakteryzować "niepewność systematyczną" i że proponowano wtedy rozmaite rozwiązania. Propozycja dość rozpowszechniona polegała na wyborze tego, co na ogół nazywano "maksymalnymi granicami". Zakładano, że te granice obejmują, według wszelkiego prawdopodobieństwa, całość obszaru, w którym mogła się znajdować wartość "prawdziwa" /nieznana/. Taki wybór mógł się wydawać na pierwszy rzut oka bardzo nęcący, ponieważ łatwo go było dokonać. Tym niemniej jednak przedstawia on sobą dwie poważne niedogodności: wielkość ta jest źle określona i w praktyce nie można jej dalej stosować w połączeniu z innymi niepewnościami /niezależnie od ich kategorii lub typu/. Rzeczywiście, nie ma w ogóle nawet prostej reguły w odniesieniu do propagacji granic, gdy daje się dwie wielkości, których indywidualne granice są znane.

Jedynym dobrym rozwiązaniem tego problemu - wydaje się - polega na tym, aby postępować zgodnie z implikacjami dobrze znanego ogólnego prawa dotyczącego "propagacji błędów". Istotne wielkości, które figurują w tej zależności, to wariancje /i kowariancje/ odpowiednich zmiennych /pomiarów/. Wskazuje to więc na to, że gdy poszukujemy "użytecznych" miar niepewności, do których można by zastosować zwykły formalizm, to powinniśmy wybrać coś, co można by uważać jako najlepsze "dysponowalne" przybliżenie odpowiednich "odchyłań standardowych".

Po pewnych dyskusjach Grupa Robocza była skłonna do przyjęcia zgodnego wniosku, że można zapatrywać się tak, jak gdyby  $u_j^2$  stanowiły wariancje. To doprowadziło do sformułowania deklaracji zawartej w paragrafie 3 Zalecenia /por. pkt. 4.7/. Uczestnicy są świadomi tego, że wdrożenie do

praktyki tego nieco enigmatycznego sformułowania będzie często bardzo trudnym zadaniem. Zalecenie ogranicza się więc do wskazania celu, który należałoby osiągnąć, nie podając explicite sposobu, za pomocą którego można by do niego dojść.

Stanowi to oczywiście kapitalny punkt w tym nowym sposobie podejścia i może oczekiwać różnych trudności w rzeczywistych warunkach. Z braku czasu nie przedyskutowano szczególnie sugestii mających na celu przezwycięzenie tych trudności; do tego istotnego punktu należy jeszcze powrócić w przyszłości. Przedstawiając bardzo schematycznie można dojść do ewentualnych oszacowań wielkości  $u_j$  w różny sposób, jak na przykład:

- Można próbować oszacować bezpośrednio  $u_j$ , na przykład, za pomocą przedziału ufności odpowiadającego poziomowi ufności od 2/3 do 70%.
- Można próbować oszacować przedział ufności przy poziomie ufności 50%, a następnie przyjąć  $u_j$  jako równe około 1,5 razy ten przedział.
- Obierając jako punkt wyjścia jakiś "pewny" przedział ufności można próbować przyporządkować mu albo dany poziom ufności /na przykład 95% lub 99%/, albo bezpośrednio daną wielokrotność odpowiedniego "odchylenia standardowego".

Postępując zgodnie z tą metodą zachodzi konieczność przyjęcia /choćby domyślnie/ pewnych założeń w sprawie danej populacji. Pozostawia się to do osobistej decyzji eksperymentatora, np. wybór rozkładu gaussowskiego lub jednostajnego. Zgodnie z regułą ogólną należy dążyć do realistycznych ocen granic niepewności, odpowiadających danemu poziomowi prawdopodobieństwa; trzeba zwłaszcza unikać sytuacji nadmiernie "rozważnych".

Jest rzeczą oczywistą, że podany wyżej podział możliwych metod mających na celu wyznaczenie  $u_j$  jest zupełnie arbi-

tralny. W istocie rzeczy te metody są równoważne i nie ma konieczności rozróżniania ich. Powinno być również oczywiste, że w większości przypadków wartość liczbowa wielkości  $u_j$  uzyskana w ten sposób nie będzie zbyt dokładna; niekiedy wszystko, czego można oczekiwać, to uzyskanie oceny rzędu wielkości. W każdym razie nie powinno to nas zbyt abсорbować, ponieważ jest to typowa właściwość oszacowania czegoś, czego rzeczywiście nie znamy. W szczególności nie powinno to nam przeszkadzać w oszacowywaniu /estymacji/, lecz będzie dobrym zwyczajem przypominać sobie zawsze, że każdy wniosek wyciągnięty w takiej sytuacji nie może być, oczywiście, bardziej wiarygodny niż wielkości, na których się on opiera.

#### 4.5. Niepewność łączna i niepewność globalna

Gdy już osiągnięto porozumienie na temat faktu, że można uważać składniki obu kategorii A i B tak jak gdyby miały one naturę losową, Grupa Robocza nie widzi żadnej przeszkody, aby nie traktować ich w ten sam sposób i w szczególności, aby nie łączyć w konkretnym przypadku. Mimo iż nie zapominamy, że wyrazy  $u_j^2$  kategorii B nie są w rzeczywistości wariancjami, panuje jednak ogólna opinia, że najlepsze co można uczynić, to zastosować do nich - przez analogię - ogólne prawo "propagacji błędów". W przypadku braku korelacji prowadzi to więc do niepewności łącznej, której kwadrat uzyskuje się dodając wszystkie składniki wynikające z  $s_j^2$  lub  $u_j^2$ . Z podanego powyżej powodu nie może być ona nazywana wariancją. Grupa Robocza ma nadzieję, że to proste rozumowanie mogłoby doprowadzić do poniesienia praktyki polegającej na dodawaniu w sposób liniowy pewnych składników przy wyznaczaniu łącznej niepewności.

Ponieważ przyjmując to jako punkt wyjścia składniki należące bądź do kategorii A, bądź do kategorii B, są traktowane na równi, jest jasne, że wszystkie problemy związane z dokładną naturą tego podziału mają mniejsze znaczenie i

rzeczywiście nie mają żadnego wpływu na wartość liczbową łącznej niepewności.

Wielu uczestników zwróciło uwagę, że trudno będzie pomyśleć, iż taką niepewność łączną będzie można stosować w pewnych istotnych przypadkach praktycznych, na przykład w certyfikatach. Poza tym przepisy prawne i specyfikacje produkcyjne mogą być oparte na danym przedziale ufności. Ten problem daje okazję do pewnych dyskusji, które można zreasumować następująco. O ile dla składników kategorii A związek z poziomem ufności opiera się na ogół na dodatkowej hipotezie /rzadko weryfikowalnej/, że można je uważać jako próbkę stanowiącą część populacji o rozkładzie normalnym, o tyle wydaje się, że trudniej będzie usprawiedliwić przyjęcie takiej hipotezy w odniesieniu do składników kategorii B. Należy jednak podkreślić, że ta trudność nie wynika w żadnym razie z wprowadzenia nowego podziału, lecz stąd, że została ona po prostu wydobyta na światło dzienne. Rzeczywiście, dotychczas można było pokrywać naszą niewiedzę dość mętnym pojęciem "maksymalna granica błędu" lub jakimś pojęciem równoważnym, którego związek z prawdopodobieństwem było chyba tak samo trudno ustalić, jak to ma miejsce dzisiaj. Ponieważ - w braku jakiegokolwiek niezawodnej kontroli - byłoby naiwnością sądzić, że natura funkcjonuje po prostu w zgodzie z życzeniem człowieka, Grupa Robocza będąc w pełni świadoma tej trudnej sytuacji proponuje przyjęcie "pozycji otwartej". Będą mogły zachodzić przypadki - w szczególności gdy składniki kategorii B mają małe znaczenie - gdzie przyjęcie hipotezy o normalności populacji będzie stanowiło rozsądne przybliżenie. To jest również prawdziwe, gdy można powołać się z ufnością na centralne twierdzenie graniczne. Na ogół jednak odpowiedzialna funkcja rozkładu nie jest na tyle dobrze znana, aby móc wyznaczyć przedziały ufności.

W świetle tej sytuacji wydaje się, że należałoby usprawiedliwić przyjęcie nieco innej pozycji, bardziej realistycznej. Zawsze gdy potrzeba większego przedziału niepewności



można uzyskać "niepewność globalną" mnożąc ww. niepewność łączną przez pewien stały czynnik  $k$ . Ponieważ nie można go dokładnie wyznaczyć, nie trzeba więc obierać wartości "dokładnych"; wydaje się, że w większości przypadków wystarczy przyjmować takie wartości konwencjonalne, jak  $k = 2$  lub  $k = 3$ . Można by stworzyć sobie pewną koncepcję wyboru przybliżonej wartości zakładając, na przykład, że można rzeczywiście identyfikować niepewność łączną z odchyleniem standardowym w próbie pochodzącej z populacji normalnej /przy odpowiedniej liczbie stopni swobody/, a więc  $k$  byłoby identyczne z dobrze znanym czynnikiem  $t$  /z rozkładu  $t$  Studenta/. Tym niemniej nie należy brać zbyt serio interpretacji tego rodzaju. Mając na uwadze wielką różnorodność ewentualnych zastosowań Grupa Robocza nie sądzi, aby spoczywał na niej obowiązek ustalenia jedynej zalecanej wartości dla czynnika  $k$ . Należy to pozostawić do decyzji eksperymentatora lub jakiegokolwiek innej instytucji /np. OIML, ISO, IEC/. Jedy- nym ścisłym wymaganiem powinno być to, że jeżeli stosuje się czynnik tego rodzaju, to jego wartość liczbowa powinna być wykazywana w sposób jednoznaczny.

Przy zastosowaniach w metrologii przemysłowej należałoby przypomnieć, że niepewność globalna - pod warunkiem przyjęcia uproszczonej hipotezy o normalności populacji - odpowiada w sposób przybliżony zwykłemu przedziałowi ufności. Zawsze gdy publikuje się istotnie ważne wyniki prowizoryczne i definitywne /na przykład w badaniach podstawowych, w certyfikatach uwierzytelnień/, należy podawać kompletną listę składników niepewności. Niekiedy należy również podawać kowariancje. W innych przypadkach wystarczy wykazywać niepewność globalną i czynnik  $k$ .

#### 4.6. Uwagi końcowe

Przewodniczący zwrócił uwagę Grupy Roboczej na ważność właściwej redakcji Zalecenia. Powinno ono zawierać - w skon-

densowanej postaci - wytyczne w sprawie zasad, na których ma się wspierać nowa metoda wyrażania niepewności.

Zalecenie zostanie przedłożone CIPM podczas jej następnego zebrania w październiku 1981 roku. CIPM może je akceptować, odrzucić lub zaproponować odpowiednie zmiany. Jeżeli będzie ono zatwierdzone, to stanie się oficjalnym zaleceniem międzynarodowym. Propozycja tekstu tego zalecenia podana jest w pkt. 4.7.; należy to traktować jako redakcję Grupy Roboczej.

Opracowanie projektu Zalecenia stanowiło jedno z głównych zadań tego posiedzenia. Ze względu na brak czasu nie było możliwości dogłębnego przedyskutowania wszystkich poruszanych kwestii. Wśród tych, które przewijały się wiele razy, należy wymienić problemy terminologiczne. W konsekwencji pewne nowe terminy, które figurują w Zaleceniu, należy uznać za prowizoryczne, w szczególności wyrażenia "niepewność łączna" lub "niepewność globalna". Ta sama uwaga odnosi się do wyrażen "kategoria A" i "kategoria B", jak również do symboli stosowanych do odpowiednich niepewności.

Jest rzeczą oczywistą, że proponowane Zalecenie nie podaje wszelkich wskazówek niezbędnych do wdrożenia nowych wytycznych do prac laboratoryjnych. W tym celu zamierza się zredagować "książkę z przepisami", w której analizowana będzie pewna liczba typowych przykładów ilustrowanych obliczeniami numerycznymi. Na razie apeluje się do laboratoriów narodowych o stosowanie tego Zalecenia i o informowanie BIPM o wynikach tej akcji. Nie ustalono żadnej daty ewentualnego dalszego spotkania Grupy Roboczej.

4.7. Zalecenie Grupy Roboczej ds. Wyrażania  
Niepewności Pomiaru przedstawione  
Międzynarodowemu Komitetowi Miar /CIPM/

Wyrażenie niepewności eksperymentalnych

ZALECENIE INC - 1 /1980/<sup>X/</sup>

1. Niepewność wyniku pomiaru zawiera na ogół wiele składników, które można podzielić na dwie kategorie w zależności od metody stosowanej do oszacowania ich wartości liczbowej:

A/ takie, które wyznacza się za pomocą metod statystycznych,

B/ takie, które wyznacza się za pomocą innych metod.

Nie zawsze istnieje prosta odpowiedniość między podziałem na kategorie A lub B i podziałem na charakter "losowy" lub "systematyczny", stosowany dawniej do klasyfikacji niepewności. Wyrażenie "niepewność systematyczna" może prowadzić do mylnej interpretacji; należy go unikać.

Każdy szczegółowy opis niepewności powinien zawierać kompletny opis składników i wskazywać dla każdego z nich metodę stosowaną przy przypisywaniu im wartości liczbowej.

2. Składniki kategorii A charakteryzuje się wariancjami w próbie  $s_1^2$  /lub "odchyleniami standardowymi" w próbie  $s_1$ / i liczbami  $\nu_1$  stopnia swobody. W pewnych przypadkach należy podawać kowariancje w próbie.

<sup>X/</sup> Na 70 - sesji /w październiku 1981 r./ CIPM przyjął Zalecenie 1 /CI-1981/ zaznaczając, "że propozycje Grupy Roboczej mogą stanowić bazę dla ewentualnego uzgodnienia dotyczącego wyrażania niepewności", i zalecając, aby zostały one poddane próbie /por. pkt. 5/.

3. Składniki kategorii B należy charakteryzować za pomocą wyrazów  $u_j^2$ , które można traktować jako przybliżenia odpowiednich wariancji, co do których zakłada się, że istnieją. Wyrazy  $u_j^2$  można uważać za wariancje, a wyrazy  $u_j$  - za odchylenia standardowe. W pewnych przypadkach kowariancje należy traktować w sposób analogiczny.
4. Niepewność łączną należy charakteryzować wartością uzyskaną przez zastosowanie zwykłej metody składania wariancji. Niepewność łączną, jak również jej składniki należy wyrażać w postaci "odchyień standardowych".
5. Jeżeli w szczególnych zastosowaniach należy pomnożyć przez pewien czynnik niepewność łączną chcąc uzyskać niepewność globalną, to powinno się zawsze podawać wartość liczbową tego czynnika.

#### 4.8. Spis dokumentów przedstawionych na posiedzeniu

1. Report on the BIPM enquiry on error statements, Rapport BIPM-80/3, 50 stron /BIPM, 1980/. Wersja francuska jest opublikowana w Procès-Verbaux CIPM 48 /1980/, Annexe III, s. C 1-36.
2. Guidelines for estimation and statement of overall uncertainty in measurement results, National Physical Laboratory of India oraz Standards and Industrial Research Institute of Malaysia, CSC/80/ MS-9/Commonwealth Science Council, Londyn, 1980/, 13 stron.
3. H. - J. v. Martens i E. Pippig: Beschreibung des Fehlers eines korrigierten Messergebnisses, Feingerätetechnik 28, s. 359-364 /1979/.
4. S. R. Wagner: On the quantitative characterization of the uncertainty of experimental results in metrology, PTB-Mitteilungen 89, s. 83-89 /1979/.
5. S. R. Wagner: Combination of systematic and random

- uncertainties, w Conference on Precision Electromagnetic Measurements, Braunschweig /czerwiec 1980/, s. 241-246.
6. J. W. Müller: Some second thoughts on error statements, Nucl. Instr. and Meth. 163, s. 241-251 /1979/.
  7. J. W. Müller: Les incertitudes de mesures, w La Physique, Encyclopédie Scientifique de l'Univers, Vol. 4 /Gauthier-Villars, Paryż, 1981/, s. 11-17.
  8. G. Colonnetti: Komentarze opracowane przez Istituto di Metrologia, Turyn, Włochy, 25 lipiec 1980, 14 stron.
  9. Komentarze opracowane przez Amt für Standardisierung, Messwesen und Warenprüfung, Berlin, NRD, 30 lipiec 1980, 9 stron.
  10. Expression des incertitudes systématiques, komentarze opracowane przez Office Fédéral de Métrologie, Berno, Szwajcaria, 23 wrzesień 1980, 10 stron.

Te dokumenty można otrzymać po zwróceniu się do Dyrektora CIPM, Pavillon de Breteuil, F-92310 Sèvres, Francja.

#### 5. ZALECENIE 1 /CI-1981/ CIPM

Wyrażenie niepewności eksperymentalnych

Międzynarodowy Komitet Miar

biorąc pod uwagę

- konieczność uzgodnienia jednolitych sposobów dotyczących wyrażania niepewności w metrologii,
- wieloletnie wysiłki podejmowane w tym celu przez różne jednostki organizacyjne,
- zachęcające postępy zmierzające w kierunku możliwego rozwiązania, które wyniknęły z dyskusji prowadzonych w łonie Grupy Roboczej na temat wyrażania niepewności, która zebrała się w BIPM w 1980 roku,

stwierdza

- że propozycje Grupy Roboczej mogą stanowić bazę ewentualnego uzgodnienia dotyczącego wyrażania niepewności,

zaleca

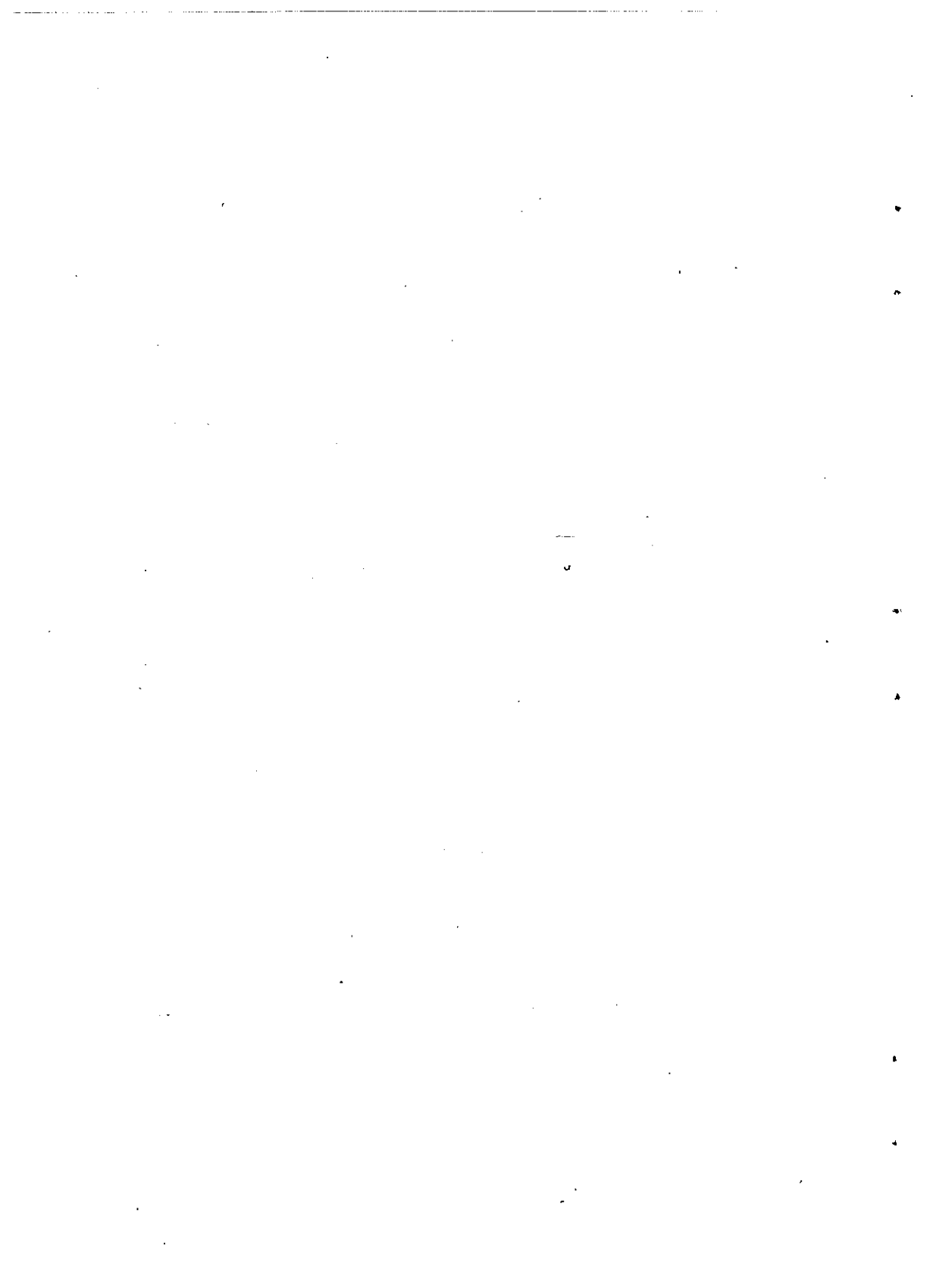
- aby propozycje tej Grupy Roboczej były szeroko rozpowszechnione wśród zainteresowanych,
- aby BIPM postarało się zastosować zasady zawarte w tych propozycjach do komparacji, które będzie ono organizować w nadchodzących latach,
- aby inne zainteresowane jednostki organizacyjne przestudiowały i poddały próbom te propozycje oraz przekazały swoje obserwacje do BIPM,
- aby w ciągu dwóch lub trzech lat BIPM dokonało ustalenia dotyczącego wdrożenia do praktyki tych propozycji.

WYKAZ LITERATURY<sup>X/</sup>

1. Berka K.: Measurement. Its Concepts, Theories and Problems. D. Reidel Publishing Company. Dordrecht, Boston, London 1983.
2. CIPM: Procès-Verbaux de séances. 2<sup>e</sup> série, tome 48, 69<sup>e</sup> session 1980 /7-9 octobre/, BIPM, Sèvres 1981, str. C1+C36.
3. CIPM: Procès-Verbaux des séances. Tome 49. 70<sup>e</sup> session - 1981, BIPM, Sèvres 1982, s. 26 i A1+A12.
4. Glossary of Terms used in Metrology, BS 5233: 1975. London, British Standard Institution, 1975.

X/ Niniejsza bibliografia została opracowana przez tłumaczy i zawiera wybrane pozycje, w których czytelnik może znaleźć podstawowe informacje dotyczące problematyki niepewności pomiarów.

5. Государственная система обеспечения единства измерений. Метрология. Термины и определения. ГОСТ 16263-70, Москва 1970.
6. Jaworski J.: Matematyczne podstawy metrologii. PWN, Warszawa 1979.
7. Obalski J.: Podstawy metrologii. Politechnika Warszawska /skrypt/, 1966.
8. Piotrowski J.: Podstawy metrologii. PWN, Warszawa 1979.
9. Polska Norma: Metrologia. Nazwy i określenia. PN-71/N-02050, PKN, 1972.
10. Polska Norma: Rachunek prawdopodobieństwa i statystyka matematyczna. Nazwy, określenia i symbole. PN-74/N-01051, PKNiM 1976.
11. Praca zbiorowa pod redakcją H. Szydlowskiego: Teoria pomiarów. PWN, Warszawa 1974.
12. Praca zbiorowa pod redakcją P.H. Sydenhama: Handbook of Measurement Science, tom I. John Wiley, New York 1982.
13. Rozenberg W. J.: Wstęp do teorii błędów systemów pomiarowych. PWN, Warszawa 1982.
14. Rumazycki L.Z.: Matematyczne opracowanie wyników eksperymentu. WNT, Warszawa 1973.
15. Strzałkowski A., Śliżyński A.: Matematyczne metody opracowania wyników pomiarów. PWN, Warszawa 1973.
16. Urban A.: Podstawy miernictwa. Część I. Politechnika Warszawska /skrypt/, 1977.
17. Vocabulaire International des Termes Fondamentaux et Généraux de Métrologie, BIPM-CEI-ISO-OIML, 1984.
18. Zajdel A.N.: Elementarna ocena błędów pomiarów. PWN, Warszawa 1967.





## ODPOWIEDŹ NARODOWEGO BIURA WZORCÓW /NBS-USA/

## 1. PRZEDMOWA

Popieramy koncepcję polegającą na tym, że specjalny Komitet powołany przez CIPM w sprawie wyrażania niepewności pomiarów przygotowuje raport o zasadach, którymi należy kierować się przy podawaniu informacji dotyczących niepewności. Te zasady powinny być przestrzegane we wszystkich międzynarodowych organizacjach metrologicznych i na różnych poziomach wewnątrz tych organizacji, tzn. w laboratoriach narodowych, laboratoriach uwierzytelniania drugiego rzędu oraz w certyfikatach przeznaczonych dla bezpośrednich użytkowników przyrządów.

Dlatego istotne jest, aby raport zawierał dyrektywy pozwalające określić niepewność, przy czym należy jednak unikać - jak tylko to jest możliwe - nie kończących się dyskusji filozoficznych dotyczących teorii statystycznych. Komitet powinien zapoznać się ze sposobami wyrażania niepewności stosowanymi w różnych laboratoriach narodowych i szukać zgodności w skali międzynarodowej, unikając procedur prowadzących do strat informacji naukowych albo uniemożliwiających późniejszą analizę doświadczeń czy wzorcowań. Można by sporządzić świadectwa wzorcowania, w których byłaby podana obszerna lista obejmująca niepewności systematyczne i losowe, sposoby ich łączenia i ewentualnie jedna liczba uzyskana na podstawie dowolnej metody składania błędów przypadkowych i systematycznych.

## 2. ODPOWIEDZI NA PYTANIA BIPM

Pytanie 1: Czy należy zalecać stosowanie odchylenia standardowego jako parametru charakteryzującego niepewność losową?

Odpowiedź: Jeżeli różnice  $(x_1 - \bar{x})$  w serii wyników pomiaru zachowują się mniej więcej tak jak "zmienna losowa X", stosowanie odchylenia standardowego w celu scharakteryzowania niepewności losowej należy uznać za całkowicie uzasadnione.

Komentarz 1A: Ponieważ "charakter losowy" jest pojęciem, które nie może być całkowicie zweryfikowane, stopień przybliżenia może być oszacowany tylko przez odniesienie do pewnej hipotezy wprowadzającej "charakter nielosowy". Warunkiem, który co najmniej musi być spełniony, jest to, aby rozkład wyników lub różnic w zależności od parametrów, które trzeba uwzględnić (ciąg odczytów, czas, temperatura, wilgotność, rodzaj przyrządów, obserwatorzy itp.) nie wykazywał na pierwszy rzut oka żadnego gromadzenia, tendencji lub innej struktury.

Komentarz 1B: Liczba stopni swobody  $\nu$  jest nierozdzieloną częścią odchylenia standardowego w próbce i powinna być również wykazywana. Gdy wynik nie jest zwykłą średnią, liczba  $\nu$  może być znacznie mniejsza, niż całkowita liczba pojedynczych obserwacji wykorzystanych w obliczeniach.

Komentarz 1C: Przy rutynowych wzorcowaniach i przy wystawianiu świadectw /certyfikatów/ nie jest dobrym zwyczajem stosowanie "odchylenia standardowego w próbce" w celu określenia niepewności losowej. Lepiej jest stosować odchylenie standardowe  $\sigma$  jako parametr pomiaru. Odchylenie standardowe  $\sigma$  jest wielkością "stabilną", w porównaniu z parametrem  $s$ , który może zmieniać się znacznie, gdy liczba  $n$  próbek w każdej serii jest mała. Gromadząc wartości  $s^2$  z dużej liczby serii można określić wartość  $\sigma^2$  dla rozważanej metody pomiaru.

Komentarz 1D: W przypadkach bardziej skomplikowanych można uzyskać odchylenie standardowe w próbce przyjmując jako punkt wyjścia pewną liczbę składowych wariancji reprezentujących różne przyczyny zmienności. Powinno się podawać listę tych przyczyn zmienności.

Pytanie 2: Czy należy zalecać stosowanie umownego prawdopodobieństwa i odpowiednich granic przedziału ufności /zamiast odchylenia standardowego/?

Odpowiedź: Wg naszych doświadczeń granice przedziału ufności są często źle interpretowane. Ponieważ można łatwo obliczyć granice przedziału ufności wychodząc z wielkości podstawowych /odchylenia standardowego w próbce i liczby stopni swobody, uzupełnionych hipotezą dotyczącą rozkładu normalnego "różnic"/, nie zalecamy ich stosowania zamiast odchylenia standardowego.

Komentarz 2A: Jest rzeczą oczywistą, że stosowanie granic ufności /lub przedziałów ufności/ jest właściwe w tych przypadkach, dla których ww. określenia są przewidziane. Tak więc w wyrażeniu na niepewność hipoteza rozkładu normalnego i wybór poziomu prawdopodobieństwa komplikują jeszcze problem.

Komentarz 2B: Zwracamy uwagę, że można uogólnić granice ufności nadając jej następującą postać:

$$\bar{x} \pm \left( \frac{s}{\sqrt{n}} \right) t \left( \nu, \frac{1-p}{2} \right)$$

gdzie  $n$  jest liczbą wyników uwzględnionych przy obliczaniu  $\bar{x}$ , a  $\nu$  liczbą stopni swobody dotyczącą estymatora  $s$ . Można wymienić dwa przypadki specjalnie ważne.

/1/ Dla  $k$  serii wyników, z których każda składa się z  $n$  pomiarów, wartość  $s$  oblicza się przyjmując jako punkt wyjście zbiór danych i uznając, że  $\nu = k(n-1)$ ; zakłada się przy tym, że wszystkie odchylenia standardowe w próbce pochodzące z każdej serii dają tę samą wartość  $\sigma$ . Przy odpowiedniej wartości  $k$  /niezbyt małej/ można uzyskać "rozsądną" szerokość przedziału ufności, nawet już przy tak małych wartościach  $n$  jak 2.

/2/ Jeżeli pomiędzy seriami występują różnice, to należy obliczać  $s$  na podstawie  $k$  średnich i przyjmować  $\nu = k - 1$ .

Komentarz 2C: Wykazano<sup>x/</sup>, że dla rozkładów symetrycznych stosowanie poziomu  $p = 0,95$  przy przedziałach ufności jest bardziej stabilne /"odporne"/ i bardziej zadawalające niż poziomu  $p = 0,99$ . Jeżeli trzeba więc wybierać poziom prawdopodobieństwa, to zalecamy stosowanie wartości  $p = 0,95$ .

Pytanie 3a: Czy istnieje istotna różnica między błędami przypadkowymi a błędami systematycznymi?

Odpowiedź: To czy jakiś błąd należy uważać jako przypadkowy czy systematyczny, zależy przede wszystkim od kontekstu. Istotna różnica pomiędzy błędami przypadkowymi a systematycznymi manifestuje się wtedy, gdy niepewność wyniku jest skutecznie wykorzystywana.

Komentarz 3A: Naszym zdaniem trudności w uzgodnieniu definicji błędu systematycznego biorą się stąd, że różni metrologowie nie rozpatrują tego problemu z tego samego punktu widzenia. Odchylenie standardowe poprawki, którą należy przyporządkować etalonowi kilograma N KG-1, bazuje na powtarzalnych wzorcowaniach w odniesieniu do innych etalonów kilograma przechowywanych w NBS; jest to miara błędu przypadkowego jeżeli chodzi o NBS, ale niepewność przypisywana ww. poprawce, powiedzmy trzy odchylenia standardowe, jest składnikiem błędu systematycznego wspólnego wszystkim etalonom kilograma uwierzytelnionym w odniesieniu do N KG-1 dla wszystkich laboratoriów w Stanach Zjednoczonych.

Komentarz 3B: Jeżeli dwa obiekty B i C są wzorcowane w odniesieniu do wzorca A, któremu przypisano niepewność  $\Delta A$ , to gdy interesująca jest różnica B-C, błąd systematyczny  $\Delta A$  znosi się, natomiast odchylenia standardowe właściwe dla B i C składają się "kwadratowo". W konsekwencji możliwość ujawnienia różnicy między B i C zwiększa się wtedy, gdy niepewność obiektu B i obiektu C jest podawana w dwu

<sup>x/</sup>J.E.WALSH: "Validity of approximate normality values for  $\mu \pm k\sigma$  areas of practical type continuous population". Annals of the Institute of Statistical Mathematics 8, s. 79-86 /czerwiec 1956/.

częściach, w części systematycznej, która może być wspólna dla obu i w częściach losowych, które są indywidualne.

Komentarz 3C: W przemyśle i w handlu, na obszarze tego samego kraju, niepewność etalonu narodowego nie ma większych konsekwencji, skoro odpowiedni błąd jest wspólny dla wszystkich urządzeń i przyrządów wzdłuż całej drogi uwierzytelniania.

Pytanie 3b: Czy należy zalecać praktyczną regułę pozwalającą określić charakter błędu?

Odpowiedź: Istnieje duża potrzeba instrukcji /dyrektyw/, które pomogłyby odróżniać jeden typ błędu od drugiego. Należałoby podać przykłady dla typowych sytuacji spotykanych w dziedzinie metrologii.

Komentarz 3A: Dyrektywy i przykłady są użytecznym narzędziem, ale nie zastąpią one głębokiego namysłu w odniesieniu do konkretnego problemu praktycznego.

Pytanie 4: Czy należy zalecać praktyczną regułę wyrażania błędów systematycznych?

Odpowiedź: Granice błędów systematycznych można podzielić na dwie kategorie: te, które oparte są na rezultatach eksperymentalnych /jak w przykładzie podanym w komentarzu 3A/ i te, które opierają się całkowicie na ocenie eksperymentatora. Uważamy, że należy opracować praktyczne reguły dla każdej kategorii.

Komentarz 4A: Chcąc, aby reguła praktyczna dotycząca wyrażania błędów systematycznego była użyteczna, należy tak ją zredagować, aby odpowiadała ona celom, których się poszukuje interpretując doświadczenie.

Komentarz 4B: W wielu przypadkach granice błędów systematycznych przyjmują formę asymetryczną, tzn. formę typu  $-a, +b/$ .

Komentarz 4C: Błąd systematyczny dotyczący wyniku może również powstawać wtedy, gdy używa się procedury estymacji lub metody kalkulacji, która wykazuje pewne obciążenia /"polaryzację"/, jak na przykład biorąc stosunek średnich zamiast średniej stosunków. Takie źródło błędu systematycznego można jednak zidentyfikować i skorygować, jego rząd wielkości i znak można oszacować i uwzględnić.

Pytanie 5: Czy należy zalecać praktyczną regułę łączenia błędów systematycznych?

Odpowiedź: Uważamy, że można by ustalić praktyczne reguły łączenia granic błędów systematycznych i że byłyby one użyteczne.

Komentarz 5A: Przy "dużej" liczbie błędów systematycznych, o mniej więcej równym rzędzie wielkości, łączenie za pomocą dodawania kwadratowego jest racjonalne w miarę tego jak można oczekiwać, że pewne składniki skompensują się wzajemnie. Gdy jeden lub dwa błędy systematyczne są dwa do trzech razy większe od pozostałych, reguła dodawania liniowego wydaje się bardziej realistyczna<sup>x/</sup>. Wybór należy jeszcze w znacznym stopniu od późniejszego wykorzystania wyniku.

Komentarz 5B: Niezależnie od tego, czy podejmie się decyzję łączenia błędów systematycznych za pomocą dodawania kwadratowego, czy za pomocą dodawania liniowego, jest równie możliwe, że dwa źródła błędów systematycznych wpływają jedno na drugie i wytwarzają błąd systematyczny znacznie większy niż suma błędów obu źródeł uwzględnianych oddzielnie. W takim przypadku tylko wyniki eksperymentalne mogą uwierzytelniać przypisane rzędy wielkości.

x/ C. Eisenhart: "Realistic evaluation of the precision and accuracy of instrument calibration system" J. Res. NBS 67C, str. 161-187 /1963/. Porównaj w szczególności tablicę 1 na stronie 184.

Komentarz 5C: W rzeczywistości łączy się tolerancje lub granice błędów systematycznych /albo błędów przypadkowych/, a nie same błędy. Być może lepiej byłoby mówić o "niepewności przypadkowej" i o "niepewności systematycznej" zamiast o "błędzie przypadkowym" i o "błędzie systematycznym". W pojęciu pewnych osób słowo "błąd" ma konotację "znanej omyłki" i stąd powstaje pytanie, co przeszkadza wyeliminować błąd całkowicie. Jest to szczególnie prawdziwe w przypadku wyrażenia "błąd systematyczny".

Pytanie 6: Czy należy zalecać praktyczną regułę łączenia błędów przypadkowych i systematycznych?

Odpowiedź: O ile tylko jeden typ błędu nie jest pomijalny w stosunku do drugiego, to łączenie obu prowadzi nieuchronnie do pewnej straty informacji. Stosownie do komentarzy do do pytania trzeciego mamy wrażenie, że użytkownik wyników jest bardziej predystynowany do formułowania swojej własnej reguły.

Komentarz 6A: Przy pomiarach na najwyższym poziomie, jak np. przy wyznaczaniu podstawowych stałych fizycznych, wg wszelkiego prawdopodobieństwa błędy systematyczne dominują nad przypadkowymi; na "ryнку" jest na ogół przeciwnie. Pomiędzy tymi dwoma przeciwieństwami rzędy wielkości błędów pochodzących z jednego czy drugiego źródła są prawdopodobnie takie same, można bowiem i powinno się ujawniać i eliminować przyczynę nienormalnie dużego błędu systematycznego. Wzorcowanie i wystawianie świadectwa /certyfikatu/ należą zazwyczaj do takiej kategorii.

Komentarz 6B: Chociaż praktyczne reguły wyrażania niepewności wynikowej mogłyby być użyteczne /patrz pytanie 7/, trudno jest opracować metodę pozwalającą łączyć granicę dwu istotnie różnych typów błędów. Trzeba by przynajmniej oddzielnie wykazywać granice błędów przypadkowych i błędów systematycznych; można by uczynić to regułą. Można by potem dyskutować o sposobie łączenia błędów dla uzyskania jedyne go wyniku.

Pytanie 7: Czy należy zalecać praktyczną regułę wyrażania niepewności wynikowej?

Odpowiedź: Reguła praktyczna wyrażania niepewności wynikowej jest użyteczna w tym sensie, że niepewność wynikowa podaje dopuszczalną niedokładność wyniku.

Komentarz 7A: Gdy niepewność wynikowa zawiera składniki, z których pewne nie są oparte na danych doświadczalnych, jest rzeczą niewłaściwą przypisywanie im określonego poziomu prawdopodobieństwa. Preferuje się granice 2 $\sigma$  lub 3 $\sigma$  dla składników losowych błędu /lub odpowiedni równoważnik w przypadku małej próbki<sup>x/</sup>/. Te konwencjonalne granice nie mają żadnego dokładnego znaczenia w kategoriach teorii prawdopodobieństwa, chyba że postać rozkładu błędów byłaby rzeczywiście dobrze uzasadniona.

Komentarz 7B: Reguła mogłaby być różna dla różnych warunków. Należałoby podać kilka typowych przykładów.

Pytanie 8: Jakie zagadnienia wydają się Wam istotne? Czy poprzednie zagadnienia powinny być według Was rozpatrywane w innym porządku?

Odpowiedź: Pytania przedstawione powyżej są pytaniami istotnymi, odpowiednimi dla przedłożenia grupie roboczej. Pytania te są ułożone w naturalnym porządku.

Komentarz 8A: Ponieważ jasno wyczuwamy, że wyrażanie niepewności, tak co do formy jak i co do treści, zależy od rodzaju wykorzystania wyniku, podzieliliśmy ogólnie te rodzaje na cztery grupy:

- badania naukowe,
- tablice stałych fizycznych,
- materiały odniesienia,
- certyfikaty sprawdzenia i uwierzytelniania.

<sup>x/</sup> B.L. Joiner: "Student-t deviate corresponding to a given normal deviate", J. Res. NBS 73C, str. 15-16 /1969/.



### 3. INFORMACJE DODATKOWE

#### 3.1. Przykłady

Podajemy cztery przykłady typów wyrażeń niepewności wykorzystywanych i zalecanych do:

##### Badań naukowych:

R.N. Goldberg: "Microcalorimetric determination of glucose in reference samples of serum". *Clinical Chemistry* 22, str. 1685-1691 /1976/.

##### Tablic stałych fizycznych

"Guide for the Presentation in the Primary Literature of Numerical Data Derived From Experiments" Unesco-UNISIST Guide, raport Grupy CODATA na temat publikacji wyników, wrzesień 1973. Przedrukowany w serii National Standard Reference Data System /NSRDS/ News, luty 1974.

##### Materiałów odniesienia

Certyfikat dla Standard Reference Material 993, Uranium - 235 Spike Assay and Isotopic Solution Standard, National Bureau of Standards, marzec 13, 1975.

##### Certyfikatów sprawdzania i uwierzytelniania

Report of Length Values, Test Number L7051-1, Dimensional Technology Section. National Bureau of Standards /wersja skrócona/.

#### 3.2. Dane bibliograficzne

Poza publikacją "NSRDS News" wykazaną powyżej pt. "Tables de constantes" chcielibyśmy wymienić następujące publikacje zawierające ogólne dyrektywy odpowiednio dla NPL i NBS.

NPL: P.J. Campion, J.E. Burns, A. Williams: "A code of practice for the detailed statement of accuracy" /HMSO, London, 1973/.

NBS: C. Eisenhart: "Expression of the uncertainties of final results". H.H. Ku: "Expression of imprecision, systematic error, and uncertainty associated with a reported value". Te dwa dokumenty zostały przedrukowane w NBS Special Publication 300, Vol. 1, Precision Measurement and Calibration: Statistical Concepts and Procedures, U.S. Government Printing Office, 1969.

Międzynarodowa Organizacja Metrologii Prawnej w swoim "Słowniku Metrologii Prawnej" podaje również dyrektywy dotyczące wyrażania niepewności. Patrz w szczególności rozdział 8. "Erreurs des Résultats de Mesurage et Erreurs des Instruments de Mesurage" i rozdział 9. "Conditions d'emploi et qualités métrologiques des Instruments de Mesurage".

#### Dodatek 2

### ODPOWIEDŹ NARODOWEGO BIURA METROLOGII /BNM - FRANCJA/

Niniejsza odpowiedź została przygotowana wspólnie z l'Établissement Technique Central de l'Armement /ETCA - Centralny Zakład Techniczny Uzbrojenia/, l'Institut National de Métrologie /INM/ du Conservatoire National des Arts et Metiers /Narodowy Instytut Metrologii Narodowej Wyższej Szkoły Sztuk i Rzemiosł/, le Laboratoire Central des Industries Électriques /LCEI - Centralne Laboratorium Przemysłu Elektrotechnicznego/ i le Laboratoire Primaire de Métrologie de Rayonnements Ionisants /LMRI - Pierwotne Laboratorium Metrologiczne Promieniowania Jonizującego/.

Pytanie 1: Czy należy zalecać stosowanie odchylenia standardowego jako parametru charakteryzującego niepewność losową?

Odpowiedź: Odchylenie standardowe daje korzyści w charakteryzowaniu niepewności losowej niezależnie od typu rozkładu. Ponieważ jednak można stąd uzyskać tylko ocenę za pomocą odchylenia standardowego z próbki, należy podawać liczbę uży-

tych wartości /liczność próbek/. W pewnych przypadkach można również podawać rozrzut wyników.

Pytanie 2: Czy należy zalecać stosowanie umownego prawdopodobieństwa i odpowiednich granic przedziału ufności /zamiast odchylenia standardowego/?

Odpowiedź: Prawdę mówiąc, nie ma wyboru między stosowaniem granic przedziału ufności a odchyleniem standardowym. W rzeczywistości gdy rozkład nie jest znany, można estymować tylko odchylenie standardowe.

Stosowanie granicy przedziału ufności okazuje się wygodne w praktyce dla scharakteryzowania niepewności wyniku końcowego, aby nie stracić jednak informacji należy również podawać liczbę stopni swobody /lub liczbę pomiarów/. Byłoby zresztą korzystne znormalizować odpowiednią wartość poziomu ufności /najlepiej 95%/. Należałoby wyeliminować stosowanie granic przedziału ufności do obliczeń pośrednich.

Pytanie 3a i 3b: Czy istnieje istotna różnica między błędami przypadkowymi a błędami systematycznymi? Czy należy zalecać praktyczną regułę, pozwalającą określić charakter błędu?

Odpowiedź: Istnieje różnica między błędami przypadkowymi a błędami systematycznymi.

Błąd systematyczny można określić jako błąd, który - podczas wielu pomiarów wykonywanych w tych samych warunkach tej samej wartości pewnej wielkości - pozostaje stały co do wartości bezwzględnej i co do znaku lub który zmienia się zgodnie z określonym prawem, gdy warunki zmieniają się.

Błąd przypadkowy /losowy/ można określić jako błąd, który zmienia się w sposób nieprzewidziany co do wartości bezwzględnej i co do znaku, gdy wykonuje się dużą liczbę pomiarów tej samej wartości pewnej wielkości w praktycznie identycznych warunkach /por. normę NFx07-007/. Można to interpretować etatystycznie w następujący sposób. Dla każdego pomiaru istnieje różnica /odstęp/ zwana "błędem" między wartością mierzoną a "wartością prawdziwą". Ten odstęp można traktować jako wartość, którą przyjmuje zmienna losowa,

której wartość oczekiwana przedstawia część systematyczną, a której odchylenie standardowe charakteryzuje część losową.

Jeżeli wartość średnia /oczekiwana/ jest znikomo mała względem odchylenia standardowego, to mówi się, że dominującą cechą błędu jest losowość. W przeciwnym przypadku mówi się, że dominującą cechą błędu jest systematyczność. W rzeczywistości mogą mieć miejsce wszystkie sytuacje pośrednie między tymi skrajnymi przypadkami.

Takie rozróżnienie między błędem systematycznym i losowym jest możliwe tylko wtedy, gdy można było wyznaczyć wartość średnią i w tym przypadku bierze się to pod uwagę uwzględniając poprawki systematyczne, co ma na celu jego zmniejszenie. Jednak na ogół wartość średnia nie jest znana i dysponujemy raczej subiektywną ocenę jej rozrzutu wokół pewnej wartości, co do której zakłada się zazwyczaj, że jest równa zeru.

Poza tym należy poczynić następujące uwagi:

- Błędy losowe i błędy systematyczne odpowiadają różnym pojęciom fizycznym; należałoby raczej zachować te dwa terminy.
- Zwiększając liczbę pomiarów można zmniejszyć wartość odchylenia standardowego - miarę rozrzutu błędów losowych, lecz w żadnym razie nie można zmniejszyć wpływu błędów systematycznych /to można tylko uzyskać stosując odpowiednie metody eksperymentalne, a nie powtarzając pomiary/.

Pytanie 4: Czy należy zalecać praktyczną regułę wyrażania błędów systematycznych?

Odpowiedź: Byłoby rzeczą pożądaną osiągnąć ujednoczenie takich wyrażen, co pozwoliłoby charakteryzować rozrzut wartości średniej /z próbki/. Wydaje się, że obecna praktyka, polegająca na stosowaniu maksymalnych granic błędu, prowadzi do wyników pesymistycznych. Jednak każda inna reguła

praktyczna i bardziej realistyczna jest umowna i zależna od wielkości mierzonej, a nawet od stosowanej metody.

Pytanie 5: Czy należy zalecać praktyczną regułę łączenia błędów systematycznych?

Odpowiedź: Jeżeli występuje wiele błędów systematycznych, to dysponujemy subiektywnymi estymatorami dla scharakteryzowania rozrzutu różnych wartości średnich. Jeżeli chce się oszacować całkowity rozrzut i jeżeli można założyć, że te średnie są niezależne, to można uwzględnić w tych warunkach i łączyć te rozrzuty tak jakby chodziło o błędy losowe. Należy przy tym uczynić zastrzeżenia w odniesieniu do następujących punktów:

- liczba błędów systematycznych powinna być dostatecznie duża /w każdym razie większa od 4/,
- będzie wtedy można obliczyć samo odchylenie standardowe, natomiast nie będzie można wyznaczyć granic przedziału ufności, co wymagałoby znajomości rozkładu wartości średnich.

Tym niemniej należy silnie podkreślić fakt, że nie przestrzegając hipotez ukrytych w tej regule łączenia błędów systematycznych zwiększa się tylko ryzyko pogorszenia jakości pomiaru przez niedocenienie wagi błędów systematycznych.

Pytanie 6: Czy należy zalecać praktyczną regułę łączenia błędów przypadkowych i systematycznych?

Jeżeli założyć, że istnieje pewna praktyczna reguła łączenia ze sobą błędów systematycznych, taka o jakiej była mowa w odpowiedzi na pytanie 5 powyżej, to można by uwzględnić jakąś regułę łączącą błędy systematyczne i przypadkowe w następujący sposób: chcąc uzyskać globalną wariancję wyników należałoby dodać estymator wariancji wartości średniej z pomiarów do wariancji uzyskanej z połączenia estymatorów wariancji odpowiednio do różnych wartości oczekiwanych /systematycznych/.

I przeciwnie, jeżeli takiej reguły nie można by przestrzegać, to wydaje nam się, że przybliżenie /choćby pesymistyczne/, które polega na wyrażaniu błędów systematycznych jako "wartości maksymalnych" i na dodawaniu do granic przedziału ufności /błędy przypadkowe/ granic obliczonych /błędy systematyczne/, pozwala eksponować wagę błędów systematycznych.

Pytanie 7: Czy należy zalecać praktyczną regułę wyrażania niepewności wynikowej?

Odpowiedź: Nawet w przypadku, gdy różne błędy mogą być ujmowane pewną regułą łączenia ze sobą błędów systematycznych /por. odpowiedź na pytanie 5/ i pewną regułą łączenia błędów systematycznych i błędów przypadkowych /por. odpowiedź na pytanie 6/, to niemożliwe jest w ogólnym przypadku przedstawić globalnej niepewności w postaci przedziału ufności odpowiadającego danemu poziomowi ufności. Sprowadza się więc to do wyrażania tej niepewności opartej na globalnej wariancji wyników. Jednak dla niektórych przypadków praktycznych /na przykład w certyfikatach uwierzytelniania/ byłoby korzystne, aby można było ujednoczyć takie wyrażenie w postaci pewnej wielokrotności estymatora odchylenia standardowego.

To zastrzeżenie, dotyczące niemożliwości określenia przedziału ufności, stanowi w efekcie pewną trudność przy prezentowaniu wyników.

W konkluzji końcowej na temat niniejszego kwestionariusza chcemy w dalszym ciągu podkreślić niebezpieczeństwa stosowania metody "czysto" statystycznej w odniesieniu do błędów systematycznych, metody, którą staraliśmy się prezentować w naszych odpowiedziach. Jest rzeczą konieczną, aby zweryfikować hipotezy, dotyczące liczby uwzględnianych błędów systematycznych i ich niezależności. W przeciwnym przypadku niedocenienie wagi błędów systematycznych przyczyniłoby się do pogorszenia wiarygodności wyników pomiaru.

Bibliografia:

A. Allisy: "Les erreurs aléatoires" /Błędy przypadkowe/, kształcenie metrologiczne, /CNAM, Paryż, 1975/.

ODPOWIEDŹ FIZYCZNO-TECHNICZNEGO  
INSTYTUTU FEDERALNEGO /PTB-RFN/

1. WSTĘP

PTB powołało ostatnio grupę roboczą, która ma się zajmować jednocześnie aspektami terminologicznymi i praktycznymi dotyczącymi zagadnień wyrażania niepewności pomiaru.

Po wielu dyskusjach okazało się, że sprawą o pierwszorzędym znaczeniu są odpowiednie terminy jako podstawa głębszej analizy na płaszczyźnie międzynarodowej. Ponieważ chodziło tylko o wstępne studia prowadzone w łonie grupy roboczej złożonej z niewielkiej liczby uczestników, wśród których jest kilku ekspertów z DIN, proponowany słownik jak również nasze odpowiedzi należy traktować jako wstępne..

Sądzymy, że pojawią się jeszcze inne zagadnienia oraz że odpowiednie priorytety zostaną ustalone; wtedy przedstawimy nasze koncepcje w sposób bardziej szczegółowy.

Wreszcie należałoby opisać i wyjaśnić cel wyrażania niepewności pomiaru. Wyjaśnią to krótki wstęp dotyczący przeznaczenia i charakterystyk trzech poziomów hierarchii metrologicznej:

a/ Metrologia o wysokim poziomie

- Cele pomiarów stanowią, na przykład, wyznaczenie stałych fizycznych z bardzo małą niepewnością i porównania międzynarodowe między narzędziami pomiarowymi etalonów pierwotnych.
- W sprawozdaniach z wzorcowania lub w publikacjach należy zamieszczać wszystkie informacje, które mogą się okazać potrzebne przy dalszym przetwarzaniu /obliczeniach/ wyników pomiaru.

**b/ Metrologia o średnim poziomie**

- Celem pomiarów są, na przykład, wyznaczenie wartości wielkości fizycznych w naukach podstawowych i stosowanych ze średnią niepewnością pomiaru i wzorcowanie stanowisk pomiarowych etalonów wtórnych.
- W zależności od dziedziny należy podawać składowe niepewności wyniku pomiaru, jak również niepewność całkowitą w postaci jednej wartości. Ta ostatnia powinna być dostępna przy ewentualnym dalszym wykorzystaniu, gdzie może ona z kolei stanowić część niepewności dotyczącą innych eksperymentów.

**c/ Metrologia powszechnego zastosowania**

- Celem pomiarów są, na przykład, wyznaczenie wartości wielkości fizycznych w rutynowych pomiarach przeprowadzanych w handlu, przemyśle i medycynie, jak również wzorcowanie narzędzi stosowanych w pracach rutynowych.
- Niezależnie od tego, czy chodzi o wynik pomiaru, czy o wzorcowanie narzędzia stosowanego w pracy rutynowej, potrzebna jest tylko niepewność całkowita /a nie jej składowe/, aby ocenić niezawodność i aby móc wyeliminować wadliwe przyrządy podczas porównań narzędzi. Nie przewiduje się, aby podana niepewność była wykorzystywana przy dalszych obliczeniach.

W poniższych odpowiedziach na pytanie kwestionariusza staramy się uwzględnić potrzeby różnych poziomów hierarchii metrologicznej.

Byłoby rzeczą bardzo pożądaną i jednocześnie praktyczną wyrażanie i wyznaczanie niepewności w sposób jednolity w każdym przypadku i zgodnie z prostymi regułami /w miarę możliwości/.



W związku ze słownikiem /glosariuszem/ uwzględniliśmy następujące dokumenty:

- British Standards Institution Glossary of Terms Used in Metrology, 1974 /Brytyjski Ośrodek Normalizacji - Słownik terminów stosowanych w metrologii/.
- IEC Publications 359, 1971.
- ISO R 645, 1967.
- VDI/VDE 2600, 1973.
- OIML: Vocabulaire de Métrologie Légale, 1969 /Słownik metrologii prawnej/.
- CCIR: Draft Report /projekt sprawozdania/ 7/179, 28 marzec 1978.

## 2. PROJEKT SŁOWNIKA DOTYCZĄCEGO NIEPEWNOŚCI POMIARU<sup>X/</sup>

1. Wielkość mierzona. Wielkość podlegająca procesowi mierzenia.
2. Wartość /wielkości/. Wielkość wyrażona iloczynem liczby i jednostki miary. Na przykład: 5,3 m, 20°C.
3. Wielkość wpływowa. Wielkość nie będąca przedmiotem pomiaru, ale wpływająca na wartość wielkości mierzonej lub wskazania przyrządu pomiarowego, lub wartość wzorca miary odtwarzającego tę wielkość.  
 Uwaga: Wielkość wpływowa może "wynikać" z otoczenia lub z samego narzędzia pomiarowego. Na przykład: temperatura, częstotliwość prądu, własne podgrzewanie przyrządu, czas odpowiedzi.

<sup>X/</sup> Por. także normę krajową PN-71/N-02050, Metrologia. Nazwy i określenia. 1971 r. /przyp. tłum./

4. Wartość poprawna lub wartość umownie prawdziwa /wielkości/

- a/ Wartość wyznaczona z dostatecznie małą niepewnością dla każdego poszczególnego przypadku.
- b/ W odniesieniu do wielkości wpływowej: wartość, której można oczekiwać, gdy wielkość wpływowa przyjmuje wartość odniesienia /tzn. wartość, do której odnosi się wzorcowanie/.

5. Błąd wskazania /przyrządu pomiarowego lub wzorca miary/. Różnica: "wartość zmierzona mniej wartość poprawna" lub "wartość wskazana mniej wartość poprawna". Podaje się dwa następujące terminy - z pewną rezerwą - które można uznać jako przydatne tylko wtedy, gdy można przyjąć pewne hipotezy upraszczające:

- 5.1. Błąd przypadkowy /błąd powtarzalności/. Różnica: "wartość zmierzona mniej wartość średnia z szeregu pomiarów". Błąd przypadkowy powstaje wskutek oddziaływania wielkości wpływowych, które ulegają zmianie podczas pomiaru. Może on być określony w sposób jednoznaczny tylko wtedy, gdy czas trwania zmian jest porównywalny z czasem trwania pomiaru.
- 5.2. Błąd systematyczny. Termin stosowany dla wyrażania ewentualnej różnicy "wartość zmierzona mniej wartość poprawna" powstałej w skutek oddziaływania stałych wielkości wpływowych lub wskutek pewnego obciążenia<sup>x/</sup> wprowadzonego w poprzednich pomiarach, których nie można już powtórzyć /Błędy systematyczne znane można skorygować, jednak niepewność pozostaje/.

<sup>x/</sup> Odpowiednik wyrażenia francuskiego "le biais", który w tym kontekście należy utożsamiać z pewną polaryzacją.

6. Niepewność /wartości zmierzonej/. Wyraża ona wartość bezwzględną ewentualnego odchylenia wartości zmierzonej od wartości poprawnej. Aby wykazać wielkość niepewności, można ją traktować jako przedział po jednej i drugiej stronie wartości zmierzonej lub najlepszego estymatora wartości; istnieje pewne prawdopodobieństwo, że w tym przedziale znajduje się wartość poprawna. Często istnieje możliwość rozróżnienia następujących dwóch składowych:
- 6.1. Niepewność losowa /niepewność powtarzalności/. Pewna wielokrotność odchylenia standardowego z próbki /eksperymentu/ wartości średniej z szeregu pomiarów. Można ją zmniejszyć zwiększając liczbę pomiarów.
- 6.2. Niepewność systematyczna. Ta niepewność może być oceniana tylko na podstawie znajomości lub doświadczenia uzyskanych poza rozważanym pomiarem.
7. Składowa niepewności /lub niepewność cząstkowa/. Niepewność powstała wskutek oddziaływania pewnej wielkości wpływowej rozpatrywanej oddzielnie /por. termin 3 powyżej/, błędów obliczeniowych, pewnego obciążenia lub fluktuacji wskazywanych wartości.
8. Niepewność całkowita. Pewne połączenie /kombinacja/ wszystkich składowych niepewności.

### 3. ODPOWIEDZI NA PYTANIA BIPM

Chcemy podkreślić, że nasza norma narodowa, dotycząca tych zagadnień jak również innych zagadnień metrologicznych /DIN 1319, części 1+3/, jest w toku całkowitej rewizji. W związku z tym można na razie traktować nasze odpowiedzi tylko jako wytyczne do dyskusji, jaka ma się toczyć w łonie PTB i DIN.

Pytanie 1: Czy należy zalecać stosowanie odchylenia standardowego jako parametru charakteryzującego niepewność losowa?

Pytanie 2: Czy należy zalecać stosowanie umownego prawdopodobieństwa i odpowiednich granic przedziału ufności /zamiast odchylenia standardowego/?

Należałoby to pozostawić do decyzji każdego, lecz ten wybór powinien być zaznaczony explicite. W obu przypadkach należy podawać liczbę stopni swobody.

Pytanie 3a: Czy istnieje istotna różnica między błędami przypadkowymi a błędami systematycznymi?

Odpowiedź zależy od typu rozważanego eksperymentu. Odpowiedź pozytywna powinna obowiązywać tylko w przypadku, gdy wielkości wpływowe, które są zmienne, można wyraźnie sklasyfikować w zależności od tego, czy stałe czasu są albo krótkie, albo długie względem czasu trwania pomiaru. W wielu przypadkach oba czasy są porównywalne i to rozróżnianie nie ma zastosowania; odchylenie standardowe eksperymentalne /w próbcie/ zmienia się w zależności od czasu trwania pomiarów.

Pytanie 3b: Czy należy zalecać praktyczną regułę pozwalającą określać charakter błędu?

Porównaj odpowiedź na pytanie 3a.

Pytanie 4: Czy należy zalecać praktyczną regułę wyrażania błędów systematycznych?

W tym kontekście nie należałoby stosować terminu "błąd systematyczny", lecz tylko "niepewność systematyczna". Po uwzględnieniu tej modyfikacji odpowiedź powinna brzmieć "tak".

Pytanie 5: Czy należy zalecać praktyczną regułę łączenia błędów systematycznych?

Tak, należałoby zalecać pewne reguły łączenia błędów systematycznych /por. terminy 7 i 8 ww. słownika/.

Pytanie 6: Czy należy zalecać praktyczną regułę łączenia błędów przypadkowych i systematycznych?

Tak, należałoby zalecać odpowiednie reguły łączenia niepewności systematycznych i losowych.

Pytanie 7: Czy należy zalecać praktyczną regułę wyrażania niepewności wynikowej?

Ta sama odpowiedź jak na pytanie 6.

Pytanie 8: Jakie inne zagadnienia wydają się Wam istotne? Czy poprzednie zagadnienia powinny być wg Was rozpatrywane w innym porządku?

Należałoby prowadzić dyskusje na temat informacji, jakie powinny być podawane w świadectwach wzorcowania w różnych dziedzinach.

#### Bibliografia

- S. Wagner: PTB-Bericht, FMRB 31/69, listopad 1969.  
 H. Reich: "Die Unsicherheit von Messungen mit Ionisationskammer-Dosimetern" /Niepewność pomiarów za pomocą dozymetrów z komorą jonizacyjną/, PTB-Mitteilungen, 86, s. 421-426 /1976/.

Dodatek 4

### ODPOWIEDŹ NARODOWEGO LABORATORIUM FIZYCZNEGO /NPL - WIELKA BRYTANIA/

#### 1. PRZEDMOWA

Zanim udzielona będzie odpowiedź na ankietę BIPM dotyczącą niepewności, NPL chce zwrócić uwagę, że - wg jego opinii - najważniejszym zadaniem, jakiego mogłaby się podjąć specjalnie powołana grupa BIPM, jest precyzyjne sformułowanie odpowiednich podstawowych pojęć w celu stworzenia solidnej podstawy, na której będzie można oprzeć reguły praktyczne. Rozważając trudności tego przedsięwzięcia i starając się za wszelką cenę osiągnąć cel, trudno powstrzymać się od wyrażenia kilku wątpliwości co do szansy powodzenia.

## 2. WSTĘP

Nie ulega wątpliwości, że wiele dyskusji o niepewności jest jałowych. Przyczyną tego jest na ogół to, że uczestnicy tych dyskusji używają tych samych słów w różnych znaczeniach. Jest rzeczą bardzo istotną, aby przed przystąpieniem do rozważań o praktycznych regułach dotyczących niepewności Grupa Robocza uzyskała najpierw zgodę co do podstawowych pojęć i terminologii - o ile nie ma ona zamiaru poświęcać czas na dyskusje o nieporozumieniach. Zanim odpowiemy na specyficzne pytania ankiety BIPM, zrobimy przegląd tych podstawowych pojęć w trzech grupach:

- a/ rozróżnianie między błędami i niepewnościami,
- b/ klasyfikacja i definicja niepewności,
- c/ terminologia dotycząca niepewności.

Pozostając w tej samej kategorii pojęć można by postawić pytanie, czy słowa tak "skompromitowane" jak "błąd, dokładność i precyzja" mogą mieć znaczenie akceptowane na szczeblu międzynarodowym.

## 3. ROZRÓŻNIENIE MIĘDZY BŁEDEM I NIEPEWNOŚCIĄ

Skoro nauka przyswaja sobie słowa z języka potocznego pożądane jest, aby każde znaczenie zawężone, które jest tym słowom nadawane, nie różniło się zbyt od powszechnego znaczenia przyjmowanego w życiu codziennym. Tak więc pospolite w języku angielskim słowo "error" /"błąd"/ jest mniej więcej synonimem "omyłki, niesłuszności, różnicy, niezgodności, rozbieżności", gdy tymczasem słowo "uncertainty" /"niepewność"/ jest mniej więcej synonimem "wątpliwości, nieokreśloności, niecisłości, nieznamomości, niekompletnej wiedzy". Stąd wyrażając w kategorii prostej teorii eksperymentalnej można powiedzieć, że błąd pomiaru jest różnicą pomiędzy wartością zmierzoną a wartością prawdziwą wielkości

mierzonej. Niepewność wyraża natomiast wątpliwość dotyczącą dokładności pomiaru, czyli obszaru, wewnątrz którego - jak sądzi się - znajduje się wartość prawdziwa.

W wielu przypadkach można przekonać się, że wielkość mierzona powinna mieć wartość prawdziwą. W innych przypadkach z wielkością mierzoną związana jest znaczna niepewność. Wielkość losowa ma jednak wartość oczekiwaną. Ale we wszystkich tych przypadkach wartość prawdziwa nie jest znana i nie może być znana; wynika stąd, że błąd pomiaru nie jest znany i nie może być znany. Tak więc dyskusja o błędach jest, ogólnie rzecz biorąc, jałowa i w większości przypadków, gdy dyskutuje się o błędach, chodzi faktycznie o niepewności; zawsze należy wyraźnie rozróżniać te dwa terminy. Nie można niczym usprawiedliwić stosowania takich wyrażen, jak "propagacja błędów", gdy to o czym chce się powiedzieć, jest "łączeniem niepewności", prowadzi to bowiem do błędnych koncepcji i do dyskusji o nieporozumieniach.

#### 4. KLASYFIKACJA I DEFINICJA NIEPEWNOŚCI

Dyskusje o niepewnościach pozostają wciąż pod silnym wpływem starych rozróżnień pomiędzy błędami przypadkowymi, które są uwarunkowane wahaniami wartości mierzonej, a błędami systematycznymi związanymi ze stałymi błędami w pomiarze. Nie jest jednak korzystne dzielić niepewności w ten sam sposób. Zamiast tego proponuje się dzielić niepewności na dwie kategorie zależnie od sposobu, w jaki się je uzyskuje, a nie od sposobu, w jaki manifestują się one w pomiarach.

Można rozważać te dwie kategorie jako kategorie wprowadzające rozróżnienie między faktami a hipotezami. Z jednej strony mamy odchylenie standardowe, granice przedziału ufności itd., które są obliczone na podstawie wyników powtarzanych pomiarów; każdy powinien mieć możliwość uzyskania dokładnie tych samych wartości niepewności, jeżeli otrzyma

szczególne wyniki eksperymentu. Z drugiej strony mamy estymacje niepewności pochodzących na przykład: od zmian wpływów zewnętrznych, z pomiarów uzupełniających /dla wyznaczenia współczynników korekcyjnych/, z badania długookresowej zmienności wyników, a czasami po prostu - z inteligentnej oceny eksperymentatora; które inny badając te same wyniki eksperymentów mógłby dojść do innej wartości dotyczącej tej niepewności. Niełatwo jest podać precyzyjną definicję tych dwu kategorii niepewności, ale Grupa Robocza powinna spróbować to zrobić. Sugeruje się następujące definicje:

- a/ niepewności dotyczące wartości mierzonej estymowane za pomocą metod statystycznych na podstawie powtarzanych wyników pomiarów,
- b/ niepewności estymowane za pomocą metod niestatycznych dla uwzględnienia skutku oddziaływania na pomiar zjawisk zarówno stałych, jak i zmiennych.

## 5. TERMINOLOGIA DOTYCZĄCA NIEPEWNOŚCI

Oczywiste jest, że podane wyżej definicje dwu kategorii niepewności są zbyt długie, aby mogły być używane jako nazwy. W przeszłości wielu badaczy proponowało różne nazwy - szczególnie badacze z NBS, których publikacje na ten temat zebrano w świetnym opracowaniu "NBS Special Publication 300". Propozycje te są na ogół zbyt skomplikowane i nie zostały powszechnie przyjęte. W 1973 r. Burns, Campion i Williams zaproponowali w liście skierowanym do redakcji czasopisma Metrologia wyżej podaną klasyfikację oraz nazwy: niepewności przypadkowe i niepewności systematyczne. Te nazwy mają jednak oczywistą wadę: są zbyt podobne do terminów błędy przypadkowe i błędy systematyczne, które odpowiadają jednak innej klasyfikacji. Tym niemniej są one krótkie i od pewnego czasu coraz częściej stosowane w publikacjach naukowych.



Na pierwszy rzut oka rozważania powyższe mogą wydawać się zbyt filozoficzne, ale tak nie jest. Zmierzają one do utworzenia solidnych podstaw, na których można wznosić nadbudowę reguł i zaleceń. Grupa Robocza powinna jednak mieć dość czasu dla osiągnięcia zgodności do do pojęć fundamentalnych. Dopiero potem mogłaby przejść do dyskusji o praktycznych regułach stosowania niepewności. Może się zdarzyć, że w ogóle nie zbuduje się systemu całkowicie jednolitego, ponieważ różne wymagania i tradycje mogą narzucać różne sposoby wyrażania i różne metody pochodne. Jeżeli jednak w wyrażeniu niepewności rzeczywiste znaczenie tego wyrażenia jest niepewne, to to wyrażenie jest bez wartości.

#### 6. ODPOWIEDZI NA PYTANIA BIPM

Nie należy zapominać, że kwestionariusz nie oddziałuje w tym stadium na indywidualne czy zbiorowe punkty widzenia co do sposobu, w jakim powinno się traktować niepewność; pyta on raczej o zdanie co do tematów dyskusji dla Grupy Roboczej.

Pytanie 1. Czy należy zalecać stosowanie odchylenia standardowego jako parametru charakteryzującego niepewność losową?

Grupa Robocza mogłaby ku ogólnemu pożytkowi wyjaśnić różnicę pomiędzy informacją o powtarzalności wskazań przyrządu wzorcowanego /na ogół charakteryzowaną przez odchylenie standardowe odczytów indywidualnych/ a informacją o niepewności losowej współczynnika wzorcowania lub pomiaru; w rzeczywistości niepewność losowa może zawierać wiele składników poza tymi, które pochodzą z rozrzutu odczytów dokonanych przy wzorcowaniu lub podczas całego szeregu pomiarów.

Pytanie 2. Czy należy zalecać stosowanie umownego prawdopodobieństwa i odpowiednich granic przedziału ufności /zamiast odchylenia standardowego/?

Istnieją dwa typowe sposoby oznaczania niepewności losowej pomiaru:

- a/ odchylenie standardowe średniej wraz z efektywną liczbą stopni swobody,
- b/ granice przedziału ufności przy danym poziomie prawdopodobieństwa /ufności/ i efektywnej liczbie stopni swobody.

Jeżeli jedno z nich jest dane i jeżeli zakłada się pewien rozkład, to można obliczyć i drugie. Wydaje się bezcelowe tracenie czasu na dyskusje na ten temat zakładając, że różne okoliczności mogą odgrywać decydującą rolę w wyborze odpowiedniej metody. Co do wyboru poziomu prawdopodobieństwa - por. odpowiedź na pytanie 7.

Pytanie 3a. Czy istnieje istotna różnica między błędami przypadkowymi a błędami systematycznymi?

Pytanie 3b. Czy należy zalecać praktyczną regułę pozwalającą określać charakter błędu?

To jest ważny punkt; mówiliśmy o nim we wstępie do naszej odpowiedzi.

Pytanie 4. Czy należy zalecać praktyczną regułę wyrażania błędów systematycznych?

Istnieje wiele metod eksperymentalnych, za pomocą których można usiłować uzyskać ocenę sposobu oddziaływania warunków zewnętrznych i wewnętrznych na pomiar. Grupa Robocza mogłaby się zastanowić nad tym, czy można zaproponować kilka ogólnych reguł dotyczących tych metod, mimo iż szczegółowa dyskusja nie jest w ogóle możliwa, metody te bowiem różnią się bardzo w poszczególnych dziedzinach nauki. Tym niemniej po przeprowadzeniu dodatkowych kontroli i zastosowaniu odpowiednich poprawek do wyniku można znaleźć jeszcze niepewności resztkowe, a mianowicie niepewności systematyczne, które trzeba oszacować i wyliczyć.

Istnieje pewien obiektywny sposób oznaczania takich niepewności; polega on na założeniu granic przedziału niepewności tak szerokich, że rozważane oddziaływanie nie może w

zaden sposób wpłynąć na wynik o największej wartości niezależnie od okoliczności. Jednakże nie jest to to o, co chodzi, gdyż rozszerza się bezużytecznie granice przedziału niepewności /rozpatrując to z jakiegokolwiek praktycznego punktu widzenia dotyczącego porównania wyników/. Od eksperymentatora oczekuje się tego, że wyznaczy on granice jak najwęższe, co do których można być przekonany, że są racjonalne. Granice przedziału niepewności systematycznej estymuje się więc odwołując się do oceny subiektywnej. I ponieważ chodzi o ocenę subiektywną, jedynym sposobem przedstawienia kryteriów przyjętych przy oszacowaniu niepewności jest posłużenie się wyrażeniem subiektywnym.

Chociaż powszechnie stosowane wyrażenie "maksymalne granice błędu" sprawia wrażenie wyrażenia obiektywnego, jest ono bardzo subiektywne, ale implikuje oszacowanie bez ryzyka, w granicach dostatecznie szerokich, być może porównywalnych z granicami przedziału ufności "99,9%". Z drugiej strony sugerowane wyrażenie "tyleż prawdopodobne, co nieprawdopodobne" implikuje granice wąskie, porównywalne w pewnym sensie z "błędem prawdopodobnym" w statystyce, tj. o granicach przedziału ufności "50%", gdy tymczasem inne sugerowane wyrażenie "dwa razy bardziej prawdopodobny niż nieprawdopodobny" byłoby porównywalne z granicami przedziału ufności "67%" lub z odchyleniem standardowym wartości średniej.

W pewnych wydziałach NPL jest używany termin "pewny ponad rozsądne wątpliwości"; jest to oczywiście kryterium stosowane przy wydawaniu werdyktów sądu przysięgłych w angielskim systemie prawnym; można uważać, że odpowiada on przedziałowi ufności 99%. Gdyby Grupa Robocza zgodziła się na ten sposób oceny błędów systematycznych, można by przedyskutować:

- a/ wybór subiektywnego poziomu prawdopodobieństwa /porównaj również odpowiedź na pytanie 7/,
- b/ wybór wyrażen, które można by wykorzystać do opisu tego subiektywnego poziomu prawdopodobieństwa.

Pytanie 5. Czy należy zalecać praktyczną regułę łączenia błędów systematycznych?

Jest co najmniej pięć różnych metod stosowanych zwykle w celu łączenia niepewności systematycznych:

- a/ dodawanie liniowe niepewności /znane większości eksperymentatorów jako oszacowanie z nadmiarem/;
- b/ sumowanie kwadratowe niepewności /według mniemania wielu eksperymentatorów jest to oszacowanie z niedomiarem/;
- c/ dodawanie liniowe jednego lub dwu składników niepewności znacznie większych od pozostałych do kwadratowej sumy mniejszych składników;
- d/ wybór wiarygodnej wartości pośredniej pomiędzy sumą liniową a sumą kwadratową;
- e/ stosowanie formuły matematycznej przy założeniu pewnego rozkładu prawdopodobieństwa błędu systematycznego w celu obliczenia wartości porównywalnej z jakimś parametrem statystycznym, takim jak odchylenie standardowe wartości średniej.

Potrzebna jest przekonywująca teoria, oparta na pojęciu niepewności /a nie błędu/, z której można by wyprowadzić metodę łączenia niepewności systematycznych. Jeśli okazałoby się to niemożliwe, to inne rozwiązanie polegałoby na przeprowadzeniu wyraźnej klasyfikacji i na opisaniu różnych metod wraz z kilkoma wskazówkami dotyczącymi zalet i obszarów zastosowania każdej z nich.

Zagadnienie to nie może być oddzielone od zagadnienia łączenia niepewności systematycznych i losowych /por. również odpowiedź na pytanie 7/.

Pytanie 6. Czy należy zalecać praktyczną regułę łączenia błędów przypadkowych i systematycznych?

Większość zainteresowanych uznałaby, że niepewności losowych i systematycznych nie należy łączyć, chyba że i jedno i drugie są estymowane na porównywalnym poziomie prawdopodobieństwa /prawdopodobieństwa subiektywnego w przypadku niepewności systematycznych/. Zagadnienie to nie może być rozpatrywane efektywnie bez uwzględnienia następnego pytania.

• Pytanie 7. Czy należy zalecać praktyczną regułę wyrażenia niepewności wynikowej?

W przypadku pomiarów o bardzo wysokim poziomie dokładności należy - zgodnie z sugestią zawartą w kwestionariuszu - błędy przypadkowe i systematyczne wykazywać oddzielnie. W tym przypadku powinno się sporządzać listę indywidualnych oszacowań dla każdego składnika niepewności; jeżeli wszystkie te informacje są dane, to użytkownik raportu może je łączyć wg dowolnej metody, która jest według niego odpowiednia.

Potrzebę wyrażania niepewności całkowitej odczuwa się przede wszystkim na niższych poziomach hierarchii wzorcowań; można postawić rozsądne pytanie, jakie jest praktyczne wykorzystanie takiej informacji. Odpowiedź w większości przypadków jest taka, że wykorzystuje się je do celów prawnych lub pseudonaukowych. Weźmy, na przykład, przypadek producenta, który kupuje drogi przyrząd pomiarowy, specjalnie wzorcowany w laboratorium metrologicznym drugiego rzędu i który wykorzystuje z kolei ten przyrząd jako etalon roboczy do produkcji dużej liczby elementów, których wymiary są krytyczne podczas użytkowania. Jeżeli uwierzytelnienie tego roboczego wzorca wykaże błąd 1% /tj. odchylenie 1% od innego wzorca roboczego/ i jeżeli certyfikat wzorcowania tego przyrządu wykazuje niepewność całkowitą  $\pm 0,1\%$ , to producent ma prawo poczuć się poszkodowany. I wtedy małe znaczenie ma dla jednej lub drugiej strony informacja o tym, że z tej wartości 0,1% przypada 0,07% na niepewność losową, a 0,07% na niepewność systematyczną i że połączono je przez "dodawanie kwadratowe".

Najczęściej producent wybiera przyrząd pomiarowy o niepewności znacznie mniejszej niż tolerancje wymagane w procesie produkcyjnym; w takim przypadku nie mają praktycznego znaczenia metody wykorzystywane do łączenia i określania niepewności losowych oraz systematycznych. Jednakże taki wybór nie jest zawsze możliwy i wtedy gdy tolerancje wymagane w procesie produkcyjnym zbliżają się do niepewności wzorcowania informacja o niepewności staje się istotna.

Uważa się często, że granice ufności o poziomie prawdopodobieństwa 99% są odpowiednie tylko przy pomiarach wykonywanych w laboratoriach pierwotnych, że gdzie indziej granice przedziału ufności 95%, a nawet 90% są do przyjęcia i że niepewności systematyczne powinny być estymowane przy tym samym subiektywnym poziomie prawdopodobieństwa. Jednakże oznacza to w zasadzie, że przeciętnie jeden przyrząd na dwadzieścia /albo jeden na dziesięć przy poziomie prawdopodobieństwa 90%/ będzie dawał poprawny wynik wykraczający /i to nie wiadomo ile/ poza granice niepewności podanego współczynnika wzorcowania. Takie prawdopodobieństwo błędu może być w pewnych przypadkach nie do przyjęcia dla producenta. Mogłoby to oznaczać, że niepewności systematyczne powinny być estymowane na podstawie takiego kryterium jak "maksymalne granice błędu" lub "pewny ponad rozsądne wątpliwości" oraz, że powinny być łączone z granicami przedziałów ufności oszacowanymi przy poziomie 99%.

Stanowi to wskazówkę dotyczącą odpowiedzi na pytania 2 i 4. Pozostaje pewna wątpliwość, czy to ma implikować, że trzeba łączyć niepewności systematyczne liniowo; w każdym razie Grupa Robocza powinna przestudiować to zagadnienie. Należałoby również przedyskutować sposoby łączenia niepewności /bezpośrednio lub obliczając odchylenie standardowe średniej i wielkości porównywalnej z niepewnością systematyczną - por. 5e powyżej/.

Jest też pewna delikatna sprawa związana z niepewnościami systematycznymi, którą, być może, Grupa Robocza będzie chciała przestudiować. Rozważmy przypadek eksperymentu, dla

którego indywidualne niepewności systematyczne zostały oszacowane przy "subiektywnym" poziomie ufności 99%. Przyjmując teraz, że ktoś chciałby potem wykorzystać wynik tego eksperymentu w połączeniu z innym eksperymentem, dla którego niepewności systematyczne zostały oszacowane przy "subiektywnym" poziomie ufności 67%. Czy można by zaproponować jakąś metodę, za pomocą której udałoby się uczynić pierwszą grupę niepewności "kompatybilną" z drugą grupą?

Podobnym zagadnieniem jest łączenie niepewności losowych związanych z tymi doświadczeniami, których wyniki mają być wykorzystane do innych doświadczeń. Czy trzeba traktować niepewności losowe tych pierwszych doświadczeń jako niepewności losowe, czy jako systematyczne w drugim doświadczeniu? Byłoby również interesujące zbadać, czy nowoczesna teoria informacji może dać pewną koncepcję dotyczącą ilości traczonej informacji w przypadku, gdyby szczegółowe zestawienie granic przedziału niepewności zostało skondensowane, dając w rezultacie granice jedynej globalnej niepewności.

Problemem mniej ważnym, jednak istotnym dla praktyki, jest uwzględnianie odczytów bardzo oddalonych od średniej.

## 7. INFORMACJE UZUPELNIAJĄCE

W publikacji N<sup>o</sup> 3003 Brytyjskiej Służby Metrologicznej "The expression of uncertainty in electrical measurements" /1977/ /"Wyrażenie niepewności w pomiarach elektrycznych"/ porażono metody wykorzystywane w służbie sprawdzeń /uwierzytelniania/. Zaznaczono zresztą, że wydaje się mało prawdopodobne znalezienie jednej metody, która byłaby odpowiednia na wszystkich poziomach hierarchii sprawdzeń. Przytacza się ten dokument jako przykład związanej /z konieczności/ formy przedstawiania niepewności, formy, która powinna być wykorzystywana na poziomie dość odległym od szczytu hierarchii; nie należy traktować tego dokumentu jako wskazania dotyczącego "polityki" NPL w tej dziedzinie.

ISSN 0209-1046

