

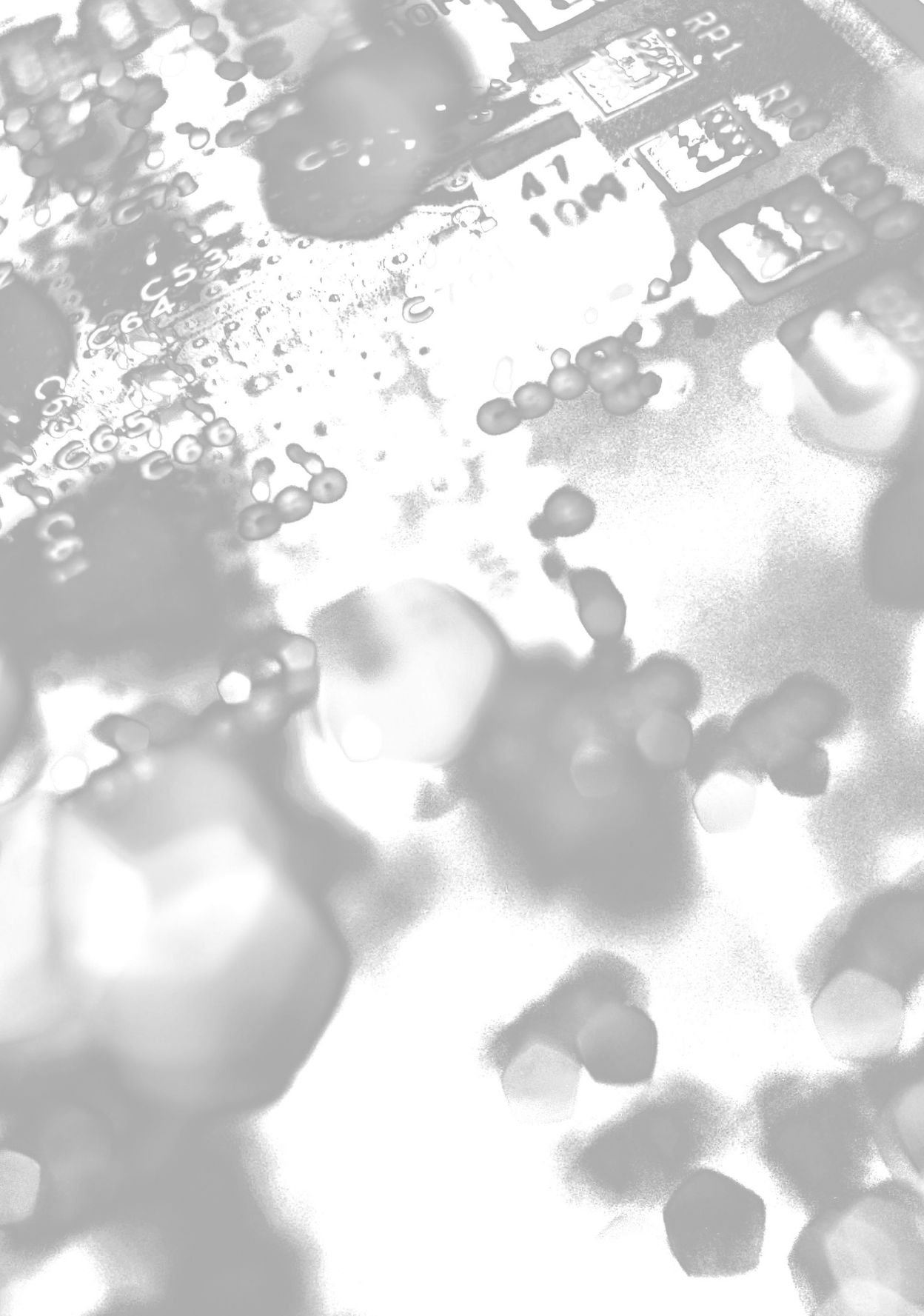
JAMES MILLER JANE MILLER

Statystyka i chemometria

w chemii analitycznej



PWN



Dane oryginału

Authorized translation from the English language edition, entitled: Statistics and Chemometrics for Analytical Chemistry; ISBN 978-0-273-73042-2; by James N. Miller, Jane C. Miller; published by Pearson Education Limited.

Copyright © Ellis Horwood Limited 1993

Copyright © Pearson Education Limited 2000, 2010

This translation of Statistics and Chemometrics for Analytical Chemistry 6/e is published by arrangement with Pearson Education Limited.

All rights reserved. No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage and retrieval system, without written permission from Pearson Education, Inc. Polish language edition published by Wydawnictwo Naukowe PWN SA. Copyright © 2016

Wszystkie prawa zastrzeżone. Żaden fragment tej książki nie może być kopiowany lub przekazywany w jakiegokolwiek formie, elektronicznie lub mechanicznie, w tym za pomocą fotokopii, nagrań albo zapamiętany w jakimkolwiek systemie wyszukiującym, bez zgody Pearson Education Limited.

Przekład: Jan Mazerski

Projekt okładki i stron tytułowych: Maryna Wiśniewska

Wydawca: Katarzyna Włodarczyk-Gil

Dyrektor Pionu Produktów i Usług: Sylwia Krawczyk

Menedżer Pionu Wydawniczego: Emilia Leśniewska

Redaktor prowadzący: Renata Ziółkowska

Redaktor: Ewa Czarnecka-Żołek

Produkcja: Mariola Grzywacka

Skład i łamanie: Grafini

Copyright © for the Polish edition by Wydawnictwo Naukowe PWN SA
Warszawa 2016

ISBN 978-83-01-18303-5

Wydanie I
Warszawa 2016

Wydawnictwo Naukowe PWN SA
02-460 Warszawa, ul. Gottlieba Daimlera 2
tel. 22 69 54 321, faks 22 69 54 288
infolinia 801 33 33 88
e-mail: pwn@pwn.com.pl; www.pwn.pl

Druk i oprawa: OSDW Azymut Sp z o.o

Spis treści

Przedmowa do pierwszego wydania	IX
Przedmowa do wydania szóstego	XI
Stosowane symbole	XIII
1 Wstęp	1
1.1 Problemy analityczne	1
1.2 Błędy w analizie ilościowej	2
1.3 Rodzaje niepewności pomiarowej	4
1.4 Błędy losowe i systematyczne w analizie miareczkowej	8
1.5 Sposób postępowania z błędami systematycznymi	12
1.6 Planowanie doświadczeń i pomiarów	16
1.7 Zastosowanie kalkulatorów i komputerów osobistych w obliczeniach statystycznych	18
Bibliografia i zasoby	20
Ćwiczenia	22
2 Analiza statystyczna powtarzanych pomiarów	23
2.1 Średnia i odchylenie standardowe	23
2.2 Rozmieszczenie powtórzeń na osi liczbowej	26
2.3 Rozkład logarytmiczno-normalny	31
2.4 Definicja „próby statystycznej”	32
2.5 Rozkład średniej z próby	33
2.6 Granice przedziału ufności średniej z dużych prób	34
2.7 Granice przedziału ufności średniej z małych prób	36
2.8 Prezentacja wyników	38
2.9 Dalsze zastosowania granic przedziału ufności	39
2.10 Przedział ufności dla średniej geometrycznej populacji o rozkładzie logarytmiczno-normalnym	40
2.11 Propagacja błędów losowych	41
2.12 Propagacja błędów systematycznych	45
Bibliografia	46
Ćwiczenia	46

3	Testy istotności	48
3.1	Wstęp	48
3.2	Porównanie średniej z pomiarów ze wzorcem	49
3.3	Porównanie dwóch średnich pochodzących z pomiarów	51
3.4	Test t dla par wiązanych	56
3.5	Testy jedno- i dwustronne	58
3.6	Test F Snedecora dla porównania odchyłeń standardowych	60
3.7	Wyniki odbiegające	62
3.8	Analiza wariancji	67
3.9	Porównanie kilku średnich	68
3.10	Przebieg obliczeń w teście ANOVA	72
3.11	Test chi-kwadrat (χ^2)	75
3.12	Sprawdzenie normalności rozkładu	78
3.13	Wnioski z testów istotności	83
3.14	Statystyka bayesowska	85
	Bibliografia	87
	Ćwiczenia	88
4	Jakość pomiarów analitycznych	93
4.1	Wstęp	93
4.2	Pobieranie próbek	94
4.3	Zastosowanie ANOVA do niezależnego oszacowania wariancji pochodzącej z różnych źródeł	96
4.4	Strategia pobierania próbek	97
4.5	Wprowadzenie do metod kontroli jakości	99
4.6	Karty Shewharta do oceny wartości średnich	100
4.7	Karty Shewharta do oceny zakresu	102
4.8	Określanie zdolności procesu, σ	105
4.9	Średnia długość przebiegu, karty sum skumulowanych	108
4.10	Strefowe karty kontroli	112
4.11	Schematy testowania biegłości	114
4.12	Porównania międzylaboratoryjne	118
4.13	Niepewność	124
4.14	Akceptujące pobieranie próbek	129
4.15	Walidacja metod analitycznych	131
	Bibliografia	134
	Ćwiczenia	135
5	Metody kalibracji w analizie instrumentalnej – regresja i korelacja	139
5.1	Wstęp. Analiza instrumentalna	139
5.2	Wykresy kalibracyjne w analizie instrumentalnej	141
5.3	Współczynnik korelacji liniowej	144
5.4	Linia regresji y od x	149
5.5	Błędy współczynnika kierunkowego i wyrazu wolnego prostej regresji	150

5.6	Obliczanie stężenia i jego błędu losowego	154
5.7	Granica wykrywalności	157
5.8	Metoda dodatku wzorca	161
5.9	Wykorzystanie linii regresji do porównania metod analitycznych	164
5.10	Ważona linia regresji	170
5.11	Przecięcie dwóch linii prostych	176
5.12	Analiza wariancji i obliczenia regresyjne	177
5.13	Wprowadzenie do metod regresji krzywoliniowej	179
5.14	Dopasowanie krzywej	183
5.15	Punkty odbiegające w regresji	187
	Bibliografia	190
	Ćwiczenia	190
6	Metody nieparametryczne i odporne	193
6.1	Wstęp	193
6.2	Mediana i wstępna analiza danych	194
6.3	Test znaków	200
6.4	Test serii Walda–Wolfowitza	203
6.5	Test rangowanych znaków Wilcoxon	205
6.6	Proste testy dla dwóch niezależnych prób	207
6.7	Testy nieparametryczne dla więcej niż dwóch prób	211
6.8	Korelacja rang	214
6.9	Nieparametryczne metody regresyjne	216
6.10	Wstęp do metod odpornych	219
6.11	Proste metody odporne. Równoważenie (trymowanie) danych i winsoryzacja	220
6.12	Dalsze odporne oszacowania położenia i rozproszenia	221
6.13	Odporna analiza wariancji	224
6.14	Odporne metody regresyjne	225
6.15	Statystyki wielokrotnego próbkowania	227
6.16	Podsumowanie	229
	Bibliografia i zasoby	230
	Ćwiczenia	231
7	Projektowanie doświadczeń i optymalizacja	233
7.1	Wstęp	233
7.2	Planowanie losowe i blokowe	235
7.3	Dwuczynnikowa analiza wariancji	237
7.4	Kwadraty łacińskie i inne plany	241
7.5	Interakcje	242
7.6	Identyfikacja istotnych czynników. Plan czynnikowy	248
7.7	Ułamkowe plany czynnikowe	254
7.8	Optymalizacja. Podstawowe założenia i metody jednej zmiennej	258
7.9	Optymalizacja naprzemienna (metoda Gaussa)	261
7.10	Metoda najszybszego wzrostu (gradientowa)	264

7.11	Optymalizacja metodą simpleksów	267
7.12	Symulowane wyżarzanie	272
	Bibliografia i zasoby	273
	Ćwiczenia	273
8	Analiza wielowymiarowa	277
8.1	Wstęp	277
8.2	Kontrola i wstępna analiza danych wielowymiarowych	279
8.3	Analiza głównych składowych (PCA)	281
8.4	Analiza wiązkowa	286
8.5	Analiza dyskryminacyjna (klasyfikacja)	291
8.6	Metoda K najbliższych sąsiadów, KNN	296
8.7	Oddzielne modele klas	298
8.8	Metody regresyjne	298
8.9	Regresja liniowa wielu zmiennych	300
8.10	Regresja głównych składowych (PCR)	302
8.11	Cząstkowa metoda najmniejszych kwadratów (PLS)	305
8.12	Naturalne metody obliczeniowe. Sztuczne sieci neuronowe (ANN)	308
8.13	Podsumowanie	311
	Bibliografia i zasoby	311
	Ćwiczenia	312
	Rozwiązania ćwiczeń	315
	Załącznik 1: Powszechnie stosowane statystyczne testy istotności	328
	Załącznik 2: Tablice statystyczne	331
	Indeks	340

Przedmowa do wydania szóstego

Od publikacji piątego wydania tej książki w 2005 roku zastosowanie podstawowych i zaawansowanych metod statystycznych w nauczaniu i praktyce analitycznej znacząco się zwiększyło pod względem zasięgu i jakości. W nowym wydaniu staraliśmy się zachować równowagę między nowymi wyzwaniem, wynikającymi z tego rozwoju, rozszerzając zawartość niektórych rozdziałów i zachowując równocześnie podstawowe założenie tego podręcznika, polegające na pragmatycznym i jeśli to tylko możliwe, niematematycznym podejściu do obliczeń statystycznych.

Wyniki pomiarów analitycznych powszechnie interpretuje się, korzystając z dobrze ugruntowanych testów istotności. Jednak, w ostatnich latach coraz większe zastosowanie znajdują metody bazujące na regule Bayesa, szczególnie w takich dziedzinach jak kryminalistyka i chemia medyczna. Podstawowe założenia i metodologia metod statystycznych opartych na tej regule odbiegają znacząco od tradycyjnego podejścia statystycznego. Dlatego zdecydowaliśmy się poszerzyć rozdział 3 o nowy podrozdział poświęcony tym zagadnieniom.

Jakość wyników analitycznych uzyskiwanych przez różne laboratoria w przypadku tych samych próbek jest z oczywistych, praktycznych względów bardzo ważna i stanowi ciągle przedmiot zainteresowania. Takie studia porównawcze są znaczącym elementem postępowania walidacyjnego poprzedzającego wprowadzenie danej metody w poszczególnych laboratoriach analitycznych. Dlatego rozdział 4 rozszerzono o nowy podrozdział poświęcony sposobom walidacji metod analitycznych. Powszechnie stosowana forma porównań międzylaboratoryjnych oraz schematy testowania doprowadzają nas często do dziwnych i niespodziewanych wyników. Są one obecnie analizowane z wykorzystaniem tzw. metod odpornych (*robust statistical methods*). W rozdziale 6 rozbudowano opis niektórych z tych metod.

Oszacowanie niepewności pomiarowej stanowi powszechnie akceptowany element wielu analiz. Dlatego na oszacowaniu udziału niepewności pochodzącej z takich źródeł jak pobieranie próbek trzeba skupiać wielki wysiłek. Omówiono to zagadnienie w rozdziale 4. Metody kalibracji stanowią rdzeń wielu nowoczesnych metod analitycznych. W rozdziale 5 rozbudowaliśmy opis tradycyjnej metody dodatków wzorca o opis metody regresji ważonej i metod regresyjnych, w których uwzględnia się, że zarówno wartości na osi x , jak i y mogą zawierać losową niepewność pomiarową.

Zagadnieniem, któremu w laboratoriach analitycznych nie poświęca się tyle uwagi, na ile ono zasługuje, jest właściwe wykorzystanie planowania doświadczeń. W tej dziedzinie stosuje się specyficzne słownictwo i podejścia typowe raczej w metodach *a posteriori*. Są to prawdopodobnie przyczyny, dla których planowanie doświadczeń jest zagadnieniem pomijanym w typowych pracach analitycznych. Jednak, wiele planów doświadczeń jest stosunkowo prostych i istnieje sporo doskonałych programów komputerowych wspomagających praktyczne zastosowanie tego podejścia. To skłoniło nas do znacznego powiększenia rozdziału 7 o zagadnienia planowania doświadczeń. Obecnie stosuje się nowe i bardzo zaawansowane metody analizy danych wielowymiarowych, w tym również do typowych zagadnień analitycznych. Problem ten zasługuje na osobne opracowanie, jednak w tym wydaniu podręcznika zdecydowaliśmy się jedynie na spore rozszerzenie rozdziału 8 opisującego te metody.

W tej publikacji kontynuujemy zwyczaj uzupełniania tekstu wieloma przykładami obliczeń przeprowadzanych z wykorzystaniem dwóch powszechnie używanych programów Excel® i Minilab®. Pierwszy z nich jest dostępny w prawie wszystkich komputerach i stosowany do gromadzenia i przetwarzania danych cyfrowych z przyrządów pomiarowych. Z kolei drugi znajduje często zastosowanie i w edukacji, i w pracach analityków praktyków. W każdym z programów proste obliczenia przedstawione w tej książce są łatwe do przeprowadzenia, a ich wyniki można czytelnie przedstawić. Ponadto wiele ogólnie dostępnych opracowań może służyć jako wprowadzenie do obu programów. Również odpowiednie makra i wtyczki w użyteczny sposób rozszerzające standardowe możliwości Excela® i Minilaba® są łatwo i bezpłatnie dostępne w sieci. Dzięki tym programom dysponuje się również możliwością graficznej prezentacji wyników, co stwarza możliwości łatwiejszego zrozumienia i interpretacji danych. Te dodatkowe możliwości użyto w „Instructors' Manual”, który jest dostępny, podobnie jak w poprzednich wydaniach. Umieszczono w nim również propozycje ćwiczeń dla grup studenckich, kompletny zestaw rysunków, które mogą być użyteczne do przygotowania prezentacji multimedialnych oraz pełne rozwiązania do ćwiczeń zawartych w książce. W samym podręczniku podano jedynie krótkie omówienie tych rozwiązań.

Jesteśmy bardzo zobowiązani licznym korespondentom, zarówno pracownikom nauki, jak i studentom, którzy nadesłali swoje konstruktywne uwagi i sugestie dotyczące treści zawartych w poprzednich wydaniach tej książki. Jesteśmy również wdzięczni za zwrócenie uwagi na niewielkie błędy i pomyłki, których nie udało się nam uniknąć. Składamy również podziękowania Królewskiemu Towarzystwu Chemicznemu za udzielenie zgody na skorzystanie z danych opublikowanych w *The Analyst*. Na końcu dziękujemy również Rufusowi Curnowowi i jego kołegom z Pearson Education oraz Nicolii Chilwersowi i Rosowi Woodwardowi za ich wyjątkowy zasób doświadczenia, cierpliwości i entuzjazmu. Wszelkie błędy, jakie pozostały w książce pomimo ich najlepszych chęci, obciążają tylko i wyłącznie nas.

James N. Miller
Jane C. Miller

grudzień 2009

1

Wstęp

Główne zagadnienia omawiane w tym rozdziale

- Błędy w pomiarach analitycznych
- Błędy ogólne, losowe i systematyczne
- Precyzja, powtarzalność, odtwarzalność, obciążenie i dokładność pomiaru
- Planowanie doświadczeń
- Stosowanie kalkulatorów i komputerów osobistych

1.1 Problemy analityczne

Chemik-analytyk spotyka się w praktyce z oznaczeniami jakościowymi i ilościowymi. Przykładem problemów jakościowych może być bardzo szkodliwa obecność boru w wodzie destylowanej używanej przy wytwarzaniu elementów mikroelektronicznych. W praktyce zagadnienie sprowadza się do odpowiedzi na pytanie, czy dana próbka wody destylowanej zawiera bor. Z kolei w kryminalistyce pojawia się często problem identyczności dwóch próbek gleby.

W innych przypadkach pojawiają się pytania natury ilościowej. Jaka jest zawartość albuminy w danej próbce osocza krwi? Jaka jest zawartość ołowiu w próbce wodzie wodociągowej? Badana stal zawiera niewielkie domieszki chromu, wolframu i manganu. Jaka jest zawartość każdej z domieszek? Są to typowe przykłady problemów jedno- lub wieloskładnikowej analizy ilościowej.

Współczesna chemia analityczna jest głównie nauką **ilościową**. W większości zagadnień odpowiedź ilościowa jest zdecydowanie bardziej wartościowa niż odpowiedź jakościowa. Analytyk może stwierdzić, że wykrył w próbce wody destylowanej obecność boru, ale ważniejsze jest ustalenie, *jakie stężenie* boru znajduje się w wodzie. Osoba zamawiająca analizę może teraz, uzbrojona w taką ilościową odpowiedź, podjąć decyzję, czy takie stężenie jest dopuszczalne, czy

trzeba rozważyć metody dalszego oczyszczania. Informacja o tym, że w próbce wody wykryto obecność boru, nie jest zwykle wystarczająca do oceny znaczenia tego faktu. W wielu przypadkach jedynie wynik ilościowy ma w ogóle jakiegokolwiek znaczenie. Na przykład, praktycznie wszystkie próbki osocza ludzkiej krwi zawierają albuminę. Tak więc istotna jest informacja, ile jej jest.

Nawet jeżeli potrzebna jest odpowiedź jakościowa, uzyskuje się ją w wyniku oznaczeń ilościowych. W praktyce analityk nigdy nie przedstawia swojego wyniku w formie stwierdzenia „Wykryłem (nie wykryłem) obecności boru w próbce wody”. Przyjmijmy, że stosujemy metodę pozwalającą na wykrycie obecności boru na poziomie powiedzmy $1 \mu\text{g/ml}$. Jeżeli test wypadnie negatywnie, poprawny wynik powinien mieć postać: „Ta próbka zawiera mniej niż $1 \mu\text{g/ml}$ boru”. Jeżeli test wypadnie pozytywnie, przedstawiamy wynik w postaci stwierdzenia, że próbka zawiera co najmniej $1 \mu\text{g/ml}$ boru. Do takiego stwierdzenia dodawane są jeszcze inne informacje, o których za chwilę. Podejście ilościowe może być również zastosowane do porównania dwóch próbek gleby. Próbki te mogą być na przykład poddane analizie wielkości cząstek. W analizie takiej określa się, jaki ułamek cząstek trafia do powiedzmy 10 przedziałów wielkości. Każda próbka jest wtedy opisana przez 10 wielkości ilościowych. Porównanie tych wielkości (rozdz. 8) może być użyte do ilościowej oceny podobieństwa próbek.

1.2 Błędy w analizie ilościowej

Ponieważ badania ilościowe odgrywają dominującą rolę w laboratorium analitycznym, musimy przyjąć, że błędy, jakie mogą wystąpić podczas oznaczeń ilościowych, mają decydujące znaczenie. Naszą podstawową zasadą powinno być stwierdzenie, że *wyniki ilościowe nie mają żadnej wartości, o ile nie towarzyszy im oszacowanie zawartych w nich błędów (oszacowanie ich niepewności)*. Ta zasada powinna obowiązywać nie tylko w analityce chemicznej, ale we wszystkich badaniach, których wyniki mają postać ilościową. Poniżej przedstawimy kilka przykładów ilustrujących powyższą zasadę. Zobaczymy również niektóre problemy statystyczne, z którymi spotykamy się w praktyce analitycznej i których rozwiązanie będzie przedstawione w kolejnych rozdziałach tego podręcznika.

Wyobraźmy sobie, że zsyntetyzowaliśmy, naszym zdaniem zupełnie nowy, odczynnik analityczny. Podczas badania metodą spektroskopową wynik pomiaru wyniósł 104 (normalnie wynik pomiaru powinien być podany w jednoznacznie zdefiniowanych jednostkach, ale w tym hipotetycznym przykładzie mogą być zastosowane całkowicie dowolne jednostki). Z danych literaturowych wynika, że w przypadku żadnego ze znanych związków nie uzyskuje się wyniku większego niż 100 przy zastosowaniu tej samej metody pomiarowej i w tych samych warunkach pomiaru. Pojawia się więc oczywiste pytanie, czy rzeczy-

wiecie odkryliśmy nowy związek chemiczny. Odpowiedź na to pytanie zależy w oczywisty sposób od tego, w jakim stopniu możemy polegać na wyniku pomiaru równym 104. Zawsze należy zapytać, jaki jest błąd (niepewność) tego wyniku pomiaru. Jeżeli z dalszych badań będzie wynikać, że niepewność wynosi 2 jednostki, tzn. prawdziwa wartość najprawdopodobniej leży w zakresie 104 ± 2 , istnieje duża szansa, że jest to rzeczywiście nowy związek. Jednak, jeżeli niepewność wyniesie 10 jednostek (tzn. zakres 104 ± 10), to jest wysoce prawdopodobne, że prawdziwa wartość jest mniejsza od 100 i nowe odkrycie jest dalekie od pewności. Tak więc wiedza o niepewności pomiaru jest krytyczna dla poprawnej interpretacji wyniku pomiaru (tego z naszego przykładu jak i każdego innego). Ze statystycznego punktu widzenia powyższy przykład należy do problemu porównania wyniku pomiaru (104) z wartością referencyjną (100). Zagadnienie to będzie omówione ze szczegółami w rozdziale 3.

Analitik zwykle wykonuje kilka pomiarów (powtórzeń) w ramach pojedynczego doświadczenia. Wartość i znaczenie takich powtórzeń będą szczegółowo omówione w następnym rozdziale. Załóżmy, że analitik wykonał 4-krotnie miareczkowanie tej samej próbki i otrzymał wyniki 24,69, 24,73, 24,77 i 25,39 ml. Warto zwrócić uwagę, że wyniki miareczkowania podane są w zaokrągleniu do drugiego miejsca po przecinku. Ten problem będzie również przedyskutowany w rozdziale 2. Wszystkie cztery wartości są różne na skutek nieuniknionej niepewności pomiarowej. Ponadto czwarta wartość (25,39 ml) wydaje się być zdecydowanie inna niż pozostałe trzy. Czy w tej sytuacji możemy ją bezpiecznie odrzucić i przedstawić wynik miareczkowania jako równy 24,73 ml (wartość średnia z trzech pomiarów)? Czy z punktu widzenia statystyki wartość 25,39 ml może być traktowana jako *punkt odbiegający*? Najważniejsze zagadnienia dotyczące punktów odbiegających i ich odrzucania są omówione w rozdziałach 3 i 6.

Inny często spotykany problem dotyczy porównania dwóch (lub więcej) zbiorów pomiarów. Przypuśćmy, że analitik oznacza zawartość wanadu w próbce stali, stosując dwie niezależne metody. Według pierwszej metody wynik średni wynosi 1,04% z oszacowaniem niepewności pomiarowej na poziomie 0,07%. W drugiej metodzie otrzymano wartość średnią 0,95% z niepewnością 0,04%. Powstaje pytanie, czy te metody dają znacząco różne wyniki, czy też są one nierozróżnialne ze względu na niepewność pomiarową. Kolejne pytanie dotyczy niepewności pomiarowej obu metod: czy różnią się one od siebie w sposób znaczący? No i najważniejsze pytanie: która wartość średnia jest bliższa prawdziwej zawartości wanadu w próbce stali? Odpowiedzi na te i podobne pytania zostaną przedstawione również w rozdziale 3.

W wielu metodach analizy instrumentalnej bazuje się na graficznej interpretacji wyników. W takim przypadku zamiast wykonać serię pomiarów tej samej próbki, wykonuje się pomiary niewielkiej grupy standardowych próbek. W tych próbkach jest znana zawartość analitu, a przedział zawartości obejmuje spodziewaną zawartość analitu w badanej próbce. Z wyników uzyskanych w przypadku próbek standardowych tworzy się teraz krzywą kalibracyjną, na podsta-

wie której ustala się na drodze interpolacji zawartość analitu w próbce badanej. Warto zdać sobie sprawę, że wszystkie wykonane pomiary (zarówno próbek standardowych, jak i badanych) są obciążone niepewnością pomiarową. Dlatego należy oszacować niepewność uzyskanej krzywej kalibracyjnej oraz niepewność odczytu na jej podstawie zawartości analitu w pojedynczej próbce testowej. Zalecane jest także oszacowanie granicy wykrywalności, czyli najmniejszej zawartości analitu, jaka może być wykryta przy zadanym poziomie ufności. Szczegółowy tok postępowania będzie omówiony w rozdziale 5.

Omówione przykłady stanowią jedynie niewielką część problemów wynikających z obecności niepewności pomiarowej w analizie ilościowej. Ale problemy te muszą zostać pokonane, jeżeli wyniki pomiarów mają mieć jakąkolwiek rzeczywistą wartość. Nasze rozważania rozpoczniemy od szczegółowego omówienia różnych rodzajów niepewności pomiarowej i ich źródeł.

1.3 Rodzaje niepewności pomiarowej

W badaniach doświadczalnych rozróżnia się trzy podstawowe rodzaje błędów: błędy **grube**, błędy **losowe** (przypadkowe) i błędy **systematyczne**. Jako tzw. błędy grube definiuje się wyniki tak dalece odbiegające od rzeczywistości, że jedynym rozsądnym rozwiązaniem jest pominięcie uzyskanego wyniku i powtórzenie pomiaru od początku. Przykładowo, źródłem takich błędów może być:

- awaria aparatury pomiarowej,
- rozlanie lub zniszczenie kluczowej próbki,
- stwierdzenie w trakcie doświadczenia, że odczynnik o wymaganej dużej czystości jest w rzeczywistości silnie zanieczyszczony.

Błędy tego typu (które zdarzają się od czasu do czasu nawet w najlepszych laboratoriach) są zwykle łatwe do wykrycia. Ponieważ wyniki obciążone błędem grubym można jedynie odrzucić, nie będziemy się nimi dalej zajmować. W poniższej dyskusji skupimy uwagę na różnicach pomiędzy błędami **losowymi** i **systematycznymi**.

Różnicę pomiędzy naturą tych błędów najlepiej poznać, analizując starannie jakiś konkretny przykład. Załóżmy, że każdy z czterech studentów (oznaczymy ich literami od A do D) wykonuje doświadczenie polegające na miareczkowaniu *dokładnie* 10,00 ml *dokładnie* 0,1 M roztworu wodorotlenku sodu za pomocą *dokładnie* 0,1 M roztworu kwasu solnego. Każdy ze studentów powtarza miareczkowanie 5 razy. Wyniki uzyskane przez studentów przedstawiono w tabeli 1.1.

Wyniki uzyskane przez studenta A posiadają dwie cechy. Po pierwsze, mają bardzo podobne wartości: wszystkie wyniki znajdują się w przedziale od 10,08 do 10,12 ml. W języku potocznym można powiedzieć, że wyniki są bardzo *po-*

Tabela 1.1 Błędy losowe i systematyczne

Student	Wynik pomiaru, ml					Uwagi
	10,08	10,11	10,09	10,10	10,12	
A	10,08	10,11	10,09	10,10	10,12	Dokładny, obciążony
B	9,88	10,14	10,02	9,80	10,21	Niedokładny, nieobciążony
C	10,19	9,79	9,69	10,05	9,78	Niedokładny, obciążony
D	10,04	9,98	10,02	9,97	10,04	Dokładny, nieobciążony

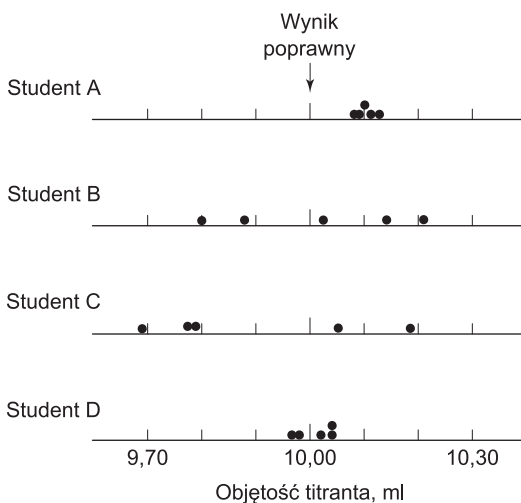
wtarzalne. Drugą cechą tych wyników jest to, że *wszystkie* one są *zbyt duże*. W tym hipotetycznym doświadczeniu (odmiennie niż w rzeczywistości) wiadomo, że prawdziwy wynik powinien wynosić *dokładnie* 10,00 ml. Widać, że na wynik pojedynczego pomiaru mają wpływ dwa typy błędów. Pierwszym z nich jest **błąd losowy** – w jego efekcie poszczególne powtórzenia różnią się między sobą i lokują się po obu stronach wartości średniej (10,10 ml w tym przypadku). Błąd losowy wpływa na **precyzję** (*precision*), czyli **powtarzalność** (*repeatability*) pomiaru. W przypadku studenta A widać wyraźnie, że błąd losowy jest niewielki, a tym samym wyniki są **precyzyjne**. Jednakże wyniki te zawierają również **błąd systematyczny** – jego wpływ przejawia się przesunięciem wyników *w tym samym kierunku* (w naszym przykładzie wszystkie wartości są za duże). Należy zauważyć, że dane doświadczenie obarczone może być kilkoma błędami systematycznymi. Niektóre z nich powodują przesunięcia wyników w kierunku wartości większych, a inne przeciwnie (rozdział 2). Suma wszystkich błędów systematycznych nosi nazwę **obciążenia** (*bias*) danego pomiaru. Brak obciążenia bywa czasami nazywany **prawdziwością** (*trueness*) pomiaru (podrozdział 4.15). Błędy losowy i systematyczny są łatwe do rozróżnienia przy analizie wyników. Mogą one mieć również odmienne przyczyny w zależności od techniki pomiarowej i używanej aparatury (podrozdział 1.4).

Wyniki uzyskane przez studenta B mają odmienną charakterystykę. Wartość średnia pięciu powtórzeń wykonanych przez studenta B (10,01 ml) jest bardzo bliska wartości prawdziwej. Nie ma więc żadnych przesłanek sugerujących istnienie obciążenia wyników (brak błędu systematycznego). Jednak rozrzut wyników jest bardzo duży, wskazując na małą ich precyzję (duży błąd losowy). Z porównania wyników uzyskanych przez studentów A i B widać jasno, że występowanie błędów losowych i systematycznych jest od siebie niezależne. Stwierdzenie to znajduje poparcie w wynikach studentów C i D. Wyniki studenta C charakteryzują się niską precyzją (zakres od 9,69 do 10,19 ml), a średnia wyników (9,90 ml) znacznie mniejsza od wartości prawdziwej wskazuje, że są one obciążone obciążeniem ujemnym. Z kolei student D uzyskał wyniki precyzyjne (zakres od 9,97 do 10,04 ml) i nieobciążone (średnia 10,01 ml). Różnice pomiędzy błędami systematycznymi i losowymi zostały podsumowane w tabeli 1.2.

Porównanie wyników uzyskanych przez poszczególnych studentów pokazano na rysunku 1.1 w postaci zestawu tzw. wykresów rozproszenia (*dot-plots*). Wykresy rozproszenia, na których poszczególne wyniki pomiarów przedsta-

Tabela 1.2 Porównanie cech błędu losowego i systematycznego

Błąd losowy	Błąd systematyczny
Ma wpływ na precyzję – powtarzalność lub odtwarzalność	Wytwarza obciążenie – ogólne odchylenie wyników od wartości prawdziwej, nawet gdy błąd losowy jest bardzo mały
Powoduje, że powtarzane pomiary lokują się po obu stronach wartości średniej	Przesuwa wszystkie wyniki w tę samą stronę – wszystkie za duże lub za małe
Może być oszacowany na podstawie powtórzeń	Nie może być wykryty na podstawie powtórzenia pomiarów
Może być zmniejszony w dobrych technikach pomiarowych, ale nie może być całkowicie wyeliminowany	Można go skorygować (wylimitować), stosując standardowe techniki lub materiały odniesienia
Źródłem jest czynnik ludzki i aparaturowy	Źródłem jest czynnik ludzki i aparaturowy

**Rysunek 1.1** Obciążenie i precyzja – wykresy rozrzutu dla danych z tabeli 1.1

wione są w postaci punktów na liniowej osi liczbowej, są prosta metodą graficzną często stosowaną w analizie rozpoznawczej (*exploratory data analysis*, EDA) zwanej także wstępną kontrolą danych (*initial data analysis*, IDA). Przykłady zastosowania wykresów rozproszenia pokazano w rozdziałach 3 i 6.

W większości pomiarów analitycznych najważniejszym problemem jest ustalenie, jak bardzo uzyskany wynik odbiega od rzeczywistego stężenia lub zawartości analitu. Problem ten nosi nazwę **dokładności** (*accuracy*) pomiaru. Dokładność pomiaru jest definiowana przez ISO jako „stopień zgodności między wynikiem pomiaru a akceptowalną wartością odniesienia” dla danego analitu. Zgodnie z tą definicją dokładność pojedynczego pomiaru zależy

zarówno od błędów losowych, jak i systematycznych. Dokładność wyniku średniego zależy również od obu źródeł błędów. Co więcej, nawet jeżeli brak jest błędu systematycznego, wynik średni rzadko kiedy jest idealnie zgodny z wartością odniesienia. Wynika to z obecności błędu losowego (niepewności pomiarowej), patrz rozdział 2 i 3. Wyniki studenta B są potwierdzeniem tej reguły. Cztery spośród pięciu pomiarów tego studenta mocno odbiegają od wartości prawdziwej 10,00 ml. Pomimo to wartość średnia ze wszystkich wyników, 10,01 ml, jest bardzo dokładna. Dowodzi to, że niedokładność pojedynczych wyników jest rezultatem błędów losowych, a nie systematycznych. Całkiem inaczej jest w przypadku wyników studenta A: zarówno pojedyncze wyniki jak i ich średnia są niedokładne, chociaż cechują się dużą precyzją. Wskazuje to wyraźnie, że niedokładność wynika z błędów systematycznych. Należy w tym miejscu zdecydowanie podkreślić, że w analityce chemicznej (w odróżnieniu od definicji słownikowych) pojęcia „precyzja” i „dokładność” mają zdecydowanie różne znaczenie.

PODSUMOWANIE: precyzja jest miarą błędu losowego, obciążenie opisuje błąd systematyczny, a dokładność, czyli zgodność pojedynczego lub średniego wyniku z wartością prawdziwą zależy od obu typów błędów.

Innym istotnym zagadnieniem terminologicznym jest różnica między **odtwarzalnością** (*reproducibility*) a **powtarzalnością** (*repeatability*). Można to zilustrować, poszerzając nieco nasz poprzedni przykład. W typowej sytuacji na przykład student A może wykonać wszystkie pięć miareczkowań jedno po drugim. Powinno mu to zająć nie więcej niż około godziny. Będzie wtedy korzystał z tego samego zestawu roztworów i szkła laboratoryjnego. Również temperatura, wilgotność i inne warunki w laboratorium nie będą się drastycznie zmieniać. W takiej sytuacji zmierzona precyzja będzie precyzją *między kolejnymi pomiarami* (*within-run precision*). Będziemy ją nazywać **powtarzalnością**. Rozważmy teraz sytuację, że z jakichś przyczyn poszczególne miareczkowania będą wykonywane przez różne osoby, w różnych laboratoriach i z użyciem różnych zestawów roztworów i szkła laboratoryjnego. Nie powinno nas dziwić, że takie wyniki będą miały większy rozrzut. Taki zestaw danych zawiera w sobie informacje o precyzji między różnymi pomiarami (*between-run precision*), tzn. **odtwarzalności** pomiarów.

- **Powtarzalność** opisuje rozrzut wyników kolejnych pomiarów prowadzonych w tych samych warunkach.
- **Odtwarzalność** opisuje rozrzut wyników pomiarów prowadzonych w różnych warunkach.
- Należy oczekiwać, że odtwarzalność metody będzie gorsza (tzn. z większym błędem losowym) niż jej powtarzalność.

Jeszcze jeden wniosek można wyciągnąć z omawianych miareczkowań. Nie budzi wątpliwości, że wyniki uzyskane przez studenta C są nieakceptowalne, a te uzyskane przez studenta D są najlepsze. Czasami zdarza się, że istnieje możliwość wyboru pomiędzy dwiema metodami: jedną precyzyjną, ale prawdopodobnie obciążoną i drugą nieprecyzyjną, ale nieobciążoną. Odwołując się do powyższego przykładu, mamy możliwość wyboru pomiędzy rodzajem wyników uzyskanym przez studenta A a rodzajem uzyskanym przez studenta B. Który rodzaj wyników należy preferować?

Nie ma jednoznacznej odpowiedzi na tak postawiony problem. W praktyce wybór konkretnej metody pomiarowej zależy jeszcze od wielu innych czynników takich jak koszt, łatwość automatyzacji, szybkości analizy. Warto sobie również uświadomić, że metoda nieobciążona może, jeżeli ma bardzo małą precyzję, prowadzić do wyników średnich dalekich od wartości poprawnych. Z drugiej strony istnieje często możliwość zamiany metody precyzyjnej, lecz obciążonej (student A) w metodę precyzyjną i nieobciążoną (student D). Dzieje się tak, gdy poznamy i usuniemy przyczynę błędu systematycznego. Błędów losowych nigdy nie można wyeliminować. Można je jednak ograniczyć, stosując odpowiednie procedury postępowania i rozwiązania techniczne. Wielkość niepewności pomiarowej (wpływ błędów losowych) można oszacować, wykorzystując powtórzenia pomiarów. Można wtedy ocenić jej istotność w interpretacji wyników doświadczeń. Te podstawowe różnice między obu rodzajami błędów będziemy dalej poznawać w kolejnym podrozdziale.

Gdy laboratorium analityczne otrzymuje próbkę do analizy w celu ustalenia stężenia jednego z jej składników, powinno przede wszystkim określić wielkość typowych błędów losowych i systematycznych. Może to wymagać dodatkowych pomiarów lub być już znane na podstawie wcześniejszej praktyki. Często osoba lub instytucja dostarczająca próbkę oczekuje, że w odpowiedzi zawarta będzie ta informacja. Dlatego odpowiedź często ma postać pojedynczego stwierdzenia, że prawdziwe stężenie zawarte jest z określonym prawdopodobieństwem w pewnym przedziale: „Z prawdopodobieństwem 95% stężenie analitu zawarte jest w przedziale od ... do ...”. Ten przedział zwany jest **niepewnością** pomiarową. Oszacowania niepewności pomiarowej są obecnie szeroko używane w analityce chemicznej. Będą one szczegółowo omówione w rozdziale 4.

1.4 Błędy losowe i systematyczne w analizie miareczkowej

Na podstawie przykładów pomiarów miareczkowych widać, że błędy losowe i systematyczne mogą występować niezależnie od siebie. Można przypuszczać, że pojawiają się one na różnych etapach pomiaru. Pełna analiza miareczkowa roztworów wodnych z użyciem wskaźnika kolorymetrycznego składa się z następujących etapów:

1. Przygotowanie mianowanego roztworu jednego z reagentów. Etap ten wymaga: a) zważenia naczynka wagowego zawierającego odczynnik w postaci ciała stałego, b) przeniesienia naważki do kolby miarowej i ponownego zważenia naczynka wagowego (wielkość naważki określamy jako *różnicę* masy naczynka z zawartością i pustego) i c) rozpuszczenia naważki i dopełnienie kolby miarowej do odpowiedniego znaku.
2. Napełnienie naczynka pomiarowego odpowiednią objętością badanego roztworu za pomocą pipety. Wymaga to odpowiedniego napełnienia i opróżnienia pipety.
3. Miareczkowanie roztworu w naczynku pomiarowym za pomocą roztworu drugiego reagenta dodawanego z biurety. Etap ten wymaga: a) napełnienia biurety i odczekania aż poziom menisku w biurecie się ustali, b) dodania kilku kropel wskaźnika do roztworu w naczynku pomiarowym, c) odczytu początkowej objętości roztworu w biurecie, d) dodawaniu małymi porcjami roztworu z biurety aż do osiągnięcia punktu końcowego miareczkowania (zmiana koloru roztworu) i e) odczytu końcowej objętości roztworu w biurecie.

Nawet tak prosta analiza składa się z 10 oddzielnych kroków. Przy czym w celu uzyskania powtórzeń pomiaru ostatnie 7 kroków należy powtórzyć kilkakrotnie. Teoretycznie, w celu uzyskania informacji o wielkości możliwych błędów losowych i systematycznych każdy z etapów powinniśmy starannie przebadać. W praktyce prościej jest oceniać oddzielnie te etapy, które wymagają ważenia (1a i 1b) i pozostałe wymagające zastosowania pomiarów objętości. Przy czym naszym celem nie jest szczegółowy opis technik doświadczalnych używanych w poszczególnych etapach ani metod kalibracji wag i naczyń miarowych. Chodzi raczej o pokazanie źródeł potencjalnych błędów i ich istotności dla jakości końcowego wyniku.

Ważnymi wielkościami mającymi wpływ na potencjalne błędy jest tolerancja wag i naczyń miarowych (kolby miarowe, pipety, biurety itp.). Standardy dotyczące tych tolerancji ustalane są przez takie instytucje jak British Standard Institute, American Society for Testing and Materials, a w Polsce przez Państwowy Urząd Miar i Wag. Tolerancja 100-gramowych wag najwyższej jakości może być nawet tak mała jak $\pm 0,25$ mg. Jednak wagi używane w rutynowych pomiarach mogą mieć tolerancję nawet 4-krotnie większą. Analogicznie, tolerancja 250 ml kolby miarowej klasy A wynosi $\pm 0,12$ ml, podczas gdy analogiczna kolba klasy B ma tolerancję dwukrotnie większą.

Jeżeli waga lub naczynie miarowe ma dopuszczalną tolerancję, ale uzyskana waga lub użyta objętość nie jest dokładnie równa deklarowanej wartości, może to być źródłem błędu systematycznego. Na przykład, jeżeli kolba miarowa ma rzeczywistą objętość 249,95 ml zamiast 250,00 ml, ten błąd odbije się na wszystkich pomiarach, w których użyto tej kolby. Przy czym powtarzanie pomiarów nie ujawni istnienia tego błędu: we wszystkich pomiarach będziemy bowiem zakładać, że jej objętość wynosi 250,00 ml, podczas gdy w rzeczywistości jest mniejsza. Jednak, jeżeli wynik uzyskany z zastosowaniem tej kolby miarowej

będzie porównany z kilkoma innymi wynikami uzyskanymi z zastosowaniem różnych kolb miarowych, może on mieć wpływ na zmienność międzylaboratoryjną, czyli odtwarzalność wyników.

Ważenie jest zwykle związane z bardzo małym błędem *losowym*. W rutynowej pracy laboratoryjnej stosowane są najczęściej wagi o błędzie losowym rzędu 0,0002 g (w następnym rozdziale wyrażenia statystyczne stosowane do określania wielkości błędu losowego są szczegółowo omówione). Jeżeli ważony obiekt ma masę ok. 1 g lub większą, to błąd losowy związany z ważeniem (wyrażony jako procent wielkości ważonej masy) nie jest większy niż 0,02%. Dobre odczynniki stosowane w pomiarach wolumetrycznych powinny (oprócz innych charakterystyk) mieć możliwie dużą masę cząsteczkową. Umożliwia to zmniejszenie błędu losowego ważenia przy przygotowywaniu roztworów o zadanej molarności. W niektórych analizach stosuje się wagi mikroanalityczne. Pozwalają one na zważenie masy rzędu kilku miligramów. Ale nawet ważenie tak małych mas związane jest z błędem losowym jedynie kilku mikrogramów.

Błąd *systematyczny* ważenia może być znaczący i wynikać z kilku dobrze określonych źródeł. Źródłem błędów systematycznych przy ważeniu może być np.:

- wilgoć zaadsorbowana na powierzchni naczynka wagowego,
- skorodowane lub zanieczyszczone odważniki,
- efekt siły wyporu pojawiający się przy ważeniu obiektów o różnej gęstości.

Przy szczególnie dokładnych pomiarach waga powinna być kalibrowana z wykorzystaniem odważników wzorcowych dostarczanych przez odpowiednie urzędy standaryzacji. Kalibracja taka umożliwi pomiar masy z rzeczywiście dużą dokładnością, tzn. $\pm 0,01$ mg w przypadku mas w zakresie od 1 do 10 g.

W celu ograniczenia możliwości wystąpienia błędów systematycznych przy ważeniu zalecane jest stosowanie kilku prostych środków ostrożności. Ważenie różnicowe (jak w omawianym przykładzie) eliminuje na przykład błędy wynikające z osadzania się wilgoci i innych zanieczyszczeń na naczynku pomiarowym (patrz również podrozdział 2.12). Stosowanie takich środków ostrożności powoduje, że prawdopodobieństwo błędu systematycznego ważenia będzie zdecydowanie mniejsze niż błędy systematyczne wynikające ze stosowania naczyń miarowych. W praktyce pomiary grawimetryczne (ważenie) wykorzystywane są przy kalibracji szklanych naczyń miarowych. Kalibracja taka polega na ważeniu w standardowych warunkach masy wody zawartej w kolbie miarowej lub dostarczanej z pipety lub biurety. Ponadto roztwory standardowe przeznaczone do pomiarów o szczególnie wysokiej jakości (rozdział 5) są przygotowywane z wykorzystywaniem technik grawimetrycznych, a nie wolumetrycznych.

W technikach wolumetrycznych błąd *losowy* wynika głównie ze stosowania szklanych naczyń miarowych. Błąd przy napełnianiu typowej kolby miarowej o pojemności 250 ml, czyli odstęp pomiędzy znakiem a meniskiem cieczy, wynosić może $\pm 0,03$ cm. Średnica szyjki takiej kolby wynosi ok. 1,5 cm, co prowadzi do błędu objętości ok. 0,05 ml, czyli 0,02% pojemności kolby

miarowej. Błąd odczytu biurety wyskalowanej co 0,1 ml wynosi ok. 0,01–0,02 ml. Podczas każdego miareczkowania dokonujemy dwóch takich odczytów, przy czym ich błędy nie sumują się w prosty sposób (rozdział 2). Jeżeli objętość dodawanego titranta wynosi ok. 25 ml, błąd względny pomiaru jest również niewielki. Warunki doświadczenia powinny być przy tym tak dobrane, aby objętość dodawanego titranta nie była zbyt mała. Jeżeli będzie ona mniejsza niż ok. 10 ml, błąd pomiaru stanie się znaczący. Ten środek ostrożności odpowiada doborowi odczynników o dużych masach cząsteczkowych w pomiarach grawimetrycznych.

Pomimo, że pomiary wolumetryczne składają się zwykle z kilku etapów, w których stosowane są naczynia miarowe, popełniany podczas nich błąd losowy jest niewielki, o ile poszczególne pomiary wykonane zostaną starannie. W praktyce wyniki poprawnie wykonanych miareczkowań mają względne odchylenie standardowe (rozdział 2) nie większe niż ok. 0,1%. Nawet obecnie taka precyzja nie jest zwykle osiągalna w analizie instrumentalnej i nie jest powszechna.

Pomiary wolumetryczne posiadają jednak wiele znaczących źródeł błędów systematycznych. Najważniejszymi z nich są: błędy kalibracji i opróżniania naczyń miarowych i tzw. błędy wskaźnika. Najpowszechniejszym błędem w rutynowej analizie wolumetrycznej jest zbyt szybkie (niecałkowite) opróżnienie pipety i zbyt krótki czas przeznaczony na ustabilizowanie się poziomu menisku w biurecie. Nieprawidłowe opróżnianie pipety ma wpływ zarówno na błędy systematyczne, jak i losowe: dodawana objętość ma większy rozrzut niż powinna.

Temperatura pomiaru może mieć dwójaki wpływ na jakość wyniku. Naczynia miarowe są zwykle kalibrowane w temperaturze 20°C. Temperatura w laboratorium analitycznym jednak różni się o kilka stopni. Ponadto wiele pomiarów biochemicznych przeprowadza się w tzw. zimnym pokoju (*cold room*) w temperaturze ok. 4°C. Temperatura pomiaru wpływa z jednej strony na pojemność naczyń miarowych, a z drugiej na gęstość stosowanych roztworów.

Błąd wskaźnikowy może być czasami bardzo istotny, nawet większy niż błąd losowy w typowym miareczkowaniu. Na przykład, podczas miareczkowania 0,1 M kwasu solnego za pomocą 0,1 M wodorotlenku sodu zakładamy, że punkt końcowy miareczkowania odpowiada $\text{pH} = 7$. W rzeczywistości jednak ustalamy ten punkt końcowy za pomocą wskaźnika barwnego takiego jak oranż metylowy. W niezależnych pomiarach wykazano, że związek ten zmienia kolor w przedziale $\text{pH} 3\text{--}4$. Tak więc, jeżeli miareczkowanie polega na dodawaniu zasady do roztworu kwasu, wskaźnik może wskazać punkt końcowy, gdy pH roztworu wyniesie np. 3,5, czyli *przed* rzeczywistym punktem końcowym. Powstający w ten sposób błąd systematyczny może sięgać 0,2%. Z kolei podczas miareczkowania zasady roztworem kwasu punkt końcowy miareczkowania uzyskany przy zastosowaniu oranżu metylowego wypadnie trochę *po* rzeczywistym punkcie końcowym. W obu przypadkach wielkość błędu może być określona, a sam błąd skorygowany za pomocą tzw. ślepej próby. Polega ona na określeniu, jaką objętość zasady lub kwasu należy dodać, aby uzyskać zmianę koloru wskaźnika *bez* kwasu (zasady).