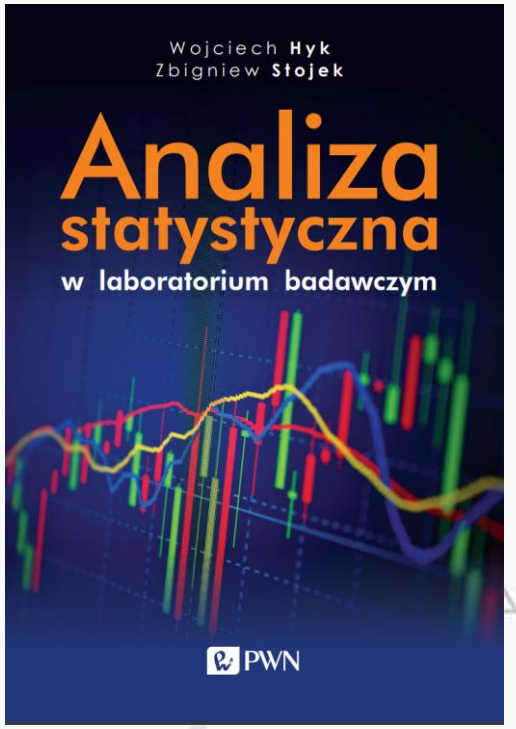
***Ł.Tymecki**, Uniwersytet Warszawski, Wydział Chemii*

Analiza Instrumentalna – laboratorium, ćwiczenie 1

Statystyczna ocena wyników pomiaru analitycznego.

na podstawie:



dr hab. Wojciech Hyk
prof. dr hab. Zbigniew Stojek
Wydawnictwo Naukowe PWN SA Warszawa 2016, 2019
ISBN 978-83-01-20824-0

STATYSTYKA

Zbieranie danych które mogą pochodzić z eksperymentów, badań ankietowych, obserwacji czy innych źródeł.

Analiza danych -narzędzia do analizy danych, takich jak średnia, mediana, odchylenie standardowe, korelacja, regresja i wiele innych. Analiza danych pozwala na wyciąganie wniosków i podejmowanie decyzji na podstawie zebranych informacji.

Interpretacja danych: Statystyka pomaga interpretować wyniki analiz, co pozwala na zrozumienie, co dane mówią o badanym zjawisku. Na przykład, testy statystyczne mogą pomóc określić, czy różnice między grupami są statystycznie istotne.

Prezentacja danych: Statystyka zajmuje się również prezentacją danych za pomocą wykresów, tabel i raportów.

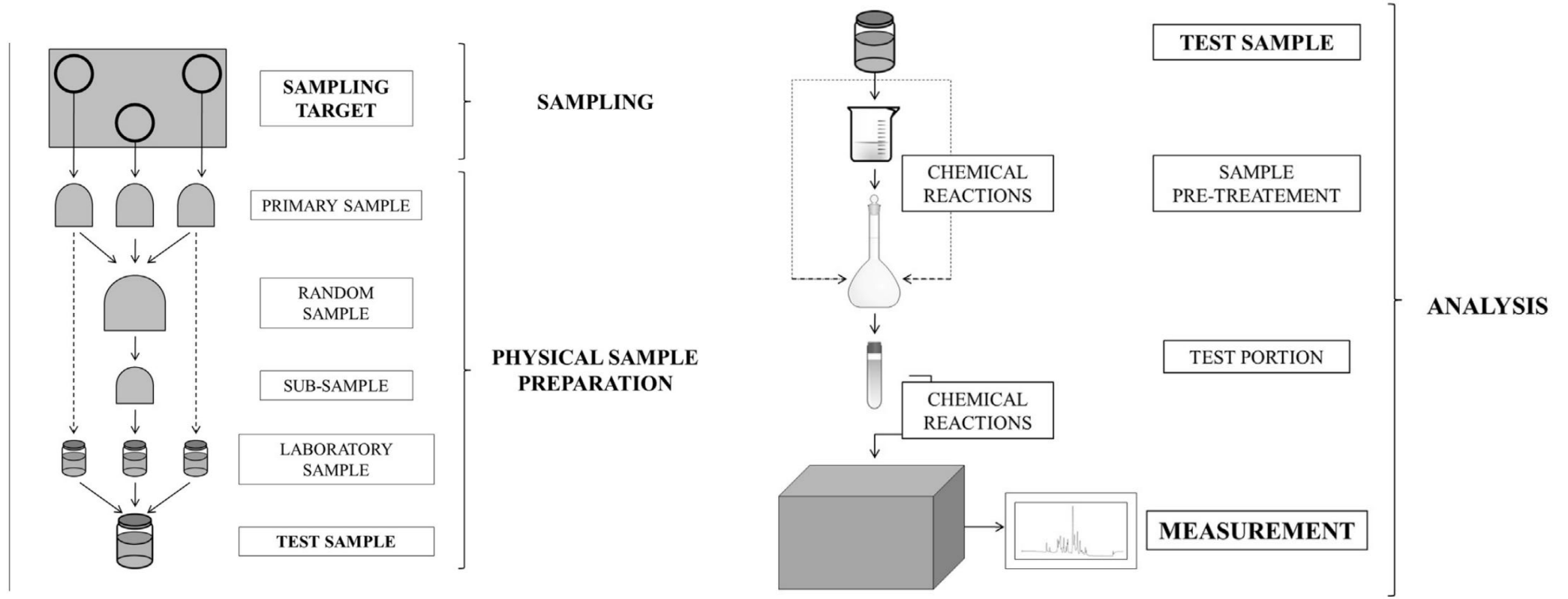
Wnioskowanie statystyczne o populacji na podstawie danych z próby. Przykłady to estymacja punktowa i przedziałowa, testowanie hipotez oraz budowanie modeli predykcyjnych.

STATYSTYKA

Ryzyka stosowania statystyki:

- użycie niewłaściwych pojęć i modeli, które nie są uzasadnione teoretycznie i źle reprezentują dane
- ograniczanie informacji z eksperymentu poprzez zastosowanie zbyt daleko idących uproszczeń
- błędy w analizie danych
- stosowanie procedur obliczeniowych bez istotnej wiedzy o ich działaniu
- niewłaściwa prezentacja danych, prowadzące do mylnych wniosków
- oszustwa- ukrywanie faktów np. dużego rozrzutu danych eksperymentalnych poprzez podanie jedynie wartości średniej
- przeszacowanie istotności statystycznej

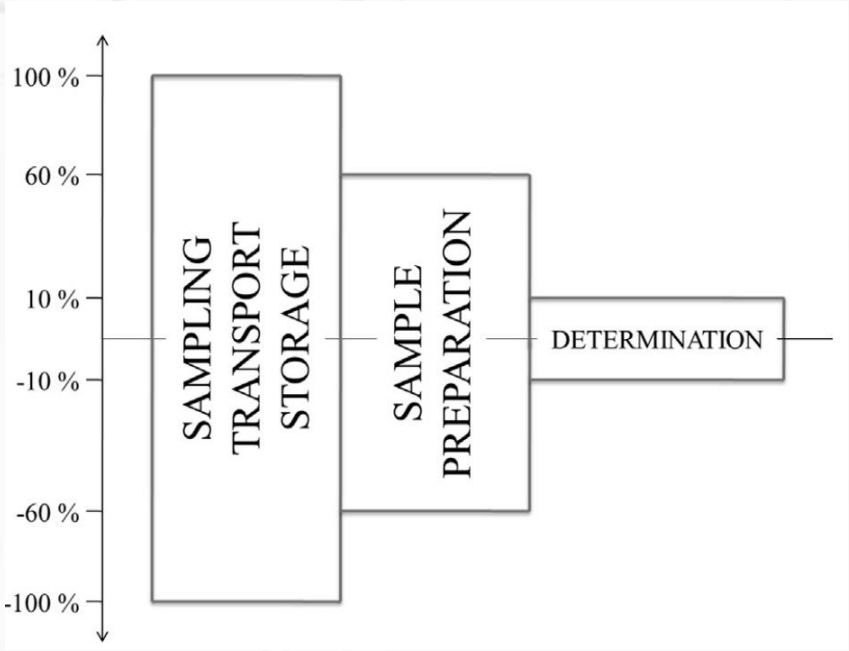
CHEMIA ANALITYCZNA



DOI: 10.1016/j.trac.2013.06.011

Co jest w próbce? – analiza jakościowa
Ile jest? – analiza ilościowa
Jak się zmienia? - analiza procesowa
Jak zbudowana ? – analiza strukturalna

WYNIK



Niepewność to wątpliwość co do wartości wyniku pomiaru. Dokładna znajomość wartości wielkości mierzonej jest nieosiągalna. Wynik pomiaru jest **estymatorem** (przybliżeniem) prawdziwej wartości wielkości mierzonej.

DLACZEGO?

niemożność wyeliminowania błędów losowych oraz z niedoskonałej korekty błędów systematycznych

POPULACJA A PRÓBKA

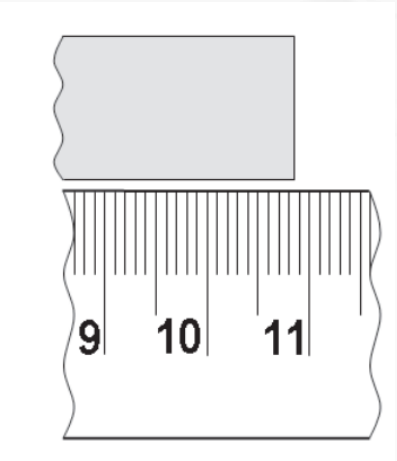
Populacja
Definicja: Populacja to całość jednostek, które są przedmiotem badania.
Przykład: Jeśli badamy wzrost wszystkich mieszkańców Polski, to populacją są wszyscy mieszkańcy Polski.
Cechy: Populacja jest zazwyczaj duża i obejmuje wszystkie możliwe jednostki, które spełniają określone kryteria.

Próbka
Definicja: Próbka to podzbiór populacji, który jest wybrany do badania. Próbka powinna być reprezentatywna dla populacji, aby wyniki badania mogły być uogólnione na całą populację.
Przykład: to próbka może składać się z 1000 losowo wybranych osób z różnych regionów kraju.
Cechy: Próbka jest mniejsza niż populacja i jest wybierana w taki sposób, aby jak najlepiej reprezentować populację.

Dlaczego używamy próbek?
Praktyczność: Badanie całej populacji może być kosztowne, czasochłonne i logistycznie trudne. Próbki pozwalają na przeprowadzenie badań w bardziej efektywny sposób.
Dokładność: Dobrze dobrana próbka może dostarczyć dokładnych i wiarygodnych wyników, które można uogólnić na całą populację.

PRÓBKA STATYSTYCZNA

W laboratorium analitycznym **próbką statystyczną** będzie zbiór wyników wielokrotnie powtórzonych pomiarów wykonanych dla wybranej próbki analitycznej lub zbiór wyników pomiarów otrzymanych dla szeregu analogicznych próbek analitycznych. Zbiór ten nazywany jest zwyczajowo **serią pomiarową**



Liczba obserwacji w próbce (n, liczność serii) - liczba powtórzonych pomiarów w serii)

$$r = \max(x_1 \ldots x_n) - \min(x_1 \ldots x_n)$$

Rozrzut wyników jest różnicą między największą i najmniejszą wartością w zbiorze n wyników Parametr ten jest używany jako miara rozproszenia wyników zwłaszcza dla próbek statystycznych o niewielkiej liczności.

67,03 / 71,76 / 74,75 / 79,69 / 67,83 / 63,80 / 73,36 / 74,41

ROZRZUT = max−min = 79,69−63,80 = **15,89**

ESTYMACJA

Estymacja to proces **szacowania** wartości nieznanych parametrów populacji na podstawie danych z próby. W statystyce wyróżniamy dwa główne rodzaje estymacji:

Estymacja **punktowa**: Polega na oszacowaniu wartości parametru populacji za pomocą jednej liczby,

Estymatory wartości środkowej:	mediana, średnia arytmetyczna, średnia ważona
Estymatory rozrzutu wyników:	wariancja odchylenie standardowe względne odchylenie standardowe współczynnik zmienności

Estymacja **przedziałowa**: Polega na oszacowaniu wartości parametru populacji za pomocą przedziału (zakresu), w którym z określonym prawdopodobieństwem znajduje się ten parametr, np. przedział ufności.

MEDIANA

W danym **szeregu uporządkowanym** liczba, która jest w **połowie** szeregu w wypadku nieparzystej liczby elementów. Dla parzystej liczby elementów – średnia arytmetyczna dwóch środkowych liczb. MEDIANA JEST DRUGIM KWARTYLEM

67,03 / 71,76 / 74,75 / 79,69 / 67,83 / 63,80 / 73,36 / 74,41

MEDIANA: 72,56

Pierwszy kwartyl (Q1): 67.03
Drugi kwartyl (Q2) (mediana): 72.56
Trzeci kwartyl (Q3): 74.75
Czwarty kwartyl (Q4): 79.69

63,80 / 67,03 / 67,83 / 71,76 / 73,36 / 74,41 / 74,75 / 79,69

KWANTYLE

dzielenie zakresu na

3:	TERCYLE
4:	KWARTYLE
5:	KWINTYLE
8:	OKTYLE
10:	DECYLE
100:	CENTYLE (PERCENTYLE)

P10 - 10. percentyl, wartość, poniżej której znajduje się 10% danych.

P50 - 50. percentyl, mediana, wartość, poniżej której znajduje się 50% danych.

P90 - 90. percentyl, wartość, poniżej której znajduje się 90% danych

ŚREDNIA ARYTMETYCZNA

Średnia arytmetyczna jest miarą (położeniem) wartości środkowej w próbce statystycznej, o ile próbka ta jest opisywana rozkładem normalnym (Gaussa).

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

Na podstawie średniej wartości próbki statystycznej wnioskujemy o wartości średniej populacji. W związku z powszechnym występowaniem błędów losowych średnia wartość próbki statystycznej praktycznie nie może być równa wartości średniej populacji.

Wartość średnia populacji jest oznaczana symbolem **μ** i jest utożsamiana zwykle z wartością prawdziwą. To pojęcie teoretyczne. W badaniach analitycznych wartość oczekiwaną najlepiej przybliża tzw. wartość rzeczywista uzyskiwana dla danej próbki na podstawie dużej liczby pomiarów przeprowadzonych przez określoną liczbę laboratoriów przy zastosowaniu różnych procedur analitycznych.

67,03 / 71,76 / 74,75 / 79,69 / 67,83 / 63,80 / 73,36 / 74,41

ŚREDNIA ARYTMETYCZNA: 71,58

wybór między medianą a średnią arytmetyczną zależy od charakterystyki danych

WARIANCJA

Wariancja jest miarą rozproszenia wyników. Dla wyników podlegających rozkładowi normalnemu, parametr ten definiują

dla populacji (gdy hipotetycznie wartość μ jest znana) (1) oraz próbki statystycznej (2)

$$\sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \mu)^2}{n} \quad (1)$$

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

wariancja próbki statystycznej podzielona przez n – 1 jest bliższa wariancji populacji. ze wzrostem n różnica między dwiema wielkościami musi zanikać, a s^2 taje się coraz lepszym estymatorem σ^2

67,03 / 71,76 / 74,75 / 79,69 / 67,83 / 63,80 / 73,36 / 74,41

WARIANCJA: 22,79

ODCHYLENIE STANDARDOWE

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \mu)^2}{n}}$$

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

SD ma tę samą jednostkę co wielkość mierzona, znacznie częściej występuje we wzorach i jest wygodniejszą formą miary rozproszenia wyników wokół średniej wartości

Względne odchylenie standardowe (RSD) jest miarą zmienności w stosunku do średniej wartości. Oblicza się je jako stosunek odchylenia standardowego do średniej wartości, wyrażony w procentach. (w ekonomii nazywany współczynnikiem zmienności)

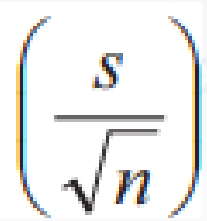
67,03 / 71,76 / 74,75 / 79,69 / 67,83 / 63,80 / 73,36 / 74,41

SD: 4.774

RSD: 6,67%

BŁĄD STANDARDOWY

Błąd standardowy (SE) dla próbki statystycznej (odchylenie standardowe średniej, standard error) jest równy ilorazowi odchylenia standardowego próbki i pierwiastka kwadratowego z liczby pomiarów



67,03 / 71,76 / 74,75 / 79,69 / 67,83 / 63,80 / 73,36 / 74,41

SE: **1.688**

Odchylenie standardowe mierzy, jak bardzo wartości w zbiorze danych różnią się od średniej. Jest to miara rozproszenia danych i jest używana, gdy chcesz zrozumieć, jak zróżnicowane są dane w zbiorze.

Błąd standardowy natomiast mierzy, jak dokładnie średnia z próby oszacowuje średnią populacji. Jest to miara niepewności związanej z oszacowaniem średniej i jest używana, gdy chcesz ocenić precyzję tego oszacowania.

Jeśli chcesz zrozumieć zmienność danych, odchylenie standardowe będzie bardziej odpowiednie. Jeśli natomiast interesuje Cię precyzja oszacowania średniej, błąd standardowy będzie lepszym wyborem

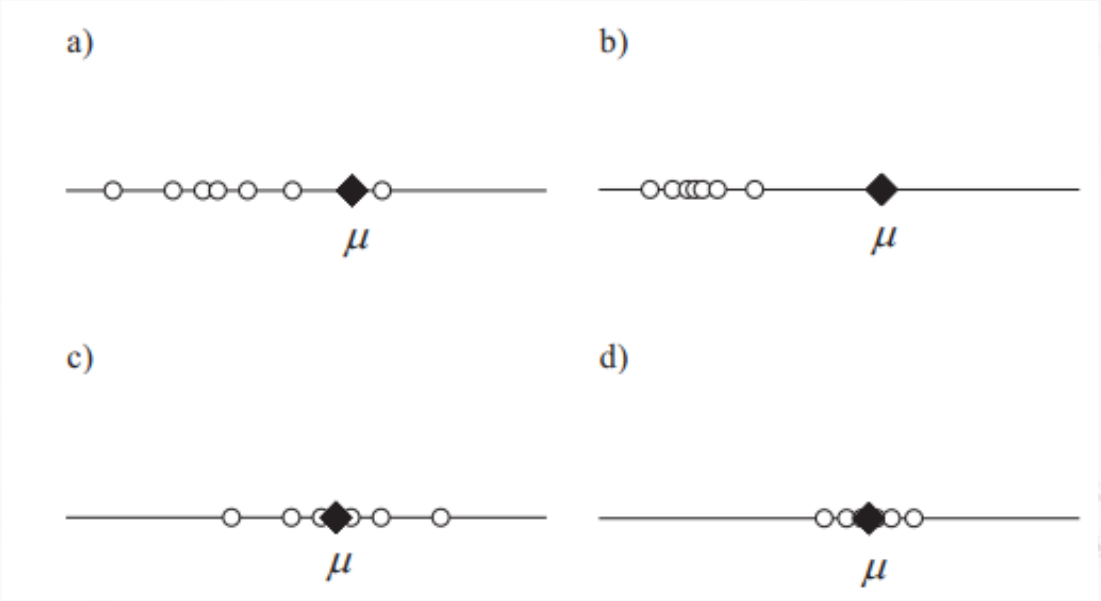
PRECYZJA/DOKŁADNOŚĆ

Precyzja określa stopień zgodności między niezależnymi wynikami pomiaru określonej próbki.
Im bliżej względem siebie położone są uzyskane wyniki pomiarów, tym większa jest precyzja pomiarów (oznaczeń).
Miarą precyzji jest odchylenie standardowe dla danej serii pomiarowej. Im mniejsza wartość odchylenia standardowego, tym bardziej precyzyjne są wyniki analityczne

Dokładność określa stopień zgodności między pojedynczym rezultatem pomiaru a wartością rzeczywistą.

Poprawność definiowana jest jako stopień zgodności między wartością średnią serii pomiarowej a wartością rzeczywistą.

dokładność jest wypadkową poprawności i precyzji. Jest ona tym większa, im większa jest precyzja i poprawność pomiarów



a) mała precyzja i poprawność,
b) duża precyzja i mała poprawność,
c) mała precyzja i duża poprawność,
d) duża precyzja i poprawność

POWTARZALNOSC /ODTWARZALNOŚĆ

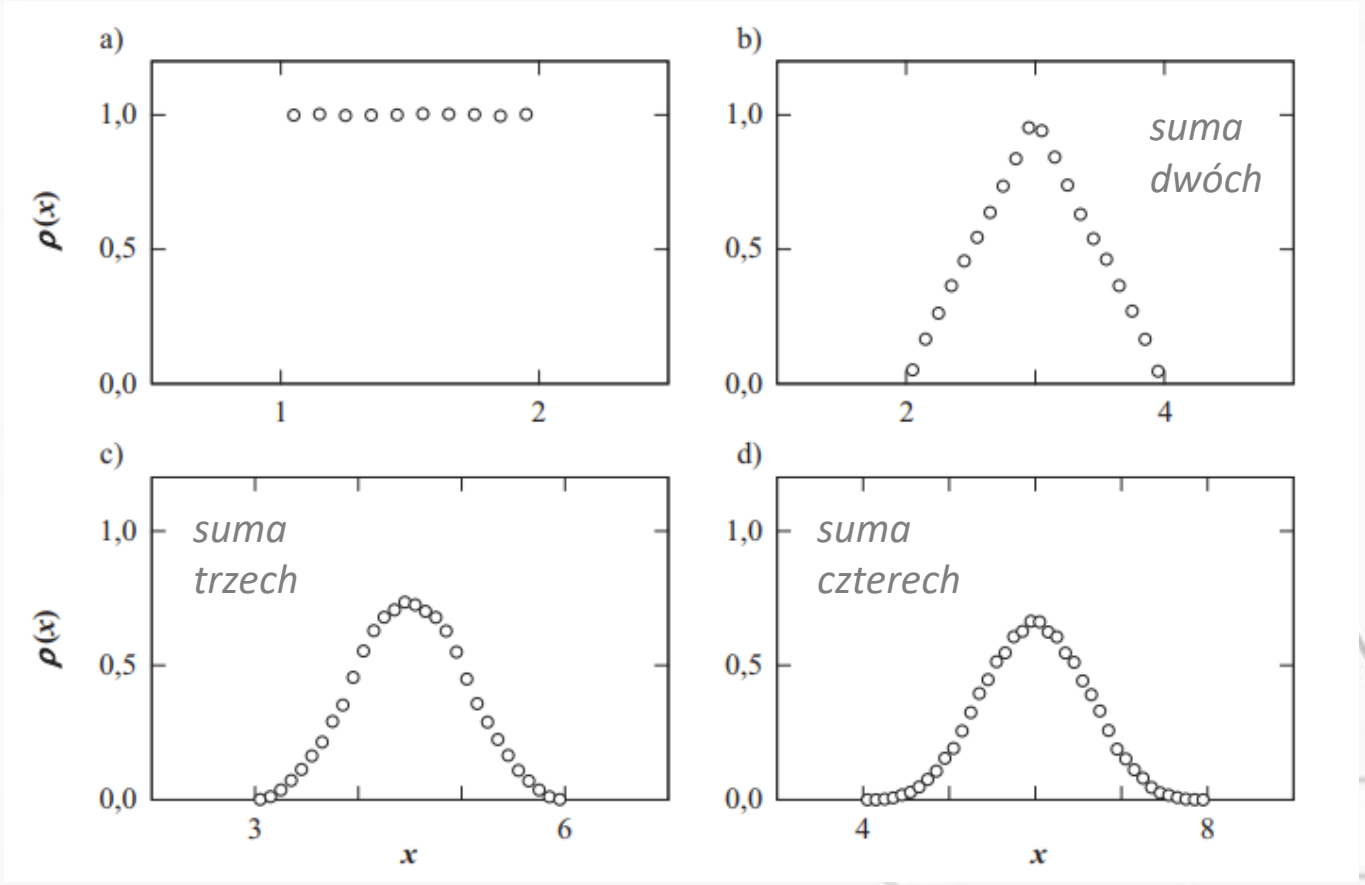
Powtarzalność oznaczeń, gdy analiza wykonywana jest w tym samym laboratorium, przez tego samego analityka, tą samą metodą, na tych samych urządzeniach, w możliwie krótkim przedziale czasowym.

Odtwarzalność oznaczeń, gdy wyniki otrzymywane są w sposób niezależny, tą samą metodą i na tej samej próbce, ale w różnych laboratoriach, przez różnych analityków na różnych urządzeniach.

GĘSTOŚĆ PRAWDOPODOBIEŃSTWA

W przypadku rzutu monetą lub kostką wystąpienie danego zdarzenia (tj. orła lub reszki, czy danej liczby oczek od 1 do 6) jest jednakowo prawdopodobne. Mówimy, że są to zdarzenia losowe, które nie podlegają rozkładowi normalnemu. Rozkład prawdopodobieństwa w tym przypadku jest funkcją stałą i jest nazywany rozkładem prostokątnym.

Wykresy gęstości prawdopodobieństwa dla liczb losowe z zakresu [1, 2] oraz sumy losowych dwóch, trzech i czterech z tego zakresu



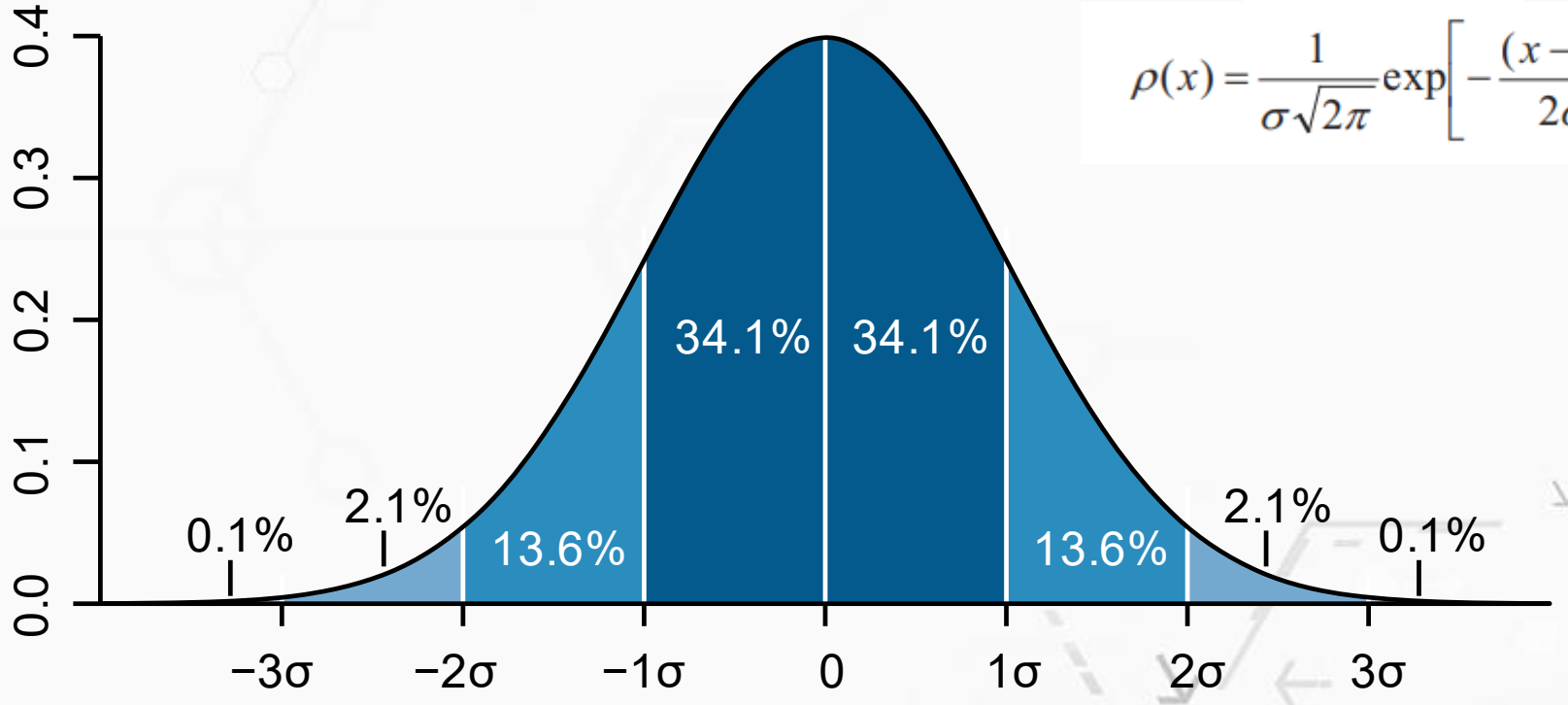
ROZKŁAD NORMALNY

Większość wielkości charakteryzujących zjawiska naturalne, podlega rozkładowi normalnemu (Gaussa)

$$\rho(x) \sim e^{-x^2}$$

$$\rho(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$$

możemy znaleźć prawdopodobną wartość mierzonej wielkości, lecz również możemy określić prawdopodobieństwo znalezienia danej wartości w obranym przedziale.

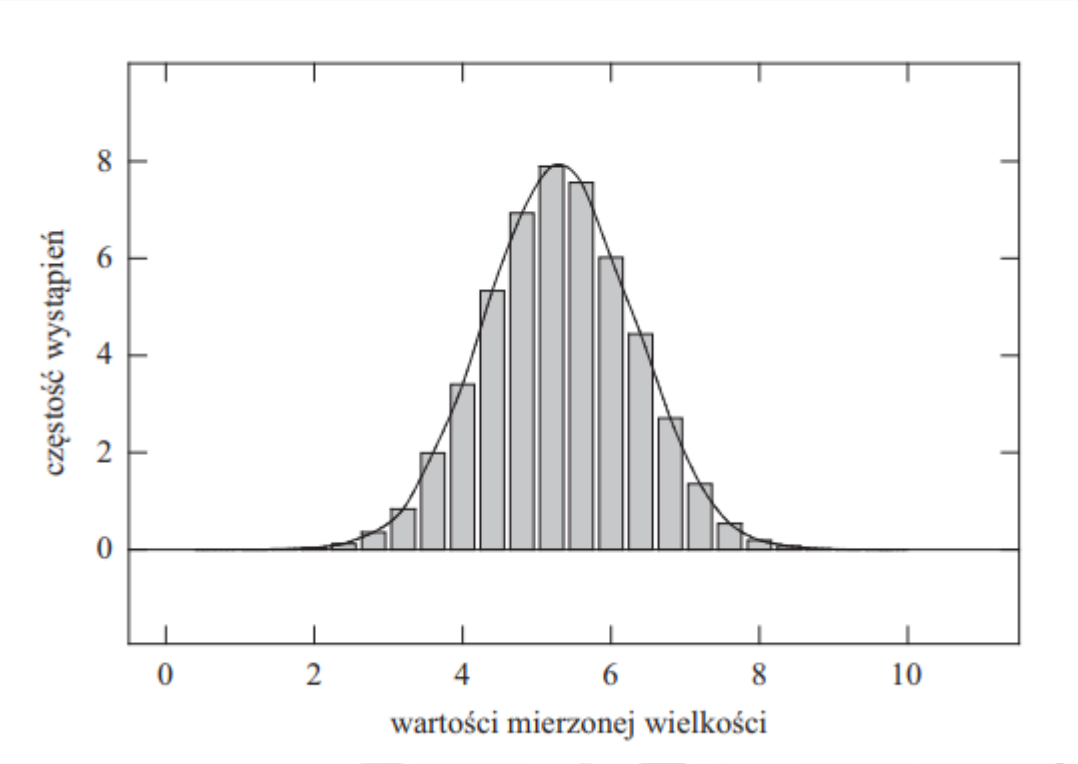


w przedziale od −σ do +σ znajduje się 68,3% wszystkich wartości. Wewnątrz granic od −1,96σ do +1,96σ znajduje się 95% wszystkich wartości, a w przedziale od −3,09σ do +3,09σ – 99,7%.

HISTOGRAM

Służy do sprawdzenia czy rozkład uzyskiwanych wyników jest normalny

W tym celu zakres zmiennej x (wielkość mierzona) dzielimy na kilkanaście podzakresów o równej szerokości i zliczamy wyniki plasujące się w kolejnych zakresach. Słupki prezentują kolejne zakresy. Wysokość słupków jest określona
przez liczbę wyników leżących w danym zakresie. kształt przykładowego histogramu wskazuje na prawie idealny rozkład normalny wyników wziętych do konstrukcji histogramu.



PRZEDZIAŁ UFNOŚCI

$$\bar{x} - z\left(\frac{\sigma}{\sqrt{n}}\right) < \mu < \bar{x} + z\left(\frac{\sigma}{\sqrt{n}}\right)$$

$$\mu = \bar{x} \pm z\left(\frac{\sigma}{\sqrt{n}}\right)$$

Na podstawie szerokości przedziału ufności ocenić można jakość wyniku.
Określamy prawdopodobieństwo znalezienia wartości prawdziwej i otrzymujemy przedział w jakim się znajduje ta wartość.

Przedział ufności wyznacza się:
przy **dużej liczbie** pomiarów (n>20) z rozkładu normalnego

$\bar{x} \pm 1{,}96\sqrt{n}$ Dla założonego prawdopodobieństwa 95%

$\bar{x} \pm 2{,}58\sqrt{n}$ Dla założonego prawdopodobieństwa 99%

przy **małej liczbie** pomiarów (n<20) z rozkładu Studenta

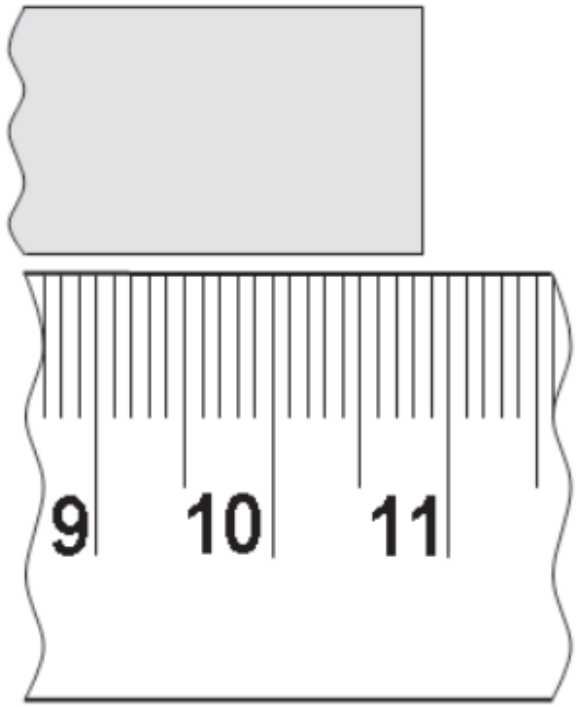
William Sealy Gosset, pseudonim „Student”

$\bar{x} \pm t\ SE$

Wartości krytyczne rozkładu *t*-Studenta

Liczba stopni swobody	Test jednostronny	
$(n-1)$	95%	99%
2	2,920	6,965
3	2,353	4,541
4	2,132	3,747
5	2,015	3,365
6	1,943	3,143
7	1,895	2,998
8	1,860	2,896
9	1,833	2,821
35	1,690	2,437
36	1,688	2,434
37	1,687	2,431
38	1,686	2,429
39	1,685	2,426
40	1,684	2,423
∞	1,645	2,326

CYFRY ZNACZĄCE



wynik pomiaru mieści się w granicach 10,8–10,9 cm.

10,85 cm (Liczba ta zawiera 4 cyfry znaczące)

Trzy pierwsze są dokładne, a czwarta jest obarczona niepewnością

W przedstawionym przykładzie zero zostało zaliczone do cyfr znaczących, gdyż występuje między cyframi różnymi od zera.

10,85 cm = 108,5 mm = 0,1085 m =108500 µm

Aby uniknąć niejednoznaczności przy zaliczaniu zer do cyfr znaczących, zaleca się zapisywanie liczb w postaci wykładniczej, tzn. wyrażonej za pomocą odpowiedniej potęgi liczby 10 (np. 108500 µm w postaci wykładniczej przyjmie postać $1{,}085 \cdot 10^5$ µm)

CYFRY ZNACZĄCE - operacje

W dodawaniu i odejmowaniu liczb (wyników pomiarów) bierzemy pod uwagę ich rozwinięcie dziesiętne. Wynik pomiaru z najmniejszą liczbą miejsc dziesiętnych determinuje rozwinięcie dziesiętne wyniku, np.

$20{,}4 + 1{,}322 + 83 = 104{,}722 \approx 105$

Dodawanie i/lub odejmowanie liczb zapisanych w postaci wykładniczej należy przeprowadzać dla tych samych potęg 10, np.

$3{,}23 \cdot 10^3 + 4{,}542 \cdot 10^1 - 6{,}844 \cdot 10^2 =$
$= 3{,}23 \cdot 10^3 + 0{,}04542 \cdot 10^3 - 0{,}6844 \cdot 10^3 =$
$= 2{,}59102 \cdot 10^3 \approx 2{,}59 \cdot 10^3$

Mnożenie i dzielenie

Wynik ma tyle cyfr znaczących, ile wynik pomiaru z najmniejszą liczbą cyfr znaczących, np.
6,221 (4 cyfry znaczące) · 5,2 (2 cyfry znaczące) =
=32,3492
≈ 32 (2 cyfry znaczące)
lub
15,5 (3 cyfry znaczące) · 27,3 (3 cyfry znaczące) · 5,4 (2 cyfry znaczące) = 2285,01 ≈
≈ 2300 = $2{,}3 \cdot 10^3$ (2 cyfry znaczące).

Logarytmowanie

Liczba cyfr znaczących w mantysie (część dziesiętna wyniku logarytmu) powinna być taka sama jak liczba cyfr znaczących w liczbie wejściowej.

$\log(0{,}00567) = -2{,}246416941 \approx -2{,}246$

ZAOKRĄGLANIE LICZB

zaokrągleń dokonujemy dopiero po zakończeniu obliczeń, obliczenia prowadzimy z taką liczbą cyfr jakimi dysponujemy

Określanie właściwej liczby cyfr znaczących w uzyskanych wynikach pomiaru zwykle sprowadza się do odrzucania niektórych cyfr, najczęściej nieznaczących. Wynik poddawany jest zaokrągleniu, które może prowadzić do modyfikacji ostatniej z pozostawianych cyfr w wyniku. Sposób zaokrąglania liczb ujęty jest następującymi regułami:

1. Jeżeli pierwsza z odrzucanych cyfr jest mniejsza od 5, to ostatnia z pozostawionych cyfr nie ulega zmianie.
2. Jeżeli pierwsza z odrzucanych cyfr jest większa od 5, to ostatnią z pozostawionych cyfr powiększa się o 1.
3. Jeżeli pierwsza z odrzucanych cyfr jest równa 5 i przynajmniej jedna z pozostałych odrzuconych cyfr jest różna od 0, to ostatnią z pozostawionych cyfr powiększa się o 1.
4. Jeżeli pierwsza z odrzucanych cyfr jest równa 5 i jest ostatnią różną od 0 cyfrą zaokrąglanej liczby, to ostatnią z pozostawionych cyfr zaokrągla się w górę do najbliższej cyfry parzystej (gdy jest to cyfra nieparzysta) lub pozostawia bez zmian (gdy jest to już cyfra parzysta).

123.4567	zaokrąglone do czterech cyfr znaczących to	123.5.
0.00456789	zaokrąglone do czterech cyfr znaczących to	0.004568
98765.4321	zaokrąglone do czterech cyfr znaczących to	98770
0.000123456	zaokrąglone do czterech cyfr znaczących to	0.0001235
456789	zaokrąglone do czterech cyfr znaczących to	456800

BŁĘDY (pomiaru)

ze względu na ich charakter oraz sposób zaburzania wyników pomiaru, dzielą się na:
błędy losowe (przypadkowe), systematyczne oraz grube.

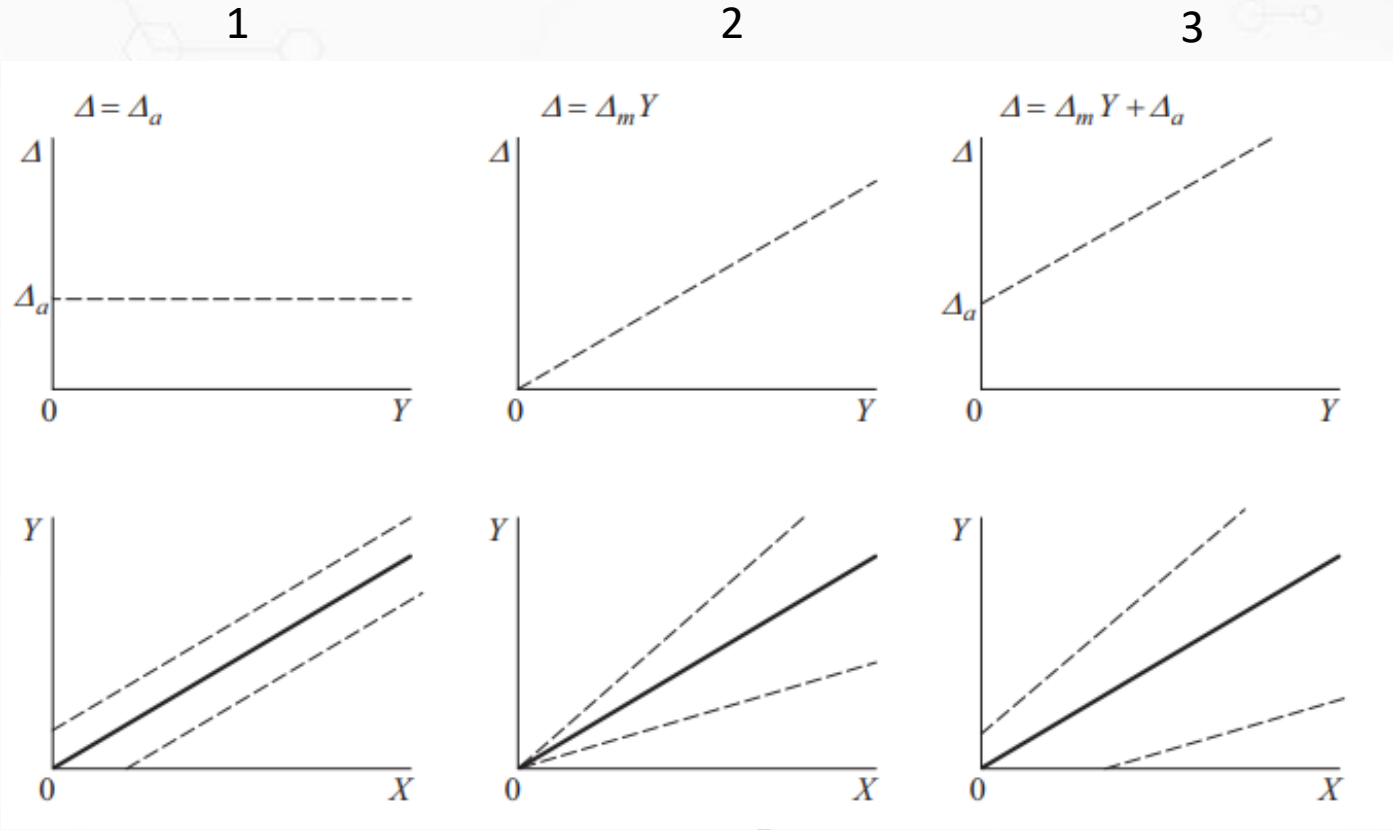
BŁĘDY LOSOWE podlegają prawdopodobieństwu, nie zmieniają rozkładu

- wszystkie wyniki pomiaru, włączając te uzyskane z bardzo dużą precyzją, nie są idealnie dokładne i posiadają przybliżony charakter;
- nie jest możliwe uzyskanie jednakowych wartości wyników w danej serii pomiarowej przy zachowaniu największej nawet dbałości eksperymentalnej (gdy tak się zdarzy to oznacza, że przyrząd pomiarowy ma czułość niedostosowaną do wykonywanych pomiarów);
- średnia arytmetyczna serii pomiarowej (n powtórzonych pomiarów) stanowi tylko estymator wartości prawdziwej wielkości mierzonej; ta jednak pozostaje zwykle nieznana.

BŁĘDY SYSTEMATYCZNE - zaburzają rozkład wyników

- nieprawidłowe ustawienia przyrządu
- niewystarczająca czystość chemiczna,
- matryca próbek,
- zaburzenia układu pomiarowego czynnikami zewnętrznymi
- niedoskonała standaryzacja lub kalibracja

1. BŁĘDY SYSTEMATYCZNE stałe
2. BŁĘDY SYSTEMATYCZNE proporcjonalne
3. BŁĘDY SYSTEMATYCZNE złożone



BŁĄD GRUBY - zaburza rozkład wyników

Istnienie **jednego** wyniku znacząco odstającego od pozostałych uzyskanych w danej serii pomiarów. Źródłem tego błędu może być nieoczekiwane zaburzenie układu pomiarowego lub czysto ludzka pomyłka (np. podczas zapisywania danej wartości). Jeśli odstępstwo danego wyniku jest wyraźnie duże, wynik taki można arbitralnie odrzucić. W razie wątpliwości należy posłużyć się odpowiednim testem **(tylko jeden raz! dla krańca zbioru)**

Test Q (Dixona) pozwala na wyeliminowanie błędu grubego w małym zbiorze danych

1. sortowanie danych (rosnąco)
2. obliczenie statystyki Q

   Q = (|podejrzana wartość- wartość najbliższa podejrzanej|/rozrzut)
3. Porównanie z wartością krytyczną (dla odpowiedniego rozmiaru próbki i poziomu ufności) - tabele
4. Jeśli Q jest większe niż wartość krytyczna, podejrzaną wartość można odrzucić

BŁĄD GRUBY

Dokonano czterech pomiarów zawartości azotanów w próbkach soku marchwiowego. Uzyskano następujące rezultaty: 0,553; 0,560; 0,551 i 0,530 mg/L.
Czy ostatni wynik może być obarczony błędem grubym? Czy należy go odrzucić?

Porządkujemy rosnąco wartości wyników
(0,530; 0,551; 0,553; 0,560).
Obliczamy wartość Q:
(0,551-0,530)/(0,560-0,530) = 0,700

Porównujemy z wartością krytyczną z tabeli
(n=4, P=95%) = 0,829
Ponieważ nasz wynik jest mniejszy od krytycznego, zostaje w zbiorze. Ustalone prawdopodobieństwo mówi że nie popełniamy błędu większego niż 5% jeśli pozostawimy wynik

Wykonano 3 dodatkowe pomiary i uzyskano:
0,550; 0,563 i 0,561 mg/L
Teraz seria po uporządkowaniu wygląda następująco:
(0,530; 0,550; 0,551; 0,553; 0,560; 0,561; 0,563)

Obliczamy wartość Q:
(0,550-0,530)/(0,563-0,530) = 0,606
Porównujemy z wartością krytyczną z tabeli
(n=7, P=95%) = 0,568
Ponieważ nasz wynik jest większy od krytycznego, powinien być usunięty ze zbioru.

BUDŻET NIEPEWNOSCI

Roztwór wzorcowy jonów kadmu(II) przygotowano przez roztworzenie 0,10028 g metalicznego kadmu (czystość: 99,99% ± 0,01%) w 100 ml roztworu kwasu. Obliczyć stężenie jonów kadmu (cCd) w przygotowanym roztworze **oraz oszacować jego niepewność wyznaczenia stężenia.**

Etapy procedury:

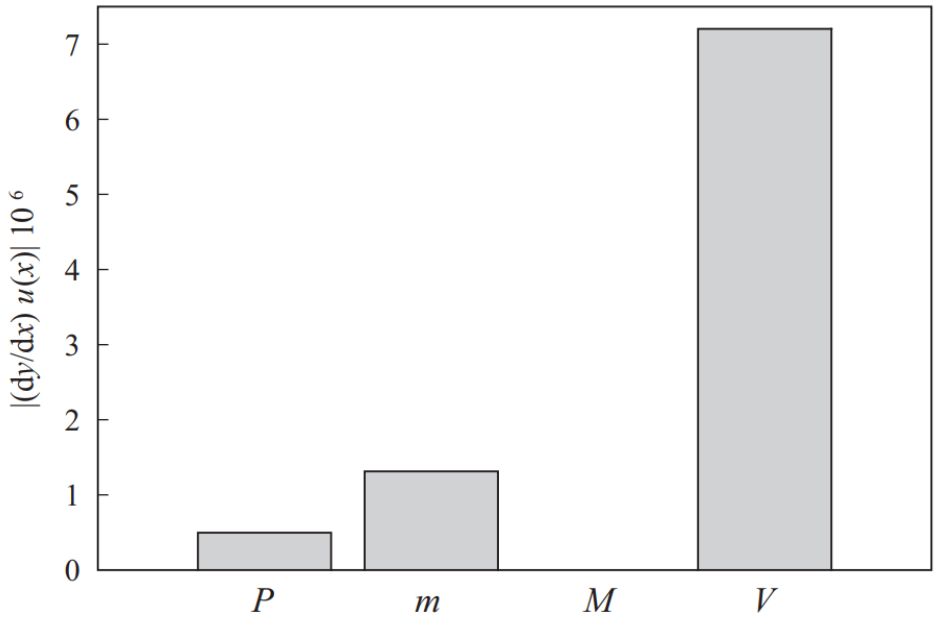
- oczyszczanie powierzchni metalu,
- ważenie metalu,
- rozpuszczanie i rozcieńczanie,
- sformułowanie wyniku.

dla każdego etapu można określić niepewność wyznaczonej wartości masa/czystość/objętość/masa molowa
Po odpowiednim oczyszczeniu metalu, użyciu wagi o błędzie wskazania 0,01mg, klasy szkła A (+-0,1mL/100)

$$c_{Cd} \pm u(c_{Cd}) = (8{,}920 \pm 0{,}007) \cdot 10^{-3}\ \text{mol/l}$$

$$c_{Cd} = \frac{mP}{M_{Cd}V}\ [\text{mol/l}]$$

gdzie:
c_{Cd} – stężenie jonów kadmu [mol/l],
m – masa próbki metalu [mg],
P – czystość metalu (wyrażona ułamkiem molowym),
M_{Cd} – masa atomowa kadmu [g/mol],
V – objętość roztworu [ml].



REGRESJA

Regresja w statystyce to metoda analizy danych, która pozwala na modelowanie i analizowanie zależności między zmiennymi. Najczęściej stosowaną formą regresji jest regresja liniowa, która bada liniową zależność między zmienną zależną (y) a jedną lub więcej zmiennymi niezależnymi (x).
W skrócie, regresja pomaga odpowiedzieć na pytania typu: "Jak zmiana jednej zmiennej wpływa na inną zmienną?”

Regresja oznacza dopasowywanie zbioru parametrów pewnej formuły matematycznej, nazywanej modelem regresyjnym, do dostępnych danych, tak aby uzyskany w ten sposób model jak najdokładniej opisywał te dane.

Współczynnik Pearsona - jest miarą siły i kierunku liniowej zależności między dwiema zmiennymi ilościowymi. Oznaczany jest symbolem r i przyjmuje wartości od -1 do 1.

-1 – korelacja idealnie ujemna
0 – brak korelacji
1 – korelacja idealnie dodatnia

R^2 - współczynnik determinacji, jest to kwadrat współczynnika korelacji Pearsona przyjmuje wartości od 0 do 1.

1 - model idealnie dopasowuje się do danych, a wszystkie punkty danych leżą na linii regresji.

0 - model nie wyjaśnia żadnej zmienności zmiennej zależnej