Analiza błędów pomiarowych

W naukach przyrodniczych kluczową rolę w weryfikacji wszelkich hipotez i teorii naukowych odgrywa eksperyment i jego wynik. Częstokroć pojedynczy wynik eksperymentalny leży u podstaw nowych teorii i odrzucenia dotychczasowych wyobrażeń o danym zjawisku, czy wręcz wyobrażenia o otaczającym nas świecie. Ale aby eksperyment naukowy mógł spełniać tak ważną rolę konieczna jest znajomość dokładności z jaką został on wykonany. Warto zdać sobie sprawę z faktu, że wszelkie wielkości fizyczne wyznaczone doświadczalnie, określone zostały z mniejszą lub większą dokładnością, nawet te podawane w tablicach fizykochemicznych jako stałe podstawowe, powszechnie uznane za wielkości „prawdziwe" i wykorzystywane we wzorach częstokroć bez wnikania jaka jest ich dokładność. Od wieków wyznaczenie „prędkości" (a w zasadzie szybkości) światła zaprzątało umysły naukowców próbujących dokonać jej pomiaru. Począwszy od prób Galileusza, który nieudanie próbował dokonać tego pomiaru mierząc opóźnienie z jakim światło pokonuje drogę pomiędzy obserwatorami na szczytach sąsiednich wzgórz, poprzez obserwacje zaćmienia jednego z księżycy Jowisza, Io, wykonane przez Römera w 1675 r., których konkluzją było przypisanie światłu skończonej prędkości czy coraz dokładniejsze pomiary pomysłowych eksperymentów Fitzeau (1849 r.), Foucaulta (1850 r.) czy Michelsona (lata 1880-1930) prędkość z jaką przemieszcza się fala świetna określano coraz dokładniej. W chwili obecnej podawana w tablicach fizykochemicznych wartość prędkości światła w próżni (c = 299 792 458 m/s) jest określana jako wartość dokładna, czyli nie obarczona błędem. Czy obecnie zaakceptowana wartość jest więc wyznaczona z nieograniczoną dokładnością? Oczywiście, że nie, choć jednocześnie osiągnięta dokładność jest na tyle duża, że od 1983 r. przyjęta wartość prędkości światła w próżni stanowi podstawę definicji metra.

W niniejszym rozdziale przedstawione zostaną podstawowe informacje związane z dokładnościami pomiarowymi, metodami ich wyznaczania i analizy.

Pomiar

Podstawowym celem eksperymentu naukowego, niezależnie od tego czy przeprowadza go z dużą dokładnością naukowiec, stosujący bardzo precyzyjną i skomplikowana aparaturę, czy też student w trakcie zajęć laboratoryjnych jest dokonanie pomiaru wielkości fizycznej, czyli wyznaczenie jej wartości (podanie wartości liczbowej wraz z jednostką) i określenie dokładności z jaką pomiar został wykonany. Wartości różnych wielkości uzyskuje się z ***pomiarów bezpośrednich*** bądź ***pośrednich***. W pomiarze bezpośrednim często odczytuje się wynik wprost ze wskazania przyrządu, przeważnie

wyskalowanego w jednostkach mierzonej wielkości. Przyrządy mogą być różnorodne, na przykład wagi, mierniki elektryczne, spektrometry, liczniki cząstek promieniowania. Częstokroć używanie przyrządów pomiarowych wymaga stosowania wzorców miar, jak np. odważniki, pojemniki miarowe (cylindry, pipety), przymiary (linijka, suwmiarka). Sposób wykonania pomiaru jest oparty na określonej metodyce, którą nazywamy ***metodą pomiarową***. Na przykład pomiar prędkości może być oparty na zjawisku Dopplera, a temperaturę można mierzyć na podstawie zjawiska termoelektrycznego.

Wśród metod pomiarowych szczególne znaczenie mają metody bezpośrednie, oparte na prawach fizycznych dających się wyrazić przez podstawowe stałe (***c, G, h, k, F,*** N_A...) i podstawowe wielkości (długość ***l***, czas ***t***, masa ***m***, temperatura ***T***, prąd elektryczny ***I***, światłość I_v, ilość substancji ***n***). W pomiarze pośrednim wartość określonej wielkości jest oznaczana na podstawie bezpośrednich pomiarów innych wielkości. Wynik pomiaru oblicza się używając wzoru wiążącego wielkość oznaczaną i wielkości mierzone. Na przykład gęstość substancji oblicza się na podstawie zmierzonych wartości masy i objętości. Pomiar pośredni często nazywa się ***oznaczaniem***.

Prezentacja wyniku pomiaru

Warto w tym momencie zwrócić uwagę na prawidłowy sposób prezentacji uzyskanego wyniku pomiarowego. Symbole wielkości piszemy czcionką pochyłą (kursywą), również ich indeksy górne i dolne, jeżeli są symbolami wielkości. Natomiast liczby i jednostki, a także symbole pierwiastków i cząstek elementarnych, piszemy czcionką prostą. Do nielicznych wyjątków należy symbol pH.

Wielkości mianowane, prezentujemy jako wartości liczbowe, wskazujące ile razy zmierzona wartość jest większa od jednostki, wraz z podaniem miana jednostki. Stosowane mogą być różnorodne przeliczniki jednostek, choć należy dążyć do posługiwania się jednolitym systemem jednostek zwanym układem SI (franc. Systeme International d'Unites), wywodzącym się jeszcze z czasów Wielkiej Rewolucji Francuskiej. Częstokroć jednak, tradycyjny w danej dziedzinie sposób prezentacji wyników jest nie tylko wymagany, ale i najbardziej czytelny. Np. wiele metod spektroskopowych wykorzystuje odmienny sposób charakteryzowania fali elektromagnetycznej poczynając od określania częstości fal radiowych (ν/Hz), liczb falowych fal w zakresie podczerwieni (k/cm^{-1}), długości fali w zakresie UV i widzialnym (λ/nm) czy jednostek energii promieniowania jonizującego (E/eV).

Z innym przykładem możemy się spotkać przy podawaniu wyniku pomiaru szybkości:

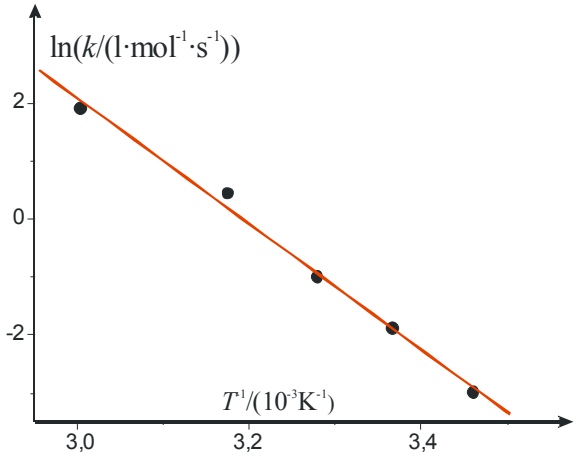
$$v = 72 \text{ km/h lub } v = 20 \text{ m/s},$$

przy czym zapis w pierwszej postaci jest charakterystyczny dla określania szybkości samochodu, a drugi szybkości wiatru.

Spotkać możemy różne sposoby zapisu wyniku pomiaru:

$\boldsymbol{v}$ = 72 km/h $\boldsymbol{v}$ = 72 [km/h] $\boldsymbol{v}$/(km/h) = 72

przy czym ostatni z zaprezentowanych sposobów, gdy wartość wielkości jest wyrażona za pomocą wartości liczbowej (niemianowanej) oraz ilorazu wielkości przez jednostkę, jest zalecany do opisu zestawień tabelarycznych i osi współrzędnych na wykresach.



Rys. 1 Zależność Arrheniusa – zależność logarytmu stałej szybkości reakcji (ln k) od odwrotności temperatury (T^{-1})

Na przykład nanosząc wartość stałej szybkości reakcji $\boldsymbol{k}$ = 0,368 l·mol^{-1}·s^{-1} w temperaturze 30^{o}C na wykres liniowej zależności Arrheniusa: ln $\boldsymbol{k}$ = ln$\boldsymbol{A}$ - $\boldsymbol{E_A/RT}$, osie współrzędnych należy opisać: ln **(**$\boldsymbol{k}$/(l·mol^{-1}·s^{-1})**)** oraz 10^3K/$\boldsymbol{T}$ (ewentualnie kK/$\boldsymbol{T}$) lub $\boldsymbol{T}^{-1}$/(10^{-3}K^{-1}) (ewentualnie $\boldsymbol{T}^{-1}$/kK^{-1}). Na wykresie otrzymamy punkt o współrzędnych: odciętej: T^{-1}/(10^{-3}K^{-1}) = 3,2987 oraz rzędnej: ln**(**$\boldsymbol{k}$/(l·mol^{-1}·s^{-1})**)** = ln 0,368 = -0,9997.

Dokładność pomiaru

Dokładność metody badawczej charakteryzuje zgodność otrzymywanych wyników, czyli zmierzonej wartości x, z ***wartością prawdziwą***, nazywaną też wartością ***rzeczywistą***. Wartość prawdziwa mogłaby zostać zmierzona w wyniku pomiaru bezbłędnego. W rzeczywistości, jedynym sposobem poznania tej wielkości jest jej ***ocena*** (***oszacowanie, estymacja***). Ocenę tę uzyskaną w wyniku pomiaru nazywa się ***wartością umownie prawdziwą, wartością poprawną*** lub ***uznaną***. Powinna ona być tak bliska wartości prawdziwej, aby różnica między nimi była pomijalnie mała z punktu widzenia celu

wykorzystania wartości poprawnej. W dalszej części niniejszego rozdziału będziemy się posługiwali pojęciem wartości prawdziwej jako głównego celu pomiaru eksperymentalnego.

Kilka ważnych czynników wpływa na poprawność estymacji, a do najważniejszych należą ***błędy pomiarowe***. Przede wszystkim mogą to być popełnione przez eksperymentatora ewidentne błędy, tzw. ***błędy grube***. Błędy grube pochodzą z pomyłek eksperymentatora, niezauważonych przez niego niesprawności przyrządów i niewłaściwych warunków pomiaru. Błędy grube pojawiają się, gdy eksperymentator nieprawidłowo odczyta wskazania przyrządu, źle zanotuje liczby lub jednostki, pomyli się w obliczeniach, wykorzysta niewłaściwe dane literaturowe itp. Jedną z przyczyn błędów grubych u początkujących eksperymentatorów jest przesadne zaufanie do sprawnego działania przyrządów i niestaranne prowadzenie notatek laboratoryjnych. Rażąco duże błędy grube dają się łatwo wykryć i usunąć.

Drugą grupę stanowią ***błędy systematyczne***. Błędy systematyczne pochodzą z niesprawności przyrządów pomiarowych, niepoprawnej ich ***kalibracji*** (***skalowania***), nieidentyczności warunków pomiaru (temperatury, ciśnienia, wilgotności, zasilania przyrządu itp.) z warunkami kalibracji przyrządów, a także indywidualnych cech eksperymentatora i nieścisłości wzorów obliczeniowych. Typowymi przykładami źródeł takich błędów mogą być np. późniący się stoper lub błędny odczyt wyniku z miernika. Każdy eksperymentator ma indywidualny sposób wykonywania pomiaru, np. odczytu wskazań przyrządów, przez co wpływa na powstanie błędu systematycznego. Błąd systematyczny rzadko bywa stały, czyli niezależny od mierzonej wielkości. Może być złożoną funkcją wielkości.

Błąd ten nie wynika z niestaranności eksperymentatora, jak błąd gruby, ale jest zależny od jego umiejętności manualnych i doświadczenia. Ocena wartości błędów systematycznych wymaga analizy wszystkich czynników aparaturowych i osobowych wpływających na wynik pomiaru. Można je w istotny sposób ograniczać wykonując pomiary metodami porównawczymi (różnicowymi, kompensacyjnymi) w stosunku do wzorców, o znanych wartościach poprawnych.

Najważniejszą grupę stanowią jednak tzw. ***błędy przypadkowe***. Charakteryzują się tym, że w serii pozornie identycznych powtórzeń pomiaru tej samej wartości mierzonej błędy te mogą być dodatnie, i ujemne, a także małe i duże. Powstają pod wpływem wielu czynników, których praktycznie nie daje się przewidzieć. Przyczyną błędów przypadkowych są niewielkie ***fluktuacje*** (wahania wokół wartości przeciętnej) temperatury, ciśnienia, wilgotności i innych parametrów zarówno w przyrządach pomiarowych i ich częściach, jak i w badanych obiektach, gdyż próbki użyte do kolejnych powtórzeń pomiaru mogą mieć

przypadkowo nieznacznie różne własności fizyczne i chemiczne. Również chwilowe zmiany przyzwyczajeń eksperymentatora, wynikające nawet z jego nastroju, mogą być przyczyną błędów przypadkowych.

Błędy przypadkowe Δx podlegają prawom statystyki matematycznej i dlatego bywają także nazwane ***błędami statystycznymi*** lub ***losowymi***. Konsekwencją przypadkowości tych błędów jest możliwość ich opisania, a także przewidywania ich wartości za pomocą analizy statystycznej wyników wielokrotnie powtórzonych pomiarów. O ile nazwa błąd pomiarowy, jest synonimem „pomyłki" w przypadku błędów grubych i systematycznych, o tyle błąd przypadkowy jest nierozerwalnie związany z istotą pomiaru i oznacza niemożliwą do uniknięcia niepewność pomiarową. Koniecznym jest więc określenie jednoznacznych reguł pozwalających tę wielkość oszacować (estymować), podobnie jak ma to miejsce w przypadku estymacji wartości mierzonej wielkości.

Niepewności pomiarowe (błędy przypadkowe)

W dalszej części tego rozdziału zajmiemy się analizą przypadkowych niepewności pomiarowych. Mają one decydujący wpływ na określenie ***dokładności*** i ***precyzji*** pomiarów, a więc i dokładności eksperymentu naukowego.

Pojęcie dokładności odnosi się zarówno do wyniku pomiaru – wartości zmierzonej, jak i do przyrządu lub metody pomiarowej. Wartość zmierzona jest dokładna, jeżeli jest zgodna z wartością prawdziwą mierzonej wielkości. Jest to oczywiście nieosiągalny ideał, ponieważ wszystkie zmierzone wartości są bardziej lub mniej niedokładne. Jednakże analiza błędów pomiarowych może wykazać, że jedne wartości są dokładniejsze od innych. Podobnie charakteryzujemy przyrządy i metody pomiarowe jako bardziej lub mniej dokładne. Niektórym przyrządom przypisuje się umowne klasy dokładności. Na dokładność pomiarową składają się zarówno błędy przypadkowe jak i systematyczne.

Pojęcie precyzji jest związane z błędami przypadkowymi i odnosi się zarówno do wartości zmierzonych, jak i do przyrządów lub metod pomiarowych. Precyzja przyrządu lub metody pomiarowej zależy od pewnej przeciętnej wartości błędu przypadkowego, którym jest obarczony każdy wynik pomiaru. Wynik pomiaru otrzymany metodą bardzo precyzyjną ma mały błąd przypadkowy, zaś otrzymany metodą mniej precyzyjną ma większy błąd przypadkowy.

Błędy pomiarów podaje się jako ***bezwzględne*** lub ***względne***. Błąd bezwzględny jest wartością bezwzględną różnicy wartości zmierzonej i wartości prawdziwej, i jest miarą odchylenia wyniku pomiarowego od wartości prawdziwej:

$$\Delta x = |x - m|$$

Błąd względny (przeważnie wyrażany w procentach) jest stosunkiem błędu bezwzględnego do modułu wartości mierzonej:

$$\delta x = \Delta x / |x|$$

Wynik pomiaru dowolnej wielkości fizycznej możemy więc zaprezentować w postaci:

$$x \pm \Delta x = x \cdot (1 \pm \delta x)$$

Oszacowanie niepewności pomiarowych może być zadaniem skomplikowanym i trudnym. W przypadku pomiarów bezpośrednich możliwe jest określenie niepewności pomiarowej jako związanej z najmniejszą podziałką skali przyrządu wykorzystywanego w danym eksperymencie. Możliwe jest uzyskanie większej precyzji przy dokonaniu liniowej ***interpolacji***, czyli oceny odcinka między działkami skali. I tak np. linijką z najmniejszą podziałką milimetrową możemy dokonać pomiaru z niepewnością ± 1 mm, gdy dysponujemy taśmą mierniczą o podziałce 1cm, niepewność pomiarowa sięgnie tegoż właśnie ± 1cm, chyba że interpolacja długości przypadającej miedzy podziałkami pozwoli nam zmniejszyć niepewność do ± 0,5 cm lub nawet do ± 0,2 cm.

Ale nawet w przypadku prostych pomiarów bezpośrednich dokładność przyrządów pomiarowych niekoniecznie musi w bezpośredni sposób przenosić się na wynik pomiaru. Przykładem może być niepewność związana z określeniem (zdefiniowaniem) początku i końca mierzonego obiektu, czy zdefiniowaniem początku i końca eksperymentu przy pomiarach czasu. W efekcie dysponowanie bardzo dokładnym przyrządem niekoniecznie zapewnia określoną dokładność pomiaru, np. w ręcznym pomiarze czasu biegu sprinterskiego na 100 m refleks i doświadczenie osoby dokonującej pomiaru mają większy wpływ na dokładność pomiaru niż dokładność stopera. W takich przypadkach dopiero wielokrotne powtórzenie eksperymentu może ujawnić rzeczywistą niepewność pomiarową. Musi zostać jednak zagwarantowany warunek, że za każdym razem mierzymy rzeczywiście tę samą wielkość (np. kolejne próbki zawierają te same stężenia substratów przy pomiarach stałej szybkości reakcji). Warto również zdać sobie sprawę z faktu, iż wielokrotne powtarzanie pomiaru nie ujawnia błędów systematycznych, choć jest skuteczną metodą analizy przypadkowych niepewności pomiarowych.

Analiza statystyczna niepewności pomiarowych

Dokonując wielokrotnie pomiaru dowolnej wielkości fizycznej spodziewamy się, że otrzymamy zbiór wartości mierzonej wielkości. Choć w serii n pomiarów wielkości x (x_1, x_2,

$x_3 ... x_n$) część wyników może ulec powtórzeniu, zbiór różniących się wartości pozwala nam ocenić zarówno wartość prawdziwą mierzonej wielkości x, jak i niepewność pomiarową, na podstawie rozrzutu, rozproszenia (wariancji) otrzymanych wyników. Najczęściej używanym przybliżeniem wielkości prawdziwej jest ***średnia arytmetyczna*** wyników z próby:

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

Najczęściej używaną miarą niepewności pomiarowej (miarą rozproszenia wyników) jest z kolei ***wariancja*** (S^2) lub ***odchylenie standardowe z próby*** (S):

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

Przykład liczbowy: obliczanie średniej arytmetycznej i odchylenia standardowego;
10 pomiarów dało następujące wyniki:8,5; 9,1; 9,2; 10,1; 10,4; 11,4; 11,6; 11,8; 12,3; 12,6

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	x_i^2
8,5	-2,2	4,84	72,25
9,1	-1,6	2,56	82,81
9,2	-1,5	2,25	84,64
10,1	-0,6	0,36	102,01
10,4	-0,3	0,09	108,16
11,4	0,7	0,49	129,96
11,6	0,9	0,81	134,56
11,8	1,1	1,21	139,24
12,3	1,6	2,56	151,29
12,6	1,9	3,61	158,76
$\sum_{i=1}^{n} x_i = 107$	$\sum_{i=1}^{n} (x_i - \bar{x}) = 0$	$\sum_{i=1}^{n} (x_i - \bar{x})^2 = 18{,}78$	$\sum_{i=1}^{n} x_i^2 = 1163{,}68$

Uwaga: $\sum_{i=1}^{n} (x_i - \bar{x})^2 = \sum_{i=1}^{n} x_i^2 - n\bar{x}^2 = 1163{,}68 - 10*(10{,}7)^2 = 18{,}78$

Średnia $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = 10{,}7$; odchylenie standardowe $S = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}} = 1{,}4$

Wyniki przedstawione w tabeli wyraźnie pokazują, że średnia arytmetyczna jest dobrą miara wielkości przeciętnej, wokół której skupione są otrzymane wartości. Suma odchyleń wartości x_i od średniej $\bar{x}$, zgodnie z definicją średniej arytmetycznej, wynosi zero. Część uzyskanych wartości x_i jest większa niż średnia, część mniejsza. W rezultacie odchylenie średnie nie może być wykorzystane jako miara rozproszenia tych wartości. Taką miarę można jednak uzyskać licząc sumę kwadratów tych odchyleń i normalizując w zależności od liczby analizowanych wartości. Z tabeli wyraźnie jednak widać, że przedział $\bar{x} \pm S$ nie obejmuje wszystkich przedstawionych w tabeli wartości. Czy możemy w takim razie uznać S za miarę niepewności pomiarowej?

Rozkład normalny

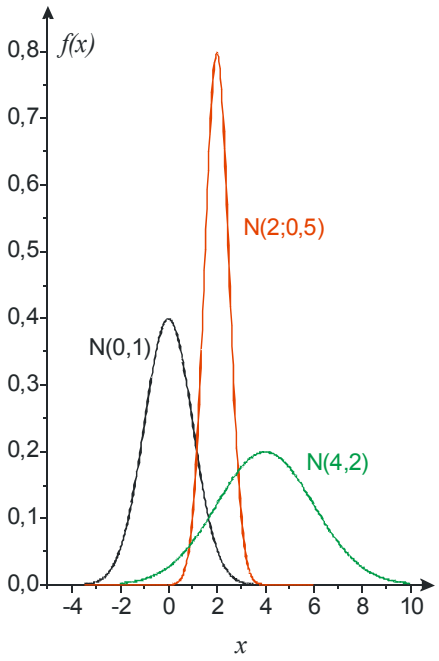
W celu uzyskana interpretacji odchylenia standardowego rozpatrzmy najważniejszy (z punktu widzenia praktycznych i teoretycznych zastosowań) rozkład prawdopodobieństwa: ***rozkład normalny*** (zwany też rozkładem Gaussa). ***Funkcja gęstości prawdopodobieństwa*** rozkładu normalnego posiada postać:

$$f_{m,\sigma}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}$$

gdzie m – jest wartością oczekiwaną (zwaną też wartością średnią), a σ - odchyleniem standardowym zmiennej losowej podlegającej temu rozkładowi. Jest ona znormalizowana:

$$\int_{-\infty}^{+\infty} f(x) \cdot dx = 1$$

co oznacza, że prawdopodobieństwo znalezienia x w całym zakresie od -∞ do +∞ wynosi 100%, a pole powierzchni pod wykresem funkcji wynosi 1.



Rys. 2 Funkcje gęstości prawdopodobieństwa rozkładów normalnych N(m, σ) o różnych wartościach średnich (m) i odchyleniach standardowych (σ)

Jako, że rozkład normalny zależy tylko od tych dwóch parametrów wystarczy symboliczny zapis N(m, σ) do jego oznaczenia. Pierwszy z parametrów określa wartość średnią rozkładu, wokół której jest on symetryczny, a drugi szerokość rozkładu. Uwaga: pola powierzchni pod zaprezentowanymi na rysunku 3 rozkładami wynoszą 1 (warunek normalizacji).

Znajomość funkcji gęstości prawdopodobieństwa pozwala określić prawdopodobieństwo znalezienia zmiennej losowej, podlegającej takiemu rozkładowi, w określonym przedziale:

$$P\{a < x < b\} = \int_a^b f_{m,\sigma}(x) \cdot dx$$

gdzie $f(x)\cdot dx$ oznacza prawdopodobieństwo znalezienia zmiennej x w przedziale od x do $x + dx$. W praktyce wygodnie jest korzystać ze standardowego rozkładu normalnego N(0,1), którego funkcja gęstości prawdopodobieństwa przyjmuje postać:

$$f(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}$$

a ***dystrybuanta*** $F(x)$:

$$F(x) = P\{u < x\} = \int_{-\infty}^{x} f(u) \cdot du$$

została stablicowana.

Podstawienie $u = (x - m)/\sigma$ pozwala dokonać zmiany zmiennych: x, który podlega rozkładowi normalnemu N(m, σ) na zmienną u, która podlega standaryzowanemu rozkładowi normalnemu N(0,1). Korzystając z powyższego podstawienia i dokonując zmiany granic całkowania możemy prawdopodobieństwo $P\{a < x < b\}$ wyrazić za pomocą wartości dystrybuanty, znalezionych w tablicach dystrybuanty rozkładu N(0,1):

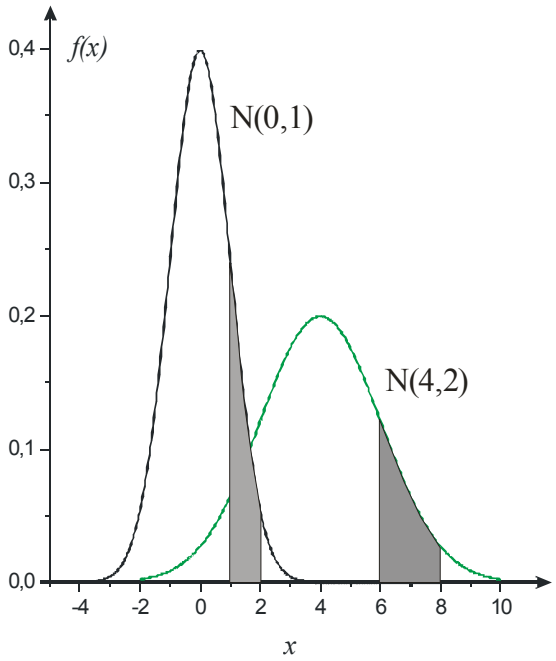
$$P\{a < x < b\} = \int_a^b f_{m,\sigma}(x) \cdot dx = \int_a^b \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}} \cdot dx = \int_{\frac{a-m}{\sigma}}^{\frac{b-m}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} \cdot du =$$

$$\int_{\frac{a-m}{\sigma}}^{\frac{b-m}{\sigma}} f_{0,1}(u) \cdot du = P\{\frac{a-m}{\sigma} < u < \frac{b-m}{\sigma}\} = F(\frac{b-m}{\sigma}) - F(\frac{a-m}{\sigma})$$

Przykład do rysunku 3:

x: N(4,2) $\Rightarrow$ u: N(0,1);

$P\{6 < x < 8\} = P\{1 < u < 2\}$- pola zacieniowane na rys. 3

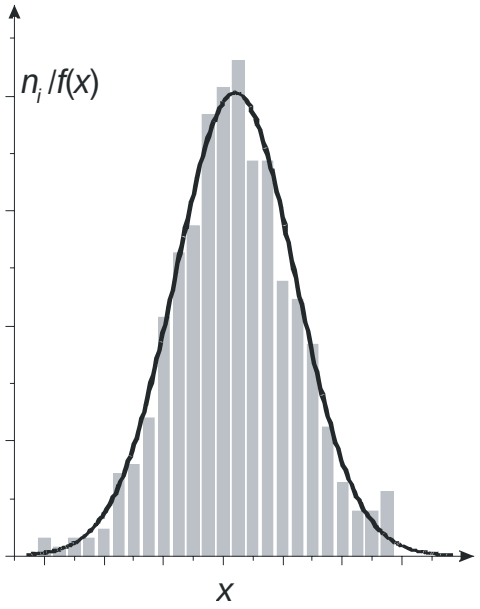


Rys. 3 Normalizacja rozkładu normalnego N(4,2) do standaryzowanego rozkładu normalnego N(0,1). Zacieniowane pole określa obszary równego prawdopodobieństwa w obu rozkładach

W szczególności prawdopodobieństwo uzyskania wielkości x w zakresie $\pm\ \sigma$ od wartości oczekiwanej m wynosi:

$$P\{ m - \sigma < x < m + \sigma\} \approx 0{,}68$$

Załóżmy, że wyniki serii pomiarów eksperymentalnych wielkości x podlegają takiemu właśnie rozkładowi normalnemu N(m, σ). Ciągła funkcja rozkładu oznacza, że mógłby on zostać osiągnięty w wyniku wykonania nieskończonej ilości pomiarów wielkości fizycznej x, której prawdziwa wartość wynosi m, a pomiar podlega wpływowi tylko błędów przypadkowych, czyli obarczony jest określoną niepewnością pomiarową. Jej miarą jest σ. Czy rzeczywiście możemy uznać taką interpretację za wiarygodną? Stosunkowo łatwo jest zaakceptować fakt, że wykresy (***histogramy***) opisujące zbiory skończone (o ograniczonej liczbie wartości np. n pomiarów x_1, x_2, x_3... x_n) będą dążyły wraz ze wzrostem ilości pomiarów do wykresów granicznych opisywanych funkcjami ciągłymi.



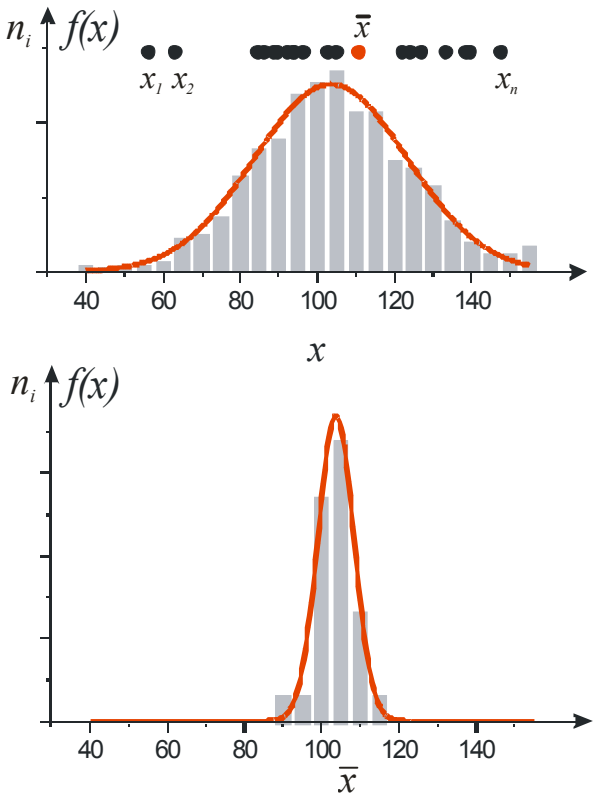
Rys. 4 Histogram skończonego zbioru wyników z dopasowaną funkcją gęstości prawdopodobieństwa rozkładu normalnego

Trudniej się jednak pogodzić z faktem, że rozkłady ciągłe jedynie asymptotycznie dążą do zera (czyli w przypadku rozkładu normalnego funkcja gęstości prawdopodobieństwa nie osiąga wartości 0 w całym przedziale od -∞ do +∞). A to oznacza określone prawdopodobieństwo uzyskania wyniku pomiaru skrajnie odbiegającego od wielkości prawdziwej. Faktycznie jednak funkcje gęstości prawdopodobieństwa są funkcjami szybko zbieżnymi do zera i na ogół, poza pewnym wąskim przedziałem, prawdopodobieństwo otrzymania wyniku staje się znikomo małe.

W rezultacie korzystanie z funkcji gęstości prawdopodobieństwa pozwala uzyskać wygodną interpretację wyników pomiarów eksperymentalnych i niepewności z nimi związanych. Można bowiem wykazać, że najlepszym estymatorem (oceną nieznanego parametru) wartości oczekiwanej m, a więc i wartości prawdziwej, jest średnia arytmetyczna serii pomiarów, a wariancji zmiennej losowej, kwadrat odchylenia standardowego z próby, S^2. Uzyskujemy w ten sposób interpretację niepewności pomiarowej, wyrażonej za pomocą odchylenia standardowego z próby, S, jako 68% prawdopodobieństwo otrzymania wyniku pomiaru w przedziale $x \pm S$ (lub 95% prawdopodobieństwo otrzymania wyniku w przedziale $x \pm 2S$, itp.)

Rozkład średniej arytmetycznej

W powyższym omówieniu ograniczyliśmy się do rozkładu normalnego, choć w niektórych przypadkach bardziej odpowiednie do stosowanych technik pomiarowych mogą być rozkłady ciągłe innego typu (inna funkcja gęstości prawdopodobieństwa) np. rozkład Poissona, rozkład logarytmiczno-normalny, rozkład χ^2. Znaczenie rozkładu normalnego i jego szerokie wykorzystanie w naukach przyrodniczych wynika z działania ***centralnego twierdzenia granicznego*** mówiącego, że suma dużej liczby zmiennych losowych niezależnych ma asymptotyczny rozkład normalny. Innymi słowy jeżeli wynik pomiaru narażony jest na wpływ wielu źródeł niewielkich i przypadkowych błędów, a błędy systematyczne są zaniedbywalne, uzyskiwane wartości najlepiej będą opisywane granicznym rozkładem normalnym.

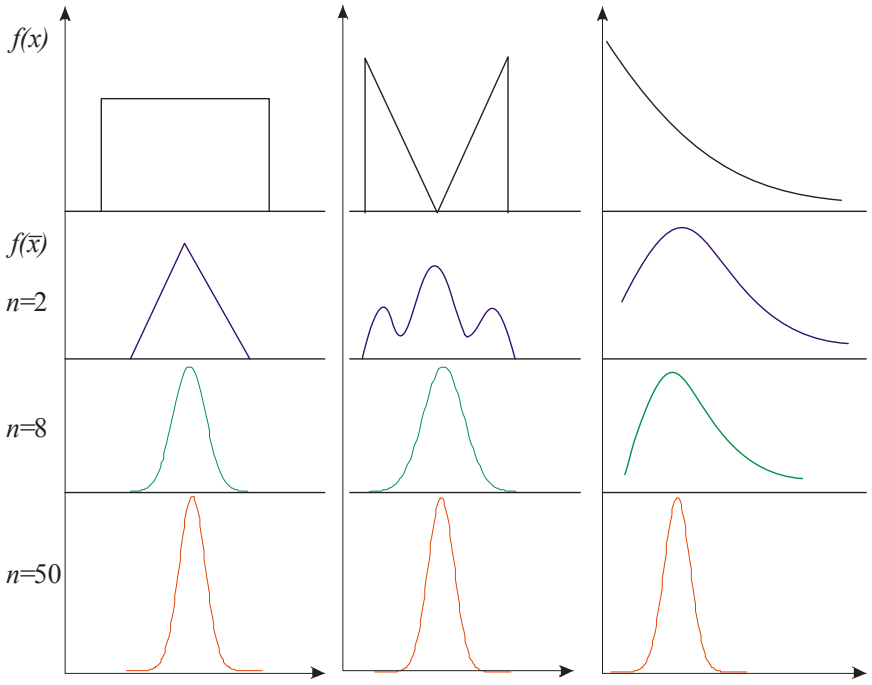


Rys. 5 Porównanie rozkładu zmiennej x (wykres górny) i rozkładu średniej (wykres dolny). Na wykresie zaprezentowano odpowiednie histogramy i funkcje gęstości prawdopodobieństwa rozkładu normalnego. Punkty x_1, x_2...x_n reprezentują przykładowy zbiór wyników serii n pomiarów zmiennej x; $\bar{x}$ średnia serii

Zastanówmy się obecnie nad prostym pytaniem o rozkład wartości średniej. Załóżmy, że wyniki serii pomiarów eksperymentalnych podlegają rozkładowi normalnemu N(m, σ). Dokonując serii n pomiarów oczekujemy, że ich wyniki ułożą się zgodnie z funkcją gęstości prawdopodobieństwa danego rozkładu. Można więc oczekiwać, że największa liczba wyników skupiona będzie wokół wartości oczekiwanej m, choć znajdziemy wśród nich wartości mniejsze i większe, oraz odbiegające mniej lub bardziej od wartości oczekiwanej m. Jednakże licząc wartość średnią wyników serii pomiarowej możemy spodziewać się, że wartość ta będzie po pierwsze odbiegać znacznie mniej od wartości oczekiwanej m, niż skrajne wyniki pojedynczych pomiarów, a po drugie tym bardziej będzie zbliżona do m, im więcej wyników jest do tej wartości zbliżonych. Im więcej pomiarów wykonamy w celu policzenia z nich wartości średniej tym jej odchylenie od wartości oczekiwanej m powinno być mniejsze. Jednocześnie wykonując kilkanaście takich serii pomiarowych nie oczekujemy, że policzone wartości średnie będą identyczne, ale raczej, że podlegać będą również rozkładowi. Rozkład wartości średniej będzie oczywiście skupiony wokół tej samej wartości

oczekiwanej m, co rozkład pojedynczego pomiaru, ale wariancja wartości średniej będzie znacznie mniejsza.

Można wykazać, że jeżeli x podlega rozkładowi normalnemu N(m, σ), to $\bar{x}$ podlega rozkładowi normalnemu N(m, $\frac{\sigma}{\sqrt{n}}$). Co więcej, niezależnie od rozkładu wielkości x (może to być rozkład opisywany dowolną funkcją gęstości prawdopodobieństwa, ze skończoną średnią i wariancją), graniczny rozkład średniej arytmetycznej $\bar{x}$ (dla dużej ilości pomiarów, $n > 50$) będzie rozkładem normalnym N(m, $\frac{\sigma}{\sqrt{n}}$). Jest to kolejny argument podkreślający ważność rozkładu normalnego i nosi nazwę twierdzenia Lindeberga-Levy'ego.



Rys. 6 Rozkłady średnich $f(\bar{x})$ uzyskiwanych z serii 2, 8 lub 50 pomiarów zmiennej x podlegającej rozkładowi $f(x)$

Podsumujmy, jakie wnioski wynikają z powyższych rozważań. Po pierwsze uzasadniają one wybór średniej jako wartości poprawnej, czyli najlepszego przybliżenia wartości prawdziwej. Po drugie, przybliżenie to jest tym dokładniejsze (obarczone mniejszym błędem) im większa liczba pomiarów została wykonana. W granicznym przypadku nieskończonej ilości pomiarów doszlibyśmy do prawdziwej wartości mierzonej wielkości fizycznej. W praktyce trzeba sobie jednak zdawać sprawę z faktu, że nasze wcześniejsze założenie o braku błędów systematycznych jest obrazem wyidealizowanym. O ile

zwiększanie liczby pomiarów pozwala zminimalizować błędy przypadkowe, statystyczne o tyle nie zmniejszy błędów systematycznych i ich wpływu na błąd pomiarowy.

Przedziały ufności – estymacja przedziałowa

Oszacowanie (estymacja) nieznanej wartości mierzonej wielkości fizycznej za pomocą pojedynczego parametru, np. poprzez wykorzystanie wartości średniej arytmetycznej jako najlepszego przybliżenia wartości prawdziwej, nazywane jest w statystyce metodą ***estymacji punktowej***. Pewną miarą niepewności estymacji z wykorzystaniem średniej arytmetycznej może być odchylenie standardowe z próby, *S*, choć jak już wspomnieliśmy powyżej rozkład średniej nie pokrywa się z rozkładem mierzonej wielkości *x*. Wygodnie jest więc w oparciu o rozkład wartości średniej dokonać ***estymacji przedziałowej*** wartości mierzonej wielkości, np. metodą ***przedziałów ufności*** stworzoną przez polskiego matematyka J. Neymana. Estymacja przedziałowa dokonuje szacunku w postaci podania przedziału wartości, który z dużym prawdopodobieństwem obejmuje wartość prawdziwą.

Jeżeli założymy, że wielkość *x* podlega rozkładowi normalnemu N(m, σ) to średnia arytmetyczna $\bar{x}$ z n - elementowej serii pomiarowej podlega rozkładowi N($m, \frac{\sigma}{\sqrt{n}}$). Dokonując podstawienia $u = \frac{\bar{x} - m}{\sigma} \cdot \sqrt{n}$ otrzymamy dla u standaryzowany rozkład N(0,1). Załóżmy, że dokonamy estymacji przedziałowej wielkości u. Możemy dokładnie policzyć z jakim prawdopodobieństwem wielkość u leży w przedziale ($-u_\alpha$, u_α):

$$P\{-u_\alpha < u < u_\alpha\} = 2F(u_\alpha)-1$$

Możemy również sytuację odwrócić. Wybierając arbitralnie prawdopodobieństwo, z którym chcemy stworzyć ten przedział znajdziemy taką wartość u_α (rys. 7), która spełni nasze wymagania:

$$P\{-u_\alpha < u < u_\alpha\} = 2F(u_\alpha)-1 = 1-\alpha$$

co jest równoważne przedziałowi ufności dla m:

$$P\{\bar{x} - u_\alpha \frac{\sigma}{\sqrt{n}} < m < \bar{x} + u_\alpha \frac{\sigma}{\sqrt{n}}\} = 1-\alpha$$

Konstrukcja przedziału ufności:

x: N(m, σ) $\Rightarrow \bar{x}$: N(m, $\frac{\sigma}{\sqrt{n}}$) $\Rightarrow$ dla $u = \frac{\bar{x} - m}{\sigma} \cdot \sqrt{n}$ $\Rightarrow$ u: N(0,1)

$\Downarrow$

$$P\{\bar{x} - u_\alpha \frac{\sigma}{\sqrt{n}} < m < \bar{x} + u_\alpha \frac{\sigma}{\sqrt{n}}\} = 1-\alpha \quad \Leftarrow \quad P\{-u_\alpha < u < u_\alpha\} = 2F(u_\alpha)-1 = 1-\alpha$$

$$\Downarrow$$

$$m \in (\bar{x} \pm \frac{u_\alpha \sigma}{\sqrt{n}})$$

dla wybranego $1-\alpha \Rightarrow u_\alpha$

np. $1-\alpha = 0{,}95 \Rightarrow u_\alpha = 1{,}96$

Przekształcenia:

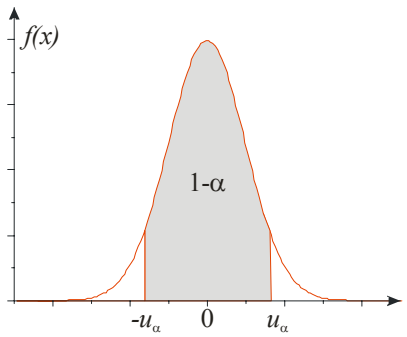
$P\{-u_\alpha < u < u_\alpha\} = 2F(u_\alpha)-1 = 1-\alpha$

$u = (\bar{x} - m)\cdot\sqrt{n}/\sigma$

$$P\{-u_\alpha < \frac{\bar{x}-m}{\sigma}\cdot\sqrt{n} < u_\alpha\} = 1-\alpha \qquad |\cdot\sigma/\sqrt{n}$$

$$P\{-u_\alpha \frac{\sigma}{\sqrt{n}} < \bar{x} - m < u_\alpha \frac{\sigma}{\sqrt{n}}\} = 1-\alpha \qquad |-\bar{x};\cdot(-1)$$

$$P\{\bar{x} - u_\alpha \frac{\sigma}{\sqrt{n}} < m < \bar{x} + u_\alpha \frac{\sigma}{\sqrt{n}}\} = 1-\alpha$$



Rys. 7 Sposób znajdywania wartości u_α przy konstrukcji przedziałów ufności

W rezultacie otrzymaliśmy przedział ufności, w którym wartość prawdziwa *m* jest zawarta z prawdopodobieństwem $1-\alpha$.

$$m \in (\bar{x} \pm \frac{u_\alpha \sigma}{\sqrt{n}})$$

Wielkość $1-\alpha$ nazywamy ***współczynnikiem ufności***, a stworzony dla wybranego współczynnika przedział nazywamy przedziałem ufności. Wartość $1-\alpha$ przyjmuje się subiektywnie, jako dowolnie duże, bliskie 1, prawdopodobieństwo. Jest ono miarą zaufania do prawidłowego szacunku. Najczęściej wybierane wartości $1-\alpha$ to 0,95 lub 0,99, a znalezione dla nich wartości u_α wynoszą odpowiednio 1,96 i 2,575.

Zaletą estymacji przedziałowej w postaci przedziałów ufności jest precyzyjne określenie niepewności pomiarowej poprzez wyznaczenie granic przedziału, w którym mierzona wartość prawdziwa jest zawarta z określonym prawdopodobieństwem. Warto

zwrócić uwagę na fakt, że precyzja estymacji przedziałowej zależy od dwóch czynników: wybranego współczynnika ufności 1-α (wpływ na wartość u_α) i liczby wykonanych pomiarów, n. Wymagając większego poziomu ufności w wykonane oznaczenie, przedział ufności ulega poszerzeniu (wzrasta wartość u_α). Można jednak temu przeciwdziałać zwiększając liczbę pomiarów. W rezultacie można z góry zaplanować liczbę pomiarów niezbędnych do osiągnięcia określonej precyzji (zwanej często ***maksymalnym błędem szacunku***, d - równym połowie wyznaczonego przedziału):

$$m \in (\bar{x} \pm d), \text{ gdzie } d = u_\alpha \frac{\sigma}{\sqrt{n}}$$

stąd dla pożądanego d należy wykonać co najmniej n pomiarów:

$$n > \frac{u_\alpha^2 \cdot \sigma^2}{d^2}$$

Podkreślmy jednak po raz kolejny, że tego typu działania prowadzące do zwiększenia precyzji pomiarowej są skuteczne jedynie w odniesieniu do błędów przypadkowych.

Nie wspomnieliśmy jednak do tej pory jaką wartość σ wykorzystać we wzorze na przedział ufności. Możemy sobie wyobrazić sytuację, że wariancja wyników pomiarowych σ^2 jest znana mimo, że przystępujemy do pomiaru nieznanej wielkości. Tak może się zdarzyć jeżeli dysponujemy właściwie skalibrowanym układem pomiarowym, np. poprzez wykonanie podobnych pomiarów wielkości fizycznych, których wartość prawdziwa jest znana. Typowym przykładem może być tutaj pracownia studencka. Z drugiej zaś strony wykonując dużą ilość pomiarów (n > 50) estymacja punktowa σ, którą uzyskujemy za pomocą odchylenia standardowego S staje się na tyle precyzyjna, że możemy w miejsce σ wykorzystać wartość odchylenia standardowego z próby, S.

Jeżeli jednak zarówno wartość prawdziwa jak i wariancja wielkości x: N(m, σ), pozostają nieznane, a liczba pomiarów jest również ograniczona (n < 50) konstrukcja przedziału ufności musi zostać zmieniona. Okazuje się, że wielkość:

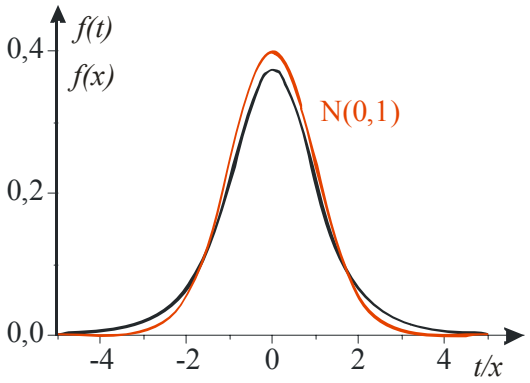
$$t = \frac{\bar{x} - m}{S} \cdot \sqrt{n}$$

podlega ***rozkładowi t – Studenta***; nazwa rozkładu pochodzi od pseudonimu naukowego „Student” angielskiego matematyka W. Gosseta, który powyższe twierdzenie udowodnił. Funkcję rozkładu gęstości prawdopodobieństwa rozkładu t – Studenta, w którym zwyczajowo w miejsce zmiennej x stosuje się symbol t, przedstawia poniższy wzór, w którym mamy tylko jeden niezależny parametr k, tzw. ***liczbę stopni swobody***:

$$f(t) = \frac{\Gamma(\frac{k+1}{2})}{\sqrt{k\pi}\cdot\Gamma(k/2)} \cdot \frac{1}{(1+\frac{t^2}{k})^{\frac{k+1}{2}}}$$

gdzie: $k = n$-1 liczba stopni swobody, $\Gamma(p) = \int_0^{+\infty} x^{p-1}\cdot e^{-x}\cdot dx$ dla p>0

Funkcja gęstości prawdopodobieństwa rozkładu t – Studenta odbiega nieznacznie od funkcji rozkładu N(0,1), szczególnie dla niskich wartości k; jest ona wolniej zbieżna do zera niż rozkład normalny. Jednocześnie gdy rośnie liczba stopni swobody różnica miedzy oboma rozkładami szybko „zanika"; praktycznie oba rozkłady stają się na tyle zbliżone dla $n > 50$, że odpowiedni przedział ufności może zostać skonstruowany wg wcześniej omówionej metody.



Rys. 8 Porównanie funkcji gęstości prawdopodobieństwa rozkładu t-Studenta ($k = 4$) i rozkładu normalnego N(0,1)

Korzystając z podstawienia $t = \frac{\bar{x} - m}{S}\cdot\sqrt{n}$ przedział ufności, skonstruowany na podstawie n – elementowego zbioru wyników pomiarów, o wartości średniej $\bar{x}$ i odchyleniu standardowym S, można przedstawić w następujący sposób:

$$m \in (\bar{x} \pm \frac{t_\alpha S}{\sqrt{n}})$$

W zależności od wybranego arbitralnie współczynnika ufności 1-α parametr t_α znajdujemy z tablic rozkładu t – Studenta, zazwyczaj skonstruowanych na odmiennej zasadzie niż tablice dystrybuanty: dla odpowiedniej wartości 1-α i liczby stopni swobody $k = n$-1 bezpośrednio podane są wartości t_α.

Podsumowując:

Wykonując n pomiarów $x_1, x_2 ... x_n$, licząc $\bar{x}$, S i wybierając poziom ufności 1-α (0,95; 0,99 ..), przedział ufności dla średniej tworzymy wg wzorów podanych w tabeli poniżej:

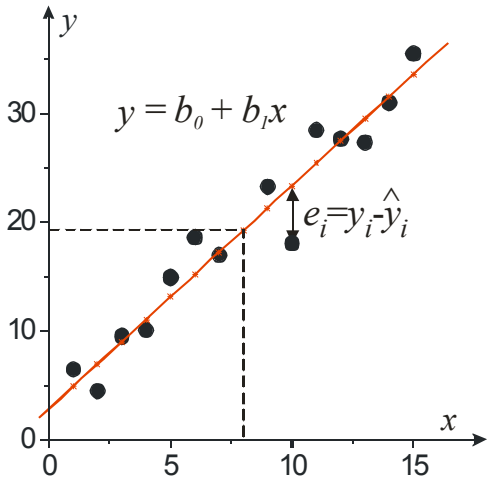
Tabela 1. Przedziały ufności

Model I x: N(m, σ), σ znane lub σ nie znane, n>50, $\sigma = S$	Model II x: N(m,σ), σ nie znane, n<50
$m \in (\bar{x} \pm \frac{u_\alpha \sigma}{\sqrt{n}})$	$m \in (\bar{x} \pm \frac{t_\alpha S}{\sqrt{n}})$
Znajdowanie u_α: P{-u_α < u < u_α} = 2F(u_α)-1 = 1-α tablice dystrybuanty rozkładu normalnego	Znajdowanie t_α: P{| t |> t_α} = α tablice wartości t_α dla różnych $k = n$-1 i α

Zależność wielkości fizycznych

Zrozumienie metodyki pomiaru wielkości fizycznej, omówionej powyżej, pozwala nam przystąpić do odkrywania podstawowych praw wiążących różne wielkości fizyczne. Określony model fizyczny badanego zjawiska, poparty odpowiednim opisem matematycznym (wzorem), pozwala zrozumieć naturę badanych zjawisk. Dla eksperymentatora dokonującego nowych odkryć naukowych czy szukającego potwierdzenia teoretycznych rozważań kluczowe jest przede wszystkim uzyskanie potwierdzenia istnienia zależności dwóch wielkości fizycznych. Doświadczenie zaprojektowane w celu potwierdzenia i znalezienia tego związku między dwiema wielkościami fizycznymi x i y musi polegać na pomiarze obu tych wielkości równocześnie. Co więcej pomiar taki nie może się ograniczyć tylko do otrzymania pojedynczej pary wartości obu wielkości, ale niezbędne jest uzyskanie co najmniej kilku takich punktów we współrzędnych zmiennych x i y. Najczęściej stosowaną metodą prowadzącą do tego celu jest takie zaprojektowanie eksperymentu, w którym jednej z tych wielkości (zwanej ***zmienną niezależną***) przyporządkowane są wartości drugiej wielkości, zwanej ***zmienną zależną***. Związek między dwiema wielkościami może być wzajemny, zmiany jednej zmiennej powodują zmiany drugiej i odwrotnie. Często jednak wyraźnie można wyróżnić zmienną zależną i niezależną, co uzasadnia wspomniany model eksperymentu, w którym eksperymentator zmienia jedną z wielkości , np. stężenie substancji i mierzy zmianę wartości drugiej wielkości np. szybkość reakcji. Gdyby udało się zmierzyć

obie te wielkości bezbłędnie, w szerokim zakresie zmian jednej z nich, niezbyt skomplikowane dopasowanie odpowiedniej funkcji matematycznej pozwoliłoby uzyskać odpowiedni opis matematyczny zjawiska. Czy jednak fakt, że pomiar obu wielkości obarczony jest niepewnością pomiarową nie powoduje takiego utrudnienia tego zadania, iż staje się ono niewykonalne? Na rysunku poniżej przedstawiono przykład eksperymentu, w którym wyraźnie widoczna jest zależność zmiennej y od zmian zmiennej x, zależność prawdopodobnie liniowa. Powstaje jednak pytanie jak najlepiej tego typu zależność liniową narysować wśród porozrzucanych na wykresie punktów eksperymentalnych.



Rys. 9 Wykres punktowy – graficzna ilustracja metody najmniejszych kwadratów

Ustalenie postaci funkcji, która opisuje zależność między zmiennymi nazywamy ***regresją***. Z praktycznego punktu widzenia ograniczymy się w dalszej dyskusji do regresji II rodzaju, czyli liniowej lub nieliniowej funkcji $y = f(x)$ znalezionej ***metodą najmniejszych kwadratów***.

Metoda najmniejszych kwadratów

Omówimy metodę najmniejszych kwadratów w oparciu o funkcję liniową, po pierwsze ze względu na jej przejrzystość, a po drugie ważność zależności liniowych w naukach przyrodniczych.

Załóżmy, że istnieje prawdziwa liniowa zależność wielkości y od x, którą możemy zapisać w postaci:

$$y = \beta_0 + \beta_1 \cdot x$$

Estymacja punktowa nieznanych parametrów β_0 i β_1 pozwala nam znaleźć najbardziej optymalne ich oszacowanie, czyli przedstawić poszukiwaną zależność liniową w postaci:

$$y = b_0 + b_1 \cdot x$$

Metoda najmniejszych kwadratów polega na takim wyborze parametrów b_0 i b_1, aby suma kwadratów reszt e_i (rys. 9) osiągnęła minimum.

$$Q = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} \left(y_i - \hat{y}_i\right)^2 = \sum_{i=1}^{n} (y_i - b_1 x - b_0)^2 = \min.$$

Oznacza to, że spośród możliwych do wyobrażenia prostych jakie moglibyśmy narysować na wykresie $y = f(x)$, wybierzemy tę, dla której suma kwadratów odchyleń punktów od prostej przyjmie wartość minimalną. Zwróćmy uwagę, że Q jest funkcją dwóch zmiennych b_0 i b_1 (wybór różnych wartości b_0 i b_1, prowadzi do otrzymania dla danego zbioru punktów o współrzędnych (x_i, y_i) różnych wartości Q). Stąd warunkiem koniecznym znalezienia minimum funkcji jest zerowanie się pochodnych cząstkowych względem obu zmiennych (jest to zarazem warunek wystarczający). W rezultacie otrzymujemy układ dwóch równań i dwóch niewiadomych:

$$\begin{cases} \dfrac{\partial Q}{\partial b_1} = 2\sum_{i=1}^{n} (y_i - b_1 x - b_0)(-x_i) = 0 \\ \dfrac{\partial Q}{\partial b_0} = 2\sum_{i=1}^{n} (y_i - b_1 x - b_0)(-1) = 0 \end{cases}$$

$$\begin{cases} b_1 \sum x_i^2 + b_0 \sum x_i = \sum x_i y_i \\ b_1 \sum x_i + n \cdot b_0 = \sum y_i \end{cases}$$

pozwalający wyznaczyć b_0 i b_1:

$$b_1 = \frac{n\sum x_i y_i - \left(\sum x_i\right)\left(\sum y_i\right)}{n\sum x_i^2 - \left(\sum x_i\right)^2} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}$$

$$b_0 = \frac{\left(\sum x_i^2\right)\left(\sum y_i\right) - \left(\sum x_i\right)\left(\sum x_i y_i\right)}{n\sum (x_i - \bar{x})^2}$$

$$b_0 = \bar{y} - b_1 \cdot \bar{x}$$

Współczynniki b_0 i b_1 noszą nazwę ***współczynników regresji liniowej*** z próby. Najprostszą miarą dokładności wyznaczenia współczynników regresji są tzw. ***błędy standardowe***:

$$S = \sqrt{\frac{1}{n-2}\sum (y_i - \hat{y}_i)^2} = \sqrt{\frac{1}{n-2}\sum (y_i - b_1 x - b_0)^2}$$ resztkowe odchylenie standardowe

$$S_{b_1} = \frac{S}{\sqrt{\sum (x_i - \bar{x})^2}} = \frac{S}{\sqrt{\sum x_i^2 - n \cdot \bar{x}^2}}$$ błąd standardowy wsp. kierunkowego

$$S_{b_0} = S \cdot \sqrt{\frac{\sum x_i^2}{n\sum (x_i - \bar{x})^2}} = S \cdot \sqrt{\frac{1}{n} + \frac{\bar{x}^2}{\sum (x_i - \bar{x})^2}}$$ błąd standardowy wyrazu wolnego

Warto w tym miejscu podkreślić fakt, że metoda najmniejszych kwadratów pozwala znacznie precyzyjniej określić niepewności pomiarowe np. za pomocą konstrukcji odpowiednich przedziałów ufności dla współczynników regresji, i dla samej krzywej, czy przedziałów tolerancji pozwalających odrzucać skrajne punkty eksperymentalne. Zainteresowanych odsyłamy do wybranych pozycji literatury. Warto również nadmienić, że wiele popularnych programów komputerowych, arkuszy kalkulacyjnych w prosty sposób pozwala te wielkości znaleźć bez większego trudu.

Przykład liczbowy: Obliczanie podstawowych parametrów charakteryzujących liniową zależność $y = b_0 + b_1 x$

	x	y	$(x_i - \bar{x})^2$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$	$\hat{y}_i$	$(y_i - \hat{y}_i)^2$
	1	8	9	64	24	9,571	2,4694
	2	13	4	9	6	11,714	1,6531
	3	14	1	4	2	13,857	0,0204
	4	17	0	1	0	16	1
	5	18	1	4	2	18,143	0,0204
	6	20	4	16	8	20,286	0,0816
	7	22	9	36	18	22,429	0,1837
Suma	28	112	28	134	60		5,4286
Średnia	4	16					

$b_1 = 60/28 = 2{,}14$; $b_0 = 16 - 2{,}14 \cdot 4 = 7{,}43$; $S = \sqrt{5{,}4286/5} = 1{,}04$; $r = 60/\sqrt{28 \cdot 134} = 0{,}98$;

$S_{b1} = 1{,}04/\sqrt{28} = 0{,}20$; $S_{b0} = 1{,}04\sqrt{\frac{1}{7} + \frac{16}{28}} = 0{,}88$;

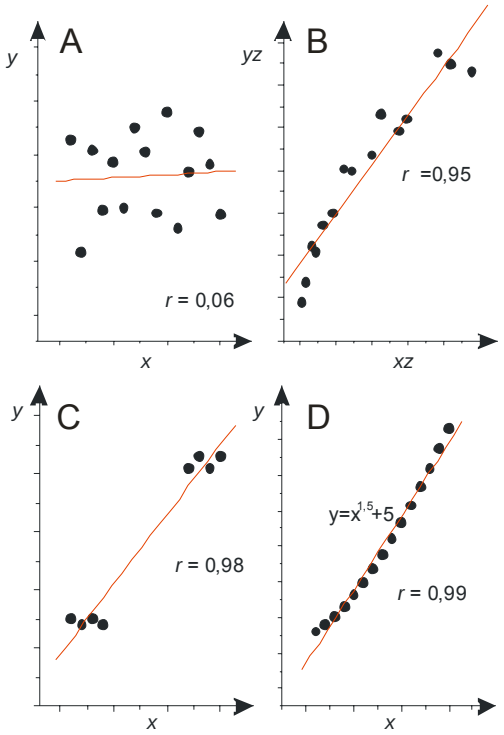
Najczęściej wykorzystywanym parametrem służącym ocenie miary liniowej zależności dwóch wielkości jest ***współczynnik korelacji z próby***:

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}$$

Współczynnik korelacji z próby może przyjmować wartości z zakresu $-1 \leq r \leq 1$, przy czym im jego wartość jest bliższa ± 1 tym silniejsze jest potwierdzenie liniowej zależności zmiennych *x* i *y* (+1 wskazuje na zależność wprost proporcjonalną, a -1 odwrotnie proporcjonalną). Gdy wartość współczynnika korelacji przyjmuje wartość bliską zeru o zmiennych *x* i *y* mówimy, że są nieskorelowane. Prawdziwe pozostają stwierdzenia:

- zmienne niezależne są nieskorelowane (twierdzenie odwrotne nie jest jednak prawdziwe),
- zmienne skorelowane są zależne.

Warto jednak pamiętać, że współczynnik korelacji z próby jest jedynie estymatorem współczynnika korelacji, dla którego powyższe stwierdzenia pozostają bezwzględnie prawdziwe. Jak dla każdego estymatora wiarygodność szacunku tak uzyskana wzrasta wraz z liczbą punktów pomiarowych, stąd należy zachować ostrożność w wykorzystaniu współczynnika korelacji z próby, jako argumentu na rzecz lub przeciw korelacji badanych zmiennych, na podstawie niezbyt licznych prób. Współczynnik korelacji z próby niesie ponadto dodatkowe niebezpieczeństwa. Wiele zależności nieliniowych może w ograniczonym przedziale być dobrze aproksymowana zależnościami liniowymi, np. dla przedstawionej na rysunku 10 zależności potęgowej (D) uzyskujemy dość wysoki współczynnik korelacji. Innym przykładem może być pojawienie się korelacji dla zmiennych nieskorelowanych (A), gdy wpływ na obie te zmienne ma trzeci czynnik (B), np. podczas, gdy *y* i *x* są nieskorelowane, wspólny czynnik *z* doprowadzi do wysokiej korelacji *yz* od *xz*. I tak np. wahania temperatury w trakcie eksperymentu wpływające zarówno na zmienną *x* jak i *y* mogą spowodować, że doszukamy się niefizycznej zależności obu tych zmiennych. Kolejnym przykładem błędnej oceny korelacji jest tzw. pozorna korelacja (C). Choć w wąskich zakresach zmienne pozostają nieskorelowane, drastyczna zmiana warunków eksperymentalnych odmiennie lokuje wyniki eksperymentalne na wykresie, w rezultacie prowadząc do uzyskania wysokiego współczynnika korelacji. Jest to przykład analogiczny do przykładu omówionego powyżej (A, B), gdy bliżej nieokreślony czynnik (bardzo często metodyczny bądź aparaturowy) jest odpowiedzialny za przesunięcie między obiema grupami punktów.



Rys. 10 Przykłady błędnego wykorzystania współczynnika korelacji z próby jako miary zależności liniowej zmiennych

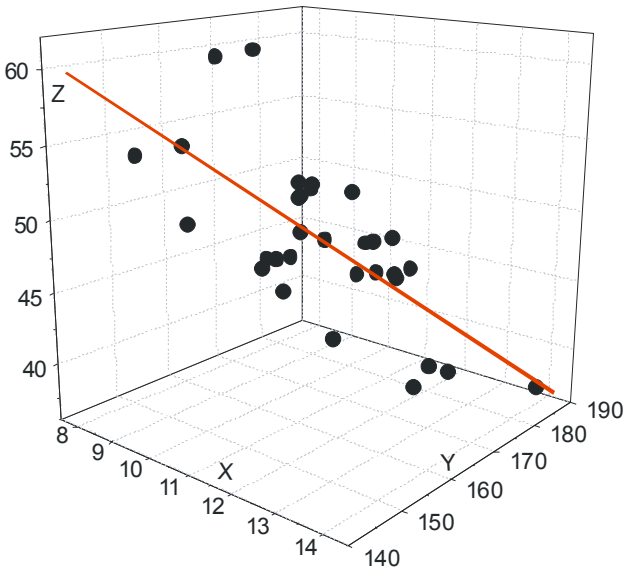
W rezultacie argumentacja na korzyść wniosku o istnieniu zależności zmiennych oparta na współczynniku korelacji z próby, powinna być wsparta logiczną interpretacją fizycznej strony tej zależności, dodatkową analizą współczynnika korelacji z próby i ewentualnie wykonaniem dodatkowych pomiarów.

Metodę najmniejszych kwadratów moglibyśmy analogicznie zastosować do wielomianów wyższego stopnia niż pierwszy. Przyrównanie do zera pochodnych cząstkowych względem nieznanych parametrów b_i prowadzi do układu k równań i k niewiadomych dla dowolnego wielomianu k – stopnia. Częściej jednak wykorzystuje się metody najmniejszych kwadratów w przypadku regresji wielorakiej, czyli zależności zmiennej zależnej y od więcej niż jednej zmiennej niezależnej, co pozwala poszukiwać bardziej złożonych zależności. Metodyka znajdywania odpowiedniego układu równań jest w tym przypadku równie prosta jak dla wielomianu i zapisana w postaci macierzowej dla $z = f(x, y)$ wygląda następująco:

$$z = A + Bx + Cy$$

$$\begin{bmatrix} n & \sum x_i & \sum y_i \\ \sum x_i & \sum x_i^2 & \sum y_i x_i \\ \sum y_i & \sum y_i x_i & \sum y_i^2 \end{bmatrix} \cdot \begin{bmatrix} A \\ B \\ C \end{bmatrix} = \begin{bmatrix} \sum z_i \\ \sum z_i \cdot x_i \\ \sum z_i \cdot y_i \end{bmatrix}$$

$$\begin{cases} A \cdot n + B\sum x_i + C\sum y_i = \sum z_i \\ A\sum x_i + B\sum x_i^2 + C\sum y_i x_i = \sum z_i \cdot x_i \\ A\sum y_i + B\sum y_i x_i + C\sum y_i^2 = \sum z_i \cdot y_i \end{cases}$$

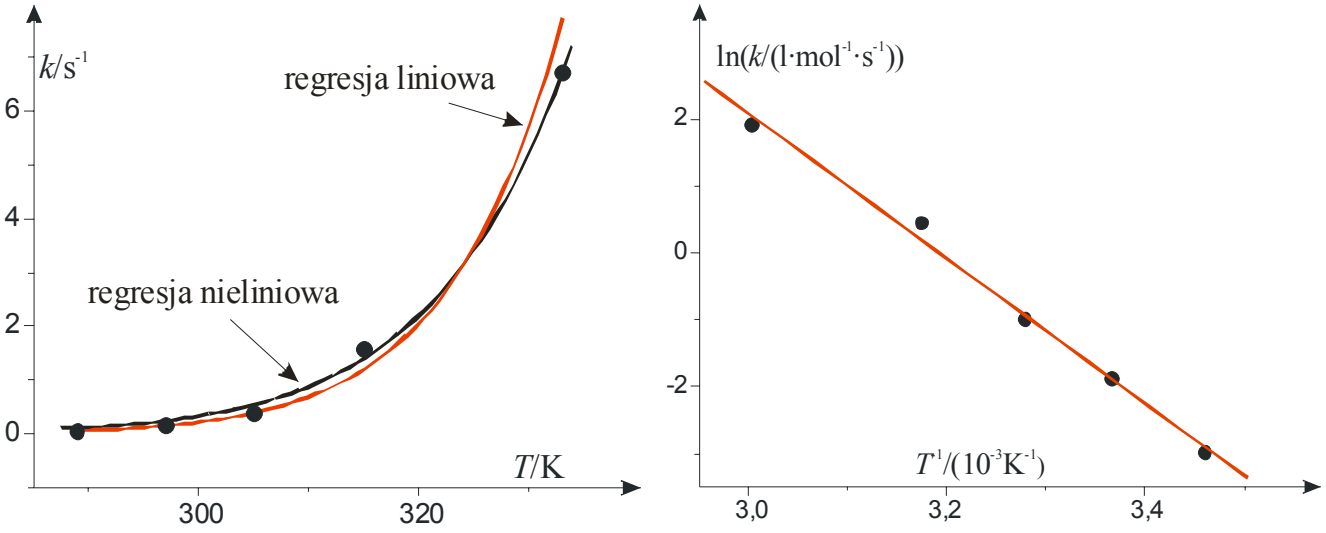


Rys. 11 Regresja wieloraka $z = Ax + By + C$

Pozwala ona określać zależność zmiennej zależnej od wielu czynników i wybrać spośród nich te, które mają dominujący wpływ na zmienność z.

Równie ważnym zagadnieniem są metody regresji liniowej stosowane do ważnych z punktu widzenia nauk przyrodniczych zależności nieliniowych, które mogą zostać sprowadzone do zależności liniowych poprzez odpowiednie podstawienie lub przekształcenie. Np. stosunkowo skomplikowana eksponencjalna zależność stałej szybkości reakcji od temperatury, znana pod nazwą zależności Arrheniusa może w wyniku zlogarytmowania obu stron zostać sprowadzona do zależności liniowej ln k od $1/T$:

$$k = A \exp(-E_A/RT) \Rightarrow \ln k = \ln A - (E_A/R)\cdot(1/T)$$



Rys. 12 Zależność Arrheniusa – porównanie wyników uzyskanych metodą regresji linearyzowanej (czerwona linia na wykresach w różnych układach współrzędnych) i regresji nieliniowej (czarna linia)

Jeżeli więc w miejsce zależności k od T metoda najmniejszych kwadratów zostanie zastosowana do zależności ln k od $1/T$, pozwoli to znaleźć ln A oraz E_A/R, i po prostych przekształceniach współczynnik przedeksponencjalny A i energię aktywacji E_A.

O ile jednak dokładność metody najmniejszych kwadratów, np. w przypadku wielomianu, prowadzi do najlepszego oszacowania nieznanych współczynników, o tyle metoda „***linearyzacji***" wprowadza dodatkowe błędy związane z charakterem zastosowanego podstawienia lub przekształcenia (rys. 12). Można te błędy ograniczyć stosując funkcje ***regresji nieliniowej*** z wykorzystaniem wysoce sprawnych i ekonomicznych metod minimalizacji funkcji metodami iteracyjnymi, jak powszechne obecnie metody Levenberga-Marquarda czy metoda Simplex. Ich stosunkowo łatwa dostępność powinna skłaniać do stosowania regresji nieliniowej w miejsce prostszej regresji nieliniowej linearyzowanej.

Działania na liczbach przybliżonych

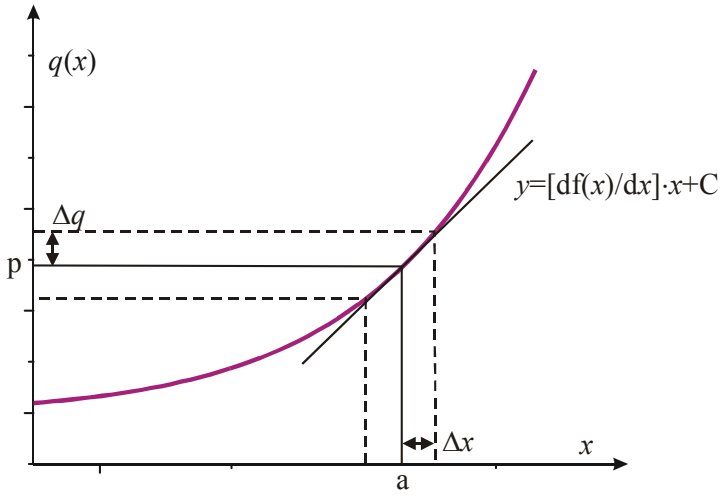
Wartości uzyskane w wyniku doświadczeń, obarczone określonymi niepewnościami pomiarowymi (błędami pomiarowymi), podlegają bardzo często dalszym przekształceniom prowadzącym do oznaczenia innych wielkości fizycznych (wyników pomiarów pośrednich, czyli takich na które składają się liczne pomiary bezpośrednie). W działaniach na liczbach przybliżonych konieczne jest stosowanie odpowiednich reguł służących określaniu dokładności wyników pomiarów pośrednich i ich zaokrąglaniu.

Liczba cyfr w rozwinięciu dziesiętnym liczby, wyniku pomiaru bezpośredniego, ograniczona jest dokładnością pomiaru (klasą używanego przyrządu) lub niepewnością pomiarową. Zawarte w takim wyniku cyfry możemy podzielić na ***cyfry znaczące***, czyli określające dokładność oznaczenia i zera służące do wyznaczenia pozycji dziesiętnych cyfr znaczących. Cyframi znaczącymi są więc wszystkie cyfry różne od zera, zera zawarte pomiędzy tymi cyframi oraz te zera na końcu liczby, których znaczenie wynika z dokładności pomiaru, np. (cyfry znaczące zaznaczone pogrubieniem):

$$0{,}0\mathbf{234} \pm 0{,}000\mathbf{2};\ \mathbf{120{,}50} \pm 0{,}0\mathbf{1};\ \mathbf{5607}00 \pm \mathbf{3}00;\ \mathbf{789} \pm \mathbf{40}$$

Prawidłowe przedstawienie wyniku pomiaru pośredniego wymaga znajomości niepewności pomiarowych wyników składających się na wynik ostateczny i reguły, wg której błędy wyników składowych przenoszą się na ostateczny wynik pomiaru. Wyobraźmy sobie

dowolne przekształcenie (funkcję) łączącą wynik pomiaru pośredniego, q, z wynikiem pomiaru bezpośredniego, x, obarczonego niepewnością pomiarową Δx. Na rysunku 13 wyraźnie widoczne jest „przeniesienie" błędu pomiarowego x na niepewność oznaczenia wielkości q.



Rys. 13 Przenoszenie błędów w zależnościach funkcyjnych

$$q = q(x) = f(x)$$

$$\Delta q = q(x + \Delta x) - q(x) = f(x + \Delta x) - f(x)$$

Ponieważ dla dostatecznie małego przedziału u wokół dowolnej wartości zmiennej x, np. $x = a$ $f(x+u)-f(x) = [df/dx]_a \cdot u$, stąd:

$\Delta q = q(x + \Delta x) - q(x) = f(x + \Delta x) - f(x) = [df/dx]_a \cdot \Delta x$

W celu prezentacji niepewności pomiarowej jako liczby dodatniej Δq, (odchylenie wyniku poniżej i powyżej określonej wartości symbolizuje znak ±) wzór powyższy powinien być zaprezentowany jako:

$$\Delta q = \left| dq/dx \right|_a \cdot \Delta x$$

W przypadku, gdy wielkość q jest funkcją wielu zmiennych otrzymamy ogólną postać ***rachunku błędu maksymalnego***, czyli przenoszenia niepewności pomiarowych w przekształceniach matematycznych:

$$\Delta q = \left| \partial q/\partial x \right|_a \cdot \Delta x + \left| \partial q/\partial y \right|_b \cdot \Delta y + \ldots + \left| \partial q/\partial z \right|_w \cdot \Delta z$$

Z powyższej zależności w prosty sposób można wyprowadzić ogólne reguły przenoszenia błędu w prostych działaniach arytmetycznych np.

- sumy (analogiczny wzór dla różnicy):

$q = x + y$

$\Delta q = |\partial q/\partial x| \cdot \Delta x + |\partial q/\partial y| \cdot \Delta y = \Delta q = |1| \cdot \Delta x + |1| \cdot \Delta y = \Delta x + \Delta y$

- iloczynu (analogiczny wzór dla ilorazu):

$q = x \cdot y$

$\Delta q = |\partial q/\partial x| \cdot \Delta x + |\partial q/\partial y| \cdot \Delta y = |y| \cdot \Delta x + |x| \cdot \Delta y \qquad |: |q| = |x \cdot y|$

$$\frac{\Delta q}{q} = \frac{|y| \cdot \Delta x}{|xy|} + \frac{|x| \cdot \Delta y}{|xy|} = \frac{\Delta x}{|x|} + \frac{\Delta y}{|y|} = \delta x + \delta y$$

Przykład liczbowy:

$x = 3{,}27 \pm 0{,}02 = 3{,}27(1 \pm 0{,}006)$

$y = 1{,}43 \pm 0{,}03 = 3{,}27(1 \pm 0{,}021)$

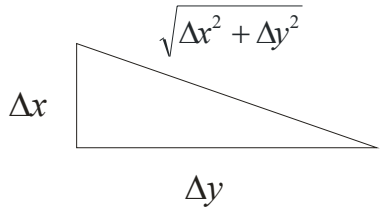
$q = x + y = 5{,}70 \pm 0{,}05$; $q = x - y = 1{,}84 \pm 0{,}05$

$q = x \cdot y = 4{,}676[1 \pm (0{,}006 + 0{,}021)] = 4{,}676(1 \pm 0{,}027) = 4{,}676 \pm 0{,}126$

$q = x / y = 2{,}287[1 \pm (0{,}006 + 0{,}021)] = 2{,}287(1 \pm 0{,}027) = 2{,}287 \pm 0{,}062$

W rachunku błędu maksymalnego, zaprezentowanym powyżej, nie jest brana pod uwagę wzajemna zależność lub niezależność rozpatrywanych wielkości i ich niepewności. W przypadku wielkości zależnych możemy mieć do czynienia z sytuacją, w której wielkości te odbiegają jednocześnie w tym samym kierunku od wielkości poprawnej, np. w górę i w rezultacie suma tych wielkości przyjmie wartość maksymalną. W przypadku wielkości niezależnych można spodziewać się, że zawyżenie jednych wielkości ($x + \Delta x$) może zostać częściowo zrekompensowane w wyniku zaniżenia innych ($y - \Delta y$). W efekcie niepewność wyniku powinna być mniejsza, niż przewidywana w rachunku błędu maksymalnego. I rzeczywiście, zakładając, że wielkości x i y pozostają niezależne i podlegają rozkładom normalnym, odpowiednio x: N(m_x, σ_x) oraz y: N(m_y, σ_y) to wielkość $q = x + y$ podlega rozkładowi normalnemu q: N(m_q, σ_q), gdzie $m_q = m_x + m_y$ i $\sigma_q = \sqrt{\sigma_x^2 + \sigma_y^2}$. Traktując σ_x i σ_y jako niepewności pomiarowe Δx i Δy możemy porównać błąd sumy uzyskany z rachunku błędu maksymalnego: $\Delta q = \Delta x + \Delta y$ i metod statystycznych: $\Delta q = \sqrt{\Delta x^2 + \Delta y^2}$.

Ponieważ:



stąd

$$\Delta x + \Delta y > \sqrt{\Delta x^2 + \Delta y^2}\ .$$

W rezultacie, zakładając niezależność niepewności pomiarowych pomiarów bezpośrednich, należy zmodyfikować wzór wynikający z rachunku błędu maksymalnego. Niepewność pomiarowa dana jest wzorem:

$$\Delta q = \sqrt{(\frac{\partial q}{\partial x} \cdot \Delta x)^2 + (\frac{\partial q}{\partial y} \cdot \Delta y)^2 + \ldots + (\frac{\partial q}{\partial z} \cdot \Delta z)^2}$$

Należy jednak pamiętać, że wszędzie tam, gdzie nie mamy pewności odnośnie niezależności zmiennych stosowanie rachunku błędu maksymalnego jest poprawniejszą metodą oznaczania błędu pomiaru pośredniego.

Przykład:

Obliczyć pojemność cieplną kalorymetru wg wzoru $K = i^2 \cdot R \cdot t/dT$ gdzie:

i - natężenie prądu = 1,75 ± 0,025A

R - opór spirali grzejnej = 45 ± 1 Ω

t - czas przepływu prądu = 600 ± 2 s

dT - przyrost temperatury kalorymetru = 20 ± 1°K

$$\Delta K = |\partial K/\partial i|_{1,75} \cdot \Delta i + |\partial K/\partial R|_{45} \cdot \Delta R + |\partial K/\partial \mathrm{t}|_{600} \cdot \Delta t + |\partial K/\partial (dT)|_{20} \cdot \Delta (\, dT)$$

$$\Delta K = |2iRt/dT|_{1,75} \cdot \Delta i + |i^2 t/dT\,|_{45} \cdot \Delta R + |\ i^2 R/dT\ |_{600} \cdot \Delta t +$$

$$+ |-i^2 Rt/(dT)^2\ |_{20} \cdot \Delta (\, dT)$$

$$K = 4134{,}4 \pm 430{,}5$$

przy czym błąd po uwzględnieniu metod statystycznych i niezależności zmiennych wyniósłby ± 255,6.

Podsumowując:

Rachunek błędu maksymalnego

Niepewność wartości funkcji jednej zmiennej	
$\Delta q = \lvert dq/dx \rvert \cdot \Delta x$	
Niepewność wartości funkcji wielu zmiennych q(x, y,...z)	
$\Delta q = \lvert \partial q/\partial x \rvert \cdot \Delta x + \lvert \partial q/\partial y \rvert \cdot \Delta y + ... + \lvert \partial q/\partial z \rvert \cdot \Delta z$	
Niepewność sumy i różnicy	
$q = x + y$	$\Delta q = \Delta x + \Delta y$
$q = x - y$	$\Delta q = \Delta x + \Delta y$
Niepewność iloczynu i ilorazu	
$q = x \cdot y$	$\delta q = \delta x + \delta y$
$q = x/y$	$\delta q = \delta x + \delta y$

Metody statystyczne

Niepewność wartości funkcji wielu zmiennych q(x, y,...z)
$\Delta q = \sqrt{(\frac{\partial q}{\partial x} \cdot \Delta x)^2 + (\frac{\partial q}{\partial y} \cdot \Delta y)^2 + \ldots + (\frac{\partial q}{\partial z} \cdot \Delta z)^2}$

Wybrane pozycje literaturowe:

1. J. B. Czermiński, A. Iwasiewicz, Z. Paszek, A. Sikorski, *Metody statystyczne dla chemików*, PWN, Warszawa 1992.
2. J. Greń, *Statystyka matematyczna*, PWN, Warszawa 1987.
3. J. Greń, *Statystyka matematyczna. Modele i zadania,* PWN, Warszawa 1978.
4. J. R. Taylor, *Wstęp do analizy błędu pomiarowego,* PWN, Warszawa 1995.
5. W. Klonecki, *Statystyka dla inżynierów,* PWN, Warszawa 1995.
6. W. Ufnalski, K. Mądry, *Excel dla chemików i nie tylko*, WNT, Warszawa 2000.
7. W. Kaczmarek, M. Kotłowska, A. Kozak, J. Kudyńska, H. Szydłowski, *Teoria pomiarów*, PWN, Warszawa 1981.
8. S. Brandt, *Metody statystyczne i obliczeniowe analizy danych,* PWN, Warszawa 1976.
9. M. A. White, *Quantity Calculus: Unambiguous Destignation of Units in Graphs and Tables*, J. Chem. Edu. 75, 607-9 (1998).

Podpisy pod rysunkami