STATYSTYCZNA OCENA WYNIKÓW ANALIZ
TEST t-STUDENTA

Wartości mierzonej funkcji (Y), ilości substancji oznaczanej (x) lub oznaczanego stężenia (c) tej substancji przedstawiane są zwykle w postaci:

a) tabelarycznej
b) wykresów graficznych
c) równań funkcji zależności pomiędzy (Y) i (c): Y = ac lub Y = ac + b

W celu obliczenia wyniku końcowego, należy opisać:

- wynik końcowy analizy w postaci średniej arytmetycznej z obliczonym przedziałem ufności i określeniem prawdopodobieństwa, z jakim ten przedział ufności został podany
- walidację metody

Parametry statystyczne opisujące wyniki

Rozkład normalny błędów pomiarowych

Wartości mierzone – wyniki analiz – podlegają nieuniknionym odchyleniom przypadkowym. Prawdopodobieństwo występowania wyników mniejszych i większych od wartości rzeczywistej (μ), jest jednakowe i stanowi rozkład odchyleń opisany ***funkcją Gaussa, której obrazem jest krzywa Gaussa (krzywa dzwonowa) przedstawiająca rozkład normalny błędów przypadkowych.*** Jej cechą jest to, że:

- największą liczebność mają wyniki, gdy wartość mierzona (x_i) ≈ rzeczywistej (μ)
- im bardziej x_i różni się od μ, tym mniejsza jest liczebność wyników
- rozrzut wyników jest symetryczny, tj. występuje jednakowa liczba odchyleń dodatnich i ujemnych
- krzywa ma maksimum przy x_i = μ, dwa punkty przegięcia przy x_i = μ ± σ (odchylenie standardowe) i po dwóch stronach zbliża się asymptotycznie do osi odciętych
- trójkąt opisany na krzywej ma podstawę o wierzchołkach w punktach -2σ i +2σ

Odchylenie standardowe

Odchylenie standardowe (σ) w rozkładzie Gaussa:

- przedstawia błąd bezwzględny dla wyniku x_i, któremu odpowiada punkt przegięcia na krzywej
- jest miarą odtwarzalności
- mała wartość tego parametru wskazuje na dużą precyzję

Miara ta dotyczy oznaczeń, w których liczba powtórzeń jest duża (n > 30). Gdy oznaczeń jest mniej niż 30 odchylenie standardowe szacuje się na podstawie parametru s:

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

Przedział ufności

Wartość rzeczywista może być ***oszacowana*** na podstawie średniej arytmetycznej z pomiarów. W tym celu określa się przedział, w którym znajduje się wartość rzeczywista z założonym z góry prawdopodobieństwem. Przedział ten nazwano przedziałem ufności (L). Prawdopodobieństwo założone (w %), że wartość rzeczywista znajduje się w przedziale ufności określa się mianem poziomu ufności (p). W obliczeniach najczęściej przyjmuje się wartość poziomu ufności 0,95 i 0,99. Oznacza to, że na 100 wyników odpowiednio 95 i 99 znajduje się w przedziale ufności.

Przedział ufności wyznacza się:

- przy dużej liczbie pomiarów (n>20) z rozkładu normalnego według wzorów:

$$L = \overline{x} \pm 1{,}96\sqrt{n} \quad dla\, p = 0{,}95$$
$$L = \overline{x} \pm 2{,}58\sqrt{n} \quad dla\, p = 0{,}99$$

- przy małej liczbie pomiarów (n<20) z rozkładu Studenta (dotyczy statystycznej oceny małych grup wyników), według wzoru:

$$L = \overline{x} \pm t\overline{s}$$

t – współczynnik z tablicy rozkładu Studenta – zależny od poziomu ufności i liczby pomiarów; - $\overline{s}$ jest odchyleniem standardowym średniej $\overline{s} = \frac{s}{\sqrt{n}}$

NIEZBĘDNE WZORY

Bezwzględną wartość różnicy pomiędzy wynikiem pomiaru ***x,*** a uznaną za prawdziwą wartością wielkości mierzonej -***μ***, nazywa się ***błędem bezwzględnym***.

$$\Delta x = |x - \mu|$$

Stosunek błędu bezwzględnego do wartości prawdziwej nazywany **jest błędem względnym**, który jest również często wyrażany w %

$$\delta x = \frac{\Delta x}{\mu} \cdot 100\%$$

Błąd względny jest kryterium oceny dokładności wyników. Z uwagi, że jest on wielkością bezwymiarową lub wyrażoną w procentach, może służyć do porównania wyników wyrażonych w różnych jednostkach. Bardzo często nieznana jest wartość przyjmowana za dokładną (μ), stąd zastępuje się ją wartością średniej arytmetycznej wartości mierzonych.

Błąd bezwzględny i względny przedstawiają wtedy następujące wzory:

$\Delta x = |\bar{x} - x_i|$ oraz $\delta x = \frac{\Delta x}{\bar{x}} \cdot 100\%$

Średnia lub **średnia arytmetyczna** serii, liczbowo równa wyrażeniu

$$\bar{x} = \frac{\sum x_i}{n}$$

gdzie
n – jest liczbą przeprowadzonych pomiarów
Średnia arytmetyczna jest najczęściej stosowanym przybliżeniem wartości oznaczanej, gdyż jest najbardziej prawdopodobne, że odchylenie tej wielkości od wartości prawdziwej jest najmniejsze, czyli wartość oznaczana jest najbliższa wartości prawdziwej

Odchylenie pojedynczego pomiaru, oznaczane symbolem *e*, jest to bezwzględna wartość różnicy między zmienną x_i średnią ẍ

$$e = |x_i - \overline{x}|$$

Medianą nazywa się wynik środkowy w przypadku nieparzystej liczby wyników otrzymany po uszeregowaniu wyników według wzrastającej wielkości, albo średnią arytmetyczną z dwóch środkowych wyników w przypadku parzystej liczby wyników
Do oszacowania stopnia rozproszenia wyników stosuje się **rozstęp** lub **rozrzut** wyników (symbol Δ). Jest to różnica między największym i najmniejszym wynikiem skrajnym

$$\Delta = x_{max} - x_{min}$$

Wariancja (V lub s^2) z matematycznego punktu widzenia jest najodpowiedniejszą miarą rozproszenia wyników. Jest wielkością addytywną, umożliwia więc oszacowanie wpływu poszczególnych czynników czy czynności na rozrzut wyników. Jest też najczęściej stosowaną miarą precyzji przy porównywaniu różnych metod analitycznych.
Wariancja, czyli kwadrat odchylenia standardowego, jest sumą kwadratów odchyleń poszczególnych wyników od wartości średniej podzielona przez ilość pomiarów minus 1

$$V = s^2 = \frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n-1} = \frac{\sum_{i=1}^{n} e^2}{n-1}$$

Odchylenie standardowe pojedynczego pomiaru *s*, najczęściej używana miara rozrzutu wyników, jest pierwiastkiem kwadratowym z sumy kwadratów odchyleń między zmiennymi a średnią serii, podzielonej przez liczbę zmiennych w serii minus 1

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

Dodatkową zaletą jest fakt, że odchylenie standardowe ma wymiar wielkości mierzonej, a więc np. g, %, ml itd. Można więc je odnieść do wartości średniej i otrzymać **względne odchylenie standardowe** (***s_r***). Gdy wyrazimy je w procentach, pozwoli na policzenie, jaki procent zmierzonej wartości stanowi odchylenie standardowe

$$s_r = \frac{s}{\overline{x}} \cdot 100\%$$

Pojęcie to ułatwia porównanie precyzji wyników otrzymanych tą samą metodą przy różnej zawartości oznaczanego składnika

Średnie odchylenie standardowe (= odchylenie standardowe średniej arytmetycznej) jest ilorazem odchylenia standardowego i pierwiastka z liczby pomiarów. Im więcej wykona się pomiarów, tym wartość średniego odchylenia standardowego powinna być mniejsze

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

Estymacja przedziałowa

Oczywiście nigdy nie mamy pewności, czy wartości średniej arytmetycznej lub mediany odpowiadają wartości rzeczywistej $\boldsymbol{\mu}$. W praktyce po wykonaniu kilku lub kilkunastu pomiarów jest mało prawdopodobne, aby średnia arytmetyczna serii pomiarowej była dokładnie równa $\boldsymbol{\mu}$. Znacznie bezpieczniej jest podać wynik końcowy analizy w postaci przedziału wartości, w którym najprawdopodobniej znajduje się wartość rzeczywista. Taki odpowiednio skonstruowany zakres nazywa się ***przedziałem ufności***, ponieważ wartość rzeczywista mierzonej wielkości zawiera się w tym przedziale z pewnym ściśle określonym prawdopodobieństwem (ufnością). Metody statystyczne pozwalające konstruować taki przedział nazywają ***estymacją przedziałową***.

Zrozumiałe jest, że im więcej zostanie wykonanych pomiarów, tym jest bardziej prawdopodobne, że μ jest zbliżone do średniej arytmetycznej tej serii pomiarowej. Można w związku z tym stwierdzić, że im więcej wykonano pomiarów, tym przedział ufności może być mniejszy (w granicznym przypadku, gdy n dąży do nieskończoności, μ jest dokładnie równe). Z drugiej jednak strony im większą chcemy mieć pewność, że wartość rzeczywista znajdzie się w naszym przedziale, tym powinien on być szerszy. Żeby mieć całkowitą (100%) pewność, że w danym przedziale ufności znajduje się rzeczywista wartość, teoretycznie powinien rozciągnąć się on w zakresie (-∞; +∞). Najczęściej prawdopodobieństwo, z jakim wyznacza się przedziały ufności, wynosi 95, 99 lub 99,7%.

Prawdopodobieństwo równe 95% oznacza, że gdy wykonuje się 100 analiz daną metodą i określa przedziały ufności z takim prawdopodobieństwem, to 95 wyników powinno być prawidłowych, a tylko w 5 przypadkach wartość rzeczywista nie znalazłaby się w tych przedziałach.

Innymi słowy jeden na dwadzieścia wyników będzie określony nieprawidłowo. Wartości prawdopodobieństwa równe 99 i 99,7% oznaczają, że odpowiednio jeden na 100 i trzy na 1000 wyznaczonych z takim prawdopodobieństwem przedziałów ufności nie pokryłoby się z wartością rzeczywistą.

Podstawowym czynnikiem mającym wpływ na szerokość przedziału ufności jest rozproszenie wyników kolejnych pomiarów. Spośród wymienionych wyżej miar tego rozproszenia do obliczania przedziałów ufności najbardziej odpowiednie jest odchylenie standardowe średnie.

Jeżeli liczba pomiarów n nie przekracza 30 (co w praktyce analitycznej prawie zawsze ma miejsce), do konstruowania przedziałów ufności stosuje się dodatkowo specjalne współczynniki zwane ***współczynnikami t-Studenta***. W takim przypadku przedział ten rozciąga się po obu stronach wartości średniej o wartość odchylenia standardowego średniego pomnożonego przez odpowiedni współczynnik

$$t\colon\ (\bar{x} - t\bar{s};\ \bar{x} + t\bar{s}\ ) = (\bar{x} \pm t\bar{s})$$

Zgodnie z powyższymi rozważaniami można zauważyć, że współczynniki te wzrastają ze wzrostem prawdopodobieństwa i maleją ze wzrostem liczby pomiarów. Przy czym, począwszy od n = 6 – 7 zmiany współczynników t są już bardzo nieznaczne. Nasuwa się stąd wniosek, że wykonywanie więcej niż 5 analiz danej próbki jest bezcelowe, ponieważ nieznacznie wpływa na poprawienie precyzji pomiaru (zmniejszenie przedziału ufności), powoduje zaś odczuwalne zwiększenie kosztów analizy.

Liczba pomiarów (n)	***Wartość prawdopodobieństwa***		
	95%	***99%***	***99,7%***
2	12,71	63,64	212,21
3	4,30	9,92	18,22
4	3,18	5,84	8,89
5	2,78	4,60	6,43
6	2,57	4,03	5,38
7	2,45	3,71	4,80
8	2,36	3,50	4,44
9	2,31	3,36	4,20
10	2,26	3,25	4,02
15	2,14	2,98	3,58
20	2,09	2,86	3,40
25	2,06	2,80	3,30
30	2,05	2,76	3,24

Tabela Najczęściej stosowane wartości współczynników t-Studenta

Przykład obróbki statystycznej...

W celu objaśnienia i przedstawienia praktycznej obróbki statystycznej rozważamy następujący przykład: wykonano siedmiokrotnie ważenie na wadze analitycznej pewnego przedmiotu, który przed każdym badaniem był suszony w 100 ^{o}C. Otrzymano masy w gramach: 0.1639, 0.1642, 0.1640, 0.1645, 0.1640, 0.1642, 0.1644

i	x_i	Odchylenie	Kwadrat odchylenia
1	0,1639	-0,000260	$6.75 \cdot 10^{-8}$
2	0,1640	-0,000190	$3{,}60 \cdot 10^{-8}$
3	0,1640	-0,000190	$3{,}60 \cdot 10^{-8}$
4	0,1642	-0,000020	$3{,}99 \cdot 10^{-10}$
5	0,1642	-0,000090	$8{,}09 \cdot 10^{-9}$
6	0,1644	-0,000230	$5{,}28 \cdot 10^{-8}$
7	0,1645	-0,000300	$8{,}99 \cdot 10^{-8}$
	1,1492	0,000000	$2{,}91 \cdot 10^{-7}$

Na podstawie powyższej tabeli można wyliczyć następujące wartości:

Średnia arytmetyczna	0,1642
Mediana	0,1642
Rozrzut	0,000559
Wariancja	$4{,}85 \cdot 10^{-8}$
Odchylenie standardowe	0,000220 g
Względne odchylenie standardowe	0.00134 = 0,134%
Odchylenie standardowe średnie	0,0000832 g

Można już teraz obliczyć przedziały ufności z określonym prawdopodobieństwem

$$t\colon\ (\bar{x} - t\bar{s};\ \bar{x} + t\bar{s}\ ) = (\bar{x} \pm t\bar{s})$$

95% → 0,1642 ± 2,45 ·0,0000832 =0,1642 ± 0,0002g

99% → 0,1642 ± 3,71 ·0,0000832 =0,1642 ± 0,0003g

99,7% → 0,1642 ± 4,80 ·0,0000832 =0,1642 ± 0,0004g

Z największym prawdopodobieństwem można więc powiedzieć, że masa ważonego przedmiotu mieści się w przedziale (0,1638 – 0,1646 g).

W praktyce analitycznej najczęściej wystarczy podać wynik oznaczenia z prawdopodobieństwem 95% - uzyskuje się dzięki temu dość wąskie przedziały ufności, a ryzyko popełnienia błędu jest i tak już niewielkie. Należy jednak jeszcze raz podkreślić, że rzeczywista zawartość żelaza może być zupełnie inna, jeśli w trakcie analizy pojawiły się duże błędy o charakterze statystycznym.