POLITECHNIKA GDAŃSKA

STATYSTYKA MATEMATYCZNA – narzędzie do opracowywania i interpretacji wyników pomiarów

Piotr Konieczka

Katedra Chemii Analitycznej
Wydział Chemiczny
Politechnika Gdańska

Statystyka matematyczna - część matematyki oparta na wykorzystaniu rachunku prawdopodobieństwa oraz ukierunkowana na badanie prawidłowości pojawiania się określonych cech w obiektach materialnych lub zjawiskach, występujących masowo, tzn. mogących pojawiać się dowolną ilość razy.

Statystyka przedstawia te prawidłowości za pomocą liczb.

Statystyka pozwala znaleźć odpowiedź na wiele pytań np.:

- Jak dokładny jest wynik oznaczenia?
- Jak wiele oznaczeń powinno być przeprowadzonych aby zwiększyć precyzję pomiaru?
- Czy badany produkt spełnia stawiane mu wymogi, normy?

Statystyka to narzędzie, którego używanie musi być prowadzone w sposób rozsądny i zrozumiały.

Zastosowanie określonej metody analitycznej sprawia, że w sposób jednoznaczny wyznaczony jest rozkład wyników pomiarów (*cechy*), które można traktować jako niezależne zmienne losowe. Wynik jest konsekwencją przeprowadzenia pomiaru. Zbiór otrzymanych wyników oznaczeń tworzy rozkład (*empiryczny*).

Analityka ⟷ Statystyka

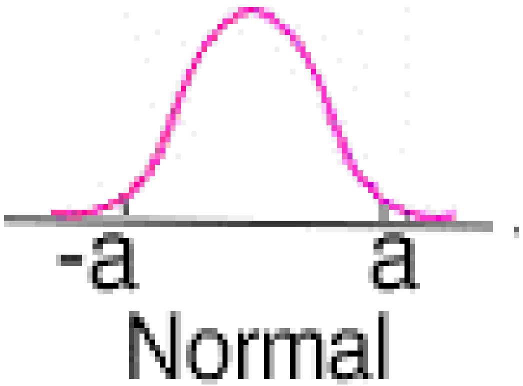
Rozkład normalny

niezwykle ważny rozkład prawdopodobieństwa w wielu dziedzinach (rozkład Gaussa)

definiowany dwoma parametrami: średnią (odpowiada za *położenie* rozkładu) i odchyleniem standardowym (*skala*).

Własności rozkładu normalnego $N(\mu_x, s)$:

- wartość oczekiwana: μ_x;
- mediana: μ_x;
- wariancja: s^2;



Rozkład jednostajny (prostokątny)

ciągły rozkład prawdopodobieństwa, dla którego gęstość prawdopodobieństwa w przedziale $\langle -a, +a \rangle$ jest stała i różna od zera, a poza nim równa zeru.

Ponieważ rozkład jest ciągły, nie ma większego znaczenia czy punkty -*a* i *+a* włączy się do przedziału czy nie. Rozkład jest określony parą parametrów -*a* i *+a.*

Własności rozkładu jednostajnego:

- wartość oczekiwana: *0*
- mediana: *0*
- wariancja: $a^2/3$

Rozkład trójkątny

Własności rozkładu trójkątnego w przedziale $\langle -a, +a \rangle$:

- wartość oczekiwana: *0*;
- mediana: *0*;
- wariancja: $a^2/6$

Znajomość rozkładu zmiennej losowej - pełna informacja na temat badanej cechy (może to być np.: stężenie, zawartość, właściwość fizykochemiczna).

Rzadko istnieje możliwość dysponowania taką pełną informacją.

Wnioskowanie na temat cechy oparte o analizę pewnej ograniczonej liczby elementów (*próbek*) reprezentujących fragment całego zbioru opisywanego rozkładem.

Wtedy należy wnioskować o badanej cesze na podstawie oszacowania niektórych jej parametrów (*parametry statystyczne*) lub na podstawie rozkładu empirycznego.

Parametry statystyczne to wielkości liczbowe służące do opisu struktury zbiorowości statystycznej w sposób systematyczny.

Wśród tych parametrów wyróżnić można cztery podstawowe grupy:

- miary położenia;
- miary rozproszenia;
- miary asymetrii;
- miary skupienia;

Wybrane właściwości średniej arytmetycznej:

- suma wartości cechy jest równa iloczynowi średniej arytmetycznej i liczebności zbiorowości;
- średnia arytmetyczna spełnia warunek:

$$x_{\min} < x_{śr} < x_{\max}$$

- (suma odchyleń poszczególnych wartości cechy od średniej równa się zero:

$$\sum_{i=1}^{n}(x_i - x_{śr}) = 0$$

- suma kwadratów odchyleń poszczególnych wartości cechy od średniej jest minimalna:

$$\sum_{i=1}^{n}\left(x_i - x_{śr}\right)^2 = \min$$

- średnia arytmetyczna jest wrażliwa na skrajne wartości cechy,
- średnia arytmetyczna z próby jest dobrym przybliżeniem (oszacowaniem, estymatorem) wartości oczekiwanej.

Mediana

Mediana (wartość środkowa) *Me* – środkowa liczba w uporządkowanej niemalejąco próbce (dla próbki o liczności nieparzystej) lub średnia arytmetyczna dwóch liczb środkowych (dla próbki o liczności parzystej).

Mediana dzieli zbiorowość na dwie równe części; połowa jednostek ma wartości cechy mniejsze lub równe medianie, a połowa wartości cechy równe lub większe od *Me*; stąd nazwa wartość środkowa.

Mediana jest odporna na wartości skrajne (odbiegające)

Miary rozproszenia (zmienności, dyspersji)

Z reguły miary rozproszenia odnoszą się do określania różnic pomiędzy obserwacjami a wartością średnią.

Do najczęściej stosowanych miar rozproszenia należą:

- rozstęp
- wariancja
- odchylenie standardowe
- odchylenie przeciętne (średnie)
- współczynnik zmienności

Rozstęp

Rozstęp to różnica pomiędzy wartością maksymalną, a minimalną cechy - jest miarą charakteryzującą empiryczny obszar zmienności badanej cechy, nie daje on jednak informacji o zróżnicowaniu poszczególnych wartości cechy w zbiorowości.

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

Wariancja

Wariancja jest to średnia arytmetyczna kwadratów odchyleń poszczególnych wartości cechy od średniej arytmetycznej zbiorowości.

$$s^2 = \frac{1}{n-1}\sum_{i-1}^{n}\left(x_i - x_{śr}\right)^2$$

Odchylenie standardowe

Definiowane jako miara rozproszenia uzyskanych poszczególnych wartości oznaczeń wokół wartości średniej:

$$s = \sqrt{\frac{\sum_{i=1}^{n}\left(x_i - x_{śr}\right)^2}{n-1}}$$

gdzie:

x_i – wartość pojedynczego wyniku oznaczenia;
$x_{śr}$ – średnia arytmetyczna z uzyskanych wyników;
n – liczba uzyskanych wyników;

Odchylenie standardowe jest równe zeru wtedy i tylko wtedy, gdy wszystkie wyniki są identyczne. W każdym innym przypadku wielkość ta jest dodatnia. Zatem im większe rozproszenie wyników, tym wartość *s* jest większa.

W tym miejscu należy zwrócić uwagę na jeden podstawowy fakt. Rozrzut wyników związany jest z każdym postępowaniem analitycznym. Możliwe jest natomiast, że zjawiska tego nie udało się zaobserwować ze względu na np. zbyt niską rozdzielczość stosowanego przyrządu kontrolno-pomiarowego.

Serie wyników pomiarów uzyskane z wykorzystaniem przyrządów kontrolno-pomiarowych o różnej rozdzielczości.

	Przyrząd 1	**Przyrząd 2**	**Przyrząd 3**
Uzyskane wyniki	17	16,8	16,83
	17	17,1	17,14
	17	16,9	16,88
	17	17,4	17,43
	17	17,3	17,27
	17	17,2	17,24
	17	17,0	16,96
Wartość odchylenia standardowego	**0**	**0,22**	**0,223**

Właściwości odchylenia standardowego

- jeżeli do każdej wartości wyniku pomiaru dodamy (lub od niej odejmiemy) stałą wartość to wartość odchylenia standardowego nie zmieni się,
- jeżeli każdą wartość wyniku pomiaru pomnożymy lub podzielimy przez dowolną stałą to wartość odchylenia standardowego zostanie także pomnożona lub podzielona przez tę stałą,
- odchylenie standardowe jest zawsze liczbą mianowaną, przy czym miano jego jest wyrażone w takich samych jednostkach jak miano wartości wyników w próbce,

Odchylenie standardowe:

a. dla znanej wartości rzeczywistej μ_x

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

b. dla nieznanej wartości rzeczywistej (oszacowanie $x_{śr}$)

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

c. względne odchylenie standardowe

$$s_R\ (RSD) = \frac{s}{x_{śr}}$$

d. odchylenie standardowe średniej arytmetycznej

$$\bar{s} = \frac{s}{\sqrt{n}}$$

e. odchylenie standardowe metody (ogólne)

$$s_g = \sqrt{\frac{1}{n-k}\sum_{i=1}^{k} s_i^2 (n_i - 1)}$$

gdzie:

n - ogólna liczba oznaczeń

k - liczba serii

dla równolicznych serii wzór upraszcza się do postaci:

$$s_g = \sqrt{\frac{1}{k} \cdot \sum_{i=1}^{k} s_i^2}$$

Współczynnik zmienności

Współczynnik zmienności (*CV*) powstaje przez pomnożenie wartości *RSD* przez 100%:

$$CV = RSD \cdot 100\%$$

Współczynnik zmienności - jest ilorazem bezwzględnej miary zmienności cechy i średniej wartości tej cechy, jest wielkością niemianowaną, najczęściej podawaną w procentach.

Współczynnik zmienności stosuje się w porównaniach zróżnicowania:

- kilku zbiorowości pod względem tej samej cechy,
- tej samej zbiorowości pod względem kilku różnych cech.

Określenie **selektywności**
(analiza próbek roztworów wzorcowych)

Wyznaczenie **liniowości**,
LOD, **LOQ**, **zakresu pomiarowego**

Wyznaczenie **powtarzalności**

Wyznaczenie **precyzji pośredniej**

Określenie **odporności**
(na podstawie powtarzalności i precyzji pośredniej)

Określenie **selektywności**
(analiza próbek rzeczywistych)

Wyznaczenie **poprawności**
(analiza próbek materiałów odniesienia)

Określenie **dokładności**
(na podstawie precyzji i poprawności)

Wyznaczenie **odtwarzalności**
(np. porównania międzylaboratoryjne)

Określenie **elastyczności**
(na podstawie odtwarzalności i precyzji pośredniej)

RAPORT WALIDACYJNY

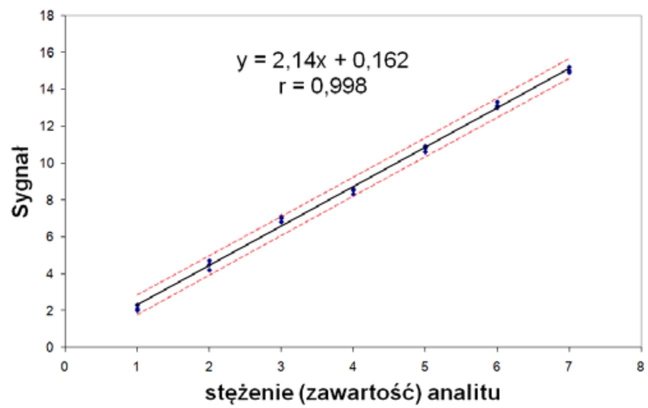
$$s = \sqrt{\frac{\sum_{i=1}^{n}\left(x_i - x_{śr}\right)^2}{n-1}}$$

$$s_R\,(RSD) = \frac{s}{x_{śr}}$$

$$CV$$

$$t_a = \frac{a - a_o}{s_a}$$

$$t_b = \frac{b - b_o}{s_b}$$



Określenie **selektywności**
(analiza próbek roztworów wzorcowych)

Wyznaczenie **liniowości**,
LOD, **LOQ**, **zakresu pomiarowego**

Wyznaczenie **powtarzalności**

Wyznaczenie **precyzji pośredniej**

Określenie **odporności**
(na podstawie powtarzalności i precyzji pośredniej)

Określenie **selektywności**
(analiza próbek rzeczywistych)

Wyznaczenie **poprawności**
(analiza próbek materiałów odniesienia)

Określenie **dokładności**
(na podstawie precyzji i poprawności)

Wyznaczenie **odtwarzalności**
(np. porównania międzylaboratoryjne)

Określenie **elastyczności**
(na podstawie odtwarzalności i precyzji pośredniej)

RAPORT WALIDACYJNY

Określenie **selektywności**
(analiza próbek roztworów wzorcowych)

Wyznaczenie **liniowości**,
LOD, **LOQ**, **zakresu pomiarowego**

Wyznaczenie **powtarzalności**

Wyznaczenie **precyzji pośredniej**

Określenie **odporności**
(na podstawie powtarzalności i precyzji pośredniej)

Określenie **selektywności**
(analiza próbek rzeczywistych)

Wyznaczenie **poprawności**
(analiza próbek materiałów odniesienia)

Określenie **dokładności**
(na podstawie precyzji i poprawności)

Wyznaczenie **odtwarzalności**
(np. porównania międzylaboratoryjne)

Określenie **elastyczności**
(na podstawie odtwarzalności i precyzji pośredniej)

RAPORT WALIDACYJNY

$$x_{śr} = \frac{\sum_{i=1}^{n} x_i}{n}$$

Weryfikacja hipotez statystycznych

Hipoteza to sąd o populacji oparty na prawdopodobieństwie, przyjęty w celu wyjaśnienia jakiegoś zjawiska, prawa lub faktu i wymagający sprawdzenia; przypuszczenie.

Weryfikacją hipotez nazywamy sprawdzanie sądów o populacji, sformułowanych bez zbadania jej całości.

Przebieg procedury weryfikacyjnej wygląda następująco:

Sformułowanie hipotezy zerowej i alternatywnej

Hipoteza zerowa – H_o - prosta postać hipotezy poddana testom;

Hipoteza alternatywna – H_1 - przeciwstawiona hipotezie zerowej;

Przebieg procedury weryfikacyjnej wygląda następująco:

Wybór odpowiedniego testu

Test służy do sprawdzania hipotezy.

Przebieg procedury weryfikacyjnej wygląda następująco:

Określenie poziomu istotności α

z reguły α=0,05 co odpowiada poziomowi prawdopodobieństwa 0,95, czyli prawdopodobieństwu 95%

Przebieg procedury weryfikacyjnej wygląda następująco:

Wyznaczenie obszaru krytycznego testu

Wielkość obszaru krytycznego wyznacza dowolnie mały poziom istotności α, natomiast jego położenie określane jest przez hipotezę alternatywną.

Przebieg procedury weryfikacyjnej wygląda następująco:

Obliczenie parametru testu na podstawie próby

Wyniki próby opracowuje się w odpowiedni sposób, zgodnie z procedurą wybranego testu i są one podstawą do obliczenia statystyki testowej.

Przebieg procedury weryfikacyjnej wygląda następująco:

Podjęcie decyzji

Wyznaczona na podstawie próby wartość statystyki porównywana jest z wartością krytyczną testu.

- Jeżeli wartość ta znajdzie się w obszarze krytycznym to hipotezę zerową należy odrzucić jako nieprawdziwą.
- Jeżeli natomiast wartość ta znajdzie się poza obszarem krytycznym oznacza to, że brak jest podstaw do odrzucenia hipotezy zerowej. Stąd wniosek, że hipoteza zerowa może być prawdziwa.

Rodzaje błędów popełnianych przy weryfikacji:

błędy I-go rodzaju - odrzucenie hipotezy zerowej H_o gdy jest ona prawdziwa;

błędy II-go rodzaju - *przyjęcie* H_o gdy jest ona fałszywa;

		hipoteza	
		prawdziwa	fałszywa
hipoteza	przyjęta	✓	błąd II-go rodzaju
	odrzucona	błąd I-go rodzaju	✓

Coraz częściej do weryfikowania hipotez statystycznych wykorzystywane są różnego rodzaju programy komputerowe (np. program *Statistica*). W takim przypadku sposób postępowania ogranicza się do wyliczenia, dla badanego zbioru danych, parametru p – po wybraniu odpowiedniego testu statystycznego. Tak obliczona wartość porównuje się następnie z przyjętą wartością poziomu istotności α. Jeżeli obliczona wartość p jest mniejsza od wartości α: $p < \alpha$ odrzuca się hipotezę zerową H_o. W przeciwnym przypadku hipotezy zerowej nie odrzuca się.

Testy parametryczne służą do weryfikacji hipotez parametrycznych, odnoszących się do parametrów rozkładu badanej cechy w populacji generalnej. Najczęściej za ich pomocą dokonuje się weryfikacji sądów o takich parametrach populacji jak średnia arytmetyczna i wariancja.

Testy nieparametryczne są stosowane do sprawdzania różnorodnych hipotez, dotyczących m.in. zgodności rozkładu cechy w populacji z określonym rozkładem teoretycznym, zgodności rozkładów w dwóch populacjach, a także losowości doboru próby.

- PRECYZJA (powtarzalność, precyzja pośrednia, odtwarzalność)
- ODPORNOŚĆ
- ELASTYCZNOŚĆ
- KALIBRACJA (liniowość, trwałość (stabilność) krzywej kalibracyjnej)

Ocena (porównanie) uzyskanej(ych) wartości odchylenia standardowego

1. Ocena na podstawie obliczonej wartości *RSD* (bądź *CV*)

2. Z zastosowaniem odpowiedniego testu statystycznego
   a. w celu sprawdzenia istotności różnicy między odchyleniem standardowym badanej populacji a wartością zadaną stosujemy test χ^2.
   b. w celu porównania precyzji dwóch niezależnych serii pomiarowych uzyskanych w trakcie analizy próbek o zawartości analitu na takim samym poziomie, stosujemy test *F-Snedecora*

c. do porównania precyzji dwóch zależnych serii pomiarowych, stosujemy test *Morgana*

d. do porównywania precyzji dla równolicznych populacji (ilość wyników uzyskanych porównywanymi procedurami) stosujemy test F_{max} *Hartleya*

e. w celu porównywania precyzji (kilka metod, serie nie koniecznie równoliczne) - test *Bartletta*

Dokładność i miary niedokładności

1. dokładność wyniku pojedynczego oznaczenia (DOKŁADNOŚĆ):

$$d_{x_i} = x_i - \mu_x = \Delta x_{syst} + \Delta x_i + \delta x_i$$

2. dokładność wyniku analizy (POPRAWNOŚĆ):

$$d_{x_{śr}} = x_{śr} - \mu_x = \Delta x_{syst} + \Delta x_{śr}$$

3. dokładność procedury analitycznej:

$$d_{x_{met}} = E(x_{met}) - \mu_x = \Delta x_{syst}$$

Wynik odbiegający $\overset{?}{\neq}$ Wynik obarczony błędem grubym

- POPRAWNOŚĆ (odzysk)
- DOKŁADNOŚĆ (poprawność i precyzja)
- KALIBRACJA (liniowość, trwałość (stabilność) krzywej kalibracyjnej)

Porównanie dokładności dwóch metod (wartości średnich)

Jeżeli porównywane metody nie różnią się w sposób statystycznie istotny pod względem precyzji (stosujemy w tym celu test *F-Snedecora*) ich dokładność porównujemy stosując test *t-Studenta*.

Jeżeli porównywane metody różnią się w sposób statystycznie istotny pod względem precyzji (stosujemy w tym celu test *F-Snedecora*) ich dokładność (POPRAWNOŚĆ) porównujemy stosując przybliżony test *C-Cochrana i Coxa* - serie mało liczne lub test *Aspina i Welcha.*

„Istnieją trzy stopnie kłamstwa:
przepowiadanie pogody
statystyka
i komunikat dyplomatyczny."

Jean Rigaux

„Statystyka nie kłamie. Kłamią jedynie statystycy.”

Janusz Leon Wiśniewski

DZIĘKUJĘ ZA UWAGĘ