Tutorato di Chimica Analitica

2016/2017

Descrittori di dispersione

(Stima della variabilità in un gruppo di dati)

Varianza, deviazione standard

(deviazione standard relativa, coefficiente di variazione)

Intervallo di variazione

(Tolleranza, differenza massimo-minimo)

La **tolleranza** è definita dai limiti di massima variazione di una misura di grandezza (valore massimo – valore minimo), e serve a stabilire l'intervallo di fiducia associato a una misura.

Gli **scarti** sono le differenze tra ogni singolo valore misurato e la media

Varianza

È la media aritmetica delle differenze tra ogni singolo dato e la media, elevate al quadrato

Varianza

Varianza della popolazione

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

x_i = valore numerico di ciascuna misurazione
μ = media della popolazione
N = numero totale dei dati che compongono la popolazione

Varianza

Varianza del campione

$$s^2 = \frac{\sum_{i=1}^{N}\left(x_i - \bar{x}\right)^2}{N-1}$$

$\boldsymbol{x_i}$ = valore numerico di ciascun campione
$\bar{x}$ = media del campione
$\boldsymbol{N}$ -1 = gradi di libertà

Deviazione standard

La deviazione standard, o scarto tipo, è un indice di dispersione delle misure sperimentali.
È la radice quadrata della varianza

Estraendo la radice quadrata, compensiamo il fatto di avere elevato al quadrato i termini nel calcolare la varianza, cosicché s viene ad essere espressa nella stessa unità di misura delle osservazioni originali.

Deviazione standard della popolazione

- **Deviazione standard della popolazione**

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

x_i = valore numerico di ogni misurazione

μ = media della popolazione

N = numero dei dati che compongono la popolazione

Deviazione standard del campione

- **Deviazione standard del campione**

$$\boxed{s = \sqrt{s^2} = \sqrt{\frac{\sum_{i=1}^{N}\left(x_i - \bar{x}\right)^2}{N-1}}}$$

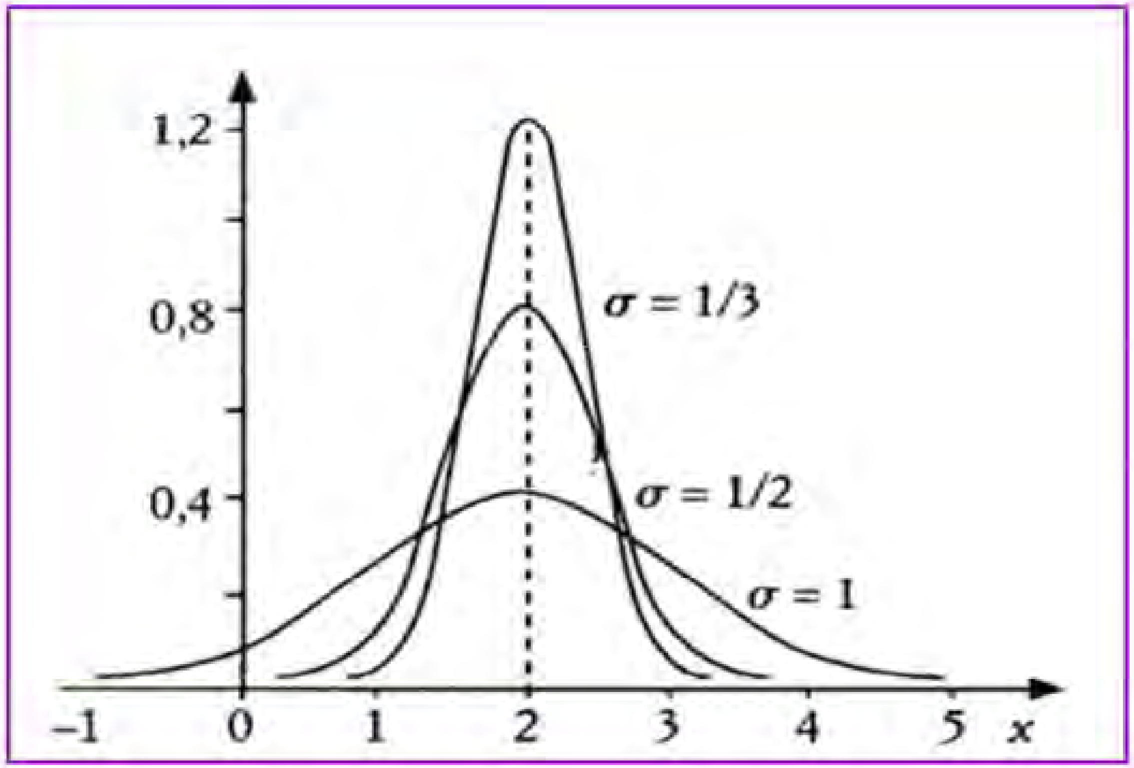
x_i = valore numerico di ciascun dato

$\bar{x}$ = media del campione

N-1 = numero dei gradi di libertà

Media (μ) e deviazione standard (σ)

La distribuzione normale è univocamente definita dalla media (μ) e dalla deviazione standard (σ).



Dato il valore della media $\mu = 2$ in questo caso, la curva a campana si allarga e si appiattisce al crescere della deviazione standard da 1/3 a 1.

Se σ è grande RISPETTO ALLA MEDIA i dati sono molto dispersi intorno alla media. All'opposto, se σ è un numero piccolo RISPETTO ALLA MEDIA, la curva a campana è più stretta attorno al valore della media.

ATTENZIONE!

Se i dati hanno distribuzione normale la popolazione è descritta in modo univoco con media e deviazione standard

Se i dati <u>non</u> hanno distribuzione normale media e deviazione standard <u>non</u> sono descrittori adatti per la popolazione

Deviazione standard relativa

È un parametro statistico adimensionale usato per descrivere la variabilità di una serie di dati e si calcola dividendo la deviazione standard per la media del campione

$$\mathrm{DSR} = \frac{\text{Deviazione standard}}{\text{Media}}$$

Deviazione standard relativa

Può essere rappresentata come percentuale (DSR% o RSD%), in tale caso si chiama anche *coefficiente di variazione*

$$\mathrm{RSD} \times 100 = RSD\% \equiv CV$$

Usando il CV per mettere in evidenza quale è il gruppo di misure in cui si ha la maggiore variabilità bisogna stare attenti a confrontare dati acquisiti nelle medesime unità di misura.

Praticamente il *coefficiente di variazione* dà una informazione più chiara sulla qualità dei dati rispetto alla deviazione standard (che misura la variazione assoluta dei dati)

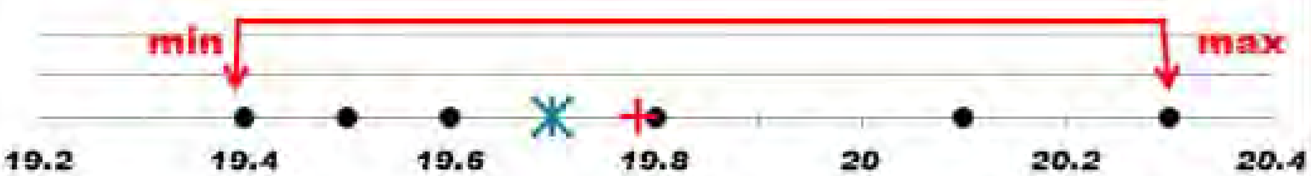
Esempio. **Supponiamo che un campione contenga 50 mg di Cu^{2+} e che due laboratori riportino la misura del rame in questo modo:**

Lab. 1: Cu^{2+}, la quantità è 50 ± 2 mg
Lab. 2: Cu^{2+}, la quantità è 50 mg con CV del 10%

Quale dei due laboratori fornisce la misura "migliore" in termini di precisione (minore dispersione del dato intorno alla media)?

Intervallo di variazione

INTERVALLO DI VARIAZIONE



I.V. = 20.3 - 19.4 = 0.9

Oppure

Min = 19.4

Max = 20.3

Mediana = 19.7

Media = 19.8

Moda, in questo caso non esiste

Intervallo di variazione

L'intervallo di variazione è il modo più semplice per dare l'informazione sulla dispersione dei dati: si tratta della differenza algebrica tra il valore più grande ed il valore più piccolo di una serie di misurazioni.

Media ponderata

Ogni dato non contribuisce in ugual misura al calcolo del valore della media.

w_i può sia rappresentare la frequenza sia l'importanza attribuita a ogni misurazione.

$$\frac{\sum_{i=1}^{N} w_i X_i}{N}$$

X_i= tutti i dati compresi da i=1 a i=N

N = numero di dati raccolti

w_i = frequenza di ciascun gruppo di dati

Esempio 1. In tabella sono riportati gli effetti di un analgesico su 20 volontari sottoposti ad uno stimolo doloroso.

Come rappresenteresti la media della stima del dolore segnalato dai tre gruppi di volontari?

Numero di volontari	Stima del dolore
2	3 (dolore estremo)
12	2 (dolore moderato)
6	1 (dolore lieve)

$$\frac{\sum_{i=1}^{N} w_i X_i}{N} = \frac{(2\times 3)+(12\times 2)+(6\times 1)}{20} = 1.8$$

DEVIAZIONE STANDARD RAGGRUPPATA

(«POOLED STANDARD DEVIATION»)

1 campione, molte analisi (30-50), ⇨ 1 media, 1 deviazione standard

n campioni, n gruppi di analisi per ciascun campione

⇩

n medie, n deviazioni standard

1 media ponderata e 1 deviazione standard raggruppata

$$s_{pooled} = \sqrt{\frac{\sum_{i=1}^{N_1}(x_i - \bar{x}_1)^2 + \sum_{j=1}^{N_2}(x_j - \bar{x}_2)^2 + \ldots + \sum_{z=1}^{N_n}(x_z - \bar{x}_n)^2 +}{(N_1 - 1) + (N_2 - 1) + \ldots + (N_n - 1)}}$$

x_i = risultati delle misure del gruppo i, i =1 a i =N_1, con media $\bar{x}_1$

x_j =risultati delle misure del gruppo i, i =1 a i =N_2, con media $\bar{x}_2$

Errori

Si usa il termine «errore» dando ad esso il significato statistico di differenza tra il valore vero e il valore osservato.

Si possono incontrare errori di tipo *casuale*, *sistematico* e *grossolano*

$$\text{Valore misura} = \text{Valore vero} + \text{Errore}_{\text{casuale}} + \text{Errore}_{\text{sistematico}} + \text{Errore}_{\text{grossolano}}$$

Errore casuale

Ha origine dalle limitazioni intrinseche degli strumenti di osservazione di cui disponiamo ed è inevitabile che compaia in ogni misura fisica o chimica.

Causa una dispersione dei dati, più o meno simmetrica, intorno al valore medio. L'effetto accumulato delle singole incertezze indeterminate provoca una fluttuazione casuale dei dati intorno al valore medio del risultato della serie

1- Per la messa a punto di un metodo analitico per la determinazione di alcol etilico nel sangue un chimico esegue la seguente serie di analisi, espresse in mg/100ml:

83,37
79,81
84,19
79,32
79,32
81,62
83,83
85,35
82,67
82,51

Calcolare media, deviazione standard e coefficiente di variazione.

Calcolare i limiti di confidenza al 95% del risultato, se questo è fornito come media di due analisi.

Tre analisti eseguono la determinazione della caffeina in uno stesso campione che ne contiene 1,21g/l; questi sono i loro risultati:

analista A:	1,382	1,323	1,360	1,187 g/l
analista B:	1056	1062	1061	1081 mg/l
analista C:	1,146	1,197	1,115	1,232 g/l

Indicare quale analista è il più accurato e quale il più preciso.

Un analista esegue 10 misure del contenuto di Hg in pesce da lago:

1,61 ppm
2,20 ppm
1,54 ppm
1,79 ppm
2,39 ppm
1,98 ppm
1,98 ppm
1,99 ppm
1,66 ppm
2,15 ppm

Quale è la precisione del metodo utilizzato?

Quante analisi dello stesso campione deve eseguire e mediare per avere un intervallo di confidenza (al 95%) inferiore a 0,8 ppm?

Per la messa a punto di un metodo analitico per la determinazione di metanolo nell'alcol etilico un chimico esegue la seguente serie di analisi:

3267 ppm
3124 ppm
3162 ppm
3146 ppm
3219 ppm
3149 ppm
3146 ppm
3231 ppm
3186 ppm
3241 ppm

Calcolare media, deviazione standard e coefficiente di variazione.

Calcolare i limiti di confidenza al (95%)del risultato, se questo è fornito come media di 3 analisi.

Per la messa a punto di un metodo analitico per la determinazione di alcol nel sangue un chimico esegue la seguente serie di analisi (espresse in μg/ml):

83,37
79,81
84,19
79,32
79,32
81,62
83,83
85,35
82,67
82,51
83,39
82,42
84,74
87,51
82,81
83,43
87,84
82,21
78,77
87,90

Calcolare la media, la deviazione standard e il coefficiente di variazione.

Inoltre calcolare i limiti di confidenza (95%) del risultato se questo è fornito come media di 2 analisi.

Tre analisti eseguono la determinazione dell'acido salicilico in uno stesso campione che ne contiene 850 ppm; i risultati ottenuti sono:

Analista n.1: 830 856 858 815 ppm

Analista n.2: 875 844 867 852 ppm

Analista n.3: 856 824 846 865 ppm

Indicare quale analista è il più accurato e quale è il più preciso.

Un analista esegue 9 misure del contenuto di Pb in un vino:
2,61 ppm
3,20 ppm
2,54 ppm
2,79 ppm
3,39 ppm
2,98 ppm
2,98 ppm
2,99 ppm
3,15 ppm
Quale è la precisione di questo metodo usato?
Quante analisi dello stesso campione deve eseguire e mediare per avere un intervallo di confidenza al 99% inferiore a 0,8 ppm?