

CHIMICA ANALITICA · DISPENSA DI STUDIO

Statistica, calibrazione e validazione dei metodi analitici

Dalla misura al metodo validato: teoria, criteri ed esempi svolti

Corso di Laurea in Biotecnologie — II anno
Università degli Studi di Teramo
Prof. Federico Fanti · Anno Accademico 2026/2027

Come usare questa dispensa

Questa dispensa accompagna le lezioni *Statistica applicata ai metodi analitici* e *Calibrazione e validazione dei metodi analitici* del corso di Chimica Analitica. È pensata per lo studio autonomo: ogni argomento delle slide è spiegato per esteso, con il ragionamento che sta dietro alle formule, esempi svolti passo per passo e i criteri numerici delle linee guida internazionali. Puoi usarla sia per prepararti all'esame sia come riferimento durante le esercitazioni di laboratorio.

Com'è organizzata

La **Parte I** (Capitoli 1–8) costruisce gli strumenti statistici: incertezza ed errori di misura, precisione ed esattezza, statistica descrittiva, distribuzioni e inferenza, test per i valori anomali, incertezza di misura e regressione lineare. La **Parte II** (Capitoli 9–17) li applica alla calibrazione e alla validazione dei metodi: standard di riferimento, metodi di calibrazione, parametri di validazione, controllo qualità, effetto matrice, stabilità, quadro normativo e aspetti pratici della routine. Le **Appendici** raccolgono il formulario (A), le tavole statistiche (B), un glossario italiano–inglese (C) e i riferimenti bibliografici (D).

Le due parti sono collegate: quando un concetto è già stato spiegato, il testo rimanda al capitolo corrispondente invece di ripeterlo. Per questo conviene studiare la Parte I prima della Parte II.

I riquadri

Nel testo trovi sette tipi di riquadri, sempre con lo stesso colore:

Definizione

Le definizioni formali dei termini chiave, secondo il VIM, la norma ISO 5725 o le linee guida ICH.

Formula

Le formule importanti, con il significato di ogni simbolo e le unità di misura.

Esempio svolto

Problemi numerici risolti passo per passo. Rifalli da solo con la calcolatrice prima di leggere la soluzione.

Attenzione

Gli errori più comuni e le trappole tipiche, anche quelle che compaiono nelle domande d'esame.

Approfondimento

Dettagli per chi vuole andare oltre: non sono indispensabili per l'esame, ma aiutano a capire.

In sintesi

I punti chiave da ricordare, alla fine di ogni capitolo.

Mettiti alla prova

Gli argomenti del quiz di autovalutazione con cui esercitarti sul capitolo appena studiato.

Come studiare con il quiz

La dispensa è collegata al quiz di autovalutazione (file *Quiz-Statistica-Calibrazione-Validazione.html*, che si apre con qualsiasi browser). Un percorso efficace è:

1. leggere il capitolo e rifare a mano gli esempi svolti;
2. aprire il quiz in modalità **Esercitazione**, scegliere la parte e gli argomenti indicati nel riquadro *Mettiti alla prova* e leggere le spiegazioni delle risposte sbagliate;
3. tornare sul testo per gli argomenti in cui hai sbagliato;
4. quando hai completato entrambe le parti, provare la **Simulazione d'esame**: 30 domande in 30 minuti, con difficoltà calibrata.

Convenzioni

I numeri decimali sono scritti con la virgola (0,026 mg/L) e le unità di misura seguono il Sistema Internazionale. I termini tecnici sono accompagnati, alla prima occorrenza, dall'equivalente inglese usato nelle linee guida (per esempio *esattezza*, in inglese *trueness*). I simboli principali sono riassunti nella tabella seguente.

Tabella 0.1 — Simboli principali usati nella dispensa

Simbolo	Significato
$\bar{x}$	media aritmetica dei dati
s, σ	deviazione standard del campione e della popolazione
n, ν	numero di dati, gradi di libertà
$\mathrm{RSD} = \mathrm{CV}\%$	deviazione standard relativa (coefficiente di variazione)
u, u_c, U, k	incertezza tipo, incertezza composta, incertezza estesa, fattore di copertura
b_0, b_1	intercetta e pendenza della retta di calibrazione
s_B	deviazione standard del segnale del bianco
r, r^2	coefficiente di correlazione e coefficiente di determinazione
LOD, LOQ, LOL	limite di rivelabilità, di quantificazione e di linearità
LLOQ, ULOQ	limite inferiore e superiore di quantificazione della curva di calibrazione
SI, SIL-IS	standard interno, standard interno marcato con isotopi stabili

Le linee guida citate sono quelle in vigore nell'anno accademico 2026/2027: ICH M10 (2022) per la bioanalisi, ICH Q2(R2) e Q14 (2023) per i metodi farmaceutici, la guida Eurachem *The Fitness for Purpose of Analytical Methods* (3ª edizione, 2025), il documento SANTE/11312/2021 v2026 per i residui di pesticidi e il Regolamento (UE) 2021/808 per i residui di farmaci veterinari.

Indice

Parte I — Statistica applicata ai metodi analitici 8
1. Misura, incertezza e dati analitici 8
1.1 Perché ogni misura ha un'incertezza 8
1.2 Come si esprime correttamente un risultato 8
1.3 Dal dato all'informazione alla conoscenza 9
1.4 Tipologie di dati 9
1.5 Il segnale analitico e la curva di calibrazione 10
2. Gli errori di misura 11
2.1 Errore e incertezza: due concetti distinti 11
2.2 Errore assoluto, errore relativo e bias 12
2.3 I tre tipi di errore 13
2.4 Errori sistematici: le fonti 13
2.5 Errori sistematici costanti e proporzionali 13
2.6 Come individuare ed eliminare gli errori sistematici 14
2.7 Errori casuali 15
3. Precisione, esattezza e accuratezza 16
3.1 Le definizioni secondo VIM e ISO 5725 16
3.2 I quattro casi: i bersagli 17
3.3 Ripetibilità, precisione intermedia e riproducibilità 18
4. Statistica descrittiva 19
4.1 Campione e popolazione 19
4.2 Indicatori di tendenza centrale 20
4.3 Indicatori di dispersione 20
4.4 Esercizio completo: Fe(III) in un campione d'acqua 22
5. Distribuzioni, inferenza e valori anomali 23
5.1 La distribuzione normale 23
5.2 Distribuzioni simmetriche e asimmetriche 24
5.3 L'errore standard della media 25
5.4 La distribuzione t di Student 25
5.5 L'intervallo di confidenza della media 26
5.6 I valori anomali 26
6. L'incertezza di misura 29
6.1 Errore e incertezza 29
6.2 Il GUM: contributi di Tipo A e di Tipo B 30
6.3 La valutazione di Tipo A 30

6.4 La valutazione di Tipo B 31
6.5 L'incertezza tipo composta 31
6.6 L'incertezza estesa 32
6.7 Il bilancio di incertezza 33
6.8 Come si riporta il risultato 33
7. Regressione lineare e curva di calibrazione 34
7.1 Calibrazione e regressione 34
7.2 Il modello di regressione lineare semplice 35
7.3 Il metodo dei minimi quadrati 35
7.4 Interpolazione, estrapolazione e predizione inversa 36
7.5 Il diagramma di dispersione e la scelta del modello 37
7.6 Scomposizione della variabilità, r^2 e r 37
7.7 Esempio completo: calibrazione spettrofotometrica del Fe(III) 40
8. Diagnostica della regressione e collegamento alla validazione 41
8.1 L'analisi dei residui 41
8.2 Omoschedasticità, eteroschedasticità e regressione pesata 42
8.3 Effetto degli errori sistematici sulla curva di calibrazione 43
8.4 Collegamento alla validazione dei metodi 44
8.5 Sintesi dei concetti chiave della Parte I 45
Parte II — Calibrazione e validazione dei metodi analitici 46
9. Standard e materiali di riferimento 46
9.1 La gerarchia degli standard analitici 46
9.2 La tracciabilità metrologica 48
9.3 Dal PS agli standard di lavoro 48
10. I metodi di calibrazione 50
10.1 Metodi assoluti e metodi relativi 50
10.2 Il bianco analitico 50
10.3 Scelta del segnale: area o altezza del picco 51
10.4 La calibrazione con standard esterno 51
10.5 La curva di calibrazione: requisiti operativi 52
10.6 La calibrazione con standard interno 53
10.7 Il metodo delle aggiunte multiple 54
10.8 Confronto tra i metodi di calibrazione 55
11. La validazione: scopo, selettività, sensibilità, LOD/LOQ e range lineare 56
11.1 Perché validare un metodo analitico 56
11.2 I parametri di validazione 57
11.3 Selettività e specificità 57
11.4 Sensibilità di calibrazione e sensibilità analitica 58

11.5 Limite di rivelabilità e limite di quantificazione 59
11.6 Range dinamico lineare e intervallo di calibrazione 60
12. Esattezza, precisione e controllo qualità 62
12.1 Esattezza e bias 62
12.2 Accuratezza = esattezza + precisione 63
12.3 Precisione: ripetibilità, precisione intermedia, riproducibilità 63
12.4 La curva di Horwitz 63
12.5 I campioni di controllo qualità (QC) 65
12.6 Prestazioni ai livelli bassi e alti 66
12.7 Coerenza su tutto l'intervallo 66
13. Recupero, effetto matrice ed efficienza di processo 67
13.1 Definizioni e schema di Matuszewski 67
13.2 Le formule 68
13.3 L'effetto matrice in LC-MS/MS 69
13.4 Il fattore di matrice normalizzato sullo standard interno 70
13.5 La valutazione nelle linee guida 70
13.6 Individuare e ridurre l'effetto matrice 70
13.7 Il recupero ai diversi livelli di concentrazione 71
14. Stabilità, robustezza, solidità, carry-over e integrità della diluizione 72
14.1 La stabilità dell'analita 72
14.2 Robustezza 74
14.3 Solidità 74
14.4 Carry-over 74
14.5 Integrità della diluizione 75
15. Il quadro normativo 76
15.1 Gli enti e la natura dei documenti 76
15.2 La bioanalisi: ICH M10 (2022) 76
15.3 ICH Q2(R2) e ICH Q14 (2023): i metodi per i prodotti farmaceutici 77
15.4 Alimenti, farmaci veterinari e laboratori in generale 77
15.5 Confronto tra i documenti 79
16. Dal metodo alla routine: aspetti pratici 79
16.1 Metodi strumentali e principali tecniche 80
16.2 Il processo analitico 80
16.3 Propagazione dell'incertezza nelle diluizioni seriali 81
16.4 Monitoraggio della curva nel tempo 82
16.5 L'intercetta e la forzatura per l'origine 82
16.6 Punti aberranti e accettazione della curva 83
16.7 Lo standard interno marcato con isotopi stabili (SIL-IS) 83

16.8 Riepilogo dei criteri ICH M10 84
16.9 Validazione completa, parziale e cross-validazione 84
16.10 Piano e rapporto di validazione 85
16.11 Errori comuni nella costruzione della curva 85
16.12 La sequenza analitica in LC-MS/MS 85
17. Sintesi: dalla calibrazione alla validazione 86
17.1 Il quadro d'insieme 87
17.2 I messaggi chiave 87
17.3 Letture consigliate 88
Appendici 89
Appendice A — Formulario 89
A.1 Statistica descrittiva (Capitolo 4) 89
A.2 Errori di misura (Capitoli 2–3) 89
A.3 Inferenza e valori anomali (Capitolo 5) 90
A.4 Incertezza di misura (Capitolo 6) 90
A.5 Regressione lineare (Capitoli 7–8) 90
A.6 Calibrazione (Capitoli 9–10 e 16) 91
A.7 Parametri di validazione (Capitoli 11–14) 92
A.8 Criteri di accettazione ICH M10 (metodi cromatografici) 92
Appendice B — Tavole statistiche 93
B.1 Valori critici della t di Student (due code) 93
B.2 Valori critici del test Q di Dixon 94
B.3 Valori critici del test di Grubbs 94
B.4 Fattore di copertura e livello di fiducia 95
Appendice C — Glossario italiano–inglese 95
Appendice D — Riferimenti 98
D.1 Linee guida e norme 98
D.2 Articoli 98
D.3 Testi di riferimento 99

Parte I — Statistica applicata ai metodi analitici

Ogni misura chimica è affetta da errori. Questa parte fornisce gli strumenti statistici per descrivere i dati, stimare la loro incertezza, riconoscere i valori anomali e costruire una retta di calibrazione affidabile: la base su cui poggia la validazione dei metodi analitici della Parte II.

1. Misura, incertezza e dati analitici

Ogni decisione basata su un'analisi chimica si fonda su un numero, e quel numero è utile solo se sappiamo quanto fidarci di esso. Una glicemia di 7,1 mmol/L supera davvero la soglia diagnostica di 7,0 mmol/L? Dipende dall'incertezza. Questo capitolo spiega come si esprime un risultato, che tipi di dati produce un laboratorio e come il segnale di uno strumento diventa una concentrazione.

1.1 Perché ogni misura ha un'incertezza

Nessuna analisi chimica dà risultati privi di incertezza. Strumento, operatore, ambiente, campione e procedura introducono piccoli scostamenti di entità e segno in parte ignoti. Il valore vero del **misurando** (la grandezza da misurare) resta quindi sconosciuto e ogni risultato è una stima. Il compito dell'analista è **minimizzare gli errori e stimarne l'entità con un'accuratezza accettabile**. Una misura affidabile segue tre passi, in quest'ordine:

1. **stabilire il massimo errore tollerabile prima di iniziare l'analisi**: dipende dallo scopo (*fitness for purpose*) e guida la scelta del metodo, della vetreria e del numero di replicati;
2. **eseguire misure replicate**, per stimare la variabilità e calcolare il valore medio;
3. **esprimere il risultato come valore ± incertezza**, con unità di misura e cifre significative corrette.

1.2 Come si esprime correttamente un risultato

Un risultato analitico si scrive nella forma seguente, completata dall'unità di misura:

$$x = \bar{x} \pm U \tag{1.1}$$

dove $\bar{x}$ è la migliore stima (di solito la media dei replicati) e U è l'**incertezza estesa**. Per esempio: glucosio nel sangue = (5,42 ± 0,08) mmol/L. L'incertezza estesa dell'equazione 1.1 è l'incertezza tipo composta u_c, che riunisce tutti i contributi (ripetibilità, taratura, pesate, volumi; Capitolo 6), moltiplicata per un **fattore di copertura** k:

$$U = k \cdot u_c \tag{1.2}$$

Nell'equazione 1.2, con $k = 2$ il livello di fiducia è **circa il 95%**, con $k = 3$ circa il 99,7%. Queste corrispondenze valgono se la distribuzione è approssimativamente normale e i gradi di libertà sono sufficienti. U **non** è la deviazione standard dei replicati, né un errore o un bias: per calcolare questi servirebbe il valore vero (Capitolo 2). L'incertezza si esprime spesso in forma relativa, come percentuale del risultato:

$$U_{\text{rel}} = \frac{U}{\bar{x}} \times 100 \tag{1.3}$$

Per la glicemia dell'esempio l'equazione 1.3 dà $U_{\text{rel}} = 0{,}08/5{,}42 \times 100 = 1{,}5\%$. Non si usa l'ampiezza totale $2U$, che darebbe 3,0%.

1.2.1 Cifre significative e arrotondamento

Le **cifre significative** sono tutte le cifre note con certezza più la prima cifra incerta. Le regole per contarle sono:

- le cifre diverse da zero e gli zeri compresi tra di esse sono significativi;
- gli **zeri iniziali** non lo sono: servono solo a collocare la virgola;
- gli **zeri finali dopo la virgola** lo sono, perché indicano una cifra misurata;
- gli zeri finali di un intero sono ambigui: si usa la notazione scientifica ($1{,}20 \times 10^3$ mL ha tre cifre significative).

Così 0,02050 g ha **quattro** cifre significative (2, 0, 5, 0), come 20,50 mg. Nei calcoli intermedi si conservano tutte le cifre e **si arrotonda solo il risultato finale**. Per un risultato con incertezza valgono quattro regole pratiche:

1. l'incertezza si arrotonda a **1–2 cifre significative** (meglio due se la prima cifra è 1 o 2);
2. il valore centrale si arrotonda **alla stessa cifra decimale dell'incertezza**;
3. si indica sempre l'**unità di misura**;
4. si specificano il **fattore di copertura** e il livello di fiducia (per esempio $k = 2$, circa 95%).

Esempio 1.1 — Arrotondare un risultato con la sua incertezza

Dati: $\bar{x}$ = 12,3467 mg/L, U = 0,0832 mg/L ($k = 2$).

1. Si arrotonda l'incertezza a una cifra significativa: 0,0832 → 0,08 mg/L, che termina ai centesimi.
2. Si arrotonda il valore centrale ai centesimi: 12,3467 → 12,35 mg/L.

Risultato: (12,35 ± 0,08) mg/L ($k = 2$, circa 95%). Con due cifre sull'incertezza va bene anche (12,347 ± 0,083) mg/L. Sono sbagliati (12,3 ± 0,08), che perde una cifra, e (12,347 ± 0,08), con decimali diversi. Sono sbagliati anche "5,42 ± 0,08" senza unità e un'incertezza con tre cifre significative come 0,00427.

1.3 Dal dato all'informazione alla conoscenza

Il **dato** è la misura grezza di un fenomeno, per esempio un'assorbanza o l'area di un picco. Di per sé non ha significato. L'**analisi**, cioè l'elaborazione e la correlazione del dato (sottrazione del bianco, curva di calibrazione, calcolo di media e incertezza), lo trasforma in **informazione**, un risultato processato: "glucosio = (5,42 ± 0,08) mmol/L". Un **modello**, cioè un quadro interpretativo, trasforma l'informazione in **conoscenza**, cioè comprensione scientifica. La composizione del sangue acquista senso solo attraverso un modello del corpo umano e delle patologie: solo allora 5,42 mmol/L diventa "glicemia a digiuno nella norma". Un errore nel dato si propaga lungo tutta la catena.

1.4 Tipologie di dati

I **dati quantitativi** sono un **valore numerico con unità di misura** ("la temperatura dell'acqua è 298,2 K"). Sono la base delle *hard sciences*, discipline fondate su dati rigorosi collegati da modelli matematici; la scienza galileiana si fonda sui dati quantitativi. I **dati qualitativi** sono etichette, descrittori o categorie ("l'acqua è calda", "gruppo sanguigno A positivo"): sono difficili da standardizzare e riprodurre, come nell'analisi sensoriale. Tra i dati di un individuo, altezza, peso e concentrazione di glucosio nel sangue sono quantitativi; colore della pelle e gruppo sanguigno sono qualitativi.

I dati quantitativi possono essere **discreti**, cioè conteggi interi (numero di colonie o di cellule), oppure **continui**, cioè qualunque valore in un intervallo (massa, concentrazione, assorbanza). Si distinguono inoltre quattro **scale di misura**:

- **nominale**: categorie senza ordine (gruppo sanguigno);
- **ordinale**: categorie ordinate (torbidità +, ++, +++; taglie S, M, L);
- **a intervalli**: zero convenzionale (temperatura in °C);
- **di rapporti**: zero assoluto, per cui i rapporti hanno senso (massa, concentrazione, temperatura in K).

Media, deviazione standard e CV% (Capitolo 4) hanno pieno senso solo sulle ultime due scale; il CV% solo sulla scala di rapporti.

1.5 Il segnale analitico e la curva di calibrazione

Il **segnale analitico** è la grandezza misurata dallo strumento che dipende dalla quantità di analita: assorbanza, intensità di emissione, area di un picco cromatografico, conteggi di ioni. Condizione necessaria è una **relazione funzionale nota** tra quantità di analita e sorgente del segnale, come la legge di Lambert–Beer in spettrofotometria. Il segnale è **proporzionale** alla concentrazione solo nell'intervallo lineare e se il segnale del bianco è nullo o sottratto. In generale, nell'intervallo lineare vale:

$$y = b_0 + b_1 x \tag{1.4}$$

con intercetta b_0 (segnale del bianco) e pendenza b_1. Se $b_0 \neq 0$ la relazione è lineare ma non proporzionale. Per costruire la **curva di calibrazione** si preparano soluzioni standard a concentrazione nota, se ne misura il segnale e si stimano b_0 e b_1 (Capitolo 7). Dal segnale y_0 di un campione si ricava poi la concentrazione invertendo l'equazione 1.4:

$$x_0 = \frac{y_0 - b_0}{b_1} \tag{1.5}$$

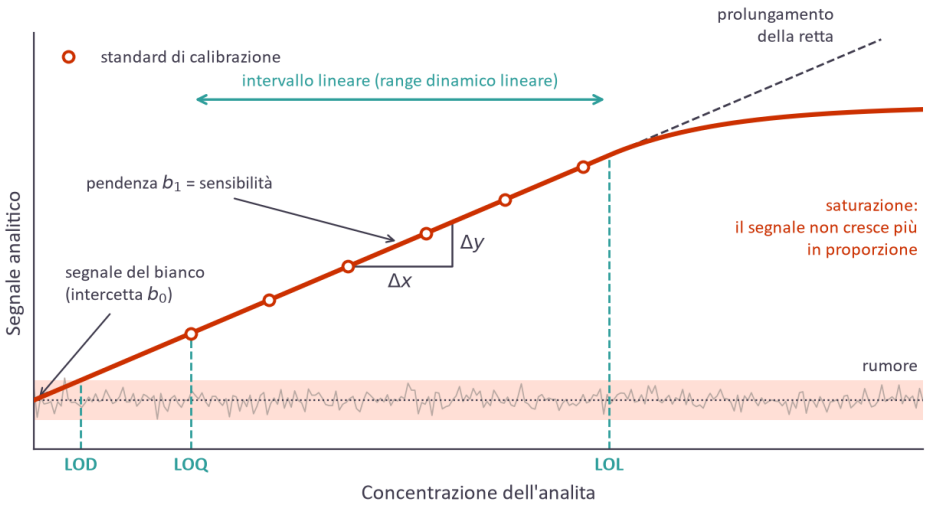


Figura 1.1 — Andamento tipico del segnale analitico in funzione della concentrazione: segnale del bianco con rumore, LOD e LOQ, intervallo lineare fino al LOL, saturazione oltre il LOL.

Nella Figura 1.1 si leggono le grandezze principali:

- la **pendenza** $b_1 = \Delta y / \Delta x$ è la **sensibilità**: più la retta è ripida, meglio il metodo distingue piccole differenze di concentrazione. Non va confusa con il limite di rivelabilità;
- **LOD** e **LOQ** sono le concentrazioni minime rilevabile e quantificabile (Capitolo 11);
- l'**intervallo lineare** va dal LOQ al **LOL** (limite di linearità);
- oltre il LOL il rivelatore **satura** e il segnale tende a un plateau: la proporzionalità non vale più.

Un campione oltre l'intervallo di calibrazione non si quantifica per estrapolazione: si diluisce e si rianalizza (Capitolo 14). Il plateau non dipende dal bianco, che sposterebbe la retta parallelamente (errore costante, Capitolo 2), né dagli errori casuali.

Esempio 1.2 — Dal segnale alla concentrazione, e quando diluire

Dati: retta corretta per il bianco e passante per l'origine, b_1 = 0,0760 L/mg, intervallo di calibrazione 0,50–10,0 mg/L.

1. Un campione con A = 0,380 dà (equazione 1.5) $x_0 = 0{,}380/0{,}0760$ = 5,00 mg/L, dentro l'intervallo.
2. Un secondo campione con A = 0,950 darebbe 12,5 mg/L, oltre lo standard più alto: il valore non è affidabile.
3. Diluito 1:5, il secondo campione dà A = 0,231, cioè 3,04 mg/L; moltiplicando per 5 si ottiene 15,2 mg/L.

Risultato: 5,00 mg/L e 15,2 mg/L. Vicino al plateau la lettura diretta sottostimava di circa il 18%.

Punti chiave del Capitolo 1

- Ogni misura ha un'incertezza; il massimo errore tollerabile si stabilisce prima di iniziare l'analisi.
- Il risultato si scrive $\bar{x} \pm U$ con unità di misura; $U = k \cdot u_c$ ($k = 2$, circa 95%) non è un errore né la deviazione standard dei replicati.
- L'incertezza ha 1–2 cifre significative e il valore centrale termina alla stessa cifra decimale; gli zeri iniziali non sono significativi.
- L'analisi trasforma il dato in informazione, il modello l'informazione in conoscenza.
- I dati sono quantitativi o qualitativi, discreti o continui, su scala nominale, ordinale, a intervalli o di rapporti.
- La pendenza della curva di calibrazione è la sensibilità; la proporzionalità vale solo nell'intervallo lineare, fino al LOL.

Mettiti alla prova

- **Incertezza ed espressione del risultato**
- **Dati e segnale analitico**

2. Gli errori di misura

Per migliorare un metodo bisogna sapere da dove vengono gli errori e come si comportano. Alcuni spostano tutti i risultati nella stessa direzione e si possono correggere; altri li disperdono in modo imprevedibile e si possono solo ridurre e quantificare; altri ancora sono sbagli da eliminare. Riconoscere il tipo di errore permette di scegliere il rimedio giusto.

2.1 Errore e incertezza: due concetti distinti

Il vocabolario internazionale di metrologia (VIM) e la guida GUM distinguono due concetti che il linguaggio comune confonde.

- L'**errore di misura** è una **differenza con segno**: valore misurato meno valore vero (o di riferimento). Poiché il valore vero è in genere ignoto, anche l'errore lo è; lo si può calcolare solo analizzando un materiale con valore noto.

- L'**incertezza di misura** è un **parametro di dispersione**, sempre positivo, che indica l'ampiezza dell'intervallo dei valori attribuibili al misurando ($\bar{x} \pm U$, Capitolo 1). Si stima senza conoscere il valore vero.

Espressioni come "barre di errore" o "errore di ± 0,1 mL" usano "errore" nel senso di incertezza: è un uso colloquiale da evitare. I due concetti sono indipendenti (Figura 2.1). Un risultato con incertezza piccola ma lontano dal valore vero segnala un errore sistematico non individuato.

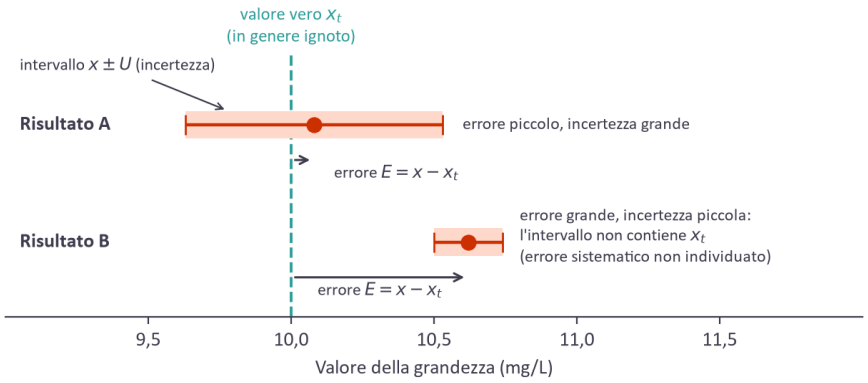


Figura 2.1 — Errore e incertezza sono concetti distinti: il risultato A ha errore piccolo e incertezza grande, il risultato B errore grande e incertezza piccola.

2.2 Errore assoluto, errore relativo e bias

L'**errore assoluto** di un risultato x_i rispetto al valore vero x_t e l'**errore relativo** percentuale sono:

$$E = x_i - x_t \tag{2.1}$$

$$E_r = \frac{x_i - x_t}{x_t} \times 100 \tag{2.2}$$

Un errore positivo indica una **sovrastima**, uno negativo una **sottostima**. Si sottrae sempre il valore vero dal valore misurato e, nell'errore relativo, si divide per il valore vero. Per esempio, uno standard da 10,40 mg/L analizzato come 10,25 mg/L ha $E_r = -0{,}15/10{,}40 \times 100 = -1{,}44\%$.

L'errore di un singolo risultato contiene una componente sistematica e una casuale. Il **bias** è l'errore sistematico della popolazione:

$$\text{bias} = \mu - x_t \tag{2.3}$$

dove μ è la media della popolazione, cioè degli infiniti risultati possibili. L'errore del singolo risultato si scompone quindi in bias più errore casuale:

$$x_i - x_t = (\mu - x_t) + (x_i - \mu)$$

In pratica il bias si stima con la media $\bar{x}$ dei replicati eseguiti su un materiale di valore noto. Si esprime anche in forma relativa, $(\bar{x} - x_t)/x_t \times 100$, o come recupero, $\bar{x}/x_t \times 100$.

Esempio 2.1 — Errore di una pesata e bias di un metodo

Dati: (a) una massa tarata di 20,0000 g pesata su una bilancia dà 20,0012 g; (b) uno standard da 10,00 mg/L analizzato cinque volte dà 10,13; 10,08; 10,15; 10,05; 10,09 mg/L.

1. (a) Equazioni 2.1 e 2.2: $E = 20{,}0012 - 20{,}0000 = +0{,}0012$ g = +1,2 mg, cioè E_r = +0,006%: la bilancia **sovrastima**.
2. (b) $\bar{x} = 50{,}50/5 = 10{,}10$ mg/L, quindi (equazione 2.3, con $\bar{x}$ al posto di μ) bias = $10{,}10 - 10{,}00$ = +0,10 mg/L (+1,0%).

Risultato: E = +1,2 mg; bias = +0,10 mg/L. Gli errori dei singoli risultati (da +0,05 a +0,15 mg/L) contengono anche la componente casuale: il bias si stima con la media.

2.3 I tre tipi di errore

- Gli **errori grossolani** sono sbagli macroscopici e facilmente individuabili: campione sbagliato, provette scambiate, contaminazione evidente. **Si eliminano** ripetendo l'analisi, non si trattano statisticamente.
- Gli **errori sistematici** (determinati) agiscono secondo **leggi definite**: hanno sempre lo stesso segno e causano sovrastima o sottostima. **Si individuano e si correggono** e determinano l'**esattezza** (Capitolo 3).
- Gli **errori casuali** (indeterminati) derivano da **fluttuazioni imprevedibili** dei parametri sperimentali e disperdono i risultati attorno alla media. **Non si eliminano**, ma si riducono e si quantificano statisticamente; determinano la **precisione**.

Aumentare i replicati e calcolarne la media **riduce l'effetto degli errori casuali, ma non il bias**. Gli errori casuali hanno segno variabile e nella media si compensano: la dispersione della media diminuisce come $s/\sqrt{n}$ (Capitolo 5). Il bias sposta invece tutti i risultati nello stesso verso: cento misure con una pipetta starata danno una media precisa ma sbagliata.

2.4 Errori sistematici: le fonti

Gli errori sistematici si classificano in base alla **causa**:

- **strumentali**: inesatta calibrazione, uso improprio di vetreria e strumenti, strumenti non idonei. Esempio: una micropipetta che eroga sempre il 2% in meno del volume impostato commette un errore strumentale e, poiché è una frazione fissa del volume, anche proporzionale (paragrafo 2.5). Si elimina tarando la pipetta per via gravimetrica;
- **di metodo**: comportamento non ideale di reattivi e reazioni, condizioni sperimentali non idonee. Esempi: precipitazione incompleta, solubilità inattesa, tempi di calcinazione inadeguati, indicatore che vira prima del punto equivalente, estrazione incompleta, soppressione della ionizzazione in LC-MS. Sono i più difficili da scoprire;
- **personali**: distrazione, scarsa conoscenza della procedura, difetti fisici e pregiudizi inconsci. Un analista daltonico che vede il viraggio in ritardo, o la tendenza a fermare la titolazione vicino al volume atteso, producono errori che si ripetono sempre con lo stesso segno: sono sistematici, non grossolani.

2.5 Errori sistematici costanti e proporzionali

Gli errori sistematici si distinguono anche per come dipendono dalla quantità di analita (Figura 2.2).

Un **errore costante** ha entità **indipendente dalla quantità di analita**. Sulla retta di calibrazione produce una **traslazione parallela**: cambia l'intercetta, non la pendenza. Esempi: le perdite per solubilità nel lavaggio in gravimetria (bias negativo) e gli assorbimenti estranei in spettrofotometria (bias positivo). L'errore assoluto è fisso, quindi **l'errore relativo diminuisce all'aumentare della quantità** di analita o della massa di campione.

Un **errore proporzionale** ha entità **proporzionale alla quantità di analita**. **Modifica la pendenza** della retta e lascia invariata l'intercetta. Esempi: perdite per estrazioni con resa incompleta (bias negativo) e un effetto matrice che amplifica il segnale (bias positivo). **L'errore relativo è uguale a ogni livello**, mentre quello

assoluto cresce con la quantità di analita. Per distinguere i due casi si confrontano campioni di massa diversa. Un materiale con 10,0 mg/g di analita, analizzato in porzioni da 0,500 g e 1,000 g (5,00 e 10,00 mg di analita), dà per esempio 4,50 e 9,00 mg. Gli errori assoluti (−0,50 e −1,00 mg) raddoppiano con la massa, mentre l'errore relativo resta −10%: l'errore è **proporzionale negativo**. Un errore costante di −0,50 mg darebbe invece 9,50 mg con la porzione maggiore (−5%).

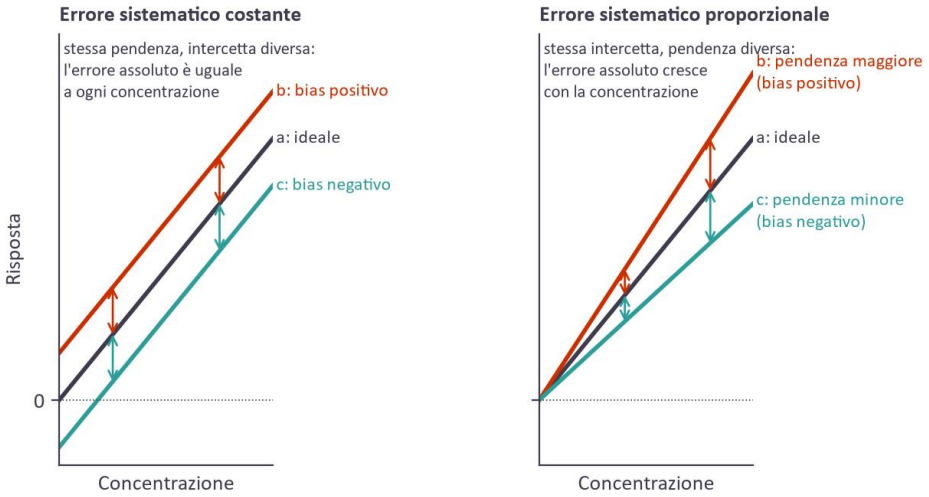


Figura 2.2 — Errore costante (a sinistra): retta traslata, stessa pendenza, errore assoluto uguale a ogni concentrazione. Errore proporzionale (a destra): stessa intercetta, pendenza diversa, errore assoluto crescente.

Esempio 2.2 — Errori costanti e proporzionali in pratica

Dati: (a) il lavaggio di un precipitato ne scioglie sempre 0,50 mg; (b) un indicatore richiede sempre 0,04 mL di titolante in eccesso; (c) un CRM con 50,0 µg/kg di piombo dà in media 47,6 µg/kg, con errore proporzionale.

1. (a) Errore costante: $E_r = -0{,}50/50{,}0 \times 100 = -1{,}0\%$ con 50,0 mg di precipitato e $-0{,}50/500 \times 100 = -0{,}10\%$ con 500 mg.
2. (b) Perché l'errore relativo non superi lo 0,1% (cioè 0,001) serve $V \geq 0{,}04/0{,}001 = 40$ mL di titolante.
3. (c) Bias relativo $= (47{,}6 - 50{,}0)/50{,}0 \times 100 = -4{,}8\%$: il metodo restituisce il 95,2% del valore vero. Per un campione con 120,0 µg/kg ci si aspetta $120{,}0 \times 0{,}952 = 114{,}2$ µg/kg.

Risultato: −1,0% e −0,10%; almeno 40 mL; 114,2 µg/kg. Se l'errore in (c) fosse costante, si sottrarrebbero 2,4 µg/kg a ogni livello (117,6 µg/kg). Per correggere un risultato affetto dall'errore proporzionale si divide per 0,952.

2.6 Come individuare ed eliminare gli errori sistematici

Un errore sistematico non si vede dalla dispersione dei replicati: serve un termine di confronto esterno.

1. **Materiali di riferimento certificati (CRM)**: campioni con valore certificato da produttori competenti (Capitolo 9). Si confronta la media dei risultati con il valore certificato.
2. **Metodo indipendente**: confronto con un metodo di riferimento o con strumentazione di provata affidabilità.
3. **Analisi del bianco**: soluzione con tutti i componenti tranne l'analita. Il suo segnale, dovuto solo a contributi estranei all'analita, si sottrae a quello del campione e corregge i contributi **additivi** (errori costanti). Non corregge gli effetti **moltiplicativi** della matrice (errori proporzionali), che richiedono calibrazione in matrice, aggiunte standard o standard interno (Capitolo 10). Un vero bianco di matrice richiede inoltre una matrice priva di analita, non sempre disponibile.
4. **Prove di recupero** (*spike*): aggiunta di una quantità nota di analita al campione reale. Sono la scelta giusta quando manca un CRM con matrice simile, perché lavorano nella matrice effettiva.

5. **Variazione della massa di campione**: rivela gli errori costanti, non quelli proporzionali.

Il recupero percentuale si calcola con l'equazione 2.4:

$$R\% = \frac{C_{\text{arricchito}} - C_{\text{campione}}}{C_{\text{aggiunta}}} \times 100 \tag{2.4}$$

Esempio 2.3 — Bianco e recupero

Dati: (a) retta corretta per il bianco, passante per l'origine, pendenza 0,0760 L/mg; campione A = 0,415, bianco A = 0,035. (b) Acqua con 2,10 mg/L di nitrato; dopo un'aggiunta di 1,00 mg/L si trovano 3,02 mg/L.

1. (a) $(0{,}415 - 0{,}035)/0{,}0760 = 5{,}00$ mg/L. Senza sottrarre il bianco si otterrebbe 5,46 mg/L: l'errore costante positivo causa una sovrastima.
2. (b) $R\% = (3{,}02 - 2{,}10)/1{,}00 \times 100 = 92\%$.

Risultato: 5,00 mg/L; recupero 92%. Il rapporto tra totale trovato e totale atteso (3,02/3,10 = 97%) è sbagliato, perché attenua l'effetto della perdita.

2.7 Errori casuali

Gli errori casuali derivano da fluttuazioni indefinite di temperatura, pH, pressione, umidità, forza ionica, rumore del rivelatore e lettura delle scale. Causano una **dispersione simmetrica** dei dati attorno al valore medio: errori positivi e negativi sono ugualmente probabili e quelli piccoli più frequenti di quelli grandi (distribuzione normale, Capitolo 5). **Non si possono eliminare**, ma si riducono operando con cura e si **quantificano** con la statistica su misure ripetute: deviazione standard (Capitolo 4), errore standard e intervallo di confidenza (Capitolo 5).

Esempio: la lettura della buretta. Una buretta è graduata ogni 0,1 mL; questa è la divisione della scala, cioè la risoluzione di lettura, da non confondere con la *sensibilità* dello strumento. La lettura 9,68 mL ha i decimi letti sulle tacche e i centesimi **stimati a occhio** tra due tacche, a circa 1/10 della divisione. La cifra stimata varia da una lettura all'altra ed è soggetta a errore casuale, ma va riportata perché è significativa.

La tolleranza della vetreria non è un errore casuale

Per **un determinato pezzo** di vetreria lo scostamento dal valore nominale è fisso: usando sempre la stessa pipetta si commette un **errore sistematico**, eliminabile con la taratura. La tolleranza dichiarata (per esempio ±0,02 mL per una pipetta da 10 mL di classe A) entra invece nell'incertezza come contributo di Tipo B (Capitolo 6). Diventa una fonte di variabilità casuale solo se si usano a caso pezzi diversi. Sono invece tipicamente casuali le piccole fluttuazioni non controllate della temperatura.

Punti chiave del Capitolo 2

- Errore (differenza con segno, in genere ignota) e incertezza (parametro di dispersione) sono concetti distinti.
- $E = x_i - x_t$; E_r si divide per il valore vero; bias $= \mu - x_t$, stimato con la media dei replicati.
- Gli errori grossolani si eliminano, i sistematici si correggono, i casuali si riducono e si quantificano. Più replicati riducono l'effetto degli errori casuali, non il bias.
- Gli errori sistematici sono strumentali, di metodo o personali.
- Un errore costante cambia l'intercetta e pesa meno sui campioni grandi; un errore proporzionale cambia la pendenza e ha errore relativo costante.
- Si individuano con CRM, metodo indipendente, bianco (solo contributi additivi), recupero e variazione della massa di campione.

Mettiti alla prova

- **Tipi di errore ed errori casuali**
- **Errore assoluto, errore relativo e bias**
- **Errori sistematici**

3. Precisione, esattezza e accuratezza

"Questo metodo è accurato?" è una domanda che richiede di separare due proprietà: quanto i risultati sono vicini tra loro e quanto la loro media è vicina al valore vero. Dipendono da errori diversi (Capitolo 2) e si migliorano con rimedi diversi. I criteri numerici usati nella validazione sono nel Capitolo 12.

3.1 Le definizioni secondo VIM e ISO 5725

Le definizioni seguono il VIM e la norma ISO 5725 (*Accuracy (trueness and precision) of measurement methods and results*). L'**accuratezza** è il concetto più ampio e comprende **esattezza** e **precisione**.

Precisione, esattezza e accuratezza

- **Precisione** (*precision*): grado di concordanza tra risultati di misurazioni ripetute sullo stesso oggetto, o su oggetti simili, **in condizioni specificate**. Dipende dagli errori casuali e si quantifica con la deviazione standard.
- **Esattezza** (*trueness*): grado di concordanza tra il valore medio di un numero molto grande di risultati e il valore vero x_t (o un valore di riferimento). Corrisponde all'assenza di errori sistematici e si esprime con il **bias**.
- **Accuratezza** (*accuracy*): grado di concordanza tra un risultato e il valore vero. **Comprende esattezza e precisione**: un risultato accurato è sia esatto sia preciso.

La **precisione** descrive la dispersione dei risultati attorno alla loro media e **non richiede il valore vero**. Si esprime con la deviazione standard s (Capitolo 4) o con la deviazione standard relativa:

$$\mathrm{RSD} = \mathrm{CV}\% = \frac{s}{\bar{x}} \times 100 \tag{3.1}$$

Più grande è s, peggiore è la precisione. La definizione parla di condizioni *specificate*: misure nelle stesse condizioni definiscono la ripetibilità, misure in condizioni diverse la precisione intermedia o la riproducibilità (paragrafo 3.3).

L'**esattezza** richiede invece un **valore di riferimento**, come il valore certificato di un CRM, e si quantifica con il bias, $\bar{x} - x_t$ (equazione 2.3), o con il recupero. Più piccolo è il bias, maggiore è l'esattezza. La differenza tra la media e il valore vero misura quindi l'esattezza, non l'accuratezza.

L'**accuratezza** riguarda il singolo risultato, il cui errore è la somma di bias ed errore casuale: è buona solo se entrambi sono piccoli. Un metodo esatto ma impreciso dà una media corretta, ma singoli risultati anche lontani dal valore vero. Un metodo preciso ma non esatto dà risultati vicini tra loro, ma tutti spostati.

Accuracy nelle linee guida

ICH Q2(R2) (2023) e ICH M10 (2022) usano *accuracy* per la vicinanza della media al valore nominale, cioè per l'esattezza ISO, e valutano la precisione a parte (Capitolo 12). In questa dispensa: **esattezza** = assenza di bias, **accuratezza** = esattezza + precisione.

Esempio 3.1 — Due metodi a confronto su un CRM

Dati: CRM con valore certificato 100,0 mg/kg, sei analisi per metodo. Metodo 1: $\bar{x}$ = 100,4 mg/kg, s = 3,1 mg/kg. Metodo 2: $\bar{x}$ = 94,8 mg/kg, s = 0,5 mg/kg.

1. Bias: metodo 1 = +0,4 mg/kg (+0,4%); metodo 2 = −5,2 mg/kg (−5,2%).
2. RSD (equazione 3.1): metodo 1 = $3{,}1/100{,}4 \times 100 = 3{,}1\%$; metodo 2 = $0{,}5/94{,}8 \times 100 = 0{,}53\%$.

Risultato: il metodo 1 ha migliore esattezza, il metodo 2 migliore precisione. Nessuno dei due è pienamente accurato. Una s piccola non basta: il metodo 2 ha colpi raggruppati ma lontani dal centro.

3.2 I quattro casi: i bersagli

Nei bersagli della Figura 3.1 il centro è il valore vero e ogni colpo un risultato. La distanza della media dei colpi dal centro rappresenta il bias, l'ampiezza del gruppo la dispersione.

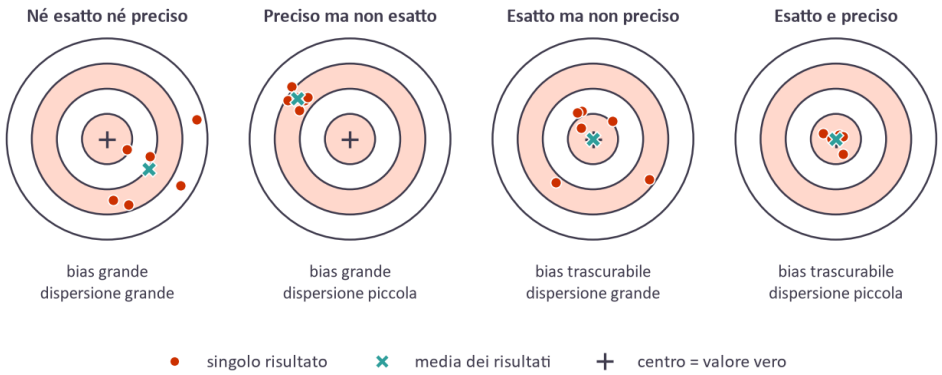


Figura 3.1 — I quattro casi di esattezza e precisione. Punti rossi: singoli risultati; croce verde: media; centro: valore vero.

1. **Né esatto né preciso**: bias e dispersione grandi; errori sistematici e casuali entrambi rilevanti.
2. **Preciso ma non esatto**: colpi raggruppati ma lontani dal centro; domina l'errore sistematico, quello casuale è piccolo.
3. **Esatto ma non preciso**: colpi dispersi ma centrati; bias trascurabile, errore casuale elevato.
4. **Esatto e preciso**: colpi raggruppati e centrati. È l'obiettivo ideale e l'unico caso accurato.

Gli errori casuali sono sempre presenti

Dire che il caso 2 ha "solo errori sistematici" e il caso 3 "solo errori casuali" è una semplificazione: gli errori casuali ci sono sempre, e si parla di errore **dominante**. Il caso 3 è "esatto ma non preciso", non "accurato ma non preciso". Inoltre in laboratorio il centro non si vede: dal gruppo di risultati si giudica la precisione, ma per l'esattezza serve un riferimento esterno.

C'è anche una differenza pratica tra i due difetti. Un bias noto e stabile si può correggere, per esempio con un fattore di recupero (Capitolo 2). L'imprecisione invece si può solo ridurre, mediando più replicati o migliorando il metodo. La correzione del bias non sostituisce la ricerca della sua causa e aggiunge a sua volta un contributo di incertezza (Capitolo 6).

3.3 Ripetibilità, precisione intermedia e riproducibilità

La ISO 5725 e il VIM distinguono tre condizioni di precisione, con variabilità crescente.

Condizioni di precisione (ISO 5725)

- **Ripetibilità** (s_r): stessa procedura, **stesso operatore, stesso strumento, stesso laboratorio, breve intervallo di tempo**. Esprime la variabilità minima del metodo.
- **Precisione intermedia** (s_I): **stesso laboratorio**, su un periodo esteso, con giorni, operatori, strumenti o lotti di reattivi diversi. Descrive la variabilità della routine.
- **Riproducibilità** (s_R): **laboratori diversi** che applicano lo stesso metodo allo stesso materiale. Si stima con studi interlaboratorio ed è essenziale per la standardizzazione dei metodi.

Il punto chiave è il **laboratorio**: finché si resta nello stesso laboratorio, cambiare operatore, strumento o giorno definisce la precisione intermedia. Si parla di riproducibilità solo con laboratori diversi (Tabella 3.1). In bioanalisi la ripetibilità corrisponde alla precisione *within-run* di ICH M10 e la precisione intermedia a quella *between-run*.

Tabella 3.1 — Che cosa varia nelle tre condizioni di precisione

Fattore	Ripetibilità	Precisione intermedia	Riproducibilità
Laboratorio	stesso	stesso	**diverso**
Operatore, strumento	stessi	possono variare	diversi
Tempo	breve	esteso (giorni diversi)	esteso
Variabilità attesa	minima	intermedia	massima

In uno **studio interlaboratorio** più laboratori analizzano lo stesso materiale omogeneo con lo stesso metodo. La deviazione standard dell'insieme dei risultati di tutti i laboratori stima la riproducibilità, perché include le differenze di operatori, strumenti, reattivi e ambiente e il piccolo bias proprio di ogni laboratorio. Una deviazione standard descrive sempre la precisione, mai l'esattezza.

Tutte e tre si quantificano con la deviazione standard nelle rispettive condizioni, o con la RSD. Ci si attende $s_r \leq s_I \leq s_R$. Nel modello della ISO 5725-2 la varianza di riproducibilità somma la varianza di ripetibilità e quella tra laboratori, s_L^2:

$$s_R^2 = s_r^2 + s_L^2 \tag{3.2}$$

Si sommano le **varianze**, non le deviazioni standard. Per i parametri del modello la riproducibilità è quindi sempre maggiore o uguale alla ripetibilità. Con dati reali la stima di s_R può però risultare minore di s_r per effetto del caso; la ISO 5725-2 prevede allora di porre $s_R = s_r$.

Esempio 3.2 — Deviazione standard di riproducibilità

Dati: studio interlaboratorio con s_r = 0,30 mg/L e s_L = 0,40 mg/L.

1. Equazione 3.2: $s_R^2 = 0{,}30^2 + 0{,}40^2 = 0{,}09 + 0{,}16 = 0{,}25$ (mg/L)².
2. $s_R = \sqrt{0{,}25} = 0{,}50$ mg/L.

Risultato: s_R = 0,50 mg/L. Sommare le deviazioni standard (0,70 mg/L) è sbagliato; 0,25 è la varianza, non la deviazione standard.

Punti chiave del Capitolo 3

- Precisione: concordanza tra risultati ripetuti in condizioni specificate; dipende dagli errori casuali, si esprime con s o RSD e non richiede il valore vero.
- Esattezza (*trueness*): concordanza tra media e valore vero; dipende dagli errori sistematici e si esprime con il bias.
- Accuratezza (*accuracy*) = esattezza + precisione.
- Nei bersagli lo spostamento del gruppo è il bias, la sua ampiezza la dispersione; gli errori casuali sono sempre presenti.
- Ripetibilità: stesse condizioni; precisione intermedia: stesso laboratorio, condizioni variate; riproducibilità: laboratori diversi.
- $s_R^2 = s_r^2 + s_L^2$: si sommano le varianze.

Mettiti alla prova

- **Precisione, esattezza e accuratezza**
- **Ripetibilità e riproducibilità**

4. Statistica descrittiva

Un risultato analitico nasce quasi sempre da più replicati dello stesso campione, simili ma non identici a causa degli errori casuali (Capitolo 2). La statistica descrittiva riassume la serie con un **valore centrale**, che rappresenta il risultato, e un **indice di dispersione**, che ne quantifica la precisione: sono le grandezze su cui si basano l'intervallo di confidenza (Capitolo 5), l'incertezza (Capitolo 6) e i criteri di validazione (Capitolo 12).

4.1 Campione e popolazione

Popolazione e campione

Popolazione: insieme di tutte le possibili misure della grandezza nelle condizioni considerate, teoricamente infinito. I suoi **parametri** si indicano con lettere greche: μ (media) e σ (deviazione standard).

Campione: insieme finito delle n misure effettivamente eseguite. Le sue **statistiche** si indicano con lettere latine: $\bar{x}$ (media) e s (deviazione standard). Le statistiche del campione sono **stime** dei parametri della popolazione.

I parametri μ e σ sono fissi ma **ignoti**; $\bar{x}$ e s sono noti ma variano da un campione all'altro. Aumentare il numero di replicati non trasforma $\bar{x}$ e s in μ e σ: le rende stime più affidabili. La distribuzione del campione si rappresenta in modo **discreto**, con un **istogramma** (barre alte quanto il numero di misure in ciascuna classe di

valori); quella della popolazione con una **curva continua**, che per gli errori casuali è la gaussiana (Capitolo 5). Ricordiamo infine che μ coincide con il valore vero solo in assenza di errori sistematici: la differenza tra i due è il bias (Capitolo 2).

4.2 Indicatori di tendenza centrale

La **media aritmetica** è la somma delle misure divisa per il loro numero:

$$\bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{n} = \frac{\sum_{i=1}^{n} x_i}{n} \tag{4.1}$$

La media usa tutti i dati ed è la stima migliore di μ per dati distribuiti normalmente; la somma degli scarti dalla media, $\sum(x_i - \bar{x})$, è sempre zero. Proprio perché usa tutti i dati, però, è **sensibile ai valori anomali** (outlier).

La **mediana** è il valore centrale dei dati **ordinati in senso crescente**: con n dispari è il valore in posizione $(n + 1)/2$; con n pari è la **media dei due valori centrali** (posizioni $n/2$ e $n/2 + 1$). Dipende solo dall'ordine dei dati, non da quanto siano lontani gli estremi: è un indicatore **robusto**, preferibile alla media quando si sospettano valori anomali o la distribuzione è asimmetrica.

La **moda** è il valore più frequente; una serie può non averne o averne più d'una. Si usa poco come risultato, ma il confronto tra media, mediana e moda rivela la forma della distribuzione (Capitolo 5).

Esempio 4.1 — Media e mediana in presenza di un valore anomalo

Si considerano i sette valori 12, 15, 20, 23, 24, 27, 200.

1. Media: $(12 + 15 + 20 + 23 + 24 + 27 + 200)/7 = 321/7 = 45{,}9$.
2. Mediana: dati già ordinati, $n = 7$ dispari, 4° valore = 23.
3. Senza il valore 200 ($n = 6$, pari): media $121/6 = 20{,}2$; mediana = media del 3° e del 4° valore, $(20 + 23)/2 = 21{,}5$.

Risultato: il valore 200 sposta la media a 45,9, lontano da tutti gli altri dati, mentre la mediana (23) resta rappresentativa; eliminandolo la media cambia di oltre 25 unità, la mediana di 1,5.

Errori frequenti con la mediana

Prendere il valore centrale della serie nell'ordine di acquisizione senza ordinarla; con n pari, scegliere uno solo dei due valori centrali; confondere la mediana con la media di minimo e massimo (che dipende proprio dagli estremi) o con la moda.

4.3 Indicatori di dispersione

L'**intervallo** (range) $w = x_{\max} - x_{\min}$ è immediato ma usa solo i due valori estremi; compare nel test Q di Dixon (Capitolo 5).

Per usare tutti i dati si parte dagli scarti $x_i - \bar{x}$; poiché la loro somma è zero, si elevano al quadrato e se ne fa la media, ottenendo la **varianza** (σ^2 per la popolazione, s^2 per il campione). La varianza ha come unità il quadrato dell'unità dei dati, per esempio (mg/L)², ma ha una proprietà preziosa: **le varianze di contributi indipendenti si sommano** (Capitolo 6). La sua radice quadrata è la **deviazione standard**, che ha la **stessa unità dei dati**.

Deviazione standard della popolazione e del campione

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \mu)^2}{n}} \tag{4.2}$$

$$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}} \tag{4.3}$$

- La 4.2 vale quando si conosce l'intera popolazione; con una serie di replicati si usa **sempre** la 4.3.
- $n - 1 = \nu$ sono i **gradi di libertà**.

Perché n − 1. Gli scarti sono calcolati rispetto a $\bar{x}$, ricavata dagli stessi dati; dato che la loro somma deve essere zero, solo $n - 1$ scarti sono indipendenti: il calcolo di $\bar{x}$ "consuma" un grado di libertà. Inoltre $\bar{x}$ è il valore che rende minima la somma dei quadrati degli scarti, per cui gli scarti da $\bar{x}$ sono in media più piccoli di quelli dall'ignota μ: dividendo per n si **sottostimerebbe sistematicamente** σ, e il denominatore $n - 1$ compensa questo effetto. La differenza conta con pochi dati: con tre replicati, dividere per n abbassa s del 18%. Con una sola misura s non si può calcolare. Attenzione ai programmi: in Excel la funzione del campione è DEV.ST.C (non DEV.ST.P); sulle calcolatrici è il tasto s o σ_{n-1}.

Per confrontare la precisione a livelli di concentrazione diversi, tra metodi o tra analiti, la deviazione standard si rapporta alla media.

Deviazione standard relativa (RSD) e coefficiente di variazione (CV%)

$$\mathrm{RSD} = \frac{s}{\bar{x}} \tag{4.4}$$

$$\mathrm{CV\%} = \frac{s}{\bar{x}} \times 100 \tag{4.5}$$

- La RSD è **adimensionale** (rapporto tra grandezze con la stessa unità); in percentuale si chiama CV%.
- Al numeratore va s, non l'errore standard $s/\sqrt{n}$. Un CV più basso indica una **precisione relativa migliore**.

Per esempio, un metodo LC-MS/MS con $s = 0{,}12$ ng/mL a 2,0 ng/mL e $s = 4{,}0$ ng/mL a 200 ng/mL ha una s oltre trenta volte maggiore al livello alto, ma un CV del 6,0% al livello basso e del 2,0% a quello alto: in termini relativi è più preciso al livello alto. Per questo i criteri di precisione delle linee guida sono espressi come CV% (per esempio CV ≤ 15%, ≤ 20% al LLOQ in ICH M10, Capitolo 12).

4.4 Esercizio completo: Fe(III) in un campione d'acqua

Esempio 4.2 — Statistica descrittiva di sette replicati

La concentrazione di Fe(III) in un'acqua è determinata con sette replicati (mg/L): 2,41; 2,38; 2,45; 2,43; 2,39; 2,44; 2,42.

1. Media (equazione 4.1): $\bar{x} = 16{,}92/7 = 2{,}41714$ mg/L; nei calcoli si usa 2,4171, come risultato si riporta 2,417 mg/L.
2. Mediana: dati ordinati 2,38; 2,39; 2,41; 2,42; 2,43; 2,44; 2,45; 4° valore = 2,42 mg/L.
3. Scarti e quadrati (Tabella 4.1): somma dei quadrati = 0,00394287 (mg/L)².
4. Varianza ($\nu = 6$): $s^2 = 0{,}00394287/6 = 0{,}00065715$ (mg/L)².
5. Deviazione standard (equazione 4.3): $s = \sqrt{0{,}00065715} = 0{,}02563$ mg/L ≈ 0,026 mg/L.
6. CV% (equazione 4.5): $0{,}02563/2{,}4171 \times 100 = 1{,}06\% \approx 1{,}1\%$ (la RSD dell'equazione 4.4 vale 0,0106).

Risultato: media 2,417 mg/L, mediana 2,42 mg/L, s = 0,026 mg/L (n = 7), CV = 1,1%.

Tabella 4.1 — Scarti dalla media per i sette replicati di Fe(III) (media = 2,4171 mg/L)

Replicato	x_i (mg/L)	$x_i - \bar{x}$ (mg/L)	$(x_i - \bar{x})^2$ (mg/L)²
1	2,41	−0,0071	0,00005041
2	2,38	−0,0371	0,00137641
3	2,45	+0,0329	0,00108241
4	2,43	+0,0129	0,00016641
5	2,39	−0,0271	0,00073441
6	2,44	+0,0229	0,00052441
7	2,42	+0,0029	0,00000841
Somma	16,92	≈ 0	**0,00394287**

Dividere per n invece che per n − 1

Dividendo la somma dei quadrati per $n = 7$ si otterrebbe $\sqrt{0{,}00394287/7} = 0{,}024$ mg/L e un CV di circa 1,0%: è la formula della popolazione (4.2), non applicabile a sette replicati. Altri errori tipici: usare $s/\sqrt{n}$ al posto di s nel CV; dimenticare il fattore 100 (RSD = 0,011 corrisponde a CV = 1,1%); arrotondare i risultati intermedi (media, scarti), che va fatto solo sul risultato finale.

La deviazione standard si riporta di norma con due cifre significative e sempre insieme a n: senza il numero di dati non si possono calcolare l'errore standard della media ($0{,}0256/\sqrt{7} = 0{,}0097$ mg/L), l'intervallo di confidenza (Capitolo 5) né il contributo di ripetibilità all'incertezza (Capitolo 6).

Punti chiave del Capitolo 4

- **Popolazione**: tutte le misure possibili, parametri μ e σ ignoti, curva continua; **campione**: n misure eseguite, statistiche $\bar{x}$ e s che li stimano, istogramma.
- La **media** usa tutti i dati ed è sensibile ai valori anomali; la **mediana** (valore centrale dei dati ordinati; per n pari, media dei due centrali) è robusta; la **moda** è il valore più frequente.
- La **varianza** ha l'unità al quadrato e si somma per contributi indipendenti; la **deviazione standard** ha la stessa unità dei dati.
- Nel campione si divide per **n − 1** (gradi di libertà), perché $\bar{x}$ è stimata dagli stessi dati: dividere per n sottostima σ.
- **RSD** $= s/\bar{x}$ (adimensionale) e **CV%** $= s/\bar{x} \times 100$ confrontano la precisione tra livelli e metodi diversi.
- Fe(III): $\bar{x} = 2{,}417$ mg/L, $s = 0{,}026$ mg/L, CV = 1,1%.

Mettiti alla prova

Per esercitarti su questo capitolo usa le domande del quiz sugli argomenti:

- **Campione, popolazione e tendenza centrale**
- **Dispersione: deviazione standard, varianza e CV%**

5. Distribuzioni, inferenza e valori anomali

In laboratorio interessa sapere quale frazione dei risultati cadrà entro certi limiti, quanto la media del campione può distare da quella della popolazione e se un replicato molto diverso dagli altri vada tenuto. Per rispondere servono la distribuzione normale e gli strumenti dell'**inferenza statistica**, che portano dal campione alla popolazione con un livello di fiducia dichiarato.

5.1 La distribuzione normale

Gli errori casuali di una misura sono la somma di molti piccoli contributi indipendenti (fluttuazioni di temperatura, rumore del rivelatore, piccole differenze di volume); per il **teorema del limite centrale** la loro somma tende a distribuirsi secondo la **distribuzione normale** o **gaussiana**:

$$f(y) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(y-\mu)^2}{2\sigma^2}\right] \tag{5.1}$$

dove y è il valore misurato ($-\infty < y < +\infty$), μ la media e σ la deviazione standard della popolazione. Nell'equazione 5.1, μ fissa la **posizione** del massimo, attorno al quale la curva è simmetrica (media = mediana = moda); σ fissa la **larghezza**: più σ è grande, più la curva è larga e bassa. Il fattore $1/(\sigma\sqrt{2\pi})$ rende l'**area totale sotto la curva uguale a 1**, cioè alla probabilità totale. La probabilità che un risultato cada tra due valori è l'area sotto la curva tra quei valori.

Le aree entro un dato numero di deviazioni standard dalla media sono sempre le stesse (Figura 5.1): entro $\mu \pm 1\sigma$ cade il **68,3%** dei risultati, entro $\mu \pm 2\sigma$ il **95,4%**, entro $\mu \pm 3\sigma$ il **99,7%**. La curva non si annulla mai: anche oltre $\pm 3\sigma$ resta lo 0,27% dei risultati. La parte esterna a un intervallo simmetrico si divide **in parti uguali tra le due code**: fuori da $\mu \pm 2\sigma$ c'è il 4,6%, cioè il 2,3% sopra $\mu + 2\sigma$ e il 2,3% sotto $\mu - 2\sigma$. Il 95% esatto corrisponde a $\mu \pm 1{,}96\sigma$, il 99% a $\mu \pm 2{,}576\sigma$. Spesso si esprime la distanza dalla media in unità di σ con la variabile standardizzata $z = (y - \mu)/\sigma$.

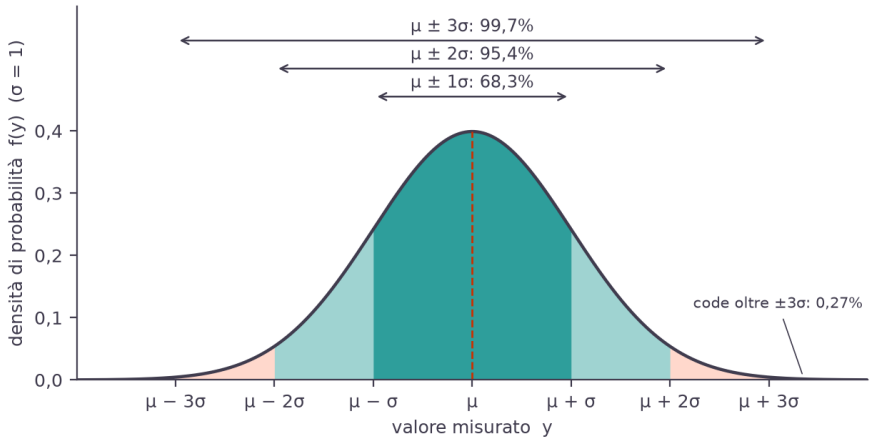


Figura 5.1 — Distribuzione normale: frazione dei risultati entro μ ± 1σ, μ ± 2σ e μ ± 3σ.

Esempio 5.1 — Percentuali della distribuzione normale

Un siero di controllo dà risultati distribuiti normalmente con $\mu = 5{,}50$ mmol/L e $\sigma = 0{,}10$ mmol/L. Che frazione dei risultati è attesa (a) tra 5,40 e 5,60 mmol/L, (b) sopra 5,70 mmol/L, (c) sotto 5,40 mmol/L?

1. (a) 5,40 e 5,60 mmol/L sono $\mu \pm 1\sigma$: **68,3%**.
2. (b) $z = (5{,}70 - 5{,}50)/0{,}10 = 2$. Fuori da $\mu \pm 2\sigma$ c'è il 4,6%, diviso tra due code: sopra $\mu + 2\sigma$ c'è il **2,3%**.
3. (c) Fuori da $\mu \pm 1\sigma$ c'è il 31,7%; sotto $\mu - 1\sigma$ ne cade la metà, il **15,9%**.

Risultato: 68,3%; 2,3%; 15,9%. L'errore tipico è dimenticare di dividere tra le due code (4,6% invece di 2,3%).

5.2 Distribuzioni simmetriche e asimmetriche

Quando i dati non sono gaussiani, la posizione relativa di media, mediana e moda ne rivela la forma (Figura 5.2):

- **distribuzione simmetrica**: media = mediana = moda; è tipica delle misure affette solo da errori casuali;
- **asimmetria positiva** (coda verso destra): pochi valori molto alti spostano la media più della mediana, mentre la moda resta sul picco: **media > mediana > moda**;
- **asimmetria negativa** (coda verso sinistra): **moda > mediana > media**.

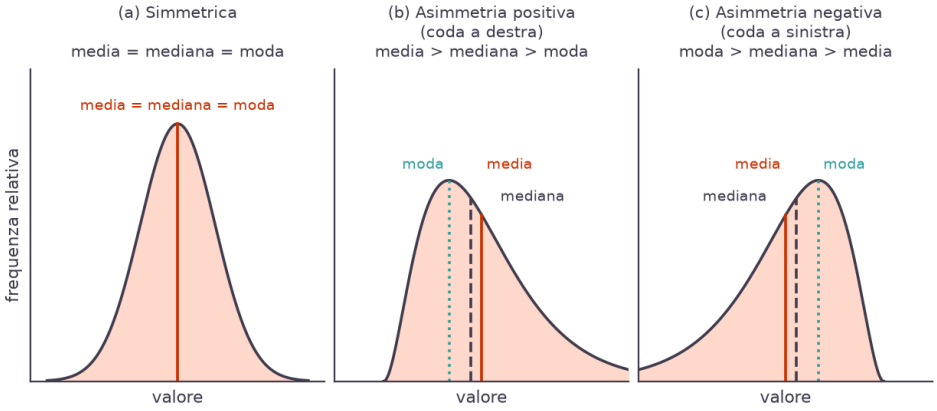


Figura 5.2 — Posizione relativa di media, mediana e moda in una distribuzione simmetrica (a), con asimmetria positiva (b) e negativa (c).

Hanno tipicamente asimmetria positiva i contaminanti in tracce in una serie di lotti (valori non negativi, vicini allo zero): nove lotti di caffè con ocratossina A pari a 0,8; 0,9; 0,9; 0,9; 1,0; 1,1; 1,3; 2,2; 3,5 µg/kg hanno

media 1,40, mediana 1,0 e moda 0,9 µg/kg. Hanno invece asimmetria negativa le rese di estrazione (quasi tutte tra il 95% e il 98%, alcune molto più basse per perdite). L'ordine media–mediana–moda è una **regola empirica** per distribuzioni con un solo picco, non una proprietà generale. Con dati asimmetrici la mediana descrive meglio il valore tipico, e gli strumenti basati sulla normalità vanno usati con cautela.

5.3 L'errore standard della media

Ripetendo più volte una serie di n replicati si ottengono medie diverse: la **media è essa stessa una variabile casuale**, distribuita attorno a μ in modo più stretto delle singole misure, perché nella media gli errori casuali si compensano in parte. La sua deviazione standard è l'**errore standard della media**:

$$s_{\bar{x}} = \frac{s}{\sqrt{n}} \tag{5.2}$$

($\sigma/\sqrt{n}$ se σ è nota). La deviazione standard s descrive la dispersione delle **singole misure** e non diminuisce aumentando n; l'errore standard descrive l'incertezza della **media** e diminuisce con $\sqrt{n}$: più misure, stima più affidabile. Per **dimezzarlo** bisogna **quadruplicare** i replicati (con $s = 0{,}40$ mg/L: 0,20 mg/L con 4 replicati, 0,10 mg/L con 16). Aumentare n non riduce invece gli errori sistematici: $\bar{x}$ converge verso μ, non verso il valore vero.

5.4 La distribuzione t di Student

Se σ fosse nota, la media cadrebbe entro $\mu \pm 1{,}96\,\sigma/\sqrt{n}$ nel 95% dei casi. In pratica σ è ignota e la sua stima s, con pochi dati, è a sua volta incerta: si usa allora la **distribuzione t di Student** con $\nu = n - 1$ **gradi di libertà**. Ha **code più ampie** della normale, quindi a parità di fiducia $t > z$: per $\nu = 3$ e il 95%, $t = 3{,}182$ contro $z = 1{,}96$. Al crescere di ν, t tende a z (per $\nu = 30$, $t = 2{,}042$). La soglia "$n < 30$" spesso citata è solo una regola pratica: la t si usa ogni volta che σ è stimata da s, sempre con dati circa normali.

Il valore di t dipende dai gradi di libertà e dal livello di fiducia, e le tavole distinguono valori **bilaterali** e **unilaterali**. Per un intervallo simmetrico la probabilità di errore (5% al 95%) si divide tra le due code: serve il **valore bilaterale** ($t_{\alpha/2,\nu}$). Il valore unilaterale al 95% coincide con il bilaterale al 90% ed è più piccolo (per $\nu = 6$: 1,943 invece di 2,447).

Tabella 5.1 — Valori critici bilaterali della t di Student (tavole complete nell'Appendice B)

ν	90%	95%	99%
2	2,920	4,303	9,925
3	2,353	3,182	5,841
4	2,132	2,776	4,604
5	2,015	2,571	4,032
6	1,943	2,447	3,707
7	1,895	2,365	3,499
8	1,860	2,306	3,355
9	1,833	2,262	3,250
10	1,812	2,228	3,169
20	1,725	2,086	2,845
∞ (= z)	1,645	1,960	2,576

5.5 L'intervallo di confidenza della media

Intervallo di confidenza della media

$$\mu = \bar{x} \pm t \cdot \frac{s}{\sqrt{n}} \tag{5.3}$$

- $\bar{x}$, s: media e deviazione standard degli n replicati; $s/\sqrt{n}$: errore standard (equazione 5.2)
- t: valore **bilaterale** della t di Student al livello di fiducia scelto (di solito 95%) con $\nu = n - 1$
- Con σ nota si usa $z \cdot \sigma/\sqrt{n}$ ($z = 1{,}96$ al 95%).

L'intervallo al 95% è costruito in modo che, ripetendo l'esperimento molte volte, il **95% degli intervalli contenga la media della popolazione** μ. Non descrive dove cadranno le singole misure (la cui dispersione è s, non $s/\sqrt{n}$) e contiene il valore vero solo in assenza di errori sistematici, perché tiene conto solo degli errori casuali.

L'intervallo **si allarga** se si aumenta il livello di fiducia (per $\nu = 5$, t passa da 2,571 al 95% a 4,032 al 99%: più fiducia richiede un intervallo più ampio) e **si restringe** all'aumentare di n, per due motivi: cresce $\sqrt{n}$ e diminuisce t. Con $s = 0{,}30$ mg/L la semiampiezza al 95% vale 0,75 mg/L con 3 replicati ($t = 4{,}303$), 0,31 mg/L con 6 ($t = 2{,}571$) e 0,19 mg/L con 12 ($t = 2{,}201$).

Esempio 5.2 — Intervallo di confidenza della media

La 25-idrossivitamina D in un siero è determinata mediante LC-MS/MS con sei replicati (ng/mL): 31,2; 30,4; 32,0; 31,5; 30,8; 31,7.

1. Media: $\bar{x} = 187{,}6/6 = 31{,}267$ ng/mL.
2. Somma dei quadrati degli scarti: 1,7533 (ng/mL)²; $s = \sqrt{1{,}7533/5} = 0{,}592$ ng/mL.
3. Errore standard: $0{,}592/\sqrt{6} = 0{,}2418$ ng/mL; gradi di libertà $\nu = 5$.
4. Al 95% (equazione 5.3): $t = 2{,}571$; semiampiezza $2{,}571 \times 0{,}2418 = 0{,}62$ ng/mL.
5. Al 99%: $t = 4{,}032$; semiampiezza $4{,}032 \times 0{,}2418 = 0{,}97$ ng/mL.

Risultato: μ = (31,27 ± 0,62) ng/mL al 95% e (31,27 ± 0,97) ng/mL al 99% (n = 6).

Errori tipici nell'intervallo di confidenza

Con i dati dell'Esempio 5.2 (valore corretto al 95%: ± 0,62 ng/mL): usare $\nu = n = 6$ dà ± 0,59; usare il t unilaterale (2,015) dà ± 0,49; usare $z = 1{,}96$ dà ± 0,47; moltiplicare t per s invece che per $s/\sqrt{n}$ dà ± 1,52 ng/mL.

5.6 I valori anomali

Un **valore anomalo** (outlier) è un risultato che sembra incompatibile con gli altri della serie. Può derivare da un errore grossolano (Capitolo 2) o essere un valore legittimo nella coda della distribuzione: in 20 misure normali la probabilità che almeno una cada oltre $\pm 2\sigma$ è di circa il 61%. Servono criteri oggettivi. I test seguenti presuppongono repliche della stessa grandezza, distribuite normalmente, ed esaminano **un solo valore sospetto**, il più lontano dagli altri; se la statistica supera il **valore critico** tabulato (funzione di n e del livello di fiducia), il valore è **statisticamente anomalo** a quel livello.

Test Q di Dixon e test di Grubbs

$$Q = \frac{|x_{\text{sosp}} - x_{\text{vicino}}|}{x_{\max} - x_{\min}} \tag{5.4}$$

$$G = \frac{|x_{\text{sosp}} - \bar{x}|}{s} \tag{5.5}$$

- **Q di Dixon**, per campioni piccoli ($3 \leq n \leq 10$): dati ordinati; x_{vicino} è il valore adiacente al sospetto; al denominatore il range di **tutti** i dati.
- **Grubbs**: distanza dalla media in unità di deviazione standard; $\bar{x}$ e s si calcolano su tutti i dati, **compreso** il sospetto. Usa tutti i dati ed è più potente del test Q per $n > 10$.
- Il valore è anomalo se $Q > Q_{\text{critico}}$ o $G > G_{\text{critico}}$ (Tabella 5.2), letti per il numero **totale** di dati n.

Tabella 5.2 — Valori critici al 95% e al 99% del test Q di Dixon (Rorabacher, 1991) e del test di Grubbs bilaterale

n	Q 95%	Q 99%	G 95%	G 99%
3	0,970	0,994	1,154	1,155
4	0,829	0,926	1,481	1,496
5	0,710	0,821	1,715	1,764
6	0,625	0,740	1,887	1,973
7	0,568	0,680	2,020	2,139
8	0,526	0,634	2,127	2,274
9	0,493	0,598	2,215	2,387
10	0,466	0,568	2,290	2,482
12			2,412	2,636
15			2,548	2,806
20			2,708	3,001

Per Grubbs si usa il test **bilaterale** quando il sospetto, individuato guardando i dati, può essere il minimo o il massimo (il caso usuale).

Esempio 5.3 — Test Q di Dixon e test di Grubbs

(a) Cinque determinazioni di zinco in un siero di controllo (µmol/L): 15,2; 15,6; 15,4; 16,9; 15,3. Il valore 16,9 è anomalo?

1. Dati ordinati: 15,2; 15,3; 15,4; 15,6; 16,9; vicino del sospetto: 15,6; range: $16{,}9 - 15{,}2 = 1{,}7$.
2. Equazione 5.4: $Q = (16{,}9 - 15{,}6)/1{,}7 = 0{,}765$.
3. Per $n = 5$: $0{,}765 > 0{,}710$ (anomalo al 95%) ma $0{,}765 < 0{,}821$ (non anomalo al 99%).

(b) Dodici determinazioni di una soluzione di controllo di paracetamolo (mg/L): 98,6; 99,1; 98,9; 99,4; 98,8; 99,0; 101,3; 99,2; 98,7; 99,3; 99,0; 98,9. Il valore 101,3 è anomalo?

1. Su tutti i 12 valori: $\bar{x} = 1190{,}2/12 = 99{,}183$ mg/L; $s = 0{,}707$ mg/L.
2. Equazione 5.5: $G = (101{,}3 - 99{,}183)/0{,}707 = 2{,}99$.
3. Per $n = 12$: $2{,}99 > 2{,}412$ (95%) e $> 2{,}636$ (99%).

Risultato: (a) Q = 0,765, valore anomalo al 95% ma non al 99%; (b) G = 2,99, valore anomalo al 95% e al 99%. In entrambi i casi, prima di escluderlo si cerca una causa sperimentale.

Errori frequenti nei test per valori anomali

- Test Q: dividere per il range dei soli valori restanti o usare la tavola con $n - 1$ invece di n.
- Grubbs: calcolare $\bar{x}$ e s **senza** il sospetto (nell'Esempio 5.3b si otterrebbe $G = 9{,}4$) o dividere per $s/\sqrt{n}$ ($G = 10{,}4$).
- Applicare il test ai punti di una curva di calibrazione: sono a concentrazioni diverse e non sono repliche; un calibratore aberrante si individua con i residui della regressione (Capitolo 8) e i criteri di accettazione della curva (Capitolo 16).

Il test fornisce un'**evidenza statistica**: dice che il valore è improbabile per la distribuzione degli altri dati, non perché sia diverso. La regola è: **un valore si elimina solo con un'evidenza statistica e, preferibilmente, una causa sperimentale identificata** (errore di trascrizione, contaminazione, diluizione sbagliata, iniezione anomala); **mai solo perché "sembra lontano"**. In pratica:

- test positivo: prima di decidere si **cerca la causa**; se la si trova, il valore si esclude documentando dato, test e causa; altrimenti si segue una procedura stabilita a priori (per esempio riportare il risultato con e senza il valore, o usare la mediana; negli studi interlaboratorio la ISO 5725-2 esclude di norma i valori anomali all'1%);
- test negativo: il valore si conserva;
- il test si applica **una volta**, al valore più estremo; sostituire il valore con la media degli altri o eliminare "per simmetria" l'estremo opposto distorce media e dispersione;
- un errore grossolano documentato (campione sbagliato) invalida il risultato indipendentemente da qualsiasi test.

Punti chiave del Capitolo 5

- Gaussiana: μ fissa la posizione, σ la larghezza, area totale 1; entro $\mu \pm 1\sigma$, 2σ, 3σ cadono il 68,3%, il 95,4% e il 99,7% dei risultati; la parte esterna si divide tra le due code.
- Simmetrica: media = mediana = moda; asimmetria positiva (coda a destra): media > mediana > moda; negativa (coda a sinistra): moda > mediana > media (regola empirica).
- Errore standard della media $s/\sqrt{n}$: per dimezzarlo servono quattro volte più replicati; non riduce gli errori sistematici.
- Con σ stimata da s si usa la t di Student con $\nu = n - 1$; $t > z$ e $t \to z$ al crescere di ν.
- IC della media: $\bar{x} \pm t \cdot s/\sqrt{n}$ con t bilaterale; contiene μ con la fiducia dichiarata; si allarga dal 95% al 99%, si restringe con n.
- Test Q di Dixon ($3 \leq n \leq 10$): scarto dal vicino / range; Grubbs: $|x_{\text{sosp}} - \bar{x}|/s$ con il sospetto incluso, più potente per $n > 10$.
- Un valore si elimina solo con evidenza statistica e, preferibilmente, una causa identificata.

Mettiti alla prova

Per esercitarti su questo capitolo usa le domande del quiz sugli argomenti:

- **Distribuzione normale e asimmetria**
- **Errore standard e intervallo di confidenza**
- **Test per valori anomali**

6. L'incertezza di misura

Un risultato senza incertezza non si può confrontare con un limite o con il risultato di un altro laboratorio. Se il piombo in un'acqua risulta 9,6 µg/L e il limite è 10 µg/L, con un'incertezza estesa di 0,3 µg/L l'acqua è conforme; con 0,8 µg/L l'intervallo (8,8–10,4 µg/L) comprende il limite e la conformità non si può affermare. Per questo la ISO/IEC 17025 (2017) richiede ai laboratori di valutare l'incertezza dei risultati.

6.1 Errore e incertezza

Errore, incertezza e grandezze collegate (VIM, GUM)

Errore di misura: differenza tra il valore misurato e il valore vero; poiché il valore vero non è noto, l'errore di un risultato è in generale **ignoto** (se lo conoscessimo, lo correggeremmo).

Incertezza di misura: parametro non negativo che caratterizza la **dispersione dei valori** ragionevolmente attribuibili al misurando; a differenza dell'errore, si **stima**.

Incertezza tipo u: incertezza espressa come deviazione standard. **Incertezza tipo composta** u_c: incertezza tipo del risultato, ottenuta combinando tutti i contributi. **Incertezza estesa** $U = k \cdot u_c$, con **fattore di copertura** k. **Incertezza relativa**: $u(x)/x$, spesso in percentuale.

Gli effetti sistematici noti (per esempio il contributo del bianco o un recupero documentato) si correggono prima di valutare l'incertezza; nel bilancio resta l'incertezza della correzione.

6.2 Il GUM: contributi di Tipo A e di Tipo B

Il **GUM** (*Guide to the Expression of Uncertainty in Measurement*, JCGM 100:2008) è la guida internazionale per la valutazione e l'espressione dell'incertezza di misura (da non confondere con il VIM, vocabolario della metrologia, con la ISO 5725 su esattezza e precisione o con ICH M10 sulla validazione bioanalitica). Distingue due categorie di contributi in base al **metodo di valutazione**:

- **Tipo A**: valutato mediante l'**analisi statistica di serie di osservazioni** ripetute (per esempio l'incertezza di ripetibilità);
- **Tipo B**: valutato con **metodi diversi dall'analisi statistica**: certificati di taratura, specifiche del costruttore, valori certificati, dati di letteratura, giudizio esperto.

La distinzione **non** coincide con quella tra errori casuali e sistematici, né indica quale contributo sia più grande o più affidabile: entrambi i tipi si esprimono come deviazioni standard e si combinano allo stesso modo.

La valutazione procede così: si scrive l'equazione con cui si calcola il risultato, si identificano le fonti di incertezza (Tabella 6.1), si esprime ciascuna come incertezza tipo (Tipo A o B), si combinano in u_c e si calcola $U = k \cdot u_c$.

Tabella 6.1 — Fonti di incertezza tipiche di una determinazione analitica

Fonte	Tipo	Origine dell'informazione
Ripetibilità	A	serie di replicati: $u_A = s/\sqrt{n}$
Taratura strumentale	B	certificato di taratura dello strumento
Pesata	B	taratura e risoluzione della bilancia
Volume finale	B	tolleranza della vetreria volumetrica, temperatura
Standard interno e soluzione di taratura	B	purezza e preparazione degli standard
Curva di calibrazione	A	dispersione dei punti attorno alla retta (Capitolo 7)

6.3 La valutazione di Tipo A

Da n osservazioni ripetute si calcola la deviazione standard s (Capitolo 4). Se il risultato riportato è la **media** degli n replicati, l'incertezza di ripetibilità è l'errore standard della media (Capitolo 5):

$$u_A = \frac{s}{\sqrt{n}} \tag{6.1}$$

con $\nu = n - 1$ gradi di libertà. Se invece si riporta una **singola determinazione**, il contributo di ripetibilità è s stessa, stimata preferibilmente da molti dati (validazione, controlli di qualità). Per esempio, con $s = 0{,}24$ mg/L e media di 4 replicati, $u_A = 0{,}24/2 = 0{,}12$ mg/L; per una singola determinazione il contributo sarebbe 0,24 mg/L.

6.4 La valutazione di Tipo B

Nel Tipo B l'informazione va convertita in incertezza tipo scegliendo una distribuzione coerente con ciò che si sa (Tabella 6.2, Figura 6.1).

Tabella 6.2 — Conversione delle informazioni di Tipo B in incertezza tipo

Informazione disponibile	Distribuzione	Incertezza tipo
certificato con U e fattore k	normale	$u = U/k$
certificato con U al 95%, senza k	normale	$u = U/1{,}96$
tolleranza $\pm a$, nessuna informazione sulla forma	rettangolare	$u = a/\sqrt{3}$
tolleranza $\pm a$, valori centrali più probabili	triangolare	$u = a/\sqrt{6}$
risoluzione digitale d	rettangolare (semiampiezza $d/2$)	$u = d/(2\sqrt{3})$

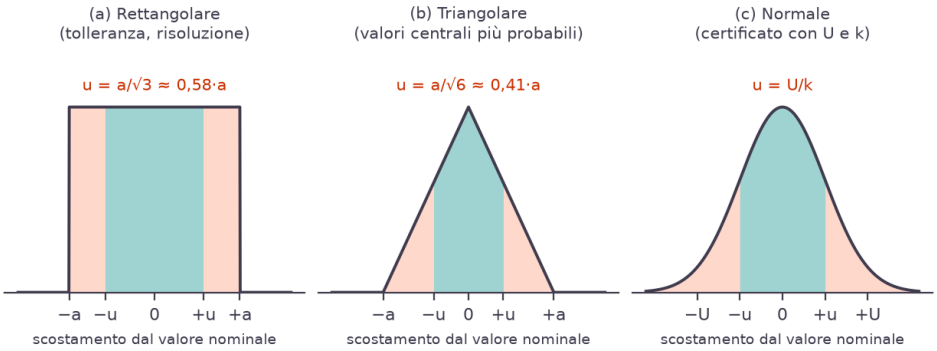


Figura 6.1 — Distribuzioni usate nelle valutazioni di Tipo B: rettangolare (a), triangolare (b) e normale (c); la zona più scura indica l'intervallo ±u.

Per esempio: il certificato di una bilancia con $U = 0{,}10$ mg ($k = 2$) dà $u = 0{,}050$ mg; un matraccio di classe A da 100 mL con tolleranza $\pm 0{,}10$ mL dà $u = 0{,}10/\sqrt{3} = 0{,}058$ mL con la distribuzione rettangolare, che è la scelta prudente quando non si sa altro (0,041 mL con la triangolare); uno standard con purezza $(99{,}5 \pm 0{,}3)\%$ dà $u = 0{,}3/\sqrt{3} = 0{,}17\%$.

6.5 L'incertezza tipo composta

Contributi **indipendenti** si combinano **in quadratura**, perché per grandezze indipendenti **si sommano le varianze** (u^2), non le deviazioni standard: i contributi non agiscono tutti insieme al massimo e nello stesso verso, e in parte si compensano. La somma lineare delle incertezze presupporrebbe proprio questo caso peggiore e **sovrastimerebbe** u_c.

Incertezza tipo composta

Somme e differenze ($y = x_1 + x_2 - x_3$), contributi espressi nell'unità del risultato:

$$u_c = \sqrt{u_A^2 + u_{B1}^2 + u_{B2}^2 + \ldots} \tag{6.2}$$

Prodotti e quozienti ($y = x_1 \cdot x_2 / x_3$), incertezze relative:

$$\frac{u_c(y)}{y} = \sqrt{\left(\frac{u(x_1)}{x_1}\right)^2 + \left(\frac{u(x_2)}{x_2}\right)^2 + \left(\frac{u(x_3)}{x_3}\right)^2} \tag{6.3}$$

- Entrambe valgono solo per contributi **indipendenti** (non correlati).
- La 6.2 richiede contributi espressi nell'unità del risultato e che vi si trasferiscono "uno a uno"; la 6.3 vale per la maggior parte dei risultati analitici (massa × purezza / volume; diluizioni in serie, Capitolo 16).

Poiché si sommano i quadrati, i contributi grandi pesano molto più di quelli piccoli: un contributo pari a un terzo del maggiore fa crescere u_c solo del 5% circa. Per ridurre l'incertezza bisogna quindi agire sul **contributo dominante**.

6.6 L'incertezza estesa

Incertezza estesa

$$U = k \cdot u_c \tag{6.4}$$

- Il risultato si esprime come $y \pm U$.
- $k = 1$: livello di fiducia di circa il 68%; $k = 2$: circa il 95%; $k = 3$: circa il 99,7%.
- Le corrispondenze valgono se la distribuzione del risultato è **circa normale** e i **gradi di libertà effettivi** sono sufficienti; con un contributo di Tipo A dominante stimato da pochi replicati si usa come k il t bilaterale al 95%.

Esempio 6.1 — Incertezza composta ed estesa con contributi nella stessa unità

Il rame in un'acqua è determinato con cinque replicati (mg/L): 0,842; 0,851; 0,838; 0,847; 0,852. Il risultato è la media; i contributi di Tipo B (purezza e preparazione degli standard, vetreria), già combinati, valgono $u_B = 0{,}0050$ mg/L.

1. Media e deviazione standard: $\bar{x} = 0{,}846$ mg/L; $s = \sqrt{142 \times 10^{-6}/4} = 0{,}0060$ mg/L.
2. Tipo A (equazione 6.1): $u_A = 0{,}00596/\sqrt{5} = 0{,}00266$ mg/L.
3. Composta (equazione 6.2): $u_c = \sqrt{0{,}00266^2 + 0{,}0050^2} = \sqrt{(7{,}1 + 25{,}0) \times 10^{-6}} = 0{,}00567$ mg/L.
4. Estesa (equazione 6.4): $U = 2 \times 0{,}00567 = 0{,}011$ mg/L.

Risultato: Cu = (0,846 ± 0,011) mg/L (k = 2, livello di fiducia di circa il 95%).

Errori tipici nel calcolo dell'incertezza

Con i dati dell'Esempio 6.1: usare s al posto di $s/\sqrt{n}$ dà $U = 0{,}016$ mg/L; sommare linearmente i contributi dà 0,015 mg/L; dimenticare k dà 0,0057 mg/L. Altri errori: omettere la radice quadrata; combinare incertezze assolute in un modello fatto di prodotti e quozienti.

6.7 Il bilancio di incertezza

Il **bilancio di incertezza** (uncertainty budget) è la tabella che elenca, per ogni fonte, tipo, incertezza e peso su u_c^2: documenta la valutazione e mostra **dove intervenire**.

Esempio 6.2 — Bilancio di incertezza per il dosaggio di un principio attivo

Il paracetamolo in compresse è determinato per HPLC-UV: una porzione di polvere (125,0 mg) è estratta e portata a volume in un matraccio da 100 mL; la calibrazione usa uno standard di purezza $(99{,}5 \pm 0{,}3)\%$. Il risultato, media di sei preparazioni indipendenti, è 498,6 mg per compressa. Il modello è un prodotto/quoziente: si combinano le incertezze relative (Tabella 6.3).

1. Ripetibilità (A): CV delle sei preparazioni 0,90%; per la media, $0{,}90/\sqrt{6} = 0{,}367\%$.
2. Curva di calibrazione (A): 0,30% dalla regressione (Capitolo 7).
3. Soluzione standard (B): 0,20%, contributo dominato dalla purezza (0,17%).
4. Volume finale (B): tolleranza (0,058 mL) e temperatura (0,048 mL), $\sqrt{0{,}058^2 + 0{,}048^2} = 0{,}075$ mL su 100 mL, cioè 0,075%.
5. Pesata (B): certificato della bilancia, $0{,}050/125{,}0 = 0{,}040\%$.
6. Composta (equazione 6.3): $\sqrt{0{,}367^2 + 0{,}30^2 + 0{,}20^2 + 0{,}075^2 + 0{,}040^2} = \sqrt{0{,}272} = 0{,}52\%$.
7. Estesa: $U = 2 \times 0{,}52\% = 1{,}04\%$, cioè $0{,}0104 \times 498{,}6 = 5{,}2$ mg.

Risultato: (498,6 ± 5,2) mg per compressa (k = 2, livello di fiducia di circa il 95%).

Tabella 6.3 — Bilancio di incertezza dell'Esempio 6.2

Fonte	Tipo	u relativa (%)	Peso su u_c^2
Ripetibilità (media di 6)	A	0,367	49,6%
Curva di calibrazione	A	0,30	33,1%
Soluzione standard	B	0,20	14,7%
Volume finale	B	0,075	2,1%
Pesata	B	0,040	0,6%
Composta / estesa ($k = 2$)		**0,52 / 1,04**	

Ripetibilità e calibrazione spiegano oltre l'80% della varianza; eliminando volume e pesata U passerebbe da 1,04% a 1,03%. Conviene quindi aumentare le preparazioni (con 12, U scende a 0,90%), non comprare vetreria più accurata; la somma lineare darebbe 1,96%, quasi il doppio. Per volume e pesata si sono incluse solo taratura, tolleranza e temperatura: la loro parte casuale è già nella ripetibilità di preparazioni indipendenti e **non va contata due volte**.

6.8 Come si riporta il risultato

Il risultato si scrive $y \pm U$ con unità, valore di k e livello di fiducia approssimato, per esempio "(498,6 ± 5,2) mg, $k = 2$, circa 95%", oppure "498,6 mg ± 1,0% ($k = 2$)". Come visto nel Capitolo 1, l'incertezza si arrotonda a una o due cifre significative e il risultato alla stessa cifra decimale. Il valore di k va sempre dichiarato,

altrimenti non si risale a u_c. L'incertezza non è una caratteristica di validazione per ICH Q2(R2) (2023) o ICH M10 (2022), ma un requisito della ISO/IEC 17025; i dati di validazione (precisione intermedia, esattezza) sono la base naturale per stimarla (Capitolo 12).

Punti chiave del Capitolo 6

- **Errore** (differenza dal valore vero, ignota) e **incertezza** (dispersione dei valori attribuibili al misurando, stimabile) sono concetti distinti.
- **GUM** (JCGM 100:2008): **Tipo A** = analisi statistica di osservazioni ripetute; **Tipo B** = altre fonti (certificati, specifiche, letteratura, giudizio esperto). La distinzione riguarda il metodo di valutazione, non la natura dell'errore.
- Ripetibilità: $u_A = s/\sqrt{n}$ se si riporta la media di n replicati; s per una singola determinazione.
- Tipo B: U/k dai certificati; $a/\sqrt{3}$ (rettangolare) o $a/\sqrt{6}$ (triangolare) dalle tolleranze.
- Contributi indipendenti si combinano **in quadratura** (si sommano le varianze): incertezze assolute per somme e differenze, relative per prodotti e quozienti; la somma lineare sovrastima.
- $U = k \cdot u_c$; $k = 2$ corrisponde a circa il 95% per distribuzioni circa normali con sufficienti gradi di libertà.
- Il bilancio di incertezza individua i contributi dominanti, su cui conviene intervenire.

Mettiti alla prova

Per esercitarti su questo capitolo usa le domande del quiz sugli argomenti:

- **Incertezza di misura (GUM)**
- **Propagazione dell'incertezza nelle diluizioni** (ripresa nel Capitolo 16)

7. Regressione lineare e curva di calibrazione

Le tecniche strumentali (spettrofotometria, HPLC, LC-MS/MS, ICP-MS) misurano un segnale, non una concentrazione. La curva di calibrazione traduce il segnale in concentrazione e la regressione lineare è il metodo statistico con cui la si costruisce; il Capitolo 8 ne verifica l'adeguatezza.

7.1 Calibrazione e regressione

Misurando il segnale di una serie di **standard** a concentrazione nota (Capitolo 9) si ottengono n coppie (x_i, y_i), da cui la **regressione** stima un'equazione che prevede una variabile a partire dall'altra. I ruoli sono fissati: la **variabile indipendente** X è la concentrazione degli standard, scelta dall'analista; la **variabile dipendente** Y è il segnale strumentale, affetto dall'errore casuale. I minimi quadrati attribuiscono infatti tutto l'errore a Y: le concentrazioni degli standard, preparate per pesata e diluizione con vetreria tarata, sono molto meno incerte del segnale. Solo dopo, con la **predizione inversa** (paragrafo 7.4), si ricava la concentrazione dal segnale.

Si parla di **regressione semplice** quando c'è una sola variabile esplicativa X (la curva di calibrazione) e di **regressione multipla** quando più variabili X_1, X_2, …, X_p determinano Y. La **correlazione** è un'altra cosa: misura soltanto il grado di associazione tra due variabili quantitative, con un indice (r, paragrafo 7.6), senza costruire un'equazione per la previsione. Nessuna delle due dimostra un rapporto di causa-effetto: in calibrazione lo garantisce la chimica della misura (per esempio la legge di Lambert-Beer). La relazione è lineare solo fino al limite di linearità (LOL), oltre il quale il rivelatore si satura (Capitolo 11).

Nelle slide e nel quiz le variabili sono scritte in maiuscolo (X_i, Y_i, $\hat{Y}_i$); qui usiamo le minuscole (x_i, y_i, $\hat{y}_i$), con lo stesso significato.

7.2 Il modello di regressione lineare semplice

Ogni osservazione è la somma di una retta e di un errore casuale:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \tag{7.1}$$

Parametri del modello

- β_0 — **intercetta**: valore atteso di Y per X = 0; in calibrazione è il segnale del bianco.
- β_1 — **pendenza**: variazione di Y per ogni aumento unitario di X; in calibrazione è la **sensibilità analitica**, in unità di segnale su unità di concentrazione (per esempio L/mg per un'assorbanza in funzione di mg/L). Il suo segno indica se la relazione è diretta ($\beta_1 > 0$) o inversa ($\beta_1 < 0$).
- ε_i — **errore casuale** dell'i-esima osservazione, non osservabile.

Le lettere greche indicano parametri della popolazione, ignoti, di cui i dati forniscono le **stime** b_0 e b_1. Allo stesso modo il **residuo** $e_i = y_i - \hat{y}_i$, calcolato dai dati, è la stima dell'errore ε_i, che non è osservabile; poiché b_0 e b_1 non coincidono con β_0 e β_1, in generale $e_i \neq \varepsilon_i$. La sensibilità non va confusa con il limite di rilevabilità, che dipende anche dal rumore del bianco (paragrafo 8.4). Poiché il bianco dà quasi sempre un piccolo segnale, l'intercetta si stima dai dati e di norma non si forza a zero (Capitolo 10).

Le stime sono affidabili se la relazione è lineare nell'intervallo, gli errori sono indipendenti, a media nulla, normali e con varianza costante (**omoschedasticità**) e l'errore su X è trascurabile; linearità e varianza costante si verificano con i residui (Capitolo 8).

7.3 Il metodo dei minimi quadrati

Il **metodo dei minimi quadrati** (*ordinary least squares*, OLS) sceglie, tra tutte le rette possibili, quella che rende minima la somma dei quadrati dei residui, cioè delle distanze **verticali** tra i punti e la retta (Figura 7.1a): $\mathrm{SQE} = \sum(y_i - \hat{y}_i)^2 \rightarrow$ minimo. La semplice somma dei residui non serve: residui positivi e negativi si compensano e la somma vale zero per qualunque retta passante per il punto medio dei dati.

Stime dei minimi quadrati

$$b_1 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})\,(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = \frac{S_{xy}}{S_{xx}} \tag{7.2}$$

$$b_0 = \bar{y} - b_1 \bar{x} \tag{7.3}$$

$$\hat{y}_i = b_0 + b_1 x_i \tag{7.4}$$

- $\bar{x}$, $\bar{y}$: medie delle concentrazioni e dei segnali degli standard; n: numero di standard.
- S_{xy}: codevianza (somma dei prodotti degli scarti dalle medie); S_{xx}: devianza di X (somma dei quadrati degli scarti).
- $\hat{y}_i$: valore previsto dalla retta per x_i.

Tre proprietà sono utili.

- **Il segno di b_1 è quello del numeratore**: S_{xx} è una somma di quadrati, sempre positiva, quindi $b_1 > 0$ se Y cresce con X e $b_1 < 0$ se diminuisce.
- **La retta passa sempre per il baricentro ($\bar{x}$, $\bar{y}$)**, come si vede riscrivendo l'equazione 7.3 come $\bar{y} = b_0 + b_1\bar{x}$: il valore previsto per $x = \bar{x}$ è $\bar{y}$, senza calcolare l'intercetta.
- **La somma dei residui è zero**: è un controllo dei calcoli, non una prova che il modello sia adeguato.

Errori frequenti sono calcolare la pendenza come $\sum x_i y_i / \sum x_i^2$ (retta forzata per l'origine) o come $\bar{y}/\bar{x}$, invertire numeratore e denominatore, oppure scrivere $b_0 = \bar{y} + b_1\bar{x}$.

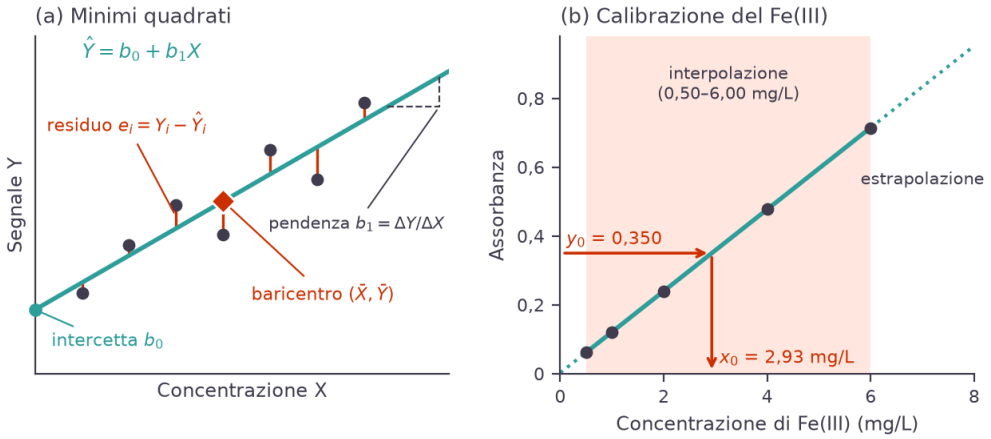


Figura 7.1 — (a) La retta dei minimi quadrati rende minima la somma dei quadrati dei residui (segmenti verticali rossi), passa per il baricentro dei dati e interseca l'asse Y nell'intercetta. (b) Calibrazione del Fe(III) dell'Esempio 7.3: tra lo standard più basso e il più alto (zona colorata) si lavora per interpolazione, oltre 6,00 mg/L si estrapolerebbe; le frecce mostrano la predizione inversa per un'assorbanza di 0,350.

7.4 Interpolazione, estrapolazione e predizione inversa

Si ha **interpolazione** quando la previsione riguarda un valore di X **interno** all'intervallo usato per stimare la retta (tra lo standard più basso e il più alto): è affidabile e raccomandata, perché lì la linearità è stata verificata. Si ha **estrapolazione** quando X è **esterno** all'intervallo: è rischiosa, perché oltre il range la linearità può non valere (saturazione in alto, rumore in basso), e nemmeno un r^2 vicinissimo a 1 lo garantisce. In unità di segnale l'intervallo interpolabile va da $\hat{y}(x_{min}) = b_0 + b_1 x_{min}$ a $\hat{y}(x_{max}) = b_0 + b_1 x_{max}$, intercetta compresa. Un campione con segnale troppo alto si diluisce fino a rientrare nel range, moltiplicando poi per il fattore di diluizione (Capitolo 14); un segnale sotto lo standard più basso si riporta come "< LLOQ" (Capitolo 11).

In routine si conosce il segnale y_0 del campione e si ricava la concentrazione risolvendo $y_0 = b_0 + b_1 x_0$: è la **predizione inversa**, l'operazione fondamentale dell'analisi quantitativa (Figura 7.1b).

Predizione inversa

$$x_0 = \frac{y_0 - b_0}{b_1} \tag{7.5}$$

- y_0: segnale del campione (o media di letture replicate); si **sottrae** l'intercetta (segnale del bianco) e si **divide** per la pendenza, che converte il segnale netto in concentrazione.
- Se il campione è stato diluito, x_0 va moltiplicato per il fattore di diluizione.

Il controllo delle unità evita gli errori più comuni (sommare b_0, moltiplicare per b_1, invertire la frazione, usare l'equazione diretta $b_0 + b_1 y_0$): $(y_0 - b_0)$ è un segnale e b_1 è segnale su concentrazione, quindi il loro rapporto è una concentrazione.

Esempio 7.1 — Intervallo interpolabile e campione da diluire

Una calibrazione HPLC con standard da 1,0 a 50,0 ng/mL dà la retta Area = 850 + 1520 · X (X in ng/mL, area in unità arbitrarie, u.a.). Un campione dà 92 000 u.a.

Passo 1 — Limiti in segnale. $\hat{y}(1{,}0)$ = 850 + 1520 × 1,0 = 2370 u.a.; $\hat{y}(50{,}0)$ = 850 + 1520 × 50,0 = 76 850 u.a.

Passo 2 — Verifica. 92 000 u.a. è fuori dall'intervallo: (92 000 − 850)/1520 = 60,0 ng/mL sarebbe un'estrapolazione e non va riportato.

Passo 3 — Diluizione 1:5 e nuova misura. Area = 19 500 u.a., interna all'intervallo: x_0 = (19 500 − 850)/1520 = 12,27 ng/mL.

Risultato: 12,27 × 5 = 61,3 ng/mL. Il valore estrapolato era più basso del 2%: oltre lo standard più alto la risposta comincia a saturarsi.

7.5 Il diagramma di dispersione e la scelta del modello

Prima di calcolare una retta si guarda il **diagramma di dispersione** (Y in funzione di X): la forma della nuvola di punti suggerisce il modello. Le situazioni tipiche sono:

- **lineare diretta**: Y cresce in proporzione a X (quasi tutte le calibrazioni nel range lineare);
- **lineare inversa**: Y diminuisce linearmente, pendenza negativa (per esempio i saggi di decolorazione come il DPPH);
- **nessuna relazione**: punti casuali attorno a una retta orizzontale, pendenza e r prossimi a zero;
- **curvilinea crescente** (detta polinomiale diretta): Y cresce rapidamente e poi sempre più lentamente verso un plateau (saturazione del rivelatore); una retta forzata dà residui sistematici;
- **curvilinea a U**: Y diminuisce e poi aumenta (per esempio la curva di van Deemter in cromatografia); anche se la relazione è forte, r è vicino a zero;
- **curvilinea decrescente** (detta polinomiale inversa): Y diminuisce rapidamente e poi sempre più lentamente, come un analita che si degrada nel tempo.

Le relazioni curve si descrivono con polinomi, esponenziali o funzioni di saturazione; in calibrazione si preferisce la retta nell'intervallo lineare, altrimenti il modello più semplice che descrive adeguatamente i dati (Capitolo 10).

7.6 Scomposizione della variabilità, r^2 e r

7.6.1 SQT = SQR + SQE

Lo scarto del segnale di ogni punto dalla media si divide in una parte **spiegata** dalla retta e in un **residuo** (Figura 7.2): $y_i - \bar{y} = (\hat{y}_i - \bar{y}) + (y_i - \hat{y}_i)$. Elevando al quadrato e sommando su tutti i punti, per la retta dei minimi quadrati il termine misto si annulla:

Scomposizione della variabilità

$$\sum (y_i - \bar{y})^2 = \sum (\hat{y}_i - \bar{y})^2 + \sum (y_i - \hat{y}_i)^2$$

$$\text{SQT} = \text{SQR} + \text{SQE} \tag{7.6}$$

- **SQT** (somma dei quadrati totale): variabilità totale dei segnali attorno alla media; dipende solo dai dati.
- **SQR** (della regressione): variabilità **spiegata** dalla retta, dovuta alle diverse concentrazioni degli standard.
- **SQE** (degli errori): variabilità **non spiegata**, $\sum e_i^2$, dovuta all'errore casuale e a eventuali difetti del modello.

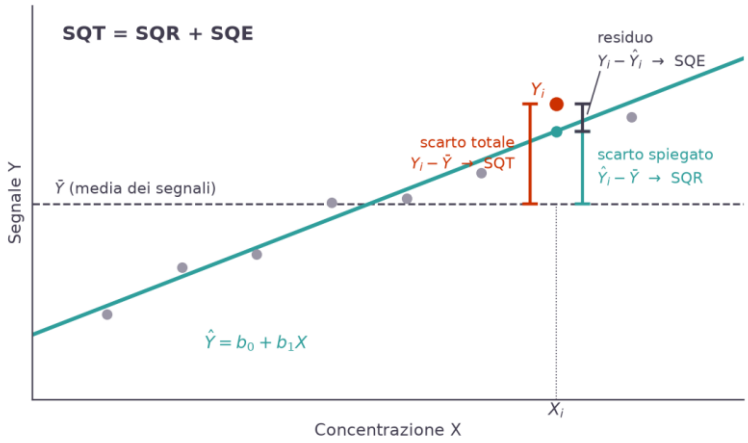


Figura 7.2 — Scomposizione dello scarto di un punto dalla media dei segnali in uno scarto spiegato dalla retta e in un residuo; sommando i quadrati su tutti i punti si ottiene SQT = SQR + SQE.

Se tutti i punti cadono sulla retta (con pendenza diversa da zero) i residui sono nulli: SQE = 0 e SQR = SQT. Se la retta è orizzontale (b_1 = 0) ogni $\hat{y}_i$ vale $\bar{y}$: SQR = 0 e SQE = SQT. Sono **somme** di quadrati, non varianze, e SQR si ottiene per differenza (SQT − SQE). La deviazione standard dei residui, $s_{y/x} = \sqrt{\text{SQE}/(n-2)}$, misura la dispersione dei punti attorno alla retta. In inglese: SST, SSR (*regression*) e SSE; alcuni testi però usano SSR per la somma dei quadrati dei residui (la nostra SQE).

7.6.2 Coefficiente di determinazione e coefficiente di correlazione

Coefficienti r² e r

$$r^2 = \frac{\text{SQR}}{\text{SQT}} = 1 - \frac{\text{SQE}}{\text{SQT}} \tag{7.7}$$

$$r = \frac{\sum(x_i - \bar{x})\,(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \cdot \sum(y_i - \bar{y})^2}} \tag{7.8}$$

- r^2: frazione della variabilità di Y spiegata dalla regressione su X, tra 0 e 1 (r^2 = 0,999: 99,9% spiegato, 0,1% non spiegato).
- r (Pearson): forza e direzione dell'associazione lineare, tra −1 e +1; ha lo stesso numeratore di b_1, quindi lo stesso segno della pendenza.
- Nella regressione semplice r^2 è il quadrato di r: noto r^2, $|r| = \sqrt{r^2}$ e il segno si ricava dalla pendenza.

Che cosa r² e r non dicono

- **Un r² elevato non dimostra la linearità**: una lieve curvatura sistematica può dare $r^2 > 0{,}99$ (Esempio 8.1). Bisogna sempre esaminare il grafico dei residui.
- **r² non misura la pendenza**: rette con sensibilità molto diverse possono avere lo stesso r^2.
- **r misura solo l'associazione lineare**: in una relazione a U simmetrica la codevianza è circa zero e r ≈ 0, anche se la relazione è forte.
- **Su intervalli ampi r² è dominato dagli standard più alti** e può superare 0,9999 anche con errori del 50% sugli standard più bassi (Esempio 8.2). Per questo ICH M10 (2022) giudica la curva sulle concentrazioni retrocalcolate, non su r^2 (Capitolo 10).

Esempio 7.2 — Scomposizione della variabilità, r² e r

Cinque standard (1–5 mg/L) danno assorbanze 0,12; 0,22; 0,35; 0,44; 0,57; la retta è $\hat{y}$ = 0,004 + 0,112 x.

Passo 1 — SQT. $\bar{y}$ = 0,340; scarti −0,22; −0,12; +0,01; +0,10; +0,23; SQT = 0,0484 + 0,0144 + 0,0001 + 0,0100 + 0,0529 = 0,1258.

Passo 2 — Residui e SQE. $\hat{y}$ = 0,116; 0,228; 0,340; 0,452; 0,564; residui +0,004; −0,008; +0,010; −0,012; +0,006 (lo standard a 4 mg/L sta sotto la retta). SQE = 0,000016 + 0,000064 + 0,000100 + 0,000144 + 0,000036 = 0,000360.

Passo 3 — SQR, r² e r. SQR = 0,1258 − 0,00036 = 0,12544; r^2 = 0,12544/0,1258 = 0,9971; r = +√0,9971 = +0,9986 (pendenza positiva).

Risultato: r² = 0,997: il 99,7% della variabilità è spiegato dalla retta. Con pendenza negativa e lo stesso r^2 sarebbe r = −0,9986.

7.7 Esempio completo: calibrazione spettrofotometrica del Fe(III)

Esempio 7.3 — Retta di calibrazione del Fe(III) e campione incognito

Cinque standard di Fe(III) danno le assorbanze di Tabella 7.1; un campione dà y_0 = 0,350.

Tabella 7.1 — Calibrazione del Fe(III): dati, calcoli intermedi, valori previsti e residui

x_i (mg/L)	y_i	$x_i - \bar{x}$	$y_i - \bar{y}$	$(x_i - \bar{x})(y_i - \bar{y})$	$(x_i - \bar{x})^2$	$\hat{y}_i$	e_i
0,50	0,062	−2,20	−0,2608	0,57376	4,84	0,06153	+0,00047
1,00	0,121	−1,70	−0,2018	0,34306	2,89	0,12091	+0,00009
2,00	0,238	−0,70	−0,0848	0,05936	0,49	0,23967	−0,00167
4,00	0,479	1,30	0,1562	0,20306	1,69	0,47719	+0,00181
6,00	0,714	3,30	0,3912	1,29096	10,89	0,71471	−0,00071
Somma	1,614			**2,47020**	**20,80**		**0,00000**

Passo 1 — Medie. $\bar{x}$ = 13,50/5 = 2,70 mg/L; $\bar{y}$ = 1,614/5 = 0,3228.

Passo 2 — Pendenza e intercetta. b_1 = 2,4702/20,80 = 0,11876 L/mg; b_0 = 0,3228 − 0,11876 × 2,70 = 0,00215.

Passo 3 — Residui. Calcolati con i coefficienti non arrotondati sommano a zero, sono piccoli (al massimo 0,0018) e cambiano segno senza formare archi o imbuti; le concentrazioni retrocalcolate (per esempio (0,062 − 0,00215)/0,11876 = 0,504 mg/L) differiscono dai valori nominali di meno dell'1%.

Passo 4 — r². SQT = $\sum(y_i - \bar{y})^2$ = 0,293367; SQE = $\sum e_i^2$ = 6,80 × 10^{-6}; SQR = 0,293360; r² = 0,99998 (r = +0,99999).

Passo 5 — Predizione inversa. y_0 = 0,350 è compreso tra $\hat{y}$(0,50) = 0,0615 e $\hat{y}$(6,00) = 0,7147: si interpola. x_0 = (0,350 − 0,00215)/0,11876 = 0,34785/0,11876 = 2,929 mg/L.

Risultato: x_0 = 2,93 mg/L. La linearità è adeguata perché r² = 0,99998 **e** i residui sono casuali: il solo r² non basterebbe.

Arrotondamenti: 2,92 o 2,93 mg/L?

Con i coefficienti arrotondati b_1 = 0,119 L/mg e b_0 = 0,002 si ottiene x_0 = 0,348/0,119 = 2,92 mg/L. La differenza viene soprattutto dalla pendenza: 0,119 è più alta dello 0,2% rispetto a 0,11876, quindi x_0 risulta più basso dello 0,2%. Con la retta arrotondata anche i residui non sommano più a zero. Regola: conservare almeno 4 cifre significative nella pendenza (0,1188 L/mg) e arrotondare solo il risultato finale (Capitolo 1). Se però un esercizio fornisce la retta già arrotondata ($\hat{y}$ = 0,002 + 0,119 x), si usano quei coefficienti e il risultato corretto del calcolo è 2,92 mg/L.

Punti chiave del Capitolo 7

- In calibrazione X è la concentrazione degli standard (variabile indipendente) e Y il segnale (variabile dipendente): $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$. La regressione serve a prevedere, la correlazione misura l'associazione; nessuna delle due prova la causalità.
- β_0 è il segnale del bianco, β_1 la sensibilità; il residuo $e_i = y_i - \hat{y}_i$ stima l'errore ε_i.
- I minimi quadrati minimizzano $\sum e_i^2$: $b_1 = S_{xy}/S_{xx}$, $b_0 = \bar{y} - b_1\bar{x}$; la retta passa per $(\bar{x}, \bar{y})$ e il segno di b_1 è quello della codevianza.
- Predizione inversa $x_0 = (y_0 - b_0)/b_1$ solo per segnali interni all'intervallo (interpolazione); sopra il range si diluisce, sotto si riporta "< LLOQ".
- SQT = SQR + SQE; r^2 = SQR/SQT (0–1); r tra −1 e +1 con il segno della pendenza; r^2 alto non prova la linearità e non misura la sensibilità.
- Arrotondare solo alla fine: Fe(III), 2,93 mg/L con i coefficienti non arrotondati.

Mettiti alla prova

Nel quiz di autovalutazione esercitati con gli argomenti:

- Modello di regressione e diagramma di dispersione
- Metodo dei minimi quadrati
- Interpolazione, estrapolazione e predizione inversa
- Scomposizione della variabilità, r^2 e r

8. Diagnostica della regressione e collegamento alla validazione

Una retta dei minimi quadrati si calcola su qualunque insieme di dati, anche quando il modello non è adatto. Bisogna quindi verificarne le ipotesi con i residui, trattare correttamente una varianza non costante e sapere come gli errori sistematici deformano la retta: sono gli stessi strumenti della validazione dei metodi (Parte II).

8.1 L'analisi dei residui

Il **residuo** è la stima dell'errore commesso nel prevedere y_i con $\hat{y}_i$:

$$e_i = y_i - \hat{y}_i \tag{8.1}$$

Un residuo positivo indica un punto sopra la retta, uno negativo un punto sotto. Per esempio, con la retta $\hat{y}$ = 0,010 + 0,0950 x uno standard da 2,00 mg/L con assorbanza 0,196 ha $\hat{y}$ = 0,200 ed e = −0,004, e contribuisce a SQE con (−0,004)² = 0,000016. Gli scarti $\hat{y}_i - \bar{y}$ e $y_i - \bar{y}$ entrano in SQR e SQT, non sono residui. La somma dei residui è zero per costruzione e **non** prova che il modello sia adeguato.

Il **grafico dei residui** (e_i in funzione di X) è lo strumento diagnostico fondamentale, perché ingrandisce deviazioni invisibili nel grafico della retta; si usano anche i residui relativi o gli scarti percentuali delle concentrazioni retrocalcolate (Capitolo 10). Se il modello è adeguato i residui si distribuiscono **casualmente** attorno a zero in una banda uniforme (Figura 8.1a). Un andamento **ad arco** (Figura 8.1b) indica una **curvatura**: il modello lineare non è adeguato e occorre restringere l'intervallo alla zona lineare o usare un modello polinomiale (di solito quadratico) o non lineare. Un andamento **a imbuto** (Figura 8.1c) indica

eteroschedasticità (paragrafo 8.2). Un singolo residuo molto grande segnala un possibile valore anomalo da indagare (Capitolo 5).

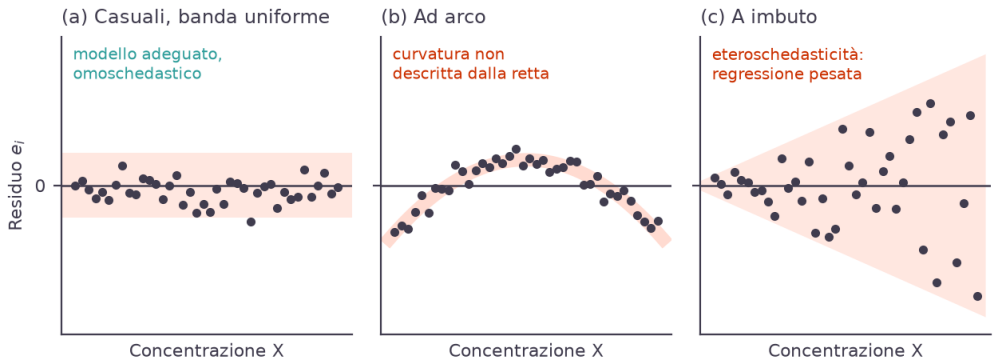


Figura 8.1 — Andamenti tipici del grafico dei residui. (a) Casuali in una banda uniforme: modello adeguato, varianza costante. (b) Ad arco: curvatura non descritta dalla retta. (c) A imbuto: varianza che cresce con la concentrazione (eteroschedasticità).

Esempio 8.1 — Una curvatura nascosta da un r² elevato

Otto standard (1, 2, …, 8 mg/L) danno assorbanze 0,098; 0,189; 0,278; 0,360; 0,437; 0,511; 0,577; 0,640. La retta dei minimi quadrati è $\hat{y}$ = 0,0375 + 0,0775 x, con r² = 0,9958.

Passo 1 — Residui. −0,0170; −0,0035; +0,0080; +0,0125; +0,0120; +0,0085; −0,0030; −0,0175: i segni (− − + + + + − −) disegnano un arco.

Passo 2 — Effetto sul risultato. Lo standard da 1 mg/L è retrocalcolato a (0,098 − 0,0375)/0,0775 = 0,78 mg/L, cioè −21,9%.

Passo 3 — Modello alternativo. La curva $\hat{y}$ = 0,100 x − 0,0025 x^2 lascia residui entro ±0,001, senza andamenti.

Risultato: nonostante r² = 0,9958 il modello lineare non è adeguato; solo i residui (o il retrocalcolo) rivelano la curvatura.

8.2 Omoschedasticità, eteroschedasticità e regressione pesata

Si ha **omoschedasticità** quando la varianza degli errori è **costante** al variare di X: è un'ipotesi della regressione ordinaria ai minimi quadrati (OLS) e nel grafico dei residui appare come una banda uniforme. Si ha **eteroschedasticità** quando la varianza **cambia** con X, di solito aumentando: i residui formano un **imbuto**. È frequente sugli ampi intervalli di concentrazione, per esempio in LC-MS/MS, dove il CV% del segnale è circa costante e quindi la deviazione standard cresce con la concentrazione.

Con dati eteroschedastici la OLS è dominata dagli standard più concentrati, che hanno gli errori assoluti maggiori: l'intercetta è stimata male e le concentrazioni basse risultano affette da grandi errori relativi, mentre r² resta altissimo. Si usa allora la **regressione pesata** (*weighted least squares*), che minimizza $\sum w_i\,(y_i - \hat{y}_i)^2$ con pesi w_i inversamente proporzionali alla varianza del segnale: i punti più variabili, cioè i più concentrati, contano meno. I pesi più usati sono $w_i = 1/x_i^2$, adatto quando la deviazione standard è proporzionale alla concentrazione (CV% costante, il caso tipico in LC-MS/MS), e $w_i = 1/x_i$, adatto quando è la varianza a crescere in proporzione alla concentrazione. Il rapporto tra i pesi degli standard estremi (1 con la OLS) è:

$$\frac{w(x_{min})}{w(x_{max})} = \left(\frac{x_{max}}{x_{min}}\right)^k \tag{8.2}$$

con k = 1 per i pesi $1/x$ e k = 2 per i pesi $1/x^2$. Si sceglie il peso più semplice che rende omogenei i residui relativi e minimizza la somma degli scarti percentuali delle concentrazioni retrocalcolate (Capitolo 10).

Esempio 8.2 — Regressione ordinaria e pesata in LC-MS/MS

Una calibrazione LC-MS/MS nel plasma (1–500 ng/mL, otto standard, segnale = rapporto delle aree analita/SIL-IS) ha CV% del segnale circa costante.

Passo 1 — Rapporto tra i pesi (equazione 8.2). 500/1 = 500: con pesi $1/x$ lo standard da 1 ng/mL pesa 500 volte quello da 500 ng/mL; con pesi $1/x^2$, 500² = 250 000 volte.

Passo 2 — Rette (da foglio di calcolo): OLS $\hat{y}$ = 0,00638 + 0,009778 x (r² = 0,9999); pesi $1/x^2$ $\hat{y}$ = 0,00119 + 0,009970 x.

Tabella 8.1 — Scarti percentuali delle concentrazioni retrocalcolate

x_i (ng/mL)	1	2	5	10	50	100	250	500
OLS (%)	−50,7	−27,3	−5,7	−0,2	−0,9	+3,8	−1,0	+0,1
Pesi $1/x$ (%)	−4,8	−4,5	+3,2	+4,0	−0,4	+3,8	−1,2	−0,2
Pesi $1/x^2$ (%)	+0,4	−2,6	+2,9	+3,1	−1,7	+2,3	−2,7	−1,7

Risultato: si sceglie la regressione pesata $1/x^2$. Con la OLS, nonostante r² = 0,9999, lo standard più basso è retrocalcolato a −50,7%, perché l'intercetta (0,00638) vale più di metà del suo segnale (0,0112). Con i pesi tutti gli scarti rientrano entro ±5% e $1/x^2$ dà la somma degli scarti assoluti più piccola (17,4% contro 22,1% con $1/x$ e 89,7% con OLS).

8.3 Effetto degli errori sistematici sulla curva di calibrazione

Un **errore costante** (Capitolo 2) aggiunge o toglie la stessa quantità di segnale a ogni livello: la retta trasla parallelamente e cambia l'**intercetta**. Un **errore proporzionale** altera il segnale di una frazione fissa: cambia la **pendenza** (sensibilità apparente), non l'intercetta. Nessuno dei due modifica r² o l'andamento dei residui.

Il verso della distorsione dipende da **dove** agisce l'errore, perché il segnale del campione viene confrontato con la retta. Con β_0, β_1 coefficienti della retta corretta e x_{vero} concentrazione vera:

- **errore costante $+c$ solo sugli standard** (per esempio un diluente contaminato usato solo per gli standard): intercetta maggiore e

$$x_0 = x_{\text{vero}} - \frac{c}{\beta_1} \tag{8.3}$$

cioè sottostima (sovrastima se $c < 0$); l'errore di segnale diventa concentrazione dividendo per la pendenza e pesa di più alle concentrazioni basse;

- **errore costante solo sul campione** (colore o torbidità della matrice, se non si sottrae il bianco del campione): $x_0 = x_{\text{vero}} + c/\beta_1$, sovrastima;
- **errore proporzionale solo sugli standard** (soluzione madre più o meno concentrata del dichiarato): la pendenza diventa $\beta_1(1+p)$ e

$$x_0 = \frac{x_{\text{vero}}}{1+p} \tag{8.4}$$

quindi una pendenza maggiore ($p > 0$) porta a sottostimare, una minore a sovrastimare; con +10% di pendenza x_0 cala del 9,1% (fattore 1/1,10), non del 10%;

- **errore proporzionale solo sul campione** (estrazione incompleta, cioè recupero R < 100%, o soppressione della ionizzazione con standard in solvente): $x_0 = R \cdot x_{\text{vero}}$, sottostima; un'esaltazione del segnale dà sovrastima (Capitolo 13);
- **errore uguale su standard e campione** (segnale dei reagenti, standard in matrice estratti come i campioni, compensazione con SIL-IS): nessuna distorsione. Un'estrazione poco efficiente ma uguale per tutti abbassa la sensibilità, non l'esattezza.

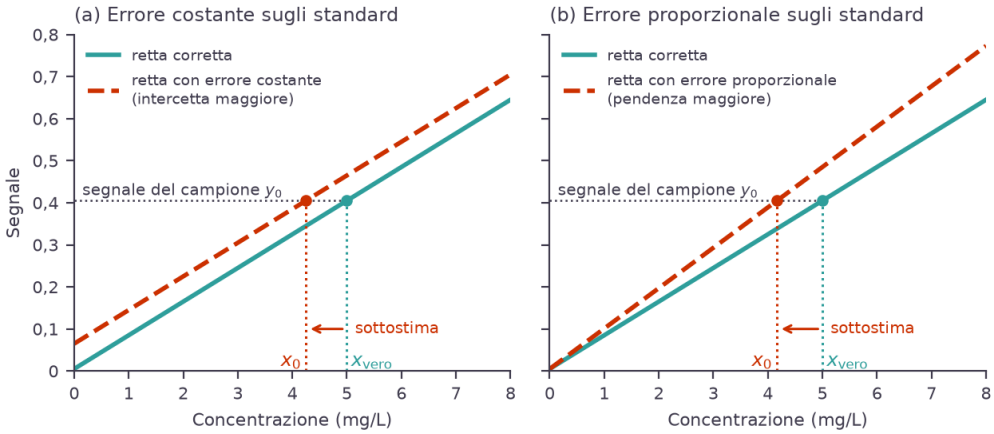


Figura 8.2 — Errore sistematico che riguarda solo gli standard, con segnale del campione invariato (errori esagerati per chiarezza): (a) un errore costante positivo alza l'intercetta; (b) un errore proporzionale positivo aumenta la pendenza. In entrambi i casi il segnale del campione incontra la retta errata a una concentrazione più bassa di quella vera.

Esempio 8.3 — Standard o campione?

La retta corretta è $\hat{y}$ = 0,005 + 0,0800 x (x in mg/L); un campione da 5,00 mg/L dà y_0 = 0,405.

Caso 1 — Errore costante +0,020 solo sugli standard. x_0 = (0,405 − 0,025)/0,0800 = 4,75 mg/L: sottostima di 0,020/0,0800 = 0,25 mg/L (−5%); per un campione da 1,00 mg/L sarebbe −25%.

Caso 2 — Stesso errore solo sul campione. x_0 = (0,425 − 0,005)/0,0800 = 5,25 mg/L (+5%); presente in entrambi, (0,425 − 0,025)/0,0800 = 5,00 mg/L.

Caso 3 — Pendenza degli standard +5%. x_0 = 0,400/0,0840 = 4,76 mg/L: −4,8% (= 1/1,05 − 1).

Caso 4 — Recupero dell'80% solo sul campione. y_0 = 0,005 + 0,0800 × 0,80 × 5,00 = 0,325; x_0 = 0,320/0,0800 = 4,00 mg/L (−20%). Con standard estratti allo stesso modo, pendenza 0,0640 L/mg e x_0 = 5,00 mg/L.

Risultato: lo stesso errore dà sottostima, sovrastima o nessuna distorsione a seconda che agisca sugli standard, sul campione o su entrambi.

8.4 Collegamento alla validazione dei metodi

- **Linearità e intervallo**: r^2, grafico dei residui, retrocalcolo degli standard e test di mancanza di adattamento (*lack of fit*), che con repliche a ogni livello confronta la distanza delle medie dalla retta con la dispersione delle repliche (Capitoli 10 e 11).
- **LOD e LOQ**: dalla deviazione standard del bianco s_B e dalla pendenza, per esempio $\text{LOD} = 3{,}3 \cdot s_B/b_1$ e $\text{LOQ} = 10 \cdot s_B/b_1$ (ICH Q2(R2), 2023). La pendenza converte il rumore, in unità di segnale, in concentrazione: a parità di rumore una pendenza maggiore dà un LOD più basso (con s_B = 0,0020 e b_1 = 0,119 L/mg, LOD = 0,055 mg/L; con pendenza doppia, 0,028 mg/L; Capitolo 11).
- **Esattezza e precisione**: bias e recupero rispetto a un valore di riferimento; s e CV% in ripetibilità, precisione intermedia e riproducibilità (Capitoli 3 e 12).
- **Recupero, selettività, effetto matrice**: prove di aggiunta (*spike*), interferenti, confronto di risposte o pendenze in solvente e in matrice (Capitoli 11 e 13).

- **Incertezza di misura**: non è una caratteristica di validazione ICH Q2(R2), ma un requisito della ISO/IEC 17025 (2017); si stima secondo la GUM (contributi di Tipo A e B) usando i dati di precisione ed esattezza della validazione, come indica la guida Eurachem (2025) (Capitolo 6).

8.5 Sintesi dei concetti chiave della Parte I

Tabella 8.2 — Concetti chiave della statistica applicata ai metodi analitici

Concetto	In breve
Incertezza	Ogni misura ha un'incertezza: risultato = valore ± U, con unità e cifre significative corrette (Capitoli 1 e 6).
Errori	Sistematici: bias = $\mu - x_t$, individuabili e correggibili. Casuali: riducibili e quantificabili, non eliminabili (Capitolo 2).
Precisione, esattezza, accuratezza	Precisione = dispersione (s, CV%); esattezza (*trueness*) = assenza di bias; **accuratezza = esattezza + precisione** (Capitolo 3).
Gaussiana	68,3%, 95,4% e 99,7% dei valori entro ±1σ, ±2σ, ±3σ; area totale = 1; intervallo di confidenza $\bar{x} \pm t \cdot s/\sqrt{n}$ (Capitolo 5).
GUM	Contributi di Tipo A (statistici) e B (altre informazioni); $U = k \cdot u_c$, con k = 2 per circa il 95% (Capitolo 6).
Regressione	$\hat{y} = b_0 + b_1 x$ ai minimi quadrati; r^2 **e** residui; $x_0 = (y_0 - b_0)/b_1$ per interpolazione (Capitoli 7 e 8).

Punti chiave del Capitolo 8

- Il residuo $e_i = y_i - \hat{y}_i$ stima l'errore; se il modello è adeguato i residui sono casuali attorno a zero. Somma nulla dei residui e r^2 alto non provano l'adeguatezza; residui ad arco indicano curvatura.
- Omoschedasticità = varianza costante (ipotesi della OLS); residui a imbuto = eteroschedasticità, da trattare con la regressione pesata ($1/x$ o $1/x^2$), che riduce l'influenza degli standard più concentrati.
- Errore costante → intercetta; errore proporzionale → pendenza. Solo sugli standard: intercetta o pendenza maggiori sottostimano il campione; solo sul campione il verso si inverte; su entrambi si compensa.
- Validazione: r^2, residui, lack of fit e retrocalcolo per la linearità; s_B/b_1 per LOD e LOQ; l'incertezza si stima dai dati di validazione ma non è una caratteristica ICH Q2(R2).
- Accuratezza = esattezza + precisione.

Mettiti alla prova

Nel quiz di autovalutazione esercitati con gli argomenti:

- Analisi dei residui e omoschedasticità
- Errori sistematici e retta di calibrazione
- Collegamento alla validazione dei metodi

Parte II — Calibrazione e validazione dei metodi analitici

Un metodo analitico produce risultati utili solo se è calibrato correttamente e se si dimostra, con prove documentate, che è adatto allo scopo. Questa parte segue il percorso dagli standard di riferimento alla curva di calibrazione, fino ai parametri e ai criteri di validazione previsti dalle linee guida internazionali, con particolare attenzione alla bioanalisi in LC-MS/MS.

9. Standard e materiali di riferimento

Nei metodi strumentali la concentrazione si ricava confrontando il segnale del campione con quello di standard a concentrazione nota: un risultato non può quindi essere più affidabile degli standard con cui è stato ottenuto. Un errore sulla purezza, sulla pesata o su una diluizione è un **errore sistematico** (Capitolo 2) che si trasferisce identico a tutti i campioni e che nessuna replica può rivelare.

9.1 La gerarchia degli standard analitici

Gli standard si organizzano in una **gerarchia** in cui ogni livello riceve il proprio valore per confronto con il livello superiore: in alto pochi materiali con l'incertezza più piccola, in basso le soluzioni di uso quotidiano (Figura 9.1).

Gli standard analitici

- **Standard primario** (PS): sostanza di purezza elevatissima e certificata, stabile, non igroscopica e di stechiometria nota, la cui quantità si stabilisce direttamente per pesata. È lo standard di livello più alto del laboratorio e non viene tarato rispetto ad altri standard della stessa grandezza.
- **Materiale di riferimento** (RM): materiale sufficientemente omogeneo e stabile rispetto a proprietà specificate, idoneo all'uso previsto (VIM 5.13).
- **Materiale di riferimento certificato** (CRM): RM accompagnato da un certificato di un ente autorevole che riporta uno o più valori di proprietà con la relativa **incertezza** e **tracciabilità**, ottenuti con procedure valide (VIM 5.14).
- **Standard secondario**: sostanza o soluzione che non soddisfa tutti i requisiti di un PS e il cui contenuto è stabilito per confronto con un PS o un CRM.
- **Standard di lavoro**: soluzione diluita, preparata dalla soluzione madre, usata nella routine per costruire la curva di calibrazione.

9.1.1 Gli standard primari e i loro requisiti

Uno standard primario si pesa, si scioglie in un volume noto e fornisce direttamente una soluzione di concentrazione nota. Esempi classici sono il ftalato acido di potassio (KHP, per standardizzare le basi), il carbonato di sodio (per gli acidi), l'ossalato di sodio (per il permanganato), il dicromato e lo iodato di potassio. L'espressione «riferimento assoluto» va intesa con cautela: anche la purezza di uno standard primario è stata determinata analiticamente e **ha un'incertezza** (per esempio 99,98% ± 0,02%). Per essere usato come tale, deve soddisfare **tutti insieme** quattro requisiti (Tabella 9.1).

Tabella 9.1 — I requisiti di uno standard primario

Requisito	Perché serve
Purezza ≥ 99,9%, con certificazione documentata	la massa pesata deve corrispondere alla massa di analita
Stabilità nel tempo e nelle condizioni di conservazione	la composizione non deve cambiare durante la conservazione (degradazione, ossidazione, reazione con la CO_2)
Non igroscopicità	durante la pesata non deve assorbire umidità, altrimenti si pesa anche acqua
Stechiometria nota (formula definita, massa molare esatta)	la quantità di sostanza si calcola come $n = m/M$

Stabilità e non igroscopicità sono requisiti distinti: la prima riguarda la conservazione nel tempo, la seconda il momento della pesata (una sostanza igroscopica fa salire la lettura della bilancia). L'NaOH, per esempio, ha una formula definita ma è igroscopico e assorbe CO_2: non è uno standard primario, e una formula definita, da sola, non basta.

9.1.2 Materiali di riferimento e materiali di riferimento certificati

Un **RM** deve essere omogeneo e stabile, ma non ha necessariamente un valore certificato (per esempio un *pool* di plasma usato come controllo interno). Solo un **CRM** ha valori **certificati con incertezza** (di solito estesa, con k = 2: Capitolo 6) e una dichiarazione di tracciabilità. Esistono CRM di **sostanza pura** e **soluzioni di calibrazione certificate**, che servono a calibrare, e **CRM di matrice** (sedimento, siero, latte in polvere), gli unici che riproducono la matrice dei campioni reali. Questi ultimi servono a **verificare l'esattezza dell'intera procedura** (Capitolo 12). Per controllare un metodo per il piombo nei sedimenti, mineralizzazione inclusa, il materiale adatto è un CRM di sedimento con piombo certificato. Una soluzione pura di piombo verifica solo la calibrazione; un sedimento non certificato permette di stimare la precisione, non l'esattezza.

Tra i principali produttori ci sono NIST (USA), JRC (UE), LGC (Regno Unito) e BAM (Germania). I requisiti di competenza dei **produttori di materiali di riferimento** sono fissati dalla **ISO 17034** (2016); la ISO/IEC 17025 (2017) riguarda invece i laboratori di prova e taratura.

9.1.3 Standard secondari

Uno standard secondario riceve il proprio valore per **standardizzazione** rispetto a un PS o a un CRM. Una soluzione di NaOH preparata per pesata approssimativa diventa uno standard secondario dopo essere stata titolata con KHP (Esempio 9.1). In bioanalisi si parte di solito da standard di riferimento commerciali con **certificato di analisi** (CoA), richiesto da ICH M10 (2022) per l'analita: la purezza riportata nel CoA va sempre usata nei calcoli.

Esempio 9.1 — Standardizzazione di NaOH con KHP

Dati. 0,5124 g di KHP (M = 204,22 g/mol, purezza 99,95%) richiedono 25,43 mL di NaOH «circa 0,1 mol/L»; la reazione è 1:1.

1. $n(\mathrm{KHP}) = \frac{0{,}5124 \times 0{,}9995}{204{,}22} = 2{,}508 \times 10^{-3}$ mol $= n(\mathrm{NaOH})$.
2. $c(\mathrm{NaOH}) = \frac{2{,}508 \times 10^{-3}\ \mathrm{mol}}{0{,}02543\ \mathrm{L}} = 0{,}09862$ mol/L.

Risultato: 0,09862 mol/L. Usando il valore nominale 0,1000 mol/L tutte le titolazioni successive sarebbero sovrastimate dell'1,4%.

9.2 La tracciabilità metrologica

Tracciabilità metrologica (VIM 2.41)

Proprietà di un **risultato di misura** per cui esso può essere collegato a un riferimento (di solito un'unità SI) attraverso una **catena ininterrotta e documentata di tarature**, ognuna delle quali contribuisce all'incertezza di misura.

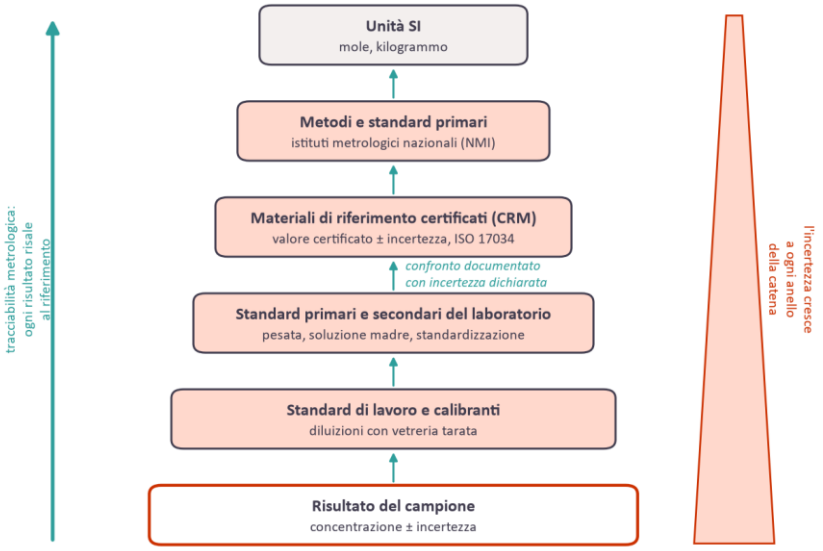


Figura 9.1 — Gerarchia degli standard e catena di tracciabilità: ogni anello è un confronto documentato con incertezza dichiarata; risalendo la catena il risultato è collegato alle unità SI, scendendo l'incertezza cresce.

Ogni anello della catena (certificato del CRM, taratura della bilancia, classe della vetreria, registro di preparazione) deve avere un'**incertezza dichiarata**. Poiché ogni passaggio aggiunge il proprio contributo, **l'incertezza cresce allontanandosi dal riferimento**, e un anello non documentato **interrompe** la tracciabilità. La tracciabilità è una proprietà del **risultato**, non del materiale: il CRM è lo strumento con cui il laboratorio la costruisce. Rende confrontabili risultati di laboratori diversi e ottenuti in tempi diversi, ma non garantisce l'esattezza del metodo: un'estrazione incompleta produce un risultato tracciabile ma sottostimato (Capitolo 13).

9.3 Dal PS agli standard di lavoro

La preparazione segue la sequenza **pesata del PS → soluzione madre → soluzione diluita → standard di lavoro** (Figura 9.2). Si pesa con una bilancia analitica tarata (meglio per differenza, evitando masse troppo piccole: con una lettura di 0,01 mg, 10 mg danno una risoluzione dello 0,1%, 1 mg dell'1%), si porta a volume in un **matraccio tarato** e si diluisce con **pipette e matracci tarati**:

$$c_{\text{madre}} = \frac{m \cdot P \cdot f}{V} \tag{9.1}$$

$$c_{\text{lavoro}} = c_{\text{madre}} \cdot \frac{V_{\text{prelievo}}}{V_{\text{finale}}} = \frac{c_{\text{madre}}}{\text{FD}} \tag{9.2}$$

dove m è la massa pesata, P la purezza come frazione, V il volume del matraccio, f il rapporto tra la massa molare dell'analita e quella della forma pesata (per esempio sale → base libera; f = 1 se coincidono) e FD il fattore di diluizione; con più diluizioni il fattore complessivo è il prodotto dei singoli fattori.

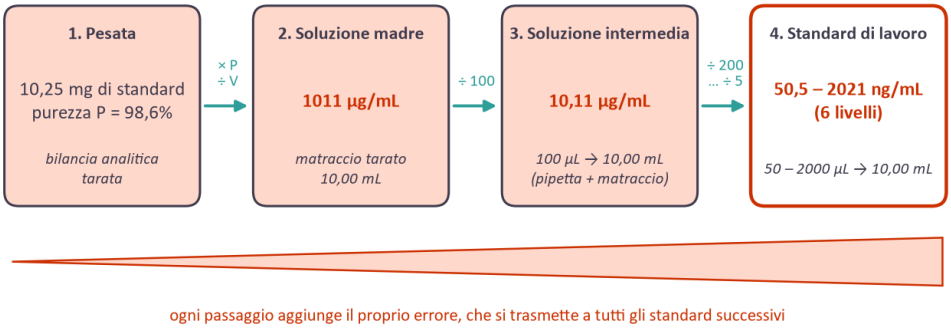


Figura 9.2 — Dalla pesata agli standard di lavoro (dati dell'Esempio 9.2): l'errore di ogni passaggio si somma ai precedenti e si trasmette a tutti gli standard successivi.

L'errore si propaga nelle diluizioni. L'errore relativo della pesata si trasmette inalterato a tutte le soluzioni derivate (le diluizioni non lo riducono), e ogni pipettatura aggiunge il proprio: la precisione della pipettatura è critica quanto quella della pesata. Conviene limitare il numero di passaggi ed evitare volumi molto piccoli; il calcolo della propagazione è nel Capitolo 16. Per poter scoprire un errore nella soluzione madre, ICH M10 raccomanda di preparare **calibranti e QC da soluzioni madre separate** (Capitolo 12).

Esempio 9.2 — Dalla pesata allo standard di lavoro

Dati. 10,25 mg di standard (purezza 98,6%) in un matraccio da 10,00 mL; 100 µL di soluzione madre portati a 10,00 mL; 500 µL di questa soluzione intermedia portati a 10,00 mL.

1. Soluzione madre (Eq. 9.1): $c = 10{,}25 \times 0{,}986 / 10{,}00 = 1{,}0107$ mg/mL = 1011 µg/mL.
2. Soluzione intermedia (FD = 100): 10,11 µg/mL.
3. Standard di lavoro (FD = 20): $10{,}107 / 20 = 0{,}5053$ µg/mL.

Risultato: 505 ng/mL. Ignorando la purezza si otterrebbero 512,5 ng/mL (+1,4%); dividendo per la purezza invece di moltiplicare, 519,8 ng/mL. Errori di questo tipo sono sistematici: spostano nello stesso verso tutti i risultati dei campioni.

Punti chiave del Capitolo 9

- Gerarchia: **standard primari**, **RM/CRM**, **standard secondari**, **standard di lavoro**; ogni livello riceve il valore per confronto con il superiore.
- Un PS richiede **tutti** i requisiti: purezza ≥ 99,9% certificata, stabilità, non igroscopicità, stechiometria nota; anche la sua purezza ha un'incertezza.
- Un **RM** è omogeneo e stabile; solo un **CRM** ha valori certificati con incertezza e tracciabilità. I produttori seguono la **ISO 17034**. I CRM di matrice verificano l'esattezza dell'intera procedura.
- Uno **standard secondario** (es. NaOH) è standardizzato con un PS (es. KHP) o un CRM.
- **Tracciabilità**: catena ininterrotta e documentata di confronti, ciascuno con incertezza dichiarata; l'incertezza cresce allontanandosi dal riferimento.
- Preparazione: pesata → madre → diluita → lavoro, con vetreria tarata; correggere per purezza (e forma salina); ogni passaggio aggiunge un errore che si trasmette a valle.

Mettiti alla prova

- **Standard e tracciabilità metrologica**
- **Preparazione degli standard; metodi assoluti e relativi** (domande sulla preparazione degli standard)

10. I metodi di calibrazione

Calibrare significa stabilire, con standard a concentrazione nota, la relazione tra segnale e quantità di analita, per poi usarla sui campioni. Se standard e campioni non rispondono allo stesso modo (matrice che sopprime il segnale, volume iniettato variabile), una curva statisticamente perfetta può dare risultati sistematicamente sbagliati: per questo la scelta del metodo di calibrazione conta quanto la regressione, trattata nei Capitoli 7 e 8.

10.1 Metodi assoluti e metodi relativi

Nei **metodi assoluti** la relazione tra grandezza misurata e quantità di analita è data da **leggi fisiche o dalla stechiometria**: gravimetria (massa di un precipitato di composizione nota), titrimetria (volume di titolante e rapporto stechiometrico), coulometria (legge di Faraday). Non serve una curva di calibrazione, ma le grandezze «note» devono esserlo davvero: bilancia e vetreria tarate e, nelle titolazioni, **titolante standardizzato con uno standard primario** (in una titolazione acido-base, NaOH con KHP: Esempio 9.1). Un errore sulla concentrazione del titolante è un errore sistematico proporzionale su tutti i risultati. Fa eccezione il titolante che è esso stesso uno standard primario preparato per pesata diretta (dicromato o iodato di potassio).

Nei **metodi relativi** la costante che lega segnale e concentrazione non è nota a priori, perché dipende da strumento, condizioni e analita: va determinata con standard, costruendo una **curva di calibrazione**. Vi appartiene la maggior parte dei metodi strumentali: spettroscopia (UV-Vis, AAS, ICP-OES, ICP-MS), cromatografia (HPLC, GC) e spettrometria di massa (LC-MS/MS).

10.2 Il bianco analitico

Il **bianco** è un campione che contiene tutto tranne l'analita e viene misurato come i campioni: il suo segnale stima il contributo non dovuto all'analita (Tabella 10.1).

Tabella 10.1 — I tipi di bianco analitico

Tipo di bianco	Composizione	Che cosa corregge
Del solvente	solo il solvente del campione	il fondo strumentale del solvente
Dei reagenti	solvente e tutti i reagenti della preparazione	le interferenze additive dei reagenti
Procedurale	sottoposto a tutta la procedura, di solito senza matrice (es. acqua reagente)	le contaminazioni introdotte durante il trattamento
Di matrice	matrice priva di analita, trattata come il campione	tutti i contributi non dovuti all'analita, matrice compresa

Il **bianco di matrice** è la migliore approssimazione del «vero bianco», perché riproduce sia la composizione del campione sia l'intero trattamento; il bianco procedurale, invece, attraversa tutta la procedura ma non contiene necessariamente la matrice. Una matrice priva di analita non sempre esiste (analiti endogeni). In bioanalisi ICH M10 (2022) distingue il **campione bianco** (matrice senza analita né standard interno) e il **campione zero** (matrice con il solo standard interno).

Che cosa corregge la sottrazione del bianco. Un contributo **additivo** (fondo, reagenti, componenti della matrice che danno segnale) si somma al segnale dell'analita senza cambiare la pendenza: sottraendo il bianco lo si elimina. Un effetto **moltiplicativo** (soppressione o esaltazione della ionizzazione in LC-MS/MS, diversa nebulizzazione in ICP, estrazione incompleta) cambia invece la **pendenza**, cioè la sensibilità: nel bianco, dove l'analita manca, non produce alcun segnale, quindi la sottrazione **non lo corregge**. Servono la calibrazione in matrice, lo standard interno o le aggiunte multiple (paragrafi 10.6–10.7; errori costanti e proporzionali:

Capitoli 2 e 8). Se i calibranti sono preparati e trattati come i campioni, il contributo del bianco è già nell'intercetta della curva e non va sottratto una seconda volta.

10.3 Scelta del segnale: area o altezza del picco

L'**area del picco** raccoglie l'intera risposta dell'analita ed è **meno sensibile alle variazioni di forma**: se il picco si allarga o presenta *tailing*, a parità di quantità iniettata l'altezza diminuisce mentre l'area resta pressoché costante. Per questo, in cromatografia e in LC-MS/MS, l'area è lo **standard de facto** per la quantificazione, salvo diversa indicazione della metodica. L'**altezza** è più semplice da misurare manualmente, ma è accettabile solo se la forma del picco è **costante e riproducibile**.

10.4 La calibrazione con standard esterno

Nella **calibrazione con standard esterno** (SE) si misurano standard a concentrazione nota preparati **separatamente** dal campione, si determina la curva per regressione e si ricava la concentrazione del campione per **interpolazione** (Figura 10.1):

$$C_{\text{campione}} = \frac{S - b_0}{b_1} \cdot \text{FD} \tag{10.1}$$

dove S è il segnale del campione, b_0 e b_1 intercetta e pendenza della curva e FD il fattore di diluizione applicato al campione (1 se non diluito).

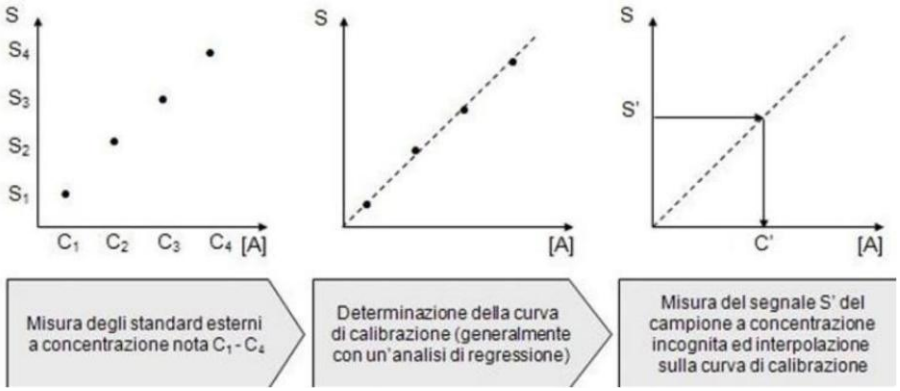


Figura 10.1 — Calibrazione con standard esterno: misura degli standard, determinazione della curva per regressione, interpolazione del segnale S' del campione. Lo schema è semplificato: una curva reale ha almeno 6 livelli e l'intercetta va stimata, non imposta a zero.

È un metodo **semplice e rapido**: una sola curva serve per tutti i campioni e il numero di standard è ridotto. Il suo **presupposto critico** è che la risposta all'analita sia **identica nello standard e nel campione**: stessa forma chimica e assenza di interferenze di matrice. Lo SE **non compensa né l'effetto matrice né la variabilità dell'iniezione**: se la matrice sopprime il segnale del 20%, il risultato è sottostimato del 20%. Preparando i calibranti in matrice bianca (*matrix-matched*) si compensa solo l'effetto medio della matrice, non le differenze tra campioni. Lo SE è quindi la scelta per l'analisi di routine di molti campioni con **matrice semplice e riproducibile**. La curva vale solo per interpolazione: un campione sopra l'ULOQ va diluito e rianalizzato (Capitolo 14).

Esempio 10.1 — Standard esterno: diluizione ed effetto matrice

Dati. Curva da 2,00 a 50,0 ng/mL: $A = 1850 \cdot C + 950$ (A in u.a., C in ng/mL). Un campione diluito 1:10 (0,100 mL portati a 1,00 mL) dà A = 46 200 u.a.

1. Soluzione misurata: $C = (46\,200 - 950)/1850 = 24{,}46$ ng/mL, all'interno dell'intervallo di calibrazione.
2. Campione originale (Eq. 10.1, FD = 10): $24{,}46 \times 10 = 244{,}6$ ng/mL.
3. Se la matrice del campione riducesse la risposta del 20% (effetto moltiplicativo), il segnale netto sarebbe 0,80 volte quello atteso e si otterrebbe $0{,}80 \times 244{,}6 = 195{,}7$ ng/mL: la sottrazione del bianco non cambierebbe nulla, perché il bianco non contiene analita.

Risultato: 245 ng/mL. Errori tipici: dimenticare FD (24,5 ng/mL), dividere per FD (2,45 ng/mL), ignorare l'intercetta (249,7 ng/mL).

10.5 La curva di calibrazione: requisiti operativi

I requisiti seguenti sono quelli di ICH M10 per i metodi cromatografici (la linea guida ha sostituito, per il suo ambito, EMA 2011 e FDA 2018).

1. **Almeno 6 livelli di concentrazione** dal LLOQ (il calibrante più basso) all'ULOQ (il più alto; Capitolo 11), **più un campione bianco e un campione zero**, che non contano tra i livelli e non entrano nella regressione.
2. **Regressione ai minimi quadrati** (Capitolo 7), eventualmente pesata ($1/x$, $1/x^2$) se i dati sono eteroschedastici (Capitolo 8). **L'intercetta non va forzata a zero**: contiene il segnale del bianco, e annullarla introduce un errore sistematico, soprattutto vicino al LLOQ (Capitolo 16).
3. **Retrocalcolo**: la concentrazione di ogni calibrante si ricalcola dalla curva e si confronta con il valore nominale (Eq. 10.2). Deve stare entro **±15%** (±20% al LLOQ) e **almeno il 75% dei calibranti**, con **almeno 6 livelli**, deve essere conforme. Un calibrante non conforme si esclude e la regressione si ricalcola; la percentuale si calcola sul totale dei calibranti. Per i metodi *ligand-binding* i limiti sono ±20% (±25% al LLOQ e all'ULOQ).
4. **Una nuova curva per ogni sequenza analitica**, perché sensibilità, fondo e prestazioni dello strumento cambiano nel tempo (sorgente, colonna, reattivi); si monitorano pendenza, segnale del bianco e QC (Capitoli 12 e 16).

$$\text{scostamento}\,\% = \frac{C_{\text{retro}} - C_{\text{nom}}}{C_{\text{nom}}} \times 100 \quad \text{con} \quad C_{\text{retro}} = \frac{S_i - b_0}{b_1} \tag{10.2}$$

Le due condizioni del punto 3 vanno lette insieme: con 8 livelli bastano 6 calibranti conformi (6/8 = 75%); con 6 livelli devono essere conformi tutti, perché 5/6 = 83% supera il 75% ma lascia solo 5 livelli.

Perché r^2 non basta. Un r^2 prossimo a 1 può mascherare deviazioni sistematiche dalla linearità, specialmente agli estremi del range, e due modelli con lo stesso r^2 possono avere residui molto diversi. Su un intervallo ampio (1–1000 ng/mL) la variabilità totale è dominata dai calibranti alti: un errore di pochi decimi di ng/mL al LLOQ non cambia r^2, ma su 1 ng/mL vale il 20–50%. Servono quindi il **grafico dei residui**, con residui distribuiti casualmente attorno allo zero (Capitolo 8), e le **concentrazioni retrocalcolate**, che esprimono lo scostamento di ogni calibrante in percentuale.

Esempio 10.2 — Retrocalcolo e regola del 75%

Dati. Curva a 8 livelli (una replica per livello). Scostamenti dei valori retrocalcolati: LLOQ (1,00 ng/mL → 1,17 ng/mL) +17,0%; 10,0 ng/mL → 8,38 ng/mL; tutti gli altri entro ±5%.

1. LLOQ: +17,0%, entro ±20%: conforme.
2. Livello 10,0 ng/mL (Eq. 10.2): $(8{,}38 - 10{,}0)/10{,}0 \times 100 = -16{,}2\%$, oltre ±15%: non conforme.
3. Conformi: 7/8 = 87,5% ≥ 75%, con 7 livelli ≥ 6.

Risultato: curva accettabile; il livello non conforme si esclude e si ricalcola la regressione. Attenzione al denominatore: un calibrante nominale 40,0 ng/mL retrocalcolato a 46,6 ng/mL dà $(46{,}6 - 40{,}0)/40{,}0 \times 100 = +16{,}5\%$ (non conforme); dividendo per 46,6 si otterrebbe +14,2% e lo si accetterebbe a torto, qualunque sia r^2.

10.6 La calibrazione con standard interno

Nella **calibrazione con standard interno** (SI) si aggiunge una quantità nota di una specie di riferimento **il prima possibile** nel processo, alla **stessa concentrazione** C_s, a **standard, QC, campioni e campione zero**, ma **non al bianco**. Il segnale di risposta è il **rapporto** tra il segnale dell'analita e quello dello SI:

$$R = \frac{S_{\text{analita}}}{S_{\text{SI}}} = b_0 + b_1 C \quad \Rightarrow \quad C = \frac{R - b_0}{b_1} \tag{10.3}$$

Le variazioni che agiscono **in proporzione** su entrambi i segnali si annullano nel rapporto (non nella differenza). Se in GC con microsiringa si inietta il 10% di volume in meno, con lo SE il risultato è sottostimato del 10%, mentre con lo SI entrambi i segnali calano del 10% e il rapporto non cambia. Uno SI aggiunto prima dell'estrazione compensa anche le perdite di preparazione; aggiunto solo prima dell'iniezione compensa la sola variabilità strumentale.

Lo SI si usa quando la **riproducibilità dell'iniezione è critica** (microsiringa in GC), quando la **sorgente del segnale è instabile** (fiamma, plasma ICP), quando la preparazione comporta **perdite variabili** e in LC-MS/MS, dove la ionizzazione dipende dalla matrice. L'effetto matrice è però compensato **solo se lo SI ne subisce uno identico** a quello dell'analita; uno SI che eluisce a un tempo diverso può subire un effetto diverso, e il rapporto non corregge le interferenze additive. La correzione è quindi **parziale**, e diventa **ottimale** con uno **SIL-IS** (standard interno marcato con isotopi stabili: analita con ^{2}H, ^{13}C o ^{15}N), che co-eluisce con l'analita, ne subisce le stesse perdite e la stessa soppressione o esaltazione, ma è distinguibile per *m/z* (dettagli e limiti nel Capitolo 16).

Requisiti dello SI: **assente nel campione** (altrimenti la sua quantità totale non sarebbe nota), **non reattivo** con analita e matrice, **segnale ben distinto** da tutti gli altri, **caratteristiche chimico-fisiche simili** all'analita.

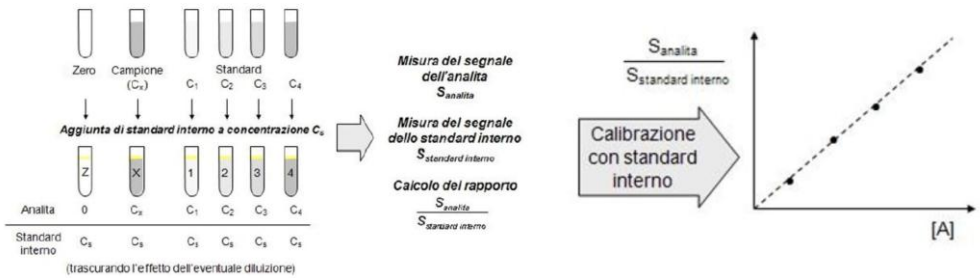


Figura 10.2 — Schema operativo dello standard interno: lo SI è aggiunto alla stessa concentrazione Cs al campione zero, al campione e agli standard (non al bianco, non mostrato); per ogni soluzione si calcola il rapporto S(analita)/S(standard interno) e si costruisce la retta del rapporto in funzione della concentrazione. Lo schema mostra 4 livelli per semplicità.

Come mostra la Figura 10.2, il campione bianco non contiene né analita né SI e il campione zero contiene solo lo SI: nessuno dei due conta tra i 6 livelli della curva. ICH M10 chiede inoltre di monitorare la risposta dello SI nei campioni, perché un valore anomalo segnala un problema in quel campione. Esempio di calcolo con la retta $R = 0{,}0250 \cdot C + 0{,}0080$ (C in ng/mL): un campione con A_{analita} = 18 450 e A_{SI} = 92 300 dà $R = 18\,450/92\,300 = 0{,}1999$ e $C = (0{,}1999 - 0{,}0080)/0{,}0250 = 7{,}68$ ng/mL (ignorando l'intercetta si otterrebbe 8,00 ng/mL; con il rapporto invertito $A_{\text{SI}}/A_{\text{analita}}$ un valore assurdo).

10.7 Il metodo delle aggiunte multiple

Nel **metodo delle aggiunte multiple** (aggiunte standard) la calibrazione si costruisce **nella matrice del campione**: aliquote uguali del campione a concentrazione incognita C_x si arricchiscono con quantità **crescenti e note** di analita puro, si portano allo stesso volume e si misurano (Figura 10.3). Se la risposta è lineare e corretta per il bianco, $S = k(C_x + C_{\text{agg}})$: la regressione del segnale in funzione della concentrazione aggiunta dà $b_1 = k$ e $b_0 = kC_x$, e la retta estrapolata interseca l'asse delle ascisse in $-C_x$:

$$C_x = \frac{b_0}{b_1} \tag{10.4}$$

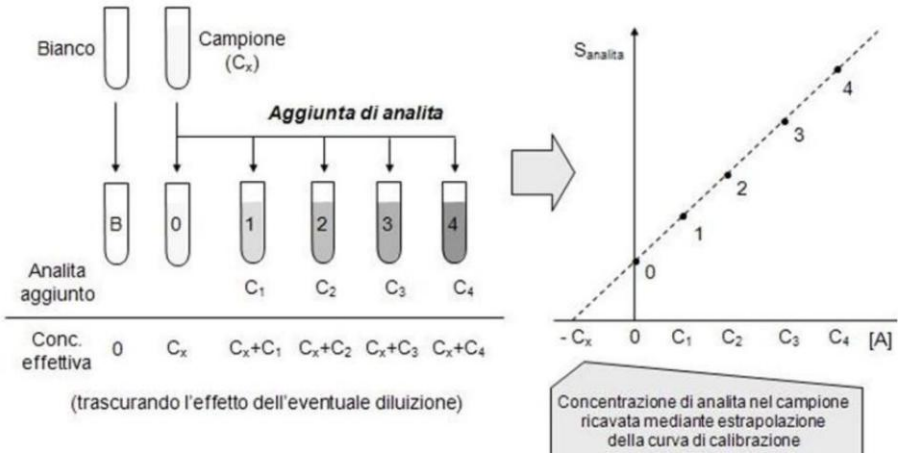


Figura 10.3 — Metodo delle aggiunte multiple: aliquote del campione arricchite con quantità crescenti di analita; la retta del segnale, estrapolata, interseca l'asse delle ascisse in −Cx (schema che trascura la diluizione).

Poiché la pendenza k è misurata nella matrice del campione, l'**effetto matrice moltiplicativo** si annulla nel rapporto b_0/b_1. C_x è la concentrazione nella soluzione misurata: per il campione originale si moltiplica per il fattore di diluizione.

Requisiti: **linearità** in tutto l'intervallo delle aggiunte (e nel tratto estrapolato) e **correzione del bianco prima dell'estrapolazione**, perché l'intercetta deve contenere solo il segnale dell'analita. **Limiti**: il metodo **non corregge le interferenze additive costanti**, che aumentano b_0 e portano a sovrastimare C_x; richiede **una curva per ogni campione**, quindi è lento e costoso; l'estrapolazione aumenta l'incertezza. È indicato per **matrici complesse e irriproducibili** in assenza di uno SI adatto.

Esempio 10.3 — Aggiunte multiple con correzione del bianco

Dati. La regressione dei segnali **grezzi** in funzione della concentrazione aggiunta dà b_0 = 0,210 e b_1 = 0,0700 L/mg; il bianco dei reagenti dà 0,035. La soluzione misurata è stata ottenuta diluendo il campione 1:5.

1. Il bianco è un contributo costante e modifica solo l'intercetta: b_0 corretto = 0,210 − 0,035 = 0,175.
2. Eq. 10.4: $C_x = 0{,}175/0{,}0700 = 2{,}50$ mg/L nella soluzione misurata.
3. Nel campione originale: $2{,}50 \times 5 = 12{,}5$ mg/L.

Risultato: 2,50 mg/L nella soluzione misurata (12,5 mg/L nel campione). Senza correzione si otterrebbe 3,00 mg/L (+20%); invertendo il rapporto (b_1/b_0) un valore privo di senso.

10.8 Confronto tra i metodi di calibrazione

Tabella 10.2 — Confronto tra i metodi di calibrazione

Caratteristica	Standard esterno	Standard interno	Aggiunte multiple
Semplicità	Alta	Media	Bassa
Correzione dell'effetto matrice	No (con calibranti in matrice, solo l'effetto medio)	Parziale, se lo SI subisce lo stesso effetto (SIL-IS: ottimale)	Sì (moltiplicativo)
Correzione della variabilità dell'iniezione	No	Sì	No
Numero di standard per campione	Basso	Medio (stessi calibranti, più lo SI in ogni soluzione)	Alto (1 curva per campione)
Applicazione tipica	Matrici semplici	LC-MS/MS, GC-MS	Matrici complesse irriproducibili

Nessun metodo corregge le interferenze additive, che richiedono il bianco. Per esempio, per un contaminante in acque reflue di impianti industriali diversi, con matrice molto variabile e senza uno SI adatto, la strategia appropriata sono le aggiunte multiple su ciascun campione: lo SE non corregge l'effetto matrice, e una curva costruita su un solo campione non vale per gli altri.

Punti chiave del Capitolo 10

- **Metodi assoluti** (gravimetria, titolazioni): nessuna curva, ma titolante standardizzato con un PS. **Metodi relativi** (quasi tutti i metodi strumentali): curva di calibrazione.
- Il **bianco** corregge solo i contributi **additivi**, non l'effetto matrice **moltiplicativo**; il bianco di matrice è la migliore approssimazione del vero bianco.
- Si quantifica con l'**area** del picco; l'altezza solo se la forma è costante.
- **Standard esterno**: semplice, una curva per tutti; presuppone risposta identica in standard e campioni; non corregge effetto matrice né iniezione.
- **Curva** (ICH M10): ≥ 6 livelli dal LLOQ all'ULOQ più bianco e zero; intercetta non forzata; retrocalcolo ±15% (±20% al LLOQ) sul nominale; ≥ 75% conformi con ≥ 6 livelli; nuova curva per ogni sequenza. r^2 **non basta**.
- **Standard interno**: stessa quantità in standard, QC, campioni e zero (non nel bianco), aggiunto il prima possibile; rapporto $S_{\text{analita}}/S_{\text{SI}}$; correzione dell'effetto matrice parziale, ottimale con SIL-IS.
- **Aggiunte multiple**: $C_x = b_0/b_1$; correggono l'effetto moltiplicativo, non quello additivo (bianco sottratto prima); una curva per campione.

Mettiti alla prova

- **Preparazione degli standard; metodi assoluti e relativi** (metodi assoluti e relativi)
- **Bianco analitico e scelta del segnale**
- **Calibrazione con standard esterno**
- **Costruzione e valutazione della curva di calibrazione**
- **Calibrazione con standard interno**
- **Metodo delle aggiunte multiple**
- **Metodi di calibrazione e metodi strumentali** (confronto tra i metodi)

11. La validazione: scopo, selettività, sensibilità, LOD/LOQ e range lineare

I risultati di un metodo analitico servono a prendere decisioni: la dose di un farmaco, la conformità di un alimento, il rilascio di un lotto. Prima di usarli bisogna dimostrare che il metodo li fornisce in modo affidabile: è la **validazione**. Questo capitolo ne presenta lo scopo e i primi parametri (selettività, sensibilità, LOD/LOQ, intervallo di lavoro); esattezza, precisione e QC sono nel Capitolo 12, recupero ed effetto matrice nel 13, stabilità e altri parametri nel 14, il quadro normativo nel 15.

11.1 Perché validare un metodo analitico

Validazione

La **validazione** è il processo che dimostra, con prove sperimentali documentate, che un metodo è adatto allo scopo per cui viene utilizzato (*fit for purpose*).

Le prestazioni richieste dipendono dallo scopo: uno screening di presenza/assenza richiede soprattutto un limite di rivelabilità basso, un metodo per uno studio di farmacocinetica richiede esattezza e precisione note su un ampio intervallo di concentrazioni. La validazione segue un **protocollo** scritto in anticipo, con esperimenti e **criteri di accettabilità** prefissati, e si conclude con un rapporto che dichiara se il metodo è idoneo.

Non va confusa con la **tracciabilità metrologica** (catena ininterrotta di tarature fino a un riferimento, Capitolo 9), con il **test di idoneità del sistema** (verifica delle prestazioni dello strumento prima di ogni sequenza) o con la **stima dell'incertezza** di ciascun risultato (Capitolo 6). Un metodo nuovo richiede una validazione completa; una modifica, compreso il trasferimento a un altro laboratorio, richiede una validazione parziale (Capitolo 16).

La linea guida di riferimento dipende dal settore (Tabella 11.1; dettagli nel Capitolo 15). Per esempio, per un metodo ICP-MS per i metalli nelle acque superficiali, che non è un ambito farmaceutico, bioanalitico o alimentare, il riferimento adatto è la guida Eurachem. In questo capitolo e nel 12 i criteri numerici citati sono quelli di ICH M10 per i metodi cromatografici.

Tabella 11.1 — Principali documenti di riferimento per la validazione

Documento	Ambito
ICH M10 (2022)	Bioanalisi (farmaci e metaboliti in matrici biologiche); adottata da EMA e FDA, sostituisce per questo ambito EMA 2011 e FDA 2018
ICH Q2(R2) e Q14 (2023)	Metodi analitici per prodotti farmaceutici
Eurachem, *The Fitness for Purpose of Analytical Methods* (3ª ed., 2025)	Laboratori chimici in generale
SANTE/11312/2021 v2026	Residui di pesticidi in alimenti e mangimi
Reg. (UE) 2021/808	Residui di farmaci veterinari negli alimenti di origine animale

11.2 I parametri di validazione

I parametri principali sono: **selettività/specificità** (11.3), **curva di calibrazione e intervallo** (Capitolo 10 e 11.6), **LOD/LOQ** (11.5), **esattezza** e **precisione** (Capitolo 12), **recupero** ed **effetto matrice** (Capitolo 13), **stabilità** (Capitolo 14). Si aggiungono **carry-over**, **integrità della diluizione**, **robustezza** e **solidità** (Capitolo 14). Non tutte le linee guida li richiedono allo stesso modo: ICH M10, per esempio, non include recupero e robustezza tra gli elementi della validazione completa (il recupero è trattato nello sviluppo del metodo).

Nessun parametro, da solo, dimostra che un metodo è valido. Un metodo con una retta eccellente può essere inutilizzabile perché un metabolita interferisce o perché i risultati al livello più basso sono troppo dispersi.

Un r^2 elevato non è una validazione

«Il metodo è validato perché la retta ha $r^2 = 0{,}999$» è una conclusione sbagliata. Il coefficiente di determinazione non dice nulla su bias, precisione, interferenze o stabilità, e ICH M10 non fissa alcun valore minimo di r^2: vanno documentati tutti i parametri.

11.3 Selettività e specificità

Selettività e specificità

La **selettività** è la capacità di un metodo di misurare l'analita in presenza di altri componenti della matrice (sostanze endogene, metaboliti, farmaci co-somministrati) senza interferenze significative. La **specificità** è la selettività assoluta (risposta solo all'analita); nella pratica i due termini sono spesso usati come sinonimi.

ICH M10 li usa per due prove distinte, con gli stessi criteri: la selettività riguarda i componenti della matrice bianca, la specificità le sostanze correlate all'analita (metaboliti, isomeri, prodotti di degradazione, farmaci concomitanti). In LC-MS/MS la selettività nasce dalla separazione cromatografica e dalla transizione MRM; in HPLC-UV una sostanza che co-eluisce e assorbe alla stessa lunghezza d'onda si somma al segnale. Un'interferenza produce di solito un errore additivo, che pesa soprattutto ai livelli bassi (Capitolo 12).

La selettività si verifica su **campioni bianchi di matrice** (matrice trattata senza analita né standard interno, SI) provenienti da **almeno 6 fonti o lotti individuali**, per esempio plasma di 6 donatori diversi. Servono lotti individuali perché un'interferenza può essere presente in un solo donatore: sei repliche dello stesso lotto non la mostrerebbero e un pool la diluirebbe. Secondo ICH M10 le risposte interferenti devono essere **non più del** ($\leq$) **20%** della risposta dell'analita al LLOQ, al tempo di ritenzione dell'analita, e **non più del 5%** della risposta dello SI nel campione al LLOQ, al tempo di ritenzione dello SI.

Criteri di selettività (ICH M10)

$$\frac{A_{\text{bianco}}(t_{R,\text{analita}})}{A_{\text{analita}}(\text{LLOQ})} \times 100 \leq 20\%$$

$$\frac{A_{\text{bianco}}(t_{R,\text{SI}})}{A_{\text{SI}}(\text{LLOQ})} \times 100 \leq 5\%$$

- A_{bianco}: area misurata nel bianco al tempo di ritenzione t_R dell'analita o dello SI;
- $A_{\text{analita}}(\text{LLOQ})$, $A_{\text{SI}}(\text{LLOQ})$: aree dell'analita e dello SI nel campione al LLOQ.

Il limite per lo SI è più severo perché il risultato si calcola dal rapporto analita/SI: un'interferenza sullo SI altera quel rapporto in tutto l'intervallo.

Esempio 11.1 — Selettività in lotti di plasma

Al LLOQ l'area dell'analita è 4 200 e quella dello SI 150 000. Tre lotti di plasma bianco danno le aree interferenti (analita; SI): lotto 1: 650; 2 900. Lotto 2: 905; 1 800. Lotto 3: 120; 8 100.

1. Limiti: 20% di 4 200 = 840; 5% di 150 000 = 7 500.
2. Lotto 1: 650/4 200 = 15,5% e 2 900/150 000 = 1,9%: conforme.
3. Lotto 2: 905/4 200 = 21,5% > 20%: non conforme per l'analita.
4. Lotto 3: 120/4 200 = 2,9%, ma 8 100/150 000 = 5,4% > 5%: non conforme per lo SI.

Risultato: i lotti 2 e 3 superano i limiti; la selettività non è dimostrata finché le interferenze non sono eliminate. Confrontando per errore 905 con l'area dello SI si otterrebbe 0,6% e il lotto 2 sembrerebbe conforme.

11.4 Sensibilità di calibrazione e sensibilità analitica

Sensibilità

La **sensibilità di calibrazione** è la pendenza della retta di calibrazione, cioè la variazione del segnale per unità di concentrazione; per una curva lineare è costante:

$$b_1 = \frac{\Delta S}{\Delta C} \tag{11.1}$$

La **sensibilità analitica** è il rapporto tra la pendenza e la deviazione standard s_S del segnale alla concentrazione considerata:

$$\gamma = \frac{b_1}{s_S} \tag{11.2}$$

La pendenza non tiene conto del rumore: amplificando il segnale crescono sia b_1 sia il rumore. La sensibilità analitica corregge questo difetto, ma **dipende dalla concentrazione anche per curve lineari**, perché s_S varia con il livello del segnale (eteroscedasticità, Capitolo 8) mentre b_1 resta costante.

Il termine «sensibilità» è spesso usato impropriamente per indicare il LOD, ma sono grandezze distinte: il LOD dipende sia dalla pendenza sia dal rumore del bianco (11.5). Nel glossario di ICH M10 *sensitivity* indica il LLOQ: leggendo un documento conviene controllare in che senso è usato il termine.

Esempio 11.2 — Pendenza più alta non significa LOD più basso

Metodo A: b_1 = 2 500 unità di area per ng/mL, s_B = 450. Metodo B: b_1 = 1 200, s_B = 120 (formula 11.3).

1. LOD_A = 3 · 450/2 500 = 0,54 ng/mL.
2. LOD_B = 3 · 120/1 200 = 0,30 ng/mL.

Risultato: B ha la pendenza minore ma il LOD migliore (0,30 ng/mL), grazie al rumore più basso. Dimenticando il fattore 3 si otterrebbe 0,10 ng/mL per B; usando 10 si otterrebbe il LOQ (1,0 ng/mL).

11.5 Limite di rivelabilità e limite di quantificazione

Il **LOD** (limite di rivelabilità) è la concentrazione minima che dà un segnale distinguibile dal bianco con un livello di fiducia definito: ha un uso **qualitativo** (presenza/assenza). Il **LOQ** (limite di quantificazione) è la concentrazione minima determinabile con precisione ed esattezza accettabili: da lì inizia l'uso **quantitativo**.

Si misura più volte un bianco (di solito almeno una decina di repliche) e se ne calcolano media $\bar{y}_B$ e deviazione standard s_B. I **segnali limite** sono $y_{\mathrm{LOD}} = \bar{y}_B + 3\,s_B$ e $y_{\mathrm{LOQ}} = \bar{y}_B + 10\,s_B$ e dividono l'asse del segnale in tre zone: non rilevabile, rilevabile ma non quantificabile, quantificabile (Figura 11.1A). Il fattore 3 è il multiplo di s_B di cui il segnale deve superare il bianco per dire, con un livello di fiducia definito, che l'analita è presente (per una distribuzione normale un bianco supera $\bar{y}_B + 3\,s_B$ solo nello 0,13% circa dei casi). Al livello y_{LOQ} il segnale netto vale $10\,s_B$, quindi la sua deviazione standard relativa è circa il 10%, compatibile con una quantificazione.

I segnali limite si convertono in concentrazioni con la retta di calibrazione, la cui intercetta coincide circa con $\bar{y}_B$: si divide il segnale netto per la pendenza (Figura 11.1B).

LOD e LOQ (concentrazioni)

$$\mathrm{LOD} = \frac{3\,s_B}{b_1} \tag{11.3}$$

$$\mathrm{LOQ} = \frac{10\,s_B}{b_1} \tag{11.4}$$

- s_B: deviazione standard del segnale del bianco; b_1: pendenza della retta.
- $\bar{y}_B + 3\,s_B$ e $\bar{y}_B + 10\,s_B$ sono **segnali** (y_{LOD}, y_{LOQ}), non concentrazioni.

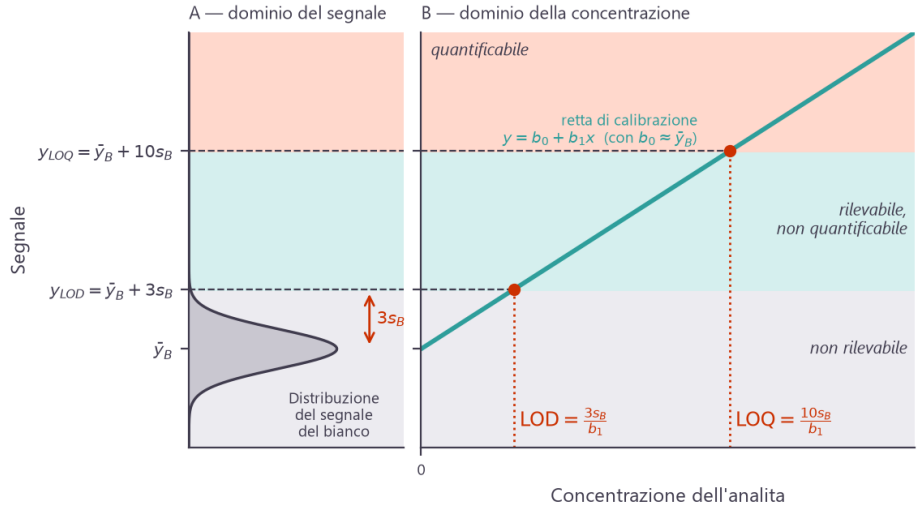


Figura 11.1 — LOD e LOQ. A: dominio del segnale; la distribuzione dei segnali del bianco definisce i segnali limite (media del bianco più 3 e più 10 deviazioni standard) e tre zone: non rilevabile, rilevabile ma non quantificabile, quantificabile. B: la retta di calibrazione converte i segnali limite nelle concentrazioni LOD e LOQ (formule 11.3 e 11.4).

In cromatografia si usa spesso il **rapporto segnale/rumore**: il LOD corrisponde a S/N ≈ 3, il LOQ a **S/N ≥ 10**. Le formule danno stime: il LOQ va poi verificato analizzando campioni a quella concentrazione.

Esempio 11.3 — LOD e LOQ da repliche del bianco

Dieci repliche di un bianco (HPLC-UV) danno aree (mAU·s): 152, 138, 165, 147, 171, 140, 158, 149, 133, 160. Pendenza b_1 = 1 250 mAU·s per µg/mL.

1. $\bar{y}_B$ = 1 513/10 = 151,3 mAU·s; s_B = 12,3 mAU·s ($n - 1 = 9$).
2. Segnali limite: y_{LOD} = 151,3 + 3 · 12,3 = 188,2 mAU·s; y_{LOQ} = 151,3 + 10 · 12,3 = 274,3 mAU·s.
3. LOD = 3 · 12,3/1 250 = 0,0295 µg/mL ≈ 30 ng/mL.
4. LOQ = 10 · 12,3/1 250 = 0,098 µg/mL ≈ 98 ng/mL.
5. Un campione con segnale 230 mAU·s cade tra y_{LOD} e y_{LOQ}: rilevato, ma non quantificabile.

Risultato: LOD ≈ 30 ng/mL e LOQ ≈ 98 ng/mL. I valori 188,2 e 274,3 sono segnali: confonderli con LOD e LOQ è l'errore più comune.

Un risultato sotto il LOD si riporta come «non rilevato»; tra LOD e LOQ come «rilevato, < LOQ»; sopra il LOQ come valore numerico. Il LOQ non va confuso con lo standard più alto della curva (ULOQ).

In bioanalisi il riferimento operativo è il **LLOQ**, il calibrante più basso, al quale ICH M10 richiede esattezza entro **±20%** e **CV ≤ 20%**. **Il LLOQ non può mai stare sotto il LOQ**. Se uno studio richiede concentrazioni più basse del LOQ, non si abbassa il LLOQ né si allargano i criteri: bisogna migliorare il metodo (più segnale, meno rumore, estrazione più pulita, campione più concentrato).

11.6 Range dinamico lineare e intervallo di calibrazione

Il **range dinamico lineare** è l'intervallo in cui il segnale è linearmente correlato alla concentrazione: va dal **LOQ** al **LOL** (limite di linearità). Con intercetta non nulla la relazione è lineare ma non proporzionale. Oltre il LOL il rivelatore si satura e la risposta si appiattisce. L'ampiezza dipende da analita e rivelatore; la linearità permette la regressione ai minimi quadrati con pochi standard (Capitoli 7 e 8).

L'**intervallo di calibrazione** (ICH M10) va dal **LLOQ** (standard più basso, mai sotto il LOQ) all'**ULOQ** (standard più alto). Si costruisce con un **campione bianco** (matrice senza analita né SI), un **campione zero** (matrice con il solo SI) e **almeno 6 livelli** di calibranti (Capitolo 10). Per un modello lineare deve stare dentro il range lineare ($\mathrm{LOQ} \leq \mathrm{LLOQ}$ e $\mathrm{ULOQ} \leq \mathrm{LOL}$): se l'ULOQ superasse il LOL, i calibranti più alti cadrebbero nel tratto curvo.

ICH M10 ammette anche modelli non lineari giustificati. I **QC** coprono tutto l'intervallo: LLOQ, basso (≤ 3 × LLOQ), medio (30–50%), alto (≥ 75% dell'ULOQ) (Figura 11.2; criteri nel Capitolo 12).

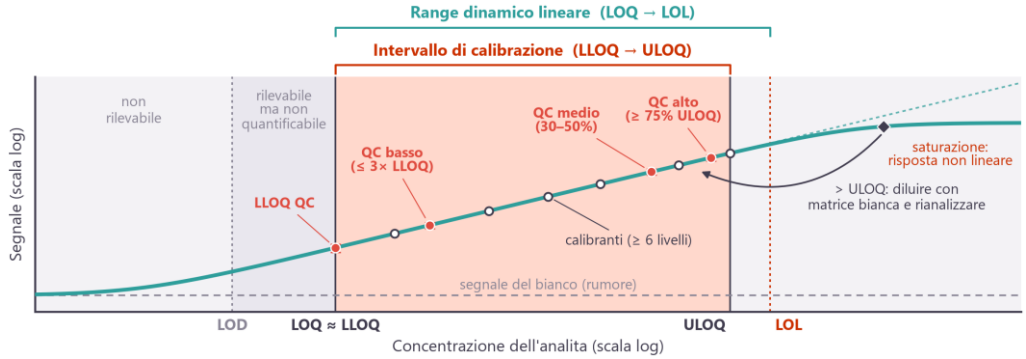


Figura 11.2 — Range dinamico lineare (dal LOQ al LOL) e intervallo di calibrazione (dal LLOQ all'ULOQ, con almeno 6 calibranti); posizione dei QC; oltre il LOL la risposta si appiattisce per saturazione. Assi in scala logaritmica.

Esattezza e precisione sono dimostrate solo tra LLOQ e ULOQ, quindi **non si estrapola mai** oltre la curva, anche se il rivelatore sarebbe ancora lineare:

- **sopra l'ULOQ** il campione si diluisce con **matrice bianca** (non con solvente, perché deve comportarsi come i calibranti) e si rianalizza; la diluizione deve essere validata (integrità della diluizione, Capitolo 14). Per esempio, con ULOQ = 200 ng/mL, un campione diluito 1 + 4 che dà 51,8 ng/mL contiene 51,8 · 5 = 259 ng/mL;
- **sotto il LLOQ** il risultato si riporta come **«< LLOQ»** (per esempio «< 0,50 ng/mL»), non come valore estrapolato né come zero, perché l'analita può essere presente anche se non quantificabile; né si può aggiungere a posteriori un calibrante più basso.

Punti chiave del Capitolo 11

- La validazione dimostra con prove documentate che un metodo è adatto allo scopo; la linea guida dipende dal settore (ICH M10, ICH Q2(R2)/Q14, Eurachem, SANTE, Reg. (UE) 2021/808). Un r^2 elevato non è una validazione.
- Selettività: bianchi da almeno 6 lotti individuali; interferenza ≤ 20% della risposta dell'analita al LLOQ e ≤ 5% della risposta dello SI.
- Sensibilità di calibrazione = pendenza b_1; sensibilità analitica = b_1/s_S, variabile con la concentrazione; nessuna delle due è il LOD.
- $\mathrm{LOD} = 3s_B/b_1$ (presenza/assenza, S/N ≈ 3) e $\mathrm{LOQ} = 10s_B/b_1$ (quantificazione, S/N ≥ 10); $\bar{y}_B + 3s_B$ e $\bar{y}_B + 10s_B$ sono segnali.
- LLOQ: ±20%, CV ≤ 20%, mai sotto il LOQ.
- Range lineare: LOQ → LOL; intervallo di calibrazione: LLOQ → ULOQ, almeno 6 livelli più bianco e zero, dentro il range lineare per un modello lineare.
- Sopra l'ULOQ si diluisce con matrice bianca (diluizione validata); sotto il LLOQ si riporta «< LLOQ»; non si estrapola.

Mettiti alla prova

Argomenti del quiz: Validazione: scopo e parametri; Selettività e specificità; Sensibilità, LOD e LOQ; Range dinamico lineare e intervallo di calibrazione.

12. Esattezza, precisione e controllo qualità

Esattezza e precisione dicono quanto ci si può fidare di un risultato: quanto la media è vicina al valore vero e quanto i risultati ripetuti sono dispersi. Questo capitolo mostra come si misurano, quali criteri fissa ICH M10 (2022) e come i campioni di controllo qualità (QC) le verificano in tutto l'intervallo di calibrazione. Le definizioni generali sono nel Capitolo 3, deviazione standard e CV% nel Capitolo 4.

12.1 Esattezza e bias

Esattezza (trueness) e bias

L'**esattezza** è il grado di concordanza tra la **media** di un numero elevato di risultati e un valore di riferimento μ (ISO 5725, VIM). Si esprime con il **bias**, differenza tra media e valore di riferimento.

Il valore di riferimento è la concentrazione nominale di un QC (matrice bianca arricchita con una quantità nota di analita), il valore certificato di un CRM (Capitolo 9) o il risultato di un metodo di riferimento.

Bias assoluto e relativo

$$\text{bias} = \bar{x} - \mu \tag{12.1}$$

$$\text{bias}\ (\%) = \frac{\bar{x} - \mu}{\mu} \times 100 \tag{12.2}$$

- $\bar{x}$: media delle repliche; μ: valore nominale o certificato.
- Il segno conta: negativo = sottostima, positivo = sovrastima.

Il bias si calcola sulla **media**: la differenza $x_i - \mu$ di un singolo risultato è il suo errore di misura, che comprende anche l'errore casuale. Anche il criterio di ICH M10 si applica alla media delle repliche: bias entro **±15%** del nominale a tutti i livelli QC, **±20% al LLOQ**. ICH M10 e ICH Q2(R2) chiamano *accuracy* proprio questa esattezza, espressa anche come percentuale del nominale ($\bar{x}/\mu \times 100$: 92,4% del nominale equivale a un bias del −7,6%).

Esempio 12.1 — Bias di un QC alto

Un QC alto (nominale 400 ng/mL) dà in 5 repliche: 372, 365, 380, 358, 370 ng/mL.

1. $\bar{x}$ = 1 845/5 = 369,0 ng/mL.
2. Bias = 369,0 − 400 = −31,0 ng/mL.
3. Bias% = −31,0/400 × 100 = −7,75% ≈ −7,8%, entro ±15%.

Risultato: bias = −7,8% (sottostima), conforme. Dividendo per il valore misurato (369,0) si otterrebbe −8,4%: il denominatore è sempre il valore di riferimento. Altri errori tipici sono perdere il segno, confondere il bias assoluto (in ng/mL) con quello relativo (in %) e calcolare il bias su un singolo risultato invece che sulla media.

12.2 Accuratezza = esattezza + precisione

Secondo ISO 5725 e il VIM l'**accuratezza** è il termine generale che descrive la vicinanza di un singolo risultato al valore vero: poiché l'errore di un risultato è la somma di bias ed errore casuale, l'accuratezza comprende sia l'esattezza (assenza di bias) sia la precisione (assenza di errore casuale rilevante). Non è sinonimo di nessuna delle due. Per esempio, cinque misure di un CRM certificato a 25,0 mg/kg danno 22,9; 23,1; 23,0; 22,8; 23,2 mg/kg: media 23,0 mg/kg, s = 0,16 mg/kg, bias −8,0%. I risultati sono **precisi ma non esatti**: una dispersione così piccola non spiega uno scostamento di 2 mg/kg, che è sistematico. Il metodo quindi non è accurato.

12.3 Precisione: ripetibilità, precisione intermedia, riproducibilità

La precisione è la concordanza tra risultati indipendenti in condizioni specificate e si esprime come CV% di misure replicate (non come bias%, r² o S/N):

$$\mathrm{RSD\%} = \mathrm{CV\%} = \frac{s}{\bar{x}} \times 100 \tag{12.3}$$

con s calcolata con $n - 1$ al denominatore. Le tre condizioni di precisione (Capitolo 3) sono:

- **ripetibilità** (intra-day): stessa procedura, stesso operatore, stesso strumento, stesso giorno; in ICH M10 è la precisione *within-run* (nella stessa sequenza);
- **precisione intermedia** (inter-day): stesso laboratorio, ma giorni, operatori e strumenti diversi; descrive la variabilità nella routine ed è la precisione *between-run* di ICH M10. Un QC analizzato in 5 giorni da due analisti su due strumenti dà la precisione intermedia;
- **riproducibilità** (inter-laboratorio): laboratori diversi con la stessa procedura; è la base per la standardizzazione e il riconoscimento normativo del metodo.

Ogni condizione aggiunge fonti di variabilità, quindi ci si attende $\mathrm{CV}_r \leq \mathrm{CV}_I \leq \mathrm{CV}_R$ (ripetibilità, intermedia, riproducibilità).

ICH M10 ragiona per sequenze (*run*), non strettamente per giorni: la precisione within-run si valuta con almeno **5 repliche** per livello QC in ogni sequenza, quella between-run con ogni livello in almeno **3 sequenze** in almeno **2 giorni**. Il criterio vale per **entrambe**: **CV ≤ 15%**, **≤ 20% al LLOQ**. Un QC medio con CV del 5% within-run e del 17% between-run non è conforme: i due CV non si mediano, e il 20% vale solo al LLOQ.

Esempio 12.2 — RSD% di sei repliche

Sei repliche di un QC medio (nominale 250 ng/mL): 241, 258, 236, 262, 249, 244 ng/mL.

1. $\bar{x}$ = 1 490/6 = 248,3 ng/mL.
2. $\sum(x_i - \bar{x})^2$ = 505,3; $s = \sqrt{505{,}3/5}$ = 10,05 ng/mL.
3. RSD% = 10,05/248,3 × 100 = 4,0%.

Risultato: RSD% = 4,0%, entro il 15%. Errori tipici: dividere per n (3,7%), usare $s/\sqrt{n}$ (1,7%), dimenticare il fattore 100 (0,040%).

12.4 La curva di Horwitz

Dall'analisi di numerosi studi interlaboratorio, Horwitz osservò che la **riproducibilità interlaboratorio** peggiora in modo prevedibile al diminuire della concentrazione, quasi indipendentemente da analita, matrice e tecnica (Figura 12.1).

Equazione di Horwitz

$$\mathrm{CV}_R\,(\%) = 2^{(1-0{,}5\log C)} \qquad (12.4)$$

- CV_R: CV di riproducibilità interlaboratorio atteso;
- C: **frazione di massa adimensionale** (1 = 100%; 10^{-6} = 1 mg/kg; 10^{-9} = 1 µg/kg);
- $\log$: logaritmo **decimale**.

Ogni volta che la concentrazione diminuisce di un fattore 100, l'esponente aumenta di 1 e il CV atteso raddoppia: 2% per un analita puro (C = 1), 4% all'1%, 8% a 100 mg/kg, 16% a 1 mg/kg, circa 45% a 1 µg/kg. Per esempio, a 100 mg/kg: $C = 10^{-4}$, $\log C$ = −4, $\mathrm{CV}_R = 2^{(1+2)}$ = 8%. A 0,5 mg/kg: $C = 5 \times 10^{-7}$, $\log C$ = −6,301, $\mathrm{CV}_R = 2^{4{,}151}$ = 17,8%. La curva serve soprattutto come termine di paragone: negli studi interlaboratorio il rapporto **HorRat** tra il CV di riproducibilità osservato e quello previsto è considerato accettabile tra circa 0,5 e 2 (per esempio, 12% osservato a 0,5 mg/kg dà HorRat = 12/17,8 ≈ 0,7).

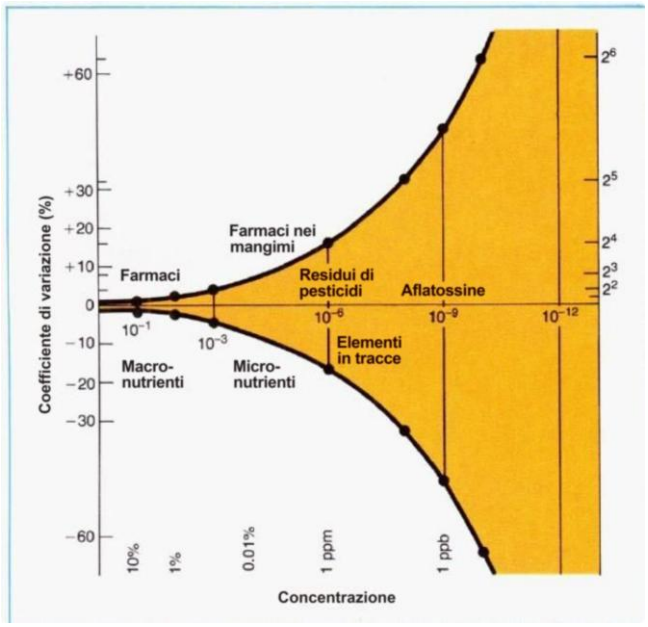


Figura 12.1 — La curva di Horwitz: CV interlaboratorio atteso (con il segno ±) in funzione della concentrazione, espressa come frazione di massa in scala logaritmica; il CV raddoppia ogni due ordini di grandezza (scala di destra: 2^2, 2^3, …, 2^6 %).

Come non usare la curva di Horwitz

- C va espresso come frazione di massa: inserire 100 (per 100 mg/kg) darebbe 1% invece di 8%; dimenticare il fattore 0,5 darebbe 32%.
- La curva descrive la **riproducibilità interlaboratorio**, non la ripetibilità né la precisione intermedia di un singolo laboratorio, che sono di norma più piccole. Il valore del 2% vale solo per $C \approx 1$.
- Il suo andamento è **coerente** con i criteri più larghi di ICH M10 al LLOQ (±20%, CV ≤ 20%), perché la precisione peggiora vicino al limite inferiore; ma i criteri ICH M10 sono intra-laboratorio e non derivano dall'equazione.
- Sotto circa $1{,}2 \times 10^{-7}$ (120 µg/kg) l'equazione sovrastima il CV osservato.

12.5 I campioni di controllo qualità (QC)

I **QC** sono campioni di matrice bianca arricchiti con una quantità nota di analita, preparati da una **soluzione madre diversa** da quella dei calibranti (così un errore nella preparazione degli standard non resta nascosto) e trattati come i campioni reali. **Non servono a costruire la curva**: la mettono alla prova. Si usano in validazione e, nella routine, per accettare o respingere ogni sequenza (Capitolo 16).

Tabella 12.1 — Livelli dei QC e criteri ICH M10 (metodi cromatografici)

QC	Posizione	Esattezza	Precisione
QC al LLOQ	al limite inferiore di quantificazione	±20%	CV ≤ 20%
QC basso	$\leq 3 \times$ LLOQ	±15%	CV ≤ 15%
QC medio	30–50% dell'intervallo di calibrazione	±15%	CV ≤ 15%
QC alto	$\geq 75\%$ dell'ULOQ	±15%	CV ≤ 15%

I QC si riferiscono all'intervallo di calibrazione (LLOQ–ULOQ), non al LOD né al LOL; per un modello lineare anche il QC alto deve cadere nel tratto lineare (Figura 11.2). Il QC basso controlla la zona in cui rumore ed effetto matrice cominciano a pesare, il QC alto quella in cui compaiono saturazione e non linearità, il QC medio il centro; poiché di solito il LLOQ è molto più piccolo dell'ULOQ, il 30–50% dell'intervallo corrisponde in pratica a circa il 30–50% dell'ULOQ.

Solo il QC al LLOQ ha criteri allargati: vicino al limite inferiore il segnale supera di poco il rumore e l'errore relativo cresce. Il LLOQ però non è il LOD: anche lì si richiede una quantificazione affidabile. Il QC alto, pur vicino alla saturazione, non ha criteri allargati. I criteri valgono per l'esattezza e la precisione sia within-run sia between-run. Nella routine si analizzano i QC basso, medio e alto almeno in duplicato, e la sequenza è accettata se almeno due terzi di tutti i QC, e almeno metà di quelli di ogni livello, sono entro ±15% del nominale (Capitolo 16).

Esempio 12.3 — I QC soddisfano ICH M10?

LLOQ = 1,00 ng/mL, ULOQ = 500 ng/mL. Medie e CV di 5 repliche: QC al LLOQ 1,00 → 1,17 ng/mL (CV 18,0%); QC basso 2,50 → 2,11 ng/mL (CV 8,5%); QC medio 200 → 212 ng/mL (CV 15,0%); QC alto 400 → 372 ng/mL (CV 6,2%).

1. Posizioni: 2,50 ≤ 3 × 1,00; 200/500 = 40%; 400/500 = 80% ≥ 75%: corrette.
2. LLOQ: (1,17 − 1,00)/1,00 × 100 = +17,0% (≤ 20%), CV 18,0% (≤ 20%): conforme.
3. Basso: (2,11 − 2,50)/2,50 × 100 = −15,6%: oltre ±15%, non conforme.
4. Medio: +6,0%; CV 15,0%, che non supera il 15%: conforme.
5. Alto: (372 − 400)/400 × 100 = −7,0%, CV 6,2%: conforme.

Risultato: il QC basso non è conforme (−15,6%); la causa della sottostima va indagata. Applicargli il ±20% «perché è vicino al LLOQ» è un errore (Figura 12.2).

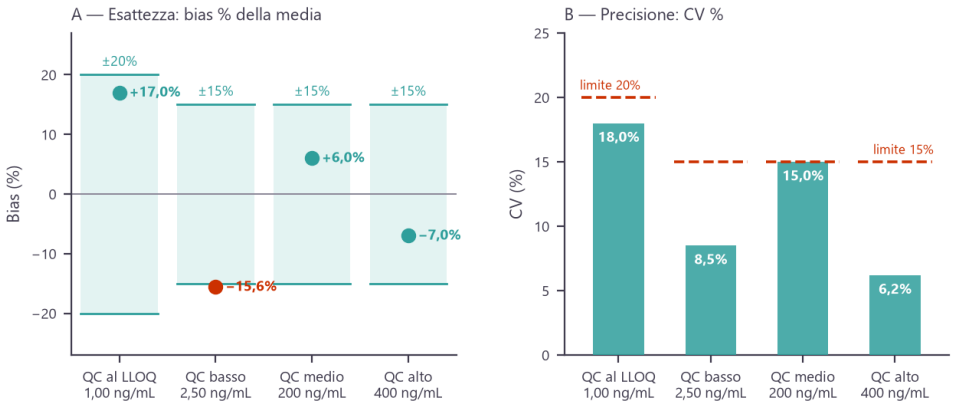


Figura 12.2 — Valutazione dei QC dell'Esempio 12.3. A: bias % della media rispetto ai limiti (±20% al LLOQ, ±15% agli altri livelli); il QC basso è fuori. B: CV % rispetto ai limiti (20% al LLOQ, 15% altrove).

12.6 Prestazioni ai livelli bassi e alti

Vicino al LLOQ il segnale è piccolo e pesano di più: rumore strumentale elevato rispetto al segnale, deriva della linea di base, effetto matrice più marcato, interferenze co-eluenti o co-estratte, instabilità o adsorbimento dell'analita. Il bias tende quindi ad aumentare avvicinandosi al LLOQ e la precisione peggiora. Un errore costante pesa molto di più in basso: un'interferenza equivalente a +0,08 ng/mL dà +16% al LLOQ di 0,50 ng/mL (0,08/0,50 × 100) ma +0,05% a un QC alto di 160 ng/mL, mentre un errore proporzionale darebbe lo stesso bias% a tutti i livelli.

Vicino all'ULOQ pesano saturazione del rivelatore e deviazioni dalla linearità (la risposta cresce meno che proporzionalmente: bias negativo) e diluizioni imprecise; il bias può quindi essere positivo o negativo. Un QC alto con bias del −18%, mentre gli altri sono entro ±5% e i calibranti più alti crescono meno del previsto, indica saturazione vicino all'ULOQ.

Il **carry-over** nasce dai campioni concentrati ma falsa soprattutto i campioni bassi iniettati dopo: un campione vicino al LLOQ sovrastimato del 30% dopo un campione molto concentrato, e corretto se reiniettato dopo un bianco, ne è il segno tipico. Si valuta iniettando un bianco dopo l'ULOQ (Capitolo 14). Alcune fonti di imprecisione, come la variabilità dell'iniezione o l'instabilità della fase mobile, agiscono invece a tutti i livelli.

12.7 Coerenza su tutto l'intervallo

Un metodo valido deve essere:

- **stabile**: esattezza, precisione e recupero non cambiano significativamente da un livello di concentrazione all'altro (non si parla qui della stabilità chimica dell'analita, Capitolo 14);
- **affidabile agli estremi**: prestazioni accettabili anche al LLOQ e all'ULOQ;
- **costante nel tempo**: risultati concordi tra giorni e operatori (precisione intermedia);
- **corretto su tutta la curva**: una curva «perfetta» con QC che falliscono ai livelli bassi indica un metodo **non valido**.

Per esempio, con $r^2 = 0{,}9995$ e calibranti entro ±10%, un QC basso con bias costante di circa −22% rende il metodo non valido: al QC basso il limite è ±15%, e r^2, dominato dai punti alti, non rivela errori relativi ai livelli bassi. Per questo i QC devono coprire tutto l'intervallo e non solo il centro della curva.

Punti chiave del Capitolo 12

- Esattezza = bias della **media**: bias% = $(\bar{x} - \mu)/\mu \times 100$, con il segno; ICH M10: ±15% (±20% al LLOQ). In ICH M10 *accuracy* indica questa esattezza.
- Accuratezza = esattezza + precisione (ISO 5725, VIM).
- Precisione = CV% (con $n - 1$); ripetibilità (stesse condizioni), precisione intermedia (stesso laboratorio, giorni/operatori/strumenti diversi), riproducibilità (laboratori diversi); $\mathrm{CV}_r \leq \mathrm{CV}_I \leq \mathrm{CV}_R$.
- ICH M10: CV ≤ 15% (≤ 20% al LLOQ) sia within-run sia between-run.
- Horwitz: $\mathrm{CV}_R = 2^{(1-0{,}5\log C)}$, con C frazione di massa e log decimale; descrive la riproducibilità interlaboratorio.
- QC: LLOQ (±20%, CV ≤ 20%), basso ≤ 3 × LLOQ, medio 30–50%, alto ≥ 75% dell'ULOQ (±15%, CV ≤ 15%); soluzione madre separata, non entrano nella curva.
- In basso pesano rumore, effetto matrice, interferenze e carry-over; in alto saturazione, non linearità e diluizioni. Il metodo deve essere coerente su tutto l'intervallo.

Mettiti alla prova

Argomenti del quiz: Esattezza, precisione e accuratezza; Livelli QC e criteri ICH M10; Recupero, esattezza e precisione nel range (domande su esattezza, precisione e coerenza ai diversi livelli).

13. Recupero, effetto matrice ed efficienza di processo

Il segnale dell'analita in un campione reale non coincide quasi mai con quello della stessa quantità in solvente puro: una parte dell'analita si perde durante la preparazione del campione, e i componenti della matrice ne alterano la risposta nello strumento. Separare i due contributi serve a capire dove intervenire e a dimostrare che il risultato resta corretto in ogni campione.

13.1 Definizioni e schema di Matuszewski

Recupero, effetto matrice ed efficienza di processo

- **Recupero (recovery, RE)**: frazione dell'analita che supera le fasi di estrazione e trattamento del campione; descrive le perdite di estrazione.
- **Effetto matrice (matrix effect)**: alterazione della risposta dell'analita dovuta a componenti interferenti, spesso non identificati, della matrice (ICH M10). Si quantifica con il **fattore di matrice (matrix factor, MF)**.
- **Efficienza di processo (process efficiency, PE)**: effetto complessivo delle perdite di estrazione e dell'effetto matrice.

Per separarli si usa lo schema di Matuszewski (2003): tre serie di campioni (*set*) alla **stessa concentrazione nominale finale** nel vial (Figura 13.1).

- **Set 1 — soluzione standard**: analita nel solvente puro, senza matrice e senza estrazione. È il segnale di riferimento.
- **Set 2 — arricchimento dopo l'estrazione** (*post-extraction spike*): matrice bianca processata, con l'analita aggiunto all'estratto finale; contiene la matrice, ma l'analita non ha subito l'estrazione.
- **Set 3 — arricchimento prima dell'estrazione** (*pre-extraction spike*): matrice bianca arricchita e poi processata come un campione reale; subisce estrazione ed effetto matrice.

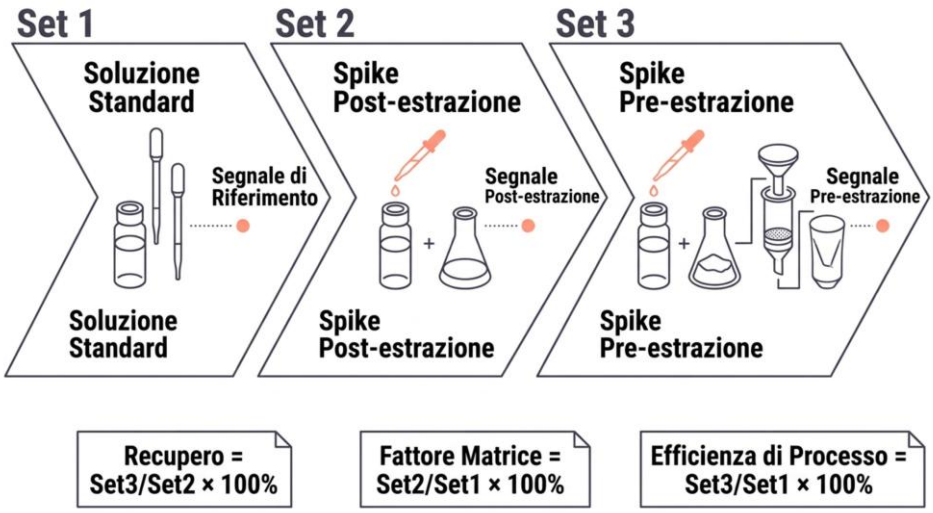


Figura 13.1 — Schema di Matuszewski. Set 1: soluzione standard in solvente puro (segnale di riferimento). Set 2: estratto di matrice bianca arricchito dopo l'estrazione. Set 3: matrice bianca arricchita prima dell'estrazione.

13.2 Le formule

Recupero, fattore di matrice ed efficienza di processo

$$\mathrm{RE}\% = \frac{A_3}{A_2} \times 100 \tag{13.1}$$

$$\mathrm{MF}\% = \frac{A_2}{A_1} \times 100 \tag{13.2}$$

$$\mathrm{PE}\% = \frac{A_3}{A_1} \times 100 \tag{13.3}$$

- A_1, A_2, A_3: aree medie del picco dell'analita nel Set 1, nel Set 2 e nel Set 3.
- MF < 100%: soppressione ionica (*ion suppression*); MF > 100%: potenziamento ionico (*ion enhancement*).

- **RE = Set 3/Set 2.** Entrambi contengono lo stesso estratto di matrice e differiscono solo per il momento dell'aggiunta dell'analita: l'effetto matrice si elide e resta la perdita di estrazione. Il recupero non è quindi "l'analita ritrovato dopo l'intero processo" (questa è la PE).
- **MF = Set 2/Set 1.** Nessuno dei due ha subito l'estrazione; differiscono solo per la presenza della matrice.
- **PE = Set 3/Set 1.** Comprende entrambi gli effetti.

Le tre grandezze sono legate: moltiplicando e dividendo per A_2,

$$\frac{A_3}{A_1} = \frac{A_3}{A_2} \times \frac{A_2}{A_1} \Rightarrow \mathrm{PE} = \mathrm{RE} \times \mathrm{MF} \tag{13.4}$$

con le grandezze espresse come frazioni: i due effetti agiscono in sequenza, quindi si moltiplicano.

Esempio 13.1 — RE, MF e PE da aree dei picchi

Aree medie al QC basso: Set 1 = 410 000; Set 2 = 307 500; Set 3 = 264 450.

1. Recupero: $\mathrm{RE}\% = 264\,450/307\,500 \times 100 = 86{,}0\%$.
2. Fattore di matrice: $\mathrm{MF}\% = 307\,500/410\,000 \times 100 = 75{,}0\%$ (soppressione del 25%).
3. Efficienza di processo: $\mathrm{PE}\% = 264\,450/410\,000 \times 100 = 64{,}5\%$.
4. Verifica: $0{,}860 \times 0{,}750 = 0{,}645$.

Risultato: RE = 86,0%, MF = 75,0%, PE = 64,5%. La perdita di segnale dipende più dalla matrice che dall'estrazione: conviene migliorare la pulizia dell'estratto o la cromatografia. Noti PE e MF, si ricava RE = PE/MF.

Errori comuni

- Invertire i rapporti: Set 1/Set 2 trasforma una soppressione (MF = 80%) in un apparente potenziamento (125%).
- Scambiare il recupero (Set 3/Set 2) con l'efficienza di processo (Set 3/Set 1).
- Sommare o dividere RE e MF invece di moltiplicarli.

13.3 L'effetto matrice in LC-MS/MS

Nella sorgente a ionizzazione elettrospray (ESI) gli ioni dell'analita si formano da goccioline cariche, con carica e superficie limitate: i componenti della matrice che **co-eluiscono** con l'analita **competono con esso per la ionizzazione**, e quelli poco volatili (sali, fosfolipidi) ostacolano l'evaporazione delle gocce. Il caso più frequente è la soppressione (MF < 100%); più raramente la matrice favorisce la ionizzazione (MF > 100%). La ionizzazione chimica a pressione atmosferica (APCI), che avviene in fase gassosa, è in genere meno sensibile, ma non immune.

L'effetto matrice è quindi una **variazione del segnale non dovuta alla concentrazione dell'analita**: le molecole nel vial sono le stesse, ma cambia la frazione che diventa ione. È di solito **moltiplicativo** (cambia la pendenza) e va distinto da un'interferenza **additiva**, che dà un segnale alla m/z e al tempo di ritenzione dell'analita anche nel bianco (selettività, Capitolo 11). Responsabili tipici sono i fosfolipidi del plasma, i sali, gli eccipienti, i campioni emolizzati o lipemici; poiché l'effetto cambia da un individuo all'altro, si valuta su più lotti di matrice (paragrafo 13.5).

Le conseguenze dipendono dalla calibrazione (Capitolo 10). Con **calibranti in solvente e senza standard interno** i campioni subiscono l'effetto e i calibranti no: con MF < 100% la concentrazione è **sottostimata**, con MF > 100% **sovrastimata**, all'incirca dello stesso fattore se il recupero è completo (MF = 130% dà circa +30%). Con **calibranti in matrice** l'effetto medio è compensato; con uno **standard interno (SI)** conta l'effetto che resta sul rapporto dei segnali (paragrafo 13.4). Anche se compensata, una forte soppressione riduce il rapporto segnale/rumore e peggiora il LLOQ.

13.4 Il fattore di matrice normalizzato sullo standard interno

Fattore di matrice normalizzato sullo SI

$$\mathrm{MF}_{\mathrm{norm}} = \frac{\mathrm{MF}_{\mathrm{analita}}}{\mathrm{MF}_{\mathrm{SI}}} \tag{13.5}$$

- $\mathrm{MF}_{\mathrm{analita}}$, $\mathrm{MF}_{\mathrm{SI}}$: area nell'estratto arricchito dopo l'estrazione / area in solvente, per l'analita e per lo SI.
- $\mathrm{MF}_{\mathrm{norm}} \approx 1$: lo SI subisce lo stesso effetto dell'analita e lo compensa nel rapporto.

Con lo SI la concentrazione si ricava dal rapporto dei segnali: conta l'effetto matrice che resta su quel rapporto. Se lo SI arriva in sorgente con gli stessi componenti co-eluiti, subisce la stessa soppressione e il MF normalizzato è vicino a 1. Lo SI ideale è lo **standard interno marcato con isotopi stabili (SIL-IS)**, che co-eluisce con l'analita e si distingue solo per la m/z (ICH M10 lo raccomanda quando possibile); un analogo deuterato può eluire leggermente prima e compensare in modo imperfetto, un analogo strutturale eluisce altrove e può subire un effetto diverso. La normalizzazione non elimina la matrice e non riguarda le perdite di estrazione.

Esempio 13.2 — MF normalizzato con un SIL-IS e con un analogo strutturale

Aree in solvente e in un estratto di plasma bianco arricchito dopo l'estrazione: analita 320 000 e 208 000; SIL-IS 150 000 e 98 250; analogo strutturale 210 000 e 197 400.

1. $\mathrm{MF}_{\mathrm{analita}} = 208\,000/320\,000 = 0{,}650$ (soppressione del 35%).
2. $\mathrm{MF}_{\mathrm{SI}}$ = 98 250/150 000 = 0,655 per il SIL-IS; 197 400/210 000 = 0,940 per l'analogo.
3. $\mathrm{MF}_{\mathrm{norm}}$ = 0,650/0,655 = 0,99 con il SIL-IS; 0,650/0,940 = 0,69 con l'analogo.

Risultato: con il SIL-IS la soppressione, pur marcata, è compensata (0,99); con l'analogo sul rapporto resta un effetto del −31%, che si trasferirebbe sul risultato.

13.5 La valutazione nelle linee guida

ICH M10 (2022) valuta l'effetto matrice sul risultato finale: QC basso e QC alto preparati in matrice da **almeno 6 fonti o lotti diversi**, in almeno 3 repliche, quantificati con la normale curva di calibrazione. Per **ciascun lotto** l'esattezza deve essere entro **±15%** del valore nominale e il **CV ≤ 15%**. La media sui lotti non conta: se un lotto (per esempio lipemico, +24%) è fuori, la prova non è superata, perché serve proprio a rivelare le differenze tra matrici. Un lotto anomalo non si esclude a posteriori: M10 raccomanda anzi, caso per caso, prove su matrici emolizzate o lipemiche e su popolazioni speciali; per matrici rare bastano meno lotti. Da non confondere con la selettività: bianchi da 6 lotti, segnale ≤ 20% della risposta al LLOQ (Capitolo 11).

La **linea guida EMA (2011)**, sostituita da M10 dal 2023, calcolava invece in 6 lotti il MF dell'analita e dello SI, a una concentrazione bassa e a una alta, e richiedeva che il **CV del MF normalizzato sullo SI** fosse ≤ 15%. Il MF resta lo strumento più usato durante lo sviluppo del metodo.

13.6 Individuare e ridurre l'effetto matrice

Per vedere **dove** si trovano le zone di soppressione lungo il cromatogramma si usa l'**infusione post-colonna** (Figura 13.2): una pompa a siringa infonde in continuo l'analita, tramite un raccordo a T, tra l'uscita della colonna e la sorgente, mentre si inietta un estratto di matrice bianca. Il segnale dell'analita infuso dovrebbe restare costante; un **avvallamento** a un certo tempo indica che lì co-eluiscono componenti che sopprimono la ionizzazione, un rialzo un potenziamento. L'analita infuso non passa dalla colonna, e carry-over o contaminazione darebbero un aumento del segnale, non una diminuzione.

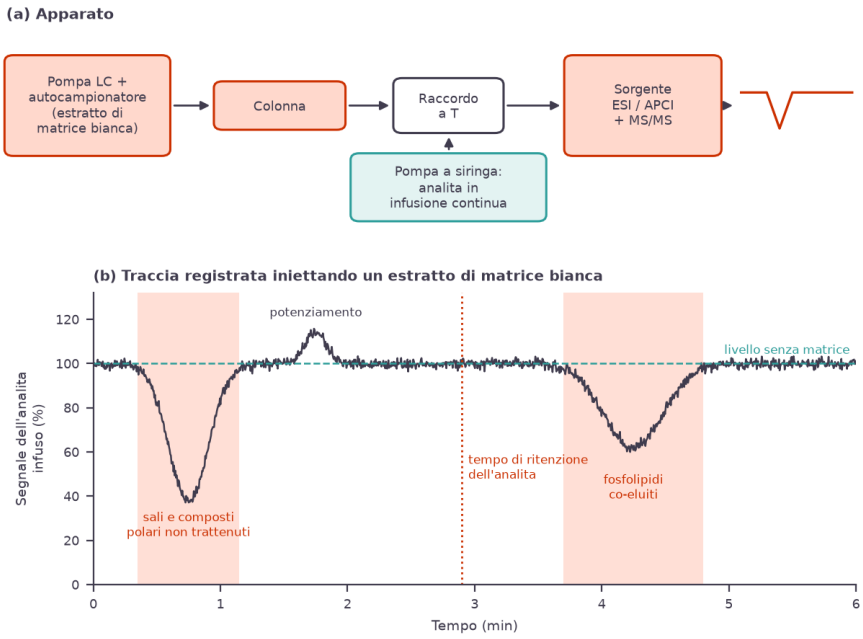


Figura 13.2 — Infusione post-colonna. (a) L'analita è infuso in continuo tra colonna e sorgente mentre si inietta un estratto di matrice bianca. (b) Gli avvallamenti indicano zone di soppressione (sali all'inizio della corsa, fosfolipidi), il rialzo una zona di potenziamento; l'analita dovrebbe eluire dove il segnale resta al livello senza matrice.

Per **ridurre l'effetto alla fonte** (meno componenti co-eluiti in sorgente) si rende più selettiva la pulizia del campione (la precipitazione delle proteine, PPT, lascia fosfolipidi e sali; l'estrazione liquido-liquido, LLE, e soprattutto in fase solida, SPE, danno estratti più puliti), si migliora la separazione cromatografica, si diluisce l'estratto o si passa da ESI ad APCI. Per **compensarlo**: SIL-IS, calibranti in matrice, metodo delle aggiunte multiple (aggiunte standard). Il SIL-IS non riduce l'effetto matrice: lo subisce come l'analita e lo annulla nel rapporto.

13.7 Il recupero ai diversi livelli di concentrazione

La prova verifica che l'estrazione funzioni **in modo equivalente a tutti i livelli QC** (basso, medio, alto). ICH M10 tratta il recupero tra le considerazioni aggiuntive (§7.3), come aspetto dello sviluppo del metodo: si valuta per analita e SI, tipicamente a tre concentrazioni, e **non deve essere del 100%, ma costante e riproducibile**. Una perdita costante colpisce allo stesso modo calibranti e campioni ed è compensata dalla calibrazione; un recupero che cambia con la concentrazione produce bias agli estremi. Un profilo 62%, 60%, 63% è accettabile; 55%, 78%, 95% no, anche se la media è più alta; un recupero misurato al solo QC medio non dimostra la costanza.

Le cause di variazione sono diverse ai due estremi:

- **livelli bassi** (perdite quasi fisse, che in percentuale pesano di più): adsorbimento sulle pareti dei contenitori, inefficienza dello SPE a basse concentrazioni, volatilizzazione durante l'evaporazione, instabilità dell'analita;
- **livelli alti**: saturazione della fase stazionaria SPE ed effetti non lineari; la co-estrazione di interferenti incide più sulla selettività e sull'effetto matrice che sul recupero.

Punti chiave del Capitolo 13

- Schema di Matuszewski: Set 1 = standard in solvente; Set 2 = estratto di matrice bianca arricchito dopo l'estrazione; Set 3 = matrice arricchita prima dell'estrazione.
- RE = Set 3/Set 2 (perdite di estrazione, l'effetto matrice si elide); MF = Set 2/Set 1 (effetto matrice); PE = Set 3/Set 1 = RE × MF.
- MF < 100%: soppressione; MF > 100%: potenziamento. Il meccanismo è la competizione per la ionizzazione, in sorgente ESI o APCI, tra analita e componenti co-eluiti; senza compensazione porta a sottostima o sovrastima.
- MF normalizzato = MF analita/MF SI; con un SIL-IS co-eluente è vicino a 1.
- ICH M10: QC basso e alto in ≥ 6 lotti, esattezza ±15% e CV ≤ 15% in ogni lotto. EMA 2011 (superata): CV del MF normalizzato ≤ 15%.
- L'infusione post-colonna mappa le zone di soppressione; pulizia del campione e cromatografia riducono l'effetto, SIL-IS e calibranti in matrice lo compensano.
- Il recupero va verificato ai QC basso, medio e alto: non serve il 100%, serve che sia costante e riproducibile.

Mettiti alla prova

Usa le domande del quiz sugli argomenti:

- **Recupero, effetto matrice ed efficienza di processo**
- **Effetto matrice**
- **Recupero, esattezza e precisione nel range** (le domande sul recupero ai diversi livelli)

14. Stabilità, robustezza, solidità, carry-over e integrità della diluizione

Un metodo esatto e preciso in validazione può fallire nella routine: l'analita si degrada mentre il campione aspetta, una piccola deriva del pH fa co-eluire un interferente, il metodo trasferito in un altro laboratorio non funziona, un campione concentrato contamina quello successivo, un campione sopra l'ULOQ va diluito. Le prove di questo capitolo verificano che il metodo resti affidabile in queste situazioni reali.

14.1 La stabilità dell'analita

Il risultato deve rappresentare la concentrazione presente nel campione **al momento del prelievo**. Fino all'iniezione, però, l'analita può degradarsi (idrolisi enzimatica, ossidazione, luce, interconversioni dipendenti dal pH) o adsorbirsi sulle pareti; nelle soluzioni stock l'evaporazione del solvente può invece aumentare la concentrazione. Le prove di stabilità riproducono, una per una, le condizioni a cui saranno esposti i campioni dello studio (Figura 14.1).

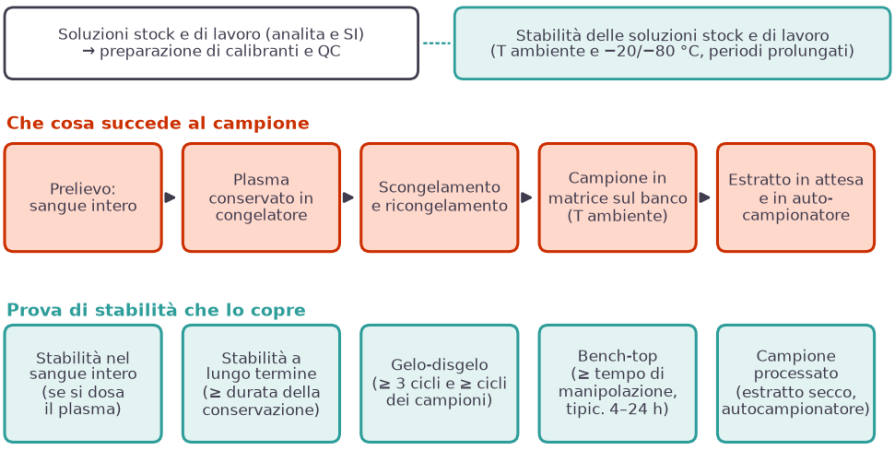


Figura 14.1 — Il percorso del campione e le prove di stabilità che lo coprono. Ogni fase in cui l'analita potrebbe degradarsi (in alto) corrisponde a una prova (in basso); la stabilità delle soluzioni stock e di lavoro si valuta a parte.

Regole generali di ICH M10 (2022):

- le condizioni della prova (temperature, tempi, matrice, anticoagulante, contenitori) devono **riflettere quelle dei campioni dello studio**; i dati di letteratura non bastano;
- in matrice si usano **QC basso e QC alto**, ciascuno suddiviso in almeno 3 aliquote, analizzate contro calibranti preparati al momento;
- **criterio**: la concentrazione media di ogni livello deve essere entro **±15% del valore nominale**.

Tabella 14.1 — Le prove di stabilità

Prova	Che cosa simula	Condizioni e durata
Soluzioni stock e di lavoro	conservazione e uso delle soluzioni di analita e SI	temperatura ambiente e −20 °C/−80 °C, per periodi prolungati; confronto con soluzioni fresche
Bench-top (breve termine)	campione **in matrice** sul banco, prima e durante il trattamento	temperatura del banco, almeno il tempo di manipolazione (tipicamente 4–24 h), continuativo
Cicli di gelo-disgelo	scongelamenti ripetuti	**almeno 3 cicli** e non meno di quelli che subiranno i campioni; ≥ 12 h di congelamento tra i cicli
Lungo termine	conservazione in congelatore	stessa temperatura dei campioni, durata **≥ conservazione** dei campioni (settimane, mesi)
Campione processato	estratto in attesa e nell'autocampionatore	condizioni d'uso, fino al termine dell'analisi
Sangue intero	intervallo tra prelievo e centrifugazione	se si analizza il plasma

Alcune precisazioni. Le **soluzioni stock** servono a preparare calibranti e QC per tutto lo studio: se si degradano, l'errore si trasferisce su tutti i risultati. La prova **bench-top** riguarda il campione in matrice, non l'estratto: la stabilità dell'estratto nell'autocampionatore è una prova distinta. I tempi devono essere **continuativi**: non si sommano, per esempio, le ore sul banco delle diverse prove di gelo-disgelo. Per il **gelo-disgelo**, un campione rianalizzato accumula cicli: se può subirne 5, si validano 5 cicli. Per il **lungo termine**, se i campioni resteranno 8 mesi a −80 °C, serve la stabilità a −80 °C per almeno 8 mesi (per i farmaci chimici la stabilità a −20 °C si può estendere a −70/−80 °C); nessun'altra prova la sostituisce. Se l'analita è instabile sul banco si lavora in bagno di ghiaccio, si aggiungono inibitori enzimatici o si limita il tempo di manipolazione a quello dimostrato.

14.2 Robustezza

Robustezza (robustness)

Misura della capacità di una procedura analitica di mantenere le prestazioni attese durante l'uso normale; si verifica con **piccole variazioni deliberate dei parametri del metodo** (ICH Q2(R2) e Q14, 2023). Domanda chiave: *il metodo rimane affidabile anche se qualcosa cambia di poco?*

Deliberate significa che l'analista introduce le variazioni di proposito, in modo pianificato (un parametro alla volta o con un disegno sperimentale), entro l'intervallo che può verificarsi in routine. Esempi tipici:

- temperatura della colonna ±2–5 °C; pH della fase mobile ±0,1;
- composizione della fase mobile (da 98/2 a 97/3); flusso ±0,1 mL/min;
- volume o tempo di estrazione, tempo di agitazione o centrifugazione;
- colonne di lotti e/o fornitori diversi (ICH Q2 le elenca tra le variazioni della robustezza);
- parametri impostabili del rivelatore: lunghezza d'onda UV, temperature, flussi dei gas e tensioni della sorgente MS.

Un metodo robusto mantiene esattezza, precisione, tempi di ritenzione e risoluzione entro limiti accettabili. Le variazioni devono essere **piccole**: pH ±2 unità, una colonna HILIC al posto di una C18 o un rivelatore diverso significano un metodo nuovo, non una prova di robustezza. Lo scopo è individuare i **parametri critici** e fissarne le tolleranze nella procedura e nei controlli di idoneità del sistema (per esempio, se con pH −0,1 la risoluzione da un interferente scende sotto 1,5, il pH è critico e va controllato più strettamente). Per questo la robustezza si studia **durante lo sviluppo** (ICH Q2(R2) e Q14); ICH M10 non la include tra gli elementi della validazione completa.

14.3 Solidità

La **solidità** (*ruggedness*) valuta se il metodo dà risultati simili quando cambiano **fattori di contesto**, che non sono parametri scritti nella procedura. Domanda chiave: *il metodo dà risultati simili se cambia chi lo usa, dove o come?* Esempi: strumentazione diversa ma equivalente, laboratori diversi (studi inter-laboratorio), operatori con diversa esperienza, reagenti e consumabili di marche diverse, giorni molto distanti, condizioni ambientali diverse. Se basta cambiare operatore o strumento per far crollare i risultati, il metodo non è solido e non può essere trasferito ad altri laboratori; per ICH M10 il trasferimento tra laboratori richiede una validazione parziale (Capitolo 16).

Il termine *ruggedness* veniva dalla farmacopea statunitense (USP); le linee guida ICH non lo usano, e oggi il concetto coincide con la **precisione intermedia** (stesso laboratorio, giorni, operatori o strumenti diversi) e con la **riproducibilità** (laboratori diversi), secondo la ISO 5725 (Capitoli 3 e 12). Alcuni esempi sono al confine: le colonne di lotti diversi sono per ICH una variazione di robustezza. In questo corso: **robustezza** = piccole variazioni deliberate dei parametri del metodo; **solidità** = cambiamenti del contesto. Un metodo può essere robusto ma non solido: supera le variazioni di flusso e di fase mobile, ma peggiora nettamente in un laboratorio partner, con altri strumenti e operatori.

14.4 Carry-over

Carry-over

Alterazione della concentrazione misurata dovuta a **residui di analita di un campione precedente** rimasti nello strumento (ago, valvola e loop dell'autocampionatore, colonna, sorgente) (ICH M10).

Il carry-over danneggia soprattutto i **campioni a bassa concentrazione iniettati dopo un campione concentrato**: un residuo trascurabile per un campione alto può essere una frazione importante del segnale di un campione vicino al LLOQ, che risulta sovrastimato. La prova riproduce il caso peggiore: si inietta un **bianco**

di matrice subito dopo il calibratore all'ULOQ. ICH M10 richiede che il segnale nel bianco non superi il **20% della risposta dell'analita al LLOQ** e il **5% della risposta dello standard interno**.

$$\text{carry-over}\% = \frac{A_{\text{bianco}}}{\bar{A}_{\text{LLOQ}}} \times 100 \leq 20\% \tag{14.1}$$

dove A_{bianco} è l'area dell'analita nel bianco dopo l'ULOQ e $\bar{A}_{\text{LLOQ}}$ l'area media al LLOQ. Il riferimento è il LLOQ, non l'ULOQ. Il carry-over si minimizza già in sviluppo (lavaggi dell'ago con solventi forti, gradiente di lavaggio della colonna); se è inevitabile, i campioni non si randomizzano e si iniettano bianchi dopo quelli molto concentrati. Un segnale in un bianco iniettato *prima* dell'ULOQ indica invece una contaminazione.

Esempio 14.1 — Valutazione del carry-over

Area media dell'analita al LLOQ = 6 250; all'ULOQ = 2 480 000; risposta media dello SI = 152 000. Nel bianco iniettato dopo l'ULOQ: 1 040 alla posizione dell'analita, 3 800 a quella dello SI.

1. Analita (Equazione 14.1): $1\,040/6\,250 \times 100 = 16{,}6\%$, entro il 20%.
2. SI: $3\,800/152\,000 \times 100 = 2{,}5\%$, entro il 5%.
3. Errore da evitare: $1\,040/2\,480\,000 \times 100 = 0{,}042\%$ usa l'ULOQ come riferimento e non dice nulla sui campioni bassi.

Risultato: carry-over conforme. Con un'area di 1 490 nel bianco si avrebbe 23,8%: non conforme.

14.5 Integrità della diluizione

Un campione **sopra l'ULOQ** non si quantifica estrapolando la curva: va **diluito con matrice bianca** (stessa specie e stesso tipo dei QC, non solvente, così che recupero ed effetto matrice restino quelli validati) e il risultato si moltiplica per il fattore di diluizione f:

$$C_{\text{campione}} = C_{\text{misurata}} \times f \tag{14.2}$$

L'**integrità della diluizione** verifica che la diluizione non comprometta esattezza e precisione. Secondo ICH M10 si preparano **QC di diluizione** in matrice **sopra l'ULOQ**, si diluiscono con matrice bianca con il fattore o i fattori che si useranno per i campioni (2×, 5×, 10× sono solo esempi) e se ne analizzano **almeno 5 repliche per fattore**. Criteri: esattezza media **±15%** e **CV ≤ 15%**. I fattori e le concentrazioni applicati ai campioni devono rientrare in quelli validati.

Esempio 14.2 — Verifica dell'integrità della diluizione

ULOQ = 1 000 ng/mL. Un QC di diluizione a 8 000 ng/mL è diluito 20 volte con plasma bianco; le 5 repliche danno 384; 401; 392; 377; 395 ng/mL.

1. Valori × 20: 7 680; 8 020; 7 840; 7 540; 7 900 ng/mL; media 7 796 ng/mL.
2. Esattezza: $(7\,796 - 8\,000)/8\,000 \times 100 = -2{,}6\%$.
3. Precisione: s = 188 ng/mL; $\text{CV}\% = 188/7\,796 \times 100 = 2{,}4\%$.
4. Errore da evitare: confrontare la media non moltiplicata (389,8 ng/mL) con 8 000 ng/mL dà −95,1%.

Risultato: integrità della diluizione 20× dimostrata. Un campione a circa 45 000 ng/mL, diluito 20×, resterebbe sopra l'ULOQ (2 250 ng/mL): serve un fattore maggiore (per esempio 100×, circa 450 ng/mL), da validare prima dell'uso con un QC di diluizione adeguato.

Punti chiave del Capitolo 14

- Stabilità: soluzioni stock (T ambiente e −20/−80 °C, periodi prolungati), bench-top (campione in matrice sul banco, tipicamente 4–24 h), gelo-disgelo (≥ 3 cicli e non meno di quelli dei campioni), lungo termine (durata ≥ conservazione), campione processato, sangue intero. In matrice: QC basso e alto, media entro ±15%.
- Robustezza: piccole variazioni deliberate dei parametri del metodo (pH ±0,1, flusso ±0,1 mL/min, temperatura ±2–5 °C); individua i parametri critici, si studia in sviluppo.
- Solidità: cambiamenti di contesto (operatore, strumento, laboratorio); oggi coincide con precisione intermedia e riproducibilità. Robusto non significa solido.
- Carry-over: colpisce i campioni bassi dopo quelli alti; bianco dopo l'ULOQ ≤ 20% della risposta al LLOQ e ≤ 5% di quella dello SI.
- Sopra l'ULOQ non si estrapola: si diluisce con matrice bianca con un fattore validato (QC sopra l'ULOQ, ≥ 5 repliche, ±15%, CV ≤ 15%).

Mettiti alla prova

Usa le domande del quiz sugli argomenti:

- **Stabilità**
- **Robustezza e solidità**
- **Carry-over e integrità della diluizione**

15. Il quadro normativo

Quali prove eseguire e con quali criteri dipende dal settore e dall'uso dei risultati: un metodo per uno studio farmacocinetico, uno per il rilascio dei lotti di un medicinale, uno per i pesticidi nella frutta e uno per gli antibiotici nel latte rispondono a documenti diversi. Scegliere il documento giusto è il primo passo del piano di validazione; i singoli criteri sono spiegati nei capitoli precedenti.

15.1 Gli enti e la natura dei documenti

La **FDA** (*Food and Drug Administration*, Stati Uniti) e l'**EMA** (*European Medicines Agency*, Unione europea) sono le autorità che valutano i dati per autorizzare i medicinali. L'**ICH** (*International Council for Harmonisation*) riunisce autorità regolatorie e industria farmaceutica per scrivere linee guida tecniche **armonizzate**, che poi le singole autorità adottano. Per alimenti e mangimi il riferimento è la Commissione europea (DG SANTE); **Eurachem** e **CITAC** pubblicano guide valide per qualsiasi laboratorio chimico.

Una **linea guida** non è legalmente vincolante: sono ammessi approcci diversi se scientificamente giustificati, anche se in pratica le autorità si aspettano che sia seguita. Un **regolamento dell'Unione europea** è invece obbligatorio e direttamente applicabile in tutti gli Stati membri.

15.2 La bioanalisi: ICH M10 (2022)

Per anni i laboratori bioanalitici hanno seguito due documenti simili ma non identici: la guida FDA *Bioanalytical Method Validation* (**2018**) e la linea guida EMA (**2011**). L'armonizzazione è arrivata con **ICH M10** (*Bioanalytical Method Validation and Study Sample Analysis*, **2022**), adottata sia dalla FDA sia dall'EMA: ha **sostituito la linea guida EMA dal 2023** e **supera la guida FDA 2018 per i metodi nel suo ambito**. Un metodo LC-MS/MS per uno studio da presentare a entrambe le agenzie si valida quindi una sola volta, secondo M10.

Ambito. M10 riguarda la quantificazione di farmaci chimici e biologici e dei loro metaboliti in **matrici biologiche** (sangue, plasma, siero, urina, tessuti) per studi **tossicocinetici** e **farmacocinetici** non clinici e per tutte le fasi della sperimentazione clinica, compresi gli studi di bioequivalenza. Sono esclusi i biomarcatori e l'immunogenicità, oltre ovviamente al controllo qualità dei medicinali, ai pesticidi negli alimenti e ai farmaci veterinari negli alimenti, che hanno documenti propri.

Famiglie di metodi. M10 fornisce criteri specifici per i **metodi cromatografici** (LC o GC, tipicamente con spettrometria di massa) e per i **saggi di legame con ligandi** (*ligand binding assays*, LBA, come gli immunodosaggi). Per i cromatografici l'esattezza è ±15% (±20% al LLOQ) e il CV ≤ 15% (≤ 20% al LLOQ); per gli LBA ±20% e CV ≤ 20% (±25% e CV ≤ 25% al LLOQ e all'ULOQ).

Parametri. La validazione completa dei metodi cromatografici comprende selettività e specificità (Capitolo 11), effetto matrice (Capitolo 13), curva di calibrazione e intervallo (Capitolo 10), esattezza e precisione (Capitolo 12), carry-over, integrità della diluizione e stabilità (Capitolo 14), oltre alla riproducibilità della reiniezione (un estratto reiniettato, per esempio dopo un guasto dello strumento, deve dare lo stesso risultato). Il **recupero** non è tra questi elementi: M10 lo tratta nelle considerazioni aggiuntive (§7.3), come aspetto dello sviluppo del metodo, chiedendo che sia valutato (tipicamente a tre livelli) e costante, non necessariamente del 100%. Nemmeno la robustezza fa parte della validazione completa.

Novità rispetto alle linee guida precedenti:

- un unico documento armonizzato per FDA, EMA e le altre autorità ICH;
- con rivelazione MS, **SIL-IS raccomandato ogni volta che è possibile**;
- **effetto matrice con QC basso e alto in almeno 6 lotti**, con esattezza ±15% e CV ≤ 15% in ogni lotto, al posto del CV del fattore di matrice normalizzato della linea guida EMA;
- criteri per **LBA e metodi cromatografici** nello stesso documento.

15.3 ICH Q2(R2) e ICH Q14 (2023): i metodi per i prodotti farmaceutici

ICH Q2(R2) è la revisione della storica linea guida sulla validazione dei metodi analitici (Q2(R1), 2005). Riguarda i metodi per il **rilascio e la stabilità di principi attivi e medicinali** (per esempio il dosaggio del principio attivo nelle compresse o delle impurezze), non la bioanalisi. Descrive le caratteristiche da validare: specificità, intervallo e risposta (linearità, ma anche modelli non lineari), accuratezza, precisione, LOD, LOQ e robustezza, quest'ultima affrontata soprattutto in sviluppo. Due precisazioni: Q2(R2) **non fissa criteri numerici universali** (li stabilisce il richiedente in base all'uso previsto), e nei documenti ICH *accuracy* indica l'**esattezza** (*trueness* della ISO), non l'accuratezza intesa come esattezza + precisione (Capitolo 3).

ICH Q14 (*Analytical Procedure Development*) è la nuova linea guida sullo **sviluppo** dei metodi, basato sulla comprensione del metodo e sulla gestione del rischio (*Analytical Quality by Design*, AQbD): studio sistematico dei parametri, spesso con disegni sperimentali, strategia di controllo, gestione delle modifiche nel ciclo di vita.

Analytical Target Profile (ATP)

Descrizione, stabilita **prima** dello sviluppo, di che cosa la procedura deve misurare e con quali prestazioni (analita, matrice, intervallo, esattezza e precisione richieste), indipendente dalla tecnica. È introdotto da ICH Q14.

I due documenti sono **complementari**: Q2(R2) dice **che cosa validare**, Q14 **come sviluppare** il metodo; Q2(R2) si raccorda all'ATP, perché la validazione dimostra che i requisiti dell'ATP sono soddisfatti. Nessuno sostituisce l'altro.

15.4 Alimenti, farmaci veterinari e laboratori in generale

Il documento **SANTE/11312/2021** (versione v2026, in vigore dal 1° gennaio 2026) è la guida europea per il controllo qualità e la validazione dei metodi per i **residui di pesticidi in alimenti e mangimi**, in particolare dei

metodi **multi-residuo** (per esempio QuEChERS con LC-MS/MS e GC-MS/MS). I criteri generali sono: **recupero medio 70–120%** a ciascun livello di arricchimento e **RSD ≤ 20%**; il LOQ è il livello più basso che li soddisfa. Recuperi medi fuori da 70–120% sono ammessi solo in casi eccezionali, se costanti (RSD ≤ 20%) e con una causa ben documentata. Qui il recupero è il rapporto tra concentrazione trovata con l'intero metodo e concentrazione aggiunta, cioè una misura di esattezza, diversa dal recupero di estrazione di M10 (Capitolo 13). La v2026 ha sostituito la v2 del 2024 lasciando invariati questi criteri.

$$R\% = \frac{C_{\text{trovata}}}{C_{\text{aggiunta}}} \times 100 \tag{15.1}$$

Esempio 15.1 — Verifica dei criteri SANTE

Un campione di mela privo di residui è arricchito a 0,010 mg/kg con un pesticida; 5 repliche danno 0,0064; 0,0069; 0,0071; 0,0066; 0,0068 mg/kg.

1. Recuperi (Equazione 15.1): 64; 69; 71; 66; 68%.
2. Media 67,6%; s = 2,70; $\text{RSD} = 2{,}70/67{,}6 \times 100 = 4{,}0\%$.
3. La precisione è buona, ma il recupero medio è sotto il 70%.

Risultato: non conforme ai criteri generali. Per un altro pesticida con recuperi 105, 78, 122, 84 e 131% la media (104,0%) sarebbe corretta, ma l'RSD (22,2%) supera il 20%: anche in questo caso il metodo non è conforme.

Il **Reg. (UE) 2021/808** disciplina i metodi analitici per il **controllo ufficiale dei residui di farmaci veterinari negli alimenti di origine animale** (carne, latte, uova, miele). Ha **sostituito la Decisione 2002/657/CE** e ne **aggiorna i criteri per CCα e CCβ**. Il **limite di decisione (CCα)** è la concentrazione oltre la quale si può concludere, con probabilità di errore α (falso non conforme), che il campione è non conforme; la **capacità di rivelazione (CCβ)** è la più piccola concentrazione che il metodo rivela, identifica e/o quantifica con probabilità di errore β (falso conforme).

La guida Eurachem *The Fitness for Purpose of Analytical Methods* (3ª ed., 2025) è una guida generale per i laboratori chimici di **qualsiasi settore** (ambiente, alimenti, clinica, industria), riferimento per l'accreditamento secondo la ISO/IEC 17025: spiega quando e come validare un metodo e come valutarne le caratteristiche. Per l'**incertezza di misura** rimanda alla guida **Eurachem/CITAC QUAM** (*Quantifying Uncertainty in Analytical Measurement*), che applica la GUM alla chimica e usa anche i dati di validazione (Capitolo 6).

15.5 Confronto tra i documenti

Tabella 15.1 — I principali documenti sulla validazione

Documento	Ente, anno	Ambito	Elementi chiave
Bioanalytical Method Validation	FDA, 2018	bioanalisi (USA)	superata da ICH M10 per i metodi nel suo ambito
Guideline on Bioanalytical Method Validation	EMA, 2011	bioanalisi (UE)	sostituita da ICH M10 dal 2023; CV del MF normalizzato ≤ 15%
ICH M10	ICH, 2022 (FDA ed EMA)	farmaci e metaboliti in matrici biologiche, studi PK/TK	metodi cromatografici e LBA; ±15% (±20% al LLOQ); effetto matrice in ≥ 6 lotti; SIL-IS
ICH Q2(R2)	ICH, 2023	metodi per principi attivi e medicinali	che cosa validare; *accuracy* = esattezza
ICH Q14	ICH, 2023	sviluppo dei metodi farmaceutici	come sviluppare; ATP; AQbD
SANTE/11312/2021 v2026	Commissione europea, 2026	residui di pesticidi in alimenti e mangimi	recupero 70–120%, RSD ≤ 20%
Reg. (UE) 2021/808	Commissione europea, 2021	controllo ufficiale dei residui di farmaci veterinari	sostituisce la Decisione 2002/657/CE; CCα e CCβ
Fitness for Purpose (3ª ed.)	Eurachem, 2025	laboratori chimici di ogni settore	validazione e verifica; incertezza con Eurachem/CITAC QUAM

Punti chiave del Capitolo 15

- ICH M10 (2022) è la linea guida armonizzata sulla validazione bioanalitica, adottata da FDA ed EMA: ha sostituito EMA 2011 (dal 2023) e supera FDA 2018 per i metodi nel suo ambito.
- M10: farmaci e metaboliti in matrici biologiche per studi PK/TK; criteri per metodi cromatografici e LBA; SIL-IS raccomandato; effetto matrice con QC in ≥ 6 lotti; il recupero è trattato tra le considerazioni aggiuntive (§7.3).
- ICH Q2(R2) dice che cosa validare per i metodi farmaceutici, ICH Q14 come svilupparli e introduce l'ATP; nei documenti ICH *accuracy* = esattezza.
- SANTE/11312/2021: pesticidi in alimenti e mangimi, recupero 70–120%, RSD ≤ 20%.
- Reg. (UE) 2021/808: residui di farmaci veterinari negli alimenti di origine animale; sostituisce la Decisione 2002/657/CE e aggiorna CCα e CCβ.
- Eurachem *Fitness for Purpose*: validazione in qualsiasi settore; incertezza con Eurachem/CITAC QUAM.

Mettiti alla prova

Usa le domande del quiz sull'argomento:

- **Linee guida e quadro normativo**

16. Dal metodo alla routine: aspetti pratici

Un metodo validato va poi usato ogni giorno senza perdere affidabilità. Questo capitolo raccoglie gli aspetti pratici che collegano calibrazione e validazione al lavoro di routine, con particolare attenzione alla bioanalisi in LC-MS/MS.

16.1 Metodi strumentali e principali tecniche

Gravimetria e volumetria sono metodi *assoluti* (Capitolo 10). I metodi strumentali offrono tre vantaggi:

- **limiti di rivelabilità e di quantificazione molto più bassi**, adatti all'analisi in tracce (ng/mL o pg/mL in LC-MS/MS). Questo vantaggio viene spesso chiamato "maggiore sensibilità", ma in senso stretto la sensibilità è la pendenza della curva, distinta da LOD e LOQ (Capitolo 11);
- **automazione**: analisi in serie con autocampionatori, meno errori umani e più produttività;
- **gestione dei dati**: acquisizione, elaborazione e archiviazione al PC con software dedicati.

Sono però **metodi relativi**: richiedono calibrazione, standard e QC, e la loro precisione peggiora alle basse concentrazioni. L'automazione non elimina errori sistematici né effetto matrice e non autorizza a quantificare fuori dall'intervallo validato.

Tabella 16.1 — Le quattro famiglie di tecniche strumentali

Famiglia	Principio	Esempi
Spettroscopia	Assorbimento o emissione di radiazione da parte di molecole o atomi	UV-Vis, IR (assorbimento molecolare); fluorescenza (emissione molecolare); AAS (assorbimento atomico); ICP-OES (emissione atomica)
Metodi elettrochimici	Misura di potenziale, corrente o conduttanza	Potenziometria (potenziale), voltammetria (corrente), conduttimetria (conduttanza), elettrolisi
Cromatografia	Separazione tra fase mobile e fase stazionaria, seguita da rivelazione spettroscopica o elettrochimica	GC, HPLC, cromatografia ionica
Spettrometria di massa	Separazione degli ioni in base al rapporto massa/carica (m/z); alta specificità e sensibilità	MS, LC-MS/MS, GC-MS

Nel triplo quadrupolo la transizione ione precursore → ione prodotto è caratteristica dell'analita: per questo l'LC-MS/MS è la tecnica di riferimento della bioanalisi. Il suo punto debole è la sorgente, sensibile all'effetto matrice (Capitolo 13).

16.2 Il processo analitico

Il processo analitico comprende cinque fasi: **campionamento → standard e bianco → calibrazione → analisi → elaborazione dei dati** (Figura 16.1). La calibrazione precede l'analisi perché è la curva a convertire i segnali dei campioni in concentrazioni.



Figura 16.1 — Le cinque fasi del processo analitico, con alcune fonti di errore tipiche di ciascuna fase.

Ogni fase introduce potenziali errori, sistematici o casuali. Per questo **la validazione deve coprire l'intero processo, non solo la fase strumentale**: iniettando solo standard in solvente resterebbero invisibili le perdite

in estrazione, l'effetto matrice e l'instabilità del campione. Il campionamento sfugge persino ai QC, che sono preparati in laboratorio da una matrice omogenea. Se due porzioni dello stesso lotto (per esempio di un mangime contaminato da micotossine, distribuite in modo disomogeneo) differiscono di un fattore 3 mentre QC e iniezioni ripetute sono conformi, la causa è un campionamento non rappresentativo.

16.3 Propagazione dell'incertezza nelle diluizioni seriali

Una soluzione di lavoro si ottiene in catena: pesata → soluzione madre → diluizioni (Capitolo 9). La concentrazione finale è un **prodotto di quozienti**, per esempio $c = \frac{m \cdot P}{V_0} \cdot \frac{V_{p,1}}{V_{m,1}} \cdot \frac{V_{p,2}}{V_{m,2}}$. Per prodotti e quozienti di grandezze indipendenti le **incertezze relative si combinano in quadratura** (GUM, Capitolo 6):

$$\frac{u_c(c)}{c} = \sqrt{\sum_{i=1}^{n} \left(\frac{u_i}{x_i}\right)^2} \tag{16.1}$$

Equazione 16.1

- x_i: ogni grandezza del calcolo (massa, purezza, volume di ogni pipetta e matraccio); u_i: la sua incertezza standard; u_i/x_i: incertezza relativa.
- Si usano le incertezze relative perché un errore dell'1% su qualsiasi fattore dà l'1% su c; le incertezze assolute si combinano così solo per somme e differenze.
- La somma lineare sovrastima; la somma dei quadrati senza radice è in %². Con n passaggi di uguale incertezza u_{rel} il risultato è $u_{\mathrm{rel}}\sqrt{n}$: cresce con n (non si divide per $\sqrt{n}$, che vale per l'errore standard di una media).
- Una tolleranza di vetreria $\pm a$ si converte in incertezza standard con $u = a/\sqrt{6}$ (distribuzione triangolare, Eurachem/CITAC) o $a/\sqrt{3}$ (rettangolare).

Esempio 16.1 — Una soluzione di lavoro in due diluizioni

Dati. Soluzione madre: 25,00 mg di standard (purezza 99,2%) in 25 mL, 0,992 mg/mL, con incertezza relativa 0,32% (pesata, purezza, matraccio). Diluizioni: 1,000 mL → 100,0 mL, poi 5,00 mL → 50,0 mL. Incertezze relative: pipetta da 1 mL 0,327%; matraccio da 100 mL 0,041%; pipetta da 5 mL 0,122%; matraccio da 50 mL 0,049%.

1. c = 0,992 mg/mL × (1,000/100,0) × (5,00/50,0) = 992 ng/mL.
2. Somma dei quadrati: 0,32² + 0,327² + 0,041² + 0,122² + 0,049² = 0,228 %².
3. $u_c(c)/c$ = $\sqrt{0{,}228}$ = 0,48%; u_c = 4,7 ng/mL; U (k = 2) = 9,5 ng/mL.

Risultato: c = (992,0 ± 9,5) ng/mL. Dominano la pipetta da 1 mL (0,327%) e la soluzione madre (0,32%): conviene usare volumi più grandi, non matracci migliori.

Le implicazioni pratiche sono: **ridurre al minimo le diluizioni seriali** (a parità di incertezza per passaggio; usare volumi minuscoli per risparmiare passaggi è controproducente); usare **pipette tarate e calibrate** con incertezza nota e **matracci tarati di classe A** (non cilindri graduati); **documentare ogni passaggio** con la sua incertezza; preparare la **soluzione madre per pesata diretta** di una massa sufficiente, corretta per la purezza. Un errore nella soluzione madre si trasmette a tutti i calibranti senza deformare la curva: per questo ICH M10 chiede che calibranti e QC derivino da **soluzioni madre separate**.

16.4 Monitoraggio della curva nel tempo

In bioanalisi regolata (ICH M10) si costruisce **una nuova curva per ogni sequenza analitica**: non si quantifica mai con la curva di una sequenza precedente. Dove invece la procedura consente di riutilizzare una curva (per esempio negli analizzatori di chimica clinica), al cambio di **lotto di reattivi** la curva va **verificata con 1–2 standard** prima di continuare a usarla. Se gli standard escono dalla tolleranza (per esempio −20%), la risposta è cambiata e si costruisce una nuova curva. Non si correggono i risultati con un fattore empirico e non si ripete la verifica finché "passa".

La **pendenza** è la sensibilità del metodo e deve restare costante nel tempo: conviene registrarla in ogni sequenza su una carta di controllo. Un calo progressivo (per esempio del 25% in alcune settimane) va indagato anche se ogni curva è accettabile, con r^2 elevato e calibranti conformi: indica problemi strumentali (sorgente sporca, rivelatore, colonna) o di reattivi e soluzioni. La contaminazione del bianco agisce invece sull'intercetta. I **QC** a più livelli, preparati indipendentemente dai calibranti, confermano in ogni sequenza l'efficienza di **procedura, strumento e operatori**. Non entrano nella regressione e non correggono l'effetto matrice.

16.5 L'intercetta e la forzatura per l'origine

L'intercetta b_0 è il segnale previsto a concentrazione zero, cioè il **segnale del bianco**: fondo strumentale più **interferenze additive**. Un'intercetta significativamente diversa da zero può indicare contaminazione del bianco o dei reattivi, interferenze additive non corrette (comprese le interferenze additive della matrice) o un problema strumentale come la deriva della linea di base. L'effetto matrice in senso stretto (soppressione o esaltazione ionica) è invece moltiplicativo e agisce sulla **pendenza**, come gli errori proporzionali (per esempio un errore nella soluzione madre). La saturazione del rivelatore incurva la retta in alto.

La significatività di b_0 si verifica con un test t di Student (Capitolo 5): $t = b_0/s_{b_0}$, dove s_{b_0} è la deviazione standard dell'intercetta fornita dal software di regressione, con $\nu = n - 2$. In modo equivalente si controlla se l'intervallo di confidenza di b_0 comprende lo zero. Forzare la retta per l'origine significa imporre b_0 = 0; la pendenza diventa:

$$b_1{}' = \frac{\sum x_i\, y_i}{\sum x_i^2} \tag{16.2}$$

Esempio 16.2 — Effetto della forzatura con un'intercetta reale

Dati. Calibranti a 0,5, 1, 2, 5, 10 e 20 µg/mL; risposta esatta $y = 3{,}0 + 10{,}0x$ (intercetta dovuta al fondo).

1. Equazione 16.2: $b_1{}'$ = 5418/530,25 = 10,218 (solo +2,2% rispetto a 10,0).
2. Un campione da 1,00 µg/mL dà y = 13,0; con la retta forzata x = 13,0/10,218 = 1,272 µg/mL, bias +27,2%.
3. A 0,50 µg/mL il bias è +56,6%; a 20,0 µg/mL è −0,7%.

Risultato: la forzatura introduce un errore sistematico artificiale concentrato ai livelli bassi (Figura 16.2).

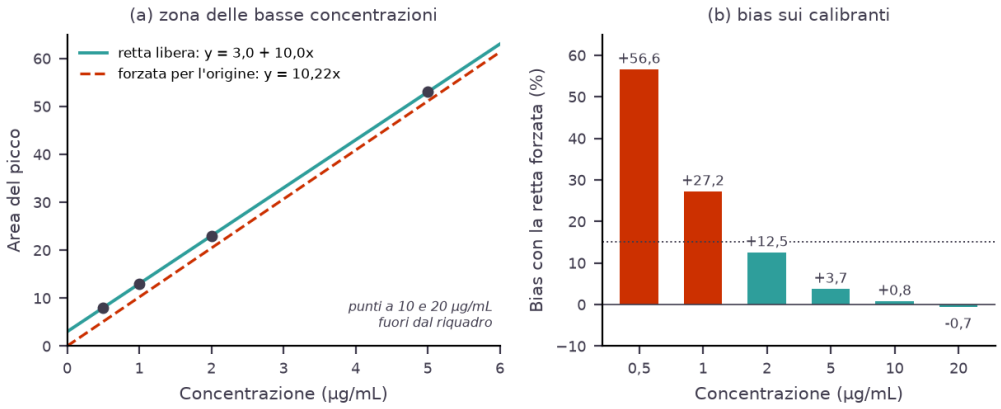


Figura 16.2 — Forzatura per l'origine con intercetta reale (Esempio 16.2): (a) la retta forzata passa sotto i punti a bassa concentrazione; (b) bias con la retta forzata, oltre ±15% ai livelli bassi.

La **regola operativa** è: l'intercetta si **stima liberamente**; se è nettamente diversa da zero si valuta se è **fisicamente giustificabile** (per esempio un fondo strumentale noto), altrimenti si **rivede la procedura analitica**. Non si forza la retta, non si escludono i calibranti bassi e non ci si accontenta di un r^2 elevato. Secondo Eurachem e ISO 8466-1 la regressione per l'origine è ammissibile solo se l'intercetta **non è statisticamente diversa da zero** e la scelta è giustificata e documentata.

16.6 Punti aberranti e accettazione della curva

I punti anomali si cercano **dopo la regressione**: grafico dei residui, **residui standardizzati** ($e_i/s_{y/x}$; oltre 3 in valore assoluto il punto è sospetto) e concentrazioni retrocalcolate. I **test di Grubbs e Dixon** valgono solo per **repliche di uno stesso livello**, non per calibranti a concentrazioni diverse.

Per ogni calibratore si calcola il bias della concentrazione retrocalcolata $x_{\text{retro}} = (y_i - b_0)/b_1$:

$$\text{bias}\% = \frac{x_{\text{retro}} - x_{\text{nom}}}{x_{\text{nom}}} \times 100 \tag{16.3}$$

Per esempio, con y = 0,0042 + 0,0315x un calibratore al LLOQ (1,00 ng/mL) con y = 0,0381 dà x_{retro} = 1,076 ng/mL e un bias del +7,6% (Equazione 16.3), conforme. Dimenticando l'intercetta si otterrebbe +21,0%.

Le regole ICH M10 sono:

- un calibratore con bias oltre **±15%** (**±20% al LLOQ**) è **non conforme e viene scartato**, senza bisogno di una causa; si ricalcola la regressione senza di esso;
- un calibratore **conforme** si esclude **solo con una giustificazione documentata** (errore di preparazione, problema di iniezione). Escluderlo per far salire r^2 è inaccettabile, anche se un residuo standardizzato elevato impone di indagare;
- la curva è accettata se almeno il **75% dei calibranti totali** (scartati compresi) è conforme e i conformi coprono **almeno 6 livelli**. Si arrotonda per eccesso: con 10 calibranti 0,75 × 10 = 7,5, quindi servono 8; con 8 ne servono 6.

16.7 Lo standard interno marcato con isotopi stabili (SIL-IS)

Il SIL-IS è l'analita marcato con ^2H, ^13C o ^15N: si comporta come l'analita ma si distingue per m/z. Si quantifica con il rapporto $R = A_{\text{analita}}/A_{\text{SI}}$, nel quale si annulla ogni fattore che moltiplica in egual misura i due segnali. Il SIL-IS compensa quindi, **dopo la sua aggiunta**: perdite in estrazione, variabilità dell'iniezione, effetto matrice (se co-eluisce) e instabilità della sorgente MS.

Non compensa: gli errori nei calibranti (una soluzione madre sbagliata distorce tutti i risultati in proporzione), ciò che avviene prima della sua aggiunta, gli errori nel volume di SI aggiunto e le interferenze additive nel canale dell'analita.

Requisiti pratici:

- aggiunta il prima possibile, idealmente prima dell'estrazione;
- concentrazione costante in campioni, calibranti e QC;
- assenza di contributi incrociati: per ICH M10 il contributo dello SI al canale dell'analita deve essere ≤ 20% della risposta al LLOQ, quello dell'analita al canale dello SI ≤ 5% della risposta dello SI;
- purezza isotopica elevata: "≥ 98%" è una regola pratica, non un criterio ICH M10.

Se il bianco di matrice è pulito ma il **campione zero** (stessa matrice e stessi reattivi, più lo SI) mostra un picco nel canale dell'analita, la causa è lo SI stesso (impurezza isotopica o cross-talk).

Il comportamento "identico all'analita" vale pienamente per ^13C e ^15N. Gli analoghi **deuterati** possono eluire leggermente prima (effetto isotopico del deuterio) e compensare l'effetto matrice in modo imperfetto; inoltre il deuterio su –OH o –NH si scambia con il solvente. Il limite principale del SIL-IS è il **costo elevato**, insieme a una **disponibilità non garantita** per tutti gli analiti.

Esempio 16.3 — Rapporto analita/SI sotto soppressione ionica

Dati. Pendenza 0,0300 mL/ng (rapporto delle aree), intercetta trascurabile; area dello SI nei calibranti circa 120 000. Un plasma lipemico contiene 8,00 ng/mL e subisce una soppressione del 50%.

1. Aree misurate: analita 14 400, SI 60 000; R = 0,240; c = 0,240/0,0300 = 8,00 ng/mL. La soppressione si annulla nel rapporto e non va corretta di nuovo.
2. Calibrando sulla sola area dell'analita (3600 per ng/mL senza soppressione): 14 400/3600 = 4,00 ng/mL (−50%).
3. SI deuterato che eluisce dove la soppressione è solo del 35%: SI = 78 000; R = 0,185; c = 6,15 ng/mL (−23%).

Risultato: solo uno SI che co-eluisce esattamente restituisce 8,00 ng/mL.

16.8 Riepilogo dei criteri ICH M10

Tabella 16.2 — Principali criteri ICH M10 (metodi cromatografici)

Parametro	Criterio generale	Al LLOQ
Esattezza (bias%)	±15%	±20%
Precisione (CV%)	≤ 15%	≤ 20%
Calibranti conformi	≥ 75% del totale, ≥ 6 livelli	—
Selettività (bianchi di ≥ 6 lotti)	≤ 20% della risposta al LLOQ e ≤ 5% della risposta dello SI	—
Effetto matrice (QC basso e alto)	±15% e CV ≤ 15% in ciascuno di ≥ 6 lotti	—
Carry-over (bianco dopo l'ULOQ)	≤ 20% della risposta al LLOQ e ≤ 5% della risposta dello SI	—
Accettazione della sequenza	≥ 2/3 dei QC entro ±15% e ≥ 50% a ogni livello	—

Selettività e carry-over hanno lo stesso limite ma campioni diversi: **bianchi di matrice di lotti diversi** per la selettività, **un bianco iniettato subito dopo l'ULOQ** per il carry-over. Per esempio, un'area di 480 nel bianco contro 2850 al LLOQ dà 16,8%, conforme. L'effetto matrice si valuta su più lotti perché la composizione della matrice varia tra individui (plasma lipemico, emolizzato), e il criterio va rispettato lotto per lotto. Per i metodi ligand-binding i limiti sono ±20% (±25% al LLOQ e all'ULOQ).

16.9 Validazione completa, parziale e cross-validazione

- **Validazione completa**: metodo **nuovo o modificato in modo significativo**. Per i metodi cromatografici ICH M10 prevede selettività, specificità, effetto matrice, curva e intervallo, esattezza, precisione, carry-over,

integrità della diluizione, stabilità e riproducibilità della reiniezione. La **robustezza** non ne fa parte: si studia nello sviluppo del metodo (ICH Q14). Il recupero si valuta, ma senza un criterio numerico.

- **Validazione parziale**: **modifiche a un metodo validato**, valutando solo i parametri potenzialmente influenzati. Esempi: **trasferimento del metodo a un altro laboratorio**, nuovo strumento equivalente, nuovo intervallo (per un nuovo ULOQ: curva, esattezza e precisione ai nuovi livelli, carry-over), cambio di matrice. Un nuovo lotto di reattivi richiede solo la verifica della curva (paragrafo 16.4).
- **Cross-validazione**: confronto tra due metodi validati, o tra laboratori diversi, sugli stessi campioni, quando i dati vanno **combinati o confrontati** (per esempio uno studio multicentrico). **ICH M10 non fissa criteri di accettazione**: il bias si valuta statisticamente. Il **grafico di Bland-Altman** riporta le differenze contro le medie e dà il bias medio $\bar{d}$ e i limiti di accordo $\bar{d} \pm 1{,}96\, s_d$. La **regressione di Deming**, a differenza dei minimi quadrati ordinari, considera l'errore di misura su entrambi gli assi: una pendenza diversa da 1 indica bias proporzionale, un'intercetta diversa da 0 bias costante. Il criterio "±20% per due terzi dei campioni" riguarda invece la rianalisi dei campioni reali (ISR).

16.10 Piano e rapporto di validazione

Il **piano di validazione** (VP) si redige **prima degli esperimenti**: obiettivi, parametri, **criteri di accettabilità**, disegno sperimentale, responsabilità e tempistiche. Fissare i criteri in anticipo impedisce di adattarli ai risultati. Il **rapporto di validazione** (VR) riporta tutti i dati sperimentali, i calcoli statistici, il confronto con i criteri e la conclusione sulla validità del metodo; va **firmato e archiviato** secondo le norme GLP/GMP. In ambito regolatorio entrambi sono **documenti ispezionabili** e la **tracciabilità di ogni dato è obbligatoria**. Anche una sequenza fallita va riportata con la sua causa: non invalida di per sé la validazione, ma ometterla compromette l'integrità dello studio.

16.11 Errori comuni nella costruzione della curva

1. **Forzare la retta per l'origine** senza aver verificato che l'intercetta sia non significativa: introduce un bias sistematico.
2. **Usare solo** r^2, che non rileva deviazioni sistematiche: si valutano residui e concentrazioni retrocalcolate. ICH M10 non fissa soglie di r^2. Una curva con r^2 = 0,992, residui casuali e tutti i calibranti conformi è accettabile.
3. **Escludere calibranti conformi senza giustificazione**: invalida la curva; ogni esclusione va documentata.
4. **Riutilizzare la curva con nuovi reattivi senza verificarla** con almeno 1–2 standard.

16.12 La sequenza analitica in LC-MS/MS

L'ordine di iniezione è standardizzato per minimizzare il carry-over e garantire la validità della curva per tutta la sequenza (Figura 16.3):

1. **bianco di matrice**, per primo, così che un suo segnale non possa derivare da carry-over;
2. **campione zero** (matrice + SI), che verifica il contributo dello SI;
3. **calibranti** dal LLOQ all'ULOQ, seguiti da un bianco per il carry-over;
4. **QC**, poi i **campioni** intervallati da altri QC, e infine i **QC di fine sequenza**, che verificano la stabilità della risposta strumentale.

Bianco e campione zero non entrano nella regressione.

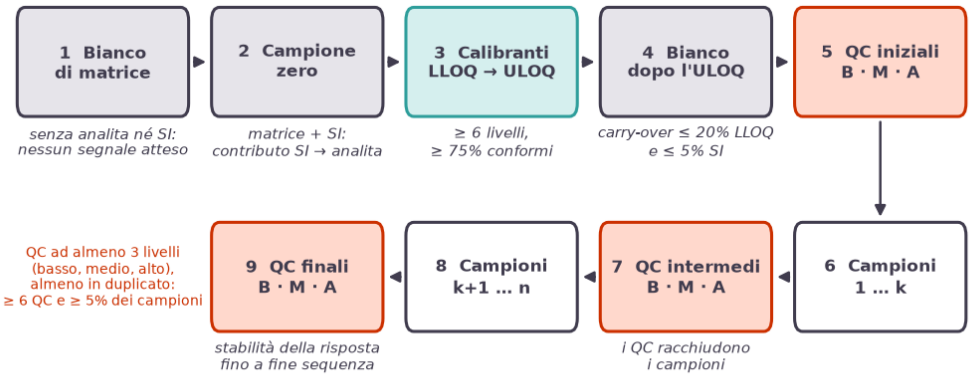


Figura 16.3 — Sequenza analitica tipica in LC-MS/MS secondo ICH M10, con i QC distribuiti lungo la sequenza.

"QC triplicati" va inteso come **QC a tre livelli**. ICH M10 chiede QC ad **almeno 3 livelli** (basso, medio, alto), **almeno in duplicato**, in numero pari ad almeno 6 e ad almeno il 5% dei campioni, distribuiti nella sequenza. Con 150 campioni, il 5% dà 7,5, quindi servono almeno 8 QC: in pratica 9, tre per livello. La sequenza è accettata se ne sono conformi almeno 6 (2/3) e almeno 2 per livello (50%): due QC bassi fuori limite la fanno respingere anche con 7 conformi su 9.

Punti chiave

- I metodi strumentali hanno LOD/LOQ bassi, automazione e gestione dei dati, ma sono relativi; la validazione deve coprire l'intero processo, dal campionamento all'elaborazione.
- Nelle diluizioni le incertezze relative si sommano in quadratura: poche diluizioni, classe A, madre per pesata diretta.
- Una curva per ogni sequenza; con nuovi reattivi una curva riutilizzata va verificata; pendenza e QC si monitorano nel tempo.
- L'intercetta (segnale del bianco) si stima liberamente; forzare l'origine è lecito solo se l'intercetta non è significativa e la scelta è giustificata.
- Calibranti non conformi si scartano, conformi solo con causa documentata; ≥ 75% del totale conforme e ≥ 6 livelli.
- Il SIL-IS compensa estrazione, iniezione, effetto matrice e sorgente, ma non errori nei calibranti; i deuterati possono non co-eluire.
- Validazione completa per metodi nuovi, parziale per modifiche e trasferimenti, cross-validazione senza criteri fissi; piano e rapporto garantiscono la tracciabilità.

Mettiti alla prova

Argomenti del quiz: *Metodi di calibrazione e metodi strumentali*; *Tecniche strumentali e processo analitico*; *Propagazione dell'incertezza nelle diluizioni*; *Monitoraggio della curva e intercetta*; *Accettazione della curva e punti aberranti*; *Standard interno marcato con isotopi stabili (SIL-IS)*; *Criteri ICH M10 e sequenza analitica*; *Tipi di validazione e documentazione*.

17. Sintesi: dalla calibrazione alla validazione

Questo capitolo riassume la Parte II: come gli strumenti della calibrazione e i parametri della validazione portano a un metodo pronto per la routine, quali messaggi conviene ricordare e dove approfondire.

17.1 Il quadro d'insieme

La Figura 17.1 riassume la Parte II in due blocchi.

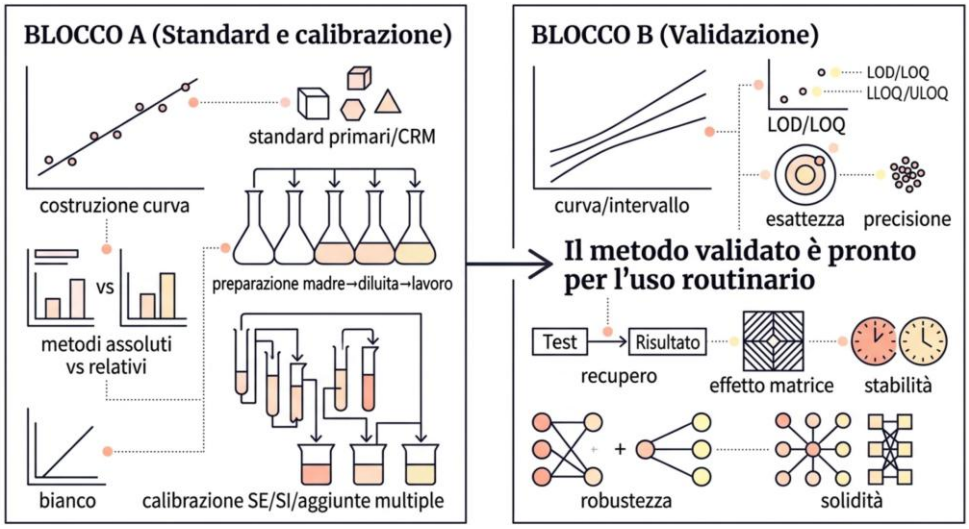


Figura 17.1 — Sintesi della Parte II. Blocco A: gli strumenti per costruire una curva affidabile. Blocco B: i parametri che dimostrano l'idoneità del metodo. Solo dopo entrambi il metodo è pronto per l'uso di routine.

Il **blocco A, standard e calibrazione**, comprende:

- standard primari e CRM, che danno la tracciabilità metrologica (Capitolo 9);
- la preparazione della catena soluzione madre → diluita → di lavoro (Capitoli 9 e 16);
- la distinzione tra metodi assoluti e relativi e il ruolo del bianco (Capitolo 10);
- la costruzione della curva e la scelta tra standard esterno, standard interno e aggiunte multiple (Capitolo 10, con le basi statistiche dei Capitoli 7–8).

Il **blocco B, validazione**, raccoglie i parametri che dimostrano che il metodo è adatto allo scopo:

- curva e intervallo, LOD/LOQ e LLOQ/ULOQ (Capitolo 11);
- esattezza e precisione (Capitolo 12);
- recupero ed effetto matrice (Capitolo 13);
- stabilità, robustezza e solidità (Capitolo 14).

Robustezza e solidità compaiono nello schema come parte del percorso di validazione. In ICH M10 e in ICH Q2(R2)/Q14 la robustezza si studia però nello sviluppo del metodo, mentre il concetto di solidità confluisce in precisione intermedia e riproducibilità (Capitolo 14).

Superati i due blocchi, il metodo entra in routine, dove resta sotto controllo. Ogni sequenza ha la sua curva e i suoi QC; pendenza e QC si monitorano nel tempo. Le modifiche richiedono una validazione parziale; i dati di metodi o laboratori diversi, se vanno combinati, una cross-validazione (Capitolo 16).

17.2 I messaggi chiave

1. **La tracciabilità metrologica è fondamentale.** Ogni standard deve essere collegato al SI attraverso una catena documentata di confronti, con incertezza dichiarata a ogni livello. Una regressione perfetta non può rivelare un errore sul valore assegnato allo standard (Esempio 17.1). La vetreria tarata rende tracciabili i volumi, non la purezza dello standard.
2. r^2 **non è sufficiente.** La curva si valuta con i residui e con le concentrazioni retrocalcolate. Un r^2 alto può nascondere deviazioni sistematiche; ICH M10 non fissa alcuna soglia di r^2.
3. **La coerenza su tutto il range è obbligatoria.** Esattezza, precisione e recupero devono rispettare i criteri dal LLOQ all'ULOQ, non solo a metà curva. Una curva "perfetta" con QC che falliscono ai livelli bassi indica un metodo non valido.

4. **Il SIL-IS è la scelta ottimale in LC-MS/MS.** Compensa insieme perdite in estrazione, variabilità dell'iniezione ed effetto matrice, purché co-eluisca con l'analita. Non corregge però gli errori nei calibranti, e gli analoghi deuterati possono co-eluire in modo imperfetto.

Esempio 17.1 — Una curva perfetta con uno standard sbagliato

Dati. Una curva costruita con uno standard di purezza non certificata, ritenuto puro al 100%, ha r^2 = 0,9999 e residui casuali. La purezza reale è il 95%.

1. Ogni calibratore contiene in realtà il 95% della concentrazione nominale; la pendenza apparente è quindi 0,95 volte quella vera.
2. Un campione a concentrazione c dà il segnale atteso per c, ma viene quantificato come $c/0{,}95$ = 1,053 c.

Risultato: tutti i risultati sono sovrastimati del 5,3%, senza alcun segno nei residui. Solo uno standard tracciabile (un CRM) o QC preparati da uno standard indipendente rivelano l'errore.

17.3 Letture consigliate

Le linee guida da conoscere sono **ICH M10 (2022)** per la bioanalisi, **ICH Q2(R2) e Q14 (2023)** per i metodi farmaceutici, la guida **Eurachem** *The Fitness for Purpose of Analytical Methods* (3ª ed., 2025) per i laboratori di ogni settore, **SANTE/11312/2021 v2026** per i residui di pesticidi e il **Reg. (UE) 2021/808** per i residui di farmaci veterinari. Per approfondire: Skoog, Holler e Crouch, *Principles of Instrumental Analysis* (tecniche strumentali); Harvey, *Modern Analytical Chemistry* (fondamenti della chimica analitica quantitativa); Matuszewski et al., *Anal. Chem.* 2003, 75, 3019–3030, l'articolo che ha introdotto lo schema a tre set (Capitolo 13); la norma ISO 5725 su esattezza e precisione. I riferimenti completi sono nell'Appendice D.

Punti chiave

- Blocco A (standard e calibrazione) e blocco B (validazione) insieme rendono un metodo pronto per la routine.
- La tracciabilità al SI è la base dell'esattezza: una buona regressione da sola non la garantisce.
- La curva si giudica con residui e retrocalcolo, non con r^2.
- I criteri valgono su tutto il range, dal LLOQ all'ULOQ.
- Il SIL-IS è la scelta ottimale in LC-MS/MS, nei limiti di ciò che il rapporto può compensare.
- In routine il metodo resta sotto controllo con curve, QC e, quando cambia qualcosa, validazione parziale o cross-validazione.

Mettiti alla prova

Argomento del quiz: *Sintesi e messaggi chiave*. Per un ripasso finale, svolgi una prova mista su tutti gli argomenti di calibrazione e validazione.

Appendici

Formulario, tavole statistiche, glossario italiano–inglese e riferimenti bibliografici per lo studio e la consultazione rapida.

Appendice A — Formulario

Le formule della dispensa raccolte per argomento. Il capitolo indicato tra parentesi contiene la spiegazione completa e gli esempi svolti.

A.1 Statistica descrittiva (Capitolo 4)

Tabella A.1 — Indici di posizione e di dispersione

Grandezza	Formula	Note
Media	$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$	stessa unità dei dati
Mediana	valore centrale dei dati ordinati	con n pari: media dei due valori centrali
Varianza campionaria	$s^2 = \frac{\sum(x_i - \bar{x})^2}{n-1}$	unità al quadrato
Deviazione standard	$s = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n-1}}$	$\nu = n - 1$ gradi di libertà
Deviazione standard relativa	$\mathrm{RSD} = \mathrm{CV}\% = \frac{s}{\bar{x}} \times 100$	adimensionale
Errore standard della media	$s_{\bar{x}} = \frac{s}{\sqrt{n}}$	diminuisce con $\sqrt{n}$

A.2 Errori di misura (Capitoli 2–3)

Tabella A.2 — Errori, bias e recupero

Grandezza	Formula	Note
Errore assoluto	$E = x_i - \mu$	μ = valore vero o di riferimento
Errore relativo	$E_{\mathrm{rel}}\% = \frac{x_i - \mu}{\mu} \times 100$	con segno
Bias (esattezza)	$b = \bar{x} - \mu$	si calcola sulla media
Bias relativo	$b\% = \frac{\bar{x} - \mu}{\mu} \times 100$	criterio ICH M10: entro ±15% (±20% al LLOQ)
Recupero di un'aggiunta	$R\% = \frac{C_{\text{arricchito}} - C_{\text{non arricchito}}}{C_{\text{aggiunto}}} \times 100$	verifica dell'esattezza

A.3 Inferenza e valori anomali (Capitolo 5)

$$\bar{x} \pm t_\nu \cdot \frac{s}{\sqrt{n}}$$

Intervallo di fiducia della media, con t_ν a due code e $\nu = n - 1$ (Tabella B.1).

$$Q = \frac{|x_{\text{sosp}} - x_{\text{vicino}}|}{x_{\max} - x_{\min}}$$

Test Q di Dixon: il dato è anomalo se $Q > Q_{\text{crit}}$ (Tabella B.2).

$$G = \frac{|x_{\text{sosp}} - \bar{x}|}{s}$$

Test di Grubbs, con $\bar{x}$ e s calcolati includendo il valore sospetto: il dato è anomalo se $G > G_{\text{crit}}$ (Tabella B.3).

A.4 Incertezza di misura (Capitolo 6)

Tabella A.3 — Valutazione e combinazione delle incertezze

Grandezza	Formula	Note
Tipo A (media di n repliche)	$u_A = \frac{s}{\sqrt{n}}$	se il risultato è la media delle n repliche
Tipo B, distribuzione rettangolare	$u = \frac{a}{\sqrt{3}}$	a = semiampiezza (tolleranza ±a)
Tipo B, distribuzione triangolare	$u = \frac{a}{\sqrt{6}}$	valori centrali più probabili
Incertezza composta (somme e differenze)	$u_c = \sqrt{u_1^2 + u_2^2 + \cdots + u_n^2}$	contributi indipendenti, stessa unità
Incertezza composta (prodotti e quozienti)	$\frac{u_c(y)}{y} = \sqrt{\sum \left(\frac{u_i}{x_i}\right)^2}$	le incertezze relative si sommano in quadratura
Incertezza estesa	$U = k \cdot u_c$	$k = 2$: circa 95% (Tabella B.4)

A.5 Regressione lineare (Capitoli 7–8)

$$b_1 = \frac{\sum (x_i - \bar{x})\,(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} \qquad b_0 = \bar{y} - b_1 \bar{x}$$

$$\hat{y}_i = b_0 + b_1 x_i \qquad e_i = y_i - \hat{y}_i$$

$$x_0 = \frac{y_0 - b_0}{b_1}$$

Concentrazione di un campione dal suo segnale y_0 (predizione inversa), valida solo dentro l'intervallo di calibrazione.

Tabella A.4 — Scomposizione della variabilità e indici di bontà

Grandezza	Formula	Significato
Somma dei quadrati totale	$\mathrm{SQT} = \sum (y_i - \bar{y})^2$	variabilità totale dei segnali
Somma dei quadrati della regressione	$\mathrm{SQR} = \sum (\hat{y}_i - \bar{y})^2$	variabilità spiegata dalla retta
Somma dei quadrati dell'errore	$\mathrm{SQE} = \sum (y_i - \hat{y}_i)^2$	variabilità residua
Relazione	$\mathrm{SQT} = \mathrm{SQR} + \mathrm{SQE}$	vale per la retta dei minimi quadrati con intercetta
Coefficiente di determinazione	$r^2 = \frac{\mathrm{SQR}}{\mathrm{SQT}} = 1 - \frac{\mathrm{SQE}}{\mathrm{SQT}}$	frazione di variabilità spiegata
Deviazione standard dei residui	$s_{y/x} = \sqrt{\frac{\mathrm{SQE}}{n-2}}$	$n-2$ gradi di libertà
Pesi (regressione pesata)	$w_i = \frac{1}{x_i^2}$ oppure $\frac{1}{x_i}$	dati eteroschedastici

A.6 Calibrazione (Capitoli 9–10 e 16)

Tabella A.5 — Calibrazione, diluizioni e standard

Grandezza	Formula	Note
Diluizione	$C_1 V_1 = C_2 V_2$	stesse unità di volume
Standard interno	$R = \frac{A_{\text{analita}}}{A_{\text{SI}}}$	la curva si costruisce con R
Aggiunte multiple	$C_x = \frac{b_0}{b_1}$	intercetta con l'asse x in $-C_x$
Bias di retrocalcolo	$\frac{C_{\text{retro}} - C_{\text{nom}}}{C_{\text{nom}}} \times 100$	±15% (±20% al LLOQ)
Propagazione in diluizioni seriali	$\frac{u_c(c)}{c} = \sqrt{\sum \left(\frac{u_i}{x_i}\right)^2}$	ogni passaggio aggiunge incertezza

A.7 Parametri di validazione (Capitoli 11–14)

Tabella A.6 — Sensibilità, limiti e precisione attesa

Grandezza	Formula	Note
Sensibilità di calibrazione	$b_1 = \frac{\Delta S}{\Delta C}$	pendenza della retta
Sensibilità analitica	$\gamma = \frac{b_1}{s_S}$	dipende dalla concentrazione
Limite di rivelabilità	$\mathrm{LOD} = \frac{3 \cdot s_B}{b_1}$	S/N ≈ 3
Limite di quantificazione	$\mathrm{LOQ} = \frac{10 \cdot s_B}{b_1}$	S/N ≈ 10
Curva di Horwitz	$\mathrm{CV\%} \approx 2^{(1-0{,}5\log C)}$	C = frazione di massa adimensionale, log in base 10

Tabella A.7 — Recupero, effetto matrice ed efficienza di processo (schema di Matuszewski)

Grandezza	Formula	Significato
Recupero	$\mathrm{RE\%} = \frac{\text{Set 3}}{\text{Set 2}} \times 100$	perdite nell'estrazione
Matrix factor	$\mathrm{MF\%} = \frac{\text{Set 2}}{\text{Set 1}} \times 100$	< 100%: soppressione; > 100%: aumento del segnale
Efficienza di processo	$\mathrm{PE\%} = \frac{\text{Set 3}}{\text{Set 1}} \times 100 = \frac{\mathrm{RE} \times \mathrm{MF}}{100}$	effetto complessivo
Carry-over	$\frac{A_{\text{bianco dopo ULOQ}}}{A_{\mathrm{LLOQ}}} \times 100$	≤ 20% (≤ 5% per lo SI)

A.8 Criteri di accettazione ICH M10 (metodi cromatografici)

Tabella A.8 — Riepilogo dei criteri ICH M10

Parametro	Criterio
Calibranti (retrocalcolo)	entro ±15% del nominale (±20% al LLOQ); almeno il 75% conformi, con almeno 6 livelli
Esattezza dei QC	media entro ±15% del nominale (±20% al LLOQ QC)
Precisione dei QC	CV ≤ 15% (≤ 20% al LLOQ QC), entro la serie e tra le serie
Accettazione di una serie analitica	almeno 2/3 dei QC entro ±15% e almeno il 50% a ogni livello
Selettività	risposta nel bianco ≤ 20% del LLOQ (analita) e ≤ 5% (standard interno), in almeno 6 lotti
Effetto matrice	QC basso e alto in almeno 6 lotti: esattezza ±15%, CV ≤ 15%
Carry-over	bianco dopo l'ULOQ ≤ 20% del LLOQ e ≤ 5% dello SI
Integrità della diluizione	esattezza ±15%, CV ≤ 15%
Livelli dei QC	LLOQ; basso ≤ 3 × LLOQ; medio 30–50% dell'intervallo; alto ≥ 75% dell'ULOQ
Cross-validazione	nessun criterio di accettazione fisso: il bias tra i metodi si valuta statisticamente

Appendice B — Tavole statistiche

Le tavole riportano i valori critici usati negli esempi e nel quiz. I valori di t, di G e del fattore di copertura sono calcolati dalle distribuzioni teoriche; i valori critici del test Q di Dixon sono tratti dalla tabella di Rorabacher (1991).

B.1 Valori critici della t di Student (due code)

Si usa la colonna del livello di fiducia desiderato e la riga dei gradi di libertà $\nu = n - 1$ (per l'intervallo di fiducia della media). Per $\nu \to \infty$ i valori coincidono con quelli della distribuzione normale.

Tabella B.1 — Valori di t_ν per intervalli di fiducia bilaterali

ν	90%	95%	99%
1	6,314	12,706	63,657
2	2,920	4,303	9,925
3	2,353	3,182	5,841
4	2,132	2,776	4,604
5	2,015	2,571	4,032
6	1,943	2,447	3,707
7	1,895	2,365	3,499
8	1,860	2,306	3,355
9	1,833	2,262	3,250
10	1,812	2,228	3,169
11	1,796	2,201	3,106
12	1,782	2,179	3,055
13	1,771	2,160	3,012
14	1,761	2,145	2,977
15	1,753	2,131	2,947
16	1,746	2,120	2,921
17	1,740	2,110	2,898
18	1,734	2,101	2,878
19	1,729	2,093	2,861
20	1,725	2,086	2,845
21	1,721	2,080	2,831
22	1,717	2,074	2,819
23	1,714	2,069	2,807
24	1,711	2,064	2,797
25	1,708	2,060	2,787
26	1,706	2,056	2,779
27	1,703	2,052	2,771
28	1,701	2,048	2,763
29	1,699	2,045	2,756
30	1,697	2,042	2,750

ν	90%	95%	99%
40	1,684	2,021	2,704
60	1,671	2,000	2,660
120	1,658	1,980	2,617
∞	1,645	1,960	2,576

B.2 Valori critici del test Q di Dixon

Il valore sospetto è anomalo se $Q_{\text{calc}} > Q_{\text{crit}}$ al livello di fiducia scelto (di norma 95%).

Tabella B.2 — Valori critici di Q (test a due code)

n	$Q_{90\%}$	$Q_{95\%}$	$Q_{99\%}$
3	0,941	0,970	0,994
4	0,765	0,829	0,926
5	0,642	0,710	0,821
6	0,560	0,625	0,740
7	0,507	0,568	0,680
8	0,468	0,526	0,634
9	0,437	0,493	0,598
10	0,412	0,466	0,568

B.3 Valori critici del test di Grubbs

$G = |x_{\text{sosp}} - \bar{x}|/s$, con media e deviazione standard calcolate includendo il valore sospetto. Valori per il test a due code.

Tabella B.3 — Valori critici di G (test a due code)

n	$G_{95\%}$	$G_{99\%}$
3	1,154	1,155
4	1,481	1,496
5	1,715	1,764
6	1,887	1,973
7	2,020	2,139
8	2,127	2,274
9	2,215	2,387
10	2,290	2,482
11	2,355	2,564
12	2,412	2,636
13	2,462	2,699
14	2,507	2,755
15	2,548	2,806
16	2,586	2,852
17	2,620	2,894

n	$G_{95\%}$	$G_{99\%}$
18	2,652	2,932
19	2,681	2,968
20	2,708	3,001

B.4 Fattore di copertura e livello di fiducia

Per una distribuzione normale (o approssimativamente normale, con un numero sufficiente di gradi di libertà) il fattore di copertura k corrisponde al livello di fiducia indicato.

Tabella B.4 — Fattore di copertura k e livello di fiducia

k	**Livello di fiducia**
1	68,27%
1,645	90,00%
1,960	95,00%
2	95,45%
2,576	99,00%
3	99,73%

Appendice C — Glossario italiano–inglese

I termini sono in ordine alfabetico italiano. La terza colonna riporta il significato in breve; il capitolo indicato contiene la spiegazione completa.

Tabella C.1 — Glossario dei termini tecnici

Italiano	**English**	**Significato (capitolo)**
Accuratezza	Accuracy	grado di concordanza tra un singolo risultato e il valore vero; comprende esattezza e precisione (3)
Aggiunte multiple (aggiunte standard)	Standard addition	calibrazione eseguita aggiungendo quantità note di analita al campione stesso (10)
Bianco di matrice	Blank matrix sample	matrice priva di analita e di standard interno, trattata come un campione (10)
Bianco dei reagenti	Reagent blank	solvente con tutti i reagenti, senza campione (10)
Bias	Bias	differenza tra la media dei risultati e il valore di riferimento (2–3)
Campione zero	Zero sample	matrice bianca con il solo standard interno (10)
Carry-over	Carry-over	trascinamento di analita da un'iniezione alla successiva (14)
Coefficiente di determinazione	Coefficient of determination (r^2)	frazione della variabilità dei segnali spiegata dalla retta (7)

Italiano	English	Significato (capitolo)
Coefficiente di variazione	Coefficient of variation (CV%)	deviazione standard relativa espressa in percentuale (4)
Controllo qualità (QC)	Quality control sample	campione a concentrazione nota usato per accettare una serie analitica (12)
Cross-validazione	Cross-validation	confronto tra due metodi validati sugli stessi campioni (16)
Curva di calibrazione	Calibration curve	relazione tra segnale e concentrazione ottenuta dagli standard (7, 10)
Deviazione standard	Standard deviation	misura della dispersione dei dati attorno alla media (4)
Effetto matrice	Matrix effect	variazione del segnale dovuta ai componenti della matrice (13)
Efficienza di processo	Process efficiency	effetto combinato di recupero ed effetto matrice (13)
Errore casuale	Random error	errore che varia in modo imprevedibile tra misure ripetute (2)
Errore grossolano	Gross error	errore macroscopico che invalida la misura (2)
Errore sistematico	Systematic error	errore che resta costante o varia in modo prevedibile (2)
Errore standard della media	Standard error of the mean	deviazione standard della media, $s/\sqrt{n}$ (5)
Esattezza	Trueness	grado di concordanza tra la media di molti risultati e il valore vero (3, 12)
Eteroschedasticità	Heteroscedasticity	varianza del segnale che cambia con la concentrazione (8)
Fattore di copertura	Coverage factor (k)	moltiplicatore che trasforma l'incertezza composta in incertezza estesa (6)
Gradi di libertà	Degrees of freedom (ν)	numero di dati indipendenti disponibili per stimare la dispersione (4–5)
Incertezza di misura	Measurement uncertainty	parametro che descrive la dispersione dei valori attribuibili al misurando (1, 6)
Incertezza estesa	Expanded uncertainty (U)	incertezza composta moltiplicata per il fattore di copertura (6)
Integrità della diluizione	Dilution integrity	prova che la diluizione con matrice bianca non altera il risultato (14)
Intervallo di calibrazione	Calibration range	intervallo dal LLOQ all'ULOQ coperto dagli standard (11)
Intervallo di confidenza	Confidence interval	intervallo che contiene il valore vero con il livello di fiducia scelto (5)
Limite di linearità (LOL)	Limit of linearity	concentrazione oltre la quale la risposta non è più lineare (11)
Limite di quantificazione (LOQ)	Limit of quantification	concentrazione minima determinabile con precisione ed esattezza accettabili (11)
Limite di rivelabilità (LOD)	Limit of detection	concentrazione minima distinguibile dal bianco (11)
LLOQ / ULOQ	Lower / upper limit of quantification	standard più basso e più alto della curva di calibrazione (11)

Italiano	English	Significato (capitolo)
Materiale di riferimento certificato (CRM)	Certified reference material	materiale con valori certificati, incertezza e tracciabilità (9)
Matrix factor	Matrix factor (MF)	rapporto tra il segnale in matrice e quello in solvente (13)
Media	Mean	somma dei dati divisa per il loro numero (4)
Mediana	Median	valore centrale dei dati ordinati (4)
Metodo assoluto / relativo	Absolute / relative method	metodo che non richiede / richiede una curva di calibrazione (10)
Minimi quadrati	Least squares	metodo che sceglie la retta minimizzando la somma dei quadrati dei residui (7)
Precisione	Precision	grado di concordanza tra misure ripetute (3, 12)
Precisione intermedia	Intermediate precision	precisione nello stesso laboratorio con giorni, operatori o strumenti diversi (3, 12)
Range dinamico lineare	Linear dynamic range	intervallo dal LOQ al LOL in cui la risposta è lineare (11)
Recupero	Recovery	frazione di analita ritrovata dopo l'estrazione (2, 13)
Residuo	Residual	differenza tra il segnale misurato e quello previsto dalla retta (7–8)
Ripetibilità	Repeatability	precisione nelle stesse condizioni, in breve tempo (3)
Riproducibilità	Reproducibility	precisione tra laboratori diversi (3)
Robustezza	Robustness	insensibilità del metodo a piccole variazioni deliberate dei parametri (14)
Selettività	Selectivity	capacità di misurare l'analita senza interferenze dalla matrice (11)
Sensibilità	Sensitivity	variazione del segnale per unità di concentrazione (pendenza) (11)
Solidità	Ruggedness	stabilità dei risultati al cambiare del contesto (laboratorio, strumento, operatore) (14)
Specificità	Specificity	selettività assoluta: risposta al solo analita (11)
Stabilità	Stability	conservazione della concentrazione dell'analita nelle condizioni d'uso (14)
Standard interno (SI)	Internal standard (IS)	composto aggiunto a concentrazione costante per correggere le variazioni (10)
Standard interno marcato con isotopi stabili	Stable isotope-labeled internal standard (SIL-IS)	analogo isotopico dell'analita, standard interno ideale in LC-MS/MS (10, 16)
Standard primario	Primary standard	sostanza di purezza elevata e nota usata come riferimento (9)
Test di Grubbs	Grubbs test	test per un valore anomalo basato su $G = \|x - \bar{x}\|/s$ (5)
Test Q di Dixon	Dixon's Q test	test per un valore anomalo in piccoli gruppi di dati (5)

Italiano	English	Significato (capitolo)
Tracciabilità metrologica	Metrological traceability	collegamento di un risultato a un riferimento tramite una catena ininterrotta di tarature (9)
Validazione	Validation	dimostrazione documentata che un metodo è adatto allo scopo (11)

Appendice D — Riferimenti

D.1 Linee guida e norme

- ICH M10 — *Bioanalytical Method Validation and Study Sample Analysis*, International Council for Harmonisation, 2022.
- ICH Q2(R2) — *Validation of Analytical Procedures*, International Council for Harmonisation, 2023.
- ICH Q14 — *Analytical Procedure Development*, International Council for Harmonisation, 2023.
- EMA — *Guideline on Bioanalytical Method Validation*, European Medicines Agency, 2011 (sostituita da ICH M10 dal 2023).
- FDA — *Bioanalytical Method Validation, Guidance for Industry*, U.S. Food and Drug Administration, 2018 (superata da ICH M10 per i metodi nel suo ambito).
- Eurachem — *The Fitness for Purpose of Analytical Methods: A Laboratory Guide to Method Validation and Related Topics*, 3ª edizione, 2025.
- Eurachem/CITAC — *Quantifying Uncertainty in Analytical Measurement* (guida QUAM, CG4), 3ª edizione, 2012.
- SANTE/11312/2021 v2026 — *Analytical Quality Control and Method Validation Procedures for Pesticide Residues Analysis in Food and Feed*, Commissione europea, in vigore dal 1° gennaio 2026 (sostituisce la v2 del 2024).
- Regolamento di esecuzione (UE) 2021/808 sul rendimento dei metodi analitici per i residui di sostanze farmacologicamente attive negli animali destinati alla produzione di alimenti.
- JCGM 100:2008 — *Evaluation of measurement data — Guide to the expression of uncertainty in measurement* (GUM).
- JCGM 200:2012 — *International vocabulary of metrology — Basic and general concepts and associated terms* (VIM, 3ª edizione).
- ISO 5725 (serie) — *Accuracy (trueness and precision) of measurement methods and results*.
- ISO 8466-1 — *Water quality — Calibration and evaluation of analytical methods — Part 1: Linear calibration function*.
- ISO 17034:2016 — *General requirements for the competence of reference material producers*.
- ISO/IEC 17025:2017 — *General requirements for the competence of testing and calibration laboratories*.

D.2 Articoli

- Matuszewski B. K., Constanzer M. L., Chavez-Eng C. M. — Strategies for the assessment of matrix effect in quantitative bioanalytical methods based on HPLC–MS/MS. *Analytical Chemistry* 75 (2003) 3019–3030.
- Horwitz W. — Evaluation of analytical methods used for regulation of foods and drugs. *Analytical Chemistry* 54 (1982) 67A–76A.
- Rorabacher D. B. — Statistical treatment for rejection of deviant values: critical values of Dixon's "Q" parameter and related subrange ratios at the 95% confidence level. *Analytical Chemistry* 63 (1991) 139–146.
- Grubbs F. E. — Procedures for detecting outlying observations in samples. *Technometrics* 11 (1969) 1–21.

- Bland J. M., Altman D. G. — Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet* 327 (1986) 307–310.

D.3 Testi di riferimento

- Skoog D. A., Holler F. J., Crouch S. R. — *Principles of Instrumental Analysis*, 7ª edizione, Cengage Learning.
- Harvey D. — *Modern Analytical Chemistry*, McGraw-Hill.
- Miller J. N., Miller J. C., Miller R. D. — *Statistics and Chemometrics for Analytical Chemistry*, 7ª edizione, Pearson, 2018.