

RANGE DINAMICO E LINEARE

Il **range** è l'intervallo di concentrazione esplorato nel corso delle misurazioni.

Il **range dinamico** è l'intervallo di concentrazione nel quale il segnale varia con la concentrazione: i limiti inferiore e superiore del range dinamico corrispondono, rispettivamente, al limite di rivelabilità ed alla più alta concentrazione alla quale un incremento di concentrazione produce ancora un incremento di segnale.

Il **range lineare** esprime l'intervallo di concentrazione nel quale il segnale varia linearmente con la concentrazione.

La costruzione del diagramma di calibrazione implica l'adozione di un metodo di regressione. Quello più generalmente adottato è il metodo di regressione lineare ordinaria dei minimi quadrati (OLLSR).

Condizioni necessarie per poter eseguire una OLSR:

- la relazione funzionale segnale/concentrazione deve essere del tipo FI: gli errori sperimentali associati alla variabile indipendente (concentrazione, quantità) devono essere trascurabili rispetto a quelli associati alla variabile dipendente (segnale);
- gli errori associati alla variabile dipendente devono essere distribuiti normalmente;
- il sistema analitico deve essere omoschedastico, ovvero la precisione non deve cambiare significativamente al variare della concentrazione;
- il segnale deve essere funzione lineare della concentrazione.

È consigliabile inoltre verificare la presenza di dati anomali (outliers).

Ovviamente, l'accettazione dei risultati ottenuti per mezzo del metodo OLSR implica la verifica delle condizioni che ne consentono l'uso!

Provvedimenti:

- se la relazione funzionale è di tipo FII (entrambe le variabili, dipendente ed indipendente, sono affette da errore sperimentale) si deve ricorrere a metodi specifici di regressione.*
- se gli errori associati alla variabile dipendente non sono distribuiti normalmente si deve ricorrere a metodi di regressione più robusti (per es. LMS: Least median of squares);
- se il sistema è eteroschedastico (la precisione varia significativamente con la concentrazione) è necessario ricorrere a trasformazioni dei dati, o a metodi di regressione pesata;
- se il segnale non è funzione lineare della concentrazione, si deve ridurre l'intervallo di concentrazione esplorato oppure, dopo aver suddiviso l'intervallo in sub-intervalli, si deve eseguire una regressione lineare in ciascuno dei sub-intervalli, oppure ancora si deve ricorrere a metodi di regressione non lineare.

* M.G. Natrella, *Experimental Statistics*, National Bureau of Standards Handbook, US Department of Commerce, National Technical Information Service, Gaithersburg, MD, USA, (1991)

I test statistici a disposizione sono numerosi.

Test per normalità:

- test di Shapiro-Wilk

Test per omoschedasticità:

- test-F alla concentrazione minima e massima
- test di Bartlett
- test di Cochran

Test per linearità:

- analisi grafica dei residui della regressione
- test-F su $s_{y/x}$ della regressione quadratica e lineare (test di Mandel)
- lack-of-fit test

Test per outliers:

- test-F su $s_{y/x}$ prima e dopo eliminazione del dato sospetto

I test più affidabili richiedono l'esecuzione di repliche di ciascuna soluzione standard.

ESEMPIO N. 8 – Nel corso della procedura di validazione del metodo spettrofotometrico VIS di analisi del nitrito nelle acque (metodo di Griess), si desidera verificare se il sistema analitico è omoschedastico. Allo scopo sono eseguite 10 repliche di due soluzioni contenenti una concentrazione di nitrito uguale al limite di rivelabilità ($2,173 \cdot 10^{-7}$ M, dall'Esempio N. 4, corrispondente al limite inferiore del range lineare) ed al limite superiore del range lineare ($6,00 \cdot 10^{-5}$ M, vicina a quella corrispondente alla massima assorbanza consentita dallo strumento). I risultati sono:

$$C_{NO_2^-} = 2,173 \cdot 10^{-7} \text{ M}$$

$$A_i: 0,005; 0,006; 0,004; 0,011; 0,008; 0,007; 0,013; 0,012; 0,005; 0,007$$

$$C_{NO_2^-} = 6,00 \cdot 10^{-5} \text{ M}$$

$$A_i: 2,984; 2,828; 2,769; 2,912; 2,934; 2,902; 2,889; 2,900; 2,997; 2,810$$

Verificare la normalità dei dati mediante test di Shapiro-Wilk e l'omogeneità della varianza mediante un test-F (P= 95% e P = 99%)

Per eseguire il test di Shapiro-Wilk i dati relativi alle due concentrazioni ($C_b = C_{NO_2^-} = 2,173 \cdot 10^{-7} \text{ M}$ e $C_a = C_{NO_2^-} = 6,00 \cdot 10^{-5} \text{ M}$) sono disposti in ordine crescente* (vedere esempio a fianco).

Quindi sono calcolati i valori di

$$S2_b = 8,960 \cdot 10^{-5} \quad S2_a = 0,0478$$

e di

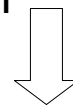
$$b_b = 8,998 \cdot 10^{-3} \quad e \quad b_a = 0,2137$$

Quindi è possibile calcolare la statistica del test:

$$W_b = 0,9037 \quad e, \quad \text{rispettivamente,} \quad W_a = 0,9560$$

Dato che il valore critico è $W_{(\alpha = 0,05; n=10)} = 0,842$, l'ipotesi di normalità non viene rifiutata in entrambi i casi.

$$A_a = \begin{bmatrix} 2.984 \\ 2.828 \\ 2.769 \\ 2.912 \\ 2.934 \\ 2.902 \\ 2.889 \\ 2.900 \\ 2.997 \\ 2.810 \end{bmatrix} \quad x_a = \begin{bmatrix} 2.769 \\ 2.810 \\ 2.828 \\ 2.889 \\ 2.900 \\ 2.902 \\ 2.912 \\ 2.934 \\ 2.984 \\ 2.997 \end{bmatrix}$$



* Il test è descritto nell'Appendice 3 del testo a lato.

Avendo accettato l'ipotesi di normalità, si può procedere con il calcolo.

Dopo aver stimato le varianze dei due set di dati

$$s_{\min}^2 = \frac{\sum_{i=1}^{10} [(A_{\min})_i - \bar{A}_{\min}]^2}{9} = 9,956 \cdot 10^{-6}$$

$$s_{\max}^2 = \frac{\sum_{i=1}^{10} [(A_{\max})_i - \bar{A}_{\max}]^2}{7} = 5,308 \cdot 10^{-3}$$

si calcola il valore di F

$$F_{\text{exp}} = \frac{s_{\max}^2}{s_{\min}^2} = 533.17$$

e lo si confronta con i valori critici (ad una coda):

$$F_{(1-\alpha,9,9)} = 3,179 \quad (\alpha = 0,05) \quad \text{e} \quad F_{(1-\alpha,9,9)} = 5,351 \quad (\alpha = 0,01)$$

Dato che il valore sperimentale è maggiore di entrambi i livelli critici, si può concludere che la varianza non è omogenea nei limiti di entrambi i livelli di fiducia.

Qualora siano disponibili repliche eseguite a più di due concentrazioni differenti, si può usare il test di Cochran. Questo però richiede campioni della stessa numerosità.

ESEMPIO N. 9 – In seguito al risultato delle prove citate nell'Esempio N. 8, sono analizzate altre due soluzioni aventi concentrazione $4,74 \cdot 10^{-6}$ M e $1,90 \cdot 10^{-5}$ M. I risultati sono:

$$C_{NO_2^-} = 4,74 \cdot 10^{-6} \text{ M}$$

$$A_i: 0,232; 0,211; 0,221; 0,207; 0,247; 0,237$$

$$C_{NO_2^-} = 1,90 \cdot 10^{-5} \text{ M}$$

$$A_i: 0,902; 0,861; 0,931; 0,891; 0,986; 0,918$$

Valutare l'omogeneità della varianza ($P = 95\%$) usando questi dati e quelli delle *prime sei misurazioni* eseguite alle due concentrazioni esplorate nell'Esempio n. 8.

Dopo aver stimato le varianze dei quattro set di dati (qui riportati in ordine crescente di concentrazione)

$$s_1^2 = 6,167 \cdot 10^{-6}$$

$$s_2^2 = 2,418 \cdot 10^{-4}$$

$$s_3^2 = 1,793 \cdot 10^{-3}$$

$$s_4^2 = 5,973 \cdot 10^{-3}$$

si calcola il valore di g

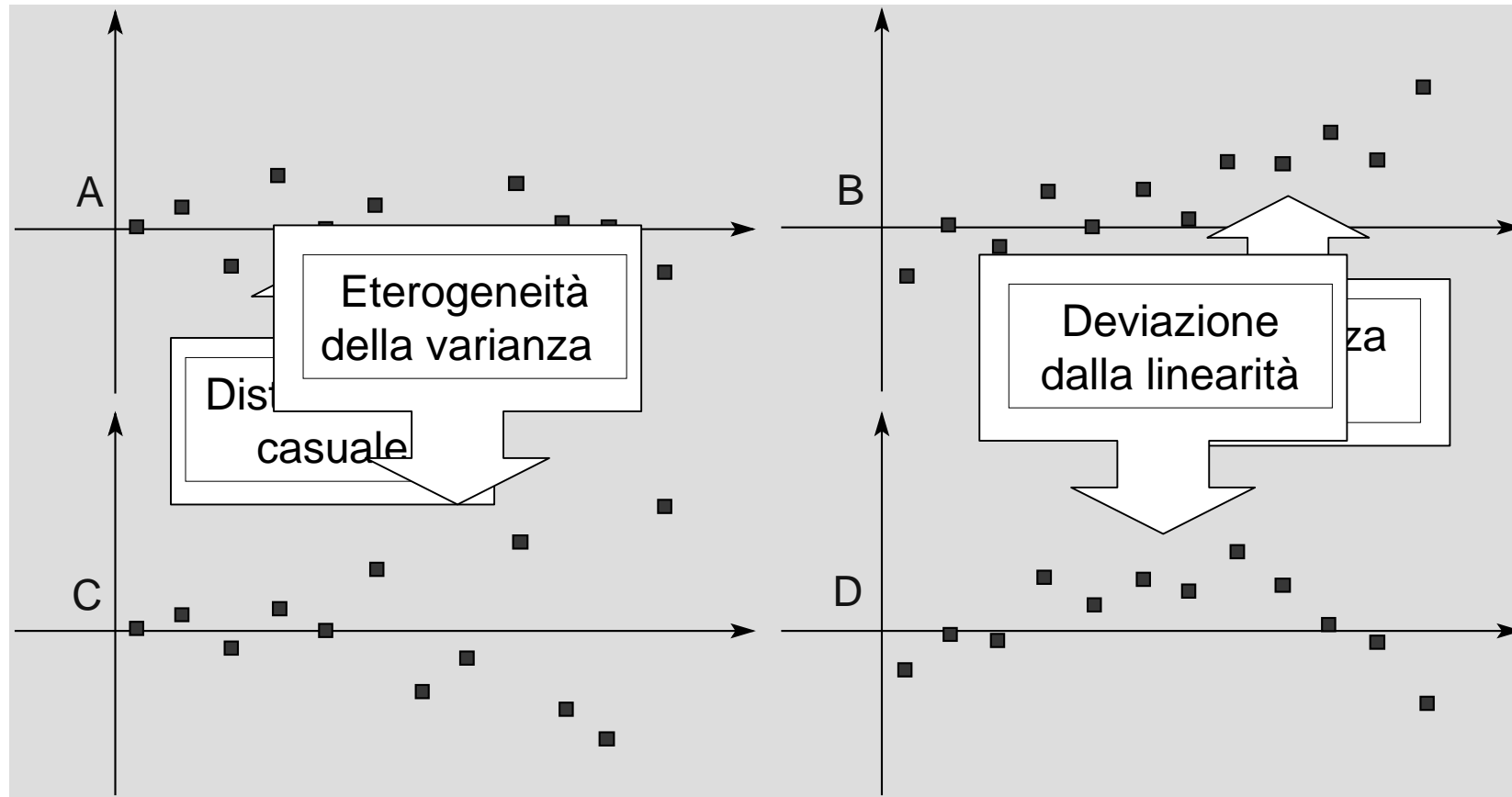
$$g = \frac{s_{\max}^2}{\sum_{i=1}^4 s_i^2} = \frac{5,973 \cdot 10^{-3}}{5,973 \cdot 10^{-3} + 1,793 \cdot 10^{-3} + 2,418 \cdot 10^{-4} + 6,167 \cdot 10^{-6}} = 0,7453$$

e lo si confronta con il valore critico per il livello di significatività prescelto, il numero di osservazioni (6) ed il numero di varianze (4). In questo caso $g_{0,95;6;4} = 0,5895$.

Dato che il valore sperimentale è maggiore di quello critico, al livello di fiducia del 95% non si può escludere che il sistema non sia omoschedastico.

I valori critici di g sono riportati nelle Tavole statistiche (Cap. 11).

Il test più semplice per la verifica della linearità è quello che sfrutta l'analisi grafica dei residui della regressione.



L'analisi è più significativa se sono eseguite misurazioni replicate dei campioni usati per costruire il diagramma di calibrazione.

Un'analisi di regressione completa, da condurre mediante misurazioni replicate di ciascun campione standard, consiste di due fasi.

Procedura:

Fase A

- analizzare una volta il bianco più almeno 6 materiali di riferimento o bianchi fortificati a diversa concentrazione;
- riportare il segnale in funzione della concentrazione e identificare approssimativamente il range lineare e gli estremi del range dinamico;

Fase B

- analizzare (tre volte) almeno 6 materiali di riferimento o bianchi fortificati a diversa concentrazione entro il range lineare (*randomizzare!*);
- calcolare i coefficienti di regressione più appropriati (coefficiente di correlazione e di determinazione, pendenza, intercetta e relativi intervalli di fiducia);
- verificare possibili outliers (test-F), linearità e omoschedasticità.

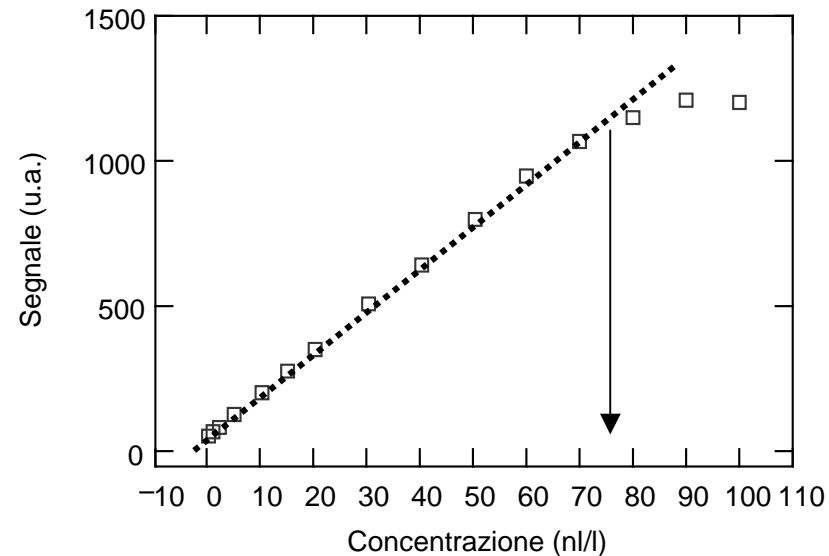
ESEMPIO N. 10 - Si desidera sviluppare un metodo rapido per la determinazione dell'etilene sviluppata da specie vegetali palustri mediante spettroscopia fotoacustica (PAS) con eccitazione mediante laser a CO₂.

Nella fase preliminare sono analizzate 13 miscele gassose contenenti concentrazioni crescenti di etilene. I risultati sono riportati qui di seguito:

$C_{C_2H_2}(nL/L)$	=	0,1	1,0	2,0	5,0	10,0	15,0
Segnale (u.a.)	=	53,75	67,21	82,14	126,80	201,0	274,9
$C_{C_2H_2}(nL/L)$	=	20,0	30,0	40,0	50,0	60,0	70,0
Segnale (u.a.)	=	350,0	508,1	642,1	799,9	948,0	1069
$C_{C_2H_2}(nL/L)$	=	80,0	90,0	100,0			
Segnale (u.a.)	=	1152	1207	1200			

Valutare graficamente il range dinamico.

Riportando in grafico i punti sperimentali, si può notare che la sensibilità (segnale/concentrazione) è circa uguale a 15 u.a. (stima grafica), ed è costante fino a livelli di concentrazione dell'ordine di 70 nL/L. Il range dinamico sembra estendersi da 0,1 nL/L a circa di 90 nL/L.



Il range lineare si estende, approssimativamente, fino a $C = 70-80$ nL/L.

ESEMPIO N. 11 – Determinazione dell'etilene mediante PAS. Avendo determinato il range dinamico e, approssimativamente, quello lineare (vedere Esempio n. 10), si costruisce il diagramma di calibrazione mediante regressione lineare ordinaria ai minimi quadrati (OLLSR). Allo scopo, sono analizzati in triplo sia il bianco tal quale che il bianco fortificato con etilene a otto livelli di concentrazione nel range lineare. I risultati medi delle tre repliche sono i seguenti:

$C_{C_2H_2}(nL/L)$	=	0,0	10,0	20,0	30,0	40,0
Segnale (u.a.)	=	29	215	346	477	646
$C_{C_2H_2}(nL/L)$	=	50,0	60,0	70,0	80,0	
Segnale (u.a.)	=	775	958	1055	1142	

Sapendo che il limite di rivelabilità ($\alpha = \beta = 0,05$, calcolato in precedenti esperimenti) è 1,5 nL/L, valutare quantitativamente il range lineare e la relazione funzionale Segnale/Concentrazione (P = 95%).

La relazione funzionale risultante dall'analisi di tutti i dati è:

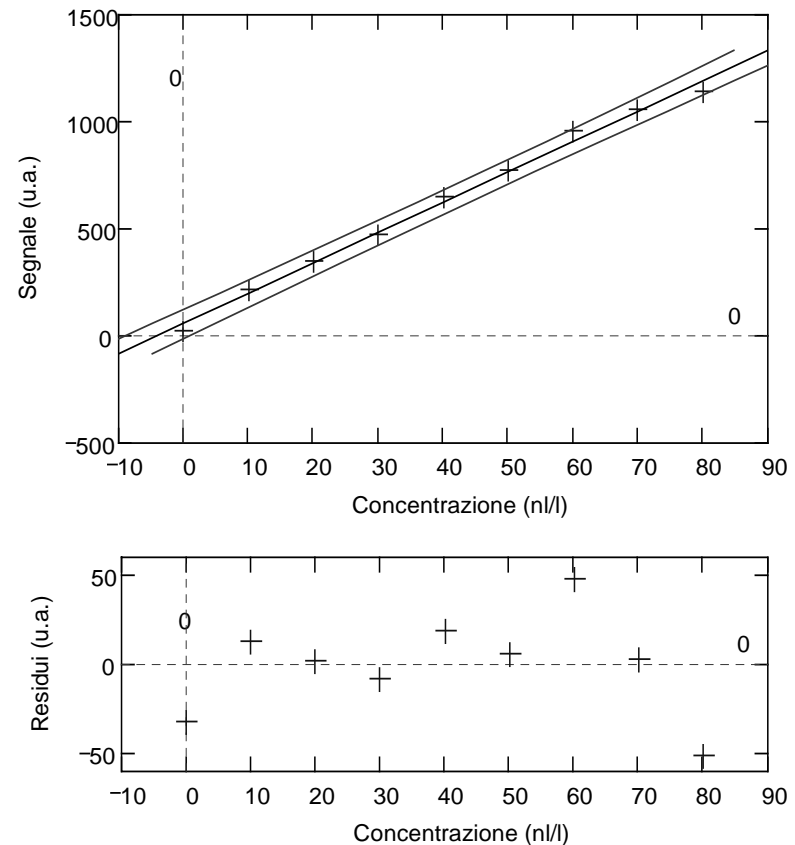
$$S(C)_{(u.a.)} = (14,16 \pm 0,94) \cdot C_{(nL/L)} + (61 \pm 45) \quad (P = 95\%; \nu = 7)$$

Il coefficiente di correlazione è $r = 0,9973$.

Esaminando il risultato dell'analisi di regressione si può notare una certa curvatura della curva di calibrazione alle concentrazioni più alte.

Anche la distribuzione intorno allo zero dei residui potrebbe indicare una certa curvatura.

Il risultato potrebbe migliorare dopo eliminazione del punto relativo a 80 nL/L.



Sulla base delle precedenti osservazioni si decide di verificare la possibilità di eliminare il punto corrispondente a 80 nL/L. Allo scopo si ripete la procedura OLLSR dopo l'eliminazione del dato.

Prima dell'eliminazione (N = 9)

$$r^2 = 0,9946$$

$$b = 14,16 \pm 0,94$$

$$a = 61 \pm 44$$

$$s_{y/x} = 30,652$$

Dopo dell'eliminazione (N = 8)

$$r^2 = 0,9974$$

$$b = 14,71 \pm 0,75$$

$$a = 48 \pm 31$$

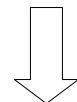
$$s_{y/x} = 19,802$$

L'eliminazione ha leggermente migliorato il risultato (confrontare i valori di r e di $s_{y/x}$ e gli intervalli di fiducia di pendenza ed intercetta). L'equazione della retta di calibrazione potrebbe essere

$$S(C)_{(u.a.)} = (14,71 \pm 0,74) \cdot C_{(nL/L)} + (48 \pm 31) \quad (P = 95\%; \nu = 6)$$

e non

$$S(C)_{(u.a.)} = (14,16 \pm 0,94) \cdot C_{(nL/L)} + (61 \pm 44) \quad (P = 95\%; \nu = 7)$$

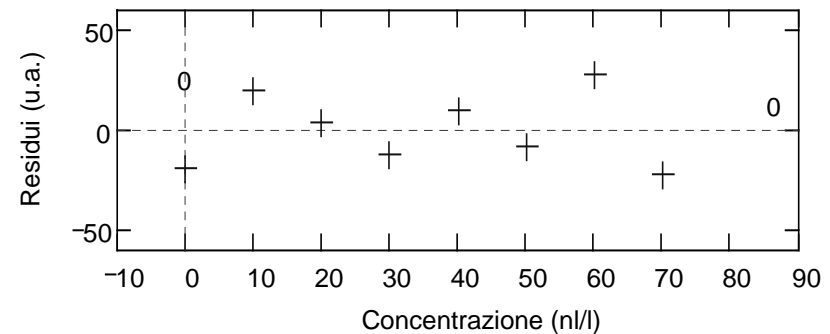
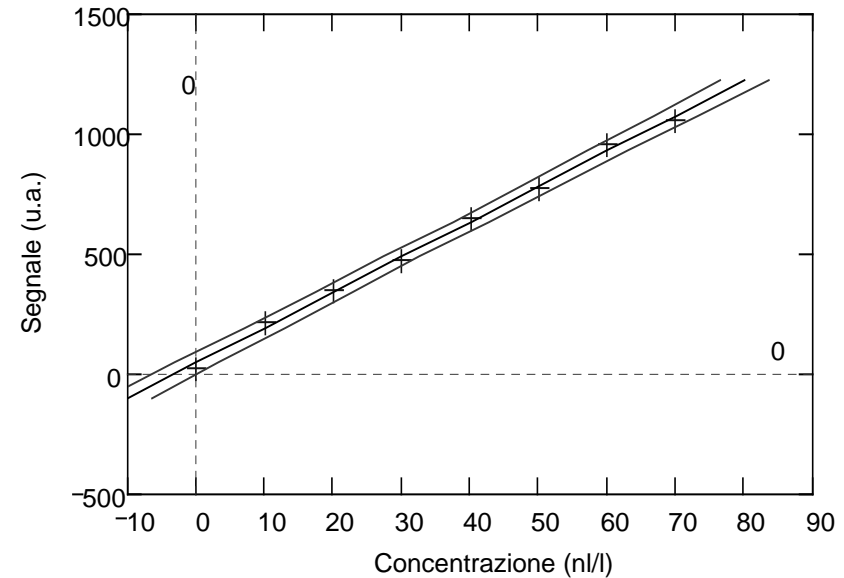


L'esame della curva di calibrazione non mostra più alcuna curvatura apprezzabile.

Anche i residui ora sembrano distribuiti casualmente intorno allo zero.

Il range lineare si estenderebbe quindi da 1,5 nL/L (dal Idr) a 70 nL/L.

Si veda però (nel testo a lato) il risultato del test-F eseguito sui due valori di $S_{y/x}$.



ESEMPIO N. 12 – Verificare la possibilità di analizzare i dati dell'Esempio N. 11 mediante regressione ordinaria non lineare ai minimi quadrati (ONLLSR).

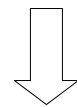
I dati sperimentali, gli stessi analizzati con il metodo OLLSR, sono i seguenti:

$C_{C_2H_2}$ (nL/L)	=	0,0	10,0	20,0	30,0	40,0
Segnale (u.a.)	=	29	215	346	477	646
$C_{C_2H_2}$ (nL/L)	=	50,0	60,0	70,0	80,0	
Segnale (u.a.)	=	775	958	1055	1142	

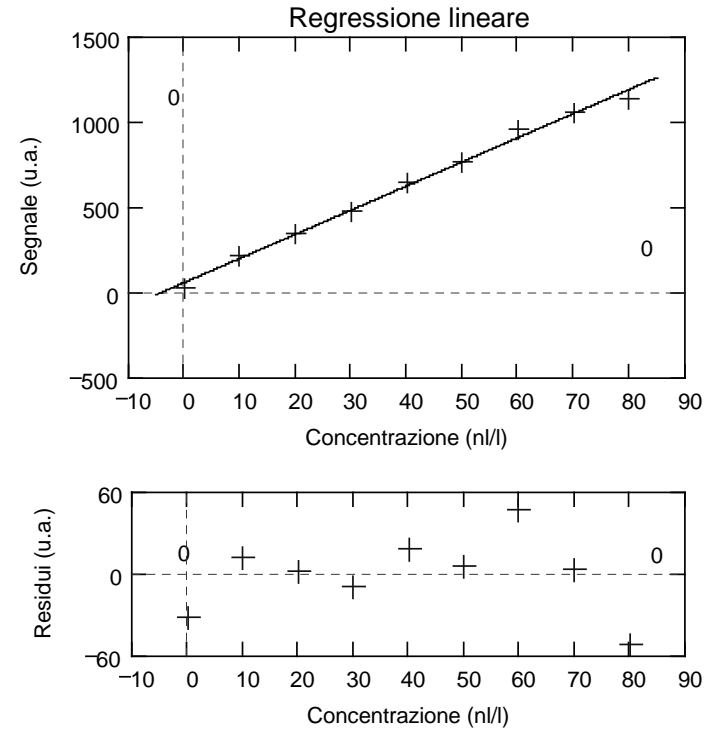
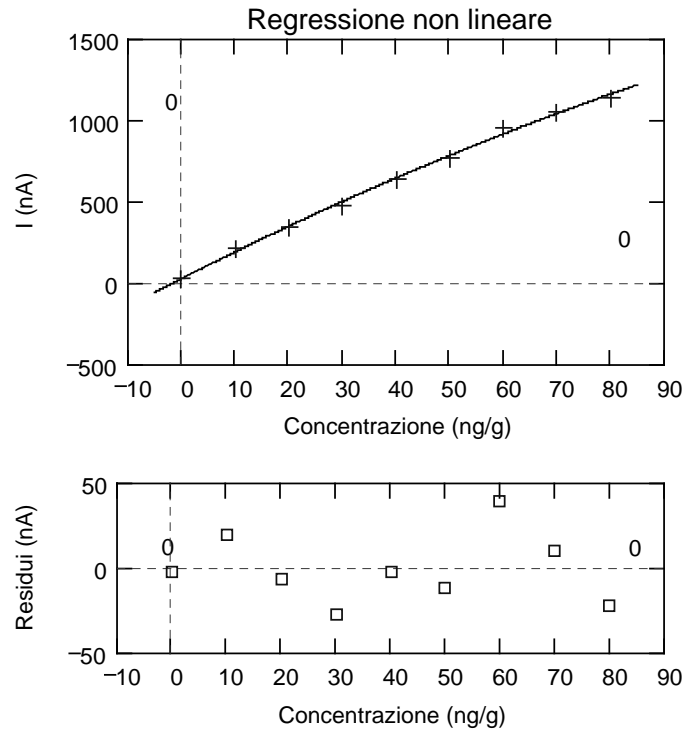
Naturalmente, in questo caso viene conservato anche il punto relativo a $C = 80,0$ nL/L.

L'equazione della curva di calibrazione risulta essere la seguente:

$$S(C) = -0,032 \cdot C^2 + 16,718 \cdot C + 30,854 \quad (P = 95\%; \nu = 8; s_{y/x} = 23.876)$$



I risultati ottenuti mediante OLLSR e ONLLSR possono essere confrontati qui sotto.



L'esame visivo dei risultati ottenuti non permette una chiara scelta del metodo migliore. Allo scopo si può usare il test di Mandel (vedere esempio successivo).

Il test di Mandel per la verifica della non-linearità si basa sul confronto delle deviazioni standard residue stimate analizzando gli stessi dati mediante regressione lineare e non lineare (s_{y/x_1} e, rispettivamente, s_{y/x_2}). Esso viene eseguito calcolando il valore di TV (test-value)

$$TV = \frac{(N-2) \cdot s_{y/x_1}^2 - (N-3) \cdot s_{y/x_2}^2}{s_{y/x_2}^2}$$

TV deve essere confrontato con il valore critico di $F_{1-\alpha,1,N-3}$ (test ad una coda) relativo al livello di fiducia considerato.

Se TV è minore o uguale a $F_{1-\alpha,1,N-3}$, la funzione di calibrazione di secondo ordine non porta a risultati qualitativamente migliori. La non linearità non è significativa.

ESEMPIO N. 13 – Utilizzando il test di Mandel, confrontare i risultati ottenuti analizzando i dati relativi alla determinazione PAS dell'etilene mediante OLLSR e della ONLLSR (P = 95%).

I due valori della deviazione standard residua sono:

$$\text{OLLSR} \quad s_{y/x_1} = 30,652$$

$$\text{ONLLSR} \quad s_{y/x_2} = 23,876$$

TV risulta uguale a:

$$\text{TV} = \frac{7 \cdot 30,652^2 - 6 \cdot 23,876^2}{23,876^2} = 5,537$$

Dato che il valore critico di F è $F_{1-\alpha,1,6} = 5,987$ si può concludere che la funzione di calibrazione di secondo ordine non porta a risultati significativamente migliori (P = 95%).

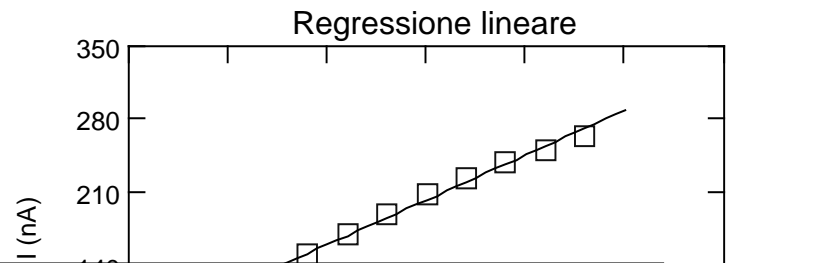
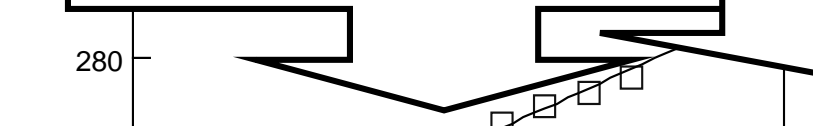
L'uso del coefficiente di correlazione per verificare la linearità non è consigliabile: secondo l'Analytical Committee della RSC, solo $r > |0,999|$ può essere indicativo di linearità.

Qui sotto sono riportati i risultati usati nel corso della costruzione di un diagramma di calibrazione nell'ambito dello sviluppo di un metodo per la determinazione del platino in bevande e alimenti liquidi mediante voltammetria catalitica di adsorbimento e ridissoluzione ad un elettrodo a goccia di mercurio pendente.

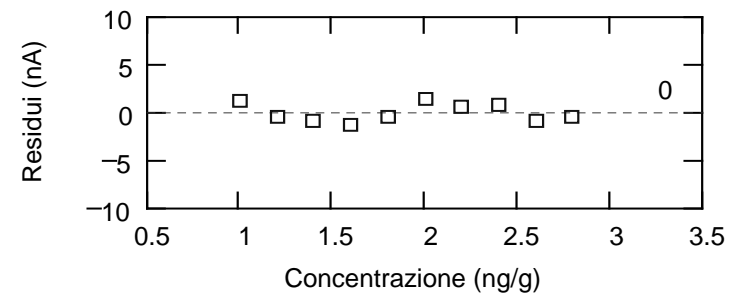
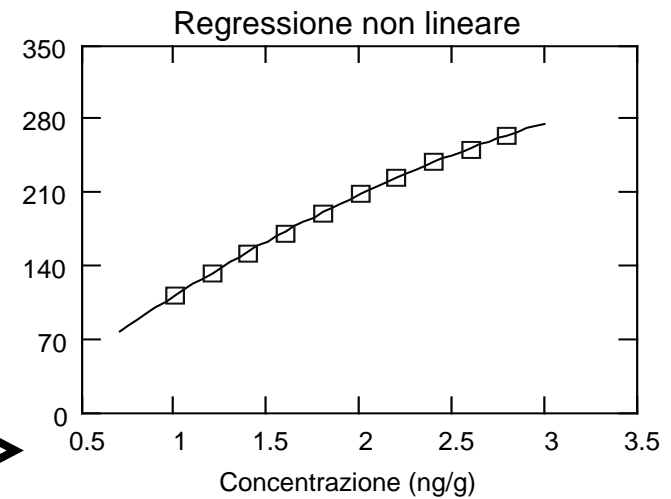
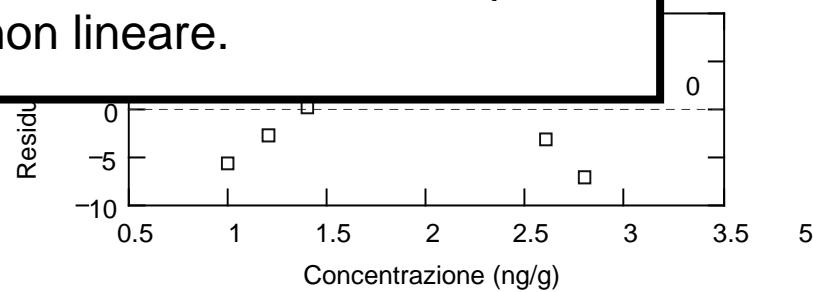
C_{Pt}(pg/g):	1,0; 1,2; 1,4; 1,6; 1,8; 2,0; 2,2; 2,4; 2,6; 2,8
I_{picco}(nA):	113; 133; 153; 172; 191; 210; 225; 240; 252; 265

Tuttavia la
suggerisce l'
lineare.

Si confronti il risultato con quello ottenuto
mediante ONLLSR.



Nonostante l'alto coefficiente
di correlazione valutato con
l'OLLSR, la più corretta
relazione funzionale è quella
non lineare.



La significatività della correlazione può essere stimata per mezzo di un apposito test-t:

$$t = |r| \frac{\sqrt{n-2}}{\sqrt{1-r^2}}$$

L'ipotesi nulla è che *non ci sia correlazione*. Quindi, se il valore calcolato è maggiore di quello critico per il livello di fiducia prescelto e N-2 gradi di libertà si evince una correlazione significativa.

ESEMPIO N. 14 – Verificare la significatività (P = 95%) della correlazione tra variabile dipendente e variabile indipendente nei due casi sotto indicati:

Caso 1: $r = 0,8453$ $n = 5$

Caso 2: $r = 0,8453$ $n = 11$

$$\text{Caso 1: } t_{\text{exp}} = |0,8453| \cdot \frac{\sqrt{5-2}}{\sqrt{1-(0,8453)^2}} = 2,740$$

$$t_{1-\frac{\alpha}{2},3} = 3,182$$

$$\text{Caso 2: } t_{\text{exp}} = |0,8453| \cdot \frac{\sqrt{11-2}}{\sqrt{1-(0,8453)^2}} = 4,746$$

$$t_{1-\frac{\alpha}{2},9} = 2,262$$

Nel caso 1 la correlazione non è significativa, nel caso 2 sì.

I diagrammi di calibrazione possono essere usati per interpolare la concentrazione di un campione incognito.

ESEMPIO N. 15 – L'analisi in triplo dell'etilene sviluppata da un campione reale ha portato al valore medio di segnale $S = 318$ u.a. Stimare la concentrazione del campione ($P = 99\%$) usando i dati degli Esempi N. 11 e N. 12.

Nel caso della OLLSR (Esempio N. 11) si ottiene

$$C_x = 18,366 \pm 2,386 = 18,4 \pm 2,4 \text{ nL/L (P = 95\%; } \nu = 6)$$

Nel caso della ONLLSR (Esempio N. 12) si ottiene

$$C_x = 17,782 \pm 2,746 = 17,8 \pm 2,8 \text{ nL/L (P = 95\%; } \nu = 6)$$

Quando si usa il metodo delle aggiunte standard, per estrapolare la concentrazione del campione incognito si usano equazioni leggermente diverse da quelle usate nel caso del metodo della curva di calibrazione.

ESEMPIO N. 16 – L'arsenico in un campione di zenzero (Zingiber officinale, provenienza Marocco) viene determinato con il metodo delle aggiunte standard mediante DPASV su elettrodo a solido di oro.

I risultati dell'analisi di 5 aliquote di mineralizzato (as received + quattro aliquote fortificate con concentrazioni crescenti di analita), sono i seguenti:

$C_{As(aggiunta)}$ (ng/g):	0,0;	4,76;	9,52;	14,28;	19,04
Area _{picco} (nW):	7,065;	19,61;	29,13;	43,05;	52,27

Valutare la concentrazione dell'arsenico nel mineralizzato (P = 95%).

Risultati della regressione:

$$r = 0,9982 \quad S_{y/x} = 1,2306$$

$$S(C) = (2,39 \pm 0,26) \cdot C + (7,4 \pm 3,0)$$

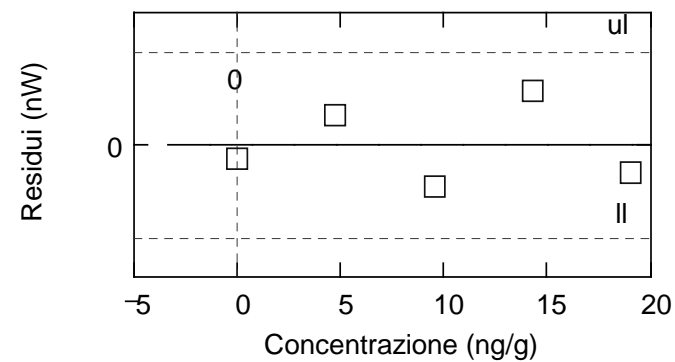
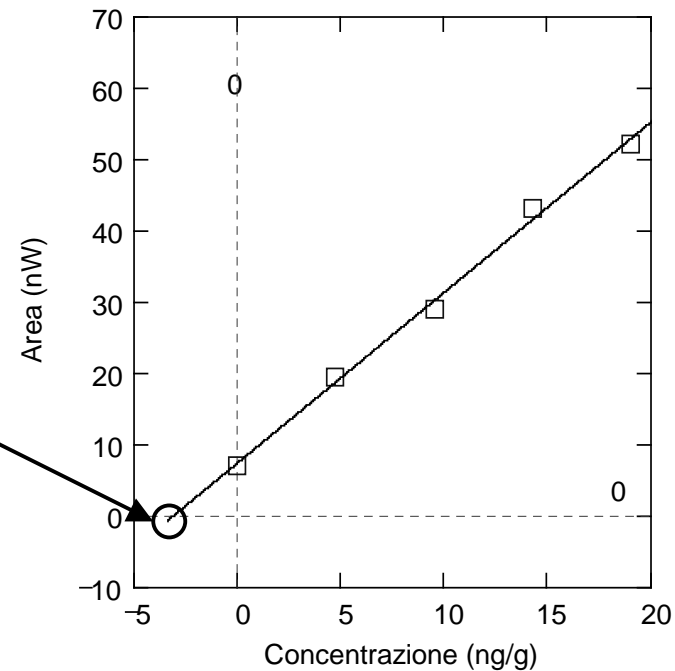
($\alpha = 0,05$; $\nu = 3$)

La freccia indica il valore di $-C_x$.

Il risultato può essere scritto come segue:

$$C_{as} = 3,1 \pm 1,6 \text{ ng/g (P = 95\%; } \nu = 3)$$

Si noti l'incertezza del risultato nonostante il buon valore del coefficiente di correlazione. L'incertezza è determinata dal basso valore della concentrazione dell'analita.



In teoria, i dati della calibrazione non dovrebbero contenere outliers. Diversi test permettono di riconoscere gli outliers.

Uno dei test possibili è il test-F illustrato nell'Esempio seguente.

ESEMPIO N. 17 – Nel corso della messa a punto di un metodo per la determinazione dell'indio, in soluzione acida per ac. solforico, mediante FAAS a 303,9 nm, l'analisi di una serie di soluzioni standard ha permesso di ottenere i seguenti dati

C_{In} ($\mu\text{g/L}$):	6,0;	12,0;	16,0;	24,0;	30,0;	38,0
A:	0,087;	0,113;	0,170;	0,223;	0,226;	0,341

Il dato $S = 0,226$ (corrispondente a $C_{In} = 30,0 \mu\text{g/L}$) è un sospetto outlier. Si desidera verificare l'ipotesi ($P = 95\%$).

Si esegue prima l'analisi dei dati mediante OLSR. L'analisi viene quindi ripetuta dopo eliminazione del dato sospetto. Qui sotto sono riportati i risultati ottenuti prima e dopo l'eliminazione:

Prima

$$v_p = 4$$

$$r_p = 0,9750$$

$$s_{y/x_p} = 0,0228$$

$$b_p = 0,0075 \pm 0,0039$$

$$a_p = 0,036 \pm 0,093$$

Dopo

$$v_d = 3$$

$$r_d = 0,9953$$

$$s_{y/x_d} = 0,0113$$

$$b_d = 0,0081 \pm 0,0015$$

$$a_d = 0,031 \pm 0,032$$

Il test-F implica il confronto dei valori di $s_{y/x}$. Si calcola il valore di F_{exp} :

$$F_{\text{exp}} = \frac{v_p \cdot s_{y/x_p}^2 - v_d \cdot s_{y/x_d}^2}{s_{y/x_d}^2} = \frac{4 \cdot 0,0228^2 - 3 \cdot 0,0113^2}{0,0113^2} = 13,284$$

Dato che F_{exp} è maggiore del valore critico al livello di fiducia prescelto, $F_{1-\alpha,1,3} = 10,128$, il dato può essere considerato un outlier e rigettato. Il risultato della calibrazione è quello ottenuto dopo la sua eliminazione.

Qui sotto si possono confrontare i due diagrammi di calibrazione.

