

# **ESTATÍSTICA BÁSICA- SUMÁRIO ( 2ª PARTE)**

Introdução

Processo de calibração

Construção da curva analítica

Regressão inversa

Validação do modelo

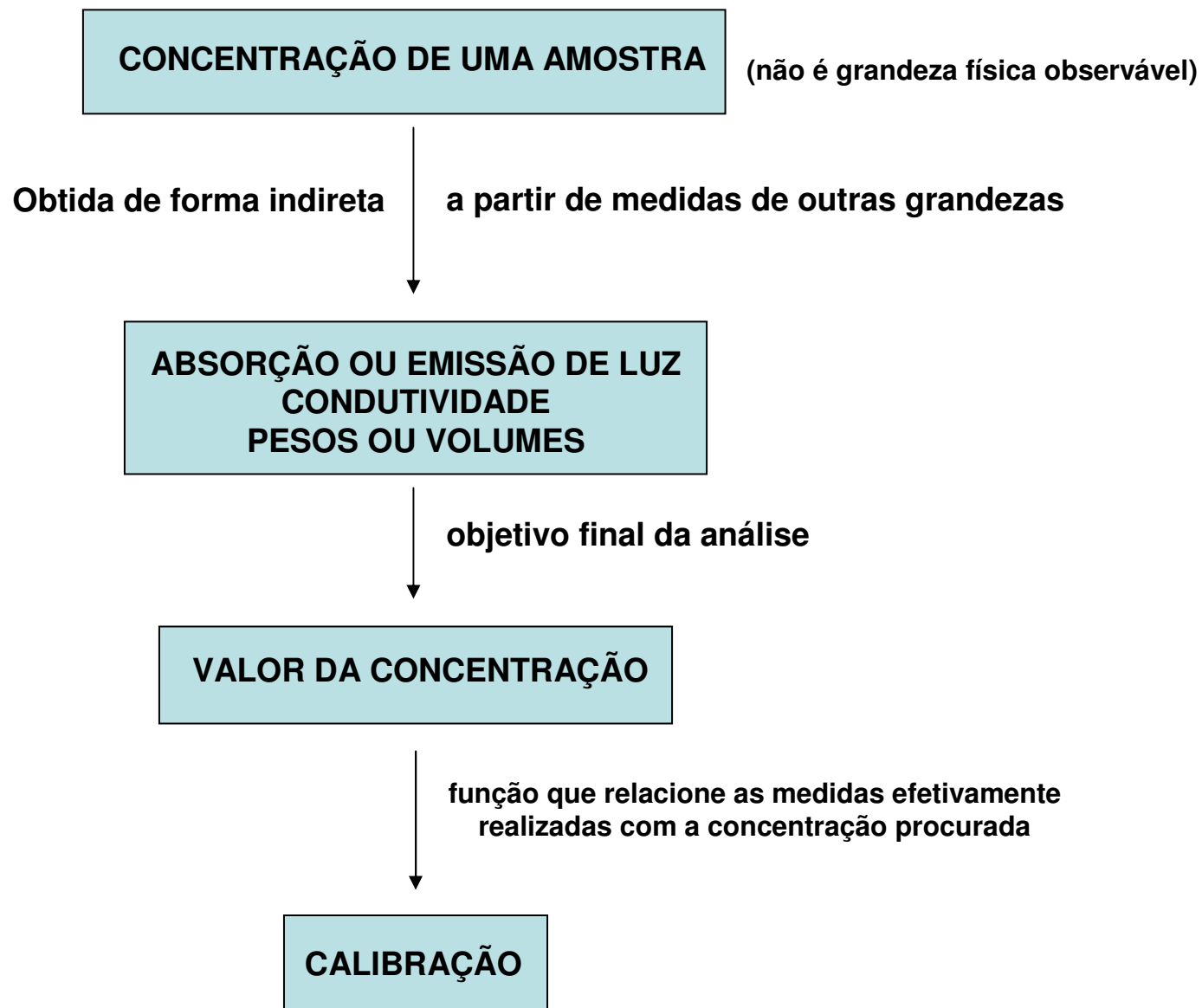
Análise de resíduos

Coeficiente de determinação

Curva de calibração por adição de padrão

Curva de calibração com uso de padrão interno

# INTRODUÇÃO



# PROCESSO DE CALIBRAÇÃO

**DESCRITIVA**

medidas feitas numa série de padrões analíticos de concentrações conhecidas são usadas para construir um modelo que relacione a grandeza medida com a concentração da espécie de interesse

**PREDITIVA**

usa-se um modelo para prever concentrações de novas amostras, a partir dos sinais analíticos medidos para elas

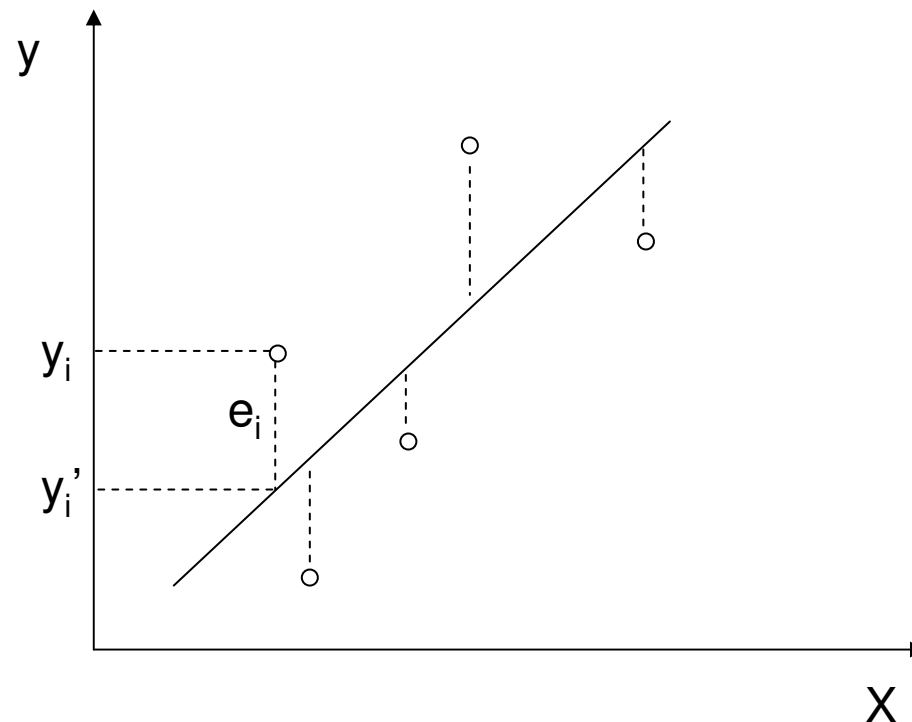
uma parte fundamental do processo analítico

**AFETA A APLICABILIDADE, A EXATIDÃO, A PRECISÃO,  
A DURAÇÃO E CUSTO GLOBAL DE UMA ANÁLISE**

## AJUSTE POR MÍNIMOS QUADRADOS

- A etapa descritiva no processo de implementação do modelo leva em consideração o ajuste da reta passando o “mais perto” possível dos pontos experimentais, já que é impossível passar exatamente sobre todos eles.

-Em termos práticos, “passar mais perto” significa minimizar a distância global dos pontos em relação a reta, isto é, minimizar o comprimento total dos segmentos verticais na figura:



- A maneira tradicional de conseguirmos esse resultado é localizar a reta de tal maneira que a soma dos quadrados dos resíduos seja mínima, razão pela qual esse método é chamado de ajuste por mínimos quadrados ou análise de regressão.

Uma relação linear entre uma variável aleatória  $y$  e uma variável não aleatória  $X$  é descrita pelo modelo

$$y = \beta_0 + \beta_1 X + \varepsilon \quad (1)$$

onde  $\beta_0$  e  $\beta_1$  são os parâmetros do modelo e  $\varepsilon$  é o erro aleatório associado à determinação de  $y$ . Na química analítica,  $y$  normalmente é o sinal ou a resposta dada por um instrumento e  $X$  a concentração da espécie química que se quer determinar.

Para cada sinal  $y_i$  observado num padrão analítico de concentração  $X_i$  o modelo linear fornece uma estimativa  $y_i'$  dada pela equação

$$y_i' = b_0 + b_1 X_i \quad (2)$$

onde  $b_0$  e  $b_1$  são as estimativas dos parâmetros  $\beta_0$  e  $\beta_1$ , obtidas na etapa descritiva da calibração. As diferenças entre os valores observados e os valores estimados são os resíduos  $e_i = y_i - y_i'$ .

Como os valores de  $y_i$  já são conhecidos de antemão, os resíduos irão depender apenas dos valores que escolhermos para  $b_0$  e  $b_1$ . No ajuste por mínimos quadrados, esses valores são aqueles que tornam o somatório  $\sum e_i^2$  o menor possível.

No método dos mínimos quadrados os valores de  $b_0$  e  $b_1$  são estimados minimizando-se a soma quadrática dos resíduos, dada por

$$SQ_r = \sum (e_i)^2 = \sum [y_i - (y_e)_i]^2 = \sum (y_i - b_0 - b_1 x_i)^2 \quad (4)$$

O somatório deve ser feito sobre todas os valores experimentais, ou seja, de  $i = 1$  até  $i = n$ , onde  $n$  é o número total de medidas. Os índices foram omitidos do símbolo do somatório para simplificar a notação.

Para minimizar a soma quadrática residual deriva-se a Eq. (4) em relação a  $b_0$  e  $b_1$  e iguala-se as derivadas a zero. Isto leva, depois de alguns rearranjos, às seguintes expressões para o cálculo de  $b_0$  e  $b_1$ :

$$b_0 = (\sum y_i - b_1 \sum x_i) / n \quad (5)$$

$$b_1 = (n \sum y_i x_i - \sum x_i \sum y_i) / [n \sum (x_i)^2 - (\sum x_i)^2] \quad (6)$$

### Nota:

- O tratamento estatístico dos resultados obtidos com o método dos mínimos quadrados exige que sejam válidas (pelo menos aproximadamente) algumas suposições sobre a natureza dos erros associados às medidas.
- Na presença de violações grosseiras dessas suposições, amostras diferentes extraídas do mesmo sistema podem levar a equações bastante diferentes, isto é a modelos sem estabilidade.

## SUPOSIÇÕES

→ **Os únicos erros nas medidas são devidos a flutuações na variável dependente  $y$  (o sinal analítico).**

Se por acaso a variável independente,  $x$  (a concentração), também estiver sujeita a erros, estes devem ser desprezíveis quando comparados com os erros  $y$ .

→ **Os erros  $\varepsilon_i$  são variáveis aleatórias com média zero e variância constante  $\sigma^2$ .** (homocedasticidade)

Esta suposição é comumente violada em algumas técnicas analíticas onde a faixa de trabalho é extensa, como o ICP-AES. Nesses casos a variância do sinal não é constante ao longo de toda a faixa de calibração, e sim proporcional a intensidade do sinal analítico. Isto significa que o método dos mínimos quadrados ordinários não deve ser empregado, pois embora estimativas dos parâmetros continuem não sendo tendenciosas, não mais terão variância mínima.

Pode-se então adotar três procedimentos alternativos:

- Restringir a faixa de calibração a um trecho aonde a variância seja constante.
- Realizar uma transformação estabilizadora da variância, então aplicar o método dos mínimos quadrados ordinários.
- Usar o método dos mínimos quadrados ponderados.

A homocedasticidade pode ser verificada pelos teste F e o teste de Bartlett. (repetições autênticas)

→ **Os erros  $\varepsilon_i$  não são correlacionados.**

Preparar os padrões analíticos separadamente, e de maneira semelhante em cada nível de concentração, para que não ocorra uma variação sistemática em determinado nível provocada pela técnica de preparação. (por exemplo, efetuar o mesmo número de diluições na solução estoque para cada nível de concentração)

→ Para a realização de testes estatísticos e a construção de intervalos de confiança, é necessário acrescentar mais uma suposição referente à forma da distribuição populacional dos erros: **Os erros  $\varepsilon$  seguem uma distribuição normal.**

## ERROS PADRÕES E INTERVALOS DE CONFIANÇA PARA OS PARÂMETROS DO MODELO

Tendo-se usado o método dos mínimos quadrados para estimar os parâmetros do modelo, pode-se estimar os erros padrões associados a esses parâmetros através das expressões:

$$S(b_0) = s [\Sigma(x_i)^2 / n \Sigma(x_i - x_m)^2]^{1/2} \quad (7)$$

e

$$S(b_1) = s / [\Sigma(x_i - x_m)^2]^{1/2} \quad (8)$$

Onde  $m$  é a média de todas as concentrações  $x_i$  usadas no ajuste do modelo e  $s$  é uma estimativa do desvio (erro) padrão da resposta  $y_i$  determinadas repetidamente na mesma concentração  $x_i$ . Quando o modelo é correto, uma estimativa não tendenciosa do erro padrão  $s$  é obtida a partir da soma quadrática residual dada pela equação (4)

$$S = (SQ_r / \nu)^{1/2} = (MQ_r)^{1/2} \quad (9)$$

Onde  $\nu$  é o número de graus de liberdade associado a estimativa  $s$ , dado pela diferença entre o número total de medidas realizadas,  $n$ , e o número de parâmetros que estão sendo estimados. Neste exemplo, como apenas dois parâmetros são estimados,  $\nu = (n-2)$ . A notação  $MQ_r$  vem de média quadrática residual, o nome que se dá ao valor calculado pela equação (9) numa análise de variância.

Admitindo que os erros se distribuam normalmente (quarta suposição), podemos agora usar a distribuição de Student para construir intervalos de confiança em torno das estimativas  $\beta_0$  e  $\beta_1$

$$\beta_0 = b_0 \pm t s(b_0) \quad (10)$$

e

$$\beta_1 = b_1 \pm t s(b_1) \quad (11)$$

Onde  $t$  é o ponto da distribuição de Student correspondente ao nível de confiança desejado com  $\nu$  graus de liberdade

Desvio padrão sobre a regressão:  $S_y = [[\Sigma(y_i - y_m)^2 - b_1^2 \Sigma(x_i - x_m)^2] / n - 2]^{1/2}$



## ERROS PADRÕES E INTERVALOS DE CONFIANÇA PARA A RESPOSTA ANALÍTICA

Considere-se uma determinação realizada numa nova amostra de concentração  $x_0$ , num sistema analítico para o qual já tenha sido feita a calibração descritiva. Pode-se usar a curva de calibração para estimar a resposta nesse ponto,  $y_0$ , através da equação 2. A presença de erros aleatórios fará com que essa estimativa seja afetada por uma certa incerteza, que pode ser quantificada através da expressão:

$$S_0 = s \{1 + (1/n) + [(x_0 - x_m)^2 / \Sigma(x_i - x_m)^2]\}^{1/2} \quad (12)$$

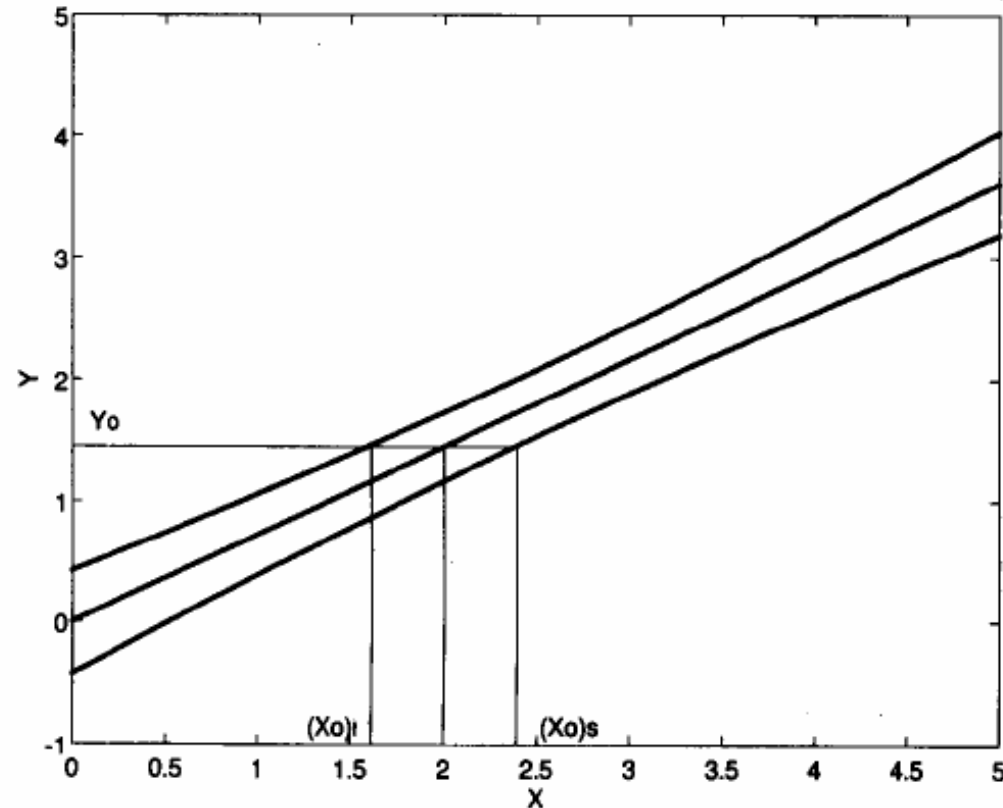
Onde  $s_0$  representa o erro padrão da estimativa  $y_0$

Imagine agora que, ao invés de uma única determinação, seja feita uma série de  $q$  determinações repetidas (autenticamente) na amostra de concentração  $x_0$ . A resposta média obtida  $(y_0)_q$ , será um resultado mais preciso, porque os erros aleatórios tenderão a se cancelar no cálculo da média. O erro padrão da estimativa dessa resposta média a partir da curva de calibração terá portanto um valor menor, dado pela expressão

$$(S_0)_q = s \{(1/q) + (1/n) + [(x_0 - x_m)^2 / \Sigma(x_i - x_m)^2]\}^{1/2} \quad (13)$$

Admitindo-se a validade das quatro suposições básicas, pode-se mais uma vez usar a distribuição de Student para obter intervalos de confiança em torno das resposta estimadas. Esses intervalos, cuja largura é determinadas pelo produto do erro padrão pelo valor t apropriado, como nas Eqs. (10) e (11), podem ser usados para avaliar a confiabilidade dos novos resultados.

A variação dos intervalos de confiança ao longo da reta ajustada define uma faixa limitada por duas curvas hiperbólicas. De acordo com a equação 13 a largura dessa faixa depende, através do termo  $(x_0 - x_m)^2$ , da localização de  $x_0$  em relação a média das concentrações empregadas no ajuste do modelo. Essa dependência implica que a largura do intervalo de confiança é mínima em  $x_0 = x_m$  e aumenta à medida que  $x_0$  distancia-se de  $x_m$ , em ambas as direções. Este é um dos motivos pelos quais deve ser cuidado ao fazer extrapolações além da faixa de calibração.



*Figura 1. Regressão inversa. As linhas curvas são as hipérboles que definem os intervalos de confiança em torno da reta de regressão. O intervalo de confiança para a concentração estimada a partir de um sinal observado  $y_0$  vai de  $(x_0)_i$  a  $(x_0)_s$ .*

## REGRESSÃO INVERSA

O que mais interessa ao químico analítico é usar a curva de calibração para prever o valor da concentração a partir de uma resposta analítica medida, além, é claro, de ter uma estimativa do erro associado a essa previsão. Esse procedimento é chamado de *regressão inversa*.

$$(x_e)_0 = (y_0 - b_0) / b_1 \quad (14)$$

Chegar a uma expressão que nos dê os intervalos de confiança exige um pouco mais de manipulações algébricas e resulta, para uma média de  $q$  determinações feitas na mesma amostra, no intervalo aproximado

$$x_0 = (x_e)_0 \pm \{ts / b_1\} \left\{ (1/q) + (1/n) + [((x_e)_0 - x_m)^2 / \Sigma(x_i - x_m)^2] \right\}^{1/2} \quad (15)$$

Esta expressão é válida quando o coeficiente angular  $b_1$  é estatisticamente significativo e bem determinado, que é a situação em que normalmente se usa uma regressão inversa.

## VALIDAÇÃO DO MODELO

A validade do modelo e a significância estatística da curva ajustada podem ser testadas por meio de **Análise Da Variância**.

$H_0$  : A equação da reta é adequada para descrever os dados

$H_a$  : A equação da reta não é adequada para descrever os dados

Tabela 5.8 – Tabela de análise da variância para o ajuste, pelo método dos mínimos quadrados, de um modelo linear nos parâmetros.  $n_i$  = número de repetições no nível  $i$ ;  $m$  = número de níveis distintos da variável independente;  $n = \sum n_i$  = número total de observações;  $p$  = número de parâmetros do modelo.

Fonte de variação	Soma Quadrática	Nº de g. l.	Média Quadrática
Regressão	$SQ_R = \sum_i \sum_j (\hat{y}_i - \bar{y})^2$	$p - 1$	$MQ_R = \frac{SQ_R}{p - 1}$
Resíduos	$SQ_r = \sum_i \sum_j (y_{ij} - \hat{y}_i)^2$	$n - p$	$MQ_r = \frac{SQ_r}{n - p}$
Falta de ajuste	$SQ_{faj} = \sum_i \sum_j (\hat{y}_i - \bar{y}_i)^2$	$m - p$	$MQ_{faj} = \frac{SQ_{faj}}{m - p}$
Erro puro	$SQ_{ep} = \sum_i \sum_j (y_{ij} - \bar{y}_i)^2$	$n - m$	$MQ_{ep} = \frac{SQ_{ep}}{n - m}$
Total	$SQ_T = \sum_i \sum_j (y_{ij} - \bar{y})^2$	$n - 1$	

### TESTE DA FALTA DE AJUSTE

$$MQ_{faj} / MQ_{ep} = F_{calc}$$

$$\text{Se } *F_{calc} < F_{vfaj, vep}$$

\*Indica ausência de falta de ajuste

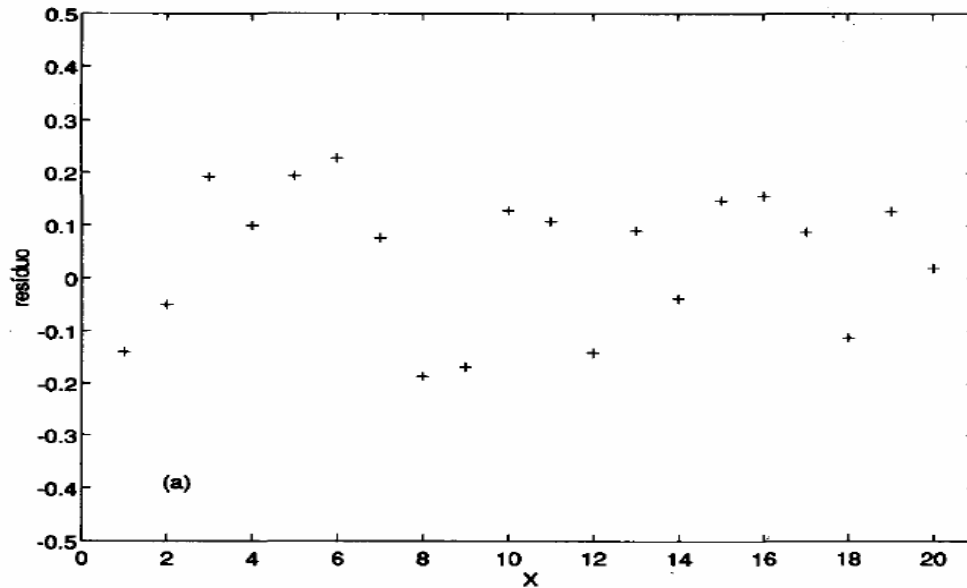
### TESTE DA SIGNIFICÂNCIA DA REGRESSÃO

$$MQ_{reg} / MQ_r = F_{calc}$$

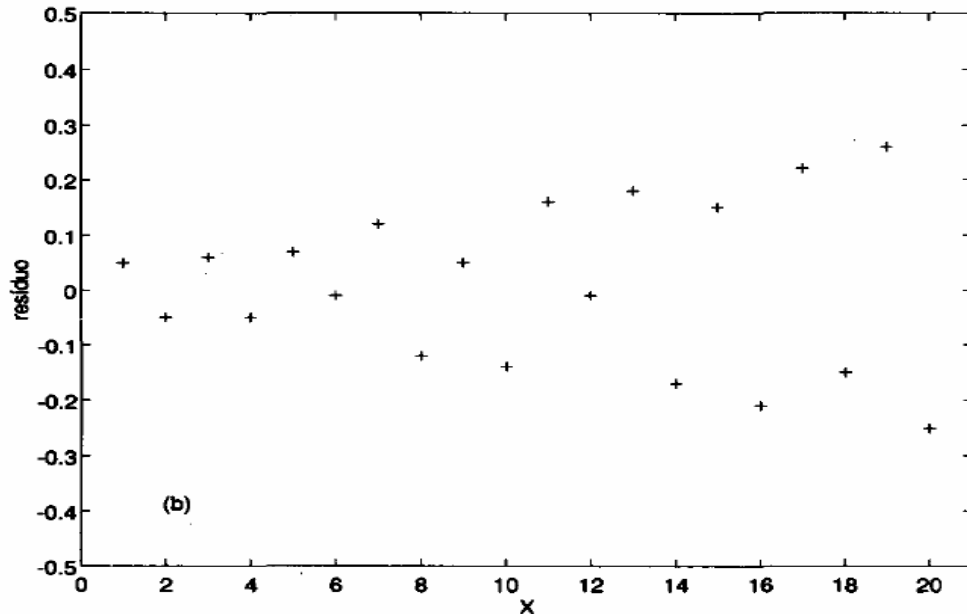
$$\text{Se } **F_{vreg, vr} > F_{calc}$$

\*\*indica a existência de uma relação linear entre as duas variáveis; quanto maior for a razão entre  $MQ_{reg} / MQ_r$ , mais significativa será a relação

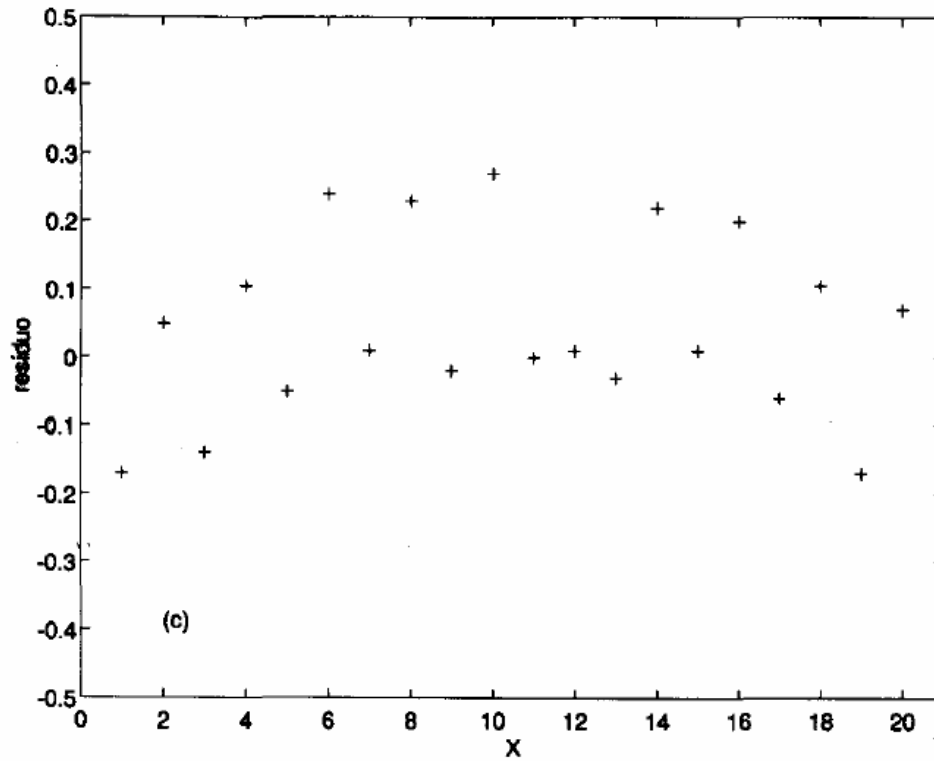
# ANÁLISE DOS RESÍDUOS



Mostra a situação ideal, em que o modelo está bem ajustado: os resíduos estão distribuídos aleatoriamente, e não há nada que indique que sua variância não é constante.



Embora a distribuição continue aleatória, pode-se perceber claramente que os resíduos crescem com o valor da concentração indicando a presença de heteroscedasticidade.



É evidente o aparecimento de uma estrutura não aleatória, de aspecto curvilíneo. Os resíduos são menores nas extremidades da faixa de calibração, e maiores na região intermediária. Isto significa que parte da estrutura sistemática dos dados não foi explicada pelo modelo adotado.

## COEFICIENTE DE DETERMINAÇÃO

O coeficiente de determinação, ou coeficiente de correlação múltipla,  $R^2$ , é a razão entre a soma quadrática explicada pela regressão e a soma quadrática total:

$$R^2 = \frac{SQ_{reg}}{SQ_t} = \frac{\sum n_i [(y_e)_i - y_m]^2}{\sum \sum (y_{ij} - y_m)^2} \quad (18)$$

- Comumente chamado de variação percentual explicada pelo modelo.
- Quanto mais próximo de 1 estiver o valor de  $R^2$ , mais o modelo consegue escrever a variação em  $y$
- A raiz quadrada do coeficiente de determinação representa o coeficiente de correlação entre os valores observados e os valores estimados pelo modelo:  $R = r(y_e, y)$
- Erro comum é usar o valor numérico de  $R^2$ , ou de sua raiz quadrada  $R$ , para avaliar e concluir, caso esse valor seja suficiente alto, que o ajuste é satisfatório. Por exemplo, um valor de 0.99 para  $R$  (isto é  $R^2 = 0.98$ ) é obviamente muito alto, mais significa apenas que 98% da variação total em torno da média foi explicada pelo modelo. É possível que os 2% restantes sejam concentrados numa única porção da curva, e isso significa falta de ajuste.
- É comum ainda, dado um valor  $R$  próximo de 1, concluir-se erroneamente que o intervalo de confiança para estimativas da concentração será estreito. (*vide equação 15*)

## UM EXEMPLO NUMÉRICO

**Tabela I.** Resultados de uma determinação de ferro por absorção molecular na região do UV-VIS, usando ortofenantrolina como agente complexante<sup>17</sup>.

Concentração dos Padrões em mg/L ( $x$ )	Absorvância ( $y$ )
0,2	0,1351
0,2	0,1519
0,2	0,1344
0,2	0,1457
1,0	0,7169
1,5	1,0846
2,0	1,4416
2,5	1,6849

Para ajustar uma reta a estes dados usa-se as equações (5) e (6), obtendo-se as seguintes estimativas

Para os parâmetros  $\beta_1$  e  $\beta_2$ :

$$b_0 = 0.011$$

$$b_1 = 0.693$$

Antes de pensar em utilizar estes valores para fazer previsões, é preciso verificar se o modelo linear está bem ajustado aos dados coletados.



**Tabela II.** Análise da variância para o ajuste do modelo  $y = \beta_0 + \beta_1 x + \varepsilon$  aos dados da Tabela I.

Fonte	Soma Quadrática	G. L.	Média Quadrática
Modelo	$SQ_{reg} = 2,90629$	1	$MQ_{reg} = 2,90629$
Residual	$SQ_r = 0,00722$	6	$MQ_r = 0,00120$
Falta de Ajuste	$SQ_{faj} = 0,00700$	3	$MQ_{faj} = 0,00233$
Erro Puro	$SQ_{ep} = 0,00022$	3	$MQ_{ep} = 0,00007$
Total	$SQ_t = 2,91351$	7	

$$MQ_{faj} / MQ_{ep} = 0.00233 / 0.00007 = 33.28$$

$$F_{vfaj, vep} = F_{3, 3} = 9.28 \text{ para } 95\% \text{ de confiança}$$

$$R^2 = SQ_{reg} / SQ_t = 2.90629 / 2.91351 = 0.9975$$

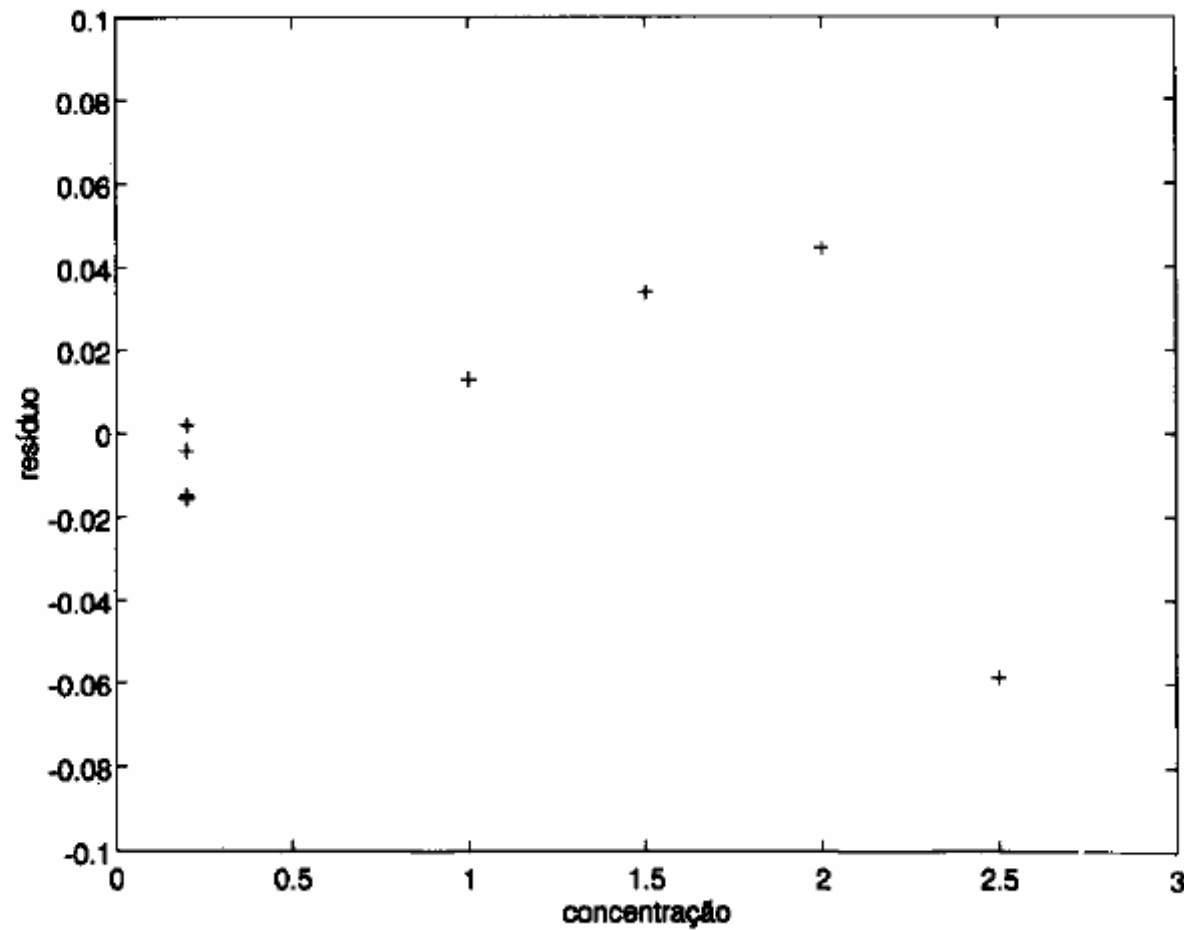
$$R^2_{max} = (SQ_t - SQ_{ep}) / SQ_t = (2.91351 - 0.00022) / 2.91351 = 0.9999$$

O modelo consegue explicar  $(0.9975 / 0.9999) \times 100 = 99.76\%$

**TABELA F : VALORES CRÍTICOS PARA UM TESTE UNILATERAL ( $\alpha = 0,05$ )**

$v_1 \backslash v_2$	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	$\infty$
1	161,4	199,5	215,7	224,6	230,2	234,0	236,8	238,9	240,5	241,9	243,9	245,9	248,0	249,1	250,1	251,1	252,2	253,3	254,3
2	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,38	19,40	19,41	19,43	19,45	19,45	19,46	19,47	19,48	19,49	19,50
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,74	8,70	8,66	8,64	8,62	8,59	8,57	8,55	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,91	5,86	5,80	5,77	5,75	5,72	5,69	5,66	5,63
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,68	4,62	4,56	4,53	4,50	4,46	4,43	4,40	4,36
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,00	3,94	3,87	3,84	3,81	3,77	3,74	3,70	3,67
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,57	3,51	3,44	3,41	3,38	3,34	3,30	3,27	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,28	3,22	3,15	3,12	3,08	3,04	3,01	2,97	2,93
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,07	3,01	2,94	2,90	2,86	2,83	2,79	2,75	2,71
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,91	2,85	2,77	2,74	2,70	2,66	2,62	2,58	2,54
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	2,79	2,72	2,65	2,61	2,57	2,53	2,49	2,45	2,40
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	2,69	2,62	2,54	2,51	2,47	2,43	2,38	2,34	2,30
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67	2,60	2,53	2,46	2,42	2,38	2,34	2,30	2,25	2,21
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60	2,53	2,46	2,39	2,35	2,31	2,27	2,22	2,18	2,13
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,48	2,40	2,33	2,29	2,25	2,20	2,16	2,11	2,07
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,42	2,35	2,28	2,24	2,19	2,15	2,11	2,06	2,01
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45	2,38	2,31	2,23	2,19	2,15	2,10	2,06	2,01	1,96
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,34	2,27	2,19	2,15	2,11	2,06	2,02	1,97	1,92
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38	2,31	2,23	2,16	2,11	2,07	2,03	1,98	1,93	1,88
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,28	2,20	2,12	2,08	2,04	1,99	1,95	1,90	1,84
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32	2,25	2,18	2,10	2,05	2,01	1,96	1,92	1,87	1,81
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30	2,23	2,15	2,07	2,03	1,98	1,94	1,89	1,84	1,78
23	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,37	2,32	2,27	2,20	2,13	2,05	2,01	1,96	1,91	1,86	1,81	1,76
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25	2,18	2,11	2,03	1,98	1,94	1,89	1,84	1,79	1,73
25	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,28	2,24	2,16	2,09	2,01	1,96	1,92	1,87	1,82	1,77	1,71
26	4,23	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22	2,15	2,07	1,99	1,95	1,90	1,85	1,80	1,75	1,69
27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,31	2,25	2,20	2,13	2,06	1,97	1,93	1,88	1,84	1,79	1,73	1,67
28	4,20	3,34	2,95	2,71	2,56	2,45	2,36	2,29	2,24	2,19	2,12	2,04	1,96	1,91	1,87	1,82	1,77	1,71	1,65
29	4,18	3,33	2,93	2,70	2,55	2,43	2,35	2,28	2,22	2,18	2,10	2,03	1,94	1,90	1,85	1,81	1,75	1,70	1,64
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16	2,09	2,01	1,93	1,89	1,84	1,79	1,74	1,68	1,62
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08	2,00	1,92	1,84	1,79	1,74	1,69	1,64	1,58	1,51
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04	1,99	1,92	1,84	1,75	1,70	1,65	1,59	1,53	1,47	1,39
120	3,92	3,07	2,68	2,45	2,29	2,17	2,09	2,02	1,96	1,91	1,83	1,75	1,66	1,61	1,55	1,50	1,43	1,35	1,25
$\infty$	3,84	3,00	2,60	2,37	2,21	2,10	2,01	1,94	1,88	1,83	1,75	1,67	1,57	1,52	1,46	1,39	1,32	1,22	1,00

**GRÁFICO DOS RESÍDUOS EM FUNÇÃO DA CONCENTRAÇÃO PARA O AJUSTE DO MODELO  $Y = \beta_0 + \beta_1 X + \varepsilon$  AOS DADOS DA TABELA I**



**Tabela III.** Análise da variância para o ajuste do modelo  $y = \beta_0 + \beta_1 x + \varepsilon$  aos dados da Tabela I, excluindo-se o nível de 2,5 mg/L.

Fonte	Soma Quadrática	G. L.	Média Quadrática
Modelo	$SQ_{reg} = 1,77495$	1	$MQ_{reg} = 1,77495$
Residual	$SQ_r = 0,00024$	5	$MQ_r = 0,00005$
Falta de Ajuste	$SQ_{faj} = 0,00002$	2	$MQ_{faj} = 0,00001$
Erro Puro	$SQ_{ep} = 0,00022$	3	$MQ_{ep} = 0,00007$
Total	$SQ_t = 1,77519$	6	

$$b_0 = -0.003 \quad ; \quad b_1 = 0.723$$

$$MQ_{faj} / MQ_{ep} = 0.14$$

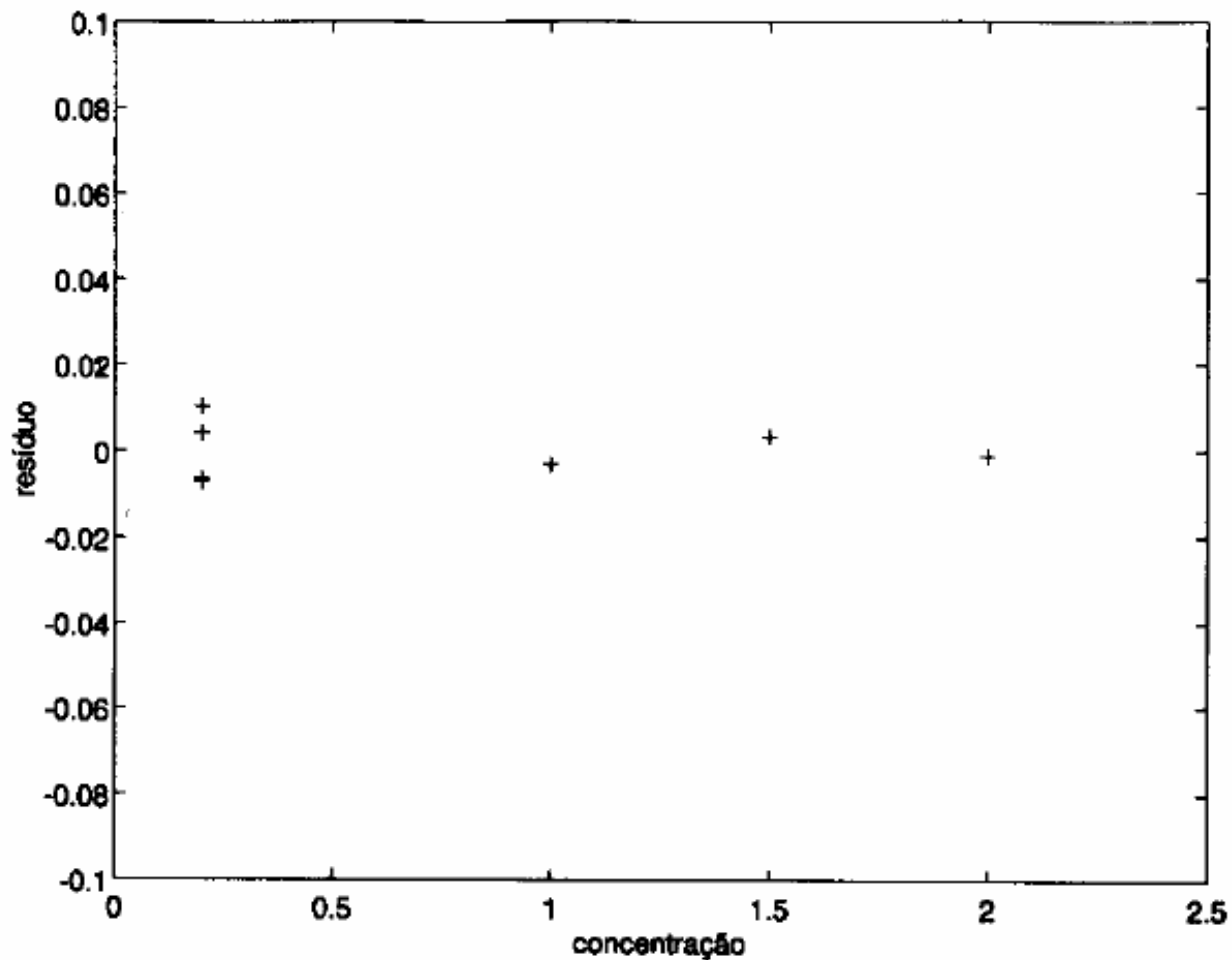
$$F_{vfaj, vep} = F_{2, 3} = 9.55 \text{ para } 95\% \text{ de confiança}$$

$$R^2 = SQ_{reg} / SQ_t = 0.9999$$

$$MQ_{reg} / MQ_r = F_{calc} = 37289$$

$$F_{1,5} = 6.61 \text{ para } 95\% \text{ de confiança}$$

**GRÁFICO DOS RESÍDUOS EM FUNÇÃO DA CONCENTRAÇÃO PARA O AJUSTE DO MODELO  $Y = \beta_0 + \beta_1 X + \varepsilon$  AOS DADOS DA TABELA I, EXCLUINDO O NÍVEL DE CONCENTRAÇÃO DE 2.5 mg/L**



RESUMO DOS RESULTADOS

<i>Estatística de regressão</i>	
R múltiplo	0.999932962
R-Quadrado	0.999865929
R-quadrado ajustado	0.999839115
Erro padrão	0.006899289
Observações	7

ANOVA					
	<i>gl</i>	<i>SQ</i>	<i>MQ</i>	<i>F</i>	<i>F de significação</i>
Regressão	1	1.774946508	1.774947	37288.64	7.06704E-11
Resíduo	5	0.000238001	4.76E-05		
Total	6	1.775184509			

	<i>Coefficientes</i>	<i>Erro padrão</i>	<i>Stat t</i>	<i>valor-P</i>	<i>95% inferiores</i>	<i>95% superiores</i>	<i>Inferior 95.0%</i>	<i>Superior 95.0%</i>
Interseção	-0.002970732	0.003851302	-0.771358	0.47535	-0.012870802	0.006929338	-0.012870802	0.006929338
Variável X 1	0.722829268	0.003743238	193.1027	7.07E-11	0.713206984	0.732451553	0.713206984	0.732451553

Considere uma amostra de concentração desconhecida para a qual tenham sido medidas em duplicatas autênticas, as absorvâncias 0.7304 e 0.7430. Para obter uma estimativa da concentração da amostra basta Substituir na equação (14) o valor médio dessas duas medidas, 0.7367:

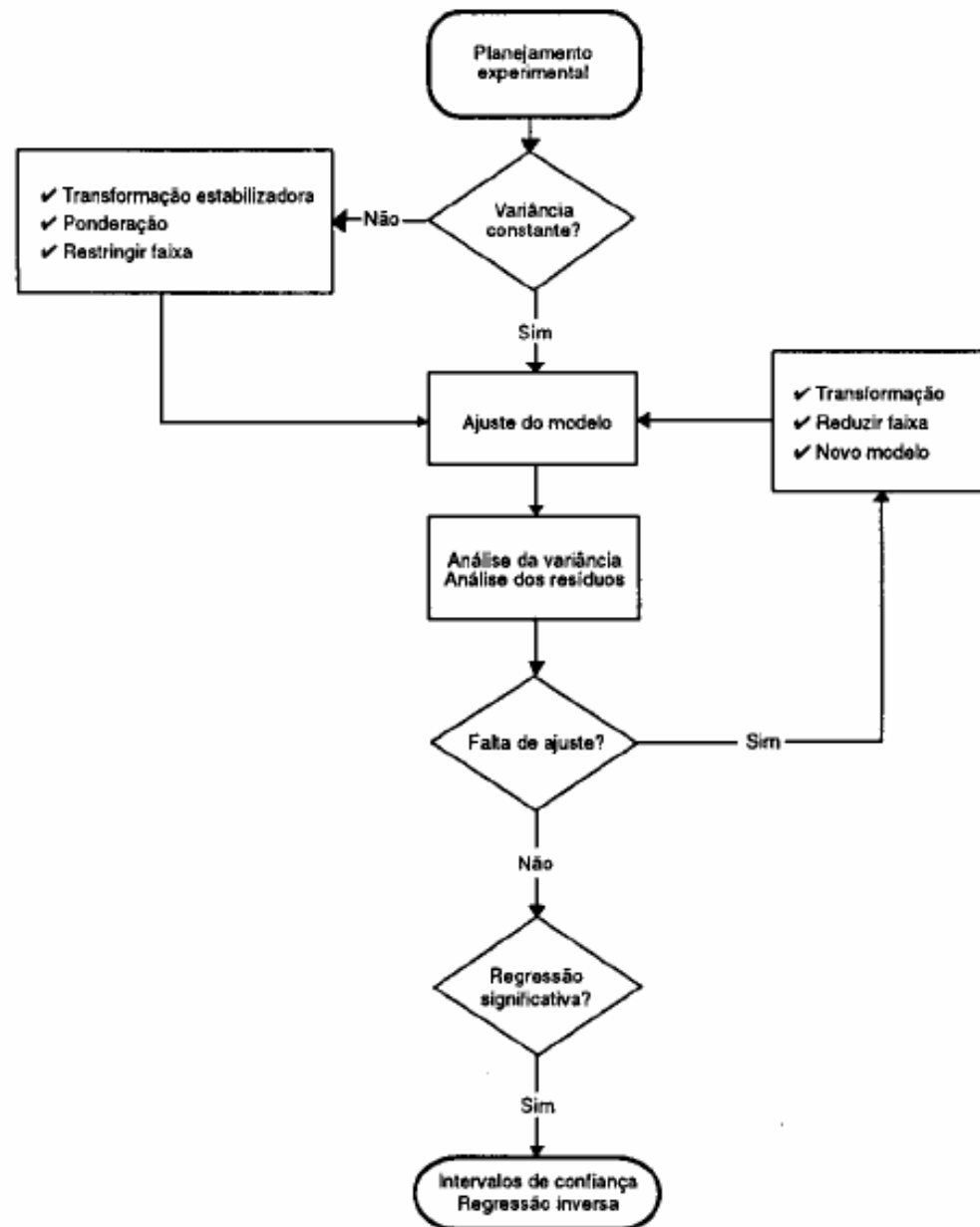
$$(\mathbf{x}_e)_0 = (0.7367 + 0.003) / 0.723 = 1.023 \text{ mg/L}$$

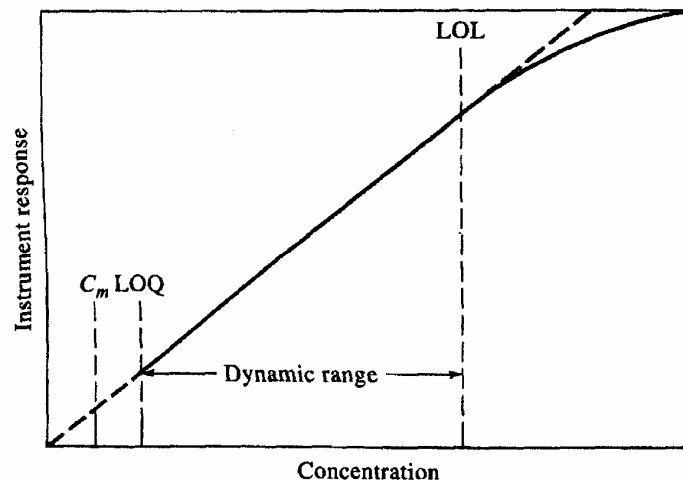
O intervalo de confiança em torno de  $(\mathbf{x}_e)_0$  é dado pela equação (15), fazendo-se  $q = 2$ , já que  $(\mathbf{x}_e)_0$  foi calculado a partir da média de duas observações. O resultado é:

$$X = 1.02 \pm 0.02 \text{ mg/L}$$

O que significa, com 95% de confiança, que a concentração de ferro na amostra deve estar entre 1.00 e 1.04 mg/L

# FLUXOGRAMA PARA O PROCEDIMENTO GERAL DE CALIBRAÇÃO





**Limite de detecção (LOD ou  $C_m$ ):** é uma mínima concentração ou massa do analito que pode ser detectado em torno de um nível de confiança conhecido. Branco mais três vezes o desvio padrão do branco (ruído). ( $LOD = 3.3 \sigma / b_1$ )

**Limite de quantificação (LOQ):** É a concentração mais baixa na qual medidas quantitativas podem ser feitas. O Mais baixo LOQ é geralmente considerado ser igual a dez vezes o desvio padrão de medidas repetidas do branco. ( $LOQ = 10 \sigma / b_1$ )

⇒ Método visual; método relação sinal ruído; método baseado em parâmetros da curva de calibração, onde  $\sigma$  pode ser o desvio padrão do branco, o erro padrão da regressão ou o erro padrão do coeficiente linear da regressão e  $b_1$  é o coeficiente angular da regressão.

**Limite de linearidade (LOL):** é a concentração na qual a curva de calibração desvia da linearidade.

Nota: “ *Entre dois métodos que apresentam igual precisão, aquele que possuir maior inclinação na curva de calibração será o mais sensível. Por outro lado, se dois métodos possuem curvas de calibração com inclinação Iguais, o que exibir melhor precisão será o mais sensível* ”

**Sensibilidade de calibração:** É a inclinação da curva de calibração em torno da concentração e interesse (IUPAC)

**Sensibilidade analítica ( $\gamma = b_1 / s$ ):** sensibilidade analítica é a inclinação da curva de calibração dividido pelo desvio padrão das medidas.



# CURVA DE CALIBRAÇÃO POR ADIÇÃO DE PADRÃO

## Motivação:

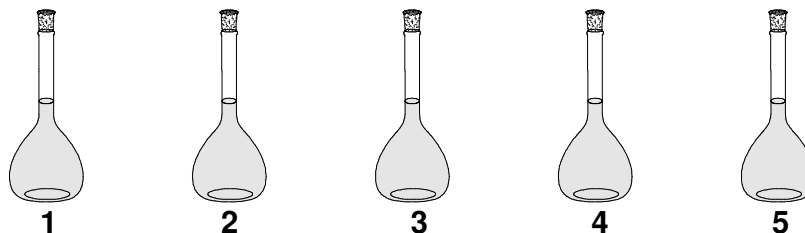
- quando a composição da amostra é desconhecida ou complexa (dificuldade em mimetizar o branco) e afeta o sinal analítico;
- baixa concentração do analito (aumentar sinal analítico);
- identificação do(s) analito(s) de interesse em presença de interferentes.

Matriz: é tudo na amostra desconhecida além do constituinte.

Efeito de matriz: é a mudança do sinal analítico causado por qualquer coisa na amostra além do constituinte.

Branco: é o sinal registrado da matriz.

## 1° CASO

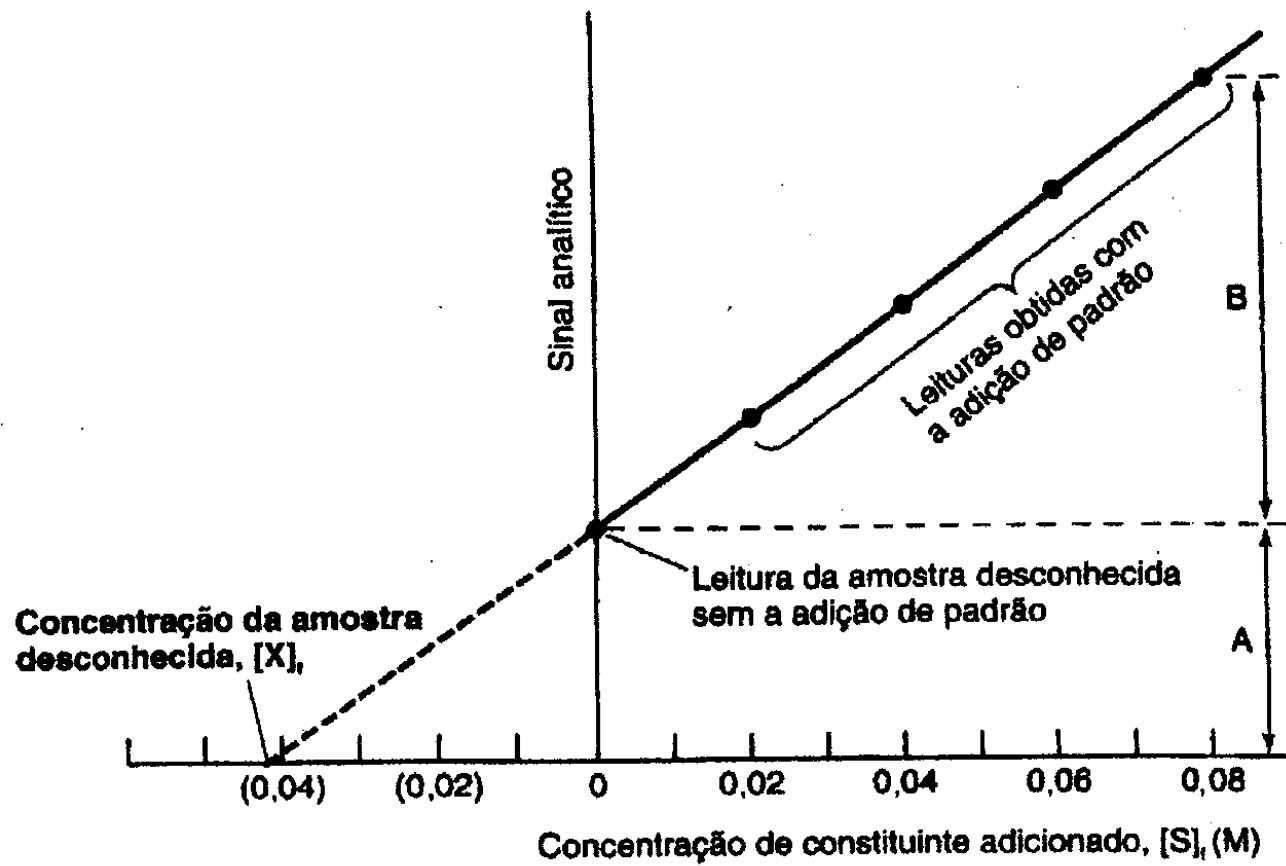


- 1- 1,0 mL de amostra (x) + 0,0 mL de padrão (s) + qsp 10,0 mL;
- 2- 1,0 mL de amostra (x) + 1,0 mL de padrão (s) + qsp 10,0 mL;
- 3- 1,0 mL de amostra (x) + 2,0 mL de padrão (s) + qsp 10,0 mL;
- 4- 1,0 mL de amostra (x) + 3,0 mL de padrão (s) + qsp 10,0 mL;
- 5- 1,0 mL de amostra (x) + 4,0 mL de padrão (s) + qsp 10,0 mL;

**OBS: suponha a solução estoque padrão contendo 100 mg L<sup>-1</sup>**

## - SINAIS OBTIDOS

1- 9 u.a. ; 2- 21 u.a. ; 3- 28 u.a. ; 4- 42 u.a. ; 5- 49 u.a. ; 6- 60 u.a.



Como a resposta do instrumento é proporcional a concentração, o que deve ocorrer se o método da adição de padrão é aplicável, pode-se escrever:

$$S = \frac{kV_s C_s}{V_t} + \frac{kV_x C_x}{V_t}$$

Onde k é uma constante de proporcionalidade. O gráfico de S em função de  $V_s$  é uma linha reta de forma

$$S = mV_s + b$$

Onde a inclinação m e o intercepto b são dados por

$$m = \frac{kC_s}{V_t} \quad e \quad b = \frac{KV_x C_x}{V_t}$$

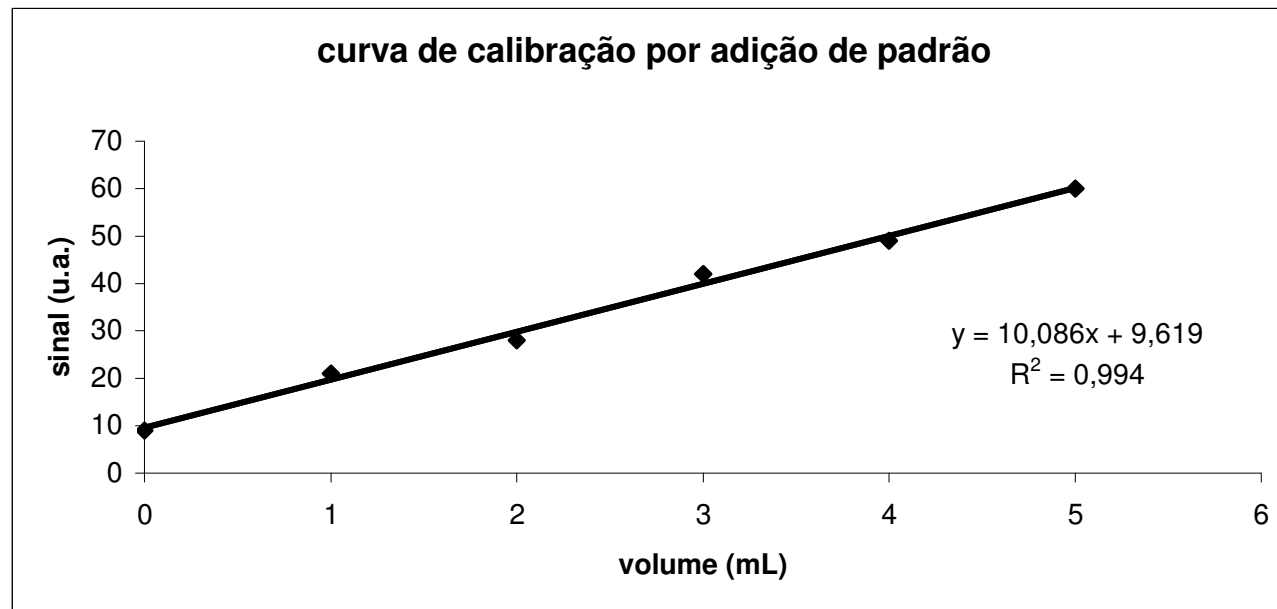
A análise dos mínimos quadrados pode ser usada para determinar m e b e relacionar com valores conhecidos de  $V_x$  e  $V_s$  no intuito de obter  $C_x$ .

$$\frac{b}{m} = \frac{\frac{KV_x C_x}{V_t}}{\frac{KC_s}{V_t}} = \frac{V_x C_x}{C_s} \Rightarrow C_x = \frac{bC_s}{mV_x}$$

O valor do desvio padrão em  $C_x$  pode ser obtido assumindo a incerteza em  $C_s$ ,  $V_s$  e  $V_t$  desprezíveis em relação aqueles em  $m$  e  $b$ . Logo, a variância relativa do resultado  $(S_x / C_x)^2$  é assumido ser a soma das variâncias relativas em  $m$  e  $b$ . Portanto,

$$S_c = C_x \sqrt{\left(\frac{S_m}{m}\right)^2 + \left(\frac{S_b}{b}\right)^2}$$

### GRÁFICO DE SINAL X VOLUME



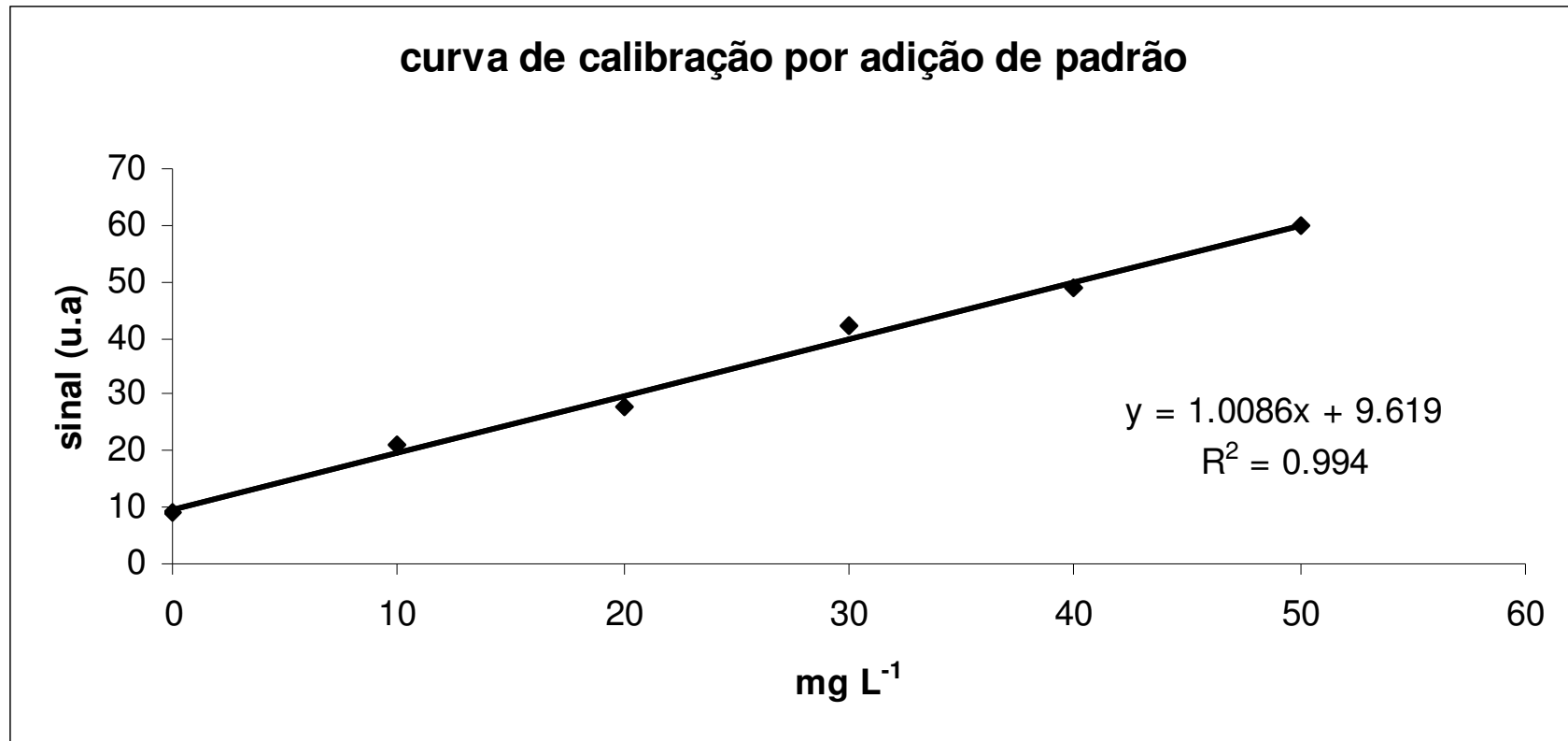
⇒ Alternativamente, o gráfico pode ser construído manualmente, e a parte linear pode ser extrapolada da esquerda da origem.

⇒ A diferença entre o volume de padrão adicionado em torno da origem (zero) e o valor do volume em torno da interseção da linha reta com o eixo, ou intercepto de x  $(V_x)_0$ , é o volume de reagente padrão equivalente a quantidade de amostra.

⇒ Adicionalmente, o intercepto em x corresponde ao zero da resposta do instrumento, e pode-se escrever da seguinte maneira:

$$S = \frac{KV_s C_s}{V_t} + \frac{KV_x C_x}{V_t} = 0 \quad \Rightarrow \quad C_x = -\frac{(V_s)_0 C_s}{V_x}$$

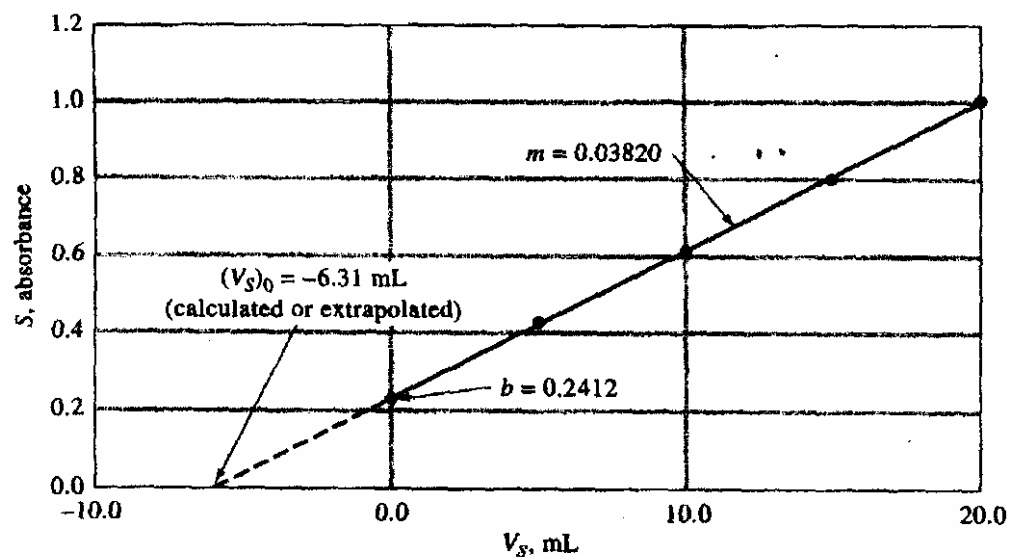
## GRÁFICO DE SINAL X CONCENTRAÇÃO



Exemplo: Alíquota de 10 mL de amostra de água natural foram pipetadas para dentro de um balão volumétrico de 50 mL. Exatamente 0,00, 5,00, 15,00 e 20,00 de uma solução padrão contendo 11,1 ppm de  $\text{Fe}^{3+}$  foram adicionados a cada, seguindo por um excesso de íon tiocianato para dar um complexo vermelho  $\text{Fe}(\text{SCN})_2^+$ . Após diluição para volume do balão, a resposta do instrumento S para cada uma das cinco soluções, medidas com um colorímetro, foram: 0,240, 0,437, 0,621, 0,809 e 1,009, respectivamente.

(a) – Qual foi a concentração de  $\text{Fe}^{3+}$  na amostra de água? **R: 7,01 ppm  $\text{Fe}^{3+}$**

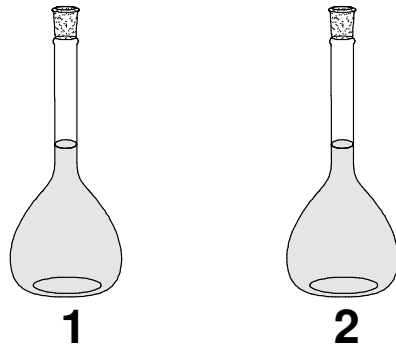
(b) – calcule o desvio padrão da inclinação e do intercepto e o desvio padrão para a concentração de  $\text{Fe}^{3+}$ . **R: 0,12 ppm  $\text{Fe}^{3+}$**



## 2° CASO

### Motivação:

- pouca quantidade de amostra disponível para análise



- 1- 1,0 mL de amostra (x) + 0,0 mL de padrão (s) + qsp 10,0 mL;  
2- 1,0 mL de amostra (x) + 1,0 mL de padrão (s) + qsp 10,0 mL;

**OBS: suponha a solução estoque padrão contendo 100 mg L<sup>-1</sup>**

**- SINAIS OBTIDOS**

**1- 9 u.a. ; 2- 21 u.a.**



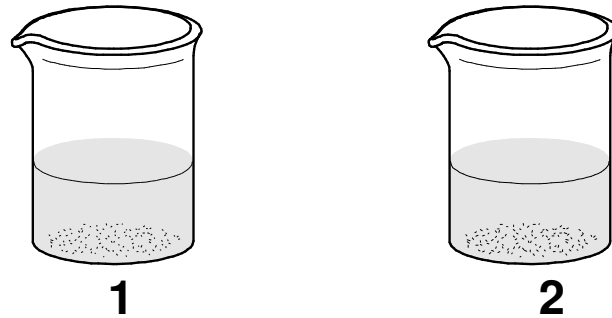
**EXEMPLO-** 10,00 g de amostra contendo um analito x foi transferido para um balão volumétrico de 250 mL e acertado o menisco com solvente apropriado. 10,00 mL de uma alíquota da solução resultante (250 mL) foi diluída para um balão de 25,00 mL, sendo realizada leitura em equipamento espectrofotométrico e obtido sinal de 0,235 (unidades de absorvância). Uma segunda alíquota de 10,00 mL da amostra (250 mL) foi diluída juntamente com 10 mL de solução padrão 1,00 ppm (mg/L) do analito x para um segundo balão volumétrico de 25,00 mL. O sinal para a amostra mais padrão foi de 0,502.

**(a)** Calcule a porcentagem em peso do analito na amostra original.

### 3° CASO

#### Motivação:

- quantidade de amostra disponível para análise menor que o 2°CASO



1- 1 mL de amostra (x)

2- 1 mL de amostra (x) + 1 mL de padrão (s)

**OBS: suponha a solução estoque padrão contendo  $100 \text{ mg L}^{-1}$**

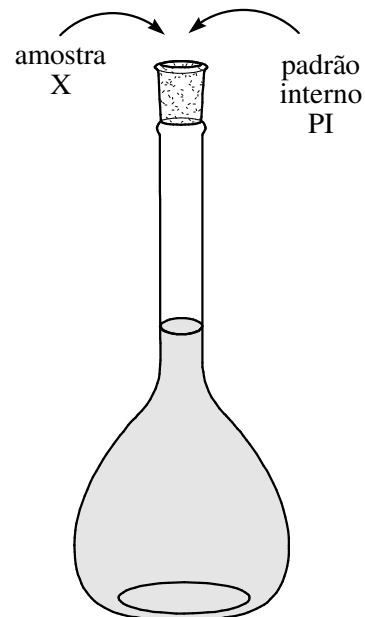
**- SINAIS OBTIDOS**

**1- X u.a. ; 2- X+ S u.a.**

# CURVA DE CALIBRAÇÃO COM USO DE PADRÃO INTERNO

## Motivação:

- corrigir variações instrumentais;
- corrigir variações de concentração decorrentes do processo de preparo de amostra;
- identificação do constituinte de interesse em presença de outros constituintes
- o constituinte utilizado como padrão interno (PI) deve ter características semelhantes ao(s) analito(s) de interesse, porém não deve fazer parte da amostra e nem coeluir com o(s) analito(s) que deverão ser determinados.



$$S_x \propto n_x \Rightarrow S_x = K_x n_x \Rightarrow S_x = K_x [x] V$$

$$S_{PI} \propto n_{PI} \Rightarrow S_{PI} = K_{PI} n_{PI} \Rightarrow S_{PI} = K_{PI} [PI] V$$

$$\frac{S_x}{S_{PI}} = \frac{K_x}{K_{PI}} \cdot \frac{[x]}{[PI]} \Rightarrow \frac{S_x}{[x]} = \frac{K_x}{K_{PI}} \cdot \frac{S_{PI}}{[PI]}$$

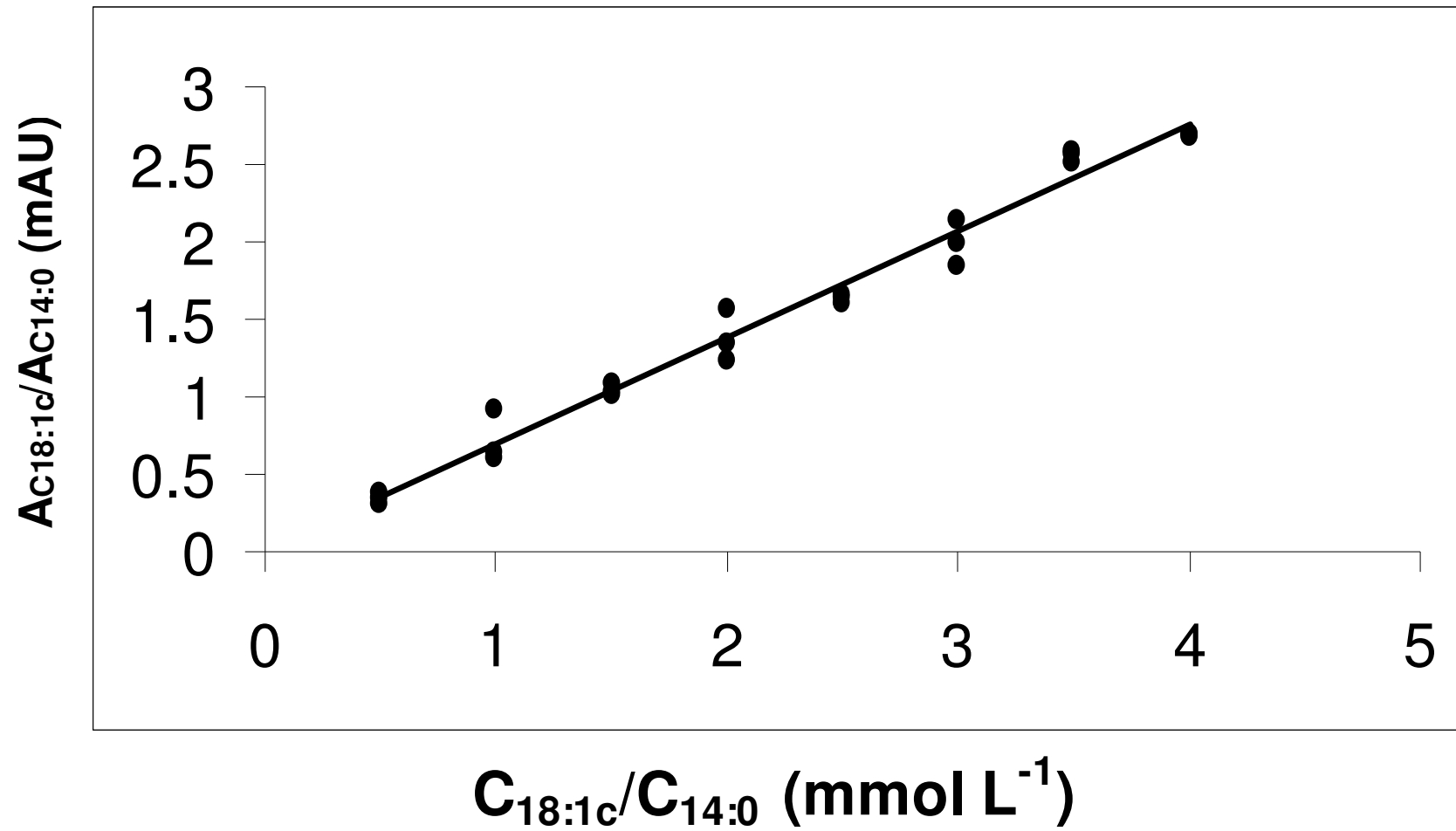


**FATOR DE RESPOSTA**

**Tabela1-** Planejamento experimental com um fator

Nível	Fator	área C18:1c * área C14:0		
	C18:1c mmol L <sup>-1</sup>	1	2	3
1	0.25	0.397	0.353	0.315
2	0.50	0.607	0.643	0.923
3	0.75	1.01	1.09	1.03
4	1.00	1.58	1.36	1.24
5	1.25	1.64	1.66	1.61
6	1.50	2.00	2.14	1.85
7	1.75	2.59	2.58	2.51
8	2.00	2.70	2.68	-----

\* Área do C14:0 em concentração fixa de 0.5 mmol L<sup>-1</sup>



RESUMO DOS RESULTADOS

<i>Estatística de regressão</i>	
R múltiplo	0.989265969
R-Quadrado	0.978647158
R-quadrado ajustado	0.977630356
Erro padrão	0.117466491
Observações	23

ANOVA

	<i>gl</i>	<i>SQ</i>	<i>MQ</i>	<i>F</i>	<i>F de significação</i>
Regressão	1	13.28060131	13.2806	962.4756	5.00291E-19
Resíduo	21	0.289765908	0.013798		
Total	22	13.57036722			

	<i>Coefficientes</i>	<i>Erro padrão</i>	<i>Stat t</i>	<i>valor-P</i>	<i>95% inferiores</i>	<i>95% superiores</i>	<i>Inferior 95.0%</i>	<i>Superior 95.0%</i>
Interseção	0.011245776	0.053886973	0.208692	0.8367	-0.100818338	0.12330989	-0.100818338	0.12330989
Variável X 1	0.684986943	0.022079408	31.02379	5E-19	0.639070293	0.730903593	0.639070293	0.730903593

$$Y (\pm 0.117) = 0.685(\pm 0.0220) [C18:1c]/[C14:0] + 0.0112(\pm 0.0530)$$

$$R^2 = 0.9786$$

<b>Source of variation</b>	<b>D.f.</b>	<b>Sum of squares</b>	<b>Mean square</b>	<b>F</b>
<b>Regression</b>	1.000	13.28	13.28	1147
<b>Lack of fit</b>	6.000	0.1100	0.01800	1.550
<b>Treatments</b>	7.000	13.39	1.910	165.0
<b>Residual</b>	15.00	0.1700	0.01200	
<b>Total</b>	22.00	13.56		

**Exemplo:** Em um experimento preliminar, uma solução contendo 0,0837 M de X e 0,0666M de S fornece picos com áreas de  $A_x = 423$  e  $A_s = 347$ . Para analisar a amostra desconhecida, 10 mL de uma solução 0,146 M de S foram adicionados a 10,0 mL da amostra desconhecida, e a mistura foi diluída a 25,0 mL em balão volumétrico. A mistura forneceu o cromatograma da figura abaixo, para a qual  $A_x = 553$  e  $A_s = 582$ . Encontre a concentração de X na amostra desconhecida.  
**R:  $[X] = 0,143 M$**

