

5 Comparação de métodos

5.1 Exatidão e Precisão do Método

A *exatidão* é o grau de concordância entre o resultado de uma medição e um *valor verdadeiro convencional* do mensurando (VIM, 2000). A *exatidão* do método é a proximidade do valor obtido do analito em relação ao *valor verdadeiro convencional*.

O termo precisão não deve ser utilizado como exatidão.

A *exatidão* de um processo de medição está relacionada aos erros aleatórios e sistemáticos e deve ser sempre avaliada antes de sua aplicação em amostras futuras. A verificação da *exatidão* pode ser um dos passos básicos para garantir a *rastreabilidade*.

A *rastreabilidade* é a propriedade de um resultado de medição ou do valor de um padrão estar relacionado a referências estabelecidas, geralmente padrões nacionais ou internacionais, por meio de uma cadeia de comparações, seguindo uma hierarquia metrológica.

O conceito internacional de *rastreabilidade* pode ser aplicado ao procedimento de medições químicas, comparando estatisticamente o método em estudo com um método de referência apropriado (EURACHEM, 1998).

A Figura 5.1 compara os resultados do método em estudo com dois tipos de referências: com o método de referência e com o material de referência certificado, MRC. No entanto, outras referências importantes podem ser usadas.

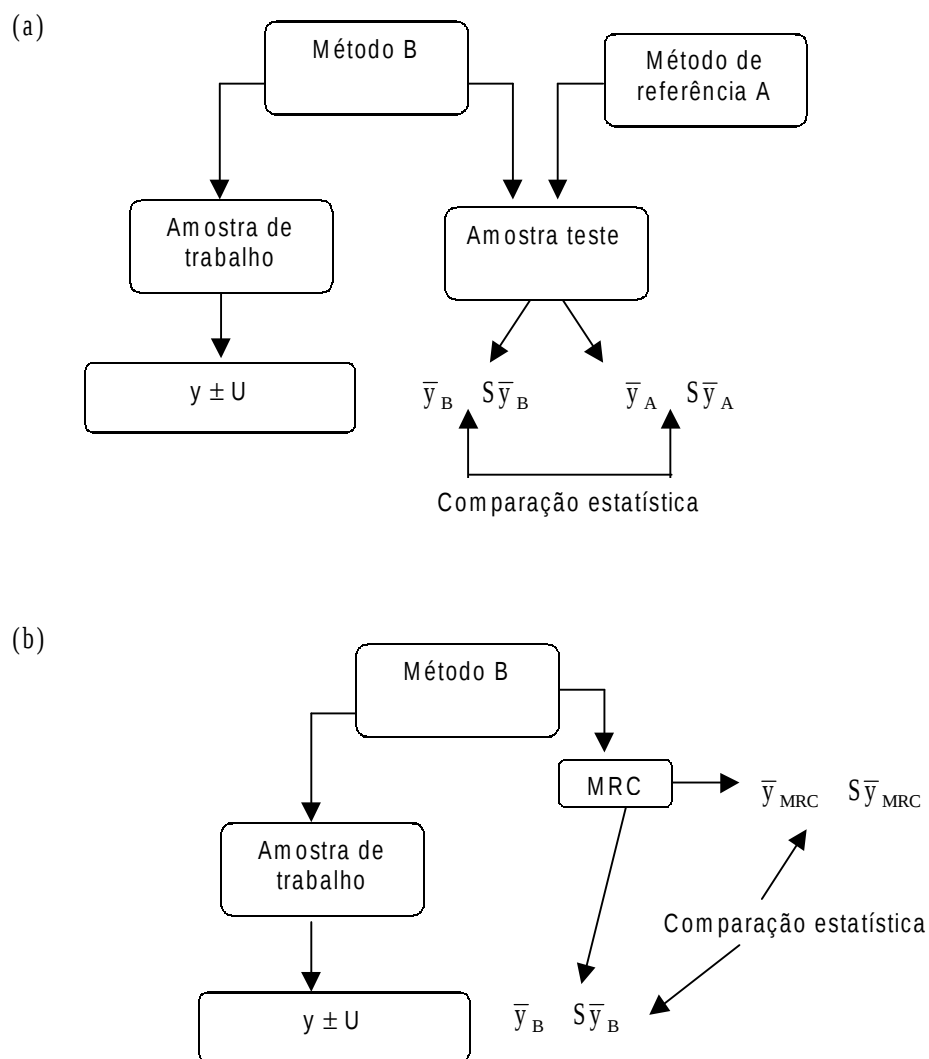


Figura 5.1. Avaliação da exatidão de um método B a ser testado. (a) Comparação com um método de referência A. (b) Comparação com o material de referência certificado, MRC.

A Figura 5.1(a) compara a *exatidão* do método B com o método de referência A. É o procedimento mais comum em muitos laboratórios, quando um método alternativo é desenvolvido para substituir um método de referência em análise de rotina.

A *exatidão* do método pode ser examinada por outros meios. Por exemplo, a Figura 5.1 (b) mostra o processo de comparação do método B com o material de referência certificado, MRC. Neste caso, o MRC é analisado com o método B pela variação de todos os fatores que podem afetar de forma representativa o resultado.

A média geral e o desvio-padrão da média do método B são comparados estatisticamente aos dos valores de referência certificados. Se a referência utilizada na avaliação da *exatidão* tiver baixo nível de *rastreabilidade*, níveis mais altos devem ser alcançados para estabelecer a conexão dessa referência a um teste interlaboratorial, a um método primário, a um MRC ou, em alguns casos, ao Sistema Internacional de Unidades, SI.(Maroto *et al.*, 1999)

Com o aumento de passos na cadeia de rastreabilidade, o resultado das medições é influenciado pelos erros inerentes aos passos adicionais e pela propagação dos mesmos.

Se a *exatidão* do método B tiver sido avaliada a partir da comparação com um método de referência A em um único laboratório, o projeto experimental deve considerar todas as fontes de variação, levando em consideração, sobretudo, que a amostra para análise deve ser homogênea, estável e representativa para amostras subsequentes.

A avaliação da *exatidão* inclui a avaliação do erro sistemático e da precisão do método.

O erro sistemático, conhecido também como *tendência* do método, é a parcela de erro que se mantém constante quando se fazem várias medições do mesmo analito, sob as mesmas condições, fazendo com que a média do conjunto de medições se afaste do *valor verdadeiro convencional*.

A *precisão* do método é o grau de concordância entre ensaios mutuamente independentes e é normalmente apresentada em forma de desvio-padrão: o desvio-padrão da parcela aleatória do erro de medição, por definição de média nula.

A edição brasileira 2000 do Vocabulário Internacional de Termos Fundamentais e Gerais de Metrologia, VIM, não considera o termo “precisão”; no entanto, feita esta ressalva, neste trabalho ele será empregado devido a ter uso consagrado na literatura de metrologia química, inclusive sendo definido nas Normas ISO.

As principais medidas de precisão incluem o desvio-padrão associado à *repetitividade*, S_r , o desvio-padrão associado à *reprodutibilidade*, S_R , (ISO 3534-1) e a *precisão intermediária*, algumas vezes designada $S_{I(A)}$ (ISO 5725-3:1994) (EURACHEM, 1998).

De acordo com o VIM, a *repetitividade* é o grau de concordância entre os resultados de medições sucessivas de um mesmo mensurando efetuadas sob as mesmas condições de medição.

Estas condições são denominadas condições de repetitividade.

Condições de repetitividade incluem:

- mesmo procedimento de medição;
- mesmo observador;
- mesmo instrumento de medição, utilizado nas mesmas condições;
- mesmo local;
- repetição em curto período de tempo.

A *repetitividade* pode ser expressa, quantitativamente, em função das características da dispersão dos resultados (VIM, 2000).

Em medições químicas, para um lote experimental, a *repetitividade* expressa o erro aleatório através de dispersão obtida de uma série de medições repetidas, realizadas durante o mesmo ensaio.

O desvio-padrão associado a *repetitividade*, S_r , pode ser obtido em um laboratório ou por estudos interlaboratoriais, indicando a variabilidade observada entre resultados de medições sucessivas de um mesmo mensurando, sob as mesmas condições de medição (*condição de repetitividade*).

O desvio-padrão associado à *reprodutibilidade* interlaboratorial, S_R , para um método particular, só pode ser estimado diretamente por estudos interlaboratoriais

e evidencia a variabilidade obtida quando diferentes laboratórios analisam a mesma amostra.

A *precisão intermediária* está relacionada com a variação dos resultados observada quando um ou mais fatores, tais como tempo, equipamento e técnico, são variados dentro de um laboratório; obtêm-se diferentes valores dependendo dos fatores que se mantêm constantes. Por exemplo, o projeto experimental da Tabela 5.1 pode ser utilizado se estão envolvidos dois operadores em operação de rotina.

As estimativas da *precisão intermediária* são normalmente feitas intralaboratorialmente, mas podem também ser obtidas interlaboratorialmente.

Tabela 5.1. Exemplo de projeto para avaliação da precisão intermediária.

Série						
Fatores	1	2	3	4	5	6
Dia	1	1	2	2	3	3
Operador	A	B	A	B	A	B

É importante notar que as diferentes amostras para cada resultado não precisam ser necessariamente obtidas da mesma forma. Por exemplo, o laboratório A pode obter os resultados em 10 dias e o laboratório B em 20 dias.

Uma das diretrizes mais avançadas na comparação de dois métodos é encontrada na ISO 5725-6, embora não seja aplicada a laboratórios individuais, pois é baseada em informações interlaboratoriais.

Para apenas um laboratório, a reprodutibilidade, não pode ser avaliada como especificado pela ISO, mas a precisão intermediária pode, desde que esta contribua para a variabilidade das medições do laboratório (Maroto, 1999).

Quando um laboratório deseja substituir um método analítico por outro, deve-se avaliar o desempenho dos dois através da comparação da precisão e tendência. Uma forma de avaliação sugerida pela ISO é através de um modelo hierárquico de 4 níveis.

Quatro fatores representam as fontes de variação que contribuem para a variabilidade das mensurações feitas em um laboratório, que são: operador, instrumento, tempo e erro aleatório.

Cada amostra é analisada em n replicatas por i operadores, que utilizam b instrumentos, em diferentes dias. Para evitar a sub-estimação do efeito do dia em que as medições são feitas, l diferentes equipamentos no mesmo dia devem ser utilizados.

5.2

Modelo estatístico básico

Assume-se que o resultado do teste y é obtido pela soma de cinco componentes:

$$y = \mu + O + I + D + \varepsilon \quad (5.1)$$

Onde:

μ é a “média” ou “valor esperado” do resultado do teste;

O é o efeito aleatório devido à mudança de operador;

I é o efeito aleatório devido à mudança de equipamento;

D é o efeito aleatório que surge quando as medições são feitas em dias diferentes;

ε é o efeito aleatório oriundo de medições feitas sob condições de repetitividade.

A precisão do método é determinada pela variância total, σ^2 , obtida pela soma das contribuições de cada fator para ela, isto é: σ^2_O , σ^2_I , σ^2_D e σ^2_r , que são estimadas por S^2_O , S^2_I , S^2_D e S^2_r respectivamente.

Assume-se que os fatores citados anteriormente não são correlacionados, de modo que a estimativa da precisão intermediária geral correspondente ao efeito devido a diferentes dias de análise, $S^2_{I(OT)}$, é obtida pela soma das componentes de variância: $S^2_O + S^2_I + S^2_D + S^2_r$. Entretanto, quando em um laboratório a análise é feita por um operador em um mesmo instrumento, então a estimativa da precisão intermediária é dada por $S^2_{I(T)} = S^2_D + S^2_r$.

5.2.1

Estimativa dos componentes de variância

As estimativas dos componentes de variância são obtidas pela condução da análise de variância. A representação e abreviações desta seção são:

i - para um operador particular, $i=1,2,...,a$;

j - para um instrumento particular, $j=1,2,...,b$;

k - para um dia particular, $k=1,2,...,c$;

b_i - número de instrumentos utilizados pelo i-ésimo operador;

c_{ij} - número de dias de análises do i-ésimo operador utilizando o j-instrumento;

n_{ijk} - número de replicatas por dia no k dia analisado pelo operador i, no j-ésimo instrumento; e

L - resultado de uma análise no k-ésimo dia, conduzida pelo i-ésimo operador, utilizado o j-ésimo instrumento, $L=1,2,...,n_{ijk}$.

Dependendo de algumas das fontes de variação diferirem quanto à quantidade, existe diferença na obtenção das componentes de análise de variância, por isso consideram-se dois casos.

1º Caso – Na análise com ambos os métodos, considera-se números diferentes de replicatas, instrumentos e operadores. O cálculo das componentes de variância é como segue:

a) Operador + instrumento + dia

$$QM_{OID} = \frac{\sum_{i=1}^a \sum_{j=1}^{b_i} \sum_{k=1}^{c_{ij}} n_{ijk} (\bar{y}_{ijk} - \bar{y}_{....})^2}{\sum_{i=1}^a \sum_{j=1}^{b_i} c_{ij} - 1} \quad (5.2)$$

$$\bar{y}_{ijk} = \frac{\sum_{L=1}^{n_{ijk}} y_{ijkL}}{n_{ijk}}, \quad \bar{y}_{ij.} = \frac{\sum_{k=1}^{c_{ij}} n_{ijk} \bar{y}_{ijk}}{\sum_{k=1}^{c_{ij}} n_{ijk}}, \quad \bar{y}_{....} = \frac{\sum_{i=1}^a \sum_{j=1}^{b_i} \sum_{k=1}^{c_{ij}} n_{ijk} \bar{y}_{ijk}}{N} \quad (5.3)$$

$$\bar{n} = \frac{N - \left(\sum_{i=1}^a \sum_{j=1}^{b_i} \sum_{k=1}^{c_{ij}} (n_{ijk})^2 \right)}{\sum_{i=1}^a \sum_{j=1}^{b_i} c_{ij} - 1}, \quad N = \sum_{i=1}^a \sum_{j=1}^{b_i} \sum_{k=1}^{c_{ij}} n_{ijk} \quad (5.4)$$

b) Dia

$$QM_D = \frac{\sum_{i=1}^a \sum_{j=1}^{b_i} \sum_{k=1}^{c_{ij}} n_{ijk} (\bar{y}_{ijk} - \bar{y}_{ij})^2}{\sum_{i=1}^a \sum_{j=1}^{b_i} (c_{ij} - 1)} \quad (5.5)$$

c) Resíduo

$$QM_r = \frac{\sum_{i=1}^a \sum_{j=1}^{b_i} \sum_{k=1}^{c_{ij}} \sum_{L=1}^{n_{ijk}} (y_{ijkL} - \bar{y}_{ijk})^2}{\sum_{i=1}^a \sum_{j=1}^{b_i} \sum_{k=1}^{c_{ij}} (n_{ijk} - 1)} \quad (5.6)$$

Os componentes de variância do modelo hierárquico são apresentados nas Tabelas 5.2 e 5.3 a seguir.

Tabela 5.2. Componentes de variância do modelo hierárquico

Componente	Estimativa	EQM	Observação
Repetitividade, S_r^2	QM_r	σ_r^2	$v = \sum_{i=1}^a \sum_{j=1}^{b_i} \sum_{k=1}^{c_{ij}} (n_{ijk} - 1)$ se $S_D^2 < 0, S_D^2 = 0$ se $S_{OID}^2 < 0, S_{OID}^2 = 0$
Dias, S_D^2	$S_D^2 = \frac{QM_D - QM_r}{\bar{n}}$	$\sigma_r^2 + \bar{n}\sigma_D^2$	
Operador+Instrumento+Dia, S_{OID}^2	$S_{OID}^2 = \frac{QM_{OID} - QM_r}{\bar{n}}$	$\sigma_r^2 + \bar{n}\sigma_{OID}^2$	
Intermediária, por diferentes tempos, $S_{I(T)}^2$	$S_{I(T)}^2 = S_D^2 + S_r^2$		
Intermediária, por diferentes operadores, instrumentos e tempo, $S_{I(OT)}^2$	$S_{I(OT)}^2 = S_{OID}^2 + S_r^2$		
Média do dia $S_{\bar{y}_{ijk}}^2$	$S_{OID}^2 + S_r^2 / \bar{n}$		$v = \sum_{i=1}^a \sum_{j=1}^{b_i} c_{ij} - 1$

2° Caso – Na análise com ambos os métodos, considera-se o mesmo número de replicatas n ($L=1,2,...,n$), mesmo número de instrumentos, b ($j=1,2,...,b$) e mesmo número de operadores a ($i=1,2,...,a$). O cálculo das componentes de variância se segue:

a) Operador + instrumento+ dia

$$QM_{OID} = \frac{n \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^{c_{ij}} n_{ijk} (\bar{y}_{ijk} - \bar{y}_{...})^2}{\sum_{i=1}^a \sum_{j=1}^b c_{ij} - 1} \quad (5.7)$$

$$\bar{y}_{ijk} = \frac{\sum_{L=1}^n y_{ijkL}}{n}, \quad \bar{y}_{ij..} = \frac{\sum_{k=1}^{c_{ij}} \bar{y}_{ijk}}{c_{ij}}, \quad \bar{y}_{...} = \frac{\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^{c_{ij}} \sum_{L=1}^n y_{ijkL}}{n \sum_{i=1}^a \sum_{j=1}^b c_{ij}} \quad (5.8)$$

b) Dia

$$QM_D = \frac{n \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^{c_{ij}} n_{ijk} (\bar{y}_{ijk} - \bar{y}_{ij..})^2}{\sum_{i=1}^a \sum_{j=1}^b (c_{ij} - 1)} \quad (5.9)$$

c) Resíduo

$$QM_r = \frac{\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^{c_{ij}} \sum_{L=1}^n (y_{ijkL} - \bar{y}_{ijk})^2}{(n-1) \sum_{i=1}^a \sum_{j=1}^b c_{ij}} \quad (5.10)$$

Tabela 5.3. Componentes de variância do modelo hierárquico

Componente	Estimativa	EQM	Observação
Repetitividade, S_r^2	QM_r	σ_r^2	$v = (n-1) \sum_{i=1}^a \sum_{j=1}^b c_{ij}$ se $S_D^2 < 0, S_D^2 = 0$ se $S_{OID}^2 < 0, S_{OID}^2 = 0$ $v = \sum_{i=1}^a \sum_{j=1}^b c_{ij} - 1$
Dias, S_D^2	$S_D^2 = \frac{QM_D - QM_r}{n}$	$\sigma_r^2 + n\sigma_D^2$	
Operador+Instrumento+Dia, S_{OID}^2	$S_{OID}^2 = \frac{QM_{OID} - QM_r}{n}$	$\sigma_r^2 + n\sigma_{OID}^2$	
Intermediária, por diferentes tempos, $S_{I(T)}^2$	$S_{I(T)}^2 = S_D^2 + S_r^2$		
Intermediária, por diferentes operadores, instrumentos e tempo $S_{I(OT)}^2$	$S_{I(OT)}^2 = S_{OID}^2 + S_r^2$		
Média do dia $S_{\bar{y}_{ijk}}^2$	$S_{OID}^2 + S_r^2 / n$		

Através das fórmulas para calcular as componentes de variância S_D^2 e S_{OID}^2 (Tabelas 5.3 e 5.4), pode-se obter valores negativos para estas estimativas.

A ISO *apud* Kuttatharmmakul *et al*, considera estas estimativas zero, como também quando isto acontece com a variância entre laboratórios S_L^2 .

Montgomery & Runger (1994) oferecem outras alternativas, através da modificação do modelo pela exclusão da fonte de variação que apresenta uma componente de variação negativa, ou pela utilização do método de máxima verossimilhança para estimação das componentes com a restrição da variância ser maior ou igual a zero.

5.3 Número de medições

Para avaliar a *tendência* do método alternativo e a diferença entre a sua precisão e a do método de referência, testa-se a hipótese de diferença entre as médias dos métodos, bem como a razão entre os parâmetros de precisão a serem especificados.

Para testar a tendência entre os métodos, especifica-se a menor tendência que é importante detectar com uma alta probabilidade, denominada de λ ; na

verdade, é a menor diferença entre as médias dos dois métodos que é importante detectar com uma alta probabilidade.

Analogamente, para comparar a precisão dos métodos, utiliza-se o menor valor que se deseja detectar com uma alta probabilidade, para a razão entre os parâmetros que representam a precisão que, na situação intralaboratorial, podem ser:

- Para comparação de *repetitividade*:

$$\rho = \frac{\sigma_{rA}}{\sigma_{rB}} \quad (5.11)$$

- Para comparação da *precisão intermediária* devido a diferentes tempos de análise:

$$\Phi_{I(T)} = \sqrt{\frac{QM_{DB}}{QM_{DA}}} \quad (5.12)$$

- Para comparação da *precisão intermediária* devido a diferentes operadores, instrumentos e tempo:

$$\Phi_{I(OT)} = \sqrt{\frac{QM_{OIDB}}{QM_{OIDA}}} \quad (5.13)$$

Devido à complexidade de se calcular o número de graus de liberdade associados a $\sigma_{I(T)}$, $\sigma_{I(OT)}$, as razões referentes a estas são dadas em termos dos quadrados médios.

5.3.1

Determinação do número mínimo de medições necessárias à obtenção de λ

Na escolha do tamanho da amostra quando se deseja detectar uma diferença específica, consideram-se as seguintes hipóteses:

$$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{cases}$$

Considere $f_{\bar{D}}(\bar{d} | \mu = \mu_0)$ e $f_{\bar{D}}(\bar{d} | \mu \neq \mu_0)$ as densidades de probabilidade da diferença entre as médias amostrais dos dois métodos.

O tamanho da amostra deve permitir que se obtenha uma variância $\sigma_{\bar{D}}^2$ tal que exista um valor $\bar{d}_0$ que divida o eixo das abscissas $\bar{d}$ com 0,05 de probabilidade à direita, sob H_0 e 0,95 de probabilidade à direita, sob H_1 (Figura 5.2).

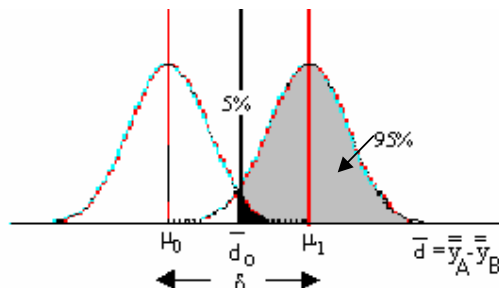


Figura 5.2. Escolha do tamanho da amostra com segurança desejada

Tem-se um teste unilateral com $\alpha=5\%$ associado a H_0 e $\beta=5\%$ associada a H_1 , com poder de 95%.

Para determinação do número mínimo de medições necessárias à obtenção de um poder $(1-\beta)$ para a detecção de uma tendência igual a λ , utiliza-se a expressão abaixo, baseada no teste t de *Student* de comparação entre duas médias, assumindo que a variância das medições de ambos os métodos sejam iguais:

$$\lambda \geq (t_{\alpha/2} + t_{\beta}) \sqrt{\frac{(a_A b_A c_A)(S_{OIDA}^2 + S_{rA}^2 / n_A) + (a_B b_B c_B)(S_{OIDB}^2 + S_{rB}^2 / n_B)}{a_A b_A c_A + a_B b_B c_B - 2}} \left(\frac{1}{a_A b_A c_A} + \frac{1}{a_B b_B c_B} \right) \quad (5.14)$$

Onde:

a_A e a_B representam o número de operadores que conduzem os métodos A e B;

b_A e b_B representam o número de instrumentos utilizados nos métodos A e B;

c_A e c_B representam o número de dias de análises utilizando os métodos A e B;

n_A e n_B representam o número de replicatas de análises utilizando os métodos A e B.

Se a precisão do método B é desconhecida, o que acontece com frequência, utiliza-se a precisão do método de referência A, de modo que:

$$\lambda \geq (t_{\alpha/2} + t_{\beta}) \sqrt{\frac{(a_A b_A c_A)(S_{\text{OIDA}}^2 + S_{\text{rA}}^2 / n_A) + (a_B b_B c_B)(S_{\text{OIDA}}^2 + S_{\text{rA}}^2 / n_B)}{a_A b_A c_A + a_B b_B c_B - 2} \left(\frac{1}{a_A b_A c_A} + \frac{1}{a_B b_B c_B} \right)} \quad (5.15)$$

Recomenda-se que ambos os métodos tenham o mesmo número de operadores, $a_A = a_B$, a mesma quantidade de instrumentos, $b_A = b_B$, e que as análises sejam conduzidas na mesma quantidade de dias, $c_A = c_B$, e com a mesma quantidade de replicatas, $n_A = n_B$. Assim, a expressão acima (5.15) se reduz a:

$$\lambda \geq (t_{\alpha/2} + t_{\beta}) \sqrt{\frac{2(S_{\text{OIDA}}^2 + S_{\text{rA}}^2 / n_A)}{a_A b_A c_A}} \quad (5.16)$$

5.3.2

Determinação do número mínimo de medições necessárias à obtenção da menor razão entre os parâmetros que representam a precisão

A menor razão entre os parâmetros que representam a precisão é dada (Tabelas 5.4(a) e 5.4(b) por):

$$\rho, \phi_{I(T)} \text{ ou } \phi_{I(OIT)} = \sqrt{F_{\beta(v_A, v_B)} \times F_{\alpha(v_A, v_B)}} \quad (5.17)$$

com os graus de liberdade para ρ , $\Phi_{I(T)}$ ou $\Phi_{I(OIT)}$ dados por:

- Para ρ : $v_A = a_A b_A c_A (n_A - 1)$ e $v_B = a_B b_B c_B (n_B - 1)$;
- Para $\Phi_{I(T)}$: $v_A = a_A b_A (c_A - 1)$ e $v_B = a_B b_B (c_B - 1)$; e
- Para $\Phi_{I(OIT)}$: $v_A = (a_A b_A c_A - 1)$ e $v_B = (a_B b_B c_B - 1)$.

**Tabela 5.4 (a). Valores de ρ (vA, vB) $\Phi I(T)$ (vA, vB) ou $\Phi I(OIT)$ (vA, vB) ,
 $\alpha=5\%$ e $\beta=5\%$**

v _B	v _A								
	5	6	7	8	9	10	11	12	13
5	5,05	4,95	4,88	4,82	4,77	4,74	4,70	4,68	4,66
6	4,39	4,28	4,21	4,15	4,10	4,06	4,03	4,00	3,98
7	3,97	3,87	3,79	3,73	3,68	3,64	3,60	3,57	3,55
8	3,69	3,58	3,50	3,44	3,39	3,35	3,31	3,28	3,26
9	3,48	3,37	3,29	3,23	3,18	3,14	3,10	3,07	3,05
10	3,33	3,22	3,14	3,07	3,02	2,98	2,94	2,91	2,89
11	3,20	3,09	3,01	2,95	2,90	2,85	2,82	2,79	2,76
12	3,11	3,00	2,91	2,85	2,80	2,75	2,72	2,69	2,66
13	3,03	2,92	2,83	2,77	2,71	2,67	2,63	2,60	2,58
14	2,96	2,85	2,76	2,70	2,65	2,60	2,57	2,53	2,51
15	2,90	2,79	2,71	2,64	2,59	2,54	2,51	2,48	2,45
16	2,85	2,74	2,66	2,59	2,54	2,49	2,46	2,42	2,40
17	2,81	2,70	2,61	2,55	2,49	2,45	2,41	2,38	2,35
18	2,77	2,66	2,58	2,51	2,46	2,41	2,37	2,34	2,31
19	2,74	2,63	2,54	2,48	2,42	2,38	2,34	2,31	2,28
20	2,71	2,60	2,51	2,45	2,39	2,35	2,31	2,28	2,25
25	2,60	2,49	2,40	2,34	2,28	2,24	2,20	2,16	2,14
50	2,40	2,29	2,20	2,13	2,07	2,03	1,99	1,95	1,92
200	2,26	2,14	2,06	1,98	1,93	1,88	1,84	1,80	1,77

**Tabela 5.4 (b). Valores de ρ (vA, vB) $\Phi_{I(T)}$ (vA, vB) ou $\Phi_{I(OIT)}$ (vA, vB) ,
 $\alpha=5\%$ e $\beta=5\%$**

v _B	v _A									
	14	15	16	17	18	19	20	25	50	200
5	4,64	4,62	4,60	4,59	4,58	4,57	4,56	4,52	4,44	4,39
6	3,96	3,94	3,92	3,91	3,90	3,88	3,87	3,83	3,75	3,69
7	3,53	3,51	3,49	3,48	3,47	3,46	3,44	3,40	3,32	3,25
8	3,24	3,22	3,20	3,19	3,17	3,16	3,15	3,11	3,02	2,95
9	3,03	3,01	2,99	2,97	2,96	2,95	2,94	2,89	2,80	2,73
10	2,86	2,85	2,83	2,81	2,80	2,79	2,77	2,73	2,64	2,56
11	2,74	2,72	2,70	2,69	2,67	2,66	2,65	2,60	2,51	2,43
12	2,64	2,62	2,60	2,58	2,57	2,56	2,54	2,50	2,40	2,32
13	2,55	2,53	2,51	2,50	2,48	2,47	2,46	2,41	2,31	2,23
14	2,48	2,46	2,44	2,43	2,41	2,40	2,39	2,34	2,24	2,16
15	2,42	2,40	2,38	2,37	2,35	2,34	2,33	2,28	2,18	2,10
16	2,37	2,35	2,33	2,32	2,30	2,29	2,28	2,23	2,12	2,04
17	2,33	2,31	2,29	2,27	2,26	2,24	2,23	2,18	2,08	1,99
18	2,29	2,27	2,25	2,23	2,22	2,20	2,19	2,14	2,04	1,95
19	2,26	2,23	2,21	2,20	2,18	2,17	2,16	2,11	2,00	1,91
20	2,22	2,20	2,18	2,17	2,15	2,14	2,12	2,07	1,97	1,88
25	2,11	2,09	2,07	2,05	2,04	2,02	2,01	1,96	1,84	1,75
50	1,89	1,87	1,85	1,83	1,81	1,80	1,78	1,73	1,60	1,48
200	1,74	1,72	1,69	1,67	1,66	1,64	1,62	1,56	1,41	1,26

O número mínimo de medições é obtido tanto pela comparação de λ , como de ρ , $\Phi_{I(T)}$ ou $\Phi_{I(OIT)}$.

5.4 Testes de avaliação

5.4.1 Comparação da precisão

Na avaliação da precisão é importante mostrar que a precisão do método B é pelo menos melhor do que a do método A. Assim, se propõe um teste F unilateral em vez do bilateral proposto pela ISO. Deseja-se testar, então:

$$\begin{cases} H_0 : \sigma_B^2 \leq \sigma_A^2 \\ H_A : \sigma_B^2 > \sigma_A^2 \end{cases}$$

5.4.1.1 Comparação da repetitividade

Para comparar a repetitividade dos dois métodos, a estatística F_r é calculada como se segue:

$$F_r = \frac{S_{rB}^2}{S_{rA}^2} \quad (5.18)$$

O valor de F_r é comparado com $F_{5\% (v_{rB}; v_{rA})}$, com $v_{rB} = (n_B - 1) \sum_{i=1}^{a_B} \sum_{j=1}^{b_B} c_{ijB}$ e

$$v_{rA} = (n_A - 1) \sum_{i=1}^{a_A} \sum_{j=1}^{b_A} c_{ijA}$$

Se $F_r > F_{5\% (v_{rB}; v_{rA})}$, o método alternativo B tem repetitividade pior do que a do método de referência A, ao nível de significância de $\alpha=5\%$ e, conseqüentemente, sua repetitividade não aceitável.

Se $F_r \leq F_{5\% (v_{rB}; v_{rA})}$, o método alternativo B tem repetitividade aceitável e diferindo por um fator ρ de repetitividade do método..

5.4.1.2 Precisão intermediária devido a diferentes tempos

Para testar a precisão intermediária devida a diferentes dias de análise, é necessário testar, inicialmente, a hipótese $\sigma_{rA}^2 = \sigma_{rB}^2$, através de:

$$F_r = \frac{S_1^2}{S_2^2} \quad (5.19)$$

Onde $S^2_{1} = \max(S^2_{rA}, S^2_{rB})$.

O valor de F_r é comparado com $F_{2,5\% (v_{r1}; v_{r1})}$, com $v_{rB} = (n_B - 1) \sum_{i=1}^{a_B} \sum_{j=1}^{b_B} c_{ijB}$

$$\text{e, } v_{rA} = (n_A - 1) \sum_{i=1}^{a_A} \sum_{j=1}^{b_A} c_{ijA} .$$

Se $F_r \leq F_{2,5\% (v_{r1}; v_{r1})}$, não existe evidência que os métodos tenham diferentes repetitividades.

Se $F_r > F_{2,5\% (v_{r1}; v_{r1})}$, os métodos têm repetitividades diferentes ao nível de significância $\alpha=5\%$. Neste caso a abordagem simplificada para comparar a precisão intermediária referente a diferentes tempos de análise não pode ser utilizada, sendo necessária a utilização da aproximação de Satherthwaite para estimar os números de graus de liberdade associados a $S^2_{I(T)}$.

Quando $n_A = n_B$ e as repetitividades dos métodos são consideradas estatisticamente iguais ($\sigma^2_{rA} = \sigma^2_{rB}$), a comparação da precisão intermediária é feita da seguinte forma:

Inicialmente, calcula-se $F_{I(T)}$ por:

$$F_{I(T)} = \frac{QMD_B}{QMD_A} = \frac{n_B S^2_{I(T)_B} - (n_B - 1) S^2_{rB}}{n_A S^2_{I(T)_A} - (n_A - 1) S^2_{rA}} \quad (5.20)$$

Em seguida, o valor de $F_{I(T)}$ é comparado com $F_{Crítico(\alpha, v_{I(T)B}; v_{I(T)A})}$, com

$$v_{I(T)B} = \sum_{i=1}^{a_B} \sum_{j=1}^{b_B} (c_{ijB} - 1) \text{ e } v_{I(T)A} = \sum_{i=1}^{a_A} \sum_{j=1}^{b_A} (c_{ijA} - 1) .$$

Se $F_{I(T)} > F_{\alpha (v_{I(T)B}; v_{I(T)A})}$, o método B tem maior precisão intermediária do que o método A, portanto sua precisão intermediária não é aceitável.

Se $F_{I(T)} \leq F_{\alpha (v_{I(T)B}; v_{I(T)A})}$, a precisão intermediária do método B é aceitável, diferenciando da do método A por um fator $\Phi_{I(T)}$.

Quando $n_A \neq n_B$ e as repetitividades dos métodos não são consideradas estatisticamente iguais ($\sigma^2_{rA} \neq \sigma^2_{rB}$), a comparação da precisão intermediária é feita como se segue:

Inicialmente calcula-se:

$$F_{I(T)} = \frac{S^2_{I(T)_B}}{S^2_{I(T)_A}} \quad (5.21)$$

$$v_{I(T)B} = \frac{(S^2_{I(T)B})^2}{(QM_{DB} / n_B)^2 / \sum_{i=1}^{aB} \sum_{j=1}^{bB} (c_{ijB} - 1)^2 + ((n_B - 1)QM_{rB} / n_B)^2 / (n_B - 1) \sum_{i=1}^{aB} \sum_{j=1}^{bB} c_{ijB}}$$

(5.22)

$$v_{I(T)A} = \frac{(S^2_{I(T)A})^2}{(QM_{DA} / n_A)^2 / \sum_{i=1}^{aA} \sum_{j=1}^{bA} (c_{ijA} - 1)^2 + ((n_A - 1)QM_{rA} / n_A)^2 / (n_A - 1) \sum_{i=1}^{aA} \sum_{j=1}^{bA} c_{ijA}}$$

(5.23)

A comparação de $F_{I(T)}$ com $F_{\alpha(v_{I(T)B}; v_{I(T)A})}$, com os graus de liberdade obtidos de (5.22) e (5.23), se faz da mesma forma como quando se utiliza a equação (5.20).

5.4.1.3

Precisão intermediária devida a diferentes operadores, instrumentos e tempos

O teste de precisão intermediária devida a diferentes operadores, instrumentos e tempo é análogo ao teste de precisão intermediária devida ao tempo. Se o número de replicatas por método for igual, $n_A = n_B$ e se $\sigma^2_{rA} = \sigma^2_{rB}$, $F_{I(OIT)}$ é calculado como se segue:

$$F_{I(OIT)} = \frac{QMD_B}{QMD_A} = \frac{n_B S^2_{I(OIT)_B} - (n_B - 1)S^2_{rB}}{n_A S^2_{I(OIT)_A} - (n_A - 1)S^2_{rA}} \quad (5.24)$$

O valor de $F_{I(T)}$ é comparado com $F_{5\%, v_{I(OIT)B}; v_{I(OIT)A}}$, com

$$v_{I(OIT)B} = \sum_{i=1}^{aB} \sum_{j=1}^{bB} c_{ijB} - 1 \text{ e } v_{I(OIT)A} = \sum_{i=1}^{aA} \sum_{j=1}^{bA} c_{ijA} - 1.$$

Quando $n_A \neq n_B$ e a repetitividade dos métodos não são consideradas estatisticamente iguais ($\sigma^2_{rA} \neq \sigma^2_{rB}$), a comparação da precisão intermediária é calculada como se segue:

$$F_{I(OIT)} = \frac{S^2_{I(OIT)_B}}{S^2_{I(OIT)_A}} \quad (5.25)$$

$$v_{I(OIT)B} = \frac{(S^2_{I(OIT)B})^2}{(QM_{OITB} / n_B)^2 / \sum_{i=1}^{aB} \sum_{j=1}^{bB} (c_{ijB} - 1) + ((n_B - 1)QM_{rB} / n_B)^2 / (n_B - 1) \sum_{i=1}^{aB} \sum_{j=1}^{bB} c_{ijB}}$$

(5.26)

$$v_{I(OIT)A} = \frac{(S^2_{I(OIT)A})^2}{(QM_{OITA} / n_A)^2 / \sum_{i=1}^{aA} \sum_{j=1}^{bA} (c_{ijA} - 1) + ((n_A - 1)QM_{rA} / n_A)^2 / (n_A - 1) \sum_{i=1}^{aA} \sum_{j=1}^{bA} c_{ijA}}$$

(5.27)

A comparação de $F_{I(OIT)}$ com $F_{5\%, v_{I(OIT)B}; v_{I(OIT)A}}$, com os graus de liberdade obtidos de (5.26) e (5.27) se faz da mesma forma que quando se utiliza a equação (5.20).

5.5

Avaliação da tendência

A avaliação da tendência é feita através da comparação das médias gerais dos métodos. A comparação no caso de métodos validados é feita usando o teste z_α . Isso pressupõe que as estimativas dos desvios-padrão usadas na comparação são obtidas de grandes amostras e, conseqüentemente, devem ser boas estimativas dos verdadeiros desvios. No entanto, como na situação intralaboratorial os desvios-padrão são relativos a pequenas amostras, o teste t de *Student* é mais apropriado (EURACHEM, 1998). Deseja-se testar então:

$$\begin{cases} H_0 : \mu_A = \mu_B \\ H_A : \mu_A \neq \mu_B \end{cases}$$

Logo,

$$t = \frac{|\bar{\bar{y}}_A - \bar{\bar{y}}_B|}{S_d} \quad (5.28)$$

onde S_d representa a estimativa do desvio-padrão das diferenças entre as médias gerais $\bar{\bar{y}}_A$ e $\bar{\bar{y}}_B$ de cada método.

Para comparar as médias gerais, combinam-se as variâncias amostrais em uma única estimativa conjunta, que tem o número de graus de liberdade maior, de

modo que o intervalo de confiança fica mais estreito e o teste fica mais sensível, tornando-se capaz de detectar diferenças sistemáticas menores.

Como só é concebível combinar variâncias amostrais se forem estimativas da mesma variância populacional, testa-se as seguintes hipóteses,

$$\begin{cases} H_0 : \sigma^2_{(\bar{y}_{ijk})_A} = \sigma^2_{(\bar{y}_{ijk})_B} \\ H_A : \sigma^2_{(\bar{y}_{ijk})_A} \neq \sigma^2_{(\bar{y}_{ijk})_B} \end{cases}$$

com os graus associados às variâncias das médias do método A e do método B, respectivamente, sendo

$$v_A = \sum_{i=1}^{aA} \sum_{j=1}^{bA} c_{ijA} - 1 \text{ e } v_B = \sum_{i=1}^{aB} \sum_{j=1}^{bB} c_{ijB} - 1 \quad (5.29)$$

Se a diferença entre as médias não for significativa, S_d é calculado por:

$$S_d = \sqrt{S_p^2 \left(\frac{1}{\sum_{i=1}^{aA} \sum_{j=1}^{bA} c_{ijA}} + \frac{1}{\sum_{i=1}^{aB} \sum_{j=1}^{bB} c_{ijB}} \right)} \quad (5.30)$$

Onde,

$$S_p^2 = \frac{(\sum_{i=1}^{aA} \sum_{j=1}^{bA} c_{ijA} - 1)S_{(\bar{y}_{ijk})A}^2 + (\sum_{i=1}^{aB} \sum_{j=1}^{bB} c_{ijB} - 1)S_{(\bar{y}_{ijk})B}^2}{v_d} \quad (5.31)$$

Com,

$$v_d = \sum_{i=1}^{aA} \sum_{j=1}^{bA} c_{ijA} + \sum_{i=1}^{aB} \sum_{j=1}^{bB} c_{ijB} - 2 \quad (5.32)$$

Quando a diferença entre as médias das variâncias é significativa, S_d é calculado como se segue:

$$S_d^2 = \sqrt{\frac{S_{(\bar{y}_{ijk})A}^2}{\sum_{i=1}^{aA} \sum_{j=1}^{bA} c_{ijA}} + \frac{S_{(\bar{y}_{ijk})B}^2}{\sum_{i=1}^{aB} \sum_{j=1}^{bB} c_{ijB}}} \quad (5.33)$$

Como número de graus de liberdade dado pela expressão:

$$v_d = \frac{(S_d^2)^2}{(S_{(\bar{y}_{ijk})A}^2 / \sum_{i=1}^{aA} \sum_{j=1}^{bA} c_{ijA})^2 / (\sum_{i=1}^{aA} \sum_{j=1}^{bA} c_{ijA} - 1) + (S_{(\bar{y}_{ijk})B}^2 / \sum_{i=1}^{aB} \sum_{j=1}^{bB} c_{ijB})^2 / (\sum_{i=1}^{aB} \sum_{j=1}^{bB} c_{ijB} - 1)} \quad (5.34)$$

Se $t > t_{\alpha/2}(v_d)$, a diferença entre as médias obtidas pelos dois métodos é estatisticamente significativa, embora a magnitude da diferença possa não ser relevante para a aplicação.

Se a estatística teste da diferença entre as médias dos métodos for maior que 2, a diferença entre as médias é estatisticamente significativa. Neste caso, para evitar a rejeição do método com uma tendência aceitável, deve-se examinar se a estimativa da tendência d pode ser considerável aceitável. A ISO *apud* Kuttatharmmakul *et al*, 1999, menciona que embora a tendência de um método seja significativa, esta é considerável aceitável se seu valor absoluto não for maior que $\lambda/2$.

No entanto, essa abordagem é controversa. Kuttatharmmakul *et al*, 1999, analisam em detalhe as diversas situações possíveis em termos de valores da estimativa d obtida (igual, maior ou menor $\lambda/2$) e da estimativa S_d do desvio-padrão σ_d (S_d igual, maior ou menor que σ_d), mostrando que essa abordagem pode levar facilmente a conclusões errôneas (rejeitar métodos com tendência aceitável e vice-versa, com riscos bem maiores que os valores α e β especificados).

Por isso, preconizam uma abordagem alternativa, que consiste em comparar um limite superior de 95%, LSC, para d com o valor de λ , para decidir sobre a aceitabilidade do método. Especificamente:

$$LSC = |\bar{\bar{y}}_A - \bar{\bar{y}}_B| + t_{(\alpha; v_d)} S_d \quad (5.35)$$

Se $LSC > \lambda$, a tendência, embora estatisticamente significativa, é aceitável, caso contrário a tendência do método é considerada inaceitável.

Outra abordagem para avaliação da tendência e precisão de um método analítico, proposta pelo Comitê de Métodos Analíticos, consiste em desdobrar os erros do processo de medição em diversos componentes.

O erro aleatório pode ser estimado a partir das dispersões das repetições. O mesmo não acontece com o erro total do laboratório, que é obtido pela tendência do método e do laboratório, sendo, portanto, necessários testes interlaboratoriais para sua estimação.

O conceito básico é ilustrado na Figura 5.3. O laboratório é visto de um nível mais alto, isto é, como membro da população de laboratórios.

Conseqüentemente, os erros sistemáticos e aleatórios de cada laboratório individual são considerados erros aleatórios pelo nível mais alto, levando em conta fatores de medição que são impossíveis de conhecer.

As principais desvantagens dessa abordagem devem-se ao fato de testes colaborativos feitos anteriormente não refletirem as condições atuais do laboratório e à dificuldade em ajustar as informações interlaboratoriais para vários mensurandos ou matrizes diversificadas em análise.

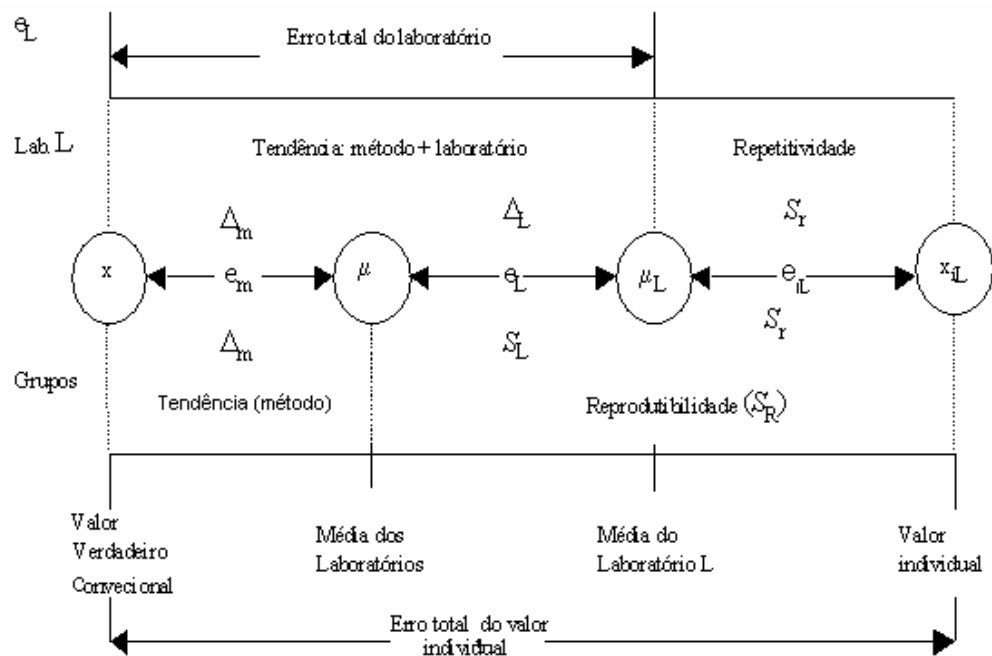


Figura 5.3. Método de partição, erro interlaboratorial e intralaboratorial (adaptado de Currie, 1999)