

ESTATÍSTICA APLICADA À QUÍMICA: DEZ DÚVIDAS COMUNS

Livia Maria Zambrozi Garcia Passari, Patricia Kaori Soares e Roy Edward Bruns*

Instituto de Química, Universidade Estadual de Campinas, CP 6154, 13083-970 Campinas - SP, Brasil

Ieda Spacino Scarminio

Departamento de Química, Universidade Estadual de Londrina, CP 6001, 86051-990 Londrina - PR, Brasil

Recebido em 28/5/10; aceito em 8/11/10; publicado na web em 18/2/11

STATISTICS APPLIED TO CHEMISTRY: TEN COMMON DOUBTS. Ten common doubts of chemistry students and professionals about their statistical applications are discussed. The use of the $N-1$ denominator instead of N is described for the standard deviation. The statistical meaning of the denominators of the root mean square error of calibration (RMSEC) and root mean square error of validation (RMSEV) are given for researchers using multivariate calibration methods. The reason why scientists and engineers use the average instead of the median is explained. Several problematic aspects about regression and correlation are treated. The popular use of triplicate experiments in teaching and research laboratories is seen to have its origin in statistical confidence intervals. Nonparametric statistics and bootstrapping methods round out the discussion.

Keywords: linear regression; median; non-parametric statistics.

INTRODUÇÃO

As aplicações da estatística se desenvolveram de tal forma que praticamente todas as áreas de pesquisa e produção se beneficiam da utilização de seus métodos. Frequentemente estudantes e pesquisadores, que aplicam métodos estatísticos como ferramenta de análise dos dados, encontram dificuldades para compreender e interpretar alguns conceitos estatísticos importantes.

O objetivo deste trabalho é reunir em um só texto, dez dúvidas e respostas que são frequentes entre químicos quando métodos estatísticos são aplicados aos seus dados. Essas informações, embora possam ser encontradas, estão dispersas em várias publicações fora da área de Química e muitas das quais num vocabulário pouco familiar aos químicos.

1. Porque o denominador do desvio padrão amostral é $N-1$?

Cursos de estatística normalmente explicam o denominador $(N-1)^{1/2}$ na equação do desvio padrão em termos de graus de liberdade,

$$s_x = \left\{ \sum (x_i - \bar{x})^2 / N - 1 \right\}^{1/2} \quad (1)$$

onde, s_x é o desvio padrão de x , x_i é a i -ésima observação, N é o número de observações e \bar{x} é a média das observações, definida como o somatório de todas as observações dividido pelo número total de observações.

Para entender a razão do denominador não ser o número total de observações, N , e sim $N-1$ imagine 5 amostras com os seguintes teores de ferro: 70,2; 71,0; 70,8; 73,5 e 70,6%. Normalmente, a estimativa da quantidade de ferro nessas amostras é a média, 71,22%. As amostras apresentam cinco resultados de porcentagem de ferro que não podem ser preditos antes da realização das análises, isto é, elas possuem cinco valores não conhecidos ou graus de liberdade para serem especificados. Mesmo conhecendo a porcentagem de ferro

em 5 amostras analisadas, não é possível prever a porcentagem da próxima amostra que será analisada.

Com a soma dos valores de todos os desvios a situação é diferente, porque este somatório é igual a zero:

$$\sum_i d_i = \sum_i (x_i - \bar{x}) = \sum_i x_i - \sum_i \bar{x} = \sum_i x_i - N\bar{x} = \sum_i x_i - \sum_i x_i = 0 \quad (2)$$

O termo $N\bar{x}$ foi substituído pelo somatório $\sum_i x_i$, uma vez que $\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i$.

Sendo assim, caso se conheça o valor da média das 5 amostras será necessário executar somente 4 análises para saber a porcentagem de ferro da quinta amostra, ou seja, temos agora 4 graus de liberdade. A restrição imposta pela Equação 2, que vem do cálculo da média, retira um grau de liberdade do conjunto de desvios. Considerando que dos N desvios só $N-1$ podem flutuar aleatoriamente, é natural que o denominador na definição da variância amostral seja $N-1$ e não N .

É importante notar que as considerações feitas acima não constituem uma prova que a Equação 1 seja uma estimativa sem tendências no desvio padrão. A prova matemática disto pode ser encontrada no livro de Montgomery.¹ Pela mesma razão, parâmetros normalmente utilizados em estatística multivariada, seja para calibração multivariada ou redes neurais, como a raiz quadrada do erro médio quadrático de calibração, tem um denominador $(N-p)^{1/2}$ e não N

$$RMSEC = \left[\frac{\sum (y_i - \hat{y}_i)^2}{N - p} \right]^{1/2} \quad (3)$$

onde, y_i é a i -ésima observação, \hat{y}_i é o i -ésimo valor previsto pelo modelo, N é o número de observações e p é o número de parâmetros do modelo de calibração.

É necessário lembrar que essa fórmula é a mesma para a raiz

*e-mail: bruns@iqm.unicamp.br

quadrada do erro médio quadrático de validação, Equação 4, salvo que neste caso o denominador é $N^{1/2}$ porque os valores de y_i não foram utilizados para determinar o modelo de calibração. Nesta equação N refere-se ao número de amostras de validação.

$$RMSEV = \left[\frac{\sum (y_i - \hat{y}_i)^2}{N} \right]^{1/2} \quad (4)$$

2. Porque utilizar a média e não a mediana?

Considerando as análises dos teores de ferro descritas no item 1, observa-se que o valor de 73,5% é consideravelmente maior do que os resultados das outras determinações. Se forem feitos testes usando os métodos de Dixon ou de Grubbs,² entre muitos outros, o valor de 73,5% seria considerado um *outlier* em relação aos outros resultados determinados para o teor de ferro.

A mediana é obtida listando-se os N valores em ordem de magnitude e selecionando o valor do centro se N for ímpar ou a média dos dois valores centrais se N for par. Sendo assim, a utilização do valor mediano é menos sensível à inclusão ou não de valores extremos como, por exemplo, o valor da mediana incluindo o teor de ferro 73,5% no cálculo é 70,6%, excluindo este valor a mediana é 70,4%. Para o valor médio utilizando os cinco valores dos teores de ferro a média é 71,22%, enquanto que rejeitando o valor suspeito a média é 70,65%, um valor bem diferente.

Podemos concluir que o valor mediano é bem mais robusto do que o valor da média para a presença de valores suspeitos no conjunto de dados. Porém, a média é preferencialmente utilizada no lugar da mediana porque existe uma equação simples para calcular o erro no valor médio, $S_{\bar{x}}$, Equação 5,

$$s_{\bar{x}} = \left(\frac{s_x^2}{N} \right)^{1/2} \quad (5)$$

onde, N é o número de observações e s_x é o desvio padrão de x , apresentado na Equação 1. Não existe uma equação, mesmo complexa, que calcule o erro no valor mediano, lembrando que o erro da média é sempre menor do que de uma medida individual.

3. Por que em métodos de regressão, a melhor reta é aquela que minimiza as soma dos quadrados das distâncias verticais entre os pontos e a reta?

A Figura 1 mostra um gráfico que ilustra o princípio do ajuste por mínimos quadrados normalmente empregado em análises de regressão.

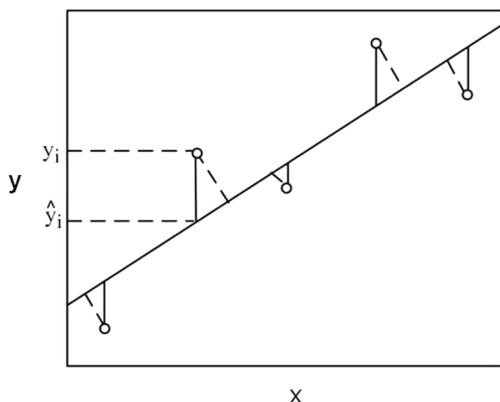


Figura 1. Reta ajustada por mínimos quadrados

Nesse método a melhor localização da reta especificada pelos coeficientes b_0 e b_1 da equação $\hat{y}_i = b_0 + b_1 X$ é aquela que minimiza a soma dos quadrados dos comprimentos dos segmentos verticais, indicado pelas linhas sólidas, que ligam os pontos experimentais à reta. Este critério implica que os valores de X sejam fixos, definidos pelo pesquisador no planejamento estatístico, e os valores de y_i as variáveis aleatórias afetadas por erros experimentais.

Se ambas as variáveis X e y forem afetadas por erros experimentais da mesma grandeza, a minimização da soma quadrática das distâncias ortogonais, representadas pelas linhas tracejadas, seria a mais apropriada.² No entanto na prática isto não ocorre, pois os químicos usam padrões na construção da curva e os erros gerados pelo procedimento de mensuração são muitos maiores do que o preparo dos padrões.

O aluno normalmente se pergunta por que não minimizar a soma das distâncias verticais em lugar da soma dos quadrados. A resposta é que minimizando as distâncias verticais, teríamos um número infinito de retas que satisfariam a condição $\sum (y_i - \hat{y}_i) = 0$, pois desvios positivos da reta cancelariam os desvios negativos. Este cancelamento poderia ser evitado, minimizando-se a soma dos valores absolutos das distâncias verticais. Na prática isto não é feito porque não existem derivadas para as Equações 6 e 7, necessárias para gerar as equações lineares que determinam os valores de b_0 e b_1 .

$$\frac{\partial \sum |y_i - \hat{y}_i|}{\partial b_0} = 0 \quad (6)$$

$$\frac{\partial \sum |y_i - \hat{y}_i|}{\partial b_1} = 0 \quad (7)$$

4. É possível determinar um modelo quadrático usando regressão linear?

O modelo quadrático é frequentemente representado pela seguinte equação genérica:

$$\hat{y} = b_0 + b_1 X + b_2 X^2 \quad (8)$$

Os primeiros dois termos do lado direito representam o modelo linear muito usado pelos químicos para calibração de métodos analíticos, determinação de quantidades físico-químicas bem como para relacionar medidas empíricas. Uma vez que a aproximação linear tem sua validade limitada, o terceiro termo ($b_2 X^2$) pode ser adicionado ao modelo, permitindo melhor ajuste dos dados.

O termo “regressão linear” usado pelos estatísticos corresponde aos parâmetros b , ou seja, às incógnitas no modelo. As variáveis independentes X são valores fixados pelo experimentador em diferentes níveis como, por exemplo, as concentrações para a curva de calibração ou temperaturas para determinar as mudanças na pressão de vapor, enquanto que as respostas, valores de y , são medidas experimentais.

Um exemplo de modelo não linear nos parâmetros estatísticos é $\hat{y} = b_0 (1 - e^{-b_1 X})$, embora esse modelo possa ser linearizado resolvendo o logaritmo.

5. Como um cientista ou engenheiro pode testar se uma curva de calibração é realmente uma reta?

Muitos modelos de calibração são baseados na suposição de que a relação entre a propriedade medida (sinal analítico, logaritmo da pressão de vapor, etc) e o nível do fator controlado pelo experimentador (concentração, inverso da temperatura, etc) é linear. Mas será que isto é sempre verdade? Como obter evidência objetiva de que a relação entre a propriedade medida e o fator controlado pelo pesquisador é realmente linear?

A resposta para essas perguntas poderá ser obtida apenas se os experimentos forem executados em replicatas, pois só assim os resultados fornecerão uma estimativa do erro experimental da propriedade que está sendo investigada. Se este erro for da mesma grandeza das diferenças entre os valores experimentais e aqueles previstos pelo modelo, podemos afirmar que a suposição sobre a linearidade está correta. Este procedimento é equivalente a fazer experimentos confirmatórios para testar um modelo. Mesmo assim é preciso fazer réplicas para determinar se os resultados confirmatórios estão dentro dos limites do erro experimental.

Na Tabela 1 encontram-se os valores da pressão de vapor de tetracloreto de carbono (CCl_4) para diferentes valores de temperatura (T) obtidos no laboratório de físico-química.³

Tabela 1. Variação da pressão de vapor do CCl_4 com a temperatura

| Ensaio | T (K) | p_{vap} (torr) |
|--------|-------|-------------------------|
| 1 | 273 | 0,044 |
| 2 | 283 | 0,075 |
| 3 | 293 | 0,122 |
| 4 | 303 | 0,190 |
| 5 | 313 | 0,288 |
| 6 | 323 | 0,422 |
| 7 | 333 | 0,601 |
| 8 | 343 | 0,829 |
| 9 | 353 | 1,124 |

Se a entalpia de vaporização for constante e não depender da temperatura e a Equação de Clausius-Clapeyron for validada nessas condições, o gráfico de $\ln p_{\text{vap}}$ vs. $(1/T)$ será uma reta.

Uma regressão linear usando $\ln p_{\text{vap}}$ como variável dependente e $(1/T)$ como a variável independente resultará na equação:

$$\ln p_{\text{vap}} = 11,20 - 3901,98 \left(\frac{1}{T} \right), \text{ com } R^2 = 0,9997.$$

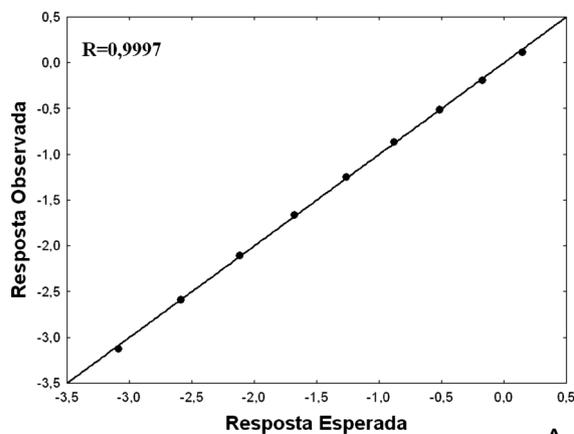
Este resultado corresponde a um calor de vaporização de $32,44 \pm 0,22 \text{ KJ mol}^{-1}$.

A Figura 2 contém o gráfico dos valores esperados pelos observados, juntamente com o gráfico dos resíduos³ deixados pelo ajuste versus os valores esperados pelo modelo linear. Apesar da excelente concordância entre os valores observados e esperados e o alto valor de R^2 , Figura 2a, há necessidade de incluir um termo quadrático no modelo apresentado, pois o gráfico dos resíduos, Figura 2b, deixa claro que eles não estão distribuídos aleatoriamente. Supondo que não existam erros sistemáticos nos resultados e que a execução dos experimentos foi feita em ordem aleatória, os resíduos não podem ser explicados como sendo devidos ao erro experimental. Concluímos que o modelo linear é falho para representar os dados da Tabela 1.

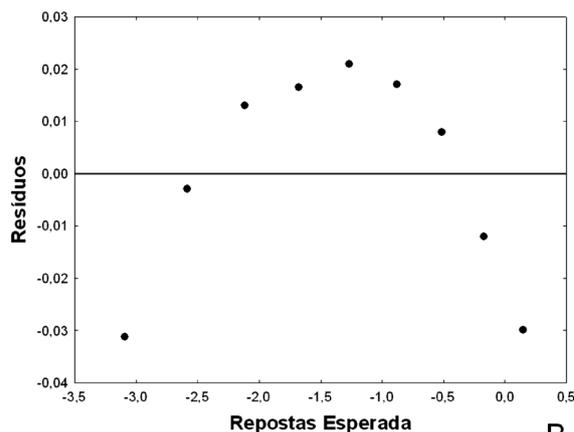
Fazendo um novo ajuste aos dados e adicionando um termo quadrático temos uma nova equação: $\ln p_{\text{vap}} = 8,00 - 1954 \left(\frac{1}{T} \right) - 300311 \left(\frac{1}{T} \right)^2$, com $R^2 = 1,0000$.

A Figura 3 mostra o gráfico dos resíduos deixados pelo ajuste do modelo quadrático. Nessa figura os resíduos estão distribuídos bem mais aleatoriamente ao redor da linha no valor zero quando comparados à distribuição dos resíduos na Figura 2b e, portanto, não há evidências fortes para suspeitar que exista falta de ajuste do modelo quadrático ajustado aos dados da Tabela 1.

Usando, portanto, o modelo quadrático e admitindo que a derivada $\frac{\partial(\ln p_{\text{vap}})}{\partial \left(\frac{1}{T} \right)} = -1,945 - 600,622 \left(\frac{1}{T} \right)$ é uma estimativa mais realista de



A



B

Figura 2. Gráfico (a) dos valores esperados pelos valores observados e (b) dos resíduos deixados pelo ajuste do modelo linear

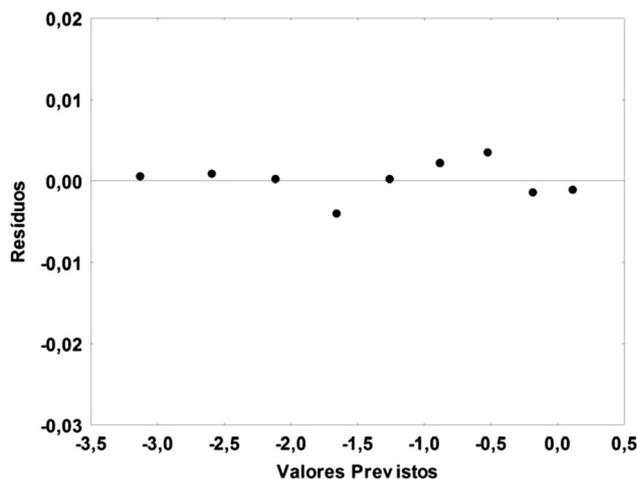


Figura 3. Resíduos do modelo quadrático ajustado aos dados da Tabela 1

$\frac{-\Delta H_{\text{vap}}}{R}$, podemos concluir que o calor de vaporização do CCl_4 na verdade varia entre 30,39 e 34,54 KJ mol^{-1} , no intervalo de temperatura considerado.

Atualmente, a verificação de linearidade da reta de calibração em química analítica é frequentemente analisada usando o coeficiente de correlação fornecido pelo programa computacional empregado para fazer a regressão. Infelizmente, não existe um teste estatístico que possa ser aplicado a este coeficiente para comprovar a linearidade da

reta num dado nível de confiança. Cada vez mais está sendo exigida a utilização do critério da falta de ajuste recomendado por Pimentel e Barros Neto⁴ e Danzar e Currie.⁵

6. Quando o coeficiente de correlação é zero, significa que não existe relação entre as duas variáveis?

Não. O coeficiente de correlação é limitado para investigar relações lineares entre as variáveis. Imagine, por exemplo, a relação entre a energia potencial de uma ligação química e sua distorção na geometria do equilíbrio. Os dados estão representados graficamente pelos pontos de uma curva anarmônica na Figura 4.

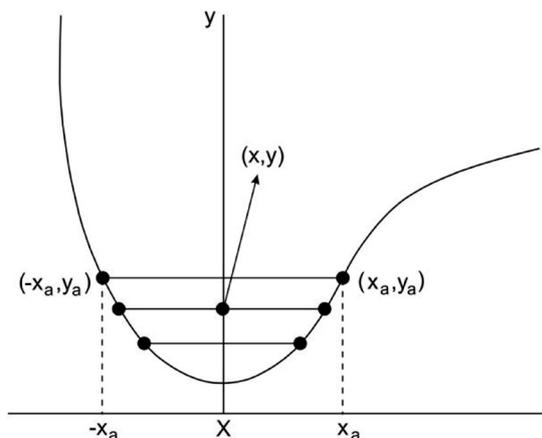


Figura 4. Relação entre a energia potencial de uma ligação química e sua distorção na geometria do equilíbrio

O coeficiente de correlação linear é dado pela equação:

$$r(x, y) = \frac{1}{N-1} \sum \left(\frac{x_i - \bar{x}}{s_x} \right) \left(\frac{y_i - \bar{y}}{s_y} \right) \quad (9)$$

onde N é o número de pontos na Figura 4 e s_x e s_y são os desvios padrão das variáveis x e y .

Aplicando a equação da correlação para os pontos deste gráfico, observa-se que as contribuições dos dois pontos, (x_a, y_a) e $(-x_a, y_a)$, na equação irão praticamente se cancelar. Este argumento é válido para qualquer par de pontos que tenham o mesmo valor de y_i . De fato, se os pontos seguissem exatamente uma parábola, a aplicação da Equação 9 resultaria em um coeficiente de correlação igual a zero. Porém, como as ligações químicas não obedecem à risca a lei de Hooke, e por isso não vibram igual a um oscilador harmônico, o coeficiente de correlação dos dados representados na Figura 4 não será exatamente zero e sim, um coeficiente de correlação pequeno que não reflete o comportamento sistemático dos dados apresentados no gráfico.

7. Porque é importante examinar os gráficos ao invés de apenas calcular os parâmetros estatísticos?

Parâmetros estatísticos como a média, o desvio padrão e o coeficiente de correlação são representações numéricas de grande quantidade de dados. O coeficiente de correlação r , por exemplo, é um número que expressa a relação entre duas variáveis, obtido pela Equação 9. Na realidade, as relações entre variáveis são normalmente muito complexas para serem representadas por um único número.

A Figura 5 mostra um gráfico de três conjuntos de pares de valores das variáveis y plotada em função das variáveis x . Os pontos representados por bolas cheias mostram uma relação evidentemente linear entre as variáveis x e y , porém, com grandes dispersões do

modelo de regressão representado pela reta. Os pontos representados por bolas vazias apresentam uma relação não linear. Observa-se que os quatro pontos na esquerda têm um arranjo linear, diferente dos pontos na direita que também apresentam uma distribuição quase linear, porém com uma inclinação diferente dos pontos da esquerda. Por fim, os pontos representados por triângulos apresentam um arranjo de pontos completamente diferente, o ponto na direita tem um valor bem distinto dos pontos na esquerda.

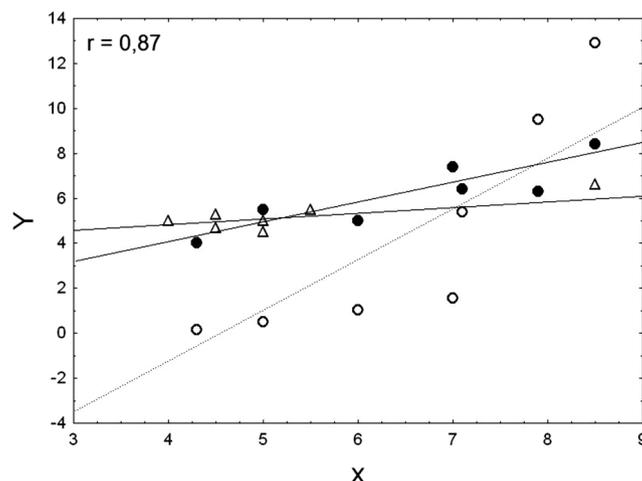


Figura 5. Três conjuntos de dados com o mesmo coeficiente de correlação, $r = 0,87$, mas representando realidades muito diferentes

Estes conjuntos de dados possuem algo em comum. Todos apresentam um coeficiente de correlação de Pearson igual a 0,87, porém, as situações físicas apresentadas nos gráficos são completamente diferentes. Os pontos representados por bolas cheias mostram uma simples relação linear, as bolas vazias uma interseção de dois modelos lineares, enquanto que os representados por triângulos mostram a possível existência de um *outlier*, ou seja, um ponto fora do padrão. Se este ponto for retirado do conjunto de dados o coeficiente de correlação dos pontos na esquerda cai para 0,27.

8. Por que nas aulas de laboratório de química as determinações são feitas em triplicata?

A realização de experimentos em triplicata é recomendada nos laboratórios de ensino porque é um compromisso aceitável entre a precisão e o trabalho. O valor médio da triplicata é a melhor estimativa do teor do analito na amostra, enquanto que o desvio padrão é a estimativa do erro experimental em uma determinação, sendo que o erro padrão no valor médio da triplicata é menor pelo fator de $1/\sqrt{3}$. Isto pode ser visto na equação que representa o intervalo de confiança do valor médio:

$$\mu = \bar{x} \pm t_{N-1} \frac{s}{\sqrt{N}} \quad (10)$$

onde, \bar{x} representa a média, s corresponde ao desvio padrão, N é o número de réplicas (3 no caso da triplicata) e t é o valor crítico da distribuição t de Student com $N-1$ graus de liberdade.

Aumentando o valor de N , o intervalo de confiança irá diminuir por causa da diminuição do valor t_{N-1} e do fator $1/\sqrt{N}$. No nível de 95% de confiança, os valores de t_{N-1} são 12,71; 4,30; 3,18; 2,78 e 2,57 quando N vai de 2 para 6. Para o mesmo intervalo de N , a expressão $1/\sqrt{N}$ vai de 1,000; 0,707; 0,577; 0,500 até 0,408. Os produtos de t_{N-1}/\sqrt{N} ficam progressivamente menores, indo de 12,71; 3,04; 1,83; 1,39 até 1,13. Sendo assim, quando se realiza uma triplicata ao invés de uma duplicata, ocorre um melhoramento da precisão de 12,71

para 3,04 (fator de 4). Entretanto, realizando uma quadruplicata ao invés de uma triplicata, ocorre um melhoramento da precisão por um fator menor do que 2, pois o valor de t_{N-1}/\sqrt{N} vai de 3,04 para 1,83.

A Figura 6 apresenta um gráfico de t_{N-1}/\sqrt{N} vs. N ilustrando como a precisão diminui quando se aumenta o número de réplicas. Nota-se que a precisão é muito pouco melhorada com a execução de cinco ou mais replicatas.

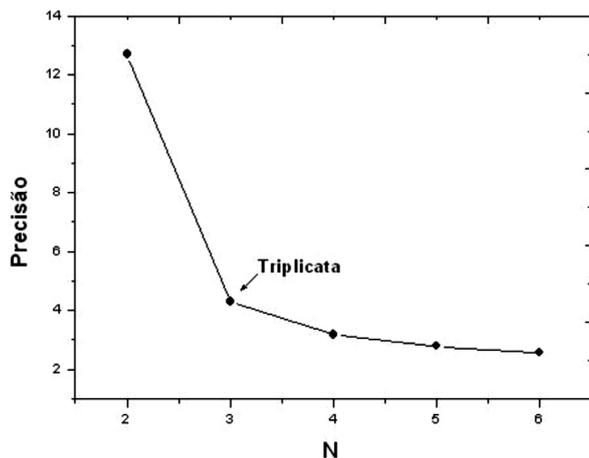


Figura 6. Gráfico de t_{N-1}/\sqrt{N} contra o número de réplicas N , ilustrando como a precisão diminui quando se aumenta o número de réplicas

9. Qual é a diferença entre a estatística paramétrica e a não paramétrica?

A grande maioria dos químicos utiliza métodos de estatística paramétrica para resolver problemas no laboratório como, por exemplo, determinar se um valor experimental é igual a um valor padrão, se os resultados provenientes de diferentes laboratórios são iguais, se o erro médio quadrático de validação de um método analítico é maior do que o erro de outro método, etc. Nestes casos, considera-se que os dados seguem uma distribuição normal, uma vez que a comparação de valores médios não é tão problemática porque estes valores tendem a seguir distribuições normais devido ao teorema do limite central,³ mesmo se os dados das determinações individuais não seguirem nenhuma distribuição conhecida. Por isto, planejamentos estatísticos de experimentos produzem resultados confiáveis. Como os valores dos parâmetros dos modelos associados a estes planejamentos são combinações lineares de resultados experimentais, eles tendem a seguir uma distribuição normal.

Os métodos de estatística não paramétrica são usados quando o pesquisador precisa testar dados que não seguem uma distribuição bem caracterizada (normal ou não). Muitos métodos básicos de estatística paramétrica têm seu análogo não paramétrico como, por exemplo, o “teste da soma dos postos de Wilcoxon” (*Wilcoxon's rank sum test*) é o equivalente não paramétrico do teste t para duas médias.⁶ O teste t pareado paramétrico corresponde ao “teste dos sinais de Wilcoxon” (*Wilcoxon's signed rank test*). O “teste de Kruskal- Wallis” (*Kruskal-Wallis test*) é o equivalente não paramétrico da análise de variância (ANOVA) para dados obtidos em ordem completamente aleatória. Para dados obtidos com aleatorização por blocos pode ser usado o método não paramétrico “teste de Friedman-R” (*Friedman-R test*). Também existe um coeficiente de correlação não paramétrico chamado “teste do coeficiente de correlação de Spearman” (*Spearman rank coefficient of correlation*) análogo ao coeficiente de correlação de Pearson.

Todos estes métodos não paramétricos utilizam os postos dos dados em lugar dos dados em si. A Tabela 2 apresenta os valores e postos dos dados x e y correspondendo aos pontos representados por bolas cheias no gráfico da Figura 5. Os postos foram determinados

colocando-se os dados em ordem crescente e atribuindo-se posto 1 para o menor valor, posto 2 o segundo menor, etc. Depois de determinar os postos para ambos os dados, x e y , calcula-se a diferença entre os postos que estão apresentados na última coluna d . O coeficiente de correlação é obtido por meio da Equação 11

$$r_s = 1 - \frac{6 \sum d^2}{N(N^2 - 1)} \quad (11)$$

onde N é o número de observações, neste caso 6 e o somatório de d^2 é igual a 4. Dessa forma temos um coeficiente de Spearman de 0,89. Nota-se que este resultado está em boa concordância com o valor do coeficiente de correlação de Pearson, 0,87, dado no item 6.

Tabela 2. Dados para calcular o coeficiente de correlação de Spearman

| Número | x | Posto | y | Posto | d |
|--------|-----|-------|-----|-------|-----|
| 1 | 4,3 | 1 | 4,0 | 1 | 0 |
| 2 | 5,0 | 2 | 5,5 | 3 | -1 |
| 3 | 6,0 | 3 | 5,0 | 2 | 1 |
| 4 | 7,0 | 4 | 7,4 | 5 | -1 |
| 5 | 7,9 | 5 | 6,3 | 4 | 1 |
| 6 | 8,5 | 6 | 8,4 | 6 | 0 |

10. Métodos que utilizam o poder do computador para fazer cálculos podem resolver problemas que não são resolvidos usando estatística clássica?

Esta questão refere-se a aplicações para as quais não existem equações analíticas para resolver o problema de interesse. Por exemplo, seria possível saber se o valor mediano de um conjunto de resultados é estatisticamente igual ao valor mediano de um grupo de controle? Usando a estatística clássica a resposta para esta questão é não, pois não existem equações analíticas que possam ser utilizadas para determinar se os dois valores medianos são realmente diferentes ou se a diferença é uma mera flutuação estatística causada pelo erro experimental. Entretanto, é possível resolver este problema empregando a metodologia do *bootstrap*⁷ que consiste em reamostrar de um conjunto de dados, diretamente ou via um modelo ajustado, a fim de criar réplicas dos dados para avaliar a variabilidade da quantidade de interesse, sem usar cálculos analíticos. Para cada grupo, usa-se um computador para criar centenas ou até milhares de amostras do mesmo tamanho por amostragem aleatória com substituição. Por exemplo, para um valor mediano de seis resultados do laboratório ($x_1, x_2, x_3, x_4, x_5, x_6$) uma amostra *bootstrap* pode ser ($x_3, x_5, x_1, x_6, x_5, x_2$). Este procedimento é feito várias vezes para cada grupo e depois é construído um histograma das diferenças dos cálculos medianos, onde podem ser determinados intervalos no nível de 95% de confiança. O mesmo histograma serve para intervalos com outros níveis de confiança.

REFERÊNCIAS

- Montgomery, D. C.; *Design and Analysis of Experiments*, 3rd ed., Wiley: New York, 1991, p. 22.
- Irvin, J. A.; Quickenden, T. I.; *J. Chem. Educ.* **1983**, *60*, 711.
- Barros Neto, B.; Bruns, R. E.; Scarminio, I. E.; *Como fazer experimentos*, 4^a ed., Artmed: Porto Alegre, 2010, p. 260-263.
- Pimentel, M. F.; Barros Neto, B. de; *Quim. Nova* **1996**, *19*, 268.
- Danzar, K.; Currie, L. A.; *Pure Appl. Chem.* **1998**, *70*, 993.
- Wagner, S. F.; *Introduction to Statistics*, Harper Collins: New York, 1992, chap. 14.
- Efron, B.; Tibshirani, R. J.; *An Introduction to the Bootstrap*, Chapman & Hall/CRC: Boca Raton, 1994.