

Algumas diretrizes estatísticas para a avaliação do erro do método na mensuração de variável quantitativa

Ary José Dias Mendes*, Ary dos Santos-Pinto**

Resumo

Introdução: o erro do método deve ser quantificado para fundamentar a validade da pesquisa e sua aplicação clínica. A determinação da confiabilidade na identificação dos pontos de referência para a obtenção de medidas constitui pré-requisito fundamental para uma interpretação segura e utilização correta dos dados obtidos. Cada medida observada pode ser analisada pelo seu valor real e pela sua componente de erro. A confiabilidade é verificada a partir de medidas repetidas. **Objetivo:** o propósito deste trabalho é descrever um método estatístico que permita que o pesquisador verifique se as mensurações que realizou num certo momento são repetidas por ele num momento posterior. Este método confirma a não ocorrência do erro sistemático e permite uma avaliação da ocorrência do erro casual, que está sempre presente em qualquer forma de mensuração. Apresenta a aplicação deste método estatístico em dados referentes a duas medidas cefalométricas e aponta uma sintaxe para resolução do erro de método a partir do uso do programa SPSS® para microcomputador.

Palavras-chave: Erro do método. Estatística. Cefalometria. Confiabilidade.

INTRODUÇÃO

A mensuração de variáveis quantitativas é um procedimento muito freqüente nos métodos de pesquisas da área da Odontologia, especialmente na Ortodontia, por se constituírem num conjunto de medidas lineares e angulares, obtidas em radiografias padronizadas, a partir de procedimentos manuais ou computadorizados, com o propósito de determinar variações e interpretar a expressão geométrica da morfologia craniofacial humana, imprescindível para o ortodontista efetuar um tratamento em seu consultório ou clínica.

Os erros inerentes à obtenção do conjunto de mensurações lineares e angulares fazem com que haja certas limitações no uso do método cefalomé-

trico^{9,11,12}. Segundo Bjork e Solow⁴, três fontes ou fatores de erro devem ser considerados: diferenças entre duas radiografias de um mesmo indivíduo; diferenças na localização de pontos cefalométricos e diferenças nas mensurações de medidas lineares ou angulares do cefalograma.

Há que se considerar um outro tipo de erro que faz menção à identificação dos pontos cefalométricos. A literatura a respeito desse tipo de erro é volumosa e muitos artigos devotados a esse tópico foram publicados^{2,4}, sendo que algumas publicações sobre investigações cefalométricas incluem uma pormenorizada análise deste^{1,6,7,9}.

Houston¹⁰ analisa a importância e o correto modo de se tratar a análise dos erros cometidos na

* Professor Adjunto de Estatística do Departamento de Ciências da Educação da Faculdade de Ciências e Letras da UNESP de Araraquara.

** Professor Adjunto de Ortodontia do Departamento de Clínica Infantil da Faculdade de Odontologia da UNESP de Araraquara.

cefalometria: a) estes são freqüentemente mal compreendidos; b) muitos trabalhos, que se supõe serem científicos, não incluem nenhuma avaliação adequada em seus métodos e c) o tópico mais importante, não somente para o investigador como também para os leitores de qualquer artigo que envolva mensurações, é a pouca atenção que se atribui ao fato de tais erros afetarem a interpretação dos resultados.

Define-se como erro de uma medida a diferença entre o valor obtido no processo de sua efetivação e o valor verdadeiro da magnitude medida. Efetuando-se várias medidas de certa grandeza sob as mesmas condições, o valor verdadeiro dessa grandeza pode ser substituído pela média dessas medidas. Assim, de um conjunto finito de medidas x de uma grandeza X , obtidas sob as mesmas condições, o erro que se comete nessas medidas x pode ser expresso como a diferença do valor médio de x , notado por $\mu(x)$, e X . Simbolicamente, tem-se:

$$\text{Erro} = \mu(x) - X \quad (1)$$

para um número finito de medidas x

Dois tipos de erro podem ocorrer em mensurações realizadas em um objeto qualquer: erro sistemático, ou constante, e erro casual, também denominado aleatório ou acidental.

Assim sendo, como para inferir-se com boa margem de segurança, a partir dos resultados obtidos em mensurações cefalométricas, torna-se necessário saber como proceder para avaliar possíveis erros cometidos nessas mensurações, objetivou-se, nesse trabalho, estudar métodos alternativos estatísticos que permitam investigar o erro do método na mensuração de variáveis quantitativas ou medidas cefalométricas ou, mais precisamente, como se obter uma reprodutibilidade ou repetição consistente para as medidas cefalométricas e caracterizar normas para o uso de técnicas estatísticas que detectem tais erros.

METODOLOGIA

Antes de apontar os modelos estatísticos alternativos aos expostos por Houston¹⁰, apropriados à

detecção dos erros inerentes a uma mensuração de uma grandeza, obtida por certo experimentador, deve-se explicar as noções dos erros sistemático e casual.

Segundo Rodrigues¹³, erro sistemático é o erro devido a fatores que, agindo da mesma maneira sobre um instrumento ou sobre um processo de mensuração ou ainda sobre o operador, afeta os resultados no mesmo modo. Na expressão (1) afetar os resultados do mesmo modo significa que o valor médio $\mu(x)$ ou superestimar ou subestimar a grandeza X .

Esse autor define também o erro aleatório como o erro em que, devido ao concurso de um grande número de fatores, a diferença entre a medida e a grandeza mensurada - expressão (1) - ora é positiva ora é negativa e que, quanto maior for o número de medidas efetuadas, a soma algébrica dessas diferenças tende mais nitidamente a anular-se e o conjunto dessas diferenças sujeita-se a uma determinada lei de distribuição de probabilidade, conhecida como distribuição Normal ou Curva de Gauss-Laplace.

A verificação da hipótese de reprodutibilidade ou repetição que o pesquisador deve apresentar em duas ocasiões distintas, isto é, a hipótese de o mesmo não apresentar erro sistemático no método de mensuração das medidas, será realizada a partir do modelo de regressão linear simples. Notando-se por X_i o valor obtido na i -ésima mensuração de certa medida cefalométrica no instante A e por Y_i o valor da i -ésima mensuração nessa mesma medida no instante B, o modelo de regressão linear simples que ajusta uma reta entre os pontos do plano cartesiano dos pares de pontos (X_i, Y_i) é expresso por:

$$Y_i = \beta_0 + \beta_1 \cdot X_i + \varepsilon_i \quad (I),$$

com: $i = 1, 2, \dots, n$ ($n =$ número de pacientes ou de radiografias); β_0 o coeficiente intercepto que a reta (I) faz com o eixo das ordenadas do sistema cartesiano; β_1 o coeficiente de inclinação que a reta (I) faz com o eixo das abcissas e ε_i é o erro.

Para se conseguir a igualdade entre as mensurações realizadas nas duas ocasiões A e B é preciso que, no modelo (I), se verifiquem as hipóteses estatísticas, a saber:

a) $H_0: \beta_0 = 0$;

b) $H_0: \beta_1 = 1$;

c) que os n valores ε_i correspondam somente ao erro aleatório, isto é, que a média deles seja nula e que se ajustem à distribuição normal.

As hipóteses relativas aos coeficientes de intercepto e de inclinação podem ser verificadas com o auxílio da estatística t-Student no nível de significância α , sendo que a regra de decisão adotada foi baseada na probabilidade $p = P(t > t_{\alpha})$.

Uma questão pode ser levantada: será que o modelo (I) ajusta-se ao conjunto de pares de pontos (X_i, Y_i) do plano cartesiano? Essa questão pode ser avaliada a partir da hipótese da não ocorrência da “falta de ajustamento”, realizada a partir da análise de variância na regressão. A verificação dessa hipótese pode ser realizada no nível de significância α na distribuição de F de Snedecor e a regra de decisão foi definida a partir da probabilidade $p = P(F > F_{\alpha})$.

Erro sistemático

Uma indagação se faz presente: será que o método de mensuração empregado evidenciou a presença do erro sistemático? Em outras palavras será que o método originou uma superestimação ou uma subestimação da verdadeira grandeza média da medida cefalométrica sob mensuração? O valor esperado do modelo (I) é dado por:

$$\mu(Y_i) = \mu(\beta_0 + \beta_1 \cdot X_i + \varepsilon_i) = \beta_0 + \beta_1 \cdot \mu(X_i) + \mu(\varepsilon_i) \quad \text{(II)}$$

Se as hipóteses relativas aos coeficientes de intercepto e de inclinação forem não rejeitadas ou admitidas e se a soma dos erros ε_i for nula, o que implica que $\mu(\varepsilon_i) = 0$, então de (II) tem-se que:

$$\mu(Y_i) = \mu(X_i) \quad \text{(III)}$$

Ou seja: a média das mensurações da medida

cefalométrica sob estudo no instante B é igual à média das mensurações dessa medida no instante A. Assim, a expressão (III) caracteriza a não ocorrência do erro sistemático no método de mensuração da medida cefalométrica, o que implica em se dizer que “a não ocorrência do erro sistemático se dará quando e somente quando: o intercepto β_0 for nulo; a inclinação β_1 for igual à unidade e a soma dos erros ε_i for nula”.

Erro casual

Será que os erros ε_i , explicitados no modelo (I), são aleatórios ou casuais? Segundo os autores Draper, Smith⁵; Hoffmann, Vieira⁸ e Rodrigues¹³ os valores ε_i poderão ser denominados erros casuais ou aleatórios se eles ajustarem-se à distribuição normal de probabilidades, com média zero e variância σ^2 .

- Média zero

A não ocorrência do erro sistemático permite afirmar que a média do erro casual é igual a zero - $\mu(\varepsilon_i) = 0$ porque $\sum \varepsilon_i = 0$.

- Variância σ^2

A variância do erro aleatório - $\sigma^2(\varepsilon_i)$ - deve ser positiva. A amplitude do erro está diretamente associada à grandeza dessa variância, que retrata a precisão das mensurações, pois quanto menor for a variância menor será a amplitude do erro aleatório e conseqüentemente maior a precisão. A questão que pode ser levantada é: o pesquisador apresenta a mesma precisão quando efetua mensurações em várias medidas? A verificação da igualdade de variâncias (ou de precisões) pode ser realizada como auxílio do teste de Bartlett na distribuição de qui-quadrado (χ^2) com $m-1$ grau de liberdade onde $m = n^\circ$ de medidas.

- Ajuste à distribuição normal de probabilidades

A verificação de que os valores do erro ε_i ajustem-se à distribuição normal de probabilidades foi realizada a partir dos testes das hipóteses: 1) “o coeficiente de assimetria é igual a zero” e 2) “o co-

eficiente de curtose é igual a três”, como auxílio da estatística t-Student. A não rejeição de ambas as hipóteses implica que os valores de ϵ_i seguem a distribuição normal.

APLICAÇÃO

Como aplicação da metodologia exposta acima, sorteou-se casualmente duas medidas, SNB e PM, utilizadas por Albuquerque Jr.³ em seu trabalho e das 150 observações mensuradas sorteou-se 10%, pelo procedimento de amostragem sistemática, sendo que a observação inicial foi a de número 3. Os dados obtidos a partir desse processo de amostragem constam na tabela 1 e, seqüencialmente, foram obtidas as tabelas 2 e 3, pelo uso do programa Statistical Package for Social Science (SPSS), pertencente ao Departamento de Clínica Infantil da Faculdade de Odontologia da UNESP de Araraquara.

A sintaxe para execução no programa SPSS foi expressa por:

```
REGRESSION
/MISSING LISTWISE
/STATISTICS COEFF OUTS R ANOVA
CHANGE
```

```
/CRITERIA=PIN(.05) POUT(.10)
/NOORIGIN
/DEPENDENT snb_b
/METHOD=ENTER snb_a
/SAVE RESID .
```

```
EXAMINE
VARIABLES=res_1
/PLOT NONE
/STATISTICS DESCRIPTIVES
/CINTERVAL 95
/MISSING LISTWISE
/NOTOTAL
```

Com os valores da tabela 1 colocados no campo Data do SPSS: 1) com o cursor à esquerda do termo REGRESSION, clica-se com o mouse no comando RUN para se obter os resultados da tabela 2 para a medida SNB; 2) com o cursor à esquerda do termo EXAMINE, clica-se como mouse o comando RUN para se obter os resultados da tabela 3 para a medida SNB. Para se obter os resultados referentes à medida PM deve-se substituir, respectivamente, snb_b por pm_b, snb_a por pm_a e res_1 por res_2 e efetuar as etapas 1) e 2) acima descritas.

Na tabela 2 para cada medida cefalométrica selecionada ao acaso verificou-se que:

- a hipótese de que o coeficiente intercepto (β_0) foi igual a zero e a de que o coeficiente de inclinação (β_1) foi igual a 1 foram não rejeitadas, porque os respectivos valores observados na estatística t de Student (t_0) foram não significantes, resultando $p > 0,05$ para cada um destes valores. Assim, em cada uma dessas medidas observou-se que os valores obtidos no instante B foram estatisticamente iguais aos obtidos no instante A, o que equivale a afirmar que o pesquisador demonstrou estar calibrado.

Tabela 1 - Valores observados das medidas (instantes A e B).

medidas	valores sorteados
SNB_A	82,34; 81,78; 70,38; 81,25; 81,77; 68,68; 81,97; 83,16; 68,85; 83,98; 79,96; 68,29; 83,95; 81,42; 69,59
SNB_B	81,96; 79,84; 69,34; 82,04; 81,27; 67,56; 82,15; 81,16; 68,85; 84,02; 81,03; 68,81; 82,73; 81,97; 68,52
PM_A	30,63; 30,64; 36,88; 28,46; 32,66; 35,92; 32,65; 29,28; 33,56; 30,00; 26,53; 32,66; 31,38; 27,29; 34,05
PM_B	32,65; 32,67; 36,93; 29,23; 33,04; 35,37; 31,01; 28,62; 32,54; 29,05; 26,09; 33,73; 32,62; 28,35; 34,25

Tabela 2 - Estimativas, erros padrão para os coeficientes β_0 e β_1 ; valores t_0 , F_0 , R^2 e p.

medida	estimativas	erro padrão	t_0	p <	F_0	p <	R^2
SNB	$\beta_0 = -0,447$	3,242	-0,138 n	0,786	580,501 s	0,0001	0,976
	$\beta_1 = 1,001$	0,042	0,024 n	0,981			
PM	$\beta_0 = 2,623$	3,235	0,811 n	0,432	81,716 s	0,0001	0,852
	$\beta_1 = 0,924$	0,102	-0,745 n	0,469			

s = valor significativo; n = valor não significativo.

Tabela 3 - Média, variância, coeficientes de assimetria e de curtose para o erro residual segundo medida.

medida	média	variância	assimetria	t_0	p <	curtose	t_0	p <
SNB	0,0000	0,929	-0,195	-0,336 n	0,742	-1,016	-0,906 n	0,380
PM	0,0000	1,200	0,024	0,041 n	0,968	-0,981	-0,875 n	0,396

n = valor não significativo.

Como ambas as hipóteses foram não rejeitadas e como a soma dos erros e do modelo (I) foi nula, segue-se que não ocorreu erro sistemático.

- a hipótese de que os pontos no plano cartesiano não se ajustaram ao modelo de regressão foi rejeitada para ambas as medidas, porque os valores observados da estatística (F_0) associados a cada uma delas foram significantes, pois obteve-se $p < 0,05$.

- o coeficiente de determinação (R^2) associado a cada uma das medidas foi relevante, pois verificou-se que o modelo explicou 97,6% da variação dos dados obtidos no instante B para a medida SNB e 85,2% para a medida PM.

Para o estudo das hipóteses que dizem respeito à identificação do erro ϵ do modelo (I) construiu-se a tabela 3, na qual pode-se constatar que:

- a média do erro ϵ foi igual a zero para cada medida selecionada ao acaso das que foram adotadas na pesquisa, e esse fato atesta que esse erro não foi sistemático;

- a hipótese de que os dados do erro ϵ se ajustem ao modelo teórico normal de probabilidade foi não rejeitada. De fato, para que os dados se ajustem à distribuição normal de probabilidades é preciso que o coeficiente de assimetria seja nulo e que o coeficiente de curtose seja igual a 3. Essas hipóteses ocorreram pois: 1) a hipótese de que o coeficiente de assimetria é igual a zero foi não rejeitada, porque o valor de t_0 , para cada medida, foi não significativo ($p > 0,05$); 2) a hipótese de que o coeficiente de curtose menos 3 é igual a zero também foi não rejeitada, para cada medida, porque o respectivo valor de t_0 foi não significativo ($p > 0,05$);

- pode-se afirmar que o erro ϵ do modelo (I), para cada medida, foi somente casual, porque apresentou-se normalmente distribuído, com média igual a zero e variância ou precisão conforme a ter-

ceira coluna da tabela 3;

- relativamente às variâncias, que revelam a precisão do erro casual cometido em cada medida, foi realizado o teste de Bartlett para se verificar a hipótese de igualdade (ou de homocedasticidade) entre elas. Esse teste apresentou o valor para o teste de Bartlett igual a 0,242, não significativo, pois $p < 0,623$ na distribuição de quiquadrado com 1 grau de liberdade. Assim, o experimentador cometeu somente o erro casual nas mensurações de cada medida com a mesma precisão, porque as variâncias associadas a cada erro casual foram estatisticamente iguais entre si.

CONSIDERAÇÕES FINAIS

Este artigo teve como objetivo descrever um método estatístico alternativo ao que é atualmente utilizado para a determinação dos erros inerentes a uma mensuração. O método proposto utilizando o programa de estatística SPSS® (SPSS Inc. Chicago, Illinois, SPSS for Windows versão 10.1) para microcomputadores favorece a determinação dos erros sistemático e aleatório, que podem estar presentes em conjuntos de dados referentes a duas medidas quantitativas contínuas, tais como valores cefalométricos obtidos em telerradiografias laterais. O método baseia-se na aplicação do modelo de regressão linear por meio de sintaxe ou conjunto de instruções inseridos no programa, de forma a permitir a verificação da hipótese de que o coeficiente de intercepto da reta de regressão é igual a zero, isto é, que esta reta passe pelo ponto de origem zero e de que o coeficiente de inclinação seja igual a 1, isto é, de que a inclinação da reta de regressão seja de 45 graus, condições que atestam a não existência de erros sistemáticos. O coeficiente de determinação associado ao

modelo permite a observação do grau de concórdância entre os dois conjuntos de medidas, atestando a precisão das mensurações.

A análise da variabilidade do erro aleatório ou casual, segundo o método proposto, permite quantificar a amplitude deste erro e conseqüentemente a precisão das mensurações realizadas. A aplicação do teste de Bartlett, para verificar a igualdade das variâncias, e a aplicação da estatística t de Student, para testar a hipótese de que o coeficiente de assimetria é igual a zero e de que o coeficiente de curtose é igual a três, servem para verificar se a distribuição do erro ajusta-se à distribuição normal de probabilidade. A distribuição normal associada a cada erro

confirma que foram cometidos somente erros casuais nas mensurações de cada medida e as variâncias obtidas para cada conjunto de medida permitem verificar se estes erros casuais foram iguais entre si (variâncias semelhantes), atestando a precisão das mensurações realizadas. Quanto menor a variância observada para cada conjunto de medidas, maior a precisão na sua obtenção, tendo como reflexo um coeficiente de determinação (R^2) maior, próximo do valor máximo 1,0, conferindo confiabilidade à interpretação dos resultados que serão obtidos a partir destas mensurações.

Enviado em: junho de 2005
Revisado e aceito: outubro de 2005

Some statistical directions for the evaluation of the method error when measuring a quantitative variable

Abstract

Introduction: the method error must be quantified to support the validity of the research and its clinical application. The determination of the reliability of the landmark identification to obtain measurements constitutes a fundamental precondition for a correct interpretation and utilization of the data. Each observed measurement can be analyzed through its actual value and its error component. The reliability is verified through repeated measurements. **Aim:** the purpose of this paper was to describe a statistical method which allows the researcher to verify if the measurements which he obtained at a certain moment are repeated by him at a posterior moment. This method confirms the non occurrence of the systematic error and allows an evaluation of the occurrence of the casual error, always presents in any form of measurement. Also, it presents the application of this statistical method in data referring to two cephalometric measurements and shows the syntax for the resolution of the method error through the use of the SPSS software for computers.

Key words: Method error. Statistics. Cephalometry. Reliability.

REFERÊNCIAS

- ADANS, J. W. Correction of error in cephalometric roentgenograms. *Angle Orthod*, Appleton, v. 10, p. 3-10, 1940.
- AHLQVIST, J.; ELIASSON, S.; WELANDER, U. The effect of projection errors on cephalometric length measurements. *Eur J Orthod*, Oxford, v. 8, p. 141-148, 1986.
- ALBUQUERQUE JR., H. R. **Avaliação da precisão - repetibilidade e reprodutibilidade - de medidas cefalométricas esqueléticas e dentárias**. 2002. Tese (Doutorado)-Faculdade de Odontologia, UNESP, Araraquara, 2002.
- BJÖRK, A.; SOLOW, B. Measurements on radiographs. *J Dent Res*, Alexandria, v. 41, p. 672-6683, 1962.
- DRAPER, N. R.; SMITH, H. *Applied regression analysis*. New York: John Wiley & Sons, 1968.
- FRANKLIN, J. B. Certain factors of aberration to be considered in clinical roentgenographic cephalometry. *Am J Orthod*, St. Louis, v. 38, p. 351-368, 1952.
- GRAVELY, J. F.; BENZIES, P. M. The clinical significance of tracing error in cephalometry. *Br J Orthod*, Oxford, v. 1, p. 95-101, 1974.
- HOFFMANN, R.; VIEIRA, S. *Análise de regressão: uma introdução à econometria*. São Paulo: Hucitec; EDUSP, 1977.
- HOUSTON, W. J. B. The application of computer aided digital analysis to orthodontic records. *Eur J Orthod*, Oxford, v. 1, p. 71-79, 1979.
- HOUSTON, W. J. B. The analysis of error in orthodontic measurements. *Am J Orthod*, St. Louis, v. 83, p. 382-390, 1983.
- MARTINS, L. P.; SANTOS-PINTO, A.; MARTINS, J. C. R.; MENDES, A. J. Erro de reprodutibilidade das medidas cefalométricas das análises de Steiner e de Ricketts, pelo método convencional e pelo método computadorizado. *Ortodontia*, São Paulo, v. 28, p. 4-17, 1995.
- MITGARD, J.; BJORK, G.; LINDER-ARONSON, S. Reproducibility of cephalometric landmarks and error of measurement of cephalometric cranial distances. *Angle Orthod*, Appleton, v. 44, p. 56-61, 1974.
- RODRIGUES, M. S. Vocabulário brasileiro de estatística. *Boletim da Faculdade de Filosofia de Ciências e Letras da USP*, São Paulo, n. 203, 1956.

Endereço de correspondência

Ary dos Santos Pinto
Rua Humaitá 1680 – Centro – C.P. 331
CEP: 14.801-903 - Araraquara/SP
Email : spinto@foar.unesp.br