

O MÉTODO DE *BOOTSTRAP* PARA O ESTUDO DE DADOS DE FADIGA DOS MATERIAIS

Mariano MARTINEZ-ESPINOSA¹
Vera Lúcia Martins SANDANIELO¹
Francisco LOUZADA-NETO²

- RESUMO: O termo *Bootstrap* se refere, em geral, a uma técnica ou método de simulação, que objetiva a obtenção de intervalos de confiança para as estimativas dos parâmetros de interesse, por reamostragem do conjunto de dados original. A técnica *Bootstrap* pode especificamente ser utilizada quando são considerados modelos de regressão polinomiais. Um caso especial desses modelos é obtido quando os valores das variáveis independentes são codificados, ao se utilizar um planejamento fatorial. Neste caso, as variáveis independentes podem ser ortogonalizadas, possibilitando estimativas dos parâmetros mais exatas. O objetivo deste trabalho é o uso da técnica *Bootstrap* residual e de pares em dados experimentais de resistência à fadiga em corpos de prova de madeira, obtidos no Laboratório de Madeiras e de Estruturas de Madeira (LaMEM) – EESC – USP – Brasil, para verificar a adequação do modelo polinomial ortogonal múltiplo com distribuição normal, no estudo da resistência à fadiga em corpos de prova de madeira e estimar os parâmetros deste modelo em diferentes situações experimentais, utilizando simulações. Os resultados das simulações pelos métodos de *Bootstrap* mostram que esta técnica pode ser utilizada para o objetivo deste trabalho.
- PALAVRAS CHAVE: *Bootstrap*; resistência à fadiga; madeira; modelo polinomial ortogonal múltiplo da distribuição normal.

1 Introdução

O procedimento *Bootstrap* é uma técnica de reamostragem, bastante utilizada em diferentes situações estatísticas. A base da técnica é a obtenção de um “novo” conjunto de dados, por reamostragem do conjunto de dados original (Efron e Tibishirani, 1993). Neste trabalho os métodos de reamostragem são utilizados especificamente para modelos de regressão, pois, pelo método de *Bootstrap*, podemos reexaminar uma análise de regressão já feita, comparando os resultados obtidos sob certas circunstâncias assumidas (Draper e Smith, 1998).

Observe que, um ponto importante da técnica *Bootstrap* não é somente avaliar as estimativas dos parâmetros, mas também obter boas estimativas dos erros padrão da distribuição gerada pelas estimativas dos parâmetros nas iterações de reamostragem

¹Departamento de Estatística, Instituto de Ciências Exatas e da Terra, Universidade Federal de Mato Grosso - UFMT, CEP 78060-900 Cuiabá, MT, Brasil. E-mail: marianom@cpd.ufmt.br / veluma@cpd.ufmt.br

²Departamento de Estatística, Universidade Federal de São Carlos - UFSCar, Caixa Postal 676, CEP 13565-905, São Carlos, SP, Brasil. E-mail: dfln@power.ufscar.br

(Lepage e Billard, 1992). Isto pode ser avaliado no ajuste de situações de estimação lineares e não lineares, tais como em modelos de regressão polinomiais (Draper e Smith, 1998). Estes modelos são utilizados em dados obtidos por meio de um planejamento estatístico de experimentos. Um caso particular de aplicação de modelos de regressão polinomial, no estudo da fadiga de madeira e derivados pode ser encontrado em Martinez e Calil (2000 e 2003).

O uso do método *Bootstrap* no estudo da resistência à fadiga em corpos de prova de madeira, é, em particular justificado por várias causas, sendo a principal a verificação do comportamento do modelo polinomial ortogonal múltiplo com distribuição normal, em diversas situações experimentais, considerando que para a realização dos ensaios, o tempo e custo em geral são elevados. Além disso, podemos obter estimativas mais precisas dos parâmetros dos modelos de regressão, dos erros padrões e intervalos de confiança.

Portanto, o objetivo deste trabalho é o uso da técnica *Bootstrap* para verificar a adequação do modelo polinomial ortogonal múltiplo com distribuição normal, em diferentes situações experimentais, no estudo da resistência à fadiga em corpos de prova de madeira, pelo uso do método de *Bootstrap* residual e *Bootstrap* de pares. Na Seção 2 apresentamos o modelo polinomial ortogonal múltiplo com distribuição normal. Na Seção 3 descrevemos, brevemente, o procedimento *Bootstrap* para o modelo de regressão múltipla. Materiais e métodos são apresentados na Seção 4. Os resultados experimentais são apresentados na Seção 5. Na Seção 6 desenvolvemos a análise dos dados.

2 Polinômio ortogonal múltiplo da distribuição normal

2.1 Definição do modelo

Uma expressão polinomial de grau d pode ser considerada como uma expansão em série de Taylor de uma função teórica verdadeira básica ($f(\vec{X}^c)$) truncada depois do termo de ordem d , onde as seguintes propriedades usualmente se cumprem (Box e Draper, 1987):

- A ordem mais alta da função aproximada, ou mais fechada da série de Taylor, pode aproximar a função verdadeira.
- Quanto menor é a região sobre a qual a aproximação é feita, melhor será a possível aproximação com uma função polinomial de um dado grau.

Na prática, podemos freqüentemente supor que, sobre regiões limitadas, um polinômio de primeiro ou segundo grau pode adequadamente representar a verdadeira função (Box e Draper, 1987). Portanto, os modelos de segunda ordem são os mais utilizados (Martínez, 2001) e sua forma geral, considerando um planejamento de experimento com três níveis e k fatores, dada pela seguinte expressão (Martínez e Calil, 2000):

$$Y_i = a_0 + \sum_{i=1}^k a_i x_i^c + \sum_{i=1}^k a_{ii} x_i^{2c} + \sum_{i=1}^{k-1} \sum_{j=2}^k a_{ij} x_i^c x_j^c + \varepsilon_i^c, \quad i < j \quad (1)$$

onde a resposta Y_i , em problemas de fadiga, representa o logaritmo de números de ciclos até a falha ($\ln(N)$), $x_1^c, x_2^c, \dots, x_k^c$ são as variáveis independentes codificadas das variáveis originais, geralmente denotadas por $x_1^*, x_2^*, \dots, x_k^*$ na área de planejamento de experimentos, que influem na resposta Y_i . Os parâmetros a_0 , a_i ($i = 1, 2, \dots, k$), a_{ij} ($i = 1, 2, \dots, k$), a_{ij} ($i = 1, 2, \dots, k; j = 1, 2, \dots, k$), são os parâmetros desconhecidos a serem estimados por meio dos dados, e ε_i^c é o erro aleatório (resíduo), onde a letra c minúscula representa o resíduo do modelo com as variáveis independentes codificadas. A codificação das variáveis independentes, em forma prática, é apresentada na Seção 4, e a justificativa teórica da codificação das variáveis independentes num planejamento de experimento pode ser encontrada em Martínez (2001).

Para estimar os parâmetros do modelo dado na equação (1) é preferível utilizar a notação matricial, pois esta facilita os cálculos, sendo assim a equação (1) pode ser escrita como:

$$\tilde{Y} = X^c \tilde{\theta} + \tilde{\varepsilon}^c, \quad (2)$$

onde $\tilde{Y} = (y_1, y_2, \dots, y_n)'$, X^c é uma matriz $n \times p$ codificada, com n linhas as quais representam os n ensaios experimentais e $p = (k+1)(k+2)/2$ colunas, $\tilde{\theta}$ é um vetor $p \times 1$ de parâmetros desconhecidos e $\tilde{\varepsilon}^c = (\varepsilon_1^c, \varepsilon_2^c, \dots, \varepsilon_n^c)'$.

2.2 Ortogonalização do modelo

Um planejamento fatorial foi considerado para o modelo da equação (2), e para este propósito os valores das variáveis independentes foram codificados. Neste caso os valores das variáveis que formam a matriz X^c , dada pelo modelo da equação (2), podem ser ortogonalizados com facilidade (Box e Draper, 1987). A matriz X^c ortogonalizada, em forma geral, pode ser representada da seguinte maneira (Martínez, 2001):

$$X = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1p} \\ 1 & x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{np} \end{bmatrix}. \quad (3)$$

As variáveis x_{ip} e x_{iq} são chamadas ortogonais uma em relação a outra se $x_{ip}x_{iq} = 0$, para $p \neq q$ (Khuri e Cornell, 1996).

Portanto, utilizando a matriz dada pela equação (3), o modelo dado pela equação (2) pode ser escrito como:

$$\tilde{Y} = X\tilde{\theta} + \tilde{\varepsilon}, \quad (4)$$

que é o modelo polinomial ortogonal múltiplo da distribuição normal, onde $\tilde{\varepsilon}$ é o erro aleatório do modelo com distribuição normal, com média zero e variância constante.

A ortogonalidade usada na aplicação estatística é justificada por várias razões, porém uma das causas principais é que a matriz $\tilde{X}'\tilde{X}$ é uma matriz diagonal, o que facilita a obtenção das estimativas dos parâmetros, do modelo polinomial ortogonal dado pela equação (4) e, além disso, evita a possibilidade de problemas de correlação entre as variáveis independentes deste modelo. Outra vantagem é que se pode facilmente decompor a soma de quadrados da regressão em componentes com apenas um grau de liberdade, cada uma das quais correspondendo a quantidade de variação de Y explicada por um dado fator (Khuri e Cornell, 1996 e Box e Draper, 1987).

Um caso especial do modelo dado pela equação (4), para $k = 2$ variáveis independentes é o modelo quadrático dado por (Martinez, 2001):

$$Y_i = \theta_0 + \theta_1 x_{1i} + \theta_2 x_{2i} + \theta_3 x_{1i}^2 + \theta_4 x_{2i}^2 + \theta_5 x_{1i} x_{2i} + \varepsilon_i, \quad (5)$$

para $i = 1, 2, \dots, n$ e $\varepsilon_i \sim N(0, \sigma^2)$, onde os parâmetros do modelo são obtidos pelo método dos mínimos quadrados. Observe que nesta situação são necessários no mínimo 7 valores, considerando que temos 6 parâmetros a serem estimados. Além disso, desde que o modelo conta com termos quadráticos em ambas as variáveis, ao menos três níveis de cada variável precisam ser usados, o que corresponde a no mínimo 9 valores para a utilização do modelo dado pela equação (5). Para este caso a matriz ortogonalizada X , sem réplicas é dada por:

$$X = \begin{bmatrix} 1 & -1 & -1 & 1 & 1 & 1 \\ 1 & 0 & -1 & -2 & 1 & 0 \\ 1 & 1 & -1 & 1 & 1 & -1 \\ 1 & -1 & 0 & 1 & -2 & 0 \\ 1 & 0 & 0 & -2 & -2 & 0 \\ 1 & 1 & 0 & 1 & -2 & 0 \\ 1 & -1 & 1 & 1 & 1 & -1 \\ 1 & 0 & 1 & -2 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}. \quad (6)$$

No entanto, em estudos experimentais de engenharia, física, química e outras áreas de pesquisa, sempre devem ser consideradas réplicas (p^*), neste caso, cada linha da matriz ortogonalizada apresentada em (6), é repetida p^* vezes (Martinez e Calil, 2003).

Uma das ferramentas mais importantes para a verificação da adequação de um modelo de regressão é a análise dos resíduos. Com esta análise, é possível verificar, se as suposições sobre os resíduos (ε_i) dos modelos dados pelas expressões (4) e (5) são satisfeitas, isto é, verificar se são cumpridas as suposições de igualdade de variância, de normalidade e de independência. A validade dessas suposições pode ser verificada por meio de gráficos.

Para avaliar a suposição de igualdade de variância, em geral é traçado o gráfico dos resíduos contra os valores ajustados. Esta suposição será válida se a dispersão dos resíduos é aleatoriamente distribuída em torno do zero.

A verificação da distribuição normal dos resíduos, também pode ser feita por meio de gráficos, sendo o gráfico de probabilidade normal bastante utilizado. Neste gráfico, se considera que a suposição de normalidade é válida se os pontos do gráfico estiverem localizados aproximadamente ao longo de uma reta.

A suposição de independência também pode ser analisada por um gráfico, geralmente utilizando o gráfico dos resíduos contra o tempo (ordem de coleta dos dados). Se os resíduos estiverem distribuídos de forma aleatória em torno do tempo (eixo horizontal), a suposição de independência será válida.

Detalhes desses gráficos podem ser encontrados em Martínez (2001). Outras técnicas podem ser encontradas nos livros de estatística e de regressão, como Draper e Smith (1998).

3 Procedimento de reamostragem para modelos de regressão

Dois métodos de reamostragem que podem ser utilizados em modelos de regressão são: o *Bootstrap* residual e *Bootstrap* de pares (Draper e Smith, 1998). Ambos os métodos são apresentados a seguir.

3.1 *Bootstrap* residual (*BR*)

Os passos principais deste método são os seguintes:

- Ajustar um modelo de regressão considerando os dados da amostra original e obter os n resíduos ($\hat{\epsilon}$).
- Selecionar uma amostra aleatória de tamanho n , dos resíduos $\hat{\epsilon}$ obtidos no passo (a), utilizando reamostragem com reposição, com probabilidade $1/n$ para cada resíduo selecionado ($\tilde{\epsilon}^{BR}$).
- Gerar os novos valores de \tilde{Y} pela seguinte equação:

$$\tilde{Y}^{BR} = X\tilde{A} + \tilde{\epsilon}^{BR}, \quad (7)$$

onde $\tilde{\epsilon}^{BR}$ são os resíduos obtidos no passo (b) e \tilde{A} é um vetor de parâmetros desconhecidos, estimados no passo (a).

- Ajustar um modelo de regressão, pelo método dos mínimos quadrados, utilizando o seguinte modelo:

$$\tilde{Y}^{BR} = X\tilde{A}^* + \tilde{\epsilon}^{BR}, \quad (8)$$

para obter a estimativa de \tilde{A}^* (\hat{A}^*).

- Repetir os passos (a), (b), (c) e (d), R vezes. Geralmente o número de iterações R é fixado em 999 (Martínez, 2001).

Então depois de R reamostragem, obtém-se a distribuição empírica de \tilde{A}^* e todas as estatísticas empíricas relacionadas a ela, tais como a média *Bootstrap* residual (média *BR*) e desvio padrão *Bootstrap* residual (desvio padrão *BR*) de cada um dos vetores

estimados. Um intervalo de confiança assintótico para \tilde{A}^* pode ser obtido de maneira usual considerando a teoria assintótica (Martínez, 2001). Além disso, podemos obter um intervalo de confiança percentil *Bootstrap* residual (intervalo percentil *BR*) para \tilde{A}^* pela ordenação de $\hat{A}_1^* \leq \hat{A}_2^* \leq \dots \leq \hat{A}_R^*$ e usar $\hat{A}_{(R+1)(\alpha^*/2)}^*$ e $\hat{A}_{(R+1)(1-\alpha^*/2)}^*$ respectivamente como limites inferior e superior do intervalo $100(1-\alpha^*)\%$ de confiança para \tilde{A}^* .

3.2 *Bootstrap* de pares (*BP*)

O *BP* é realizado por reamostragem dos pares originais de valores (Y_i, x_i) originais, onde Y_i é a i -ésima observação e x_i é a i -ésima linha da matriz X . Os passos principais deste método são:

- Reamostrar com reposição n pares dos valores originais (Y_i, x_i) , com probabilidade $1/n$ para cada par, obtendo $(\tilde{Y}_i^{BP}, X_i^{BP})$.
- Ajustar um modelo de regressão pelo método dos mínimos quadrados, considerando o seguinte modelo:

$$\tilde{Y}^{BP} = X^{BP} \tilde{A}^{**} + \tilde{\epsilon}^{bb}, \quad (9)$$

para obter a estimativa de \tilde{A}^{**} (\hat{A}^{**}).

- Repetir os passos (a) e (b), R vezes ($R = 999$).

De maneira semelhante ao método *BR*, depois de R reamostragens, obtém-se a distribuição empírica de \tilde{A}^{**} e todas as estatísticas empíricas relacionadas a ela, tais como a média *Bootstrap* de pares (média *BP*) e desvio padrão *Bootstrap* de pares (desvio padrão *BP*) de cada um dos vetores estimados. Um intervalo de confiança assintótico para \tilde{A}^{**} pode ser obtido de maneira usual, considerando a teoria assintótica (Martínez, 2001). Além disso, podemos obter um intervalo de confiança percentil *Bootstrap* de pares (intervalo percentil *BP*) para \tilde{A}^{**} pela ordenação de $\hat{A}_1^{**} \leq \hat{A}_2^{**} \leq \dots \leq \hat{A}_R^{**}$ e usar $\hat{A}_{(R+1)(\alpha^*/2)}^{**}$ e $\hat{A}_{(R+1)(1-\alpha^*/2)}^{**}$ respectivamente como limites inferior e superior do intervalo $100(1-\alpha^*)\%$ de confiança para \tilde{A}^{**} .

4 Materiais e métodos

Os aspectos relacionados aos materiais e métodos utilizados no desenvolvimento deste trabalho, para verificar a adequação do modelo polinomial ortogonal múltiplo da distribuição normal, pelo método de *Bootstrap*, são dados a seguir:

- Espécies de madeira:** Pinus Caribea variedade Hondurensis (*Pinus caribea* var. *hondurenses*).

- **Método de amostragem:** o método de amostragem utilizado na retirada dos corpos de prova foi o aleatório simples, selecionando aleatoriamente 54 corpos de prova da espécie considerada (Martínez, Calil e Sales, 2000).
- **Dimensão e confecção dos corpos de prova:** para determinação da resistência à fadiga em madeira, na tração, foram utilizados corpos de prova com as dimensões estabelecidas em Macêdo (1996 e 2000) e adotados pela NBR 7.190/97.
- **Execução dos ensaios:** os ensaios estáticos e cíclicos foram realizados na máquina universal de ensaios DARTEC M1000/RC do Laboratório de Madeiras e de Estruturas de Madeira (LaMEM) da EESC – USP, Brasil. Os ensaios estáticos para os corpos de prova de madeira seguiram as recomendações dadas na NBR 7190/97 e Macedo (1996). Os dados obtidos dos ensaios estáticos dos corpos de prova gêmeos foram utilizados para estabelecer os níveis de tensão máxima e mínima para os ensaios de fadiga na tração. Desta maneira, os níveis de tensão, máximos para os carregamentos cíclicos foram de 90%, 75% e 60% da resistência de ruptura do material (f_{t0}), estimados nos ensaios estáticos dos corpos de prova gêmeos, e o nível mínimo da tensão (σ_{\min}) foi de 10% do máximo ($0,10 \sigma_{\max}$). Nos ensaios cíclicos foram utilizados ciclos de tensões flutuantes de forma senoidal, considerando 3 diferentes frequências, de 1; 5 e 9 HZ, para os corpos de prova. O procedimento e os resultados dos ensaios estáticos utilizados são parte do trabalho de Macedo (1996) sobre ensaios estáticos de tração paralela em madeira.
- **Planejamento do experimento:** um planejamento composto central ortogonal foi utilizado para o estudar o número total de ciclos até a falha, N (variável resposta), considerando-se seis réplicas (Macedo, 1996) e usando combinações de dois fatores (tensão e frequência), cada um testado com três níveis. Os níveis para cada fator, assim como os códigos para cada nível são dados na Tabela 1, onde -1, 0 e 1 são códigos obtidos pelas seguintes equações (Martínez e Calil, 2000):

$$x_1^c = \frac{x_1^* - \bar{x}_1^*}{s_1^*}, \quad x_2^c = \frac{x_2^* - \bar{x}_2^*}{s_2^*}, \quad \text{onde:}$$

$$\bar{x}_1^* = \frac{60+75+90}{3} = 75, \quad \bar{x}_2^* = \frac{1+5+9}{3} = 5, \quad s_1^* = 75 - 60 = 15 \quad e \quad s_2^* = 5 - 1 = 4.$$

Tabela 1 - Número de ciclos, sob três níveis de tensões e três tipos de frequências

Código dos níveis.	x_i^c	-1	0	1
Tensão (S) em %.	x_1^*	60	75	90
Frequência (f) em HZ	x_2^*	1	5	9

5 Resultados

A Tabela 2 apresenta os resultados experimentais considerando a notação dada na Tabela 1. Os dados representam os números de ciclos correspondentes à vida à fadiga (N) nos corpos de prova de madeira sem emendas, com os fatores tensão (S) e frequência (f), para a espécie considerada. Os dados foram realizados no Laboratório de Madeiras e de Estruturas de Madeira (LaMEM) da EESC-USP, Brasil (Martinez; Calil e Sales, 2004).

Tabela 2 - Dados de resistência à fadiga em corpos de prova de madeira sem emendas, para a espécie de Pinus Caribea variedade Hondurensis, obtidos no LaMEM (Martinez; Louzada e Calil, 2004)

Ensaio (i)	Número de ciclos até a falha (N)					Fatores		
	N_{i1}	N_{i2}	N_{i3}	N_{i4}	N_{i5}	N_{i6}	S	f
1	211480	430899	204017	229270	252510	400197	60	1
2	93432	118610	178388	66818	72917	87134	75	1
3	91	57	21	139	113	87	90	1
4	900005	703389	1794615	1149067	1382568	791034	60	5
5	360056	124050	309413	163539	253120	219832	75	5
6	187	417	348	617	584	381	90	5
7	793417	1106942	1814582	1264336	1330109	1016584	60	9
8	150375	241314	454026	264474	297954	373929	75	9
9	103	325	649	619	669	925	90	9

Nota: N_{ip} , para $p^* = 1, \dots, 6$ são as réplicas.

6 Análise dos dados

A análise dos dados, para os números de ciclos correspondentes à vida da madeira sob fadiga, em corpos de prova, sem emendas para as espécies e adesivos considerados, se inicia com a verificação da adequação do modelo dado pela Expressão (5), conforme indicado na Seção 2.2, isto é, realizando uma análise dos resíduos deste modelo.

Assim, para verificar a igualdade de variância, foi utilizado um gráfico de resíduos contra os valores estimados. Este gráfico é apresentado na Figura 1, e mostra que eles apresentam forma de funil, indicando um aumento da variância, e portanto, deve ser feita uma transformação dos dados, para ver se é possível estabilizá-la.

Para verificar a distribuição normal, utilizamos um gráfico de escores normais. A Figura 2 ilustra este gráfico, no qual pode ser observado que os pontos não estão localizados ao longo de uma reta, o que indica que os resíduos do modelo não seguem uma distribuição normal.

Portanto, da análise dos resíduos, pode-se concluir que o modelo dado pela Expressão (5) não é adequado, sendo necessário fazer uma transformação nos dados da variável resposta. Neste caso, a transformação logarítmica ($y = \ln(N)$) foi a mais adequada.

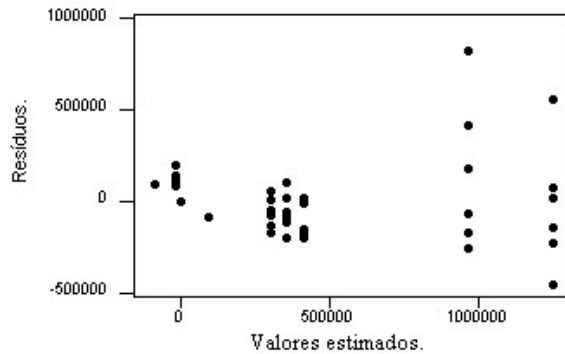


Figura 1 - Gráfico dos resíduos contra os valores estimados, para os dados da Tabela 2.

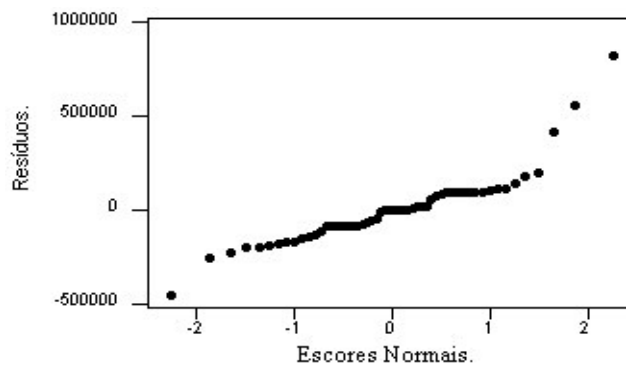


Figura 2 - Gráfico dos resíduos contra os escores normais, para os dados da Tabela 2.

O gráfico dos resíduos contra valores estimados e o gráfico dos resíduos contra os escores normais do novo modelo (modelo com os dados transformados) são apresentados nas Figuras 3 e 4, respectivamente.

Da Figura 3 pode-se concluir que a variância é constante. Porém, observa-se também a presença de dois valores com resíduos pequenos (terceira réplica do terceiro ensaio, com $N = 21$, frequência 1 e tensão de 90; primeira réplica do nono ensaio, com $N = 103$, frequência 9 e tensão de 90). Neste trabalho, os valores dos resíduos considerados pequenos ou grandes são destacados com um quadrado, e correspondem aos valores dos resíduos que não pertencem ao intervalo de confiança 95%, isto é, $(0 \pm 1,96 \times s)$, onde s é o desvio padrão amostral.

Na Figura 4, se observa um comportamento aproximadamente linear, ou que indica que os resíduos seguem uma distribuição normal aproximada, a menos dos valores destacados anteriormente na Figura 3, porém não considerados problemáticos para a utilização do modelo proposto dado pela Expressão (5).

Assim, ajustamos, um modelo polinomial ortogonal múltiplo da distribuição normal dado pela Expressão (5), com os dados do N transformados e os valores dos fatores codificados para os dados da Tabela 2, dado por:

$$\hat{y} = 10,344 - 4,001x_1^c + 0,723x_2^c - 0,885x_1^{2c} - 0,189x_2^{2c} + 0,091x_1^c x_2^c, \quad (10)$$

onde x_1^c representa a tensão e x_2^c a frequência.

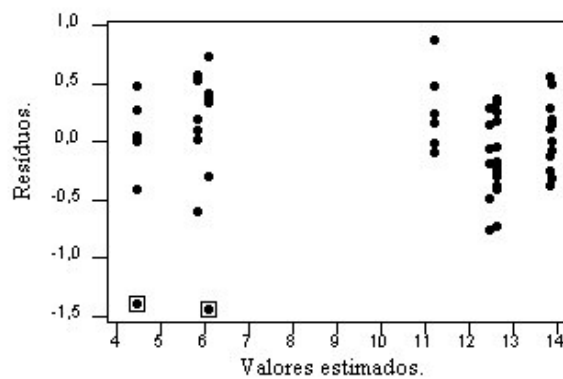


Figura 3 - Gráfico dos resíduos contra os valores estimados, para os dados da Tabela 2 (transformados).

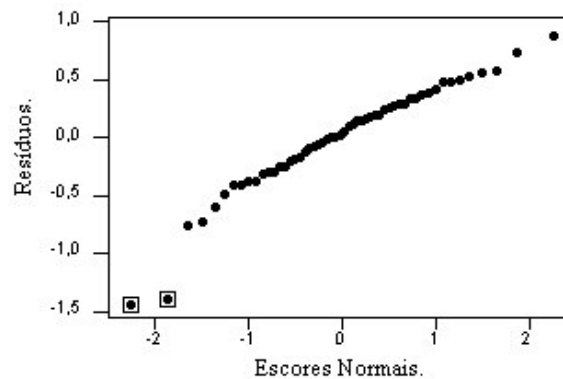


Figura 4 - Gráfico dos resíduos contra os escores normais para os dados da Tabela 2 (transformados).

Este novo modelo ajustado, tem um desvio padrão amostral de 0,4851 e um coeficiente de determinação de 98,4% ($R^2 = 98,4\%$). Além destes valores, na Tabela 3 apresentam-se os coeficientes das variáveis independentes, com seus respectivos desvios padrões (DP), valor da razão t (t_o) e numero de observações (n).

Da análise da Tabelas 3, observa-se que os efeitos lineares (x_1^c) e quadráticos (x_1^{2c}) da S no logaritmo do número total de ciclos até a falha ($\ln(N)$) são altamente significativos nos níveis usuais, sendo o efeito linear (x_1^c) da S relativamente grande quando comparado com o efeito quadrático, observe que a f também tem efeito linear (x_2^c) e quadrático significativo, no número total de ciclos até a falha. O efeito da interação entre a S e a frequência ($x_1^c x_2^c$) não é significativa, isto é, não existe interação entre os dois fatores. Observe que, com os intervalos de confiança apresentados na Tabela 5 devem ser obtidas as mesmas conclusões.

Tabela 3 - Coeficientes de regressão das variáveis independentes, DP , t_o , n , R^2 e s , para os dados do modelo ajustado da Expressão (10)

Variável	Coefficiente	DP	t_o
Constante	10,3445	0,0660	156,72
x_1^c	-4,00138	0,08084	-49,50
x_2^c	0,72345	0,08084	8,95
x_1^{2c}	-0,88545	0,04667	-18,97
x_2^{2c}	-0,18929	0,04667	-4,06
$x_1^c x_2^c$	0,09080	0,09901	0,92
$n = 54 \quad R^2 = 0,984 \quad s = 0,4851$			

Uma vez que o modelo ajustado, expressão (10), é adequado, como pode ser observado nos gráficos dos resíduos. Ele permite estimar o número de ciclos em função da S e da f para os dados obtidos, ou para extrapolar valores não determinados.

Para avaliar o comportamento do modelo polinomial ortogonal múltiplo da distribuição normal dado pela Expressão (5), em dados experimentais de resistência à fadiga em corpos de prova de madeira, em diversas situações experimentais foram feitas simulações, pelos métodos de BR e BP , utilizando as Expressões (8) e (9). Uma importante justificativa dessas simulações é o tempo e custo elevado da realização dos ensaios de fadiga. Observe que, uma vez tendo um conjunto de dados original obtido experimentalmente, é possível obter um “novo” conjunto de dados por reamostragem, podendo assim estimar os parâmetros deste modelo em diferentes situações experimentais. Cabe destacar que, no procedimento indicado foram utilizadas 999 iterações, todas de tamanho 54, pois foi utilizado um planejamento fatorial com 9 réplicas e os valores da matriz de planejamento ortogonalizados, tanto para o método BR como para o método BP .

Os resultados obtidos pelos métodos de BR e BP , considerando as equações (8) e (9), estão sumarizados nas Tabelas 4 e 5. A Tabela 4 apresenta os estimadores de mínimos quadrados (EMQ), os estimadores *Bootstrap* residual (EBR) e os estimadores *Bootstrap* de pares (EBP), com seus respectivos desvios padrão e vícios relativos (VR) para os EBR e EBP . A Tabela 5 mostra os intervalos 90% de confiança assintóticos, considerando a teoria assintótica (AS), e os estimadores BR e BP . Também a Tabela 5 mostra os intervalos de confiança 90% percentil obtidos pelos métodos de BR (PBR) e BP (PBP)

respectivamente. Para obter os estimadores pontuais e por intervalos, utilizando os métodos de *BR* e *BP*, foram consideradas 999 iterações. As estimativas por intervalo dos parâmetros, obtidas pelos métodos *BR* e *BP*, são mais fechadas que as obtidas de maneira usual (Martínez, 2001). Não mostramos os histogramas das distribuições empíricas das estimativas dos parâmetros obtidos pelos métodos de *BR* (*EBR*) e *BP* (*EBP*), dos quais observa-se que todos apresentam distribuições aproximadamente normal.

Tabela 4 - *EMQ*, *EBR*, *EBP*, *VREBR* e *BREBR*, para todos os parâmetros, utilizando os dados da Tabela 2

Método	Parâmetro					
	θ_0	θ_1	θ_2	θ_3	θ_4	θ_5
<i>EMQ</i>	10,3445 (0,0660)	-4,0014 (0,0808)	0,7234 (0,0808)	-0,8854 (0,0466)	-0,1893 (0,0466)	0,0908 (0,0990)
<i>EBR</i>	10,3415 (0,0629)	-4,0019 (0,0744)	0,7259 (0,0749)	-0,8861 (0,0434)	-0,1890 (0,0452)	0,0904 (0,0895)
<i>EBP</i>	10,3467 (0,0679)	-3,9981 (0,0890)	0,7187 (0,0885)	-0,8831 (0,0435)	-0,1875 (0,0429)	0,0911 (0,1181)
<i>VREBR</i>	-0,03	0,01	0,33	0,07	-0,17	-0,50
<i>VREBP</i>	0,02	-0,08	-0,66	-0,26	-0,96	0,33

Nota: Os valores entre parênteses são os respectivos desvios padrão.

Tabela 5 - Intervalos de confiança 90% *AS*, *BR*, *BP*, *PBR* e *PBP*, para todos os parâmetros, utilizando os dados das Tabelas 2 e 4

Parâmetro	Intervalo de confiança 90%				
	<i>AS</i>	<i>BR</i>	<i>BP</i>	<i>PBR</i>	<i>PBP</i>
θ_0	(10,236;10,453)	(10,238;10,449)	(10,236;10,458)	(10,236;10,443)	(10,231;10,457)
θ_1	(-4,134;-3,869)	(-4,124;-3,880)	(-4,144;-3,982)	(-4,130;-3,881)	(-4,149;-3,856)
θ_2	(0,591;0,856)	(0,603;0,849)	(0,574;0,864)	(0,601;0,846)	(0,568;0,856)
θ_3	(-0,962;-0,809)	(-0,957;-0,815)	(-0,955;-0,812)	(-0,958;-0,813)	(-0,955;-0,814)
θ_4	(-0,266;-0,113)	(-0,263;-0,115)	(-0,258;-0,117)	(-0,259;-0,111)	(-0,260;-0,118)
θ_5	(-0,072;0,253)	(-0,056;0,237)	(-0,103;0,284)	(-0,058;0,234)	(-0,100;0,284)

Conclusões

A utilização de um planejamento fatorial no estudo de dados de resistência à fadiga em corpos de prova de madeira e derivados é justificado por varias causas, entre elas: menor número de ensaios, menor tempo nos ensaios e, por conseguinte, menor custo, menor variação e melhor precisão nos resultados.

Como os efeitos da tensão e da frequência são estatisticamente significativos, o pesquisador deve utilizar um modelo estatístico aos dados, para estimar o número total de

ciclos até a falha como função da frequência e da tensão, e não unicamente em função da tensão como é usual na prática.

Para analisar os dados de fadiga em corpos de prova de madeira e derivados deve-se usar uma transformação logarítmica dos dados originais da variável resposta, pois a variação desses dados em geral é muito grande. Também, devem ser codificadas as variáveis independentes, tensão e frequência, pois os valores dessas variáveis podem ser facilmente ortogonalizados, e, assim sendo, pode-se utilizar o modelo polinomial ortogonal múltiplo da distribuição normal, para ajustar os dados de resistência à fadiga em corpos de prova de madeira e derivados.

A aplicação da técnica de *Bootstrap* residual e de pares apresentada fornece um método alternativo para obtenção dos estimadores pontuais e por intervalos, dos parâmetros do modelo polinomial ortogonal múltiplo, em dados experimentais de resistência à fadiga em corpos de prova de madeira. Do ponto de vista prático, a metodologia apresentada é de grande interesse no estudo da resistência à fadiga em corpos de prova de madeira e derivados, pois uma vez tendo um conjunto de dados original obtido experimentalmente é possível obter um “novo” conjunto de dados, por reamostragem deste conjunto de dados, podendo-se assim estimar os parâmetros desse modelo em diferentes possíveis situações experimentais. Como os ensaios de fadiga, em geral, são de tempo e custo elevados, utilizando a técnica de *Bootstrap*, os mesmos podem ser reduzidos notavelmente, uma vez que, em diferentes situações experimentais, utiliza-se simulações. A metodologia apresentada pode ser implementada e executada em computadores comuns. O software para fazer os cálculos dos métodos *Bootstrap* do modelo polinomial ortogonal múltiplo da distribuição normal é propriamente feito para a área de fadiga em madeira e seus derivados, no entanto o mesmo pode ser adaptado para outros materiais. O leitor interessado pode ter uma versão executável do programa escrevendo para o primeiro autor.

MARTINEZ-ESPINOSA, M.; SANDANIELO, V. L. M.; LOUZADA-NETO, F. Bootstrap procedure for studies on material fatigue data. *Rev. Mat. Est.*, São Paulo, v.24, n.2, p.37-50, 2006.

- *ABSTRACT: The term bootstrap is generally used to refer to a simulation technique that aims at obtaining interval confidences for the estimation of parameters of interest by resampling the original data set. The Bootstrap technique can be specifically used when polynomial regression models are considered. A special case of these models is obtained when the values of the independent variables are coded by using a factorial design. In this case, the independent variables are orthogonal, leading to more accurate parameter estimates. The objective of this study is to use both residual and pairs bootstrap procedure on experimental data on fatigue strength in wood specimens, obtained from the Laboratory of Wood and Timber Structures (LaMEM)-EESC-USP-Brazil for checking the adequacy of the normal multiple orthogonal polynomial model in the study of fatigue strength in wood specimens and estimate the parameters in this model at different experimental situations by using simulation. The results of the simulations by the bootstrap procedures show that the bootstrap technique can be used for the objective of this study.*
- *KEYWORDS: Bootstrap; fatigue strength; wood; normal multiple orthogonal polynomial model.*

Referências

- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. *NBR 7190/97*: Projeto de Estruturas de Madeira. Rio de Janeiro, 1997. p.107.
- BOX, G. E. P.; DRAPER, N. R. *Empirical model – Building and response surfaces*. New York: John Wiley, 1987. p.669.
- DRAPER, N. R.; SMITH, H. *Applied regression analysis*. New York: John Wiley & Sons, 1998. p.706.
- EFRON, B.; TIBISHIRANI, R. J. *An introduction to the bootstrap*. New York: John Wiley & Sons, 1993. p. 642.
- KHURI, A. I.; CORNELL, J. A. *Response surfaces: designs and analyses*. New York: Marcel Dekker, 1996. p. 668.
- LEPAGE, R.; BILLARD, L. *Exploring the limits of bootstrap*. New York: John Wiley & Sons, 1992. p.632.
- MACEDO, A. N. *Estudo de emendas dentadas em madeira laminada colada (MLC): Avaliação de métodos de ensaios*. 1996. 132f. Dissertação (Mestrado em Estruturas) – Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 1996.
- MACEDO, A. N. *Fadiga em emendas dentadas em madeira laminada colada (MLC)*. 2000. 216f. Tese (Doutorado em Estruturas) - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2000.
- MARTÍNEZ, M. E. *Desenvolvimento de um Modelo Estatístico para Aplicação no Estudo da Fadiga em Emendas Dentadas de Madeira*. 2001. 191f. Tese (Doutorado em Ciência e Engenharia de Materiais) – Interunidades – EESC/IFSC/IQSC, Universidade de São Paulo, São Carlos, 2001.
- MARTÍNEZ, M. E.; CALIL, C. J. Statistical fatigue experiment design in medium density fiberboard. *J. Mater. Res.*, Pittsburgh, v.3, n.3, p.84-91, 2000.
- MARTÍNEZ, M. E.; CALIL, C. J. Statistical design and orthogonal polynomial model to estimate the tensile fatigue strength of wooden finger joints. *Int. J. Fatigue*, Guildford, v.25, p.237-243, 2003.
- MARTÍNEZ, M. E.; CALIL, C. J.; SALES, A. Un método de muestreo para la determinación de las propiedades físicas y mecánicas de la madera. *Rev. Madera Cienc. Tecnol.*, Concepción, v. 2, n. 1, p.5-20, 2000.
- MARTÍNEZ, M. E.; LOUZADA, F. N.; CALIL, C. J. A Multiple orthogonal polynomial Birnbaum-Saunders model for fatigue. *Rev. Mat. Estat.*, São Paulo, v.22, n.2, p.53-72, 2004.

Recebido em 18.06.2005.

Aprovado após revisão em 24.06.2006.