

Métodos de estimação em regressão linear múltipla: aplicação a dados clínicos

Methods of Estimation in Multiple Linear Regression: Application to Clinical Data

EMÍLIO AUGUSTO COELHO-BARROS^{1,a}, PRISCILA ANGELOTTI SIMÕES^{2,b},
JORGE ALBERTO ACHCAR^{1,c}, EDSON ZANGIACOMI MARTINEZ^{1,d},
ANTÔNIO CARLOS SHIMANO^{2,e}

¹DEPARTAMENTO DE MEDICINA SOCIAL, FACULDADE DE MEDICINA DE RIBEIRÃO PRETO,
UNIVERSIDADE DE SÃO PAULO, SÃO PAULO, BRASIL

²DEPARTAMENTO DE BIOMECÂNICA, MEDICINA E REABILITAÇÃO DO APARELHO LOCOMOTOR,
FACULDADE DE MEDICINA DE RIBEIRÃO PRETO, UNIVERSIDADE DE SÃO PAULO, SÃO PAULO,
BRASIL

Resumo

Nesse artigo, tem-se o interesse em avaliar diferentes estratégias de estimação de parâmetros para um modelo de regressão linear múltipla. Para a estimação dos parâmetros do modelo foram utilizados dados de um ensaio clínico em que o interesse foi verificar se o ensaio mecânico da propriedade de força máxima (EM-FM) está associada com a massa femoral, com o diâmetro femoral e com o grupo experimental de ratas ovariectomizadas da raça *Rattus norvegicus albinus*, variedade Wistar. Para a estimação dos parâmetros do modelo serão comparadas três metodologias: a metodologia clássica, baseada no método dos mínimos quadrados; a metodologia Bayesiana, baseada no teorema de Bayes; e o método *Bootstrap*, baseado em processos de reamostragem.

Palavras chave: regressão linear múltipla, mínimos quadrados, análise bayesiana, *bootstrap*.

Abstract

In this paper, we show different parameters estimation forms for multiple linear regression model. We used clinical data, where the interest was to

^aPós-graduando. E-mail: eacbarros@gmail.com

^bPós-graduando. E-mail: priangelottisimo@yaho.com.br

^cProfessor. E-mail: achcar@fmrp.usp.br

^dProfessor. E-mail: edson@fmrp.usp.br

^eProfessor. E-mail: ashimano@fmrp.usp.br

verify the relationship among the mechanical assay maximum stress with femoral mass, femoral diameter and group of ovariectomized Wistar rats. We used three inference methods: Classic inference, based on the least square method; bayesian inference, based on the Bayes theorem; and bootstrap inference, based on resampling processes.

Key words: Multiple linear regression model, Least square method, Bayesian inference, Bootstrap inference.

1. Introdução

A análise de regressão é uma técnica estatística para investigar e modelar a relação entre variáveis, sendo uma das mais utilizadas na análise de dados. Pode-se citar inúmeras aplicações de análise de regressão na área da saúde (Freedman et al. 2004, Lyles & Kupper 1997, Chen & Wang 2004).

Um dos objetivos da análise de regressão é estimar os parâmetros desconhecidos do modelo. Mesmo existem várias técnicas de estimação desses parâmetros, nesse artigo serão considerados o método dos mínimos quadrados (Montgomery et al. 2001), o método bayesiano (Box & Tiao 1973) e o método *bootstrap* (Efron & Tibshirani 1993).

Em relação ao experimento clínico que será apresentado nesse artigo, alguns comentários devem ser feitos. Sabe-se que o tecido ósseo é formado por células que possuem em suas estruturas intercelulares, cálcio, protoplasma e fibras de colágenos, correspondendo a 70% do peso ósseo; o restante, por matriz extracelular (Bilezikian et al. 1994). Aproximadamente 80% da massa óssea esquelética é constituída por osso cortical e 20% por esponjoso, que ocupa a maior superfície do esqueleto (Szejnfeld 2000).

Durante a vida de uma pessoa, o sistema esquelético ajusta-se para manter as integridades estruturais dos ossos, que no cotidiano está sujeito a várias condições de carregamento mecânico. Conseqüentemente, a resposta estrutural, em parte, deve-se ao passado histórico de cargas impostas sobre o esqueleto e a necessidade presente (Keller et al. 1986).

A osteogênese pode ser estimulada por pequenas deformações na arquitetura óssea, provocadas por forças mecânicas aplicadas durante uma actividade física normal. Já a imobilização leva a um processo de reabsorção óssea (Turner 1999). Vários problemas podem acometer a estrutura óssea; entre estes destaca-se a diminuição da prática esportiva.

Com o aumento da expectativa de vida do homem, as doenças crônicas degenerativas destacam-se como factor limitante da qualidade de vida. De acordo com a Organização Mundial da Saúde (OMS), a osteoporose é uma das patologias que mais acometem a população actualmente. Esta se define como “uma doença esquelética sistêmica caracterizada por redução e massa e deterioração microarquitetural do tecido ósseo, com conseqüente aumento da fragilidade óssea e susceptibilidade à fratura” (Conference 1993).

Por ser uma patologia de etiologia multifatorial, muitas variáveis podem proporcionar o desenvolvimento de um quadro osteoporótico, baseando-se nas considerações de Robbins (1993) e Carneiro (1996), pode-se citar: sexo, biotipo, etnia, cafeína, tabagismo, etilismo, sedentarismo e imobilização prolongada, menopausa precoce e multiparidade, baixa ingestão de cálcio, ingestão exagerada de proteínas e fosfato. A osteoporose está principalmente relacionada com a deficiência dos hormônios ovarianos nas mulheres pós menopáusicas, afetando um terço delas (Johnell 1996). Dessa forma, o hormônio estrógeno possui importante função na homeostase do esqueleto, regulação da frequência de novos ciclos de remodelamento e na manutenção do equilíbrio entre formação e reabsorção óssea.

Sabe-se que com o aumento da idade, a eficiência dos osteoblastos diminui e intensifica a produção das células reabsorptivas-osteoclastos. Atualmente a osteoporose vem sendo tratada e diagnosticada precocemente, pois assim, um tratamento mais efetivo pode ser feito, visto que, com o quadro dessa patologia instalada, o osso não se encontrará em condições normais. A actividade física é a contramedida mais estudada para analisar ou prevenir a atrofia muscular e perda da massa óssea (Tuukanen et al. 1994, Van der Wiel et al. 1995). O treinamento físico estimula o crescimento do osso, em termos de largura, aumenta o diâmetro, o perímetro, o volume e o peso.

Com actividade física, a tábua óssea externa (cortical) e a parte esponjosa (trabecular) dos ossos treinados aumentam de espessura e há maior tolerância à solicitação externa (Eastell 2003). As adaptações ósseas dependem do tipo de exercício realizado, pois estímulos moderados demonstram que há um aumento pouco significativo da massa óssea. Em contrapartida, a actividade de suportar peso (na presença da gravidade) tem geralmente sido considerada como influência positiva na saúde do osso (Szejnfeld 2000).

Em particular, atividades como correr, saltar, que geram resistência externa no corpo humano entre 3 a 5 vezes o peso corporal e 7 a 10 vezes o peso corporal no impacto, respectivamente, são utilizadas para estimular o processo de remodelação óssea. Vários autores têm centralizado suas pesquisas em protocolos de treinamento físico que comprovem a eficiência da actividade física na prevenção e tratamento da osteoporose (Rennó et al. 2001). Há pesquisas que relatam uma ausência de influência significativa na arquitetura óssea (Pohlman et al. 1985).

Como não estão totalmente elucidados os mecanismos pelos quais o esqueleto responde à actividade física, realizam-se estudos com maior frequência sobre a terapêutica da osteoporose. De acordo com estudos realizados por Yeh et al. (1993), aumentaram a densidade mineral óssea e o conteúdo mineral de ratas ovariectomizadas e normais, ambos grupos treinados. Peng et al. (1997) mostraram através de estudos comparativos entre um grupo ovariectomizado e treinados, e outro não ovariectomizado e treinado, que o primeiro grupo citado obteve uma diminuição significativa da reabsorção óssea e um aumento da densidade óssea, em comparação com o segundo grupo.

As respostas ósseas à actividade física regular podem ser observadas em humanos ou animais experimentais. Diversos estudos em ossos humanos e de outros animais mostram que o desuso (paralisia, imobilização e ausência de descarga

de peso) e a deficiência hormonal resultam em perda de massa óssea (Bain & Rubin 1990, Smith & Gilligan 1989).

2. Metodologia

Geralmente, tem-se o interesse em avaliar a relação de uma variável de interesse Y (variável dependente ou variável resposta) em relação a k variáveis X_j (variável independente ou covariável), $j = 1, 2, \dots, k$. Portanto, um possível modelo para avaliar essa relação pode ser dado por

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + \varepsilon_i, \quad i = 1, \dots, n \quad (1)$$

em que, n é o número de indivíduos, y_i é a observação da variável dependente para o i -ésimo indivíduo, $X_i = (x_{i1}, x_{i2}, \dots, x_{ik})'$ é um vetor de observações das variáveis independentes para o i -ésimo indivíduo, $\beta = (\beta_0, \beta_1, \beta_2, \dots, \beta_k)'$ é um vetor de coeficientes de regressão (parâmetros) e ε_i é um componente de erro aleatório. Assume-se que esses erros são independentes e seguem distribuição normal com média zero e variância desconhecida σ^2 .

O modelo (1) é chamado de regressão linear múltipla, pois envolve mais de um coeficiente de regressão. O adjetivo “linear” indica que o modelo é linear em relação aos parâmetros $\beta = (\beta_0, \beta_1, \beta_2, \dots, \beta_k)$, e não porque y é uma função linear dos x 's.

O objetivo deste artigo é estudar o uso do modelo de regressão linear múltipla e comparar três métodos de estimação dos parâmetros do modelo: método dos mínimos quadrados, bayesiano e *bootstrap*.

Nesta pesquisa foram utilizadas 60 ratas da raça *rattus norvegicus albinus*, variedade Wistar, com aproximadamente 260 g, que corresponde à fase adulta jovem. O presente estudo baseou-se na utilização do modelo experimental de osteopenia advinda do procedimento cirúrgico denominado ovariectomia. Os animais foram divididos em seis grupos; em cada um foram alocados dez ratas (tabela 1).

2.1. Procedimentos experimentais

Os animais, após uma semana à coleta de sangue, foram submetidos a ovariectomia bilateral. A técnica cirúrgica aplicada nesse estudo seguiu a metodologia de Zarrow et al. (1964). Após 72 horas da cirurgia, os animais foram submetidos a um programa de treinamento físico em esteira rolante. O protocolo de treinamento que foi aplicado aos animais dos grupos C e F foi de caráter evolutivo de 60 minutos/dia, 4 sessões por semana durante 12 semanas, com velocidade de 15 m/min nas 6 primeiras semanas e 19 m/min nas demais semanas.

Os animais passaram por um período de adaptação à esteira durante 10 dias, este período iniciou-se com 10 minutos de exercitação, na velocidade de 8 m/min. Houve, a cada sessão de adaptação, o incremento da velocidade até a sétima sessão, mantendo-se constante até a sexta semana de exercitação e, o tempo evoluiu cinco minutos por dia.

TABELA 1: Divisão dos grupos experimentais.

Grupos	Procedimentos experimentais
A (Controle I)	Animais submetidos à coleta de sangue e eutanásia com 195 dias de idade.
B (Ovariectomizado)	Animais submetidos à coleta de sangue e após 1 semana a ovariectomia e eutanasiados com 195 dias de idade.
C (Ovariectomizado e treinado, antes da instalação osteopenia)	Animais submetidos à coleta de sangue e após 1 semana a ovariectomia. Após 1 semana iniciaram treinamento em esteira por 12 semanas e no fim do experimento foram eutanasiados com 195 dias de idade.
D (Controle II)	Animais submetidos à coleta de sangue e após 1 semana a ovariectomia e eutanasiados com 255 dias de idade.
E (Ovariectomizado)	Animais, submetidos à coleta de sangue e após 1 semana a ovariectomia e eutanasiados com 255 dias de idade.
F (Ovariectomizado e treinado, depois da instalação osteopenia)	Animais submetidos à coleta de sangue e depois de 1 sem. a ovariectomia. Após 9 semanas iniciaram treinamento em esteira por 12 semanas e no fim do experimento foram eutanasiados com 255 dias de idade.

O treinamento físico para os animais do grupo C iniciou-se 72 horas após a cirurgia. Os animais do grupo F iniciaram a exercitação em esteira posteriormente a instalação da osteopenia. Segundo Kallu (1991), a osteopenia instala-se um mês depois da ovariectomia. O exercício foi realizado em uma esteira motorizada Insight[®]-modelo EP-131, para ratos, sem oscilação da inclinação; sucedendo-se no plano horizontal.

Ao término do treinamento físico (24 horas), os animais foram pesados e submetidos à eutanásia. Os fêmures direito foram retirados, limpos das partes moles, identificados, pesados na balança eletrônica Marte[®]-modelo AS2000C com precisão de 0.01g. Em seguida, realizaram-se as medidas antropométricas de diâmetro maior, diâmetro menor e comprimento total da peça através do paquímetro digital da marca Mitutoyo[®].

Subseqüentemente, os fêmures direito foram submetidos a ensaio mecânico de flexo-compressão na máquina universal de ensaio EMIC[®]-10000N, no laboratório de bioengenharia da Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo.

Como foi visto, foram utilizadas 60 ratas no experimento; porém algumas ratas morreram antes da finalização do procedimento experimental ou não se adaptaram ao mesmo. Os dados obtidos podem ser observados na tabela 2.

2.2. Modelo estatístico

Para verificar a associação entre a variável dependente EM-FM (y), em relação às variáveis independentes grupo, massa e diâmetro femorais, considere o seguinte

TABELA 2: Dados experimentais.

Grupo	Massa(g)	Diâmetro(cm)	EM-FM(N)	Grupo	Massa(g)	Diâmetro(cm)	EM-FM(N)
A	1.33	2.44	150.00	D	1.38	2.38	128.96
A	1.20	2.38	124.59	D	1.52	2.39	139.54
A	1.01	2.14	107.34	D	1.28	2.49	162.85
A	1.13	2.25	75.06	D	1.20	2.26	123.20
A	1.18	2.11	105.49	D	1.17	2.17	121.24
A	1.37	2.26	154.49	D	1.15	2.25	118.29
A	1.04	2.10	112.43	D	1.16	2.33	121.25
A	1.22	2.43	100.39	D	1.17	2.22	117.64
A	1.09	2.35	120.09	D	1.29	2.32	131.82
A	1.11	2.25	159.89	D	1.38	2.53	126.60
B	1.03	2.32	85.60	E	1.04	1.99	142.04
B	1.44	2.49	135.40	E	0.96	2.07	97.94
B	1.34	2.50	137.29	E	1.15	2.21	86.89
B	1.17	2.51	99.17	E	1.21	2.35	83.09
B	1.13	2.21	101.91	E	1.18	2.28	110.78
B	1.18	2.36	83.93	E	1.40	2.62	96.89
B	1.18	2.20	109.09	E	1.13	2.34	106.39
B	1.40	2.37	103.21	F	1.04	2.26	105.15
C	1.48	2.38	109.37	F	1.43	2.90	114.68
C	1.16	2.35	118.08	F	1.41	2.73	117.66
C	1.30	2.35	121.71	F	1.24	2.27	102.11
C	1.16	2.16	113.16	F	1.17	2.37	80.72
C	1.03	1.91	83.32				
C	0.92	2.23	106.75				
C	1.18	2.35	108.78				

modelo de regressão linear múltipla (ver (1)),

$$y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_4 X_{i4} + \beta_5 X_{i5} + \beta_6 X_{i6} + \beta_7 X_{i7} + \varepsilon_i \quad (2)$$

em que X_{i1} , X_{i2} , X_{i3} , X_{i4} , X_{i5} são variáveis *dummy* que representam o grupo da i -ésima rata, em que

$$\begin{aligned}
 X_{i1} &= \begin{cases} 1, & \text{se grupo B;} \\ 0, & \text{caso contrário.} \end{cases} & X_{i2} &= \begin{cases} 1, & \text{se grupo C;} \\ 0, & \text{caso contrário.} \end{cases} \\
 X_{i3} &= \begin{cases} 1, & \text{se grupo D;} \\ 0, & \text{caso contrário.} \end{cases} & X_{i4} &= \begin{cases} 1, & \text{se grupo E;} \\ 0, & \text{caso contrário.} \end{cases} \\
 X_{i5} &= \begin{cases} 1, & \text{se grupo F;} \\ 0, & \text{caso contrário.} \end{cases}
 \end{aligned}$$

X_{i6} representa a massa femoral da i -ésima rata; X_{i7} representa o diâmetro femoral da i -ésima rata e ε_i representa o erro aleatório do i -ésimo indivíduo esses erros são independentes e seguem distribuição normal com média zero e variância desconhecida σ^2 .

2.3. Método dos mínimos quadrados

O método dos mínimos quadrados (MMQ) pode ser utilizado para estimar os coeficientes de regressão em (1). Suponha que $n > k$ observações são avaliadas, em que k é o número de covariáveis, e seja y_i a i -ésima variável resposta observada e x_{ij} a i -ésima observação da j -ésima variável independente ($i = 1, \dots, n, j = 1, \dots, k$). Assume-se que os erros ε_i são independentes e seguem distribuição normal com média zero e variância desconhecida σ^2 .

Reescrevendo (1), tem-se

$$y_i = \beta_0 + \sum_{j=1}^k \beta_j x_{ij} + \varepsilon_i, \quad i = 1, 2, \dots, n \tag{3}$$

A função de mínimos quadrados é dada por

$$S(\beta_0, \beta_1, \dots, \beta_k) = \sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n \left(y_i - \beta_0 - \sum_{j=1}^k \beta_j x_{ij} \right)^2 \tag{4}$$

A função S deve ser minimizada com respeito à $\beta_0, \beta_1, \dots, \beta_k$. Os estimadores de mínimos quadrados (EMQ) de $\beta_0, \beta_1, \dots, \beta_k$ devem satisfazer,

$$\frac{\partial S}{\partial \beta_0} \Big|_{\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_k} = -2 \sum_{i=1}^n \left(y_i - \hat{\beta}_0 - \sum_{j=1}^k \hat{\beta}_j x_{ij} \right) = 0 \tag{5}$$

e,

$$\frac{\partial S}{\partial \beta_j} \Big|_{\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_k} = -2 \sum_{i=1}^n \left(y_i - \hat{\beta}_0 - \sum_{j=1}^k \hat{\beta}_j x_{ij} \right) x_{ij} = 0, \quad j = 1, 2, \dots, k \tag{6}$$

De (5) e (6), tem-se as equações de mínimos quadrados,

$$\begin{aligned} n\hat{\beta}_0 + \hat{\beta}_1 \sum_{i=1}^n x_{i1} + \hat{\beta}_2 \sum_{i=1}^n x_{i2} + \dots + \hat{\beta}_k \sum_{i=1}^n x_{ik} &= \sum_{i=1}^n y_i \\ \hat{\beta}_0 \sum_{i=1}^n x_{i1} + \hat{\beta}_1 \sum_{i=1}^n x_{i1}^2 + \hat{\beta}_2 \sum_{i=1}^n x_{i1}x_{i2} + \dots + \hat{\beta}_k \sum_{i=1}^n x_{i1}x_{ik} &= \sum_{i=1}^n x_{i1}y_i \\ \vdots & \\ \hat{\beta}_0 \sum_{i=1}^n x_{ik} + \hat{\beta}_1 \sum_{i=1}^n x_{ik}x_{i1} + \hat{\beta}_2 \sum_{i=1}^n x_{ik}x_{i2} + \dots + \hat{\beta}_k \sum_{i=1}^n x_{ik}^2 &= \sum_{i=1}^n x_{ik}y_i \end{aligned} \tag{7}$$

É possível notar que há $p = k + 1$ equações, uma para cada coeficiente de regressão. Logo, as soluções dessas equações serão os EMQ para $\beta_0, \beta_1, \dots, \beta_k$.

Escrevendo (3) na forma matricial, tem-se

$$y = X\beta + \varepsilon$$

em que,

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad X = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix}, \quad \varepsilon = \begin{bmatrix} \varepsilon_0 \\ \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

Tem-se o interesse em encontrar o vetor de EMQ $\hat{\beta}$, no qual minimiza,

$$\begin{aligned} S(\beta) &= \sum_{i=1}^n \varepsilon_i^2 = \varepsilon' \varepsilon = (y - X\beta)'(y - X\beta) \\ &= y'y - \beta' X'y - y' X\beta + \beta' X' X\beta \\ &= y'y - 2\beta' X'y + \beta' X' X\beta \end{aligned} \quad (8)$$

Os EMQ devem satisfazer,

$$\left. \frac{\partial S}{\partial \beta} \right|_{\hat{\beta}} = -2X'y + 2X'X\hat{\beta} = 0$$

logo,

$$X'X\hat{\beta} = X'y \quad (9)$$

Tem-se que (9) são as equações de mínimos quadrados identicamente as mostradas em (7). Portanto o EMQ de β é dado por

$$\hat{\beta} = (X'X)^{-1}X'y \quad (10)$$

assumindo que $(X'X)^{-1}$ existe.

2.3.1. Intervalos de confiança para os coeficientes de regressão

Para construir um intervalo de confiança para o coeficiente de regressão β_j , assumir que os erros ε_i são normalmente e identicamente distribuídos com média zero e variância σ^2 . Logo, as observações y_i são normalmente e identicamente distribuídas com média $\beta_0 + \sum_{j=1}^k \beta_j x_{ij}$ e variância σ^2 . Como o EMQ de $\hat{\beta}$ é uma combinação linear das observações, então β é normalmente distribuídos com vetor de médias β e matriz de covariância $\sigma^2(X'X)^{-1}$. Isso implica que a distribuição marginal do coeficiente de regressão $\hat{\beta}_j$ é normal com média β_j e variância $\sigma^2 C_{jj}$, em que C_{jj} é o j -ésimo elemento da diagonal da matriz $(X'X)^{-1}$. Conseqüentemente,

$$\frac{\hat{\beta}_j - \beta_j}{\sqrt{\hat{\sigma}^2 C_{jj}}} \sim t_{(n-p)}, \quad j = 0, 1, \dots, k \quad (11)$$

em que, $t_{(n-p)}$ representa a distribuição t-student com $n - p$ graus de liberdade, n é o número de observações, p é o número de parâmetros do modelo de regressão e

$$\hat{\sigma}^2 = \frac{y'y - \beta' X'y}{n - p}$$

Portanto, um intervalo com $100(1 - \alpha)\%$ de confiança para o coeficiente de regressão β_j , $j = 0, 1, \dots, k$, é dado por

$$IC(95\%, \beta_j) = \hat{\beta}_j \pm t_{(\alpha/2, n-p)} \sqrt{\hat{\sigma}^2 C_{jj}} \tag{12}$$

2.4. Método bayesiano

Na inferência clássica, os parâmetros do modelo de regressão são vistos como quantidades fixas, mesmo que desconhecidas. Na inferência bayesiana, o conhecimento que se tem sobre os parâmetros do modelo pode ser quantificado através de uma distribuição de probabilidade. Essa distribuição de probabilidade é denominada distribuição a priori do parâmetro, em que esta reflete a informação que o pesquisador tem sobre os parâmetros. As inferências para os parâmetros do modelo de regressão são obtidas pela distribuição a posteriori de cada parâmetro.

Seja θ um vetor de parâmetros a serem estimados. Logo, pelo teorema de Bayes, tem-se a seguinte distribuição de probabilidade a posteriori para θ ,

$$\pi(\theta | y) = \frac{\pi(\theta)f(y | \theta)}{\int_{\theta} \pi(\theta)f(y | \theta)d\theta} \tag{13}$$

em que θ é contínuo, $\pi(\theta)$ é a distribuição a priori conjunta para θ e

$$f(y | \theta) = L(\theta) = \prod_{i=1}^n f(y_i | \theta)$$

a função de verossimilhança de θ , é

$$\pi(\theta | y) = c\pi(\theta)L(\theta)$$

em que

$$c^{-1} = \int_{\theta} \pi(\theta)L(\theta) d\theta$$

é uma constante normalizador; portanto,

$$\pi(\theta | y) \propto \pi(\theta)L(\theta) \tag{14}$$

Um estimador bayesiano para θ é obtido minimizando o erro esperado (função risco) com respeito a distribuição a posteriori para θ . Para isso deve considerar-se uma função de perda $(\hat{\theta}, \theta)$. Considerando o caso especial de função de perda quadrática dada por

$$(\hat{\theta}, \theta) = (\hat{\theta} - \theta)^2$$

a estimativa de Bayes para θ será o valor de $\hat{\theta}$ para qual $E[(\hat{\theta} - \theta)^2 | y]$ seja mínima.

Portanto, o estimador de Bayes para θ , com respeito a perda quadrática, é dado por

$$\hat{\theta} = E(\theta | y) \tag{15}$$

Na inferência Bayesiana, os intervalos para os parâmetros do modelo (intervalos de credibilidade) são obtidos directamente da distribuição a posteriori dos parâmetros. Seja θ um vetor de parâmetros a serem estimados, fixando uma probabilidade $1 - \alpha$. O intervalo de credibilidade para θ com probabilidade de cobertura $1 - \alpha$ é dado por (θ_*, θ^*) , tal que

$$\int_{-\infty}^{\theta_*} \pi(\theta | y) d\theta = \frac{\alpha}{2}; \quad \int_{\theta^*}^{\infty} \pi(\theta | y) d\theta = \frac{\alpha}{2}$$

Normalmente, as integrais que devem ser resolvidas em (14) e (15) podem ser de difícil solução. Neste caso, deve utilizar-se algum método numérico de aproximação para integrais. Um dos métodos mais utilizados para a resolução de integrais é o método de Monte Carlo em cadeia de Markov (MCMC). Trata-se de um método de simulação para se obter amostras da distribuição a posteriori de interesse (Chib & Greenberg 1995, Gelfand & Smith 1990).

2.4.1. Análise bayesiana

Para a análise bayesiana dos dados da tabela 2, e considerando o modelo definido em (2), foram consideradas as seguintes distribuições a priori para $\beta = (\beta_0, \beta_1, \beta_2, \dots, \beta_7)$ e σ^2 ,

$$\begin{aligned} \beta_j &\sim N(a_j; b_j^2), & a \text{ e } b \text{ conhecidos, } j = 0, 1, \dots, 7 \\ \sigma^2 &\sim IG(c; d), & c \text{ e } d \text{ conhecidos} \end{aligned} \quad (16)$$

em que, $N(a; b^2)$ denota uma distribuição normal com média a e variância b^2 e $IG(c; d)$ denota uma distribuição gamma inversa com média $\frac{d}{c-1}$ e variância $\frac{d^2}{(c-1)^2(c-2)}$.

Além disso, foi assumido independência a priori entre os parâmetros. Assim, a distribuição a priori conjunta é dada por,

$$\begin{aligned} \pi(\beta, \sigma^2) &\propto \prod_{j=0}^7 \exp \left[-\frac{1}{2b_j^2}(\beta_j - a_j)^2 \right] \times \\ &(\sigma^2)^{-(c+1)} \exp \left(-\frac{d}{\sigma^2} \right), \quad k = 0, 1, \dots, 7 \end{aligned} \quad (17)$$

Considerando o modelo (2), a função de verossimilhança para $\beta = (\beta_1, \dots, \beta_7)$ e σ^2 dado os dados observados y_i e as covariáveis X_{ij} , $i = 1, 2, \dots, 47$, $j = 0, 1, \dots, 7$, é dada por,

$$\begin{aligned} L(\beta, \sigma^2 | y_i, X_{ij}) &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma}} \exp \left[-\frac{1}{2} \frac{(y_i - \mu_i)^2}{\sigma^2} \right] \\ &= \frac{1}{(\sqrt{2\pi\sigma})^n} \exp \left[-\frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \mu_i)^2 \right] \end{aligned} \quad (18)$$

em que,

$$\mu_i = \beta_0 + \sum_{j=0}^7 \beta_j X_{ij}$$

Como visto em (14), a distribuição a posteriori conjunta para os parâmetros é obtida combinando-se a distribuição a priori (17) com a função de verosimilhança (18) a partir da fórmula de Bayes.

As distribuições a posteriori condicionais, são dadas por

i) Para β_j , em que, $j = 0, 1, \dots, 7$,

$$\pi(\beta_j | \beta_{(j)}, \sigma^2, y, x_j) \propto \exp \left[-\frac{1}{2b_j^2}(\beta_j - a_j)^2 - \frac{1}{2\sigma^2} \sum_{i=1}^n \left(y_i - \beta_0 - \sum_{j=0}^7 \beta_j X_{ij} \right)^2 \right]$$

em que, $\beta_{(j)} = (\beta_1, \dots, \beta_{j-1}, \beta_{j+1}, \dots, \beta_7)$; y é o vetor dos dados observados e x_j é o vetor das covariáveis, $j = 0, 1, \dots, 7$.

ii) Para σ^2 ,

$$\pi(\sigma^2 | \beta_j, \sigma^2, y, x_j) \propto (\sigma^2)^{-(c+1)} \frac{1}{(\sqrt{2\pi}\sigma)^n} \exp \left[-\frac{1}{2\sigma^2} \sum_{i=1}^n \left(y_i - \beta_0 - \sum_{j=0}^7 \beta_j X_{ij} \right)^2 - \frac{d}{\sigma^2} \right]$$

Para a análise bayesiana dos dados, foi considerado $a_j = 0$, $b_j^2 = 1000000$, $j = 0, 1, \dots, 7$, $c = d = 0.01$, nas distribuições a priori (16). Essa escolha dos hiperparâmetros foi motivada para se ter distribuições a priori aproximadamente não-informativas e tal que a convergência do algoritmo de simulação (*Gibbs* com *Metropolis-Hastings*) seja observada.

2.5. Método *bootstrap*

Alternativamente aos intervalos de confiança assintótico e bayesiano, o intervalo de confiança para β pode ser obtido via simulação *bootstrap*. Amostras *bootstrap* não paramétricas são obtidas reamostrando, com reposição, directamente dos dados disponíveis $(y_1, x_{11}, x_{12}, \dots, x_{1k}), \dots, (y_n, x_{n1}, x_{n2}, \dots, x_{nk})$ (Efron & Tibshirani 1993, DiCiccio & Efron 1996).

2.5.1. Intervalo de confiança *bootstrap* para os coeficientes de regressão

Nesta seção são apresentados os passos utilizados na obtenção de intervalo de confiança *bootstrap* para os coeficientes de regressão β . Será apresentado o intervalo de confiança, proposto por Efron (1982), baseado nos percentis da distribuição *bootstrap* de β , *p-bootstrap*. Uma recente discussão a respeito dos intervalos de

confiança obtidos via simulação *bootstrap* é apresentada em Carpenter & Bithell (2000).

Seja $U = (y, x_1, x_2, \dots, x_k)$, em que, $y = (y_1, \dots, y_n)$ é o vetor de observações da variável dependente e $x_j = (x_{1j}, \dots, x_{nj})$, $j = 1, \dots, k$, são os vetores de observações das variáveis independentes, x_j é uma variável contínua para todo j .

Intervalos *p*-bootstrap

- I. Amostragem, com reposição de U , uma amostra *bootstrap* $(y_1^*, x_{11}^*, \dots, x_{1k}^*), \dots, (y_n^*, x_{n1}^*, \dots, x_{nk}^*)$.
- II. Da amostra *bootstrap* $(y_1^*, x_{11}^*, \dots, x_{1k}^*), \dots, (y_n^*, x_{n1}^*, \dots, x_{nk}^*)$, obter o EMQ de β , definido em (10), representado por $\hat{\beta}^*$.
- III. Repetir os passos I e II um número B grande de vezes.
- IV. De $\hat{\beta}^* = (\hat{\beta}_{(1)}^* \leq \hat{\beta}_{(2)}^* \leq \dots \leq \hat{\beta}_{(B)}^*)$, para algum valor de α , ($0 < \alpha < 1$), o intervalo de confiança *p*-bootstrap $100 \times (1 - \alpha)\%$ é dado por

$$(\hat{\beta}_{(q_1)}^*; \hat{\beta}_{(q_2)}^*) \quad (19)$$

em que, $q_1 = \lceil B \times (\alpha/2) \rceil$ e $q_2 = B - q_1$.¹

Como exemplo, em IV, para ($\alpha = 0.05$ e $B = 1000$) $\implies (q_1 = 25$ e $q_2 = 975)$. Logo, o intervalo de confiança *p*-bootstrap de 95% é dado por $(\hat{\beta}_{(25)}^*; \hat{\beta}_{(975)}^*)$. Os intervalos de confiança para quaisquer outros parâmetros de interesse são obtidos de maneira similar.

Quando uma variável independente x_j não é contínua, deve-se fazer o processo de reamostragem (passos I, II e III) dentro de cada nível da variável (Wu 1986, Tibshirani 1988). Outras alternativas ao intervalo de confiança *p*-bootstrap são discutidas, por exemplo, em Efron & Tibshirani (1993), Davison & Hinkley (1997) e Carpenter & Bithell (2000).

3. Resultados e discussões

Para obtenção das estimativas dos parâmetros do modelo (2), baseado na metodologia clássica usando o MMQ, foi utilizado o comando `lm()` do software R (Apêndice A.) (R Development Core Team 2006). Na tabela 3 têm-se os estimadores de mínimos quadrados para os parâmetros do modelo com seus respectivos erros padrão assintóticos. Após a realização de uma análise de resíduos, foi observado um bom ajuste do modelo aos dados.

Para obtenção das estimativas dos parâmetros, baseado no método bayesiano via amostrador de *Gibbs*, foi utilizado o software Winbugs (Apêndice A.)

¹ $\lceil x \rceil$ indica o menor número inteiro maior ou igual ao argumento x .

TABELA 3: Resultados obtidos através do método de mínimos quadrados.

Parâmetro	EMQ	Erro-padrão	Intervalo de confiança 95%
β_0	61.534	39.570	(-18.504; 141.572)
β_1	-17.343	8.848	(-35.240; 0.554)
β_2	-12.590	9.048	(-30.891; 5.711)
β_3	3.114	8.451	(-13.980; 20.208)
β_4	-16.800	9.026	(-35.057; 1.457)
β_5	-21.557	10.880	(-43.564; 0.450)
β_6	48.841	27.643	(-7.072; 104.754)
β_7	1.055	23.499	(-46.476; 48.586)

(Spiegelhalter et al. 1995). Os parâmetros foram estimados via algoritmos *Gibbs-Sampling* e *Metropolis-Hastings*. Foram geradas 364000 amostras, das quais 4000 foram descartadas (“*burn in*”) com a finalidade de eliminar o efeito dos valores iniciais usados no algoritmo de simulação. Além disso, foram consideradas as iterações $60^a, 120^a, 180^a, \dots$, resultando em uma amostra final de 6000 observações para cada parâmetro. A convergência do algoritmo foi verificada através de gráficos temporais das amostras geradas e utilizando técnicas usuais existentes na literatura (Gelman & Rubin 1992). Na tabela 4, tem-se os sumários a posteriori obtidos para os parâmetros do modelo.

Para obtenção das estimativas dos parâmetros foi desenvolvida uma rotina, baseada no método *bootstrap*, em linguagem S utilizando o software R (Apêndice A.). Na tabela 5, pode-se observar o intervalo de confiança *p-bootstrap* considerando $B = 100000$ reamostras do conjunto de dados originais (tabela 2). O número de reamostras *bootstrap* sugerido na literatura parece crescer com o avanço computacional, seja em nível de hardware ou de software. Efron & Tibshirani (1993) comentam que “ $B = \infty$ ” é o número ideal para se estimar erros padrão. Naturalmente, na prática, B deve ser um número finito restrito ao poder computacional disponível.

TABELA 4: Resultados obtidos através do método bayesiano.

Parâmetro	Média a Posteriori	Erro-Padrão	Intervalo de credibilidade 95%
β_0	60.910	40.260	(-17.020; 141.300)
β_1	-17.340	9.230	(-35.490; 0.767)
β_2	-12.550	9.215	(-30.780; 5.815)
β_3	3.215	8.660	(-14.090; 20.480)
β_4	-16.730	9.472	(-35.740; 2.312)
β_5	-21.640	11.140	(-43.620; -0.107)
β_6	48.930	28.410	(-7.159; 105.000)
β_7	1.255	23.930	(-45.140; 49.260)

A amplitude (R) e a forma (F), para os 3 tipos de intervalos de confiança, são apresentados na tabela 6. Estas medidas são importantes para comparar os 3 procedimentos de construção dos intervalos de confiança. Sejam $\hat{\theta}_{LS}$ e $\hat{\theta}_{LI}$ os limites superior e inferior do intervalo de confiança para os parâmetros do modelo. A forma para o intervalo de confiança dos parâmetros do modelo, calculada por

TABELA 5: Resultados obtidos através do método *bootstrap*.

Parâmetro	EMQ*	Erro-Padrão*	p-intervalo de confiança 95%
β_0	59.418	39.929	(-16.977; 138.593)
β_1	-17.174	8.274	(-35.186; 1.499)
β_2	-12.101	8.504	(-29.326; 5.038)
β_3	3.130	7.879	(-13.139; 20.827)
β_4	-16.784	8.400	(-38.441; 5.715)
β_5	-21.742	10.297	(-42.164; -1.144)
β_6	50.075	27.740	(6.240; 97.348)
β_7	1.269	23.728	(-44.259; 45.286)

* Baseado nas $B = 100000$ reamostras *bootstrap*.

$F = (\hat{\theta}_{LS} - \hat{\theta}) / (\hat{\theta} - \hat{\theta}_{LI})$, mede a assimetria do intervalo em torno de $\hat{\theta}$; $F > 1$ indica maior distância entre $(\hat{\theta}_{LS} - \hat{\theta})$ e $(\hat{\theta} - \hat{\theta}_{LI})$. Como os intervalos de confiança assintóticos são simétricos, implica que $F = 1$ (Efron & Tibshirani 1993).

TABELA 6: Amplitude (R) e forma (F) para os intervalos de confiança 95% para os parâmetros do vetor β .

Parâmetro	Intervalo					
	Assintótico		Bayesiano		p-bootstrap	
	R	F	R	F	R	F
β_0	160.076	1.000	158.320	1.032	155.570	1.036
β_1	35.794	1.000	36.257	0.998	36.685	1.037
β_2	36.602	1.000	36.595	1.007	34.364	0.995
β_3	34.188	1.000	34.570	0.998	33.966	1.088
β_4	36.514	1.000	38.052	1.002	44.156	1.039
β_5	44.014	1.000	43.513	0.980	41.020	1.009
β_6	111.826	1.000	112.159	1.000	91.108	1.078
β_7	95.062	1.000	94.400	1.035	89.545	0.967

4. Conclusões

Dos resultados da tabela 6, pode-se observar que os intervalos de confiança *bootstrap* são bem mais precisos, isto é, com menores comprimentos do que os intervalos de confiança assintóticos e os intervalos de credibilidade bayesianos. Porém, a alternativa bayesiana se mostrou mais eficiente se comparada com o método clássico.

A comparação entre métodos de estimação deve ser feita utilizando simulação. Portanto, não é possível afirmar que, em geral, o método *bootstrap* retorna estimativas mais precisas se comparadas com as metodologias clássicas e bayesianas. Porém, pode-se concluir que, para esse exemplo clínico, o método *bootstrap* retornou estimativas mais precisas.

Em relação ao método bayesiano, é importante salientar que é possível obter resultados mais precisos e próximos da realidade, utilizando informações a priori assumindo o conhecimento prévio de profissionais da saúde (Dunson & Herring 2005), lembrando que se deve tomar muito cuidado na escolhas de prioris informativas.

Considerando o método *bootstrap*, pode-se concluir que existem evidências de que o grupo e a massa femoral influenciam na medida do ensaio mecânico da propriedade de força máxima (EM-FM). Essa conclusão foi tirada baseada nos intervalos de confiança dos parâmetros. Como pode observar-se (tabela 5), os intervalos de confiança para os parâmetros β_5 e β_6 não contem o valor zero.

Agradecimientos

Os autores agradecem os pareceristas por seus valiosos comentários e sugestões. A pesquisa de Emílio Augusto Coelho Barros e Priscila Angelotti Simões recebe auxílio financeiro da FAPESP (Fundação de Amparo a Pesquisa do Estado de São Paulo). Jorge A. Achcar agradece ao CNPq pelo apoio financeiro.

[Recibido: noviembre de 2007 — Aceptado: mayo de 2008]

Referências

- Bain, S. D. & Rubin, C. T. (1990), 'Metabolic Modulation of Disuse Osteopenia: Endocrine-Dependent Site Specificity on Bone Remodeling', *J Bone Miner Res* **5**, 1069–1075.
- Bilezikian, J. P., Raisz, L. G. & Rodan, G. A. (1994), *Principles of Bone Biology*, Academic Press.
- Box, G. E. P. & Tiao, G. C. (1973), *Bayesian Inference in Statistical Analysis*, Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont. Addison-Wesley Series in Behavioral Science: Quantitative Methods.
- Carneiro, R. A. (1996), 'Osteoporose problema mundial', *ARS CVRANDII Clínica Médica* **29**, 5.
- Carpenter, J. & Bithell, J. (2000), 'Bootstrap Confidence Intervals: When, which, what? A Practical Guide for Medical Statistician', *Statistics in Medicine* **19**, 1141–1164.
- Chen, Z. & Wang, Y. G. (2004), 'Efficient Regression Analysis with Ranked-Set Sampling', *Biometrics* **60**(4), 997–1004.
- Chib, S. & Greenberg, E. (1995), 'Understanding the Metropolis-Hastings Algorithm', *The American Statistician* **49**(4), 327–335.
- Conference, C. D. (1993), *Diagnosis, Prophylaxis and Treatment of Osteoporosis*, American Journal of Medicine.
- Davison, A. C. & Hinkley, D. V. (1997), *Bootstrap Methods and their Application*, Vol. 1 of *Cambridge Series in Statistical and Probabilistic Mathematics*, Cambridge University Press, Cambridge. With 1 IBM-PC floppy disk (3.5 inch; HD).

- DiCiccio, T. J. & Efron, B. (1996), 'Bootstrap Confidence Intervals', *Statistical Science* **11**(3), 189–228.
- Dunson, D. B. & Herring, A. H. (2005), 'Bayesian Latent Variable Models For Mixed Discrete Outcomes', *Biostatistics* **6**(1), 11–25.
- Eastell, R. (2003), 'Management of Osteoporosis due to Ovarian Failure', *Med. Pediatric Oncol* **41**(3), 222–227.
- Efron, B. (1982), *The Jackknife, the Bootstrap and other Resampling Plans*, Vol. 38 of *CBMS-NSF Regional Conference Series in Applied Mathematics*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, United States.
- Efron, B. & Tibshirani, R. J. (1993), *An Introduction to the Bootstrap*, Vol. 57 of *Monographs on Statistics and Applied Probability*, Chapman and Hall, New York, United States.
- Freedman, L. S., Fainberg, V., Kipnis, V., Midthune, D. & Carroll, R. J. (2004), 'A new Method for Dealing with Measurement Error in Explanatory Variables of Regression Models', *Biometrics* **60**(1), 172–181.
- Gelfand, A. E. & Smith, A. F. M. (1990), 'Sampling-Based Approaches to Calculating Marginal Densities', *Journal of the American Statistical Association* **85**, 398–409.
- Gelman, A. & Rubin, D. B. (1992), 'Inference From Iterative Simulation Using Multiple Sequences', *Statistical Science* **7**, 457–511.
- Johnell, O. (1996), 'Advances in Osteoporosis: Better Identification of Risk Factors can Reduce Morbidity and Mortality', *J. Int. Med* **239**, 299–304.
- Kallu, D. N. (1991), 'The Ovariectomized Rat Model of Postmenopausal Bone Loss', *Endocrinology* **124**(1), 7–16.
- Keller, T., Spengler, D. M. & Carter, D. R. (1986), 'Geometric, Elastic, and Structural Properties of Maturing Rat Femora', *Journal of Orthopedic Research* **4**, 57–67.
- Lyles, R. H. & Kupper, L. L. (1997), 'A Detailed Evaluation of Adjustment Methods for Multiplicative Measurement Error in Linear Regression with Applications in Occupational Epidemiology', *Biometrics* **53**(3), 1008–1025.
- Montgomery, D. C., Peck, E. A. & Vining, G. G. (2001), *Introduction to Linear Regression Analysis*, Wiley Series in Probability and Statistics: Texts, References, and Pocketbooks Section, third edn, Wiley-Interscience, New York.
- Peng, Z. Q., Vaananen, H. K. & Tuukanen, J. (1997), 'Ovariectomy-Induced Bone Loss can be Affect by Different Intensities of Treadmill Running Exercise in Rats', *Calcified Tissue International* **60**, 441–448.

- Pohlman, R. L., Darby, L. A. & Lechner, A. J. (1985), 'Morphometry and Calcium Contents in Apendicular and Axial Bones of Exterior Ovariectomized Rats', *American Journal Physiology* **248**, 12–17.
- R Development Core Team (2006), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.
*<http://www.R-project.org>
- Rennó, A. C. M., Driusso, P. & Ferreira, V. (2001), 'Atividade física e osteoporose: uma revisão bibliográfica', *Fisioterapia em movimento* **13**(2).
- Robbins, S. (1993), *Patologia estrutural e funcional*, Guanabara.
- Smith, E. L. & Gilligan, C. (1989), 'Mechanical Forces and Bone', *J Bone Miner Res* **6**, 139–173.
- Spiegelhalter, D. J., Thomas, A., Best, N. G. & Gilks, W. R. (1995), *BUGS: Bayesian Inference Using Gibbs Sampling, Version 0.50.*, MRC Biostatistics Unit Cambridge.
- Szejnfeld, V. L. (2000), *Osteoporose: diagnóstico e tratamento*, Sarvier.
- Tibshirani, R. (1988), 'Correction to Discussion of: "Jackknife, Bootstrap and other Resampling Methods in Regression Analysis"', *Ann. Statist.* **16**(1), 479.
- Turner, R. (1999), 'Mechanical Signaling in the Development of Postmenopausal Osteoporosis', *Lupus* **8**, 388–392.
- Tuukanen, J., Peng, Z. & Vaananen, H. K. (1994), 'Effect of Running Exercise of the Bone Loss Induced by Orchidectomy in the Rats', *Calcified Tissue International* **55**, 33–37.
- Van der Wiel, H. E., Lips, P., Graafmans, W. C., Danielsen, C. C., Nauta, J., Van Lingen, A. & Mosekilde, L. (1995), 'Additional Weight-Bearing During Exercise is more Important than Duration of Exercise for Anabolic Stimulus of Bone: A Study of Running Exercise in Female Rats', *Bone* **16**(1), 73–80.
- Wu, C. F. J. (1986), 'Jackknife, Bootstrap and other Resampling Methods in Regression Analysis', *Ann. Statist.* **14**(4), 1261–1350. With discussion and a rejoinder by the author.
- Yeh, J. K., Aoia, J. F., Tierney, J. M. & Sprintz, S. (1993), 'Effect of Treadmill Exercise on Vertebral and Tibial Bone Mineral Content and Bone Mineral Density in the Aged Adult Rat: Determined by Dual Energy X-Ray Absorptiometry', *Calcified Tissue International* **52**, 234–238.
- Zarrow, N. X., Yochim, J. M. & McCarthy, J. L. (1964), *Experimental Endocrinology: A Sourcebook of Basic Techniques*, Academic Press, New York, United States.

Apêndice A. Programas

Abaixo têm-se as linhas de comando do software R, utilizadas para obter os EMQ e as amostras *bootstrap*.

```

fit2 <- lm(EM.FM~Grupo+Massa.F+ Diam,data=dados)
s.fit1 <- summary(fit2)

B <- 100000
C <- matrix(0,ncol=8,nrow=B)
S <- matrix(0,ncol=8,nrow=B)
T. <- matrix(0,ncol=8,nrow=B)

d1 <- dados[dados$Grupo=="Cont1",]
d2 <- dados[dados$Grupo=="Cont2",]
d3 <- dados[dados$Grupo=="OVXa",]
d4 <- dados[dados$Grupo=="OVXd",]
d5 <- dados[dados$Grupo=="OVXTa",]
d6 <- dados[dados$Grupo=="OVXTd",]

for(i in 1:B)
{
  a1 <- sample(1:10,replace=T)
  a2 <- sample(1:10,replace=T)
  a3 <- sample(1:7,replace=T)
  a4 <- sample(1:8,replace=T)
  a5 <- sample(1:7,replace=T)
  a6 <- sample(1:5,replace=T)

  b1 <- d1[a1,]
  b2 <- d2[a2,]
  b3 <- d3[a3,]
  b4 <- d4[a4,]
  b5 <- d5[a5,]
  b6 <- d6[a6,]

  b <- rbind(b1,b2,b3,b4,b5,b6)
  fit <- lm(EM.FM~Grupo+Massa.F+Diam,data=b)

  C[i,] <- fit$coefficients
  S[i,] <- summary(fit)[[4]][,2]
  T.[i,] <- (C[i,]-fit2$coefficients)/(S[i,])
}

```

Abaixo têm-se as linhas de comando do software Winbugs, utilizadas para obter as estimativas bayesianas.

```
model
{
for(i in 1:ind)
{
EMFM[i]~dnorm(mu[i],sigma2)
mu[i] <- beta0+beta1*OVXd[i]+beta2*OVXa[i]+beta3*Cont2[i]+
          beta4*OVXTa[i]+beta5*OVXTd[i]+beta6*MassaF[i]+
          beta7*Diam[i]
}
sigma2~dgamma(1.0E-2,1.0E-2)
beta0~dnorm(0,1.0E-6)
beta1~dnorm(0,1.0E-6)
beta2~dnorm(0,1.0E-6)
beta3~dnorm(0,1.0E-6)
beta4~dnorm(0,1.0E-6)
beta5~dnorm(0,1.0E-6)
beta6~dnorm(0,1.0E-6)
beta7~dnorm(0,1.0E-6)
}
```
