

**O USO DO PRINCÍPIO DA
PSEUDO - VEROSSIMILHANÇA
EM ENSAIOS CLÍNICOS COM AUSÊNCIA
DE GRUPO CONTROLE**

Erika Tiemi Fukunaga

DISSERTAÇÃO APRESENTADA AO
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
DA UNIVERSIDADE DE SÃO PAULO
PARA OBTENÇÃO DO GRAU DE MESTRE EM ESTATÍSTICA

Área de Concentração: **Estatística**

Orientador: **Prof. Dr. Heleno Bolfarine**

Durante a elaboração deste trabalho a autora recebeu apoio financeiro do CNPq
e do Hospital Universitário da Universidade de São Paulo.

São Paulo, 12 de Dezembro de 1997

**O USO DO PRINCÍPIO DA PSEUDO-VEROSSIMILHANÇA
EM ENSAIOS CLÍNICOS
COM AUSÊNCIA DE GRUPO CONTROLE**

Este exemplar corresponde à redação final
da dissertação devidamente corrigida e
defendida por Erika Tiemi Fukunaga
e aprovada pela comissão julgadora.

São Paulo, 06 de Janeiro de 1998.

Banca examinadora:

- Prof. Dr. Heleno Bolfarine (Orientador) - IME - USP
- Prof.^a. Dr.^a Mônica Carneiro Sandoval - IME - USP
- Prof. Dr. Filidor Edilson Vilca Labra - IMECC - UNICAMP

AGRADECIMENTOS

Após a longa jornada até o término deste trabalho fica difícil lembrar de todas as pessoas que de alguma forma contribuíram para torná-lo possível, gostaria portanto de colocar neste primeiro parágrafo um agradecimento geral, a todos que de alguma forma me possibilitaram chegar até aqui.

Gostaria de agradecer a todos os professores da minha graduação, pela excelente formação adquirida neste Instituto. Em particular, ao meu orientador o Prof. Heleno Bolfarine, por despertar em mim durante o curso do Bacharelado o gosto pela Estatística e a vontade de um aprimoramento acadêmico, dado por sua orientação tanto na Iniciação Científica como durante todo o período do mestrado.

Gostaria de agradecer ao Prof. Stephen Senn por toda atenção, mesmo que à distância, no sentido de esclarecer possíveis dúvidas sobre seu artigo.

Ao funcionários do Hospital Universitário, pelo convívio e aprimoramento da minha formação. Em particular aos médicos da Divisão de Clínica Médica da qual fiz parte e as secretárias desta divisão, Cida e Lourdes.

Aos amigos que sempre tiveram tempo de ouvir, discutir e sugerir alguma coisa neste trabalho: Péricles, Francisco, Audrey, Márcio José e Marcos Rubens.

Aos amigos que dentro deste Instituto dividiram comigo o período do mestrado, em especial : Cristina, Liane, Aldy e Eliana. E aos amigos que mesmo à distância sempre me incentivaram e torceram por mim, em particular Max e Ronaldo.

À Cláudia Kimura, que mesmo sem conhecer o significado da palavra vida acadêmica, sempre me incentivou e apoiou as minhas escolhas.

E finalmente, à minha família que sempre apoiou os meus estudos o que me permitiu chegar ao mestrado.

Resumo

Em pesquisas médicas é muito comum a utilização de ensaios clínicos relacionados com a comparação de dois grupos onde um deles é o grupo controle e o outro o grupo de tratamento. Os indivíduos participantes do estudo são escolhidos ao acaso para estar no grupo de tratamento ou controle. É comum também situações onde duas medidas são tomadas em cada grupo, sendo uma antes e outra após a aplicação do tratamento. Contudo, existem situações em que não é possível a obtenção de um grupo controle, portanto torna-se necessário um método em que se possa avaliar o efeito do tratamento levando-se em conta a ausência do grupo controle. O objetivo deste trabalho é apresentar um possível modelo que aborda este tipo de problema e comparar 3 métodos de estimação dos parâmetros através de estudos de simulação.

Abstract

In medical research it is common to consider clinical trials related to comparing two groups, one is the control group and the other the treatment group. The individuals participating in such study are assigned at random to the control or treatment groups. It is also common the situations where two measurements are obtained from each subject, one priori to treatment and the other after to treatment. However, there are situations where it is not possible to work with a control group so that it is necessary to consider an approach that can be used to evaluate the effect of the treatment in the absence of the control group. The main object of this dissertation is to present an approach that can be used in such cases and to compare three different estimation approaches to estimate the model parameters.

CONTEÚDO

Introdução	1
Capítulo 1 - Uma abordagem inicial.....	4
1.1 O problema	4
1.2 O modelo adotado.....	7
1.3 O método dos momentos	10
1.4 Exemplo.....	24
1.5 Críticas	26
Capítulo 2 - Uma nova proposta.....	28
2.1 O método de máxima verossimilhança.....	28
2.2 A Função de Verossimilhança	29
2.3 Variância dos Estimadores.....	33
2.4 Aspectos Computacionais.....	37
Capítulo 3 - Uma solução aproximada para o caso da amostra selecionada	40
3.1 O Princípio da Pseudo - Verossimilhança	41
3.2 Estimadores de máxima verossimilhança.....	42
3.3 Considerações	48
Capítulo 4 - Estudos de simulação.....	49
4.1 O método dos momentos	49
4.2 Comparação entre os métodos	55
Conclusão	63
Anexo I - Derivadas para a log-verossimilhança	65
Anexo II - Programas dos estudos de simulação	85
Referências.....	112

INTRODUÇÃO

Em muitas situações em estudos clínicos existe o interesse em verificar se um dado tratamento é eficiente. Nestes casos é comum a utilização de 2 grupos de estudo selecionados aleatoriamente, onde o primeiro recebe o tratamento em questão e o segundo recebe alguma substância inócua (no caso de drogas, por exemplo) ou é apenas acompanhado. Estes grupos são, respectivamente, denominados: grupo de tratamento e grupo controle. A finalidade do grupo controle é verificar a possível existência do efeito de regressão para a média, ou seja, que um efeito de alteração na variável observada no paciente seja devida a reação do seu organismo e não ao efeito do tratamento. Embora o grupo controle seja por este motivo extremamente importante nem sempre é possível obtê-lo, quer por motivos éticos quer por dificuldades em se estabelecer e acompanhar um grupo controle adequado. Quando não é possível a existência de um grupo controle adequado, torna-se necessário algum método em que se possa avaliar o efeito do tratamento levando-se em conta o efeito da regressão para a média.

O primeiro a discutir o problema foi James (1973), onde ele apresenta um modelo que considera a regressão para a média e utilizando o método dos momentos encontra estimadores para o efeito da regressão e do tratamento e propõe estimadores para suas respectivas variâncias através de aproximações por série de Taylor.

No entanto Senn e Brown (1985) criticam o método proposto por James (1973), pois naquele método o cálculo por aproximações para as estimativas das variâncias dos estimadores utiliza certas suposições que podem ser questionadas. Pode-se citar, por exemplo, o fato de James (1973) utilizar resultados que seriam válidos para a distribuição normal usual mas não para a normal truncada. Senn e Brown (1985) propõem ainda um outro método de estimação através do método de máxima verossimilhança, onde são também apresentadas as possíveis verossimilhanças associadas ao problema que foram obtidas por Cohen (1955). Dada a complexidade do modelo, o método de máxima verossimilhança requer procedimentos iterativos.

O objetivo deste trabalho é apresentar o problema levantado por James (1973) onde é discutida a ausência de um grupo controle em situações experimentais e o método por ele proposto, e discutir as críticas feitas a seu trabalho por Senn e Brown (1985). Apresentamos também, a proposta alternativa feita por Senn e Brown (1985) através do método de máxima verossimilhança. Finalmente, apresentamos uma aproximação para o método de máxima verossimilhança que é estudado em Chen e Cox (1992), utilizando o princípio da Pseudo - Verossimilhança. Mostramos através de simulações que tais aproximações são razoáveis no caso de “screening trials”, ou seja, em estudos envolvendo um grande tamanho para a amostra onde parte desta amostra é selecionada a partir de um truncamento.

Finalmente, cabe ressaltar que apesar deste trabalho apresentar uma possível abordagem para a ausência de grupos controle em situações experimentais, em nenhum momento isto deve ser encarado como um estímulo a não utilização de ensaios experimentais envolvendo grupo controle e sim como uma possível alternativa de análise em casos onde não é possível a obtenção de tais grupos.

Capítulo 1

Uma abordagem inicial

1.1 O problema

Em estudos na área médica é muito comum a utilização de ensaios clínicos na avaliação de um tratamento, sendo o ensaio do tipo tratamento e controle com medidas pré e pós tratamento extremamente utilizado. No entanto, nem sempre é possível a obtenção de um grupo controle. Portanto, quando analisamos dados provenientes de situações experimentais em que são obtidas medidas antes e depois de um tratamento sem um paralelo grupo controle, é importante que o pesquisador esteja ciente de uma possível mudança na variável observada, que se deve a regressão em direção à média. Este fenômeno pode confundir a interpretação dos dados, sendo muitas vezes negligenciado. Salientamos que mesmo que o tratamento não seja eficiente, pode haver uma redução nos valores observados após o tratamento. Um exemplo clássico a respeito deste fenômeno é a relação entre alturas de pais e filhos, onde pais altos tendem a ter em média filhos mais baixos e pais baixos tendem a ter em média filhos mais altos.

James (1973) enfatiza em seu artigo a importância do grupo controle, pois este permitiria uma distinção maior entre o verdadeiro efeito do tratamento e o efeito da regressão em direção à média.

No sentido de ilustrar a situação que será abordada por James (1973), iremos apresentar a seguir a estrutura dos dados do ensaio clínico estudado. Imagine a seguinte situação: existe o interesse em se avaliar uma dada variável X (por exemplo, taxa de colesterol), neste caso são escolhidas um determinado número de pessoas N (tamanho da amostra), onde é medida a variável de interesse para estas pessoas. Então, elas serão divididas em 2 grupos : normais (pessoas cujo valor de $X <$ ponto de truncamento) e risco (pessoas cujo valor de $X \geq$ ponto de truncamento), onde o ponto de truncamento (denominado x_0) é definido por algum critério arbitrário. Em seguida é aplicado um tratamento (por exemplo, a utilização de algum medicamento), para as pessoas consideradas na faixa de risco, e faz-se uma segunda medida da variável de interesse (esta segunda medida será denominada Y) para estas pessoas. É preciso ressaltar que não existe um grupo controle, ou seja, todas as pessoas que se encontrarem na faixa de risco receberão tratamento.

Tentando tornar mais visual a explicação dada acima, montamos o esquema apresentado a seguir:

ESTRUTURA DOS DADOS

N : tamanho da amostra

n : número de elementos na faixa de risco

m : número de elementos na faixa considerada normal

FAIXA	PRÉ		PÓS
<p>RISCO (n elementos)</p>	<p>$X_{(N)}$ ----- ----- $X_{(N-n+1)}$</p>	<p>Tratamento →</p>	<p>$Y_{(N)}$ ----- ----- $Y_{(N-n+1)}$</p>
<p>ponto de truncamento</p> <p>NORMAIS (m elementos)</p>	<p>$X_{(N-n)}$ $X_{(N-n-1)}$ ----- ----- ----- ----- ----- ----- $X_{(2)}$ $X_{(1)}$</p>		

O objetivo deste capítulo é apresentar a abordagem proposta por James (1973) para contornar a ausência de grupo controle em dados truncados e discutir os problemas existentes com tal abordagem.

1.2 O modelo adotado

Inicialmente James (1973) apresenta um levantamento bibliográfico sobre a questão da regressão para a média, onde ele cita por exemplo, o trabalho desenvolvido por Galton (1886) com crescimento de ervilhas. Segundo James (1973), Galton (1886) observou em seu trabalho que plantas de ervilhas que apresentavam crescimento elevado produziam descendentes em média mais baixos e que plantas que apresentavam baixo crescimento produziam descendentes que eram em média mais altos, ou seja, estava ocorrendo uma regressão em direção ao valor médio. Para explicar o fenômeno da regressão para a média Galton (1886) utilizou a regressão linear simples com distribuição normal bivariada (ver por exemplo Elian, 1988), onde a equação de regressão (dado x) é dada por:

$$y = \mu_y + \rho \sigma_y / \sigma_x (x - \mu_x) + e \quad (1.1)$$

com :

x : variável independente

y : variável dependente

μ_x : média populacional da variável x

μ_y : média populacional da variável y

σ_x : desvio-padrão populacional da variável x

σ_y : desvio-padrão populacional da variável y

ρ : correlação entre as variáveis x e y

e : erro aleatório, com : $e \sim N(0, \sigma_y^2 (1 - \rho^2))$.

No entanto, para a maioria das populações genéticas é esperado que sejam satisfeitas algumas suposições adicionais, como : $\mu_x = \mu_y = \mu$ e $\sigma_x = \sigma_y = \sigma$. Portanto, considerando as suposições acima a equação (1.1) dado x, fica simplificada da seguinte forma:

$$y = \mu + \rho (x - \mu) + e. \quad (1.2)$$

James (1973) adota o modelo (1.2) que considera o fenômeno da regressão para a média e incorpora o efeito do tratamento através do seguinte modelo:

$$y = \mu + \rho \gamma (x - \mu) + \varepsilon, \quad (1.3)$$

com x dado, onde :

x : medida anterior ao tratamento

y : medida posterior ao tratamento

μ : média populacional para a variável x não truncada

ρ : efeito de regressão para a média

γ : efeito do tratamento

ε : erro aleatório, com : $\varepsilon \sim N(0, \sigma^2(1 - \rho^2))$

σ^2 : variância populacional para a variável x não truncada.

James (1973) força a entrada do parâmetro de tratamento (γ), quebrando a estrutura da normal bivariada e “repartindo” o efeito de regressão em 2 partes: uma devida ao fenômeno da regressão para a média e outra devida ao efeito do tratamento. Portanto, para o modelo descrito em (1.3) pode-se notar com relação ao efeito de tratamento, os seguintes pontos:

- $|\gamma| = 1$ implica que não existe efeito do tratamento e portanto qualquer alteração na variável observada se deve exclusivamente ao efeito da regressão para a média, ou seja, recaímos no caso da normal bivariada, descrito anteriormente.

- $|\gamma| > 1$ implica que o tratamento está contribuindo para aumentar o valor da variável observada, fato este que pode estar sendo encoberto ou compensado pelo efeito da regressão para média.

- $|\gamma| < 1$ implica que o tratamento está contribuindo para regredir o valor da variável observada em direção ao valor médio, ou seja, o tratamento está apresentando resultados satisfatórios, lembrando que parte dessa regressão se deve ao efeito da regressão para a média.

NOTA: Assumiremos a partir deste momento um truncamento superior para o desenvolvimento da teoria, lembrando que tudo que for feito poderia ser aplicado (com as devidas alterações) para um truncamento inferior.

1.3 O método dos momentos

O método dos momentos (ver por exemplo Dudewicz e Mishra,1988) consiste na utilização de momentos amostrais, no caso \bar{x} , S_x^2 e S_y^2 , para estimar os parâmetros de interesse do modelo, dados pelo vetor $\underline{\theta} = (\mu, \sigma^2, \rho, \gamma)'$.

É preciso lembrar antes de utilizar o método dos momentos, que a variável x ou “medida antes” será truncada no ponto x_0 (valor arbitrário definido pelo pesquisador ou por estudos já realizados) sendo necessária a obtenção da média e da variância teórica para a variável truncada x_T . Tais expressões são apresentadas em James (1973), Tabela 1, e de uma forma mais geral, esses resultados podem ser encontrados em Johnson e Kotz (1970). Tomando os resultados da Tabela 1 em James (1973), temos que :

$$E [x / x \geq x_0] = \mu + k_0 \sigma = \mu_{x_1} \quad (1.4)$$

e

$$\text{Var} [x / x \geq x_0] = \sigma^2 [k_0 (z_0 - k_0) + 1] = \sigma_{x_1}^2, \quad (1.5)$$

onde :

$$z_0 = (x_0 - \mu) / \sigma,$$

$$k_0 = \phi(z_0) / [1 - \Phi(z_0)],$$

$\phi(z_0)$: valor da função de densidade de probabilidade da normal padrão em z_0 ,

$\Phi(z_0)$: valor da função de distribuição acumulada da normal padrão em z_0 .

Também será necessária a variância de y dado o valor do truncamento x_0 , de modo que assumindo o modelo dado por (1.3), supondo que x e ε são independentes, e utilizando a expressão (1.5) temos que:

$$\begin{aligned} \text{Var} [y / x \geq x_0] &= \text{Var} [\mu + \rho \gamma (x - \mu) + \varepsilon / x \geq x_0] & (1.6) \\ &= \rho^2 \gamma^2 \text{Var} [x / x \geq x_0] + \text{Var} [\varepsilon / x \geq x_0] \\ &= \rho^2 \gamma^2 \sigma^2 [k_0 (z_0 - k_0) + 1] + \sigma^2 (1 - \rho^2) \\ &= \sigma^2 [\rho^2 \gamma^2 [k_0 (z_0 - k_0) + 1] + (1 - \rho^2)] \\ &= \sigma_{Y/X}^2. \end{aligned}$$

Considerando as expressões (1.4), (1.5), (1.6) e seus correspondentes valores amostrais \bar{x} , S_x^2 , S_y^2 e tomando o estimador usual de regressão $\hat{\beta}$ de β (β : coeficiente angular, que no caso do modelo (1.3) é dado por $\gamma\rho$) obtido pelo método de mínimos quadrados (ver por exemplo Draper e Smith, 1981), o método dos momentos produz o seguinte sistema de equações de estimação:

$$\bar{x} = \hat{\mu} + k_0 \hat{\sigma}, \quad (1.7)$$

$$S_x^2 = \hat{\sigma}^2 [k_0(z_0 - k_0) + 1], \quad (1.8)$$

$$S_y^2 = \hat{\sigma}^2 [\hat{\rho}^2 \hat{\gamma}^2 [k_0(z_0 - k_0) + 1] + (1 - \hat{\rho}^2)], \quad (1.9)$$

$$\hat{\beta} = \hat{\rho} \hat{\gamma}. \quad (1.10)$$

Para resolver o sistema formado pelas equações (1.7), (1.8), (1.9) e (1.10) será suposto que z_0 é conhecido, o que é equivalente a conhecer a porcentagem da população truncada pelo ponto x_0 . Com esta suposição adicional James (1973) apresenta a seguinte solução para o sistema:

$$\hat{\mu} = \bar{x} - k_0 \hat{\sigma}, \quad (1.11)$$

$$\hat{\sigma}^2 = S_x^2 / [k_0(z_0 - k_0) + 1], \quad (1.12)$$

$$\hat{\rho} = [\hat{\beta}^2 [k_0(z_0 - k_0) + 1] - (S_y^2 / \hat{\sigma}^2) + 1]^{1/2}, \quad (1.13)$$

$$\hat{\gamma} = \hat{\beta} / \hat{\rho}. \quad (1.14)$$

No entanto convém lembrar que na prática $z_0 = (x_0 - \mu) / \sigma$ normalmente não é conhecido, e portanto neste caso o método dos momentos já não apresenta uma solução tão simples quanto a obtida acima. Em termos práticos é razoável assumir que o pesquisador da área tenha uma idéia razoável da porcentagem da população truncada e portanto utiliza-se essa porcentagem (noção ou “chute”) como se fosse a verdadeira (uma aproximação). Observou-se por simulação (vide Capítulo 4) que se o afastamento do valor fornecido pelo pesquisador para z_0 não for muito grande em relação ao seu verdadeiro valor, a diferença entre as verdadeiras estimativas (calculadas com z_0 verdadeiro) e as estimativas aproximadas (calculadas com z_0 fornecido pelo pesquisador) não são tão grandes.

Admitindo os estimadores encontrados em (1.11), (1.12), (1.13), (1.14), James (1973) obtém expressões para as variâncias destes estimadores. Note que :

$$\begin{aligned} \text{Var}(\hat{\sigma}^2) &= \text{Var}(S_X^2 / [k_0(z_0 - k_0) + 1]) \\ &= [k_0(z_0 - k_0) + 1]^{-2} \text{Var}(S_X^2). \end{aligned} \quad (1.15)$$

Para calcular a $\text{Var}(S_X^2)$ é preciso lembrar o seguinte resultado:

$$\frac{(n-1)S_X^2}{\sigma_X^2} \sim \chi^2_{(n-1)}. \quad (1.16)$$

Utilizando o resultado dado por (1.16) temos então:

$$\text{Var} \left(\frac{(n-1)S_x^2}{\sigma_{x_i}^2} \right) = \text{Var}(\chi^2_{(n-1)}) , \quad (1.17)$$

$$\text{Var} \left(\frac{(n-1)S_x^2}{\sigma_{x_i}^2} \right) = 2(n-1) ,$$

$$\text{Var} (S_x^2) = \frac{2\sigma_{x_i}^4}{(n-1)} .$$

Então retomando a expressão (1.15) e utilizando a variância deduzida na expressão (1.17), temos que:

$$\begin{aligned} \text{Var}(\hat{\sigma}^2) &= [ko(zo - ko) + 1]^{-2} \text{Var}(S_x^2) & (1.18) \\ &= [ko(zo - ko) + 1]^{-2} \frac{2\sigma_{x_i}^4}{(n-1)} \\ &= \frac{2\sigma_{x_i}^4}{[ko(zo - ko) + 1]^2 (n-1)} . \end{aligned}$$

Obtendo agora a variância de $\hat{\mu}$ temos :

$$\begin{aligned} \text{Var}(\hat{\mu}) &= \text{Var}(\bar{x} - ko \hat{\sigma}) & (1.19) \\ &= \text{Var}(\bar{x}) + (ko)^2 \text{Var}(\hat{\sigma}) . \end{aligned}$$

Obs: Estamos supondo na passagem acima que \bar{x} e S_x^2 são independentes (pensando na marginal associada a variável x).

Utilizando uma aproximação de primeira ordem por Série de Taylor (ver por exemplo Leite e Singer, 1990) para uma função $h(z)$, temos :

$$h(z) \cong h(z_0) + h'(z_0)(z - z_0), \quad (1.20)$$

ou

$$h(z) \cong h(E(z)) + h'(E(z))(z - E(z)) \quad (\text{usando: } z_0 = E(z)),$$

$$h(z) - h(E(z)) \cong h'(E(z))(z - E(z)),$$

$$[h(z) - h(E(z))]^2 \cong [h'(E(z))(z - E(z))]^2,$$

de modo que :

$$E[h(z) - h(E(z))]^2 \cong E[h'(E(z))(z - E(z))]^2,$$

ou seja,

$$\text{Var}(h(z)) \cong E[h'(E(z))(z - E(z))]^2 \quad (h'(E(z)) = \text{constante}),$$

ou

$$\text{Var}(h(z)) \cong h'(E(z))^2 E[(z - E(z))]^2,$$

ou ainda,

$$\text{Var}(h(z)) \cong h'(E(z))^2 \text{Var}(z).$$

Adotando:

$$z = \hat{\sigma}^2,$$

$$h(z) = \sqrt{z} = \hat{\sigma},$$

$$h'(z) = \frac{1}{2\sqrt{z}} = \frac{1}{2\hat{\sigma}},$$

$$E(z) = E(\hat{\sigma}^2) = \sigma^2,$$

$$\text{Var}(z) = \text{Var}(\hat{\sigma}^2),$$

temos então pela aproximação dada por (1.20) que:

$$\text{Var}(h(z)) \cong [h'(E(z))]^2 \text{Var}(z) \quad (1.21)$$

$$\begin{aligned} \text{Var}(\hat{\sigma}) &\cong \left[\frac{1}{2\sigma} \right]^2 \frac{2\sigma_{x_i}^4}{(ko(z_0 - ko) + 1)^2 (n-1)} \\ &= \frac{\sigma_{x_i}^4}{2\sigma^2 (n-1)(ko(z_0 - ko) + 1)^2}. \end{aligned}$$

Então retomando a expressão (1.19) temos que:

$$\begin{aligned} \text{Var}(\hat{\mu}) &= \text{Var}(\bar{x}) + (ko)^2 \text{Var}(\hat{\sigma}) \quad (1.22) \\ &\cong \frac{\sigma_{x_i}^2}{n} + \frac{(ko)^2 \sigma_{x_i}^4}{2\sigma^2 (n-1)(ko(z_0 - ko) + 1)^2}. \end{aligned}$$

Obteremos agora a variância para $\hat{\rho}$, onde para a obtenção desta variância será suposto que todas as covariâncias existentes são nulas. Temos então:

$$\text{Var}(\hat{\rho}) = \text{Var}([\hat{\beta}^2 [k_0(z_0 - k_0) + 1] - (S_y^2/\hat{\sigma}^2) + 1]^{1/2}) \quad (1.23)$$

e portanto,

$$\text{Var}(\hat{\rho}^2) = \text{Var}([\hat{\beta}^2 [k_0(z_0 - k_0) + 1] - (S_y^2/\hat{\sigma}^2) + 1]). \quad (1.24)$$

Para facilitar a obtenção da variância vamos renomear algumas expressões, ou seja,

$$I = [\hat{\beta}^2 [k_0(z_0 - k_0) + 1], \quad (1.25)$$

$$II = (S_y^2/\hat{\sigma}^2). \quad (1.26)$$

Retomando então a expressão (1.24) e usando as expressões (1.25) e (1.26), temos:

$$\text{Var}(\hat{\rho}^2) = \text{Var}[I - II + 1] \quad (1.27)$$

$$= \text{Var}(I) + \text{Var}(II) - 2 \text{Cov}(I, II).$$

Lembrando que $\text{Cov}(I, II) = 0$ (por suposição), a expressão (1.27) fica assim simplificada:

$$\text{Var}(\hat{\rho}^2) = \text{Var}(I) + \text{Var}(II). \quad (1.28)$$

Obteremos primeiro as variâncias de I e II individualmente e depois faremos a junção destas variâncias. Então,

$$\begin{aligned} \text{Var}(I) &= \text{Var}[\hat{\beta}^2 [k_0(z_0 - k_0) + 1]] \\ &= [k_0(z_0 - k_0) + 1]^2 \text{Var}[\hat{\beta}^2]. \end{aligned} \quad (1.29)$$

Utilizando a aproximação de primeira ordem por Série de Taylor dada por (1.20), adotaremos que:

$$z = \hat{\beta},$$

$$h(z) = \hat{\beta}^2,$$

$$h'(z) = 2\hat{\beta},$$

$$E(z) = E(\hat{\beta}) = \beta,$$

$$\text{Var}(z) = \text{Var}(\hat{\beta}) = \frac{\sigma^2}{\sum_{i=1}^n (x_i - \bar{x})^2}.$$

Portanto, pela aproximação temos que:

$$\begin{aligned} \text{Var}(\hat{\beta}^2) &\cong [2\beta]^2 \text{Var}(\hat{\beta}) \\ &= [2\gamma\rho]^2 \frac{\sigma^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \\ &= \frac{4\gamma^2\rho^2\sigma^2}{\sum_{i=1}^n (x_i - \bar{x})^2}. \end{aligned} \quad (1.30)$$

Retomando a expressão (1.29) e utilizando o resultado obtido em (1.30) temos que:

$$\begin{aligned} \text{Var (I)} &= [ko (zo - ko) + 1]^2 \text{Var} [\hat{\beta}^2] & (1.31) \\ &\cong [ko (zo - ko) + 1]^2 \frac{4\gamma^2 \rho^2 \sigma^2}{\sum_{i=1}^n (x_i - \bar{x})^2} . \end{aligned}$$

Obteremos agora a variância para a expressão II (dada em (1.26)):

$$\begin{aligned} \text{Var (II)} &= \text{Var} (S_y^2 / \hat{\sigma}^2) & (1.32) \\ &= \text{Var} (S_y^2 \frac{(ko(zo - ko) + 1)}{S_x^2}) \\ &= (ko(zo-ko)+1)^2 \text{Var} (\frac{S_y^2}{S_x^2}) . \end{aligned}$$

Para obter a variância da expressão $\frac{S_y^2}{S_x^2}$, utilizaremos uma aproximação de primeira ordem por Série de Taylor para duas variáveis (ver por exemplo Guidorizzi, 1986), que seguindo raciocínio análogo ao desenvolvido em (1.20), é dada por:

$$\text{Var}(g(m)) \cong 2 g^{(1)} g^{(2)} \text{Cov} (m_1, m_2) + (g^{(1)})^2 \text{Var} (m_1) + (g^{(2)})^2 \text{Var} (m_2) , \quad (1.33)$$

onde:

$$\underline{m} = (m_1, m_2) = (S_X^2, S_Y^2),$$

$$g(\underline{m}) = \frac{S_Y^2}{S_X^2},$$

$$E(\underline{m}) = E(S_X^2, S_Y^2) = (\sigma_{X^2}^2, \sigma_{Y^2}^2),$$

$$\text{Var}(m_1) = \text{Var}(S_X^2) = \frac{2\sigma_{X^2}^4}{(n-1)} \quad (\text{ver resultado dado em (1.17)}),$$

$$\text{Var}(m_2) = \text{Var}(S_Y^2) = \frac{2\sigma_{Y^2}^4}{(n-1)} \quad (\text{idem ao item anterior}),$$

onde

$$g^{(1)} = g'(\underline{m}) = \left. \frac{\partial g(\underline{m})}{\partial m_1} \right|_{E(\underline{m})} = -\frac{\sigma_{Y^2}^2}{\sigma_{X^2}^4},$$

$$g^{(2)} = g'(\underline{m}) = \left. \frac{\partial g(\underline{m})}{\partial m_2} \right|_{E(\underline{m})} = \frac{1}{\sigma_{X^2}^2}.$$

Lembrando que $\text{Cov}(m_1, m_2) = 0$ (por suposição), temos que a expressão (1.33)

pode ser escrita como:

$$\text{Var}(g(\underline{m})) \cong (g^{(1)})^2 \text{Var}(m_1) + (g^{(2)})^2 \text{Var}(m_2), \quad (1.34)$$

$$\begin{aligned} \text{Var}\left(\frac{S_Y^2}{S_X^2}\right) &\cong \left(-\frac{\sigma_{Y^2}^2}{\sigma_{X^2}^4}\right)^2 \frac{2\sigma_{X^2}^4}{(n-1)} + \left(\frac{1}{\sigma_{X^2}^2}\right)^2 \frac{2\sigma_{Y^2}^4}{(n-1)} \\ &= \frac{4\sigma_{Y^2}^4}{\sigma_{X^2}^4 (n-1)}. \end{aligned}$$

Retomando a expressão (1.32) e utilizando (1.34) temos então:

$$\begin{aligned} \text{Var}(\Pi) &= [ko(zo - ko) + 1]^2 \text{Var}\left(\frac{S_Y^2}{S_X^2}\right) \\ &\cong [ko(zo - ko) + 1]^2 \frac{4\sigma_{Y/X}^4}{\sigma_{X_i}^4(n-1)}. \end{aligned} \quad (1.35)$$

Então reunindo $\text{Var}(I)$ (dada por (1.31)) e $\text{Var}(\Pi)$ (dada por (1.35)) na expressão dada por (1.28) teremos o seguinte:

$$\begin{aligned} \text{Var}(\hat{\rho}^2) &= \text{Var}(I) + \text{Var}(\Pi) \\ &\cong [ko(zo - ko) + 1]^2 \left\{ \frac{4\gamma^2 \rho^2 \sigma^2}{\sum_{i=1}^n (x_i - \bar{x})^2} + \frac{4\sigma_{Y/X}^4}{\sigma_{X_i}^4(n-1)} \right\}. \end{aligned} \quad (1.36)$$

Utilizando mais uma vez a aproximação de primeira ordem por Série de Taylor (dada em (1.20)) para a expressão (1.23) e o resultado obtido em (1.36), temos que :

$$\begin{aligned} \text{Var}(\hat{\rho}) &\cong \left[\frac{1}{2|\rho|} \right]^2 \text{Var}(\hat{\rho}^2) \\ &\cong \left[\frac{1}{4\rho^2} \right] [ko(zo - ko) + 1]^2 \left\{ \frac{4\gamma^2 \rho^2 \sigma^2}{\sum_{i=1}^n (x_i - \bar{x})^2} + \frac{4\sigma_{Y/X}^4}{\sigma_{X_i}^4(n-1)} \right\} \\ &= \frac{\gamma^2 \sigma^2 (ko(zo - ko) + 1)^2}{\sum_{i=1}^n (x_i - \bar{x})^2} + \frac{\sigma_{Y/X}^4}{\rho^2 \sigma^4(n-1)}. \end{aligned} \quad (1.37)$$

Finalizando, falta obter a variância de $\hat{\gamma}$. Então, utilizando a aproximação dada por

(1.33), temos que:

$$\underline{m} = (m_1, m_2) = (\hat{\rho}, \hat{\beta}),$$

$$g(\underline{m}) = \frac{\hat{\beta}}{\hat{\rho}},$$

$$E(\underline{m}) = E(\hat{\rho}, \hat{\beta}) = (\rho, \beta),$$

$$\text{Var}(m_1) = \text{Var}(\hat{\rho}),$$

$$\text{Var}(m_2) = \text{Var}(\hat{\beta}),$$

com

$$g^{(1)} = g'(E(m)) = \frac{\partial g(m)}{\partial m_1}(E(m)) = -\frac{\beta}{\rho^2} = -\frac{\gamma\rho}{\rho^2} = -\frac{\gamma}{\rho}$$

e

$$g^{(2)} = g'(E(m)) = \frac{\partial g(m)}{\partial m_2}(E(m)) = \frac{1}{\rho}$$

Então lembrando que $\text{Cov}(m_1, m_2) = 0$ (por suposição) temos que:

$$\text{Var}(g(m)) \cong (g^{(1)})^2 \text{Var}(m_1) + (g^{(2)})^2 \text{Var}(m_2), \quad (1.38)$$

ou seja,

$$\begin{aligned}
\text{Var}(\hat{\gamma}) &\cong \left(-\frac{\gamma}{\rho}\right)^2 \text{Var}(\hat{\rho}) + \left(\frac{1}{\rho}\right)^2 \text{Var}(\hat{\beta}) \\
&= \frac{\gamma^2}{\rho^2} \left(\frac{\gamma^2 \sigma^2 (ko(z_0 - ko) + 1)^2}{\sum_{i=1}^n (x_i - \bar{x})^2} + \frac{\sigma_{Y|X}^4}{\rho^2 \sigma^4 (n-1)} \right) + \\
&\quad \frac{1}{\rho^2} \left(\frac{\sigma^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \right) \\
&= \frac{\gamma^4 \sigma^2 (ko(z_0 - ko) + 1)^2}{\rho^2 \sum_{i=1}^n (x_i - \bar{x})^2} + \frac{\gamma^2 \sigma_{Y|X}^4}{\rho^4 \sigma^4 (n-1)} + \frac{\sigma^2}{\rho^2 \sum_{i=1}^n (x_i - \bar{x})^2} .
\end{aligned}$$

Resumindo os resultados obtidos em (1.18), (1.22), (1.37) e (1.38) para as variâncias dos estimadores obtidos através do método dos momentos (James (1973)), temos o seguinte:

$$\text{Var}(\hat{\mu}) \cong \frac{\sigma_{X_i}^2}{n} + \frac{(ko)^2 \sigma_{X_i}^4}{2\sigma^2 (n-1)(ko(z_0 - ko) + 1)^2} , \quad (1.39)$$

$$\text{Var}(\hat{\sigma}^2) = \frac{2\sigma_{X_i}^4}{(k_0(z_0 - k_0) + 1)^2 (n-1)} , \quad (1.40)$$

$$\text{Var}(\hat{\rho}) \cong \frac{\gamma^2 \sigma^2 (ko(z_0 - ko) + 1)^2}{\sum_{i=1}^n (x_i - \bar{x})^2} + \frac{\sigma_{Y|X}^4}{\rho^2 \sigma^4 (n-1)} , \quad (1.41)$$

$$\begin{aligned}
\text{Var}(\hat{\gamma}) &\cong \frac{\gamma^4 \sigma^2 (ko(z_0 - ko) + 1)^2}{\rho^2 \sum_{i=1}^n (x_i - \bar{x})^2} + \frac{\gamma^2 \sigma_{Y|X}^4}{\rho^4 \sigma^4 (n-1)} + \\
&\quad \frac{\sigma^2}{\rho^2 \sum_{i=1}^n (x_i - \bar{x})^2} .
\end{aligned} \quad (1.42)$$

Obs: As variâncias apresentadas pelas expressões (1.41) e (1.42) não apresentam a mesma forma do artigo de James (1973) pois foram simplificadas. Com alguma manipulação algébrica é possível mostrar que elas são idênticas.

Estimadores para as variâncias teóricas apresentadas nas expressões (1.39), (1.40), (1.41) e (1.42) podem ser obtidos substituindo-se os parâmetros por suas respectivas estimativas obtidas a partir das expressões (1.11), (1.12), (1.13) e (1.14).

1.4 Exemplo

Para ilustrar o método proposto por James (1973), apresentaremos um exemplo estudado naquele artigo. No exemplo, foi considerado um estudo que pretendia avaliar o efeito de uma cirurgia associada a certa enfermidade em pacientes que apresentavam pressão sangüínea alta devido a uma patologia renovascular. Foram considerados no exemplo 341 pacientes, onde foram medidas as pressões sangüíneas antes e depois da intervenção cirúrgica. Assumiu-se que a proporção da população que foi truncada é de 25% e para este conjunto de dados obteve-se as seguintes medidas:

$$\bar{x} = 110.1,$$

$$S_x^2 = 171.61,$$

$$S_y^2 = 184.96,$$

$$\hat{\beta} = 0.265.$$

Utilizando-se o método do momentos (expressões dadas por (1.11) a (1.14)) com base nas medidas citadas acima, obtidas a partir dos dados observados, e considerando-se as expressões para as variâncias das estimativas apresentadas em (1.39) a (1.42), obtemos a Tabela 1.1.

Tabela 1.1 - Resultados do método de momentos (James (1973))

parâmetros	μ	σ^2	ρ	γ
estimativa	76.0	720.8	0.87	0.30
variância	2.21	39.09	0.0032	0.128

Considerando o ajuste do modelo pode-se perceber que o valor médio apresenta decréscimo substancial ($0.261 = 0.87 \times 0.30$) devido ao efeito da cirurgia, quando comparado aqueles que teriam apenas um decréscimo natural devido ao efeito de regressão (0.87) para o valor médio. Pelo dados da Tabela 1.1, rejeita-se a hipótese $H_0 : \gamma = 1$, pois um intervalo com aproximadamente 95 % de confiança para γ é dado por $0.30 \pm 2 \sqrt{0.128}$, que claramente não contém $\gamma = 1$.

1.5 Críticas

Senn e Brown (1985) criticam o método dos momentos descrito na seção (1.3) pois as variâncias dos estimadores apresentadas naquela seção são calculadas com base em certas suposições que podem ser questionadas. Estas suposições são:

$$\text{i) } \frac{(n-1)S_x^2}{\sigma_{x_1}^2} \sim \chi^2_{(n-1)},$$

ii) \bar{x} e S_x^2 são independentes ,

$$\text{iii) } \text{Var}(\hat{\beta}) = \frac{\sigma^2}{\sum_{i=1}^n (x_i - \bar{x})^2} .$$

Os resultados i) e ii) seriam válidos se x (ou a medida antes do tratamento) tivesse uma distribuição normal usual. Contudo, convém lembrar que estamos trabalhando com a distribuição normal truncada (apenas são tratados pacientes com valor da característica de interesse acima do valor de truncamento), de modo que os resultados apresentados acima não são válidos.

Com relação ao resultado iii), James (1973) utiliza a variância populacional para a variável x sem truncamento ao invés de utilizar a variância associada a distribuição de y dado x dada pelo modelo (1.3) (no caso, $\text{Var}(y/x) = \sigma^2(1 - \rho^2)$). Além disso, durante o cálculo de obtenção das variâncias são feitas sucessivas aproximações por Série de Taylor e é suposto que as covariâncias existentes são todas nulas.

Conseqüentemente as variâncias apresentadas por James (1973) para os estimadores obtidos através do método dos momentos podem apresentar aproximações inadequadas para as verdadeiras variâncias. Expressões mais adequadas poderiam ser obtidas levando-se em conta a distribuição da normal truncada, considerando-se a distribuição correta de $\hat{\beta}$, e calculando-se aproximações para as covariâncias consideradas nulas. Porém, para cada uma dessas considerações existe uma certa complexidade para ser levada em conta, principalmente para os parâmetros de maior interesse, o que já não torna o método atraente.

Com base neste fato Senn e Brown (1985) propõem uma análise alternativa ao método dos momentos, ou seja, o método de máxima verossimilhança, método este que será abordado no capítulo 2 deste trabalho.

Capítulo 2

Uma nova proposta

2.1 O método de máxima verossimilhança

Considerando os problemas abordados no capítulo anterior, Senn e Brown (1985) propõem uma nova abordagem para o problema de ausência de grupo controle em ensaios experimentais. Esta abordagem consiste em utilizar a função de verossimilhança associada a amostra observada e encontrar os estimadores de máxima verossimilhança através de métodos iterativos, pois como será visto, para o modelo (1.3) em questão (descrito na seção 1.2), não existe uma solução explícita para as equações de verossimilhança.

O objetivo deste capítulo é apresentar as verossimilhanças obtidas por Cohen (1955) para a normal bivariada truncada no que diz respeito a variável x e que foram utilizadas em conjunto com o modelo (1.3) por Senn e Brown (1985). Além disso, apresentaremos estimativas para as variâncias assintóticas dos estimadores e discutiremos os aspectos computacionais relacionados com a implementação dos métodos iterativos .

2.2 A Função de Verossimilhança

Associada com a regressão linear simples com distribuição normal bivariada truncada, Cohen (1955) considera 4 possibilidades diferentes para a seleção da amostra e discute para cada uma delas a função de verossimilhança decorrente. Os 4 tipos de amostras citadas por Cohen (1955), estão descritos abaixo, segundo a ótica do nosso problema:

i) Amostra truncada: o processo de coleta de dados ocorre até que sejam observados n indivíduos apresentando valores, para a característica de interesse x , maiores que um valor de truncamento escolhido (x_0); para estes indivíduos é aplicado o tratamento em questão e são registrados apenas os dados dos indivíduos que se submeteram ao tratamento (medida antes do tratamento, x , onde $x \geq x_0$ e medida posterior ao tratamento, y). Contudo, não se tem nenhum registro a respeito de quantos indivíduos foram descartados, ou seja, indivíduos com valores $x < x_0$.

ii) Amostra censurada: neste tipo de amostragem são consideradas as medidas antes (x ; $x \geq x_0$) e depois do tratamento (y) apenas para os indivíduos que se submeteram ao tratamento (x ; $x \geq x_0$). No entanto, neste caso é conhecido o número de indivíduos que foram descartados pelo truncamento ($x < x_0$), onde este número é denotado por m .

iii) **Amostra selecionada** : neste tipo de amostragem são considerados todos os dados ($N = n + m$ elementos) referentes a variável de interesse antes do tratamento (x), ou seja, $x < x_0$ (com m elementos) ou $x \geq x_0$ (com n elementos), e também serão registrados os valores após o tratamento (y), para os indivíduos que submeteram-se ao tratamento, ou seja, indivíduos com $x \geq x_0$.

iv) **Amostra completa** : neste tipo de amostragem são considerados todos os valores (independentemente do valor de truncamento) para as medidas x e y (antes e depois do tratamento). Contudo, este tipo de amostragem não é interessante para o problema em questão, pois não haveria o interesse em princípio em acompanhar pessoas supostamente sadias.

Cohen (1955) apresenta as funções de verossimilhança associadas a cada um dos 4 tipos de amostragem citados anteriormente, levando em conta a seguinte estrutura:

$$f(x,y) = f(x) f(y/x), \quad (2.1)$$

onde :

$f(x,y)$: função densidade de probabilidade conjunta para as variáveis x e y ,

$f(x)$: função densidade de probabilidade associada a variável x ,

$f(y/x)$: função densidade de probabilidade condicional de y dado x .

Cohen (1955) supõe que $f(x,y)$ provém de uma distribuição normal bivariada, e portanto a estrutura $(f(y/x))$ se enquadraria no modelo de regressão para média dado por (1.1). Então, de acordo com o processo de amostragem definido com base no truncamento, Cohen (1955) associa uma $f(x)$ adequada, pois o processo de amostragem interfere somente na estrutura probabilística da variável x , definindo assim uma verossimilhança apropriada a cada tipo de amostragem, para a densidade conjunta $f(x,y)$.

Notemos que o modelo (1.3) considerado por James (1973) não pode ser colocado na estrutura de uma distribuição normal bivariada, a não ser quando $\gamma = 1$. Contudo, podemos considerar cada uma das distribuições no lado direito de (2.1) como uma distribuição normal univariada, onde $f(x)$ está associada a distribuição da quantidade de interesse antes do tratamento e $f(y/x)$ está associada a distribuição da quantidade de interesse após o tratamento. Senn e Brown (1985) utilizam as expressões dadas em Cohen (1955), levando em conta o modelo (1.3) e as suposições a ele associadas (ver seção 1.2), apresentando a seguinte função de verossimilhança para o modelo:

$$L(\underline{\theta}) = G \times \exp\left[-\frac{1}{2} \sum_{i=1}^n (x_i - \mu)^2 / \sigma^2\right] / (2\pi\sigma^2)^{n/2} \quad (2.2)$$

$$\times \exp\left\{-\frac{1}{2} \sum_{i=1}^n [(y_i - \mu) - \gamma\rho(x_i - \mu)]^2 / [\sigma^2(1 - \rho^2)]\right\} / [2\pi\sigma^2(1 - \rho^2)]^{n/2},$$

onde G (vide tabela 2.1) é uma função que depende do tipo de amostragem a ser utilizado.

Tabela 2.1 - Função G

TIPO DE AMOSTRA	G
Truncada	$\{1/[1 - \Phi(z_0)]\}^n$
Censurada	$[\Phi(z_0)]^m$
Selecionada	$[1/(2\pi\sigma^2)]^{m/2} \exp[-\frac{1}{2} \sum_{c=1}^m (x_c - \mu)^2 / \sigma^2]$

Conforme enfatizado acima, não foi considerada a verossimilhança associada a amostra completa, dado que tal amostra não seria de interesse no caso de ensaios clínicos sem grupo controle.

Com base nas verossimilhanças descritas pela expressão (2.2) e utilizando-se a função G adequada, pode-se implementar através de métodos iterativos a maximização da função de verossimilhança correspondente, encontrando-se possíveis estimadores para os parâmetros do modelo (1.3) descrito na seção (1.2).

2.3 Variância dos Estimadores

Após obter os estimadores de máxima verossimilhança através de métodos iterativos, podemos obter estimativas para as variâncias destes estimadores, utilizando o inverso da matriz de Informação Observada. Nesta seção iremos descrever como obter tais estimativas.

Em primeiro lugar precisamos relembrar algumas definições e resultados importantes, que garantem a utilização da matriz de Informação Observada como forma de se obter estimativas para a variância dos estimadores de máxima de verossimilhança. Temos então:

log-verossimilhança (ℓ) : Dado $\underline{\theta}$ (vetor de parâmetros), seja $L(\underline{\theta})$ uma função de verossimilhança (como a descrita em (2.2)) ; então a log-verossimilhança é simplesmente o logaritmo neperiano de L, ou seja, $\ell(\underline{\theta}) = \log L(\underline{\theta})$.

função score (u) : A função (vetor) score é definida como a primeira derivada da log-verossimilhança em relação ao vetor de parâmetros de interesse $\underline{\theta}$. Então,

$$\underline{u}(\underline{\theta}) = \partial \ell(\underline{\theta}) / \partial \underline{\theta} \quad ; \quad (2.3)$$

matriz de Informação Esperada ou de Fisher (I_F) : A matriz de Informação Esperada é definida como a matriz que contém a esperança do negativo das segundas derivadas da log-verossimilhança em relação ao vetor de parâmetros $\underline{\theta}$. Então,

$$I_F(\underline{\theta}) = E [- \frac{\partial^2 \ell(\underline{\theta})}{\partial \underline{\theta} \partial \underline{\theta}'}] ; \quad (2.4)$$

matriz de Informação Observada (I_O) : A matriz de Informação Observada é definida como a matriz que contém o negativo das segundas derivadas da log-verossimilhança avaliada no vetor de parâmetros estimados pelo método de máxima verossimilhança ($\hat{\underline{\theta}}$).

Então,

$$I_O(\hat{\underline{\theta}}) = - \frac{\partial^2 \ell(\underline{\theta})}{\partial \underline{\theta} \partial \underline{\theta}'} \Bigg|_{\underline{\theta} = \hat{\underline{\theta}}} . \quad (2.5)$$

Com base nas definições acima vamos citar alguns resultados importantes (ver por exemplo Dudewicz e Mishra (1988) e Sen e Singer (1993)) que serão utilizados:

1) Sendo θ : parâmetro (escalar) e $\hat{\theta}$: estimador de máxima verossimilhança de θ temos que:

$$n^{-1} I_O(\hat{\theta}) \xrightarrow{P} I_F(\theta) . \quad (2.6)$$

2) Sendo g : função contínua e tendo $A \xrightarrow{P} B$, temos que:

$$g(A) \xrightarrow{P} g(B). \quad (2.7)$$

3) Sendo $\hat{\underline{\theta}}$ o estimador de máxima verossimilhança de $\underline{\theta}$ (vetor de parâmetros), temos que:

$$\sqrt{n} (\hat{\underline{\theta}} - \underline{\theta}) \xrightarrow{D} N(0, I_F^{-1}(\underline{\theta})). \quad (2.8)$$

Portanto, por (2.8) a matriz de covariâncias assintótica de $\hat{\underline{\theta}}$ é dada por:

$$\text{Var}_A(\hat{\underline{\theta}}) = \frac{1}{n} I_F^{-1}(\underline{\theta}). \quad (2.9)$$

Seja $\underline{\theta}$ (vetor de parâmetros), temos então que: se $I_F(\underline{\theta})_{ij}$ é o elemento (i, j) da matriz de Informação Esperada ou de Fisher $I_F(\underline{\theta})$ (definida em (2.4)) e sendo $I_O(\hat{\underline{\theta}})_{ij}$ o correspondente elemento (i, j) na matriz de Informação Observada $I_O(\hat{\underline{\theta}})$ (definida em (2.5)), iremos supor com base em (2.6) que:

$$n^{-1} I_O(\hat{\underline{\theta}})_{ij} \xrightarrow{P} I_F(\underline{\theta})_{ij}, \text{ para todo } i \text{ e } j. \quad (2.10)$$

Portanto, usando (2.10) temos que :

$$n^{-1} I_O(\hat{\underline{\theta}}) \xrightarrow{P} I_F(\underline{\theta}). \quad (2.11)$$

Então adotando g : função inversa, $A = n^{-1} I_O(\hat{\underline{\theta}})$, $B = I_F(\underline{\theta})$ e considerando o resultado dado em (2.11), teremos por (2.7) que:

$$(n^{-1} I_O(\hat{\underline{\theta}}))^{-1} \xrightarrow{P} (I_F(\underline{\theta}))^{-1}. \quad (2.12)$$

Neste caso, lembrando (2.9) e usando (2.12) temos que:

$$\text{Var}_\lambda(\hat{\underline{\theta}}) = \frac{1}{n} I_F^{-1}(\underline{\theta}) \cong \frac{1}{n} (n I_O^{-1}(\hat{\underline{\theta}})) = I_O^{-1}(\hat{\underline{\theta}}). \quad (2.13)$$

Portanto, com base nos resultados descritos acima temos que as estimativas aproximadas da variância assintótica dos estimadores de máxima verossimilhança podem ser obtidas através da matriz de Informação Observada dada em (2.5).

Após a obtenção dos estimadores de máxima verossimilhança (por métodos iterativos) e suas respectivas variâncias (obtidas por (2.13)), pode-se então testar hipóteses sobre os parâmetros de interesse (no caso γ e ρ) através de Intervalos de Confiança ou por meio de testes de hipótese como o teste da Razão de Verossimilhança (ver por exemplo Dudewicz e Mishra ,1988).

2.4 Aspectos Computacionais

Lembrando que qualquer que seja o tipo de amostragem utilizado (ver tipos de amostras na seção 2.2) para o modelo dado por (1.3) não podem ser encontradas soluções explícitas para os estimadores de máxima verossimilhança (vide expressão (2.2) e possíveis funções G). Então, qualquer que seja o tipo de amostra, algum método numérico deve ser utilizado para a obtenção dos estimadores de máxima verossimilhança. Neste caso é importante ressaltar algumas questões que são inerentes aos aspectos computacionais:

- Escolher um bom algoritmo de maximização, por exemplo, o método de Newton-Rapson (algoritmo padrão);

- Escolhido o algoritmo existem outros aspectos computacionais que precisam ser levados em conta como:

- **Valor inicial:** Todo algoritmo precisa de valores iniciais para os parâmetros envolvidos no modelo, onde a importância destes valores está associada a *rapidez da convergência* do algoritmo;

- **Convergência:** Os valores obtidos pelos algoritmos numa dada iteração são comparados com os valores da iteração anterior, e considera-se que o procedimento convergiu quando a diferença absoluta entre a estimativa atual e a estimativa anterior para os parâmetros são pequenas (critério fixado a priori). No entanto, *nem sempre existe a convergência ou muitas vezes ela ocorre lentamente.*

- **Custos:** A utilização de qualquer método iterativo requer a utilização de recursos computacionais, onde um fator que pode comprometer a utilização de qualquer método numérico é *o tempo necessário para conseguir a convergência das estimativas.*

Portanto, para obter os estimadores de máxima verossimilhança para os parâmetros do modelo definido em (1.3) teremos que considerar todos os aspectos citados acima.

Senn e Brown (1989) propõem generalizações do modelo (1.3) (truncamento inferior e superior) e apresentam um modelo aditivo para o problema da ausência de grupo controle. Além disso eles consideram o problema do valor inicial principalmente para o caso das amostras truncada e censurada, pois eles não acham razoável a suposição feita por James (1973) de que a proporção de população que foi truncada seja conhecida pelo pesquisador.

Com a finalidade de uma futura implementação numérica encontram-se disponíveis no Anexo I deste trabalho as primeiras e segundas derivadas para as log-verossimilhanças definidas pelo tipo de amostra selecionada.

Considerando os problemas apresentados pelos métodos computacionais, principalmente pelo custo computacional que qualquer método iterativo encerra, Chen e Cox (1992) conseguiram uma aproximação através do Princípio da Pseudo-Verossimilhança para o caso de amostra selecionada, apresentando soluções fechadas para os estimadores. Este método será abordado no próximo capítulo deste trabalho.

Capítulo 3

Uma solução aproximada para o caso da amostra selecionada

Conforme visto na seção anterior, o método de máxima verossimilhança considerado por Senn e Brown (1985) apresenta dificuldades quanto a obtenção das estimativas, onde se destaca a dificuldade com relação aos aspectos computacionais, como: a proposta para os valores iniciais; convergência; custo computacional; etc. Procurando contornar estas dificuldades, Chen e Cox (1992) apresentam uma solução para um caso particular, ou seja, no caso da amostra selecionada. Tal solução utiliza uma abordagem através do Princípio da Pseudo - Verossimilhança (Gong e Samaniego, 1981).

Neste capítulo apresentaremos o Princípio da Pseudo - Verossimilhança e o seu desenvolvimento para o caso da amostra selecionada apresentado por Chen e Cox (1992), no caso de ensaios do tipo “screening trials”, ou seja, em estudos envolvendo um grande número de participantes onde uma sub-amostra dos indivíduos é selecionada através do truncamento para receber o tratamento de interesse.

Apenas para ilustrar a situação suponha, por exemplo, que exista o interesse de algum grande centro de pesquisa médica em avaliar algum tipo de tratamento para a taxa de colesterol. Com esta finalidade, será realizado um ensaio clínico com 10.000 pacientes onde será medida a taxa de colesterol de todos os pacientes e serão submetidos ao tratamento todos aqueles que tiverem taxa de colesterol muito alta (valor fixado por critérios médicos). Após o tratamento será feita uma nova medição quanto a taxa de colesterol dos pacientes submetidos ao tratamento e tentará se avaliar se o tratamento foi eficiente ou não. Esta é uma situação onde o método proposto por Chen e Cox (1992) poderia ser aplicado.

3.1 O Princípio da Pseudo - Verossimilhança

O método de máxima verossimilhança é um dos mais conhecidos e utilizados na estimação dos parâmetros de um modelo estatístico, devido ao fato de apresentar boas propriedades assintóticas, como consistência, eficiência e normalidade assintótica. No entanto, este método pode apresentar problemas de estimação principalmente quando não existem soluções explícitas para os estimadores, ou seja, quando é necessária a utilização de algum método numérico de maximização ou quando existem parâmetros “ nuisance “ (ou de incômodo) a serem estimados.

Na tentativa de resolver os problemas de estimação na presença de parâmetros “nuisance“ muitas abordagens tem sido propostas. Uma abordagem possível é a utilização do Princípio da Pseudo - Verossimilhança, que consiste basicamente em substituir os parâmetros “nuisance “ por estimativas consistentes e resolver as equações do sistema reduzidas aos parâmetros de interesse. Este método e alguns resultados a respeito encontram-se discutidos em Delgado (1995).

Então baseando-se no Princípio da Pseudo - Verossimilhança Chen e Cox (1992) apresentam uma simplificação para os estimadores de máxima verossimilhança propostos por Senn e Brown (1985), no caso do modelo (1.3). Portanto, no sentido de ilustrar uma possível aplicação do Princípio da Pseudo - Verossimilhança iremos apresentá-lo na próxima seção.

3.2 Estimadores de máxima verossimilhança

Considerando o caso onde a amostra é do tipo selecionada, temos de (2.2) com a função G adequada para este tipo de amostragem, a seguinte expressão para a função de verossimilhança:

$$L(\underline{\theta}) = [1 / (2\pi\sigma^2)]^{m/2} \exp[-\frac{1}{2} \sum_{c=1}^m (x_c - \mu)^2 / \sigma^2] \times \quad (3.1)$$

$$[1 / (2\pi\sigma^2)]^{n/2} \exp[-\frac{1}{2} \sum_{i=1}^n (x_i - \mu)^2 / \sigma^2] \times$$

$$\exp\{-\frac{1}{2} \sum_{i=1}^n [(y_i - \mu) - \gamma\rho(x_i - \mu)]^2 / [\sigma^2(1 - \rho^2)]\} / [2\pi\sigma^2(1 - \rho^2)]^{n/2}.$$

Inicialmente para desenvolver o cálculo dos estimadores iremos supor que temos os parâmetros “ nuisances “ μ e σ^2 conhecidos. Com esta suposição inicial e considerando que iremos derivar a log-verossimilhança ($\ell(\underline{\theta})$) em função dos parâmetros de interesse (ρ e γ) obtemos a seguinte simplificação para a expressão (3.1), em termos da log-verossimilhança:

$$\ell(\underline{\theta}) \propto -\frac{n}{2} \log(1 - \rho^2) \quad (3.2)$$

$$-\frac{1}{2} \sum_{i=1}^n [(y_i - \mu) - \gamma\rho(x_i - \mu)]^2 / [\sigma^2(1 - \rho^2)].$$

Calculando a primeira derivada da log-verossimilhança dada por (3.2), encontramos o vetor escore (definido em (2.3)) através do qual podemos obter os estimadores de máxima verossimilhança para este modelo. Então, para a log-verossimilhança apresentada em (3.2) temos as seguintes equações (componentes) para a função escore:

$$\frac{\partial \ell(\theta)}{\partial \rho} = \frac{n\rho}{(1-\rho^2)} + \frac{\gamma}{\sigma^2(1-\rho^2)} \sum_{i=1}^n (x_i - \mu)[(y_i - \mu) - \gamma\rho(x_i - \mu)] \quad (3.3)$$

$$- \frac{\rho}{\sigma^2(1-\rho^2)^2} \sum_{i=1}^n [(y_i - \mu) - \gamma\rho(x_i - \mu)]^2$$

e

$$\frac{\partial \ell(\theta)}{\partial \gamma} = \frac{\rho}{\sigma^2(1-\rho^2)} \sum_{i=1}^n (x_i - \mu)[(y_i - \mu) - \gamma\rho(x_i - \mu)]. \quad (3.4)$$

Resolvendo o sistema composto pelas equações (3.3) e (3.4), encontramos os estimadores de máxima verossimilhança para o modelo proposto com amostragem selecionada. No entanto, para resolver o sistema acima é importante lembrar que:

$$\hat{\beta} = \hat{\gamma}\hat{\rho} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}, \quad (3.5)$$

que é o estimador usual (por mínimos quadrados) de β (coeficiente angular) na regressão linear simples.

Então, usando a expressão (3.5) e resolvendo o sistema formado por (3.3) e (3.4) teremos os seguintes estimadores de máxima verossimilhança para os parâmetros de interesse:

$$\hat{\rho} = \pm \sqrt{\left\{1 - \frac{\sum_{i=1}^n [(y_i - \mu) - \hat{\beta}(x_i - \mu)]^2}{n\sigma^2}\right\}} \quad (3.6)$$

e

$$\hat{\gamma} = \frac{\hat{\beta}}{\hat{\rho}}, \quad (3.7)$$

onde $\hat{\rho}$ deve apresentar o mesmo sinal que $\hat{\beta}$.

Considerando que estes são os estimadores de máxima verossimilhança para os parâmetros de interesse podemos encontrar a variância dos estimadores através da matriz de Informação Observada conforme descrito na seção 2.3 do capítulo 2 deste trabalho. Com base na função ou vetor escore podemos obter as segundas derivadas, que são dadas pelas expressões abaixo:

$$\frac{\partial^2 \ell(\theta)}{\partial \rho^2} = \frac{n(1 + \rho^2)}{(1 - \rho^2)^2} - \frac{\gamma^2}{\sigma^2(1 - \rho^2)} \sum_{i=1}^n (x_i - \mu)^2 \quad (3.8)$$

$$+ \frac{4\gamma\rho}{\sigma^2(1 - \rho^2)^2} \sum_{i=1}^n (x_i - \mu)[(y_i - \mu) - \gamma\rho(x_i - \mu)]$$

$$- \frac{(1 + 3\rho^2)}{\sigma^2(1 - \rho^2)^3} \sum_{i=1}^n [(y_i - \mu) - \gamma\rho(x_i - \mu)]^2,$$

$$\frac{\partial^2 \ell(\theta)}{\partial \gamma^2} = -\frac{\rho^2}{\sigma^2(1 - \rho^2)} \sum_{i=1}^n (x_i - \mu)^2 \quad (3.9)$$

e

$$\frac{\partial^2 \ell(\theta)}{\partial \rho \partial \gamma} = \frac{1}{\sigma^2(1 - \rho^2)^2} \left\{ (1 + \rho^2) \sum_{i=1}^n (x_i - \mu)(y_i - \mu) - 2\gamma\rho \sum_{i=1}^n (x_i - \mu)^2 \right\}, \quad (3.10)$$

lembrando que a expressão (3.5) pode ser escrita da seguinte forma:

$$\sum_{i=1}^n (x_i - \mu)(y_i - \mu) = \hat{\gamma} \hat{\rho} \sum_{i=1}^n (x_i - \mu)^2. \quad (3.11)$$

Então utilizando o resultado (3.11) e tomando $\underline{\theta} = \underline{\hat{\theta}}$ obtemos as seguintes expressões (simplificadas) para as segundas derivadas dadas por (3.8), (3.9) e (3.10):

$$\begin{aligned} \frac{\partial^2 \ell(\underline{\hat{\theta}})}{\partial \hat{\rho}^2} &= \frac{n(1 + \hat{\rho}^2)}{(1 - \hat{\rho}^2)^2} + \frac{\hat{\gamma}^2}{\sigma^2(1 - \hat{\rho}^2)} \sum_{i=1}^n (x_i - \mu)^2 \\ &\quad - \frac{(1 + 3\hat{\rho}^2)}{\sigma^2(1 - \hat{\rho}^2)^3} \sum_{i=1}^n [(y_i - \mu) - \hat{\gamma} \hat{\rho}(x_i - \mu)]^2, \end{aligned} \quad (3.12)$$

$$\frac{\partial^2 \ell(\underline{\hat{\theta}})}{\partial \hat{\gamma}^2} = -\frac{\hat{\rho}^2}{\sigma^2(1 - \hat{\rho}^2)} \sum_{i=1}^n (x_i - \mu)^2 \quad (3.13)$$

e

$$\frac{\partial^2 \ell(\underline{\hat{\theta}})}{\partial \hat{\rho} \partial \hat{\gamma}} = -\frac{\hat{\gamma} \hat{\rho}}{\sigma^2(1 - \hat{\rho}^2)} \sum_{i=1}^n (x_i - \mu)^2. \quad (3.14)$$

Portanto, temos que uma estimativa para a matriz de covariância de $(\hat{\gamma}$ e $\hat{\rho})$ é dada por:

$$Var_A(\hat{\theta}) \cong I_O^{-1}(\hat{\theta}) = \begin{pmatrix} I_{11} & I_{12} \\ I_{21} & I_{22} \end{pmatrix}^{-1}, \quad (3.15)$$

onde os termos da matriz de Informação Observada $I_O(\hat{\theta})$, são obtidos utilizando (3.12), (3.13) e (3.14) conforme o esquema abaixo:

$$I_{11} = - \frac{\partial^2 \ell(\hat{\theta})}{\partial \rho^2}, \quad (3.16)$$

$$I_{12} = I_{21} = - \frac{\partial^2 \ell(\hat{\theta})}{\partial \rho \partial \gamma}, \quad (3.17)$$

e

$$I_{22} = - \frac{\partial^2 \ell(\hat{\theta})}{\partial \gamma^2}. \quad (3.18)$$

Então utilizando essas estimativas (obtidas através de (3.15), (3.16), (3.17) e (3.18)) podemos construir Intervalos de Confiança e Testes de hipótese (como por exemplo o teste da Razão de Verossimilhança) para os parâmetros de interesse γ e ρ .

No entanto, cabe lembrar que os parâmetros μ e σ^2 não são conhecidos como foi suposto em princípio. Portanto, é neste ponto que entra o Princípio da Pseudo-Verossimilhança descrito na seção (3.1) deste capítulo. Chen e Cox (1992) propõem a utilização de estimadores conhecidamente consistentes para μ e σ^2 (pensando na distribuição marginal de x) através de sua substituição na função de verossimilhança apresentada em (3.1) como se fossem os verdadeiros valores para os parâmetros. Os estimadores propostos são dados por:

$$\hat{\mu} = \sum_{i=1}^{n+m} \frac{x_i}{n+m} \quad (3.19)$$

e

$$\hat{\sigma}^2 = \sum_{i=1}^{n+m} \frac{(x_i - \hat{\mu})^2}{n+m-1} \quad (3.20)$$

Enfatizamos que Chen e Cox (1992) recomendam este procedimento apenas para estudos de grande porte, ou seja, onde o número de elementos que participam do ensaio na primeira etapa seja relativamente grande em relação ao número de pacientes que efetivamente recebem o tratamento. Então, é razoável aceitar que $\hat{\mu}$ e $\hat{\sigma}^2$ (obtidos por (3.19) e (3.20)) estejam próximos dos verdadeiros valores de μ e σ^2 . Após a substituição de $\hat{\mu}$ e $\hat{\sigma}^2$ na expressão (3.1) como se fossem os verdadeiros valores para os parâmetros, todo o procedimento desenvolvido neste capítulo poderá ser realizado com base na Pseudo-Verossimilhança, ou seja, é possível realizar a estimação e a inferência para os parâmetros γ e ρ .

3.3 Considerações

Chen e Cox (1992) obtém através da utilização do Princípio da Pseudo - Verossimilhança uma simplificação para a verossimilhança dada por (3.1), e desta forma conseguiram simplificar a obtenção dos estimadores de pseudo - máxima verossimilhança (foram obtidas soluções fechadas para estes estimadores) no caso da amostra selecionada para o modelo dado por (1.3) que foi definido por James (1973) e cujas verossimilhanças foram apresentadas por Senn e Brown (1985).

É preciso não esquecer que Chen e Cox (1992) recomendam a utilização deste Princípio no caso de “ screening trials “ (estudos de separação), ou seja, em estudos envolvendo um grande número de participantes, no sentido de se obter boas aproximações (através de estimadores consistentes) para μ e σ^2 .

Visando avaliar a eficiência através do Erro Quadrático Médio (E.Q.M.) dos três métodos de estimação descritos durante este trabalho, foi realizado um estudo de simulação cujos resultados serão apresentados no próximo capítulo.

Capítulo 4

Estudos de simulação

4.1 O método dos momentos

Na primeira parte deste estudo, procuramos avaliar o efeito nos estimadores do método de momentos quando um valor incorreto para o ponto de truncamento é escolhido. Apesar de em muitas situações o pesquisador dispor de alguma informação sobre este valor não parece correto afirmar que ele disponha do valor correto desta quantidade. Então, tentando verificar o quanto um valor errôneo para a porcentagem de truncamento compromete as estimativas obtidas pelo método dos momentos proposto por James (1973), realizou-se um estudo de simulação. É preciso ressaltar que estas simulações têm um caráter meramente ilustrativo, pois não foi estabelecido nenhum método de comparação das estimativas, e elas foram realizadas apenas no intuito de se ter uma idéia do quanto o ponto de truncamento altera as estimativas obtidas.

Com base nos motivos descritos acima implementou-se um programa (vide Anexo II) no aplicativo S-plus e foi feito um estudo de simulação com os seguintes parâmetros: $\mu = 76$, $\sigma^2 = 4$, $\rho = 0.8$ e $\gamma = 0.3$. Neste estudo, foram considerados os seguintes pontos de truncamento: 5 %, 10 %, 15 %, 20 % e 25 % e os seguintes tamanhos de amostra: 200, 500, 1000 e 2000. Para ilustrar os resultados obtidos através deste estudo de simulação iremos apresentar *algumas tabelas* com os resultados das simulações:

Tabela 4.1 - Resultados da simulação com N = 2000 e truncamento = 5 %

Truncamento	$\hat{\mu}$	$\hat{\sigma}^2$	$\hat{\rho}$	$\hat{\gamma}$
1 %	73.79081	5.580237	0.7930551	0.2395627
2 %	74.74838	4.862461	0.7577344	0.2507295
3 %	75.30610	4.442815	0.7307112	0.2600020
4 %	75.70060	4.145066	0.7074325	0.2685576
5 %	76.00586	3.914048	0.6862773	0.2768362
6 %	76.25478	3.725207	0.6664524	0.2850712
7 %	76.46486	3.565452	0.6474980	0.2934162
8 %	76.64657	3.426974	0.6291158	0.3019896
9 %	76.80662	3.304737	0.6110955	0.3108948
10 %	76.94962	3.195303	0.5932782	0.3202315
11 %	77.07882	3.096222	0.5755363	0.3301032
12 %	77.19665	3.005686	0.5577615	0.3406230
13 %	77.30494	2.922319	0.5398566	0.3519201
14 %	77.40510	2.845055	0.5217305	0.3641466
15 %	77.49828	2.773047	0.5032932	0.3774865
16 %	77.58536	2.705614	0.4844521	0.3921676
17 %	77.66711	2.642198	0.4651080	0.4084781
18 %	77.74412	2.582336	0.4451504	0.4267915
19 %	77.81693	2.525642	0.4244522	0.4476038
20 %	77.88595	2.471788	0.4028616	0.4715922

obs: o valor que aparece em negrito corresponde as estimativas com o verdadeiro valor do truncamento, enquanto os outros valores apresentados são as estimativas para possíveis “ chutes “ para o valor de truncamento nas proximidades do verdadeiro valor .

Tabela 4.2 - Resultados da simulação com N = 2000 e truncamento = 10 %

Truncamento	$\hat{\mu}$	$\hat{\sigma}^2$	$\hat{\rho}$	$\hat{\gamma}$
1 %	71.98010	8.181697	0.8595211	0.2443743
2 %	73.13958	7.129299	0.8367890	0.2510129
3 %	73.81490	6.514017	0.8196948	0.2562476
4 %	74.29259	6.077461	0.8051904	0.2608636
5 %	74.66222	5.738744	0.7921964	0.2651424
6 %	74.96362	5.461867	0.7801896	0.2692228
7 %	75.21801	5.227636	0.7688712	0.2731860
8 %	75.43803	5.024600	0.7580518	0.2770851
9 %	75.63183	4.845377	0.7476023	0.2809580
10 %	75.80498	4.684926	0.7374297	0.2848337
11 %	75.96143	4.539655	0.7274641	0.2887357
12 %	76.10410	4.406911	0.7176509	0.2926839
13 %	76.23522	4.284679	0.7079460	0.2966962
14 %	76.35651	4.171395	0.6983130	0.3007890
15 %	76.46933	4.065819	0.6887211	0.3049781
16 %	76.57478	3.966949	0.6791431	0.3092792
17 %	76.67376	3.873968	0.6695553	0.3137080
18 %	76.76702	3.786200	0.6599359	0.3182807
19 %	76.85517	3.703075	0.6502647	0.3230144
20 %	76.93875	3.624115	0.6405228	0.3279272

Tabela 4.3 - Resultados da simulação com N = 2000 e truncamento = 15 %

Truncamento	$\hat{\mu}$	$\hat{\sigma}^2$	$\hat{\rho}$	$\hat{\gamma}$
5 %	74.19725	5.666092	0.8018842	0.3522796
6 %	74.49673	5.392720	0.7905200	0.3573438
7 %	74.74951	5.161454	0.7798180	0.3622479
8 %	74.96813	4.960989	0.7695977	0.3670585
9 %	75.16070	4.784035	0.7597366	0.3718228
10 %	75.33275	4.625615	0.7501464	0.3765764
11 %	75.48821	4.482183	0.7407607	0.3813478
12 %	75.62998	4.351119	0.7315280	0.3861608
13 %	75.76026	4.230435	0.7224068	0.3910365
14 %	75.88078	4.118586	0.7133631	0.3959939
15 %	75.99288	4.014345	0.7043680	0.4010509
16 %	76.09766	3.916727	0.6953964	0.4062250
17 %	76.19602	3.824924	0.6864265	0.4115334
18 %	76.28868	3.738266	0.6774383	0.4169936
19 %	76.37628	3.656194	0.6684137	0.4226236
20 %	76.45932	3.578234	0.6593358	0.4284424
21 %	76.53827	3.503981	0.6501883	0.4344702
22 %	76.61351	3.433086	0.6409557	0.4407284
23 %	76.68536	3.365249	0.6316230	0.4472406
24 %	76.75412	3.300206	0.6221749	0.4540322
25 %	76.82005	3.237726	0.6125962	0.4611315

Tabela 4.4 - Resultados da simulação com N = 2000 e truncamento = 20 %

Truncamento	$\hat{\mu}$	$\hat{\sigma}^2$	$\hat{\rho}$	$\hat{\gamma}$
10 %	74.91090	4.833930	0.7492271	0.3108611
11 %	75.06982	4.684039	0.7397999	0.3148224
12 %	75.21474	4.547072	0.7305258	0.3188191
13 %	75.34793	4.420953	0.7213630	0.3228687
14 %	75.47113	4.304066	0.7122773	0.3269872
15 %	75.58573	4.195131	0.7032397	0.3311894
16 %	75.69285	4.093117	0.6942251	0.3354900
17 %	75.79339	3.997179	0.6852113	0.3399033
18 %	75.88812	3.906619	0.6761784	0.3444440
19 %	75.97766	3.820851	0.6671080	0.3491272
20 %	76.06256	3.739380	0.6579832	0.3539689
21 %	76.14327	3.661782	0.6487875	0.3589859
22 %	76.22018	3.587695	0.6395054	0.3641964
23 %	76.29363	3.516803	0.6301215	0.3696201
24 %	76.36393	3.448831	0.6206204	0.3752786
25 %	76.43132	3.383537	0.6109869	0.3811957
26 %	76.49604	3.320709	0.6012053	0.3873977
27 %	76.55829	3.260157	0.5912595	0.3939143
28 %	76.61826	3.201713	0.5811327	0.4007786
29 %	76.67610	3.145226	0.5708073	0.4080284
30 %	76.73196	3.090562	0.5602645	0.4157064

Tabela 4.5 - Resultados da simulação com N = 2000 e truncamento = 25 %

Truncamento	$\hat{\mu}$	$\hat{\sigma}^2$	$\hat{\rho}$	$\hat{\gamma}$
15 %	75.08437	4.920200	0.7664554	0.2019514
16 %	75.20037	4.800554	0.7597183	0.2037423
17 %	75.30926	4.688034	0.7530095	0.2055575
18 %	75.41185	4.581822	0.7463152	0.2074013
19 %	75.50882	4.481230	0.7396228	0.2092779
20 %	75.60077	4.385678	0.7329210	0.2111916
21 %	75.68817	4.294669	0.7261993	0.2131464
22 %	75.77146	4.207777	0.7194479	0.2151466
23 %	75.85101	4.124632	0.7126578	0.2171965
24 %	75.92714	4.044912	0.7058201	0.2193006
25 %	76.00012	3.968333	0.6989265	0.2214636
26 %	76.07021	3.894645	0.6919686	0.2236904
27 %	76.13763	3.823628	0.6849383	0.2259864
28 %	76.20257	3.755083	0.6778276	0.2283571
29 %	76.26521	3.688834	0.6706282	0.2308086
30 %	76.32571	3.624721	0.6633321	0.2333473
31 %	76.38421	3.562602	0.6559308	0.2359803
32 %	76.44083	3.502347	0.6484158	0.2387153
33 %	76.49570	3.443837	0.6407781	0.2415606
34 %	76.54892	3.386966	0.6330085	0.2445255
35 %	76.60059	3.331635	0.6250974	0.2476202

Conforme pode ser observado nas tabelas 4.1 a 4.5 nota-se que um pequeno afastamento do verdadeiro valor do truncamento não altera de forma significativa as estimativas obtidas pelo método dos momentos. Portanto, parece ser razoável considerar valores (sugestões) para esta porcentagem quando o pesquisador dispõe de suficiente informação (ou conhecimento) sobre a população. Contudo, lembramos que o estudo apresentado é meramente ilustrativo e que foram apresentadas apenas algumas tabelas para as simulações pois as outras apresentaram resultados bastante similares.

4.2 Comparação entre os métodos

Com a finalidade de comparar os 3 métodos descritos anteriormente, através do Erro Quadrático Médio (E.Q.M.), foi realizado um estudo de simulação. Neste estudo foi implementado um programa através do aplicativo S-plus (vide Anexo II), onde este programa apresenta a estimativa para os 3 métodos (momentos, verossimilhança e pseudo-verossimilhança) e calcula os respectivos E.Q.M.'s. No caso, do método de máxima verossimilhança é preciso lembrar que não existe solução explícita para as equações de verossimilhança (vide Anexo I), portanto foi adotado um método numérico de maximização, sendo escolhido o método de Newton-Rapson por ser um algoritmo padrão. Neste método foi adotado como valor inicial para os parâmetros as estimativas obtidas pelo método dos momentos. O processo de simulação e os resultados obtidos encontram-se a seguir:

Em primeiro lugar foram escolhidos 2 conjuntos de parâmetros definidos abaixo:

Grupo 1: Conjunto de parâmetros (escolha arbitrária) utilizados durante o desenvolvimento do programa de simulação.

$$\mu = 10 \quad \sigma^2 = 1 \quad \rho = 0,8 \quad \gamma = 0,2$$

Grupo 2 : Conjunto de parâmetros utilizados no artigo de Senn e Brown (1989).

$$\mu = 0 \quad \sigma^2 = 1 \quad \rho = 0,7 \quad \gamma = 0,7$$

Com base nestes 2 grupos de parâmetros foram geradas amostras, considerando-se os seguintes aspectos: *tamanho de amostra e porcentagem de truncamento da população*. Para tamanho de amostra foram considerados 4 valores: 1000, 5000, 10000 e 15000. Agora para a porcentagem de truncamento foram considerados 2 valores: 1% e 5%. Então, considerando as variações possíveis que totalizam 8 para cada grupo de parâmetros, foram geradas 1000 amostras para cada variação, sendo que para cada amostra foram obtidos os estimadores dos parâmetros pelos 3 métodos e através das 1000 amostras foram estimados os E.Q.M.'s. Os resultados obtidos por este estudo de simulação encontram-se nas tabelas 4.6 e 4.7 apresentadas a seguir:

Tabela 4.6 - Tabela do Erro Quadrático Médio para os 3 métodos (grupo1)

Porcentagem de truncamento	Tamanho da amostra	μ			σ^2			ρ			γ		
		momento	verossim	pseudo	momento	verossim	pseudo	momento	verossim	pseudo	momento	verossim	pseudo
1 %	1.000	0,80225	0,00129	0,00103	0,70570	0,00157	0,00201	0,02266	0,07050	0,51546	0,64709	0,00464	0,24146
	5.000	0,16130	0,00022	0,00020	0,11023	0,00040	0,00039	0,01685	0,01838	0,38535	0,04224	0,00122	0,11476
	10.000	0,08117	0,00010	0,00010	0,05429	0,00022	0,00021	0,01745	0,01809	0,24574	0,01809	0,00107	0,02032
	15.000	0,05541	0,00006	0,00006	0,03724	0,00014	0,00014	0,01757	0,01802	0,15287	0,01238	0,00098	0,01700
5 %	1.000	0,08057	0,00097	0,00094	0,09734	0,00205	0,00195	0,01680	0,01779	0,32512	0,02537	0,00169	0,05514
	5.000	0,01628	0,00019	0,00019	0,01945	0,00041	0,00042	0,01754	0,01777	0,02397	0,00511	0,00102	0,00522
	10.000	0,00804	0,00009	0,00009	0,00963	0,00021	0,00021	0,01756	0,01769	0,01648	0,00320	0,00090	0,00320
	15.000	0,00517	0,00007	0,00007	0,00601	0,00013	0,00013	0,01767	0,01771	0,01694	0,00233	0,00086	0,00233

Tabela 4.7 - Tabela do Erro Quadrático Médio para os 3 métodos (grupo2)

Porcentagem de truncamento	Tamanho da amostra	μ			σ^2			P			γ		
		momento	verossim	pseudo	momento	verossim	pseudo	momento	verossim	pseudo	momento	verossim	pseudo
		1 %	1.000	0,64245	0,00091	0,00091	0,54599	0,00134	0,00220	0,03855	0,04795	0,04131	1,30586
	5.000	0,15623	0,00021	0,00021	0,10578	0,00045	0,00043	0,02580	0,02820	0,02977	0,11000	0,01925	0,50846
	10.000	0,08245	0,00010	0,00010	0,05328	0,00020	0,00019	0,02444	0,02689	0,02383	0,05375	0,01813	0,40967
	15.000	0,05460	0,00007	0,00007	0,03549	0,00014	0,00014	0,02562	0,03452	0,01932	0,04190	0,02331	0,08168
5 %	1.000	0,07590	0,00093	0,00095	0,09534	0,00203	0,00213	0,02680	0,02839	0,02459	0,07916	0,02029	0,43310
	5.000	0,01544	0,00019	0,00019	0,01862	0,00040	0,00040	0,02550	0,02601	0,02097	0,02800	0,01755	0,02693
	10.000	0,00821	0,00010	0,00010	0,00964	0,00021	0,00021	0,02550	0,02592	0,02321	0,02126	0,01719	0,02051
	15.000	0,00596	0,00007	0,00007	0,00691	0,00014	0,00014	0,02560	0,02581	0,02380	0,01990	0,01717	0,01916

Observando os resultados obtidos neste estudo de simulação que estão apresentados nas tabelas 4.6 e 4.7 podemos notar que:

- Com relação aos parâmetros “ nuisance “ μ e σ^2 , o EQM para o método dos momentos apresenta valores sempre superiores quando comparados com os métodos de máxima verossimilhança e pseudo-verossimilhança, no entanto é preciso não esquecer que no método dos momentos são utilizados somente os dados relativos aos pacientes que se submeteram ao tratamento mesmo no caso de amostra selecionada, então este resultado já era considerado esperado. Quanto aos métodos de máxima verossimilhança e pseudo-verossimilhança eles apresentam resultados bem próximos.

- Já para o efeito de regressão para a média denominado por ρ temos que o método que, em geral, apresenta o menor EQM é o método dos momentos seguido pelo método de máxima verossimilhança e posteriormente pelo método de pseudo-verossimilhança. Notando que, à medida que o tamanho da amostra cresce os valores para o EQM se aproximam para os 3 métodos, atingindo valores bem próximos.

- Agora para o efeito do tratamento (γ) que é o parâmetro de real interesse neste tipo de estudo, podemos observar que, o método de máxima verossimilhança apresenta sempre o menor EQM, seguido do método dos momentos e posteriormente pelo método de pseudo-verossimilhança. Observamos também que a medida que o tamanho da amostra cresce, o EQM para o método de pseudo-verossimilhança se aproxima do EQM para o método dos momentos, reforçando a idéia de que o método de pseudo-verossimilhança só deve ser usado para grandes tamanhos de amostra.

Portanto se o objetivo do ensaio clínico é obter uma melhor estimativa para o parâmetro de tratamento baseado no EQM o método que apresenta as melhores estimativas é o método de máxima verossimilhança, o que teoricamente seria o esperado devido as boas propriedades dos estimadores de máxima verossimilhança em relação aos outros 2 métodos apresentados.

Contudo, existe aqui um fator que ainda não foi apresentado, e trata-se do *número de amostras aproveitadas*, ou seja, ao estimarmos o EQM foram utilizadas apenas as amostras onde foi possível obter estimativas para todos os parâmetros do modelo. Então, foram desprezadas no método dos momentos e pseudo-verossimilhança as amostras onde haviam indefinições (por exemplo: raiz quadrada de um número negativo) e no caso de máxima verossimilhança foram desprezadas as amostras que não convergiram pelo método de Newton-Rapson.

Então neste caso é preciso avaliar também a taxa de aproveitamento das amostras, ou seja, em relação às 1000 amostras geradas para cada uma das 8 variações de cada grupo de parâmetros, quantas efetivamente contribuíram para a estimativa do EQM de cada método? Com a finalidade de avaliar este problema apresentamos a seguir nas tabelas 4.8 e 4.9 a taxa de aproveitamento das amostras (em porcentagem) para cada um dos métodos e em cada grupo de parâmetros.

Tabela 4.8 - Tabela de avaliação do desempenho dos 3 métodos (grupo1)

Porcentagem de truncamento	Tamanho da amostra	Pessoas Tratadas (*)	taxa de aproveitamento das amostras (%)		
			momento	verossim.	pseudo-veros.
1 %	1.000	10	97,9	6,1	60,4
	5.000	50	100,0	30,8	93,9
	10.000	100	100,0	49,2	99,6
	15.000	150	100,0	60,3	100,0
5 %	1.000	50	100,0	47,5	99,5
	5.000	250	100,0	89,5	100,0
	10.000	500	100,0	96,0	100,0
	15.000	750	100,0	99,7	100,0

(*): Este número é o valor aproximado de pessoas que receberam o tratamento em cada variação do estudo, já que o número exato de pessoas selecionadas para o tratamento varia de amostra para amostra.

Tabela 4.9 - Tabela de avaliação do desempenho dos 3 métodos (grupo2)

Porcentagem de truncamento	Tamanho da amostra	Pessoas Tratadas (*)	taxa de aproveitamento das amostras (%)		
			momento	verossim.	pseudo-veros.
1 %	1.000	10	90,9	8,1	39,7
	5.000	50	100,0	31,7	82,3
	10.000	100	100,0	46,9	94,7
	15.000	150	100,0	60,0	97,8
5 %	1.000	50	100,0	46,6	96,0
	5.000	250	100,0	82,6	100,0
	10.000	500	100,0	90,3	100,0
	15.000	750	100,0	93,5	100,0

(*): Este número é o valor aproximado de pessoas que receberam o tratamento em cada variação do estudo, já que o número exato de pessoas selecionadas para o tratamento varia de amostra para amostra.

Observando as tabelas 4.8 e 4.9 podemos notar que :

- Para pequenos tamanhos de amostra (pequeno número de pessoas tratadas) a taxa de aproveitamento das amostras principalmente pelo método de verossimilhança é muito pequena (abaixo de 10 %), ou seja, para a maioria das amostras não foi possível conseguir uma estimativa por este método com o programa utilizado. O método que apresentou a melhor taxa de aproveitamento foi o método dos momentos, onde mesmo com pequenos tamanhos de amostra conseguiu-se obter as estimativas para a maioria das amostras geradas (acima de 90 %).

- À medida que o tamanho da amostra aumenta a taxa de aproveitamento para o método de verossimilhança e de pseudo-verossimilhança vai crescendo e então, para grandes tamanhos de amostras consegue-se taxas de aproveitamento próximas do método dos momentos.

Portanto, este é um fator que precisa ser levado em conta ao ser adotado qualquer método, pois na prática o pesquisador disporá apenas de *uma única amostra* para a estimativa dos seus parâmetros.

CONCLUSÃO

Com base nos métodos apresentados na primeira parte deste trabalho nos Capítulos 1, 2 e 3 e no estudo de simulação desenvolvido no Capítulo 4, podemos apresentar algumas conclusões:

No caso da amostra selecionada, podemos concluir através do estudo de simulação que o método de máxima verossimilhança é o que apresentou as melhores estimativas, em relação ao EQM, para o parâmetro ligado ao efeito de tratamento. Contudo, ele apresenta sérios problemas de convergência para pequenos tamanhos de amostra quando calculado através do método de Newton-Rapson. Então, sempre que for possível obter uma estimativa por este método ela deve ser utilizada.

No caso de pequenas amostras, o método mais indicado é o método dos momentos pois ele apresenta a melhor taxa de aproveitamento das amostras, lembrando que o método de máxima verossimilhança pode ser utilizado como uma melhoria da estimativa obtida pelo método dos momentos nos casos onde for possível obter tal estimativa.

Já para grandes tamanhos de amostra, a questão que se coloca é a disponibilidade de recursos computacionais, pois embora o avanço computacional seja considerável atualmente nem sempre podemos contar com a facilidade de programação e implementação de métodos numéricos de maximização, requerido pelo método de máxima verossimilhança. No caso de haver restrição de uso de um suporte computacional o método de pseudo-verossimilhança apresenta uma solução simples devido a facilidade de obtenção dos estimadores, visto que, este método apresenta fórmulas fechadas para os estimadores dos parâmetros de interesse e que poderiam portanto serem implementadas facilmente numa simples planilha de cálculo. Além disso, apesar deste método apresentar um EQM similar

ao método dos momentos ele apresenta melhores propriedades assintóticas e conta com toda a estrutura teórica para inferência dos estimadores de máxima verossimilhança. É preciso lembrar ainda que, embora o método dos momentos apresente boas estimativas para os parâmetros de interesse, ele apresenta dificuldades quanto ao cálculo de estimativas para as variâncias de seus estimadores (vide críticas apresentadas no Capítulo 1 deste trabalho).

No caso de outros tipos de amostras (truncada e censurada), o método dos momentos sempre pode ser utilizado, pois ele não leva em conta o tipo de amostragem utilizado. Quanto ao método de máxima verossimilhança, a função de verossimilhança para estes 2 tipos de amostras foi apresentada no Capítulo 2 deste trabalho, sendo necessário implementar algum método numérico de maximização, e não esquecendo que existe a dificuldade adicional de se estimar a função densidade de probabilidade para a distribuição Normal com μ e σ^2 desconhecidos.

Em relação ao trabalho desenvolvido durante esta dissertação existe a possibilidade de explorar em novos trabalhos alguns aspectos, como:

- Avaliar o método de máxima verossimilhança para as amostras truncada e censurada e tentar verificar a possível utilização do método de pseudo-verossimilhança para estes tipos de amostragem.

- No caso de pequenas amostras tentar explorar outros métodos numéricos de maximização que tenham uma maior taxa de aproveitamento das amostras, visto que o método de Newton-Rapson não se mostrou tão eficiente.

ANEXO I

DERIVADAS PARA A LOG – VEROSSIMILHANÇA

Neste anexo encontram-se disponíveis as derivadas para a log-verossimilhança no caso da amostra selecionada (vide expressão de verossimilhança (2.2) e G adequada). Tais derivadas foram obtidas através do aplicativo Maple, devido a facilidade de obtenção de derivadas neste aplicativo. Portanto as derivadas apresentadas a seguir representam as saídas obtidas através do Maple:

1) EQUAÇÃO DA LOG - VEROSSIMILHANÇA ($\ell(\underline{\theta})$) PARA AMOSTRA SELECIONADA :

$$eq = \ell(\underline{\theta}) = \log(L(\underline{\theta}))$$

$$eq := - \frac{1}{2} m \ln(2 \text{ Pi } s^2) - \frac{1}{2} \frac{\sum_{c=1}^m \left(\frac{x[c]}{s} + \mu \right)^2}{s^2}$$

$$- \frac{1}{2} n \ln(2 \text{ Pi } s^2) - \frac{1}{2} \frac{\sum_{t=1}^n \left(\frac{x[t]}{s} + \mu \right)^2}{s^2}$$

$$- \frac{1}{2} n \ln(2 \pi s^2 (1 - r)) - \frac{1}{2} \left(\frac{y[t]}{\sqrt{s^2 (1 - r)}} + \frac{y[t] - \mu}{\sqrt{s^2 (1 - r)}} \right)^2$$

$$+ \frac{1}{2} \left(\frac{g r x[t] - 2 y[t] g r x[t]}{\sqrt{s^2 (1 - r)}} + \frac{2 y[t] g r \mu + n \mu}{\sqrt{s^2 (1 - r)}} \right)^2$$

$$+ \frac{1}{2} \left(\frac{2 \mu g r x[t] - 2 n g r \mu + g r x[t]}{\sqrt{s^2 (1 - r)}} \right)^2$$

$$+ \frac{1}{2} \left(\frac{- 2 g r x[t] \mu + n g r \mu}{\sqrt{s^2 (1 - r)}} \right)^2$$

onde:

- mu = μ x[c]: vetor de dados do grupo de pessoas cuja
- s2 = σ^2 medida ficou abaixo do valor de truncamento.
- r = ρ
- g = γ x[t]: vetor de dados do grupo de pessoas cuja
- Pi = π medida ficou acima do valor de truncamento.

2) Componentes da Função Escore (primeira derivada da log-verossimilhança):

$$eq1 = \frac{\partial \ell(\theta)}{\partial \mu}$$

Pelo comando abaixo obtemos eq1 no Maple:

```
> eq1:=diff(eq,mu);
```

$$eq1 := -\frac{1}{2} \frac{\sqrt{\frac{\sqrt{c-1} \sqrt{-2x[c]} + 2m\mu}{s^2}}}{s^2} - \frac{1}{2} \frac{\sqrt{\frac{\sqrt{t-1} \sqrt{-2x[t]} + 2n\mu}{s^2}}}{s^2} - \frac{1}{2} \left(\frac{\sqrt{\frac{\sqrt{n} \sqrt{-2y[t]} + \sqrt{\frac{2y[t]gr}{s^2}} + 2n\mu + \sqrt{\frac{2grx[t]}{s^2}}}{s^2}}}{s^2} \right)$$

$$-4ngr\mu + \frac{\sqrt{\frac{\sqrt{n} \sqrt{-2grx[t]} + 2n\sqrt{gr}\mu}{s^2}}}{s^2} / (s^2(1-r)^2)$$

$$\text{eq2} = \frac{\partial \ell(\theta)}{\partial \sigma^2}$$

> eq2:=diff(eq,s2);

$$\text{eq2} := -\frac{1}{2} \frac{\frac{\sum_{c=1}^m \frac{1}{x[c]} + \frac{\sum_{c=1}^m x[c] - 2 \sum_{c=1}^m \mu}{\sum_{c=1}^m \mu}}{s2^2} + \frac{1}{2} \frac{\frac{\sum_{t=1}^n \frac{1}{y[t]} + \frac{\sum_{t=1}^n y[t] - 2 \sum_{t=1}^n \mu}{\sum_{t=1}^n \mu}}{s2^2}}$$

$$+ \frac{1}{2} \frac{\frac{\sum_{t=1}^n \frac{1}{y[t]} + \frac{\sum_{t=1}^n y[t] - 2 \sum_{t=1}^n \mu}{\sum_{t=1}^n \mu}}{s2^2} + \frac{1}{2} \left(\frac{\sum_{t=1}^n \frac{1}{y[t]} + \frac{\sum_{t=1}^n y[t] - 2 \sum_{t=1}^n \mu}{\sum_{t=1}^n \mu}}{s2^2} \right)$$

$$+ \frac{\sum_{t=1}^n \frac{1}{y[t]} + \frac{\sum_{t=1}^n y[t] - 2 \sum_{t=1}^n \mu}{\sum_{t=1}^n \mu}}{s2^2} - 2 y[t] \mu + \frac{\sum_{t=1}^n \frac{1}{y[t]} + \frac{\sum_{t=1}^n y[t] - 2 \sum_{t=1}^n \mu}{\sum_{t=1}^n \mu}}{s2^2} - 2 y[t] g r x[t] + \frac{\sum_{t=1}^n \frac{1}{y[t]} + \frac{\sum_{t=1}^n y[t] - 2 \sum_{t=1}^n \mu}{\sum_{t=1}^n \mu}}{s2^2} - 2 y[t] g r \mu$$

$$\frac{\sqrt{n} \left(\frac{2 \mu g r x[t]}{\sqrt{t}} - 2 n g r \mu \right) + \sqrt{n} \left(\frac{g r x[t]}{\sqrt{t}} \right)}{\sqrt{t}}$$

$$\frac{\sqrt{n} \left(\frac{-2 g r x[t] \mu}{\sqrt{t}} + n g r \mu \right) / (s^2 (1 - r))}{\sqrt{t}}$$

$$\text{eq3} = \frac{\partial \ell(\theta)}{\partial \rho}$$

> eq3:=diff(eq,r);

$$\text{eq3} := \frac{n r \left(\frac{y[t]}{\sqrt{t}} \right) + \left(\frac{-2 y[t] \mu}{\sqrt{t}} \right)}{1 - r}$$

$$\frac{\sqrt{n} \left(\frac{-2 y[t] g r x[t]}{\sqrt{t}} + \left(\frac{2 y[t] g r \mu}{\sqrt{t}} + n \mu \right) \right)}{\sqrt{t}}$$

$$\frac{\sqrt{n}}{\sqrt{\mu g r x(t) - 2 n g r \mu}} + \frac{\sqrt{n}}{\sqrt{g r x(t)}}$$

$$\frac{\sqrt{n}}{\sqrt{-2 g r x(t) \mu + n g r \mu}} \cdot \frac{r}{(s^2 (1 - r))} - \frac{1}{2}$$

$$\frac{\sqrt{n}}{\sqrt{-2 y(t) g x(t)}} + \frac{\sqrt{n}}{\sqrt{2 y(t) g \mu}} + \frac{\sqrt{n}}{\sqrt{2 \mu g x(t)}}$$

$$- 2 n g \mu + \frac{\sqrt{n}}{\sqrt{2 g r x(t)}} + \frac{\sqrt{n}}{\sqrt{-4 g r x(t) \mu}}$$

$$\frac{2}{+ 2 n g r \mu} \cdot \frac{2}{(s^2 (1 - r))}$$

$$\text{eq4} = \frac{\partial \ell(\theta)}{\partial \gamma}$$

> eq4:=diff(eq,g);

$$\text{eq4} := -\frac{1}{2} \left(\frac{\sqrt{n}}{\sqrt{t-1}} \left(\frac{\sqrt{y[t] r x[t]}}{\sqrt{t-1}} + \frac{\sqrt{2 y[t] r \mu}}{\sqrt{t-1}} \right) \right)$$

$$+ \left(\frac{\sqrt{n}}{\sqrt{t-1}} \left(\frac{\sqrt{2 \mu r x[t]}}{\sqrt{t-1}} - \frac{2 n r \mu}{\sqrt{t-1}} + \frac{\sqrt{2 g r x[t]}}{\sqrt{t-1}} \right) \right)$$

$$\left(\frac{\sqrt{n}}{\sqrt{t-1}} \left(\frac{\sqrt{-4 g r x[t] \mu}}{\sqrt{t-1}} + \frac{2 n g r \mu}{\sqrt{t-1}} \right) \right) / (s2 (1 - r))$$

3) MATRIZ DE SEGUNDAS DERIVADAS PARA A LOG-VEROSSIMILHANÇA:

$$EQ11 = \frac{\partial^2 \ell(\theta)}{\partial \mu^2}$$

> eq11:=diff(eq1,mu);

$$eq11 := - \frac{m}{s^2} - \frac{n}{s^2} - \frac{1}{2} \frac{2n - 4ng r + 2ng r^2}{s^2(1-r)^2}$$

$$EQ12 = \frac{\partial^2 \ell(\theta)}{\partial \mu \partial \sigma^2}$$

> eq12:=diff(eq1,s2);

$$eq12 := \frac{1}{2} \frac{\left(\frac{m}{s^2} - 2 \frac{x[c]}{s^2} + 2 \frac{\mu}{s^2} \right)}{s^2} + \frac{1}{2} \frac{\left(\frac{n}{s^2} - 2 \frac{x[t]}{s^2} + 2 \frac{\mu}{s^2} \right)}{s^2} + \frac{1}{2} \left(\dots \right)$$

$$\frac{1}{2} \frac{\left(\frac{n}{s^2} - 2 \frac{y[t]}{s^2} + \left(\frac{n}{s^2} - 2 \frac{y[t]}{s^2} \right) g r + 2 \frac{\mu}{s^2} + \left(\frac{n}{s^2} - 2 \frac{g r x[t]}{s^2} \right) \right)}{s^2} + \dots$$

$$-4ng\sqrt{r\mu} + \frac{\sqrt{-2grx[t]} + 2ng\sqrt{r\mu}}{\sqrt{s^2(1-r)}}$$

$$\text{EQ13} = \frac{\partial^2 \ell(\theta)}{\partial \mu \partial \rho}$$

> eq13:=diff(eq1,r);

$$\text{eq13} := - \frac{\sqrt{-2y[t]} + \sqrt{2y[t]}g\sqrt{r} + 2n\mu}{\sqrt{s^2(1-r)}} + \frac{\sqrt{-2y[t]} + \sqrt{2y[t]}g\sqrt{r} + 2n\mu}{\sqrt{s^2(1-r)}}$$

$$-4ng\sqrt{r\mu} + \frac{\sqrt{-2grx[t]} + 2ng\sqrt{r\mu}}{\sqrt{s^2(1-r)}}$$

$$-1/2 \left(\frac{\sqrt{-2y[t]}g + \sqrt{2y[t]}g}{\sqrt{s^2(1-r)}} - 4ng\mu \right)$$

$$\frac{\sqrt{n} \left(\sqrt{-4grx[t]} + \sqrt{4ng r \mu} \right) / (s^2 (1-r))}{\sqrt{t=1}}$$

$$E14 = \frac{\partial^2 \ell(\theta)}{\partial \mu \partial \gamma}$$

> eq14:=diff(eq1,g);

$$\text{eq14} := -1/2 \left(\frac{\sqrt{n} \left(\sqrt{-4grx[t]} + \sqrt{4ng r \mu} \right) / (s^2 (1-r))}{\sqrt{t=1}} + \frac{\sqrt{n} \left(\sqrt{-4grx[t]} + \sqrt{4ng r \mu} \right) / (s^2 (1-r))}{\sqrt{t=1}} \right) - 4ng r \mu$$

$$\frac{\sqrt{n} \left(\sqrt{-4grx[t]} + \sqrt{4ng r \mu} \right) / (s^2 (1-r))}{\sqrt{t=1}}$$

$$\text{EQ22} = \frac{\partial^2 \ell(\theta)}{\partial (\sigma^2)^2}$$

```
> eq22:=diff(eq2,s2);
```

$$\text{eq22} := \frac{1}{2} \left[\frac{\frac{\sqrt{m} \sqrt{x[c] + \mu}}{\sqrt{c-1}} - 2 \frac{\sqrt{m} \mu}{\sqrt{c-1}}}{s2^2} + \frac{\frac{\sqrt{m} \sqrt{x[c] + \mu}}{\sqrt{c-1}} - 2 \frac{\sqrt{m} \mu}{\sqrt{c-1}}}{s2^3} + \frac{\frac{\sqrt{n} \sqrt{y[t] + \mu}}{\sqrt{t-1}} - 2 \frac{\sqrt{n} \mu}{\sqrt{t-1}}}{s2^2} \right]$$

$$- \frac{\frac{\sqrt{n} \sqrt{y[t] + \mu}}{\sqrt{t-1}} - 2 \frac{\sqrt{n} \mu}{\sqrt{t-1}}}{s2^3} - \left(\frac{\sqrt{n} \sqrt{y[t] + \mu}}{\sqrt{t-1}} - 2 \frac{\sqrt{n} \mu}{\sqrt{t-1}} \right)$$

$$+ \frac{\frac{\sqrt{n} \sqrt{y[t] + \mu}}{\sqrt{t-1}} - 2 \frac{\sqrt{n} \mu}{\sqrt{t-1}}}{s2^3} - 2 \frac{\sqrt{n} \sqrt{y[t] + \mu}}{\sqrt{t-1}} \frac{\sqrt{n} \sqrt{y[t] + \mu}}{\sqrt{t-1}} + \frac{\sqrt{n} \sqrt{y[t] + \mu}}{\sqrt{t-1}} - 2 \frac{\sqrt{n} \mu}{\sqrt{t-1}}$$

$$\frac{\sqrt{n}}{\sqrt{t=1}} \left(\frac{2 \mu g r x[t]}{\sqrt{t=1}} - 2 n g r \mu \right) + \frac{\sqrt{n}}{\sqrt{t=1}} \left(\frac{g r x[t]}{\sqrt{t=1}} \right)$$

$$\frac{\sqrt{n}}{\sqrt{t=1}} \left(-2 g r x[t] \mu + n g r \mu \right) / (s2 (1 - r))$$

$$\text{EQ23} = \frac{\partial^2 \ell(\theta)}{\partial \sigma^2 \partial \rho}$$

> eq23:=diff(eq2,r);

$$\text{eq23} := \frac{\sqrt{n}}{\sqrt{t=1}} \left(\frac{y[t]}{\sqrt{t=1}} + \frac{-2 y[t] \mu}{\sqrt{t=1}} + \frac{-2 y[t] g r x[t]}{\sqrt{t=1}} \right)$$

$$\frac{\sqrt{n}}{\sqrt{t=1}} \left(\frac{2 y[t] g r \mu}{\sqrt{t=1}} + n \mu \right) + \frac{\sqrt{n}}{\sqrt{t=1}} \left(\frac{2 \mu g r x[t]}{\sqrt{t=1}} - 2 n g r \mu \right)$$

$$\frac{\sqrt{n}}{\sqrt{t=1}} \left| \frac{\sqrt{2g r x(t)} + \sqrt{-2g r x(t) \mu} + n g r \mu}{\sqrt{t=1}} \right|$$

$$\frac{\sqrt{n}}{\sqrt{t=1}} \left| \frac{\sqrt{(s^2(1-r)) + 1/2} \left(\sqrt{-2y(t) g x(t)} + \sqrt{2y(t) g \mu} \right)}{\sqrt{t=1}} \right|$$

$$\frac{\sqrt{n}}{\sqrt{t=1}} \left| \frac{\sqrt{2 \mu g x(t)} - \sqrt{2 n g \mu} + \sqrt{2 g r x(t)}}{\sqrt{t=1}} \right|$$

$$\frac{\sqrt{n}}{\sqrt{t=1}} \left| \frac{\sqrt{-4 g r x(t) \mu} + 2 n g r \mu}{\sqrt{(s^2(1-r))}} \right|$$

$$\text{EQ24} = \frac{\partial^2 \ell(\theta)}{\partial \sigma^2 \partial \gamma}$$

> eq24:=diff(eq2,g);

$$\text{eq24} := \frac{1}{2} \left(\frac{\sqrt{n}}{\sqrt{t-1}} \left(\frac{\partial}{\partial \gamma} \left(\frac{1}{\sigma^2} \right) \right) - 2 y[t] r x[t] \right) + \left(\frac{\sqrt{n}}{\sqrt{t-1}} \right)^2 y[t] r \mu$$

$$+ \left(\frac{\sqrt{n}}{\sqrt{t-1}} \right)^2 \mu r x[t] - 2 n r \mu + \left(\frac{\sqrt{n}}{\sqrt{t-1}} \right)^2 g r x[t]$$

$$+ \frac{\sqrt{n}}{\sqrt{t-1}} \left(-4 g r x[t] \mu + 2 n g r \mu \right) / (s^2 (1-r))$$

$$\text{EQ33} = \frac{\partial^2 \ell(\theta)}{\partial \rho^2}$$

```
> eq33:=diff(eq3,r);
```

$$\text{eq33} := \frac{n}{(1-r)^2} + 2 \frac{n r}{(1-r)^2} - 4 \left(\frac{y[t]}{\sqrt{t}} + \frac{y[t]}{\sqrt{t}} \right) - 2 y[t] \mu$$

$$+ \left(\frac{-2 y[t] g r x[t]}{\sqrt{t}} + \frac{2 y[t] g r \mu}{\sqrt{t}} \right) + n \mu$$

$$+ \left(\frac{2 \mu g r x[t]}{\sqrt{t}} - 2 n g r \mu \right) + \frac{g r x[t]}{\sqrt{t}}$$

$$+ \frac{-2 g r x[t] \mu}{\sqrt{t}} + \frac{n g r \mu}{\sqrt{t}} \frac{r}{(s^2 (1-r))} - 2 \left(\frac{2^2}{\sqrt{t}} \right)$$

$$\frac{\sqrt{n}}{\sqrt{t=1}} \sqrt{\frac{2y[t]gx[t]}{n}} + \frac{\sqrt{n}}{\sqrt{t=1}} \sqrt{\frac{2y[t]g\mu}{n}} + \frac{\sqrt{n}}{\sqrt{t=1}} \sqrt{\frac{2\mu gx[t]}{n}}$$

$$\frac{\sqrt{n}}{\sqrt{t=1}} \sqrt{2ng\mu + \frac{2g^2rx[t]}{n}} + \frac{\sqrt{n}}{\sqrt{t=1}} \sqrt{-4g^2rx[t]\mu}$$

$$\frac{\sqrt{n}}{\sqrt{t=1}} \sqrt{2ng^2rx\mu} + \frac{\sqrt{n}}{\sqrt{t=1}} \sqrt{(s^2(1-r))^2 - \frac{y[t]}{n}}$$

$$\frac{\sqrt{n}}{\sqrt{t=1}} \sqrt{\frac{2y[t]\mu}{n}} + \frac{\sqrt{n}}{\sqrt{t=1}} \sqrt{\frac{2y[t]grx[t]}{n}} + \frac{\sqrt{n}}{\sqrt{t=1}} \sqrt{\frac{2y[t]gr\mu}{n}}$$

$$\frac{\sqrt{n}}{\sqrt{t=1}} \sqrt{2\mu grx[t]} - \frac{\sqrt{n}}{\sqrt{t=1}} \sqrt{2ng^2rx\mu} + \frac{\sqrt{n}}{\sqrt{t=1}} \sqrt{g^2rx[t]}$$

$$\frac{\sqrt{n}}{\sqrt{\left(\frac{1}{2} \right)^2 - 2 g r x[t] \mu + n g r \mu}} \sqrt{\left(\frac{1}{2} \right)^2 - 2 g r \mu} / (s2 (1 - r))$$

$$\frac{\sqrt{n}}{\sqrt{\left(\frac{1}{2} \right)^2 - 2 g x[t] \mu + n g \mu}} \sqrt{\left(\frac{1}{2} \right)^2 - 4 g x[t] \mu + 2 n g \mu} / (s2 (1 - r))$$

$$\text{EQ34} = \frac{\partial^2 \ell(\theta)}{\partial \rho \partial \gamma}$$

> eq34:=diff(eq3,g);

$$\text{eq34} := - \frac{\sqrt{n}}{\sqrt{\left(\frac{1}{2} \right)^2 - 2 y[t] r x[t] \mu + n y[t] r \mu}} \sqrt{\left(\frac{1}{2} \right)^2 - 2 y[t] r \mu} / (s2 (1 - r))$$

$$\frac{-2nr\mu + \sqrt{2grx(t)}}{\sqrt{t=1}} + \frac{\sqrt{2grx(t)}}{\sqrt{t=1}} - \frac{4grx(t)\mu}{\sqrt{t=1}}$$

$$\frac{+2ng r^2 \mu) r / (s^2 (1 - r)) - 1/2 (\sqrt{2 y(t) x(t)})}{\sqrt{t=1}}$$

$$\frac{+ \sqrt{2 y(t) \mu} + \sqrt{2 x(t) \mu} - 2n\mu + \sqrt{4grx(t)}}{\sqrt{t=1}}$$

$$\frac{+ \sqrt{-8\mu grx(t)} + 4ng r^2 \mu) / (s^2 (1 - r))}{\sqrt{t=1}}$$

$$\text{EQ44} = \frac{\partial^2 \ell(\theta)}{\partial \gamma^2}$$

```
> eq44:=diff(eq4,g);
```

```

      / n          \ / n          \
      |-----| |-----|
      | \      2 2| | \      2    |      2 2
      | )  2 r x[t] | + | )  - 4 r x[t] mu | + 2 n r mu
      | /          | | /          |
      |-----| | |-----|
      \t = 1      / \t = 1      /
eq44 := - 1/2 -----
                                 2
                               s2 (1 - r )

```

Portanto, a matriz de segundas derivadas (M) ficou definida pelos elementos dados por:

$$M = \begin{bmatrix} eq11 & & & \\ eq12 & eq22 & & \\ eq13 & eq23 & eq33 & \\ eq14 & eq24 & eq34 & eq44 \end{bmatrix}$$

obs :

1) É IMPORTANTE NÃO ESQUECER QUE A MATRIZ DE SEGUNDAS DERIVADAS PARA A LOG-VEROSSIMILHANÇA É SIMÉTRICA.

2) NO CASO DE AMOSTRAS TRUNCADA E CENSURADA EXISTE UMA DIFICULDADE ADICIONAL PARA SE ENCONTRAR AS DERIVADAS PARA A LOG-VEROSSIMILHANÇA, POIS ELAS DEPENDEM DA FUNÇÃO DE DENSIDADE ACUMULADA (VIDE FUNÇÃO G PARA ESTAS AMOSTRAS) E PORTANTO É NECESSÁRIO SE CONHECER O PONTO z_0 OU OBTER UMA MANEIRA DE ESTIMAR ESTA FUNÇÃO VISTO QUE μ E σ^2 SÃO DESCONHECIDOS.

3) NESTE TRABALHO NÃO NOS PREOCUPAMOS COM AS AMOSTRAS TRUNCADA E CENSURADA POIS O OBJETIVO DO TRABALHO ERA COMPARAR OS 3 MÉTODOS DESCRITOS NO CASO DA AMOSTRA SELECIONADA.

ANEXO II

PROGRAMAS DOS ESTUDOS DE SIMULAÇÃO

Para a realização dos estudos de simulação foram desenvolvidos 2 programas no aplicativo S-plus for windows (versão 3.2) para micro do tipo PC, onde tais programas encontram-se disponíveis neste anexo:

1) ESTUDO DA INFLUÊNCIA DO z_0 PARA O MÉTODO DOS MOMENTOS:

PROGRAMA 1 :

```
# SIMULACAO PARA AVALIAR COMO O AFASTAMENTO DO VERDADEIRO VALOR
# DE  $z_0$  INTERFERE NA ESTIMACAO DOS PARAMETROS ASSOCIADOS AO MODELO
# DE REGRESSAO PARA MEDIA COM EFEITO DE TRATAMENTO (JAMES (1973)),
# POIS NA PRATICA  $z_0$  NAO E CONHECIDO E PORTANTO E IMPORTANTE TER-SE
# UMA IDEIA DE COMO  $z_0$  INTERFERE NAS ESTIMATIVAS DO MODELO.
```

DEFINICAO DOS PARAMETROS POPULACIONAIS:

```
# N      = tamanho da amostra
# med    = media populacional
# dp     = desvio-padrao
# trunc  = % de truncamento
# co     = correlacao populacional
# trat   = efeito do tratamento
```

INICIALIZACAO DOS VALORES:

```
N      <- 0
med    <- 0
dp     <- 0
trunc  <- 0
co     <- 0
trat   <- 0
```

```

# VALORES ASSOCIADOS A SIMULACAO:
l1 <- 0
l2 <- 0
nsim <- 10

# ESCOLHA DOS PARAMETROS:
N <- 2000
med <- 76
dp <- 2
trunc <- 0.80
co <- 0.80
trat <- 0.30
l1 <- 10
l2 <- 30

posm <- 1
s <- 1

simul <- matrix(0,nsim*(l2-l1+1),4)
tam <- vector("numeric", nsim)

for (cl in 1 : nsim)
{

# GERACAO DOS DADOS SEGUNDO O MODELO:
# GERA OS DADOS:

gc <- 0
gt <- 0

amostra <- vector("numeric",N)
amostra <- rnorm(N,med,dp)
amostra <- sort(amostra)

```

```

# ESPECIFICA O PONTO DE TRUNCAMENTO:
xo <- 0
xo <- qnorm(trunc,med,dp)

# SEPARACAO DO GRUPO DE ELEMENTOS NAO TRATADOS:
xinc <- vector ("numeric", N)
i <- 1
while (amostra[i] < xo)
  {xinc[i] <- amostra[i]
  i <- i+1 }
gc <- i-1
xc <- vector("numeric", gc)
for (l in 1 : gc)
  { xc[l] <- xinc[l] }

# SEPARACAO DO GRUPO DE ELEMENTOS TRATADOS:
gt <- N - gc
xt <- vector ("numeric",gt)
k <- 1
for ( j in i : N )
  { xt[k] <- amostra[j]
  k <- k+1 }

# GERACAO DOS VALORES Y SEGUNDO O MODELO ADOTADO:
valory <- vector ("numeric", gt)
Vary <- 0
Vary <- dp^2*(1 - co^2)
for ( l in 1 : gt )
  { Mediay <- 0
  Mediay <- med+trat*co*(xt[l] - med)
  valory[l] <- rnorm(1,Mediay,Vary) }

```

```

# METODO DOS MOMENTOS (JAMES(1973))
for (c2 in l1 : l2)
{
medx <- 0
varx <- 0
vary <- 0
zo <- 0
k1 <- 0
k2 <- 0
ko <- 0
trunct <- 0

medx <- mean(xt)
varx <- var(xt)
vary <- var(valory)
trunct <- (1-c2/100)
zo <- qnorm(trunct,0,1)
k1 <- dnorm(zo,0,1)
k2 <- 1-pnorm(zo,0,1)
ko <- k1/k2

# CALCULO DO COEFICIENTE DA REGRESSAO:
ajuste.fit <- lm(valory~xt)
b <- 0
cof <- coef(ajuste.fit)
b <- cof[2]

# ESTIMADORES PROPOSTOS POR JAMES(1973):
simul[posm,2] <- varx/(ko*(zo-ko)+1)
simul[posm,1] <- medx-ko*sqrt(simul[posm,2])
simul[posm,3] <- sqrt(b^2*(ko*(zo-ko)+1) - vary/simul[posm,2] + 1)
simul[posm,4] <- b/simul[posm,3]
posm <- posm + 1 }
tam[s] <- gt
s <- s + 1
}

```

```
sink ("a:\\simulj\\simulj19.txt")
nsim
trunc
N
tam
gt
simul
sink()

# LIMPANDO AS VARIÁVEIS:
rm(N,med,dp,trunc,co,trat,l1,l2,nsim)
rm(posm,simul)
rm(gc,gt,amostra)
rm(xo,xinc,i,xc,xt,k)
rm(valory,Vary,Mediay)
rm(medx,varx,vary,zo,k1,k2,ko,trunct)
rm(ajuste.fit,b,cof)
rm(s,tam)
```

2) COMPARAÇÃO ENTRE OS 3 MÉTODOS DE ESTIMAÇÃO:

PROGRAMA 2:

```
# PROGRAMA PARA COMPARACAO ENTRE OS 3 METODOS DESCRITOS DURANTE
# O TRABALHO: METODO DOS MOMENTOS (JAMES(1973)) , METODO DE MA-
# XIMA VEROSSIMILHANCA (SENN E BROWN(1985)) E METODO DA PSEUDO-
# VEROSSIMILHANCA (CHEN E COX (1992))
```

DEFINICAO DOS PARAMETROS POPULACIONAIS:

```
# namost = numero de amostra a serem geradas
# N      = tamanho da amostra
# med    = media populacional
# dp     = desvio-padrao
# trunc  = % de truncamento
# co     = efeito da regressao para media
# trat   = efeito do tratamento
```

INICIALIZACAO DOS VALORES:

```
namost <- 0
N      <- 0
med    <- 0
dp     <- 0
trunc  <- 0
co     <- 0
trat   <- 0
```

ESCOLHA DOS PARAMETROS:

```
namost <- 1000
N      <- 1000
med    <- 10
dp     <- 1
trunc  <- 0.95
co     <- 0.80
trat   <- 0.20
```

```

# DEFINIÇÃO DAS MATRIZES DOS PARÂMETROS:
Mmedia <- matrix(0,namost,3)
Mvar   <- matrix(0,namost,3)
Mco    <- matrix(0,namost,3)
Mtrat  <- matrix(0,namost,3)

# inter = numero máximo de interações para o Newton-Rapson
inter  <- 0
inter  <- 150

# zero: conta o numero de estimativas que foram zero para o
# efeito de regressao e o efeito de tratamento
zero <- 0

# ncm: numero de amostras em que os estimadores nao estao bem
# definidos para o metodo dos momentos
ncm <- 0

# ncp: numero de amostras em que os estimadores nao estao bem
# definidos para o metodo de pseudo-verossimilhanca
ncp <- 0

# ncnr: numero de amostras em que os estimadores nao estao bem
# definidos para o metodo de maxima-verossimilhanca
ncnr <- 0

# vetor para convergência de Newton-Rapson:
converge <- vector("numeric", namost)

# funcao que calcula o determinante de uma matriz 3x3 (por Sarrus):
det3x3 <- function(x) (x[1,1]*x[2,2]*x[3,3] + x[1,2]*x[2,3]*x[3,1] +
x[1,3]*x[2,1]*x[3,2] - x[1,1]*x[2,3]*x[3,2] - x[1,2]*x[2,1]*x[3,3] -
x[1,3]*x[2,2]*x[3,1])

```



```

# funcao que calcula o determinante de uma matriz 4x4 (por Laplace):

det<- function(x) {

D1 <- matrix (0,3,3)
D2 <- matrix (0,3,3)
D3 <- matrix (0,3,3)
D4 <- matrix (0,3,3)

D1[1,1] <- x[2,2]
D1[1,2] <- x[2,3]
D1[1,3] <- x[2,4]
D1[2,1] <- x[3,2]
D1[2,2] <- x[3,3]
D1[2,3] <- x[3,4]
D1[3,1] <- x[4,2]
D1[3,2] <- x[4,3]
D1[3,3] <- x[4,4]

c1 <- det3x3(D1)

D2[1,1] <- x[2,1]
D2[1,2] <- x[2,3]
D2[1,3] <- x[2,4]
D2[2,1] <- x[3,1]
D2[2,2] <- x[3,3]
D2[2,3] <- x[3,4]
D2[3,1] <- x[4,1]
D2[3,2] <- x[4,3]
D2[3,3] <- x[4,4]

c2 <- - det3x3(D2)

```

```
D3 [1,1] <- x[2,1]
D3 [1,2] <- x[2,2]
D3 [1,3] <- x[2,4]
D3 [2,1] <- x[3,1]
D3 [2,2] <- x[3,2]
D3 [2,3] <- x[3,4]
D3 [3,1] <- x[4,1]
D3 [3,2] <- x[4,2]
D3 [3,3] <- x[4,4]
```

```
c3 <- det3x3(D3)
```

```
D4 [1,1] <- x[2,1]
D4 [1,2] <- x[2,2]
D4 [1,3] <- x[2,3]
D4 [2,1] <- x[3,1]
D4 [2,2] <- x[3,2]
D4 [2,3] <- x[3,3]
D4 [3,1] <- x[4,1]
D4 [3,2] <- x[4,2]
D4 [3,3] <- x[4,3]
```

```
c4 <- - det3x3(D4)
```

```
x[1,1]*c1 + x[1,2]*c2 +x[1,3]*c3 + x[1,4]*c4
```

```
}
```

```

# funcao que calcula a inversa de uma matriz 4x4:
inversa4X4<- function(x) {

# calculo do determinante da matriz:
deter <- det(x)

# definicao das matrizes cofatoras:
M11 <- matrix (0,3,3)
M12 <- matrix (0,3,3)
M13 <- matrix (0,3,3)
M14 <- matrix (0,3,3)
M21 <- matrix (0,3,3)
M22 <- matrix (0,3,3)
M23 <- matrix (0,3,3)
M24 <- matrix (0,3,3)
M31 <- matrix (0,3,3)
M32 <- matrix (0,3,3)
M33 <- matrix (0,3,3)
M34 <- matrix (0,3,3)
M41 <- matrix (0,3,3)
M42 <- matrix (0,3,3)
M43 <- matrix (0,3,3)
M44 <- matrix (0,3,3)

# calculo das matrizes cofatoras:
M11[1,1] <- x[2,2]
M11[1,2] <- x[2,3]
M11[1,3] <- x[2,4]
M11[2,1] <- x[3,2]
M11[2,2] <- x[3,3]
M11[2,3] <- x[3,4]
M11[3,1] <- x[4,2]
M11[3,2] <- x[4,3]
M11[3,3] <- x[4,4]

```

```

c11 <- det3x3(M11)

M12[1,1] <- x[2,1]
M12[1,2] <- x[2,3]
M12[1,3] <- x[2,4]
M12[2,1] <- x[3,1]
M12[2,2] <- x[3,3]
M12[2,3] <- x[3,4]
M12[3,1] <- x[4,1]
M12[3,2] <- x[4,3]
M12[3,3] <- x[4,4]
c12 <- - det3x3(M12)

M13[1,1] <- x[2,1]
M13[1,2] <- x[2,2]
M13[1,3] <- x[2,4]
M13[2,1] <- x[3,1]
M13[2,2] <- x[3,2]
M13[2,3] <- x[3,4]
M13[3,1] <- x[4,1]
M13[3,2] <- x[4,2]
M13[3,3] <- x[4,4]
c13 <- det3x3(M13)

M14[1,1] <- x[2,1]
M14[1,2] <- x[2,2]
M14[1,3] <- x[2,3]
M14[2,1] <- x[3,1]
M14[2,2] <- x[3,2]
M14[2,3] <- x[3,3]
M14[3,1] <- x[4,1]
M14[3,2] <- x[4,2]
M14[3,3] <- x[4,3]
c14 <- - det3x3(M14)

```

```
M21[1,1] <- x[1,2]
M21[1,2] <- x[1,3]
M21[1,3] <- x[1,4]
M21[2,1] <- x[3,2]
M21[2,2] <- x[3,3]
M21[2,3] <- x[3,4]
M21[3,1] <- x[4,2]
M21[3,2] <- x[4,3]
M21[3,3] <- x[4,4]
c21 <- - det3x3(M21)
```

```
M22[1,1] <- x[1,1]
M22[1,2] <- x[1,3]
M22[1,3] <- x[1,4]
M22[2,1] <- x[3,1]
M22[2,2] <- x[3,3]
M22[2,3] <- x[3,4]
M22[3,1] <- x[4,1]
M22[3,2] <- x[4,3]
M22[3,3] <- x[4,4]
c22 <- det3x3(M22)
```

```
M23[1,1] <- x[1,1]
M23[1,2] <- x[1,2]
M23[1,3] <- x[1,4]
M23[2,1] <- x[3,1]
M23[2,2] <- x[3,2]
M23[2,3] <- x[3,4]
M23[3,1] <- x[4,1]
M23[3,2] <- x[4,2]
M23[3,3] <- x[4,4]
c23 <- - det3x3(M23)
```

```
M24 [1,1] <- x[1,1]
M24 [1,2] <- x[1,2]
M24 [1,3] <- x[1,3]
M24 [2,1] <- x[3,1]
M24 [2,2] <- x[3,2]
M24 [2,3] <- x[3,3]
M24 [3,1] <- x[4,1]
M24 [3,2] <- x[4,2]
M24 [3,3] <- x[4,3]
c24 <- det3x3(M24)
```

```
M31 [1,1] <- x[1,2]
M31 [1,2] <- x[1,3]
M31 [1,3] <- x[1,4]
M31 [2,1] <- x[2,2]
M31 [2,2] <- x[2,3]
M31 [2,3] <- x[2,4]
M31 [3,1] <- x[4,2]
M31 [3,2] <- x[4,3]
M31 [3,3] <- x[4,4]
c31 <- det3x3(M31)
```

```
M32 [1,1] <- x[1,1]
M32 [1,2] <- x[1,3]
M32 [1,3] <- x[1,4]
M32 [2,1] <- x[2,1]
M32 [2,2] <- x[2,3]
M32 [2,3] <- x[2,4]
M32 [3,1] <- x[4,1]
M32 [3,2] <- x[4,3]
M32 [3,3] <- x[4,4]
c32 <- - det3x3(M32)
```

```
M33 [1,1] <- x[1,1]
M33 [1,2] <- x[1,2]
M33 [1,3] <- x[1,4]
M33 [2,1] <- x[2,1]
M33 [2,2] <- x[2,2]
M33 [2,3] <- x[2,4]
M33 [3,1] <- x[4,1]
M33 [3,2] <- x[4,2]
M33 [3,3] <- x[4,4]
c33 <- det3x3(M33)
```

```
M34 [1,1] <- x[1,1]
M34 [1,2] <- x[1,2]
M34 [1,3] <- x[1,3]
M34 [2,1] <- x[2,1]
M34 [2,2] <- x[2,2]
M34 [2,3] <- x[2,3]
M34 [3,1] <- x[4,1]
M34 [3,2] <- x[4,2]
M34 [3,3] <- x[4,3]
c34 <- - det3x3(M34)
```

```
M41 [1,1] <- x[1,2]
M41 [1,2] <- x[1,3]
M41 [1,3] <- x[1,4]
M41 [2,1] <- x[2,2]
M41 [2,2] <- x[2,3]
M41 [2,3] <- x[2,4]
M41 [3,1] <- x[3,2]
M41 [3,2] <- x[3,3]
M41 [3,3] <- x[3,4]
c41 <- - det3x3(M41)
```

```
M42[1,1] <- x[1,1]
M42[1,2] <- x[1,3]
M42[1,3] <- x[1,4]
M42[2,1] <- x[2,1]
M42[2,2] <- x[2,3]
M42[2,3] <- x[2,4]
M42[3,1] <- x[3,1]
M42[3,2] <- x[3,3]
M42[3,3] <- x[3,4]
c42 <- det3x3(M42)
```

```
M43[1,1] <- x[1,1]
M43[1,2] <- x[1,2]
M43[1,3] <- x[1,4]
M43[2,1] <- x[2,1]
M43[2,2] <- x[2,2]
M43[2,3] <- x[2,4]
M43[3,1] <- x[3,1]
M43[3,2] <- x[3,2]
M43[3,3] <- x[3,4]
c43 <- - det3x3(M43)
```

```
M44[1,1] <- x[1,1]
M44[1,2] <- x[1,2]
M44[1,3] <- x[1,3]
M44[2,1] <- x[2,1]
M44[2,2] <- x[2,2]
M44[2,3] <- x[2,3]
M44[3,1] <- x[3,1]
M44[3,2] <- x[3,2]
M44[3,3] <- x[3,3]
c44 <- det3x3(M44)
```



```

adjunta <- matrix (0,4,4)

adjunta[1,1] <- c11
adjunta[1,2] <- c21
adjunta[1,3] <- c31
adjunta[1,4] <- c41
adjunta[2,1] <- c12
adjunta[2,2] <- c22
adjunta[2,3] <- c32
adjunta[2,4] <- c42
adjunta[3,1] <- c13
adjunta[3,2] <- c23
adjunta[3,3] <- c33
adjunta[3,4] <- c43
adjunta[4,1] <- c14
adjunta[4,2] <- c24
adjunta[4,3] <- c34
adjunta[4,4] <- c44

adjunta/deter }

# ESTIMACAO DOS PARAMETROS PELOS 3 METODOS:
# ESPECIFICA O PONTO DE TRUNCAMENTO:
xo <- 0
xo <- qnorm(trunc,med,dp)

for (a in 1 : namost)
{

amostra <- vector("numeric", N)
amostra <- rnorm(N,med,dp)
amostra <- sort(amostra)

gt <- 0
gc <- 0

```

```

# SEPARACAO DO GRUPO DE ELEMENTOS NAO TRATADOS:
xinc <- vector ("numeric", N)
i <- 1
while (amostra[i] < xo)
  {xinc[i] <- amostra[i]
  i <- i+1 }
gc <- i-1
xc <- vector("numeric", gc)
for (l in 1 : gc)
  { xc[l] <- xinc[l] }

# SEPARACAO DO GRUPO DE ELEMENTOS TRATADOS:
gt <- N - gc
xt <- vector ("numeric",gt)
k <- 1
for ( j in i : N )
  { xt[k] <- amostra[j]
  k <- k+1 }

# GERACAO DOS VALORES Y SEGUNDO O MODELO ADOTADO:
valory <- vector ("numeric", gt)
Vary <- 0
Vary <- dp^2*(1 - co^2)
for ( l in 1 : gt )
  { Mediai <- 0
  Mediai <- med+trat*co*(xt[l] - med)
  valory[l] <- rnorm(1,Mediai,Vary) }

```

```

# Método da Pseudo-Verossimilhança (Chen e Cox (1992))
# ESTIMAÇÃO DOS PARÂMETROS NUISANCE:
mo <- 0
varo <- 0
mo <- mean(amostra)
varo <- var(amostra)

# ESTIMAÇÃO DO PARÂMETRO DE REGRESSÃO LINEAR:
b <- 0
ajuste.fit <- lm(valory ~xt)
cof <- coef(ajuste.fit)
b <- cof[2]

# ESTIMAÇÃO DOS PARÂMETROS DE INTERESSE:
aux1 <- 0
aux2 <- 0
aux3 <- 0
ve1 <- vector ("numeric",gt)
ve2 <- vector ("numeric",gt)
ve3 <- vector ("numeric",gt)
ve1 <- valory - mo
ve2 <- xt-mo
ve3 <- (ve1-b*ve2)^2
aux1 <- sum(ve3)
aux2 <- 1-aux1/(gt*varo)
if (aux2 < 0) ncp <- ncp + 1 else {
aux3 <- sqrt(aux2)
if (aux3 > 1) ncp <- ncp + 1 else {
Mmedia[a,3] <- mo
Mvar [a,3] <- varo
Mco [a,3] <- ifelse(b > 0, aux3, -aux3)
Mtrat [a,3] <- b/(Mco[a,3]) } }
rm(ajuste.fit,cof,mo,varo)
rm(aux1,aux2,aux3,ve1,ve2,ve3)

```

```

# METODO DOS MOMENTOS (JAMES(1973))
# OBS: O METODO APRESENTA PROBLEMAS (NAO ESTA CORRETO) PORTANTO
# SERA APENAS UTILIZADO COMO ILUSTRACAO PARA O CHUTE INICIAL,
# POIS ESTAMOS SUPONDO QUE ZO E CONHECIDO, O QUE NA PRATICA
# NAO E VERDADE

```

```

medx <- 0
varx <- 0
vary <- 0
zo <- 0
k1 <- 0
k2 <- 0
ko <- 0

```

```

medx <- mean(xt)
varx <- var(xt)
vary <- var(valory)
zo <- (xo - med)/dp
k1 <- dnorm(zo,0,1)
k2 <- 1-pnorm(zo,0,1)
ko <- k1/k2

```

```

# ESTIMADORES PROPOSTOS POR JAMES(1973):

```

```

vo <- 0
mo <- 0
coro <- 0
trato <- 0

```

```

vo <- varx/(ko*(zo-ko)+1)
aux1 <- b^2*(ko*(zo-ko)+1) - vary/vo + 1

```

```

if (vo < 0 | aux1 < 0) ncm <- ncm + 1 else {
mo <- medx-ko*sqrt(vo)
coro <- sqrt(aux1)
trato <- b/coro
}

```

```

if (coro > 1) ncm <- ncm + 1 else {
Mmedia[a,1] <- mo
Mvar [a,1] <- vo
Mco [a,1] <- coro
Mtrat [a,1] <- trato } }
rm(medx,varx,vary,zo,k1,k2,ko,b,aux1)

# NEWTON-RAPSON :
cont <- 1
while (cont < inter+1 & converge[a]==0) {
# DEFINICAO DO VETOR DE PARAMETROS:
vparo <- vector ("numeric", 4)
vpar <- vector ("numeric", 4)
vparo[1] <- mo
vparo[2] <- vo
vparo[3] <- coro
vparo[4] <- trato

# INICIALIZACAO DOS VALORES DE INTERESSE:
ve4 <- vector ("numeric",gc)
ve5 <- vector ("numeric",gt)
ve6 <- vector ("numeric",gt)
ve7 <- vector ("numeric",gc)
ve8 <- vector ("numeric",gt)
ve9 <- vector ("numeric",gt)
ve10 <- vector ("numeric",gt)
ve11 <- vector ("numeric",gt)
ve12 <- vector ("numeric",gt)
ve13 <- vector ("numeric",gt)

vc1 <- 0
vc2 <- 0
vc3 <- 0
vc4 <- 0
vc5 <- 0
vc6 <- 0

```

```

vc7 <- 0
vc8 <- 0
vc9 <- 0
vc10 <- 0
vc11 <- 0
vc12 <- 0
vc13 <- 0

# CALCULO DAS QUANTIDADES BASICAS PARA O USO DE NEWTON-RAPSON:
vc1 <- trato*coro
vc2 <- vc1 - 1
vc3 <- 1 - coro^2
ve4 <- xc - mo
vc4 <- sum(ve4)
ve5 <- xt - mo
vc5 <- sum(ve5)
ve6 <- (valory - vc1 * xt)
vc6 <- sum(ve6)
ve7 <- ve4^2
vc7 <- sum(ve7)
ve8 <- ve5^2
vc8 <- sum(ve8)
ve9 <- valory - mo
vc9 <- sum(ve9)
ve10 <- ve9 - vc1*ve5
vc10 <- sum(ve10)
ve11 <- ve10^2
vc11 <- sum(ve11)
ve12 <- ve9*ve5
vc12 <- sum(ve12)
ve13 <- ve12-vc1*ve8
vc13 <- sum(ve13)

# FUNCAO ESCORE:
u <- vector("numeric", 4)
u[1] <- (vc4 + vc5 - (gt*mo*vc2^2)/vc3 - (vc2*vc6)/vc3)/vo

```

```

u[2] <- (- gc - 2*gt + (vc7+vc8+(vc11/vc3))/vo )/(2*vo)
u[3] <- (gt*coro)/vc3 + (trato*vc13)/(vo*vc3)-(coro*vc11)/(vo*(vc3^2))
u[4] <- (coro*vc13)/(vo*vc3)

# MATRIZ DE DERIVADAS:
MD <- matrix(0,4,4)
m12 <- 0
m13 <- 0
m14 <- 0
m23 <- 0
m24 <- 0
m34 <- 0
MD[1,1] <- (-gc -gt * ( 1 + (vc2^2)/vc3 ) ) / vo
m12 <- - (vc4 + vc5 - (vc2 * vc10)/vc3 ) / (vo^2)
MD[1,2] <- m12
MD[2,1] <- m12
m13 <- ( (2 * coro * (trato^2)*vc5) - (trato*(vc9+vc5)) ) / (vo * vc3) -
(2 * coro * (vc2*vc9-vc1*vc2*vc5) ) / (vo*(vc3^2))
MD[1,3] <- m13
MD[3,1] <- m13
m14 <- (coro * ((2*vc1 - 1) * vc5 - vc9)) / (vo * vc3)
MD[1,4] <- m14
MD[4,1] <- m14
MD[2,2] <- (gc + 2 * gt)/(2 * (vo^2)) - (vc7 + vc8 + (vc11/vc3))/(vo^3)
m23 <- (- (trato * vc13)/vc3 + (coro * vc11)/(vc3^2) ) / (vo^2)
MD[2,3] <- m23
MD[3,2] <- m23
m24 <- - (coro * vc13) / (vc3 * (vo^2))
MD[2,4] <- m24
MD[4,2] <- m24
MD[3,3] <- (gt * (1 + (coro^2)))/(vc3^2) - ((trato^2) * vc8)/(vo*vc3) +
- ((1 + 3 * (coro^2)) * vc11)/(vo * (vc3^3))
m34 <- ((1 + (coro^2)) * vc12 - 2 * coro * trato * vc8)/ (vo * (vc3^2))
MD[3,4] <- m34
MD[4,3] <- m34
MD[4,4] <- - ((coro^2) * vc8) / (vo * vc3)

```

```

# verifica se ha' alguma indefinicao na matriz MD
indef <- 0
if(MD[1,1] == Inf | MD[1,1] == -Inf) indef <- 1 else {
if(MD[1,2] == Inf | MD[1,2] == -Inf) indef <- 1 else {
if(MD[1,3] == Inf | MD[1,3] == -Inf) indef <- 1 else {
if(MD[1,4] == Inf | MD[1,4] == -Inf) indef <- 1 else {
if(MD[2,1] == Inf | MD[2,1] == -Inf) indef <- 1 else {
if(MD[2,2] == Inf | MD[2,2] == -Inf) indef <- 1 else {
if(MD[2,3] == Inf | MD[2,3] == -Inf) indef <- 1 else {
if(MD[2,4] == Inf | MD[2,4] == -Inf) indef <- 1 else {
if(MD[3,1] == Inf | MD[3,1] == -Inf) indef <- 1 else {
if(MD[3,2] == Inf | MD[3,2] == -Inf) indef <- 1 else {
if(MD[3,3] == Inf | MD[3,3] == -Inf) indef <- 1 else {
if(MD[3,4] == Inf | MD[3,4] == -Inf) indef <- 1 else {
if(MD[4,1] == Inf | MD[4,1] == -Inf) indef <- 1 else {
if(MD[4,2] == Inf | MD[4,2] == -Inf) indef <- 1 else {
if(MD[4,3] == Inf | MD[4,3] == -Inf) indef <- 1 else {
if(MD[4,4] == Inf | MD[4,4] == -Inf) indef <- 1 } } } } } } } } } } } } }
} } }

if(indef == 1) cont <- inter + 1 else {
if (abs(det(MD)) < 0.000001) cont <- inter+1 else {
# ALGORITMO DE NEWTON-RAPSON :
vpar <- vparo - inversa4X4(MD)%*%u
mo <- vpar[1]
vo <- vpar[2]
coro <- vpar[3]
trato <- vpar[4]
pc1 <- vector("numeric",4)
pc2 <- 0
pc1 <- abs(vpar-vparo)
pc2 <- sum(pc1)
if (pc2 < 0.0000001) converge[a] <- 1
cont <- cont + 1 }
} }

```



```

if (converge[a]==1 & abs(vpar[3]) > 0.00000001 ) {
if (vpar[3] < -1 | vpar[3] > 1 | vpar[2] < 0) ncnr <- ncnr + 1 else {
Mmedia[a,2] <- vpar[1]
Mvar [a,2] <- vpar[2]
Mco [a,2] <- vpar[3]
Mtrat [a,2] <- vpar[4] } }
if (converge[a]==1 & abs(vpar[3]) < 0.00000001 ) zero <- zero + 1

```

```

# LIMPANDO AS VARIÁVEIS:

```

```

rm(gt,gc)
rm(amostra,xinc,xc,xt)
rm(k,i,cont)
rm(valory,Vary,Mediay)
rm(mo,vo,coro,trato)
rm(vpar,vparo)
rm(ve4,ve5,ve6,ve7,ve8,ve9,ve10,ve11,ve12,ve13)
rm(vc1,vc2,vc3,vc4,vc5,vc6,vc7,vc8,vc9,vc10,vc11,vc12,vc13)
rm(m12,m13,m14,m23,m24,m34)
rm(u,MD)
}

```

```

# CALCULO DO EQM PARA OS TRES METODOS:

```

```

MEQM <- matrix(0,4,3)
M1 <- matrix(0,namost,3)
M2 <- matrix(0,namost,3)
M3 <- matrix(0,namost,3)
M4 <- matrix(0,namost,3)

M1 <- (Mmedia - med)^2
M2 <- (Mvar - dp^2)^2
M3 <- (Mco - co)^2
M4 <- (Mtrat-trat)^2

```

```

MEQM[1,1] <- (sum(M1[,1]) - ncm*med^2)/(namost - ncm)
MEQM[1,3] <- (sum(M1[,3]) - ncp*med^2)/(namost - ncp)
MEQM[2,1] <- (sum(M2[,1]) - ncm*dp^4)/(namost - ncm)
MEQM[2,3] <- (sum(M2[,3]) - ncp*dp^4)/(namost - ncp)
MEQM[3,1] <- (sum(M3[,1]) - ncm*co^2)/(namost - ncm)
MEQM[3,3] <- (sum(M3[,3]) - ncp*co^2)/(namost - ncp)
MEQM[4,1] <- (sum(M4[,1]) - ncm*trat^2)/(namost - ncm)
MEQM[4,3] <- (sum(M4[,3]) - ncp*trat^2)/(namost - ncp)

eqmvm <- 0
eqmvv <- 0
eqmvc <- 0
eqmvt <- 0

for (d in 1 : namost) {

if (converge[d] == 1) {
eqmvm <- eqmvm + M1[d,2]
eqmvv <- eqmvv + M2[d,2]
eqmvc <- eqmvc + M3[d,2]
eqmvt <- eqmvt + M4[d,2] }
}

nconverge <- sum(converge)
MEQM[1,2] <- (eqmvm - (zero+ncnr)*med^2)/(nconverge - (zero+ncnr))
MEQM[2,2] <- (eqmvv - (zero+ncnr)*dp^4)/(nconverge - (zero+ncnr))
MEQM[3,2] <- (eqmvc - (zero+ncnr)*co^2)/(nconverge - (zero+ncnr))
MEQM[4,2] <- (eqmvt - (zero+ncnr)*trat^2)/(nconverge - (zero+ncnr))

# taxa de convergencia para Newton-Rapson:
TC <- (nconverge - (zero+ncnr))/namost

# GUARDANDO OS RESULTADOS DA SIMULACAO EM ARQUIVO:
sink("/home/posmae/erika/tese/simula/sim")

```

Mmedia

Mvar

Mco

Mtrat

namost

N

trunc

med

dp

co

trat

ncm

ncp

ncnr

zero

TC

MEQM

sink()

LIMPEZA DAS VARIÁVEIS:

rm(N, med, dp, trunc, co, trat, xo, inter, nconverge, TC, namost, zero)

rm(eqmvv, eqmvt, det, det3x3, inversa4X4, pc1, pc2, ncm, ncp, ncnr)

rm(Mmedia, Mvar, Mco, Mtrat, MEQM, converge, M1, M2, M3, M4)

OBS:

- 1) Apesar do programa 2 ter sido desenvolvido no S-Plus (v3.2) para micro não seria viável devido a questões de memória rodar as simulações no micro devido ao porte do estudo de simulação, portanto este programa após ter sido desenvolvido no micro foi transferido para a uma estação de trabalho onde foi realizado o estudo de simulação.

- 2) Embora este seja o programa 2 completo para a comparação dos métodos (poderia ser rodado desta forma), durante o processo de simulação este programa foi dividido em partes (foram analisadas 50, 100 e 200 amostras de cada vez de acordo com o tamanho de N), visando uma otimização do tempo de realização das simulações, pois é preciso ressaltar que o S-Plus, apesar de apresentar inúmeras facilidades de programação, tem um custo computacional muito alto para rodar simulações deste porte.

REFERÊNCIAS:

- Chen, S. and Cox, C. (1992). Use of Baseline Data for Estimation of Treatment Effects in the Presence of Regression to the Mean. **Biometrics** 48, 593-598.
- Cohen, A.C. (1955). Restriction and selection in samples from bivariate normal distributions. **Journal of American Statistics Association** 50, 884-893.
- Delgado, J.J.F. (1995). **Estimação por Pseudo Máxima Verossimilhança**. São Paulo: Instituto de Matemática e Estatística - USP. Tese de mestrado. 56p.
- Draper, N.R. and Smith, H. (1981). **Applied Regression Analysis**. 2 ed. New York: John Wiley & Sons. 709p.
- Dudewicz, E.J. and Mishra, S.N.(1988). **Modern Mathematical Statistics**. New York: John Wiley & Sons. 838p.
- Elian, S.N. (1988). **Análise de Regressão**. São Paulo: Instituto de Matemática e Estatística - USP . 232p.
- Galton, F. (1886). Regression towards mediocrity in hereditary stature. **Journal of the Anthropological Institute** 15, 246-263.

- Gong, G.H. and Samaniego, F.J. (1981). Pseudo maximum likelihood estimation: theory and applications. **Annals of Statistics** 9, 861-869.
- Guidorizzi, H.L. (1986). **Um curso de cálculo**. Vol 2. Rio de Janeiro: FTC. 995p.
- James, K.E. (1973). Regression toward the mean in uncontrolled clinical studies. **Biometrics** 29, 121-130.
- Johnson, and Kotz, (1970). **Continuous univariate distributions - I**. vol.2. 2 ed. New York: John Wiley & Sons. 300p.
- Leite, G.L. e Singer, J.M. (1990). **Métodos Assintóticos em Estatística - Fundamentos e Aplicações**. São Paulo: Associação Brasileira de Estatística - IX Simpósio Nacional de Probabilidade e Estatística. 130p.
- Senn, S. J. and Brown, R. A. (1985). Estimating treatment effect in clinical trials subject to regression to the mean. **Biometrics** 41, 555-560.
- Senn, S. J. and Brown, R. A. (1989). Maximum Likelihood estimation of treatment effects for samples subject to regression to the mean. **Communication in Statistics - Theory and Methods** 18, 3389-3406.

Sen, P. K. and Singer, J. M. (1993). **Large sample methods in Statistics: an introduction with applications.** New York: Chapman & Hall. 382p.