

Presidente da República  
**Fernando Henrique Cardoso**

Ministro do Planejamento, Orçamento e Gestão  
**Martus Antônio Rodrigues Tavares**

## **INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA - IBGE**

Presidente  
**Sérgio Besserman Vianna**

Diretor de Planejamento e Coordenação  
**Nuno Duarte da Costa Bittencourt**

### **ÓRGÃOS TÉCNICOS SETORIAIS**

Diretoria de Pesquisas  
**Maria Martha Malard Mayer**

Diretoria de Geociências  
**Trento Natali Filho**

Diretoria de Informática  
**Paulo Roberto Ribeiro da Cunha**

Centro de Documentação e Disseminação de Informações  
**David Wu Tai**

Escola Nacional de Ciências Estatísticas  
**Kaizô Iwakami Beltrão**

Ministério do Planejamento, Orçamento e Gestão  
Instituto Brasileiro de Geografia e Estatística – IBGE

# REVISTA BRASILEIRA DE ESTATÍSTICA

Volume 59      número 211      janeiro/junho 1998

© IBGE. 1999

**Revista Brasileira de Estatística, ISSN 0034-7175**

Órgão oficial do IBGE e da Associação Brasileira de Estatística - ABE.

Publicação semestral que se destina a promover e ampliar o uso de métodos estatísticos (quantitativos) na área das ciências econômicas e sociais, através de divulgação de artigos inéditos.

Temas abordando aspectos do desenvolvimento metodológico serão aceitos, desde que relevantes para os órgãos produtores de estatísticas.

Os originais para publicação deverão ser submetidos em três vias (que não serão devolvidas) para:

Pedro Luis do Nascimento Silva  
Editor Responsável - RBES - IBGE,  
Av. República do Chile, 500 - Centro  
20031-170 - Rio de Janeiro, RJ.

Os artigos submetidos às RBES não devem ter sido publicados ou estar sendo considerados para publicação em outros periódicos.

A Revista não se responsabiliza pelos conceitos emitidos em matéria assinada.

**Editor Responsável**

Pedro Luis do Nascimento Silva (IBGE)

**Editor de Estatísticas Oficiais**

Djalma Galvão Carneiro Pessoa (IBGE)

**Editor de Metodologia**

Hélio dos Santos Migon (UFRJ)

**Editores Associados**

Gilberto Alvarenga Paula (USP)  
Kaizô Iwakami Beltrão (IBGE)

Lisbeth Kaiserlian Cordani (USP)  
Renato Martins Assunção (UFMG)  
Wilton de Oliveira Bussab (FGV-SP)

**Impressão**

Centro de Documentação e Disseminação de Informações-CDDI/  
IBGE, em meio digital, em 1999.

**Capa**

Renato J. Aguiar - Divisão de Criação - DIVIC/CDDI

**Ilustração da Capa**

Marcos Balster - Divisão de Criação - DIVIC/CDDI

---

Revista brasileira de estatística/IBGE, - v. 1, n.1 (jan./mar.  
1940)- , - Rio de Janeiro: IBGE, 1940-

v.

Trimestral (1940-1986), semestral (1987- ).  
Continuação de: Revista de economia e estatística.  
Índices acumulados de autor e assunto publicados no v.43  
(1940-1979) e v.50 (1980-1989).  
Co-edição com a Associação Brasileira de Estatística a partir  
do v. 58.  
ISSN 0034-7175 = Revista brasileira de estatística.

I. Estatística - Periódicos. I. IBGE. II. Associação Brasileira de  
Estatística.

IBGE. CDDI. Div. de Biblioteca e Acervos Especiais CDU 31 (05)  
RJ-IBGE/88-05 (rev. 98) PERIÓDICO

Impresso no Brasil/Printed in Brazil

# SUMÁRIO

---

NOTA DO EDITOR	5
----------------	---

---

ARTIGOS	
---------	--

---

BIOAMBIENTE E QUALIDADE DE VIDA: Perspectivas Bioestatísticas	7
Pranab K. Sen	

---

ANÁLISE DOS ERROS AMOSTRAIS DE INDICADORES EDUCACIONAIS OBTIDOS ATRAVÉS DE DADOS DA PESQUISA NACIONAL POR AMOSTRA DE DOMICÍLIOS	47
Fernando A. S. Moura Ruben Klein	

---

MÉTODOS BAYESIANOS EM MODELOS DE CONFIABILIDADE DE SOFTWARE USANDO PROCESSOS DE POISSON NÃO-HOMOGÊNEOS	59
Karin Storani Jorge Alberto Achcar	

---

ALGUNS MODELOS DE ANÁLISE DE VARIÂNCIA EM SÉRIES TEMPORAIS UTILIZANDO TRANSFORMADA DE FOURIER - UMA APLICAÇÃO	81
Clélia Maria de Castro Tolo	

---

POLÍTICA EDITORIAL	113
--------------------	-----

## NOTA DO EDITOR

Inicialmente agradecemos a colaboração dos Professores Lisbeth Cordani e Júlio M. Singer que traduziram o artigo do Prof. Sen submetido em inglês.

No momento, a revista conta com 25 artigos em diversos estágios do processo de avaliação e editoração.

Especificamente, nove destes artigos aguardam resposta de pareceristas; três foram encaminhados aos autores para modificações sugeridas pelos pareceristas; e três ainda estão com o editor para escolha de pareceristas.

Dos dez restantes, dois estão prontos para publicação e oito estão tendo pareceres examinados pelos editores.

Estamos na expectativa de receber artigos apresentados nas duas escolas realizadas recentemente pela ABE, a de Modelos de Regressão e de Séries Temporais e Econometria.

Saudações,

Pedro Luis do Nascimento Silva

Editor Responsável

# BIOAMBIENTE E QUALIDADE DE VIDA: Perspectivas Bioestatísticas<sup>1</sup>

Pranab K. Sen \*

## RESUMO

Vista de uma perspectiva mais ampla, a qualidade de vida humana em uma comunidade ou sociedade depende não apenas do padrão de vida em um sentido monetário convencional, mas também do bioambiente que governa um conjunto de fatores sociais, econômicos, religiosos, culturais, de saúde e psicologia, ecológicos e ambientais. Estas perspectivas são avaliadas na formulação e interpretação de Qualidade de Vida e, como refere Sen (1996), a verdadeira ênfase às bases bioestatísticas é dada neste contexto.

**Palavras-chave:** Coeficiente de Gini; custo/benefício; qualidade de vida; envelhecimento; Escala de Karnofsky; índices de pobreza; medidas estatísticas; países do Terceiro Mundo; riscos competitivos; saúde e risco ambiental; saúde mental; saúde pública; utilidade.

## 1. INTRODUÇÃO

---

Neste final de século, as luzes dos avanços foram basicamente roubadas pelo advento da *tecnologia da informação*. A eletrônica e os computadores invadiram todos os cantos do globo e influenciaram todos os níveis da vida, da ciência e da sociedade. E ainda grandes desafios emergiram de cada esfera de vida, mais notavelmente nos setores de saúde pública e ambiental. O ecoambiente de nosso planeta mãe está efetivamente ameaçado por problemas que representam um risco para a vida, não devido apenas aos desequilíbrios ecológicos e desastres ambientais, mas também ao advento de doenças ou distúrbios novos (ou desconhecidos até então), como o HIV, que podem afetar drasticamente a Qualidade de Vida (QDV) de todo o biosistema, incluindo a humanidade.

---

<sup>1</sup> Apresentado na Reunião da Região Brasileira da Sociedade Internacional de Biometria, em Recife, PE, em 10 de julho de 1997. O presente artigo é uma adaptação atualizada de uma apresentação na Conferência sobre Qualidade de Vida em Hurghada, Egito, 14 de dezembro de 1994, posteriormente publicada (Sen, 1996). A tradução para a língua portuguesa foi patrocinada pela Associação Brasileira de Estatística-ABE-, em junho de 1998, e realizada por Tone I. Sguizzardi, com revisões feitas por Lisbeth K. Cordani e Júlio da Motta Singer.

\* Endereço para correspondência: Universidade da Carolina do Norte, Chapel Hill, Estados Unidos.

A proteção de nosso *bioambiente* é uma necessidade para evitar a nossa extinção e assegurar uma propagação saudável de nossos biosistemas na Terra. Esta imensa tarefa (social e também científica) somente pode ser executada através de uma abordagem multidisciplinar, em que as ciências clínicas podem ter um papel aparentemente predominante, embora epidemiologia, nutrição, educação em saúde e comportamento saudável, saúde materno-infantil, bem como as ciências da saúde ambiental e outras disciplinas de saúde pública, sejam igualmente indispensáveis neste contexto. Além disso, na quantificação e avaliação da dimensão absoluta destes problemas que colocam a vida em risco, bem como na mensuração dos riscos e perigos para a saúde devidos a eles, a bioestatística desempenha um papel absolutamente fundamental. A interface entre a informação e as ciências estatísticas proporciona uma riqueza de ferramentas gerais para a quantificação, modelagem e análise dos fenômenos subjacentes que foram reconhecidos principalmente de forma apenas qualitativa, e a bioestatística fornece a chave para a adaptabilidade de uma abordagem interdisciplinar quantitativa.

No presente estudo, vamos nos restringir principalmente ao tema básico: formulação e interpretação da qualidade de vida em um sentido mais amplo, em que são incorporadas perspectivas individuais e também populacionais em situações espaciais, temporais e bioambientais. Interpretações convencionais de QDV relacionam-se a alguns pontos de vista objetivos, específicos, como o *padrão de vida* em sentido monetário ou o estado das condições físicas ou mentais de pessoas afligidas por alguma doença/distúrbio importante (como câncer, AIDS, etc.) ou alguma doença crônica (como diabetes). Sob o ponto de vista populacional, uma variedade de causas podem ser identificadas, e o quadro composto depende de nosso bioambiente em constante mutação. Há uma profunda necessidade de conhecer os padrões espaciais e também os temporais que influenciam o bioambiente, e que são consistentes com os recursos naturais disponíveis, devendo ser feita a melhor utilização para manter e melhorar a QDV sob uma perspectiva global. Esta linha de ataque, sugerida por Sen (1996), nos permite traçar quadros de QDV mais realistas e fazer com que possam ser utilizados também para finalidades administrativas. O impacto bioambiental tem sido adequadamente enfatizado no atual estudo e é retomado à luz da abordagem de Sen (1996).

Como o padrão de vida em um sentido monetário, também diversas perspectivas sociais, culturais, religiosas, educacionais, psicológicas, ecoambientais e de saúde ambiental têm algumas medidas de QDV, ainda que muitas delas enfoquem mais o ponto de vista individual do que o populacional. Seguindo Sen (1996), a próxima seção incorpora perspectivas de três grandes caminhos: (i) socioeconômico-culturais, (ii) de saúde e psicologia, e (iii) ecoambientais, em uma avaliação de conceitos de QDV com a devida ênfase em seu impacto sobre os indivíduos envolvidos, bem como sobre as populações. Estas medidas de QDV podem variar de países industrializados para países do Terceiro Mundo, ainda que haja um esforço global comum para fazer o melhor uso possível de recursos disponíveis para manter e melhorar a QDV. Na Seção 3, as medidas convencionais de padrão

de vida são então revistas em uma forma sistemática, tendo em vista a incorporação de alguns destes conceitos também nos outros dois setores. A Seção 4 é dedicada ao setor de saúde e saúde mental, onde há uma profunda necessidade de tais conceitos e medidas de qualidade de vida relacionadas com a saúde (QDVRs). A Seção 5 lida com o setor de saúde ambiental e traz a relevância do bioambiente para as perspectivas de QDV. Na Seção 6, estas diversas medidas são alinhadas para formar algumas medidas de QDV global. Os conceitos de *gap ratio*<sup>2</sup> de qualidade (GRQ), *deficiência de qualidade de vida* (DQDV) e *deficiência de qualidade de vida ajustada* (DQVA) desempenham importantes papéis neste contexto. A seção final chama a atenção para a necessidade de tais estudos nos países do Terceiro Mundo, juntamente com algumas questões não resolvidas.

## **2. UMA VISÃO PRISMÁTICA DE QDV**

Uma comunidade (sociedade/população) tem, em geral, pessoas em todos os níveis de vida, com diferentes características socioeconômicas, culturais e religiosas, e que estão expostas a uma variedade de problemas, tanto de saúde quanto ambientais. Visto desta perspectiva mais ampla, parece que, para desenvolver uma medida abrangente de QDV para toda a comunidade, há uma necessidade genuína de se amalgamarem as medidas de QDV para vários setores (alguns dos quais requerendo formulações adequadas), e enfatizar o seu encadeamento. Assim, a QDV de uma sociedade deveria refletir a qualidade de vida de cada setor de seus habitantes, bem como o impacto de cada força/fator que a governa.

Estes setores que formam a sociedade não estão necessariamente demarcados apenas por setores sociais, econômicos, culturais e religiosos. E, em vista dos recursos naturais (e tecnológicos) que prevalecem, o impacto de cuidados, gerenciamento e prática com relação à saúde, juntamente com tendências ecológicas e progressos ambientais, devem ser cuidadosamente incorporados a esta formulação abrangente. Além disso, medidas mais específicas, como o padrão de vida no sentido socioeconômico ou QVA (qualidade de vida ajustada) em um problema específico de risco de saúde (como câncer), não devem ser excessivamente enfatizadas, mas devem ser incorporadas a um sistema mais abrangente que leva em conta todos os fatores relevantes com sua devida importância.

A interpretação de QDV pode variar bastante de um (sub)setor a outro, e, assim sendo, como em Sen (1996), precisamos dar a devida atenção a todos estes subsectores que têm relevância para QDV/QVA. As complexidades sociais, culturais e religiosas de uma sociedade moderna (independentemente de ser oriental ou ocidental), o advento da industrialização na maior parte do mundo (incluindo os países do Terceiro Mundo), o empenho do homem em conquistar a natureza através da revolução biotecnológica contemporânea, e as extensas ameaças ambientais têm tornado a vida neste planeta muito mais complexa e frágil do que há apenas algumas décadas. Confrontados com este amplo cenário, precisamos

<sup>2</sup> N.T.: ao longo do texto, a expressão *gap ratio* foi traduzida como razão de defasagem.

traçar um quadro de QDV mais complexo e isto nos motiva a representar a sua construção na Figura 1.

Figura 1: A interface QDV



As perspectivas culturais, religiosas e sociológicas na QDV já são reconhecidas há algum tempo, ainda que tenham sido dominadas pelo padrão econômico de perspectivas de vida que dependem principalmente dos diferenciais entre pobreza e riqueza. Neste contexto de economia social, a economia do desenvolvimento e a teoria econômica acabaram por se misturar, até certo ponto, para incorporar os impactos sociais, religiosos e culturais até onde fosse possível. Ainda há um amplo espaço para mais coordenação destes fatores vitais para uma avaliação de QDV mais significativa e adequada. Atualmente, reconhece-se amplamente que, neste contexto, os aspectos de envelhecimento de uma sociedade bem como suas perspectivas de saúde mental podem acrescentar um colorido diferente ao quadro de QDV e estes fatores estão também relacionados de uma forma intrincada com a ecologia e o setor ambiental.

De modo geral, a "ecoambientometria" tem afinidades com as ciências da saúde e clínicas, ainda que a relação não tenha ainda sido totalmente delineada. Os crescentes impactos da biotecnologia, industrialização e toxicologia precisam realmente ser cuidadosamente observados dentro deste contexto. Com o tripé bioestatística, ciências

ambientais e epidemiologia em evidência, a evolução das ciências de saúde pública se expressou através de uma mensagem bastante clara: precisamos trabalhar com este complexo de uma forma integrada mais do que tratar os setores isoladamente.

Por outro lado, como a Figura 1 pode sugerir, seria mais simples tratar primeiro estes três setores separadamente, examinar sua individualidade bem como os fatores comuns e incorporar estes elementos na formulação estatística das medidas QDV abrangentes. Duas ferramentas, a saber: o *planejamento estatístico* e *inferência*, desempenham um papel básico neste contexto e nas próximas seções vamos discutir estas questões em detalhes. Neste tratamento, vamos desenvolver a noção de *gap ratio* de qualidade e deficiência de QDV (DQDV) a partir de uma metodologia estatística confiável, de forma a ter uma base adequada para a avaliação da qualidade de vida ajustada (QVA). Esta avaliação nos permite vencer o hiato crescente entre os países industrializados e os países em desenvolvimento em seus esforços para compreender QDV em um sentido amplo e interpretá-la de uma maneira racional e comparável. Esta é a meta primária deste estudo estatístico.

### **3. IMPACTOS SOCIAIS, ECONÔMICOS, CULTURAIS E RELIGIOSOS SOBRE QDV**

Em uma perspectiva socioeconômica, a qualidade de vida tem sido tradicionalmente julgada por um padrão convencional de medidas em que o conceito de renda real ou riqueza desempenha um papel básico. A avaliação de riqueza (renda real) tem sido há muito tempo uma importante medida comparativa, não apenas com finalidades de avaliação de receita ou taxaço, mas também para avaliar a solvência monetária global. Isto nunca foi uma tarefa fácil, mesmo para os países mais desenvolvidos economicamente.

Definição e mensuração de renda real envolvem a arbitragem de várias tendências sociais, culturais, religiosas e econômicas, e até mesmo de perspectivas de saúde e meio ambiente, de forma que a demarcação das assim chamadas linhas de pobreza e riqueza pode não ser tão simples, qualquer que seja o meio. Por exemplo, as rendas reais dos domicílios nos setores agrícola, industrial e administrativo podem envolver normas diferentes, e, neste aspecto, os diferenciais urbanos e rurais entre os padrões de renda podem ser relevantes. Assim, a receita, os salários, as remunerações e outras compensações monetárias em uma fazenda podem não ser adequados para avaliar a renda real de forma estatisticamente confiável. Além disso, o sistema de valores pode ser consideravelmente diferente entre um país socialista e um país capitalista, e ainda mais em um país do Terceiro Mundo.

Neste contexto, o sistema de atendimento à saúde ou um sistema de seguridade social predominantes podem tornar as situações bastante diferentes. As crenças e práticas religiosas também fazem uma grande diferença. Os mórmons nos Estados Unidos são conhecidos por terem um estilo de vida diferente, harmonizado com suas práticas e crenças religiosas, como os judeus ortodoxos em Israel e outros exemplos.

Em um país do Terceiro Mundo como a Índia, as religiões moldaram profundas atitudes filosóficas em relação aos estilos de vida (mais no passado do que agora).

Na maioria dos países latino-americanos, as populações racialmente integradas e culturalmente homogêneas, em sua maioria religiosas, parecem diferir consideravelmente quanto aos seus estilos de vida em relação à Espanha, Portugal ou outro país europeu, e mais acentuadamente quanto aos países da América do Norte. Estes diferenciais podem induzir a uma grande variação na interpretação da relação pobreza/riqueza, e na demarcação das assim chamadas linhas de pobreza ou riqueza, quando baseadas apenas na variável renda. Por esta razão, parece necessário levar em consideração várias tendências sociais, culturais e religiosas para a definição e interpretação do padrão de vida. Outras medidas de QDV relativas a outros setores também podem ter impacto neste delicado problema de mensuração. As considerações estatísticas são da maior importância nesta tarefa de avaliação.

Pressupondo que a renda/riqueza reais possam ser avaliadas de forma razoavelmente adequada em uma organização socioeconômica, a *pobreza* é definida como sendo a *extensão em que os indivíduos em uma sociedade (ou comunidade/país) caem abaixo de um padrão mínimo aceitável de vida*, de modo que pode ser quantificada em termos da proporção e distribuição da renda das pessoas pobres. Por outro lado, a *riqueza* é definida em termos da proporção e distribuição de renda das pessoas muito ricas, e, limitando estas duas classes, intermediariamente encontram-se as pessoas da *classe média* que estão de conformidade com um padrão mediano neste espectro de padrão de vida. Como foi enfatizado anteriormente, a demarcação da pobreza, da classe média e da riqueza não é uma tarefa fácil até mesmo para os países mais avançados, e através do simples ajuste das taxas de transação monetária ou fatores do tipo PIB não se pode tomar uma resolução simples neste contexto. Por isso, são necessárias salvaguardas para induzir e ajustar os impactos dos vários fatores culturais e religiosos nesta mensuração da renda real e definição do padrão de vida.

Em um sentido convencional, as pessoas da classe média devem ter um padrão mínimo aceitável de vida, enquanto que a classe rica está bem acima da linha. Assim, o ponto crucial do problema é examinar a extensão da pobreza não apenas em termos de padrões monetários, mas também com relação à qualidade de vida geral dos pobres quando comparados com as pessoas da classe média. Vistas deste ângulo, educação, recreação e outras amenidades, que são normais para as pessoas da classe média, também deveriam ser levadas em consideração no ajuste da renda real das pessoas pobres, de modo que uma medida convencional de seus padrões de vida refletisse sua QDV em uma extensão satisfatória.

Suponhamos agora que, com as salvaguardas adequadas na interpretação e mensuração da renda real ( $Y$ ) para uma sociedade/comunidade, tenhamos uma distribuição de renda  $F(y)$ ,  $y \geq 0$  e que  $\rho$  e  $\pi$  sejam as linhas de pobreza e de riqueza, também fixadas em uma base realística confiável. Então:

$$\alpha = \alpha_\rho = F(\rho) \quad \text{e} \quad \gamma = \gamma_\pi = 1 - F(\pi) = \bar{F}(\pi) \quad (3.1)$$

serão, respectivamente, as proporções de pessoas pobres e ricas. Estas são medidas bastante grosseiras de pobreza e riqueza, uma vez que não levam em consideração a desigualdade dentro da classe.

Observem que a distribuição de renda dos pobres é dada por  $F_\alpha(y) = \alpha^{-1} F(y)$  ou 1, consoante seja  $y \leq \rho$  ou não. Assim sendo, a *renda média dos pobres* é dada por

$$\mu_\alpha = \int_0^\rho y dF_\alpha(y) = \rho - \alpha^{-1} \int_0^\rho F(y) dy. \quad (3.2)$$

Portanto, a razão de defasagem de renda dos pobres é dada por

$$\beta_\alpha = (\rho - \mu_\alpha) / \rho = (\rho\alpha)^{-1} \int_0^\rho F(y) dy. \quad (3.3)$$

Na expressão acima,  $\beta_\alpha$  mede a defasagem relativa, ao comparar  $\mu_\alpha$  com a linha de pobreza  $\rho$ , ainda que possa não captar inteiramente a desigualdade de renda entre os pobres, o que pode ser retratado por alguma medida de dispersão. Neste contexto, o *coeficiente de Gini* tem emergido como uma medida bastante prática. Para uma distribuição de renda  $F$  definida sobre  $R^+ = (0, \infty)$ , o coeficiente de Gini  $G(F)$  é uma medida da desigualdade de renda em uma escala relativa, e é definido como

$$\begin{aligned} G(F) &= \{E|Y_1 - Y_2|\} / \{EY_1 + EY_2\} \\ &= [1/(2EY_1)] E|Y_1 - Y_2|, \end{aligned} \quad (3.4)$$

onde  $Y_1, Y_2$  são variáveis aleatórias independentes, com mesma função de distribuição  $F$ . Observe que  $G(F)$  é invariante quanto à escala.

Como notação, seja  $G_\alpha = G(F_\alpha)$  e  $G_\gamma = G(F_\gamma)$ . Então, tipicamente, um índice de pobreza é baseado no terno  $(\alpha, \beta_\alpha, G_\alpha)$ . Com base em um conjunto de axiomas (econômicos), A. K. Sen (1976) apresentou os seguintes índices de pobreza:

$$\begin{aligned} \xi_A^{(1)} &= \alpha\beta_\alpha = \text{proporção de pobres ajustada pela razão de defasagem de renda;} \\ \xi_A^{(2)} &= \alpha \{ \beta_\alpha + (1 - \beta_\alpha)G_\alpha \} = \alpha \{ G_\alpha + (1 - G_\alpha)\beta_\alpha \}. \end{aligned} \quad (3.5)$$

Entre outras medidas alternativas, Sen (1986) propôs

$$\xi_A^{(3)} = \alpha \beta_\alpha^{1-G_\alpha}, \quad (3.6)$$

em que foi usada a média geométrica para substituir a média aritmética em (3.5), sendo que os aspectos de robustez destes índices também foram estudados. Segue-se que

$$0 < \xi_A^{(1)} \leq \xi_A^{(3)} \leq \xi_A^{(2)} \leq \xi_A^{(1)} (2 - \beta_\alpha) \leq \alpha. \quad (3.7)$$

Pode-se observar que, para medir a dispersão relativa de uma distribuição, há outras alternativas ao coeficiente de Gini, e, entre elas, os dois coeficientes que se seguem são os mais usados na prática.

(a) Coeficiente de variação: Seja  $\sigma^2$  a variância e  $\mu$  ( $>0$ ) a média de uma variável aleatória não negativa. Então o coeficiente de variação é definido por

$$v = \sigma / \mu = (1/EY_1) \{E(Y_1 - \mu)^2\}^{1/2}. \quad (3.8)$$

(b) Coeficiente de desvio absoluto: O desvio absoluto médio em relação à média  $\mu$  é definido como  $\xi = E |Y_1 - \mu|$  e o coeficiente de desvio absoluto é dado por

$$v^* = \xi/\mu = (1/EY_1) E |Y_1 - \mu|. \quad (3.9)$$

Também é possível substituir a média  $\mu$  pela mediana  $\mu^0$  em  $\xi$  e definir  $v^*$  como  $v^0 = E |Y_1 - \mu^0|/\mu^0$ . Observe que  $\sigma \geq E |Y - \mu| \geq E |Y - \mu^0|$ , e conseqüentemente  $v \geq v^* > (\mu^0/\mu) v^0$ . Então, sempre que  $\mu^0 \geq \mu$  (como é provável que seja para a distribuição de renda dos pobres quando há uma proporção considerável de pessoas pobres com renda nula), temos  $v \geq v^* \geq v^0$ . Também se observa que  $E |Y_1 - Y_2| \leq \{E(Y_1 - Y_2)^2\}^{1/2} = \sqrt{2} \sigma$ , de forma que  $G(F) \leq \sigma / \sqrt{2} \mu = (1/\sqrt{2}) v$  ou  $v \geq \sqrt{2} G(F) \geq G(F)$ . De modo análogo a (3.6), podemos então introduzir os seguintes índices de pobreza:

$$\begin{aligned} \xi_A^{(4)} &= \alpha \beta_\alpha^{1-(1/\sqrt{2})v_\alpha}, \\ \xi_A^{(5)} &= \alpha \beta_\alpha^{1-(1/\sqrt{2})v_\alpha^*}, \\ \xi_A^{(6)} &= \alpha \beta_\alpha^{1-(1/\sqrt{2})v_\alpha^0} \end{aligned} \quad (3.10)$$

onde  $v_\alpha$ ,  $v_\alpha^*$  e  $v_\alpha^0$  referem-se às medidas  $v$ ,  $v^*$  e  $v^0$  para a distribuição de renda ( $F_\alpha$ ) dos pobres.

Observe que

$$v_\alpha \geq v_\alpha^* \geq v_\alpha^0 \Rightarrow \xi_A^{(4)} \geq \xi_A^{(5)} \geq \xi_A^{(6)}, \quad (3.11)$$

e assim, deste ponto de vista, por comparação com (3.7), podemos considerar  $\xi_A^{(6)}$  como um competidor para  $\xi_A^{(3)}$ . Lembre que a grandeza amostral correspondente a  $v_\alpha^0$  envolve a mediana e o desvio absoluto médio em relação à mediana, e por isso precisa de condições menos rigorosas sobre os momentos de  $F_\alpha$ . Neste caso, na definição de  $\beta_\alpha$ , também podemos substituir  $\mu_\alpha$  por  $\mu_\alpha^0$  e, em  $v_\alpha^0$ , substituir  $E|Y - \mu_\alpha^0|$  pela mediana de  $|Y - \mu_\alpha^0|$ , de modo que a estatística resultante apoiar-se-á em condições mínimas sobre os momentos de  $F_\alpha$ .

Então, fazendo  $\beta_\alpha^0 = (\rho - \mu_\alpha^0) / \rho$  e  $v_\alpha^{0*} = \text{med}\{|Y - \mu_\alpha^0|\} / \mu_\alpha^0$ , podemos considerar o índice de pobreza

$$\xi_A^{0*} = \alpha (\beta_\alpha^0)^{1-(1/\sqrt{2})v_\alpha^{0*}} \quad (3.12)$$

e defender seu uso do ponto de vista de robustez. Entretanto, para a distribuição de renda dos pobres, temos um limite superior finito ( $\rho$ ) e um limite inferior (0) e então as condições sobre momentos são, em geral, não restritivas. Mas os efeitos de contaminação nas caudas são muito menores com  $\xi_A^{0*}$  do que com  $\xi_A^{(3)}$  (ou mesmo com  $\xi_A^{(6)}$ ). Assim, no caso de ocorrência de erros de medida para a distribuição de renda dos pobres, nós recomendaríamos o uso de  $\xi_A^{0*}$  ao invés dos outros índices.

Vamos agora, brevemente, delinear o caso dos índices de riqueza. Tendo em vista o fato de que a distribuição de renda  $F_\gamma(y)$  ( $= [F(y) - F(\pi)] / [1 - F(\pi)]$  ou 0, consoante  $y \geq \pi$  ou não), é truncada à esquerda, uma formulação um pouco diferente dos *índices de riqueza* pode ser desejável. Note que a renda média dos ricos é igual a

$$\mu_\gamma = \int_\pi^\infty y dF_\gamma(y) = \pi + \gamma^{-1} \int_\pi^\infty F(y) dy. \quad (3.13)$$

Como  $\pi < \mu_\gamma$ , a razão de defasagem de renda dos ricos foi definida por Sen (1988) como

$$\beta_\gamma = (\mu_\gamma - \pi) / \mu_\gamma = 1 - \gamma \pi \left[ \int_\pi^\infty y dF(y) \right]^{-1} \quad (3.14)$$

e esta expressão tem uma clara aparência de *média harmônica*. Tendo em vista o truncamento à esquerda, uma *função utilidade não negativa*  $u(t_1, t_2)$ , não crescente em cada  $t_1, t_2$ , foi incorporada à formulação de um coeficiente de Gini que leva em conta a utilidade (correspondente a (3.4)):

$$G_{F,u}^* = E [u(Y_1, Y_2) |Y_1 - Y_2|] / E [u(Y_1, Y_2) (Y_1 + Y_2)]. \quad (3.15)$$

Em particular, fazendo  $u(t_1, t_2) = (t_1 t_2)^{-1}$ , temos o *coeficiente de Gini harmônico*

$$G_{F,H}^* = E [|Y_1 - Y_2| / (Y_1 Y_2)] / E [(Y_1 + Y_2) / (Y_1 Y_2)] \quad (3.16)$$

e esta medida, para a distribuição de renda  $F_Y$ , é indicada por  $G_{Y,u}^*$ . Também definimos a renda média harmônica dos ricos por

$$\mu_{Y,H} = \gamma [ \int_{\pi}^{\infty} y^{-1} dF(y) ]^{-1}, \quad (3.17)$$

e isto leva a uma medida alternativa para a razão de defasagem de renda dos ricos:

$$\beta_{Y,H} = 1 - (\pi / \gamma) [ \int_{\pi}^{\infty} y^{-1} dF(y) ]. \quad (3.18)$$

Então, de modo análogo a (3.5) - (3.6), podemos considerar os índices de riqueza:

$$\xi_Y^{(1)} = \gamma \beta_{Y,H}; \quad \xi_Y^{(2)} = \gamma \{ \beta_{Y,H} + G_{Y,H}^* (1 - \beta_{Y,H}) \}; \quad \xi_Y^{(3)} = \gamma \{ \beta_{Y,H} \}^{-1} G_{Y,H}^*. \quad (3.19)$$

Recomendamos Sen (1988) para uma discussão detalhada e alguma motivação. Neste contexto, também mencionamos uma medida bastante relevante, conhecida como coeficiente de *Gastwirth*, definida por

$$G_{F,0} = E [|Y_1 - Y_2| / (Y_1 + Y_2)]. \quad (3.20)$$

Diferentemente dos coeficientes de Gini, embora  $G_{F,0}$  não possa ser diretamente obtido da *Curva de Lorenz* (para  $F$ ), ele tem uma interessante propriedade de invariância, ou seja, tanto para  $Y$  como para  $Y^{-1}$ ,  $G_{F,0}$  é o mesmo. Da mesma forma que em (3.8) a (3.12),

também podemos considerar outros índices de riqueza. Também sob este aspecto, o análogo de (3.12), baseado na renda mediana das pessoas ricas  $\mu_Y^o$  e na mediana de  $|Y - \mu_Y^o|$  para a mesma distribuição é mais robusto. Tendo em vista o fato de que o limite superior da distribuição de renda dos ricos é  $+\infty$ , de modo que os momentos podem nem sequer existir, recomendaríamos esta medida ao invés daquelas apresentadas em (3.19). Para as pessoas da classe média, tais índices não são muito apropriados e outras medidas de padrão de vida podem ter interpretações estatísticas melhores.

Observamos que na avaliação da verdadeira renda ou riqueza existe uma característica qualitativa proeminente na qual os índices acima considerados estão baseados. Portanto, pode ser mais razoável introduzir um índice que permita acomodar estas tendências qualitativas de um modo flexível. Isto é particularmente desejável quando a variável renda não é tão bem definida, e, no caso da saúde, como veremos mais tarde, isto pode constituir um problema sério que pode afetar a interpretação dos índices já considerados. Mais ainda, no processo real de coleta e registro de dados, ao invés de uma variável resposta contínua, podemos ter um conjunto de intervalos (de classes) disjuntos. Então, é preciso incorporar as modificações devido a situações como essa.

Uma maneira de tratar esses modelos com observações fora do padrão usual seria utilizar o clássico índice de Gini-Simpson (Simpson, 1949) que, embora tenha sido originalmente proposto como uma medida de biodiversidade, pode ser usado neste contexto sob modificações convenientes (Sen, 1998). Consideremos um modelo de dados categorizados relacionado a  $c$  ( $\geq 2$ ) categorias, com as correspondentes probabilidades  $P_1, \dots, P_c$ . O índice de diversidade de Gini-Simpson é definido como

$$I_{GS} = 1 - \sum_{j=1}^c P_j^2 = \sum_{j=1}^c P_j(1 - P_j),$$

que corresponde à probabilidade de que dois indivíduos, escolhidos ao acaso nesta estrutura categorizada, não pertençam à mesma categoria.

Embora esta medida tenha algumas propriedades interessantes, ela não associa nenhum valor de utilidade às diferentes categorias. Para contemplar este aspecto, Sen (1998) considera uma função utilidade  $u(i)$ ,  $i=1, \dots, c$  e propõe o seguinte índice de Gini-Simpson orientado para utilidade:

$$I_{GSU} = \sum_{j=1}^c u(j)P_j(1 - P_j).$$

Na situação mais simples em que as categorias se referem a "fractais" da distribuição de renda relativa (à linha de pobreza) definida no intervalo  $(0,1)$ , os pesos  $u(j)$  podem ser

tomados como  $1 - \mu_j$ , onde  $\mu_j$  se refere ao valor central da renda na categoria correspondente. Então podemos fazer

$$I_{GSU} = \beta_0 - \sum_{j=1}^c (1 - \mu_j) P_j^2,$$

onde  $\beta_0$  é a razão de defasagem associada à renda. Deste modo, podemos considerar aspectos qualitativos mantendo a abordagem quantitativa. Para maiores detalhes, consultar Sen (1998).

Com base na discussão acima, podemos depreender que estes índices dependem da distribuição de renda subjacente de uma maneira relativamente complexa (isto é, não simplesmente através das proporções  $\alpha$  ou  $\gamma$ , ou através da renda média  $\mu_\alpha$  ou  $\mu_\gamma$ ) e, deste modo, eles tentam ajustar melhor a desigualdade de renda dentro dos respectivos setores. Esta característica é muito mais importante em nossa interpretação dos índices QDV, a ser detalhada nas três próximas seções. Todavia, a interpretação de pobreza e riqueza deixa muito espaço para os ajustes em relação à saúde, ambiente, psicologia e impactos culturais, e iremos abordar alguns destes fatores nas duas próximas seções.

#### 4. PERSPECTIVAS EM QDVRS

Em alguns países industrializados, foram propostas medidas de QDVRS (qualidade de vida relacionada à saúde) no Setor de Saúde (cuidados, serviços e gestão) como ferramentas úteis de avaliação de qualidade de vida para indivíduos já atingidos por uma doença ou distúrbio específico, para os quais a QDV seria utilizada para avaliação das conseqüências. A Divisão de Saúde Mental da Organização Mundial de Saúde-OMS- realizou um projeto para medir (e melhorar) a QDV e adotou a seguinte interpretação (OMSQDV, 1993/94):

*Qualidade de Vida é definida como sendo uma percepção dos indivíduos de sua posição na vida, no contexto da cultura e do sistema de valores em que vivem e em relação a suas metas, expectativas, padrões e preocupações.*

É um conceito de grande alcance, afetado de forma complexa pela saúde física, estado psicológico, nível de independência, relações sociais e suas relações com as características evidentes de seu ambiente. Basicamente, o instrumento OMSQDV é projetado para medir a qualidade de vida relacionada com a saúde e com cuidados com a saúde.

Embora a interpretação da QDV da OMS seja certamente um passo novo e ousado para uma visão mais abrangente de QDV, ela possui algumas desvantagens. Uma medida como esta relaciona-se especificamente com uma doença ou com um distúrbio no setor de saúde e se aplica a indivíduos que são afetados por este problema. Por esta razão, o conceito de qualidade de vida ajustada (QVA) é muitas vezes adaptado pela comunidade médica como uma medida dos ajustes médicos e psicológicos que são necessários para induzir uma QDV

palpável para pacientes que estão acometidos por estes problemas de saúde. Ainda assim, há vários tipos de doenças e distúrbios que podem ter diferentes tipos de QVA. Além disso, indivíduos podem estar simultaneamente afetados por mais de um destes fatores. Ainda nesta situação, os fatores socioeconômicos, culturais, religiosos e ambientais podem ter um profundo impacto sobre a QDV e por isso estes fatores precisam ser levados em consideração em uma avaliação quantitativa de um sistema com um grande setor qualitativo. Mais ainda, estas medidas devem referir-se à comunidade / sociedade / populações em que estes indivíduos estão envolvidos.

Pode-se notar que a interpretação de OMSQDV basicamente enfatiza a *percepção dos indivíduos* de sua posição na vida no contexto da cultura e de sistemas de valor em que vivem e em relação a suas metas, expectativas, padrões e preocupações. O instrumento OMSQDV é planejado para quantificar os padrões QDV relacionados com a saúde e com cuidados com a saúde, incorporando saúde física, estado psicológico, nível de independência, relação social e outras considerações socioambientais de uma maneira mais ampla. Neste contexto, ainda que o *padrão de vida* venha à tona em cada um destes setores, sua ênfase primária sobre o diferencial pobreza-riqueza é perdida.

Há algumas décadas, médicos e pesquisadores clínicos observaram a inadequação de medidas tradicionais de mortalidade e morbidade em sua avaliação do status de saúde relacionado com algumas doenças crônicas. Estas avaliações foram particularmente pertinentes em tomada de decisões para tratamento individual, em um contexto específico (isto é, em doenças ou grupos de pacientes) como câncer, doença cardiovascular e idade avançada. Isto levou à evolução de algumas medidas de *pouca saúde* em base individual para tipos específicos de doenças ou distúrbios. A *Escala Karnofsky* em câncer (Karnofsky & Burchenal, 1949) e a Escala da Associação Americana de Reumatismo (Escala ARA) para artrite (Steinbrocker, Traeger & Battman, 1949) são exemplos notáveis de precursores de QDV. Este setor de cuidados com a saúde tem sido a principal área enriquecida por estudos de QDVRS, e alguns dos recentes desenvolvimentos aí ocorridos são citados na bibliografia deste artigo. O *Centro Nacional de Recursos de Serviços de Saúde* dos EUA -National Center for Health Services Resources, USA- desenvolveu esquemas para estudar a eficiência e os custos econômicos dos serviços de saúde nos Estados Unidos e logo estudos semelhantes foram iniciados em outros países desenvolvidos. A situação é bastante diferente nos países do Terceiro Mundo onde, em vista da escassez de recursos, há uma necessidade ainda maior destes levantamentos. Contudo, a OMS também está comprometida com os países do Terceiro Mundo.

O enfoque primário destas medidas de QDVRS tem sido sobre os usos alternativos de recursos e seus benefícios. Em países industrializados, apesar da disponibilidade de mão-de-obra e de recursos tecnológicos adequados, o custo dos serviços de saúde está subindo com uma rapidez alarmante, muito além do que a maior parte da população pode pagar e, assim, os aspectos de custo/benefício dos cuidados com a saúde estão na linha de frente das

considerações das agências reguladoras e dos médicos. Neste contexto, qualquer uso alternativo de recursos e seus benefícios exige, naturalmente, um exame minucioso dos pontos de vista médico e social, bem como do econômico, e as medidas de QDV estão no centro destas explorações.

A primeira e mais importante das tarefas na avaliação da QDV para os indivíduos afetados por uma doença ou distúrbio específico, mesmo quando tratado isoladamente, é o desenvolvimento de ferramentas estatísticas apropriadas. As principais dificuldades primárias neste aspecto resultam do fato de que geralmente há muitas variáveis resposta, que são relevantes para as avaliações de QDV: muitas vezes há interação entre estas variáveis resposta que podem até mesmo ter uma natureza competitiva. Assim, a questão de *sinergismo* permanece pertinente em questões de QDV. Além disso, muitas destas variáveis resposta podem ser binárias/politômicas ou ordinais, para as quais os *erros de medida* e/ou classificação errônea de estados podem ser prevalentes. Nesse caso, devem ser usados modelos e ferramentas estatísticas não convencionais. Além das variáveis resposta, a escolha das variáveis *auxiliares* e *concomitantes* deve ser feita a partir de um número às vezes excessivo delas. Nem todas as variáveis podem qualificar-se como variáveis concomitantes, e, por outro lado, podem trazer mais complexidades para o modelo estatístico e técnicas de análise. Do ponto de vista de coleta de informações, é claro que é preferível um conjunto maior, mas, do ponto de vista de análise estatística, isto pode pedir uma amostra de tamanho muito maior para superar o nível de flutuações amostrais.

Em segundo lugar, no atual momento, as avaliações de QDV são geralmente feitas com relação a objetivos específicos que muitas vezes não são formulados com exatidão antes do experimento ou da coleta de dados. Isto pode levantar uma séria preocupação com relação à *validade* e à *confiabilidade* das ferramentas estatísticas convencionais que geralmente são incorporadas nesta busca. Por exemplo, os questionários padrão dos levantamentos de QDV consistem de um conjunto de questões numa escala qualitativa de 5 pontos (geralmente com uma ordenação parcial) e agrupadas para diversos setores amplos: (i) bem-estar físico, (ii) bem-estar social/familiar, (iii) bem-estar funcional, (iv) bem-estar emocional e (v) relação com o pessoal médico e paramédico de atendimento, entre outros. Naturalmente, a importância relativa destes setores pode variar consideravelmente de uma doença ou distúrbio a outro, e também com idade/sexo e outros fatores socioeconômicos que afetam os indivíduos sob consideração. Por isso, há necessidade de enfatizar adequadamente as questões mais importantes e relevantes para uma doença ou distúrbio específico sob estudo.

É pouco provável que o mesmo questionário possa ser inteiramente apropriado para duas ou mais doenças/distúrbios, mesmo que pertençam a uma classe comum. Por exemplo, os sintomas de câncer de pulmão, próstata e mama/útero/ovário são diferentes bem como as correspondentes variáveis resposta e os protocolos de tratamento. Em consequência disto, uma primeira e importante tarefa é identificar, principalmente a partir de variáveis médicas e terapêuticas, os fatores mais relevantes e, sob esta ótica, organizar o questionário de tal

forma a obter a maior cobertura possível. Observe que, para ser receptivo e eficiente, um questionário deve ser o mais simples possível e, ao mesmo tempo, o mais curto possível. Neste aspecto, o *planejamento estatístico* ou a *construção* de um questionário é uma tarefa importante. Infelizmente, até recentemente, as agências / sistemas de saúde ou de cuidados médicos não davam muita atenção a esta tarefa vital, mas do ponto de vista estatístico não podemos negar que isto é essencial.

Um outro aspecto importante da interpretação da OMSQDV é aquele que se refere a um indivíduo já acometido por uma doença/distúrbio específico e por isso espera-se que tenha uma QDV abaixo do padrão bem como uma *vida residual* reduzida. Esta característica da avaliação de QDV pode por isso relacionar-se a uma variedade de fatores indicados a seguir:

(i) *Variáveis resposta multivariadas* cobrindo diversos aspectos do distúrbio.

(ii) Um *esquema de acompanhamento* para traçar o progresso ao longo do tempo, resultando em *projetos com medidas repetidas* ou *estudos longitudinais*.

(iii) Devido à possibilidade de *não conformidade* (recusa ou desistência) ou a outras razões administrativas, a perda ou retirada de observações são geralmente comuns, e esta *censura* ou falta de informação pode não ser puramente *aleatória*. Assim, *esquemas de censura aleatória* podem não ser pertinentes.

(iv) Como foi explicado anteriormente, devido à natureza das variáveis resposta e seus padrões de distribuição, o recurso a alguns *modelos paramétricos* típicos pode não ser adequado para tirar conclusões estatísticas. Mesmo os modelos *lineares*, *log-lineares* ou *lineares generalizados* podem muitas vezes ser inadequados ou não apropriados neste contexto. Os assim chamados modelos *semiparamétricos* também estão sujeitos a críticas similares, ainda que em menor extensão. Ainda que os modelos *não-paramétricos* pareçam ter um escopo maior para adoção neste contexto, deve ser claramente lembrado que, em vista das suposições de regularidade não usuais requeridas, cuidados adicionais poderão ser necessários para validar tais procedimentos. Isto pode resultar em um espaço paramétrico muito maior, e, assim sendo, pode requerer uma amostra muito maior para que as conclusões estatísticas possam ser tiradas com uma precisão razoável. Por exemplo, para um mesmo indivíduo, as observações em momentos sucessivos ao longo do tempo podem não ser independentes nem identicamente distribuídas, e, além disso, em cada momento no tempo pode haver múltiplas mensurações que são estatisticamente dependentes entre si. Por isso, um modelo estatístico abrangente deve incorporar cuidadosamente todas estas possibilidades. Por isso, *modelos multivariados não-paramétricos* podem ser potencialmente mais úteis neste caso.

(v) As avaliações de QDV ao longo de um período de tempo (como geralmente é o caso) devem refletir também o padrão de mudança ao longo do tempo, de modo a facilitar os ajustes clínicos e/ou administrativos para atualizar, tanto quanto possível, o quadro QDV/QVA. Muitas vezes, a ética e a prática médica correm em sentido contrário ao das intuições estatísticas (porque são baseadas em perspectivas diferentes). Como as tendências

biomédicas são dominantes nas avaliações de QDV, cabe aos estatísticos o desafio representado pela tarefa de adquirir estes conhecimentos com vistas ao desenvolvimento de uma metodologia apropriada levando em conta tais complexidades.

(vi) No atual estágio de desenvolvimento, estas avaliações de QDV são direcionadas principalmente para os recursos de serviços alternativos de saúde e seu custo-benefício. Contudo, vista da perspectiva da população/sociedade, a maioria das medidas simples de QDV disponíveis na literatura podem requerer modificações e generalizações para acomodar um objetivo mais complexo.

Tendo em mente todos estes fatores, consideremos a seguinte formulação de um modelo para avaliação de QDV relacionado a uma única doença/distúrbio de um ponto de vista individualista. Falando de forma mais ampla, temos um cenário de avaliação de QDV com *quatro fases*:

1. *Estágio pré-terapia*: avaliações de QDV muitas vezes levam a um rumo terapêutico adequado.

2. *Estágio de terapia em andamento*: estar atento a toxicidade/efeitos colaterais e opções de emergência.

3. *Estágio pós-terapia I*: enquanto ainda sob cuidados médicos, avaliação das ações terapêuticas sobre QDV.

4. *Estágio pós-terapia II*: acompanhamento de longo prazo sobre QDV.

Em cada estágio, a avaliação de QDV envolve numerosas variáveis resposta juntamente com variáveis auxiliares e concomitantes adequadas. Estas variáveis podem não ser todas de natureza contínua e algumas vezes elas nem mesmo são quantitativas. Discutiremos as questões estatísticas relevantes em uma outra seção. Ainda assim, há algumas considerações importantes pertinentes a este protocolo de avaliação, e relacionamos algumas delas a seguir:

(i) Avaliação da QDV para um indivíduo em relação a um grupo apropriado de indivíduos experimentando um distúrbio similar.

(ii) Avaliação da QDV para um indivíduo em relação à população pertencente ao mesmo grupo coorte com base em covariáveis adequadas (como idade, sexo, etc.).

(iii) A ênfase em QDV/QVA vs. a *duração da vida* após a incidência de um distúrbio/doença específico. Há questões éticas genuínas relativas ao prolongamento do tempo de vida, sob condições mínimas de qualidade de vida. Neste contexto, destacam-se as perspectivas para as doenças mentais.

Vamos passar agora para um outro problema importante de QDV, mas estatisticamente mais complexo. Este quadro de QVA relaciona-se com um estilo de vida ajustado, baseado na qualidade de diversas características físicas, sociais, psicológicas e de saúde mental, após a ocorrência de um distúrbio ou presença de uma doença. Os conceitos de razão de defasagem *de renda* e coeficientes de Gini de desigualdade de renda desempenham um papel vital no ajuste da proporção de pessoas pobres, com a finalidade de

indução de uma medida mais significativa de pobreza. Um quadro similar também vale para os *índices de riqueza* neste diferencial pobreza/riqueza.

Conceitualmente, uma interpretação similar de índices QDV deve ser apropriada. Mas, enquanto a renda real de uma pessoa/família pode muitas vezes ser quantificada em termos de uma variável contínua, em estudos de QDV o cenário pode ser muito mais complexo. Com relação a cada uma das características mencionadas anteriormente, a capacidade de uma pessoa comportar-se funcionalmente de modo adequado ou não é geralmente uma variável binária ou algumas vezes politômica. Assim, se levarmos em conta todos os fatores relevantes determinantes de um quadro específico de QDV, temos em geral um vetor de resposta politômica múltipla. Mesmo para um vetor de múltiplas respostas dicotômicas, a interpretação e a formulação de medidas adequadas de associação, dispersão e tendência central (das características quantitativas inerentes) podem exigir teoria e metodologia estatísticas mais sofisticadas.

Além disso, a formulação de uma razão de defasagem de renda e medidas do tipo Gini para desigualdade/concentração para tais modelos politômicos múltiplos pode encontrar outros impasses conceituais. O conceito de *renda total* determinante das medidas de pobreza/riqueza pode não ser totalmente relevante no quadro de QDV, ainda que da perspectiva de custo/benefício tal conceito pode ser racionalizado até um certo ponto. Tipicamente modelos de *dados longitudinais* ou de *esquemas de acompanhamento* podem emergir de estudos QDV/QVA e uma possível *fuga das suposições* pode acrescentar mais complexidade a esta formulação.

Além disso, a avaliação da QDV geralmente é realizada através de questionários bem planejados que servirão de base às quantificações estatísticas. Se a escala ARA ou a escala Karnofsky, geralmente adotadas no contexto de uma doença/distúrbio específico, forem utilizadas para chegar a uma medida global de QDV a partir de respostas compostas, obtidas no contexto de outra doença, tais escalas podem não ser muito apropriadas mesmo se esta última for de alguma forma relacionada a distúrbios para os quais elas foram geralmente concebidas. Ainda que o mesmo questionário (ou um muito similar) seja usado para avaliar a QDV para dois distúrbios/doenças (possivelmente relacionados), os pesos relativos a serem designados aos vários componentes podem depender do caso específico. Assim, para planejar estes questionários é essencial levar em conta as várias formas de disfunções que provavelmente surgirão nos estágios pós-terapia. Por isso, o planejamento estatístico padrão adotado em estudos agrícolas, biométricos ou médicos é muitas vezes inapropriado em estudos QDV/QVA.

Em vista disso, os esquemas de análise estatística para avaliações de QDV podem, em geral, seguir algumas rotas não padronizadas e mais complexas. Para ilustrar este ponto, vamos considerar o típico modelo *composto doença/distúrbio*. Como os infortúnios, que têm uma tendência de ocorrer em grupo, várias doenças ou distúrbios que afligem os indivíduos são muitas vezes concomitantes. Por isso, ao lidar com uma fonte primária de

doença/distúrbio, poderá ser mais adequado levar em conta possíveis fontes secundárias, para que seu impacto sobre o quadro de QDV relativo a esta fonte primária possa ser avaliado em uma maior extensão e, desta forma, possam ser feitos os ajustes adequados. Do ponto de vista do planejamento estatístico, bem como da análise, esta é em geral uma tarefa altamente complexa, e cuidados adequados se fazem necessários para a sua realização.

Podemos observar que as perspectivas de saúde estão estreitamente relacionadas com fatores de saúde mental ou psicológicos via perspectivas sociais, culturais, religiosas e econômicas. Mesmo a educação, ou alfabetização em um sentido mais amplo, é muito pertinente neste aspecto e assim a avaliação QDV/QVA em uma perspectiva de saúde precisa levar tais tendências em conta. Esta é a principal razão pela qual os fatores de saúde e psicológicos precisam ser combinados de uma maneira significativa. Com uma perspectiva diferente sobre a vida e com um ambiente econômico diferente, pode ser mais fácil ou mais difícil ajustar para QDV de forma que um quadro QVA pode depender muito desses fatores de retaguarda. Neste aspecto, o quadro difere drasticamente entre as nações industrializadas de um lado e os países do Terceiro Mundo de outro, onde estas tendências são sufocantes. Isto pode exigir uma determinação de QDV com maiores complexidades e também com maiores incertezas. Ainda que a maioria dos fatores econômicos e tecnológicos possam ser bem delineados, a situação com os fatores sociais, culturais e religiosos pode ser bem mais complexa.

Para ilustrar este ponto, tomamos a liberdade de olhar para o quadro de QDV de países do Leste Europeu alguns anos atrás, quando a antiga União Soviética ainda detinha o controle econômico e político sobre toda a região. Como tendência geral, a tecnologia do consumidor não acompanhou a tecnologia científica e, em muitos casos, a industrialização ainda não estava em sua plena expansão. Isto criou problemas para a manutenção de um padrão de vida que pudesse ser comparável com o dos países ocidentais e, além disso, para os intelectuais a agonia mental deve ter sido ainda maior. Não houve setor da população que tivesse ficado realmente imune a este problema. Garantidos uma liberdade apenas limitada do uso da palavra e outros direitos humanos, o sentimento geral das pessoas comuns era de muito descontentamento e a situação da saúde mental parecia ser a pior possível. As pessoas estavam desesperadas por abandonar seu país natal e mudar-se, com grave risco, para se fixarem em lugares alternativos sem qualquer preocupação com as vantagens e desvantagens relativas. Em muitos casos, o vício do fumo ou de outras drogas intensificou-se drasticamente e também o alcoolismo surgiu como um importante problema social. Estes, por sua vez, aceleraram os problemas sociopsicológicos em uma escala de massa. Neste contexto, ficaram evidentes os impactos deste tipo de inquietação mental e social sobre a sobrevivência (sem falar em padrão de vida) e sobre a habilidade em combater diversos tipos de problemas de saúde que surgiam.

A unificação das duas Alemanhas está relacionada com esta mudança no padrão de vida bem como na qualidade de vida para os dois componentes, embora em direções opostas

para as duas partes. Enquanto a Alemanha pôde absorver este tumulto, o mesmo não ocorreu em outros países. Um fator notável neste fluxo de eventos é a ressurreição do papel da religião na vida diária, que estava faltando na era soviética, e isto por sua vez criou inquietação e grandes conflitos em outros países, como na antiga Iugoslávia. Em um caso assim, qualquer avaliação da QDV/QVA em saúde ou em outro setor pode, por isso, precisar de um exame mais detalhado das perspectivas psicológicas resultantes do caos social ou político. Qualidade de educação no nível das escolas elementar e avançada é também um importante fator que afeta este cenário. Ao inclinar-se mais para a ciência e a educação tecnológica, juntamente com o uso cada vez mais intenso da televisão ("TV-proneness"), um número cada vez maior de sociedades está sacrificando seus valores tradicionais por um estilo de vida diferente, ensejando desta forma o aparecimento de outros problemas indesejáveis. Em muitos casos, temos pouca experiência anterior para lidar com estes tumultos sociais.

Eu gostaria de ilustrar este contexto com um evidente problema de saúde mental que poderia não ser totalmente devido ao fator "envelhecimento", ao qual comumente associamos os problemas de depressão psiquiátrica ou clínica. Com o advento turbulento da moderna eletrônica e da tecnologia de computação, o ambiente de trabalho está mudando por toda a parte e mesmo as nações industrializadas não estão imunes a este distúrbio. Professores, operários, secretárias e os que ganham a vida em numerosas outras áreas, que cresceram há alguns anos em um mundo diferente do ponto de vista computacional, estão agora enfrentando uma tremenda barreira: jovens estagiários melhor equipados tanto com computadores pessoais quanto com aparelhos eletrônicos modernos, mesmo não tendo a mesma experiência sólida de trabalho, estão assumindo os empregos e, portanto, a QDV ocupacional está mudando drasticamente.

Ainda que um bom número de membros da velha guarda esteja fazendo o melhor possível para lidar com um mundo ocupacional em mudança, aprendendo mais sobre as amenidades modernas, obviamente há outros que ficam de fora desse ambiente. Estas são as pessoas que mais provavelmente serão presa fácil para a depressão, muitas vezes caminhando em direção a uma crise nervosa completa. Estão mais propensos a ficarem viciados em cigarros, bebidas alcoólicas e a desenvolverem sérios problemas familiares (entre eles, maus-tratos contra as esposas e filhos). Se quisermos olhar o estilo de vida destas pessoas, podemos obter um quadro diferente de sua QDV. Nesse quadro podem estar misturados alguns fatores econômicos (relacionados a problemas de ambiente de trabalho), embora a maioria seja da esfera da saúde mental. A enorme tensão a que se pode estar submetido num ambiente de trabalho também pode afetar a capacidade de criar uma família adequadamente e pode até mesmo afetar a capacidade reprodutiva. Por isso, ao julgar a QDV de uma comunidade ou sociedade, não podemos ignorar estes fatores sociomentais. No Japão, a taxa de incidência de suicídios é bastante elevada entre os estudantes do segundo grau, enquanto que nos Estados Unidos o mesmo padrão é prevalente entre pessoas da área

médica. A “pobreza econômica” provavelmente não é a razão primordial para este fenômeno, mas provavelmente é a “pobreza mental” que se desenvolve devido a estresse ou esforço ocupacional, e com frequência resulta em uma depressão séria e outros problemas de saúde mental. Certamente isto não está restrito às pessoas mais velhas.

Ainda que os fatores socioeconômicos dominem este cenário, a atitude perante a vida também permanece como um importante fator neste contexto. Olhando para trás, para os países socialistas, podemos observar que a perspectiva monetária da vida e a tentação representada pelo estilo de vida baseado na moderna tecnologia de consumo sofreram uma influência indevida das transformações sociais. Ainda que possa ser muito complexo descrever este quadro em termos de uma ou algumas variáveis, podemos incorporar um importante fator nesta busca. Mantendo estes fatores socioeconômicos em mente, introduzimos aqui o conceito de *frugalidade*, que costumava prevalecer no Oriente e em países do Terceiro Mundo e que também pode ser caracterizado como *contraconsumismo*. Frugalidade é diferente de *avareza*, bem como de pobreza mental ou econômica. A pobreza material expõe a angústia física de uma pessoa e pode geralmente levar também à angústia mental. Por outro lado, uma pessoa frugal, mesmo em angústia econômica ou física, pode não estar física ou mentalmente desamparada. Assim, a frugalidade pode ser identificada como *contentamento* da atitude mental e da visão da vida.

Em contrapartida, durante os últimos trinta anos, o consumismo invadiu todos os níveis da vida em todos os setores sociais e econômicos e isto tem sido desencadeado pelo assim chamado aumento (não ajustado) no ganho per capita em todo o mundo. O consumismo afetou uma faceta das sociedades desenvolvidas, mas nos países do Terceiro Mundo ele está criando um desastre. Ironicamente, a geração da *ganância* e *corrupção* em grande escala em todo o mundo pode ter sido desencadeada pelo próprio consumismo que as nações adiantadas justificam em nome do desenvolvimento econômico. Não é verdade que o consumismo tornou as pessoas impulsivamente aquisitivas e inquietas, aumentando portanto a probabilidade de distúrbios mentais e tensões físicas?

As estruturas econômicas e sociais rurais em um país do Terceiro Mundo, baseadas primariamente nos conceitos básicos de família e comunidade próximas, estão à beira de um colapso, as relações interpessoais estão em seus piores níveis e a insegurança está amplamente disseminada por toda a sociedade. Sem uma sensação de frugalidade, como podemos estabelecer um equilíbrio entre nossos recursos naturais e tecnológicos de um lado e o padrão de vida e a qualidade de vida de outro?

O colapso do assim chamado socialismo da Europa Oriental e a transformação da China continental em um estado onde impera um pseudoconsumismo levantaram sérias preocupações nos países do Terceiro Mundo e o modelo ideal não está à vista. A frugalidade tem uma contribuição inestimável para os conceitos e o processo de envelhecimento em nossa sociedade. Ainda que os impactos psicológicos da frugalidade sejam esmagadores, não podemos negar seus impactos sobre a QDV/QVA também em outros sentidos. Para uma

pessoa que sofre de uma doença crônica, a QDV/QVA pode depender também de sua atitude em relação à vida e, nesta situação, juntamente com outros fatores psicológicos, a frugalidade é um importante ingrediente.

Nenhuma das religiões de todo o mundo tem sido uma verdadeira defensora do consumismo, e, ainda assim, a ênfase que costumava haver na frugalidade está declinando com uma rapidez alarmante. Conflitos sociais irrompendo em quase todos os cantos do mundo têm geralmente uma base religiosa dominante, mas, olhando de uma perspectiva mais ampla, fica bastante aparente que as verdadeiras religiões podem não ter contribuído para estes conflitos. É muito mais relevante a maneira pela qual interpretamos estes fatores religiosos e os adotamos na vida do dia-a-dia, posto que dão origem a estes conflitos. Por isso pode ser pertinente examinar o impacto da *tolerância* social e religiosa sobre a nossa frugalidade e qualidade de vida em geral. Não podemos deixar de ver este aspecto em nossa avaliação da QDV/QVA mesmo quando olhamos apenas para o setor saúde. Por isso, o quadro de QDV/QVA pode variar drasticamente de uma sociedade frugal para outra com maior consumismo embutido.

Os últimos anos testemunharam o crescimento de alguma literatura estatística sobre a avaliação da QDV em ensaios clínicos, doenças crônicas e câncer e algumas outras doenças correlatas. A bibliografia OMSQDV é uma excelente fonte destas referências. Nestas avaliações tem havido alguma ênfase nas determinantes sociais e culturais e elas nos proporcionam uma melhor compreensão da QDVRS. Mas ainda há uma necessidade genuína de incorporar outros fatores psicológicos e socioeconômicos com vistas à sua unificação de uma maneira mais abrangente para interpretações mais amplas e o desenvolvimento de medidas mais viáveis. Trataremos delas adiante.

## **5. PERSPECTIVAS BIOAMBIENTAIS**

---

Primeiramente, precisamos elaborar um pouco o próprio conceito e a interpretação do bioambiente. Como na Seção 1, interpretamos bioambiente como as bases ambientais e ecológicas compostas que afetam todos os biosistemas sobre a Terra, quer direta ou indiretamente, de forma visível ou latente. Como tal, as ciências da ecologia e do ambiente são dois importantes setores das ciências bioambientais, e pertencem a este domínio: a saúde ambiental, o risco ocupacional e outras disciplinas correlatas da saúde pública. Qualquer sistema, quer mecanicista ou biológico, tem um "período de vida". Um sistema mecanicista tem sua vida prolongada quando há a substituição de componentes inativos ou fracos por componentes novos. Em um sistema biológico, contudo, esta reposição pode não ser viável. Por isso, deve ser criado um bioambiente adequado para manter e até mesmo para melhorar a qualidade da vida. A vida moderna tem levado a algumas preocupações ambientais sérias e, sob a perspectiva da QDV, precisamos dar a atenção adequada a elas.

É inegável que os fatores ambientais e ecológicos dominam fortemente o cenário de QDV/QVA, ainda que não possam ser diretamente responsabilizados no sentido usual. Em um contexto específico de doença/distúrbio, a gravidade do problema pode muitas vezes ser julgada adequadamente. Por outro lado, os próprios fatores eco

ambientais podem não aparecer em um modelo de regressão convencional e, ao invés disso, relacionam-se a diversos riscos de saúde, muitas vezes de maneira muito complexa mas ainda latente e, conseqüentemente, para entender seu impacto poderá ser necessário um modelo estatístico muito mais complexo para apresentar o problema na perspectiva correta e para realizar a análise estatística relevante.

Ilustramos este ponto com um modelo de *risco ocupacional* simples, que seria a QDV dos mineiros de carvão (ou a QDV da *doença do pulmão negro*). Depois de muitos anos de agitação por parte dos mineiros de todo o mundo, seus problemas pulmonares receberam a devida atenção por parte dos governos e das agências de saúde. A mineração e a perfuração do subsolo produzem muito pó de carvão e outros gases tóxicos, de forma que o ar que os mineiros respiram durante o trabalho geralmente pode estar contaminado por partículas de carbono, monóxido e dióxido de carbono, gases tóxicos e poeira. Além disso, a contaminação da água também é bastante comum. Por isso, a toxicologia da *inalação* e da *ingestão* tem impactos importantes e um possível sinergismo também não pode ser esquecido.

Neste contexto existem boas evidências de efeitos *xenobióticos*. A prática do fumo, geralmente prevalente entre estes mineiros, e os fornos de queima de carvão para cozinhar e o aquecimento durante o inverno podem acelerar os efeitos inalatórios. Partículas de carbono, monóxido de carbono, CO<sub>2</sub> e outros tóxicos são inalados e a exposição prolongada a estes efeitos resulta em um bloqueio das vias aéreas, com as partículas mais finas movimentando-se através da faringe, laringe e traquéia, ganhando seu caminho até os alvéolos através dos bronquíolos. Nos tecidos alveolares estas partículas podem bloquear, em grande parte, a troca de oxigênio inspirado e de CO<sub>2</sub> e vapor de água expirados, desencadeando assim o mau funcionamento dos pulmões. O sistema respiratório também está intimamente ligado ao sistema cardiovascular e o aumento da toxicidade do sangue pode causar outros problemas sérios de saúde. Os problemas respiratórios geralmente permitem o desenvolvimento da tuberculose e do câncer pulmonar, e as infecções virais podem desencadear o seu progresso. Parar o trabalho depois de uma exposição razoável a esta contaminação pode não resolver por completo este problema e a *genotoxicidade* tem sido identificada como um fator determinante nesta situação.

Ao se considerar estas doenças (como câncer de pulmão, tuberculose, etc.) na avaliação das questões básicas da QDV/QVA, bem como o tempo de vida residual, pode ser bastante pertinente dar a devida ênfase a estas determinantes ocupacionais de risco. A maioria dos regulamentos sanitários instituídos nos setores de mineração dos países desenvolvidos ocupa-se destas questões de forma adequada, mas o quadro está longe de ser satisfatório nos países do Terceiro Mundo. Por isso, nos países desenvolvidos, este fator

particular pode também facilitar a regulamentação das operações de mineração de forma a minimizar os seus efeitos desastrosos.

A *asbestose* é uma outra doença que também tem sido identificada não só como um risco ocupacional para os trabalhadores na indústria de amianto mas também como um grave poluente ambiental para todos os outros que estão expostos a estas partículas em seus locais de trabalho, estudo ou moradia. Aparelhos eletrodomésticos comuns, como torradeiras, aquecedores de alimentos, secadores de cabelo, etc., que contêm placas de amianto, também podem promover estes efeitos. O grave potencial de exposição a gases tóxicos das pessoas envolvidas em trabalhos de *pintura por aerossol* também foi identificado. Por isso, há necessidade de identificar os fatores ocupacionais no estudo da QDV/QVA de diversas doenças/distúrbios que com eles têm afinidade.

Parece que, mesmo se olharmos o quadro da QDVRS em um caso específico, estas determinantes bioambientais precisam ser avaliadas não apenas para identificar a extensão de riscos que são elimináveis pelo seu controle adequado, mas também para desenvolver práticas alternativas que deverão ter efeitos menos perigosos. Estas, por sua vez, podem ser relacionadas a recursos naturais disponíveis, bem como à consciência do público em geral para o melhor uso de recursos naturais que são ocupacional e ambientalmente adotáveis.

A principal relevância do bioambiente em estudos da QDV é o seu profundo impacto sobre a viabilidade bem como a qualidade de vida de todas as espécies vivas neste planeta. Por diversas razões ecológicas, muitos animais e plantas do passado foram extintos e não é improvável que as existentes (incluindo a espécie humana) sigam o mesmo caminho, se não estivermos atentos a esta preocupação básica. Para olhar esta possibilidade destruidora de forma mais realista, precisamos avaliar as principais direções ao longo das quais é mais provável que ocorram os desastres bioambientais e isto confere relevância às ciências de saúde ambiental dentro da esfera dos estudos de QDV. Obviamente, um estudo deste tipo precisa ser realizado tendo como base causas mais abrangentes ao invés de mais individualizadas, e é o que pretendemos discutir agora através de alguns desastres bioambientais que efetivamente ocorreram com sérios impactos em nossa QDV.

Se olharmos, como deveríamos, o quadro de QDV/QVA sob uma perspectiva populacional mais ampla, torna-se necessário levar em conta o complexo de fatores ecológicos e determinantes ambientais para a população sob estudo, e isto pode nos proporcionar uma melhor compreensão das doenças/distúrbios que são desencadeados por estes fatores subjacentes. Assim, por exemplo, suspeitou-se que o lançamento de substâncias químicas na área do Love Canal próximo das Cataratas do Niágara, Estado de Nova Iorque, há alguns anos, fosse a principal causa da alta na taxa de incidência de certos tipos de câncer. Um efeito mais visível ficou evidente entre as vítimas do lançamento de bombas atômicas em Hiroshima, no Japão, há 51 anos. Em uma extensão menor, agora sabemos que o vazamento do reator nuclear de Chernobyl, Rússia, há dez anos, produziu um distúrbio similar para muitas pessoas, não apenas aquelas que trabalhavam no reator, mas também

para as que viviam nas proximidades. O vazamento de gás tóxico, que ocorreu em Bhopal, Índia, não só matou mais de 2 500 pessoas como também induziu um distúrbio catastrófico em um número incontável de pessoas que viviam na época ao redor da área. Vazamentos de petróleo em alto-mar podem ser um desastre para a vida marinha e um desequilíbrio ecológico também pode ser resultante deste problema. A queima de um grande número de campos petrolíferos, durante a Guerra do Golfo em 1991, poderia ter desencadeado a produção de uma grande quantidade de agentes tóxicos e sua liberação para a atmosfera, afetando mesmo os locais mais distantes. Os efeitos genotóxicos ainda precisam ser avaliados por completo em tais ocorrências. Muitas vezes, estes desastres são denominados de calamidades naturais, mas, em uma era de imensa toxicidade e poluentes ambientais, há uma necessidade genuína de se identificarem os efeitos latentes destes fatores, e isto pode ser incorporado na manutenção e melhoria de nossa QDV/QVA global.

Os poluentes e tóxicos ambientais não são inteiramente novos. Mas constantemente estão ganhando força, principalmente devido aos esforços humanos em mudar os estilos de vida através da industrialização e falta de previsão. A *malária*, transmitida pelo mosquito na maior parte do mundo, a *febre tifóide*, em toda a região tropical, e a *diarréia* e outras doenças intestinais agudas têm um significativo componente ambiental. Em muitos casos, a *contaminação da água* é a principal fonte de contaminação e propagação de doenças e, no passado, estas doenças tinham muitas vezes conseqüências prolongadas. O lançamento de substâncias químicas em cursos d'água elevou a toxicidade a um nível alarmante e, além disso, a contaminação por mercúrio destruiu a população de peixes em muitas áreas. Menos perceptivelmente, estes peixes ou outros frutos do mar contaminados provavelmente são levados ao consumo humano, antes que alarmes sejam dados quanto ao seu impacto. Mas, quando isto acontecer, seus efeitos sobre a saúde já terão raízes profundas.

O grupo dos medicamentos *antibióticos* disponíveis para o consumo humano há cerca de 50 anos parecia altamente eficiente contra muitas destas doenças, ainda que existam, em geral, alguns efeitos colaterais sérios. Além disso, a potência destas drogas diminuiu de forma constante depois do uso repetido, e problemas mais sérios estão surgindo com o seu uso prolongado. A *varíola* costumava ser uma doença terrível, mas, graças aos esforços incansáveis da OMS, agora estamos livres dela. Mas o uso amplamente disseminado de *DDT* e outros pesticidas altamente tóxicos para combater algumas destas doenças criou alguns outros riscos para saúde, que não eram conhecidos anteriormente.

O tratamento *hormonal* de frangos e bovinos para aumentar a produção de carne também tem sido ligado a outros problemas de saúde. Mesmo os *corantes* ou pigmentos para alimentos não estão livres de riscos para a saúde. O uso amplamente disseminado de canos de chumbo para água resultou em muitos casos em uma toxicidade lenta mas séria, afetando milhões de pessoas de forma invisível. A presença de chumbo em tintas para paredes causou efeitos similares. A substituição dos canos de chumbo por plástico (ou material de PVC) também levantou questões sérias relativas a toxicidade e efeitos carcinogênicos. O

lançamento de *resíduos químicos* em aterros inadequados ou cursos d'água levou à contaminação do subsolo e da água, além de muitos problemas sérios de carcinogenicidade. A presença de fumaça no ambiente tem sido identificada como uma importante ameaça em muitas áreas urbanas. Mesmo o problema de destruição de *dejetos humanos* pode ter impactos sérios sobre o quadro de QDV/ QVA.

As refinarias de petróleo aumentaram ainda mais a quantidade de enxofre veiculado pelo ar, as deposições de carbono veiculado pelo ar provenientes das usinas termoelétricas aumentaram ainda mais a incidência da doença do pulmão negro, enquanto que as usinas nucleares têm efeitos radioativos potenciais desastrosos. O lançamento de substâncias químicas e a exploração de petróleo na plataforma oceânica têm levado a sérios desastres ecológicos para a biologia marinha e a camada de ozônio cada vez mais fina, além do *efeito estufa*, são agora ameaças globais. Automóveis, aviões e chaminés de fábricas têm aumentado muito a ocorrência de doenças respiratórias e a toxicologia por inalação é também um importante marco neste aspecto. Suspeita-se, ainda que não se acredite, que os produtos diesel possam ter efeitos carcinogênicos comparativamente graves, mas em muitos lugares não temos nenhuma escolha melhor. O uso de álcool com gasolina como combustível, substituindo os produtos baseados exclusivamente em petróleo, também pode ter sérios efeitos sobre a saúde.

Em geral, a avaliação do risco para a saúde humana resultante destes fatores ambientais tem sido feita em estudos experimentais (*estudos com animais*) ou estudos de observação (isto é, em *evidências epidemiológicas*). De qualquer forma, há certas limitações e iremos discuti-las posteriormente. Há muitos fatores ecoambientais que são relevantes para a nossa existência e bem-estar. Por isso, sob uma perspectiva mais ampla, precisamos abordar todos estes fatores com a devida importância, procurando controlá-los tanto quanto possível, determinar de forma objetiva seu impacto sobre o nosso estilo de vida, avaliar os riscos de seus efeitos sobre a saúde e, desta forma, melhorar o quadro de QDV/QVA. Claramente, estamos no terreno dos modelos com *fatores e respostas múltiplos*, com efeitos latentes e também com algumas respostas não observadas, e assim, devido a estas complexidades, estes modelos são altamente não convencionais. Vamos discuti-los mais na próxima seção.

## **6. QUESTÕES E PERSPECTIVAS ESTATÍSTICAS**

---

Foi observado anteriormente que, no presente estado de desenvolvimento, as medidas de QDV enfocam principalmente os usos alternativos de recursos e de seus benefícios, principalmente no setor de cuidados e serviços (gestão) de saúde. Com base em nossas discussões nas seções anteriores, está claro que uma medida global de QDV não deve estar limitada a uma única fonte de doença ou distúrbio por mais importante que possa ser do ponto de vista de custo/benefício. Estas doenças/distúrbios não devem ser tratadas isoladamente com a prescrição de soluções caso a caso. Contudo, todos estes casos específicos

constituem importantes ingredientes para uma avaliação global e por isso devem ser pesados de maneira coerente para se poder traçar um quadro global, abrangente, com a devida atenção às determinantes bioambientais.

Assim, uma avaliação de QDV/QVA deve abranger o espectro de todas as principais incapacidades, distúrbios e doenças e, com a devida ênfase em seu impacto sobre toda a comunidade/sociedade, deve relacionar as soluções globais específicas e mais abrangentes. Desta forma, uma avaliação de QDV abrangente pode ser altamente sensível a fatores socioeconômicos, culturais e religiosos, juntamente com as causas atribuíveis para as quais estas medidas são geralmente buscadas. Assim sendo, há necessidade de uma salvaguarda adequada para incluir todos os fatores possíveis, as respostas e a qualidade estatística dos conjuntos de dados adquiridos, baseados nos quais devem ser tiradas as conclusões estatísticas.

A QDV é realmente um importante componente das perspectivas socioeconômicas nesta tarefa imensa, especialmente em países do Terceiro Mundo, onde pobreza, elevadas taxas de analfabetismo e outros obstáculos sociais podem limitar seriamente o escopo da implantação e da avaliação adequada de medidas de QDV na saúde e em outros setores. Já notamos que a QDV pode ser interpretada do ponto de vista sociocultural no sentido de que a atitude em relação à vida pode, por si só, ser o fator mais significativo, e o nível de educação, fatores socioambientais e crenças religiosas podem ser todos importantes fatores indicadores neste contexto. Como tal, estas perspectivas podem variar consideravelmente de nações industrializadas para os países do Terceiro Mundo, bem como dentro de cada setor. Como tal, qualquer avaliação de QDV deveria dar a adequada atenção a todos estas condições socioeconômico-culturais.

A principal diferença entre as perspectivas do indivíduo e as da comunidade nas avaliações de QDV é que a sociedade/comunidade deve adotar um plano que possa ser administrado para a maioria de seus membros e não apenas para uma fração que pode pagar as elevadas despesas. Por isso, o quadro de QDV precisa ser avaliado de uma perspectiva mais ampla, ou seja, da comunidade. Deve ser observado aqui que, como a comunidade pode ser suscetível a uma bateria de doenças, distúrbios e diversas incapacidades com intrincados impactos sobre as suas diversidades de idade/sexo e socioeconômicas, o quadro real precisa ser traçado em maior profundidade, com os perfis populacionais dominando o cenário. Assim sendo, mesmo que uma avaliação de QDV esteja confinada apenas ao setor de saúde, os impactos relativos e interativos de todos estes fatores precisam ser adequadamente avaliados de forma estatisticamente válida.

Ainda que uma avaliação de QDV como essa possa ser feita mais convenientemente para cada tipo de distúrbio/incapacidade/doença, ao *reunirmos* todos eles para uma medida global de QDV, há várias questões estatísticas que merecem um exame cuidadoso. As ferramentas estatísticas convencionais nem sempre poderão ser usadas neste contexto. Há um outro fator importante que precisamos manter em mente: em geral, os quadros de QDV

precisam ser comparados para diferentes setores de uma população, seja com base na afinidade regional ou disparidade econômica, bem como em uma base cronológica para projetar o quadro para um futuro imediato ou mesmo mais remoto. Isto pode, por isso, sugerir a relevância dos modelos *espaciais* e de *séries de tempo*. Assim, se quisermos comparar o quadro de QDV global para duas áreas geográficas diferentes ou para dois períodos diferentes de tempo, possivelmente com pirâmides idade-sexo e/ou mapas de incidência doença/distúrbio diferentes, uma abordagem convencional envolvendo uma combinação de *números-índices* pode não ser muito apropriada: a atribuição dos pesos relativos pode, em geral, ser uma tarefa delicada.

Como passo intermediário, uma medida composta de QDV para uma comunidade/sociedade também pode ser formulada para cada um dos amplos setores delineados na Figura 1. Mas então eles precisam ser reunidos de forma adequada para se chegar a uma medida global genuína. No caso anterior, todos os outros fatores não incluídos no setor específico mas com um bom impacto sobre ele serão incorporados como variáveis auxiliares de modo a extrair mais informações estatísticas a partir do conjunto de dados adquirido. Neste último caso, entretanto, um modelo estatístico e um esquema de análise mais complexos, *multifatoriais*, poderão ser necessários para se adequar a esta finalidade. Em futuras publicações, pretendemos abordar as técnicas estatísticas gerais que surgem nestas avaliações de QDV. No restante da seção, apresentamos uma formulação estatística esquemática, evitando em grande parte estas complexidades.

Tipicamente, um *domínio*  $D$  de setores  $\{D_i\}$  para o qual as avaliações de QDV devem ser feitas, é expresso como

$$D = \bigcup_{i \in I} D_i \quad (6.1)$$

onde  $D_i$  refere-se ao  $i$ -ésimo setor, e  $I = \{i\}$  é um *conjunto* (finito) de *índices* para tais setores. Para cada  $i$ , nós indexamos os elementos do conjunto  $D_i$  por um conjunto  $J_i$  e designamos seus membros por  $\{d_{ij} : j \in J_i\}$  para  $i \in I$ . Lembre que, mesmo dentro de um dado setor  $D_i$ , dada a suscetibilidade a múltiplas causas ou fatores, os elementos  $d_{ij}$  podem representar uma combinação de mais do que um evento elementar (doenças/distúrbios, etc.). Em uma avaliação abrangente de QDV, podemos, portanto, começar com as medidas primárias de QDV para os subsetores individuais  $d_{ij}$ , isto é,

$$\pi_{ij}, \quad j \in J_i, \quad i \in I \quad (6.2)$$

e precisamos *reuni-los* de maneira conveniente. A tarefa estatística básica pode ser executada da seguinte maneira:

*Passo 1.* Como formular  $\pi_{ij}$  mantendo em mente os fatores estatísticos anteriormente mencionados além dos fatores operacionais? Neste contexto, precisamos dar a

devida ênfase aos outros subsetores  $d_{ij}$  dentro do mesmo  $D_i$  e também pertencendo a outros setores, ainda que com graus variáveis de importância.

*Passo 2.* Como incorporar as medidas  $\pi_{ij}$ ,  $j \in J_i$ , para formular uma medida abrangente, digamos  $\Pi_i$ , para o setor  $D_i$ ? Mais uma vez, a informação contida em outros setores não pode ser esquecida nesta formulação.

*Passo 3.* Quais são os prós e contras para a reunião destas medidas abrangentes  $\Pi_i$ ,  $i \in I$ , para formular uma medida de QDV global (digamos  $\Pi^\circ$ ) para todo o domínio  $D$ ?

A formulação de uma medida de QDV, mesmo para um subsetor individual  $d_{ij}$ , requer a incorporação de diversos fatores independentes e variáveis resposta e é geralmente feita com base em *modelos paramétricos* adequados, relacionados com suposições de distribuições específicas sobre a variável resposta e seu modelo de dependência com relação aos fatores independentes. Estas formas paramétricas das distribuições da variável resposta ou mesmo as relações entre as variáveis independentes-dependentes são justificáveis com base na simplicidade da análise estatística e nas suas propriedades de eficiência, mas exigem a suposição básica de que as formas paramétricas da suposição estejam muito próximas das verdadeiras. Ou, em outras palavras, se houver algum afastamento em relação às suposições nas quais o modelo é baseado, estes procedimentos paramétricos podem, em geral, não ser muito *robustos*.

Visando à maior robustez, também podem ser recomendados os modelos e esquemas de análise estatísticos alternativos, baseados em metodologia *não-paramétrica* ou *semi-paramétrica*. Em geral, pode ser difícil defender um modelo paramétrico específico, especialmente quando algumas das variáveis resposta ou auxiliares podem ser binárias ou politômicas. Modelos de *regressão logística* e outros *modelos lineares generalizados (MLG)* podem ser usados como alternativas a alguns dos modelos paramétricos convencionais. Em termos de robustez, os modelos não-paramétricos genuínos são mais apropriados do que alguns destes semiparamétricos, especialmente quando grandes conjuntos de dados são obtidos para neutralizar a rigidez das suposições dos modelos paramétricos e para justificar a adoção dos não paramétricos.

Ainda que a QDV tenha um sabor predominantemente qualitativo, as medidas estatísticas estão mais direcionadas para uma quantificação em termos de alguns fatores interpretáveis. Neste sentido, a situação pode ser comparada ao padrão de avaliação da vida, discutido na Seção 3, em que os índices de pobreza e de riqueza vão além da proporção de pessoas pobres e ricas em toda a população, ainda que haja algumas diferenças sutis que serão discutidas a seguir. Motivados por estas considerações, podemos prosseguir.

Como foi discutido anteriormente, dados os subsetores  $d_{ij}$ ,  $j \in J_i$ ,  $i \in I$ , seja  $\alpha_{ij}$  a proporção da população (de uma comunidade/sociedade/país) pertencente ao setor  $d_{ij}$ , de modo que para cada  $i \in I$ ,

$$\alpha_i = \sum_{j \in J_i} \alpha_{ij} \quad (6.3)$$

é a proporção da população pertencente ao setor  $D_i$  para  $i \in I$ . Tipicamente, para cada  $d_{ij}$ , uma avaliação de QDV envolve numerosas variáveis resposta juntamente com algumas variáveis auxiliares e concomitantes, das quais nem todas precisam ser de natureza contínua, podendo ser até mesmo qualitativas.

Com referência às escalas Karnofsky e ARA mencionadas anteriormente, podemos argumentar a favor de um conjunto de escores baseados sobretudo em alguma escala de  $k$  ( $\geq 2$ ) pontos. Para estes dados categorizados ordenados (incluindo os dados binários correlacionados como caso especial), há várias medidas estatísticas que podem ser adaptadas para se obter escores quantitativos apropriados. Neste contexto, portanto, é preciso ter cuidado para interpretar tais escores adequadamente (estatisticamente) em cada caso. Esta é uma das áreas em que há necessidade de se desenvolver uma metodologia estatística mais efetiva. Para uma resposta simples binária ou dicotômica, pode ser adotado um modelo *logito* ou *probit* para quantificar a variável resposta de maneira adequada e, para uma variável resposta politômica, modelos similares (mas, mais complexos) podem ser considerados. Contudo, para uma resposta múltipla, de natureza binária ou politômica, estamos diante de uma situação muito mais complicada. Em geral, há um número excessivo de parâmetros de associação de segunda ordem ou maior, e um modelo completo deste tipo poderá por isso necessitar de um número enorme de observações para assegurar um controle razoável sobre os erros estatísticos. Todavia, alguns desenvolvimentos recentes na metodologia de análise de sobrevivência multivariada relacionam-se a modelos potencialmente utilizáveis em nosso contexto, que serão discutidos em uma futura comunicação. Para respostas quantitativas, o problema de associar escores de QDV é comparativamente mais simples. Em qualquer um dos casos, procedemos como descrito a seguir.

Vamos considerar as variáveis resposta  $r_{ij}$  ( $\geq 1$ ) para o subsetor  $d_{ij}$ . Se as distribuições destas variáveis resposta forem aproximadamente contínuas, então poderemos simplesmente transformar os valores obtidos em escores; assim, quanto mais elevado for o valor obtido, maior será o escore. Se, por outro lado, uma variável resposta é binária ou politômica quanto à sua natureza, temos uma variável resposta categorizada ordenada. Aqui também podemos quantificar os escores em uma ordem crescente. Esta prática é bastante comum em *psicometria* e *teoria dos testes mentais*.

Suponhamos, por exemplo, que uma variável resposta possa assumir um dentre  $k$  resultados categorizados ordenados, indexados como  $\{1, \dots, k\}$ , para algum  $k \geq 2$ . Vamos supor que haja uma variável subjacente ( $Y$ ) com uma distribuição contínua ( $F$ ), de tal forma que haja  $k$  intervalos disjuntos  $I_1, \dots, I_k$ , tal que  $Y \in I_j$  implica que a resposta

categorizada observada é  $j$  ( $j = 1, \dots, k$ ). Escrevemos  $I_j = (a_{j-1}, a_j]$ ,  $1 \leq j \leq k$ , onde  $a_0 = -\infty < a_1 < \dots < a_k = +\infty$ . Além disso, seja  $P_j = P\{Y \in I_j\} = F(a_j) - F(a_{j-1})$ ,  $1 \leq j \leq k$ , de forma que  $P_1 + \dots + P_k = 1$ . Os escores serão então definidos como

$$\alpha_j = \left\{ \int_{a_{j-1}}^{a_j} y dF(y) \right\} / P_j, \quad j = 1, \dots, k. \quad (6.4)$$

Lembre que  $EY = \sum_{j=1}^k \alpha_j P_j = \mu_y$ . Em particular, se  $F$  for normal com média  $\mu_y$  e desvio-padrão

$\sigma_y$ , podemos escrever

$$\alpha_j = \mu_y + \sigma_y \{ \Phi[(a_{j-1} - \mu_y)/\sigma_y] - \Phi[(a_j - \mu_y)/\sigma_y] \} / P_j, \quad 1 \leq j \leq k, \quad (6.5)$$

onde  $\Phi(\cdot)$  representa a função densidade de uma distribuição normal padrão. Em psicometria e estatística educacional, estes valores são conhecidos como escores *normais* (ou Z-escores). Os escores *logito* podem ser definidos de forma similar considerando  $F$  como tendo uma distribuição logística. Desta maneira, para cada uma das variáveis resposta  $r_{ij}$  do subsetor  $d_{ij}$ , os escores quantitativos são obtidos de maneira conveniente, e são incorporados na quantificação que precisamos para formular as medidas de QDV. No caso categorizado ordenado, os  $a_j$ 's podem não ser conhecidos. Mas, usando a notação

$$F_j = \sum_{i \leq j} P_i, \quad 1 \leq j \leq k; \quad P_0 = 0, F_0 = 0,$$

então para os escores normais (ou Z-escores), temos  $a_j = \mu_y + \sigma_y F^{-1}(F_j)$ , e

$$\alpha_j = \mu_y + \sigma_y \{ \Phi(F^{-1}(F_{j-1})) - \Phi(F^{-1}(F_j)) \} / P_j, \quad 1 \leq j \leq k, \quad (6.6)$$

onde, em geral,  $\mu_y$  é tomado como 50 e  $\sigma_y = 10$  (de modo a obter uma amplitude (0,100)). Desta maneira, evitamos os cálculos de  $a_j$ . Esta última fórmula simplifica os cálculos a partir dos dados da amostra. Os valores médios para os escores serão denotados por  $\mu_{ij,s}$ ,  $s = 1, \dots, r_{ij}$ . Lembre que  $d_{ij}$  relaciona-se a um subsetor com uma combinação específica de doenças/distúrbios/incapacidades.

Pareando a constituição demográfica do subsetor  $d_{ij}$ , (isto é, idade/sexo/raça, etc.) consideramos um controle *normal* (não sujeito a esses problemas) para o qual os escores podem ser obtidos de uma maneira similar, e simbolizamos os escores médios correspondentes por  $\mu_{ij,s}^0$ ,  $s = 1, \dots, r_{ij}$ . Então, tipicamente, em um estudo de QDV/QVA as pessoas no subsetor  $d_{ij}$ , sendo afligidas por alguma doença/distúrbio/incapacidade, têm uma qualidade de vida menor, e assim a distribuição associada aos seus escores deveria estar

concentrada mais próxima ao limite inferior do que a distribuição associada aos escores dos controles normais. Assim, se  $F_{ij,s}(\cdot)$  e  $F_{ij,s}^o(\cdot)$  representarem, respectivamente, a distribuição de escores no subsetor  $d_{ij}$  e seu controle normal, teríamos

$$F_{ij,s}(x) > F_{ij,s}^o(x), \quad x \geq 0, \quad s = 1, \dots, r_{ij}. \quad (6.7)$$

Se concebermos um modelo subjacente, como já o fizemos, então uma maneira de caracterizar esta ordenação estocástica é relacionar as duas distribuições por um deslocamento (em uma escala logarítmica), de modo que possamos considerar as restrições *a priori*

$$\mu_{ij,s} < \mu_{ij,s}^o, \quad \forall s=1, \dots, r_{ij}, \quad (6.8)$$

embora não sejamos capazes de demarcar uma linha análoga à linha de pobreza na Seção 3, de tal forma que a distribuição  $F_{ij,s}$  tenha seu limite superior abaixo desta linha e que  $F_{ij,s}^o$  tenha seu limite inferior acima dela. As restrições em (6.8) tornam viável considerar a idéia da razão de defasagem de renda da Seção 3, e incorporá-la na definição de razão de defasagem de qualidade de vida (GRQ) no contexto de QDV, fazendo

$$\beta_{ij,s} = 1 - \mu_{ij,s} / \mu_{ij,s}^o, \quad s=1, \dots, r_{ij}, \quad j \in J_i, \quad i \in I. \quad (6.9)$$

Entretanto, devemos estar atentos ao fato de que a razão de defasagem de renda mencionada na Seção 3 é por definição uma fração não negativa, mas que, no contexto de QDV, o valor de GRQ nem sempre é positivo. Retomaremos este ponto mais tarde. Portanto, uma analogia completa de GRQ com a razão de defasagem de renda da Seção 3 não é geralmente defensável no contexto de QDV. Há necessidade de novos trabalhos neste contexto para melhorar a interpretabilidade de GRQ nos estudos de QDV.

Há um outro problema estatístico encontrado na combinação destas medidas  $r_{ij}$  de GRQ para um dado  $d_{ij}$  em uma única medida de GRQ para o mesmo setor. Com esta finalidade, para cada  $d_{ij}$ , definimos uma razão de defasagem composta de qualidade (GRCQ)  $\beta_{ij}^o$  como a média ponderada dos  $\beta_{ij,s}$  individuais,  $s = 1, \dots, r_{ij}$ , onde os pesos devem ser criteriosamente determinados pela importância relativa das escalas ( $r_{ij}$ ) pertinentes a esta condição. Estes pesos são naturalmente baseados em considerações médicas, custo/benefício e outras, e assim podem variar consideravelmente de um subsetor a outro. Assim, estabelecemos

$$\beta_{ij}^o = \sum_{s=1}^{r_{ij}} w_{ij,s} \beta_{ij,s} \quad \text{com} \quad \sum_{s=1}^{r_{ij}} w_{ij,s} = 1, \quad \forall (i,j). \quad (6.10)$$

Como em (3.5), podemos estabelecer uma medida de QDV ajustada para GRQ para o subsetor  $d_{ij}$  como

$$\pi_{ij}^{(1)} = \alpha_{ij} \beta_{ij}^o, \quad \text{para } j \in J_i, i \in I. \quad (6.11)$$

Neste estágio, podem ser apontados tanto a analogia como o contraste com os índices de pobreza da Seção 3. Há um conceito de renda total (real) ou de riqueza de uma comunidade/sociedade/país, que nos permite interpretar a razão de defasagem de renda de uma maneira adequada, ainda que haja alguma arbitrariedade na própria interpretação da renda ou riqueza reais. No contexto de saúde ou de perspectivas bioambientais, algumas vezes pode ser possível traçar um paralelo com a QDV total, mas muitas vezes isto é mais arbitrário. Por outro lado, as restrições em (6.8) fornecem uma justificativa do GRQ em (6.9) e, portanto, pode ser feita uma analogia entre (6.11) e (3.5).

Para tornar este ponto claro, suponhamos que vamos considerar a pobreza monetária e seu dual, a pobreza mental. No caso monetário, o conceito de razão de defasagem de renda é adotável em uma extensão muito maior. Por outro lado, no caso de pobreza mental, como foi explicado anteriormente, a frugalidade e a atitude geral em relação à vida podem tornar o quadro bastante diferente. Ainda assim, pode ser possível introduzir variáveis resposta adequadas para estes fatores qualitativos, e, em vista disso, podem ser consideradas medidas convenientes de GRQ. Além disso, como a contrapartida normal deve ser definida com pareamento demográfico adequado, as restrições *a priori* em (6.8) ainda podem ser justificadas. Mas isto, então, dependeria bastante da definição destas contrapartidas normais ao subsetor  $d_{ij}$ . Mais uma vez, várias considerações presentes no particular estudo de QDV precisam ser levadas em conta nesta escolha. Também é possível introduzir medidas de QDV mais refinadas, análogas a (3.6) e (3.7). Mas, pelas razões explicadas acima, e por motivos de simplicidade de apresentação, não vamos entrar nestas complicações.

A seguir, com base nas medidas em (6.11), podemos definir

$$\pi_i^{(1)} = \sum_{j \in J_i} \pi_{ij}^{(1)}, \quad i \in I, \quad (6.12)$$

e a medida QDV global para toda a população como

$$\pi^{(1)} = \sum_{i \in I} \pi_i^{(1)}. \quad (6.13)$$

Comentamos sobre a racionalidade da não negatividade de  $\beta_{ij,s}$  em (6.9). Gostaríamos de elaborar isto um pouco mais. Em primeiro lugar, é possível que em alguns casos (6.8) não seja verdadeira. Por exemplo, para os índices de riqueza definimos razão de defasagem de renda em um sentido harmônico. Um argumento similar pode ser apresentado para alguns outros casos em que (6.8) não seja violado. Por outro lado, olhando do ponto de vista de custo/benefício, se para alguns  $(i, j, s)$ ,  $\mu_{ij,s} > \mu_{ij,s}^0$ , a qualidade de vida naquela escala é melhor do que no caso normal, e por isso não há GRQ. Por este motivo, podemos definir GRQ de uma forma mais geral:

$$\beta_{ij,s} = (1 - \mu_{ij,s} / \mu_{ij,s}^0)^+, \quad \forall s, j, i, \quad (6.14)$$

onde  $a^+ = \max \{a, 0\}$ . Em segundo lugar, se um  $\beta_{ij}$  particular for nulo ou quase nulo, isto implica que, com relação àquele item-escore, não há deterioração perceptível de QDV. Uma interpretação similar se mantém para  $\beta_{ij}^0$  que será nulo (ou muito pequeno) quando todos os componentes  $\beta_{ij,s}$  o forem. Julgando sob este ângulo,  $\pi_{ij}^{(1)}$  em (6.11) é realmente uma medida da *deficiência de QDV* (DQDV) devido ao conjunto de distúrbios, etc., identificado no subsetor  $d_{ij}$  para  $j \in J_i$ ,  $i \in I$ . Neste aspecto, a similaridade entre os índices de pobreza estudados na Seção 3 e (6.11) pode ser bastante clara, ainda que (6.11) tenha uma forma mais complexa devido à *reunião* das medidas constituintes que não são necessariamente estatisticamente independentes. Neste sentido, temos a seguinte interpretação natural para DQDV:

$$\begin{aligned} \text{DQDV} = \text{Ajuste médio de QDV devido à condição} \\ d_{ij}, \text{ para } j \in J_i, i \in I. \end{aligned} \quad (6.15)$$

Em terceiro lugar, esta interpretação de DQDV em (6.15) estende-se diretamente para (6.12) e (6.13). Há, entretanto, uma verdadeira necessidade de mais trabalho estatístico a este respeito. A preocupação básica é: faz sentido combinar os termos  $\beta_{ij,s}$  em uma medida global ( $\beta_{ij}^0$ )? Isto também traz a questão: assegurada esta possibilidade, qual é a maneira conveniente de implementá-la? As respostas a estas duas perguntas podem ser encaradas como afirmativas, ainda que a implementação dependa da escolha hábil das escalas relativas a cada condição  $d_{ij}$ , de modo que transmitam informações combináveis. Além do mais, como explicado anteriormente, a construção de medidas estatísticas adequadas pode muitas vezes exigir esquemas mais sofisticados de modelagem e análise estatísticas. Finalmente, as combinações em (6.12) e (6.13) pertencem ao domínio da chamada *metaanálise*. Apresentamos a seguir certas considerações básicas a este respeito.

É bem possível que as medidas de QDV em (6.11) possam ser consideravelmente diferentes de um subsetor para outro dentro de um setor comum  $D_i$  e também (mais perceptível) de um  $D_i$  para outro. Ainda assim, estão inter-relacionadas em um sentido estatístico bem definido. Por isso, a filosofia básica da metaanálise se estende também a este complexo campo de aplicações. Por outro lado, para diferentes  $d_{ij}$  ou  $D_i$ , geralmente as escalas de medida, baseadas em questionários possivelmente diferentes, podem ser diferentes, e nenhum questionário será igualmente apropriado para todas estas condições. Por isso, sob o ponto de vista estatístico, é desejável desenvolver questionários padronizados que permaneceriam úteis e interpretáveis para todo um setor  $D_i$ , e poderiam facilmente ser estendidos também para outros setores. Isto nos permitiria incorporar ferramentas de meta-análise estatística em uma extensão satisfatória.

Esta tarefa é desafiadora e exige, por sua vez, uma coordenação completa da visão estatística com a de outros cientistas e pesquisadores de outras disciplinas engajadas na busca de avaliações de QDV. Neste aspecto, a demarcação dos subsectores  $d_{ij}$ , bem como dos setores  $D_i$ , com vistas a definir exatamente o domínio global  $D$  pertencente a uma tarefa de avaliação de QDV, precisa ser realizada de forma criteriosa. A avaliação de QDV para os (sub)sectores individuais pode necessitar de questionários um pouco diferentes e por isso requerer técnicas estatísticas diferentes e, assim, ao reunir a informação destes (sub)sectores individuais, pode haver necessidade de uma tarefa estatística muito abrangente na incorporação da metaanálise para uma avaliação global. Como explicado anteriormente, uma tarefa estatística ainda mais importante é a atribuição de pesos relativos (importância) para os diferentes (sub)sectores de modo a justificar (6.12) e (6.13) de um ponto de vista estatístico consistente. É claro que há necessidade de mais pesquisas estatísticas nesta direção.

A variação temporal da QDV é de grande interesse estatístico, tendo também um forte impacto sobre ações administrativas relacionadas à QDV. Por exemplo, a qualidade de vida de pessoas infectadas pelo vírus da AIDS (HIV) passou por mudanças regulares nos últimos cinco anos, e no momento existem algumas indicações que, num futuro próximo, a crise poderá ser tratada de uma maneira muito melhor do que o foi na última década, época em que se proliferou de uma forma monstruosa. Neste contexto, os  $\alpha_{ij}$ ,  $\beta_{ij}$ , etc., definidos de (6.3) a (6.6), são dependentes no tempo, e então os  $\pi_{ij}^{(1)}$ , definidos em (6.11), são geralmente expressos como  $\pi_{ij}^{(1)}(t)$ ,  $j \in J_i$ ,  $i \in I$ , e  $t \in T$ , onde  $T$  refere-se ao conjunto de índices associados aos períodos de tempo. Modificação similar pode ser feita nas equações (6.12) e (6.13). Nesse caso, teremos um modelo multivariado de séries temporais. Embora possam ser incorporados aqui modelos convencionais de séries de tempo, em vista do grande número de subsectores  $\{d_{ij}\}$ , podemos ter um modelo com muitas dimensões, sendo que o modelo convencional pode perder a sua aplicabilidade. Alternativamente, modelos de Cadeias de Markov (com parâmetro de tempo discreto/contínuo) podem ser incorporados para reduzir efetivamente a dimensão do espaço paramétrico, aumentando, portanto, a precisão

estatística. Ainda a este respeito, a estacionariedade (com relação ao tempo) pode não se verificar, e daí trabalhar com a não-estacionariedade em modelos com alta dimensionalidade pode trazer complicações tanto na modelagem quanto na análise estatística. Este assunto será exposto futuramente de modo mais sistemático.

Uma vez que estas medidas de QDV para os (sub)setores individuais ou para o domínio todo sejam definidas, devem ser completadas algumas tarefas estatísticas mais estratégicas, relacionadas às medidas de QVA, bem como ao tempo de vida residual médio (TVRM) e conceitos afins. Ainda neste contexto, a definição de QVA pode envolver aspectos mais qualitativos e, desta forma, poderão ser necessárias medidas estatísticas mais sofisticadas para quantificá-los de maneira conveniente, mas ainda interpretável. O quadro TVRM pode variar drasticamente de um (sub)setor a outro. Por exemplo, pessoas que são HIV positivas e que estão sob tratamento para câncer diagnosticado podem ter TVRM bastante diferente. Mesmo dentro do setor de câncer, diferentes tipos de doenças podem ter TVRM diferentes. Assim sendo, nesta avaliação, pode haver necessidade de um maior cuidado na identificação precisa da doença/distúrbio/incapacidade e, à luz desta informação, o estudo de QVA ou TVRM precisa ser formulado com as salvaguardas estatísticas adequadas, consistentes com os objetivos gerais do estudo de QDV.

Felizmente, durante o passado recente, muitas pesquisas estatísticas foram realizadas na área geral de *confiabilidade e análise de sobrevivência* onde TVRM desempenha um importante papel. A diferença nas abordagens de QDV e análise de sobrevivência pode ser mais fundamental em vista da natureza dos estudos observacionais envolvidos, dos objetivos gerais destes estudos, bem como das suposições básicas de regularidade necessárias. Assim sendo, há também uma necessidade verdadeira de desenvolver ferramentas estatísticas mais apropriadas para o atual propósito. Pretendemos apresentá-las em futuras comunicações.

## **7. OBSERVAÇÕES FINAIS**

---

Nesta seção, queremos fazer algumas observações gerais relativas às perspectivas do Terceiro Mundo na avaliação de QDV/QVA. Com base em nossa discussão nas seções anteriores, podemos observar que uma solução satisfatória da tarefa de avaliação de QDV somente pode ser conseguida através de uma abordagem interdisciplinar. Ainda que fatores sociais, culturais, econômicos, de saúde, psicológicos e ecoambientais sejam fundamentais em uma tarefa de avaliação como esta, as considerações estatísticas são indispensáveis. Para poder lidar com este problema, o estabelecimento de uma organização interdisciplinar, particularmente em um país do Terceiro Mundo, para concretizar a tarefa de avaliação de QDV pode ser um grande passo por si só. E, acima de tudo isso, a implantação de conceitos, ferramentas e medidas estatísticas sofisticadas pode ser ainda mais difícil.

As avaliações de QDV são em grande parte baseadas em estudos de observação com dados reunidos de diferentes fontes e muitas vezes pode não haver controle suficiente sobre a

qualidade da informação que pode ser extraída destes conjuntos de dados. Neste contexto, as tarefas de monitoramento e tratamento dos dados são vitais. Mais e mais as agências responsáveis pela coleta de dados relevantes estão se dando conta desta exigência básica, e esperamos que nos países do Terceiro Mundo isto não seja uma exceção.

Do ponto de vista de coordenação interdisciplinar, a divisão dos subsetores  $d_{ij}$  dentro dos setores  $D_i$ , bem como o domínio global  $D$ , pode variar consideravelmente de um país (sociedade) para outro país (sociedade), e isto precisa ser instituído com uma profunda compreensão das perspectivas socioeconômicas, culturais, religiosas, socioambientais, ecológicas, de saúde e psicológicas dominantes. Ainda sob este prisma o quadro pode ser bastante diferente em um país do Terceiro Mundo quando comparado a um país industrializado, e também dentro do setor do Terceiro Mundo, entre os próprios países. Um gerenciamento adequado desta tarefa exige a montagem de várias agências reguladoras que, por sua vez, devem ser interativas na coleta de informações relevantes e em sua combinação. Na maioria das nações industrializadas, estas agências ou instituições reguladoras existem e há bastante interação. Nos países do Terceiro Mundo a situação é bastante diferente. Na maioria dos casos esta infra-estrutura ainda está para ser desenvolvida.

Nenhuma avaliação de QDV pode ser feita sem um planejamento (estatístico) efetivo de um estudo, desenvolvendo um protocolo para a coleta de informações estatísticas relevantes, verificando a qualidade e a confiabilidade dos conjuntos de dados (estatísticos) adquiridos e, finalmente, implantando ferramentas de uma análise estatística confiável, válida e interpretável para a obtenção de conclusões estatísticas válidas, que possam ser explicadas para outros cientistas e pesquisadores, para as suas tarefas de avaliação. Neste aspecto, a interpretação da OMSQDV (relacionada principalmente com o setor saúde), a interpretação da FAO (com ênfase na nutrição e economia rural) e a abordagem da UNESCO (incluindo diversos aspectos culturais, religiosos, educacionais e socioeconômicos) são passos recomendáveis para uma formulação adequada de uma medida global comum, que permitiria a obtenção de um quadro abrangente. Mais uma vez, estas organizações têm uma ampla infra-estrutura que lhes permite abarcar uma variedade de situações em todo o mundo e, ainda assim, há uma necessidade básica de os países do Terceiro Mundo terem suas próprias agências que, com a cooperação capaz e o auxílio destas organizações internacionais, podem realizar planos mestre mais penetrantes para a coleta, monitoramento e assimilação de dados complicados, que possam se adequar melhor aos seus objetivos internos. Em qualquer tentativa de uma avaliação internacional de QDV, os países do Terceiro Mundo não podem ser deixados de fora ou serem tratados inadequadamente. Assim sendo, os países do Terceiro Mundo têm pela frente uma tarefa respeitável para a atualização da estrutura de dados de QDV. Não há dúvidas de que a estatística é a tecnologia chave nesta busca.

A avaliação de QDV/QVA nos diferentes setores pode ter objetivos um pouco diferentes e pode também estar sujeita a interpretações um pouco diferentes. Isto não deveria, contudo, nos desencorajar de buscar uma medida global com interpretação consistente e

flexibilidade suficiente para se adaptar aos vários setores. Para fazer isto de forma efetiva, precisamos, entretanto, de uma compreensão clara dos objetivos básicos das medidas de QDV/QVA relacionadas aos vários setores. Neste empreendimento interdisciplinar é importante, principalmente para os estatísticos, compreender a natureza dos problemas não apenas do ponto de vista estatístico mas também dos pontos de vista operacional, gerencial e de custo/benefício. Esta compreensão poderá ficar um pouco mais clara se levarmos em conta a diversidade do quadro QDV/QVA em uma perspectiva mais ampla, usando então os princípios estatísticos no planejamento, modelagem, coleta de dados e esquemas de análise. Neste aspecto precisamos desenvolver ferramentas estatísticas mais apropriadas e também estender a abordagem estatística. Por isso, com uma nota otimista, eu gostaria de concluir dizendo que, se análises estatísticas estiverem presentes em todas as esferas das tarefas de avaliação QDV/QVA, poderemos ser úteis ao bem estar da humanidade deste planeta ameaçado.

### AGRADECIMENTOS

Este artigo foi adaptado e traduzido do artigo (Sen, 1996) apresentado na palestra de abertura da Conferência QDV em Hurgada, Egito, em dezembro de 1994. Uma coletânea dos trabalhos apresentados apareceu no *Journal of Applied Statistical Science*, vol. 4, nº 4, 1996. O autor agradece a todos os participantes da Conferência QDV por seus comentários e críticas úteis. Os comentários detalhados e notas de correção do Professor Sati Mazumdar ajudaram o autor na reformatação deste artigo em uma forma mais apresentável. Em uma *Discussão* anexa, o Professor Shoutir Chatterjee gentilmente apontou algumas questões básicas com a abordagem atual, que é seguida por uma breve resposta do autor.

### REFERÊNCIAS

- AARONSON, N. C. (1989). Quality of life assessment in clinical trials: methodological issues. *Control. Clinic. Trials* 10, 195S-208S.
- ANDERSON, R. T., AARONSON, N. C., & WILKIN, D. (1993). Critical review of the international assessments of quality of life. *Qual. Life Res.* 2, 369-395.
- CELLA, D. F. (1992). Quality of Life: The concept. *J. Palliative Care* 8, 8-13.
- COX, D. R., FITZPATRICK, R., FLETCHER, A. E., GORE, S. M., SPIEGELHALTER, D. J., & JONES, D. R. (1992). Quality of Life Assessment: Can we keep it simple? *J. Roy. Statist. Soc. A* 155, 353-393.
- ERICKSON, P., KENDALL, E. A., ANDERSON, J. P., & KAPLAN, R. M. (1992). Using composite health state measures to assess the nation's health. *Medical Care* 30, MS166-175.
- FAIRCLOUGH, D., & GELBER, R. D. (1995). Quality of Life: Statistical Issues and Analysis. In *Quality of Life and Pharmacoeconomics in Clinical Trials*, 2nd ed. (ed. B. Spilker), Raven, N.Y.
- FALLOWFIELD, L. (1990). *The Quality of Life: The Missing Measurement in Health Care*. Souvener Press, London.
- FITZPATRICK, R., FLETCHER, A., GORE, S., JONES, D., SPIEGELHALTER, D., & COX, D. (1992). Quality of Life measures in health care, I: Applications and issues in assessment. *British Med. J.* 305, 1047-1077; II: Design, analysis and interpretation. *Ibid.* 305, 1145-1148.

- GASTWIRTH, J. L. (1975). A new index of income inequality. *Proc. Inter. Statist. Inst.*, Vienna, 1, 368-372.
- GELBER, R. D., GELMAN, R. S., & GOLDBIRSH, A. (1989). A quality of life oriented endpoint for comparing therapies. *Biometrics* 45, 781-795.
- GLASZIOU, P. P., SIMES, R. J., & GELBER, R. D. (1990). Quality adjusted survival analysis. *Statist. Med.* 9, 1259-1276.
- GORE, S. (1988). Integrated reporting of quality and length of life - a statistician's perspective. *Euro. Heart J.* 9, 228-234.
- GOTAY, C. C., KORN, E. L., MCCABE, M. S., MOORE, T. D., & CHESON, B. D. (1992). Building Quality of Life assessment into cancer treatment studies. *Oncology* 6, 25-28.
- KARNOFSKY, D., & BURCHENAL, J. (1949). The clinical evaluation of chemotherapeutic agents in cancer. In *Evaluation of Chemotherapeutic Agents* (ed. C. McLeod). Columbia Univ. Press. New York, pp. 191-205.
- KORN, E. L. (1993). On estimating the distribution function for quality of life in cancer clinical trials. *Biometrika* 80, 535-542.
- LANE, D. A. (1987). Utility, decision and quality of life. *J. Chronic Disease* 40, 585-591.
- LOEWY, J. W., KAPADIA, A. S., HSI, B., & DAVIS, B. R. (1992). Statistical Methods that distinguish between attributes of assessment: Prolongation of life versus quality of life. *Med. Dec. Making* 12, 83-92.
- MCDOWELL, I., & NEWELL, C. (1987). *Measuring Health: A Guide to Rating Scales and Questionnaires*. Oxford Univ. Press, N. Y.
- MEHREZ, A., & GAFRI, A. (1989). Quality adjusted life years, utility theory and health years equivalents. *Med. Dec. Making* 9, 142-149.
- MIETTINEN, O. S. (1987). Quality of life from the epidemiological perspective. *J. Chronic Diseases* 40, 641-643.
- MOSTELLER, F. (1987). Implications of measures of quality of life for policy development. *J. Chronic Diseases* 40.
- OLWENY, C. L. (1992). Quality of life in developing countries. *J. Palliative Care* 8, 25-30.
- ORLEY, J., & KUYKEN, W. (1994). *International Quality of Life Assessment in Health Care Settings*. Springer-Verlag, New York.
- PEARLMAN, R. A., & OHLMANN, R. F. (1988). Quality of life in chronic diseases: Perceptions of elderly patients. *J. Gerontology* 43, M25-30.
- SCHUMAKER, M., OLSCHESKI, M., & SCHULGEN, G. (1991). Assessment of quality of life in clinical trials. *Statist. Med.* 10, 1915-1930.
- SEN, A. K. (1976). The measurement of poverty: an axiomatic approach. *Econometrica* 44, 219-232.
- SEN, P. K. (1986). The Gini Coefficient and the poverty indexes: some reconciliation. *J. Amer. Statist. Assoc.* 81, 1050-1057.
- SEN, P. K. (1988). The harmonic Gini coefficient and affluence indexes. *Math. Soc. Sc.* 8, 65-76.
- SEN, P. K. (1996). Statistics and Quality of Life: Third world perspectives. *J. Appl. Statist. Sc.* 4, 231-252.
- SEN, P.K. (1998). Utility-oriented Simpson-type indexes and inequality measures (to appear).
- SIMPSON, E. H. (1949). The measurement of diversity. *Nature*, 163, p.688.
- SPIPKER, B. (1990). *Quality of Life Assessments in Clinical Trials*. Raven. N. Y.
- STEINBROCKER, O., TRAEGER, C., & BATTMAN, R. (1949). Therapeutic criteria in rheumatoid arthritis. *J. Amer. Med. Assoc.* 140, 659-662.
- TANDON, P. K. (1990). Application of global statistics in analyzing quality of life data. *Statist. Med.* 8, 819-827.
- WALKER, S.R., & ROSSER, R.M. (1993). *Quality of Life Assessment: Key Issues in the 1990's*. Kluwer, UK.
- WORLD HEALTH ORGANIZATION (1993). *WHOQOL Study Protocol*. WHO, Geneva.
- WHOQOL GROUP (1993). Study protocol for the World Health Organization Project to develop a quality of life assessment instrument (WHOQOL). *Qual. Life. Res.* 22, 153-159.
- WHOQOL GROUP (1994). The development of the World Health Organization Quality of Life assessment instrument (the WHOQOL). In *Quality of Life Assessment in Health Care Setting* (eds. J. Orley & W. Kuyken). Springer-Verlag, Heidelberg.

## ABSTRACT

The quality of human life in a community or society, judged from a broader perspective, depends not only on the standard of living in a conventional monetary sense but also on the bio-environment that governs a battery of social, economic, religious, cultural, health and psychological, ecological and environmental factors. These perspectives are appraised in the formulation and interpretation of Quality of Life, and, as in Sen (1996), due emphasis on biostatistical foundations is paid in this context.

## DISCUSSÃO

Shoutir K. Chatterjee, Universidade de Calcutá, Índia

Neste trabalho, o autor estende a idéia da razão de defasagem de renda para a medida da pobreza (monetária) com a finalidade de formular medidas de *deficiência de qualidade de vida* (DQDV), relativamente a várias características de desempenho para cada subclasse de uma população. Consegue, desta forma, uma única média ponderada para DQDV para cada subclasse. Ponderando-as com os tamanhos relativos das subclasses, chega a uma medida abrangente de DQDV dada por (6.13).

Dois pontos parecem relevantes com relação a esta extensão:

1) A analogia de GRQ em (6.9) com a razão de defasagem de renda em (3.3) não é perfeita uma vez que a primeira relaciona  $\mu_{ij,s}$  com  $\mu_{ij,s}^o$  enquanto que o último relaciona  $\mu_\alpha$  com o ponto de limite  $\rho$  entre os *pobres* e os *não-pobres*. Na verdade, a cauda superior da distribuição da característica considerada em relação à classe anormal se sobrepõe à cauda inferior da distribuição para a classe normal no caso (6.9), enquanto que os dois grupos estão separados por  $\rho$ , no caso (3.3). Se a pobreza fosse definida com o sentido subjetivo da *ausência de bem-estar* e os pobres e não-pobres definidos neste sentido, alguém com uma renda monetária mais elevada poderia ser pobre e outro com uma renda monetária mais baixa poderia ser não-pobre. Na busca da medida do bem-estar em termos monetários, poder-se-ia estabelecer uma medida como (6.9). Convencionalmente, entretanto, a pobreza é definida em termos monetários e (3.3) é usada.

2) No caso da distribuição de renda há alguma renda total populacional que é distribuída entre seus membros. É por isso que a medida  $G_\alpha$  de desigualdade e medidas como (3.6) - (3.7) têm relevância. No caso da qualidade de vida, tal como discutida na Seção 6, não há conceito de *qualidade total* distribuída entre os membros. Assim, medidas como (3.6)-(3.7) não são apropriadas aqui e a observação que se segue a (6.11) está um pouco deslocada.

## RESPOSTA DO AUTOR

Os dois pontos levantados pelo Professor Chatterjee são muito importantes. A diferença básica entre a avaliação da pobreza em um sentido monetário e qualidade de vida em um sentido mais amplo é a falta de quantificação adequada de QDV, bem como a escolha

de um *ponto-limite* entre as classes normal e anormal. Se, ao invés de pobreza em um sentido monetário olharmos o problema da avaliação da *pobreza mental*, vamos encontrar um problema bastante similar onde o conceito de riqueza total (mental) pode ser difícil de formular. A escolha de traços relevantes e sua qualificação adequada são as duas tarefas básicas em um empreendimento como esse. Contudo, é inegável que as pessoas variam muito também quanto aos seus traços mentais, muitas vezes drasticamente, e pode não haver muita concordância estatística entre a pobreza mental e a monetária. Contudo, o ponto comum a todas estas avaliações é o reconhecimento do desejo de uma medida quantitativa de deficiência para lidar com as imperfeições de diferentes classes de pessoas que não são vistas como normais (com relação a diferentes características de desempenho). Felizmente, todos os pesquisadores neste campo mais amplo apontaram para a necessidade de quantificação de QDV e as necessidades aqui consideradas estão voltadas para esta resolução.

Considerando que renda total ou riqueza de uma população é subordinada à mensuração de sua desigualdade de renda, de um ponto de vista custo/benefício, os *recursos totais* disponíveis para lidar com os cuidados e serviços de saúde de uma população podem servir como um ponto de referência básico. Assim, pelo menos no setor de saúde, as medidas de QDV/QVA devem ser avaliadas à luz do total destes recursos. Entretanto, as medidas deste total de recursos podem não estar disponíveis para todos os setores e, além disso, a demarcação dos pontos de limite para classificar as classes normais e causas específicas pode não ser muito razoável. Essencialmente, de um ponto de vista estatístico, somos levados a várias distribuições para várias subclasses, em que, *a priori*, temos alguma ordenação. Por este motivo, ao invés de procurar estabelecer pontos limite, as medidas de DQDV aqui consideradas referem-se a situações com ajuste das distribuições para as diversas subclasses. Na análise de custo/benefício estas medidas trazem uma percepção significativa. Na ausência de um conceito de recursos totais numa abordagem geral, é desejável que haja flexibilidade na escolha dos pesos  $w_{ij,s}$  (na Seção 6) requeridos para a definição das medidas. O conceito de DQDV, do modo como foi delineado, tem um sentido prático do ponto de vista custo/benefício e, com relação à escolha dos pesos  $w_{ij,s}$ , uma interação maior entre estatísticos e pesquisadores neste campo amplo deve pavimentar o caminho para uma melhor resolução do problema.

# ANÁLISE DOS ERROS AMOSTRAIS DE INDICADORES EDUCACIONAIS OBTIDOS ATRAVÉS DE DADOS DA PESQUISA NACIONAL POR AMOSTRA DE DOMICÍLIOS<sup>1</sup>

Ruben Klein (LNCC)\*

Fernando A. S. Moura (UFRJ)\*\*

## RESUMO

Fletcher e Ribeiro (1988) desenvolveram indicadores educacionais, utilizando dados de uma Pesquisa Nacional por Amostra de Domicílios - PNAD. Denominaram esta metodologia de Profluxo. O objetivo deste artigo é estimar e analisar a precisão das taxas de participação e de conclusão de uma coorte em uma série. Verifica-se que as estimativas em nível de Brasil têm precisão muito boa. Estuda-se também a precisão das estimativas para um estado com uma amostra "grande" como São Paulo e um com uma amostra "pequena" como o Acre urbano.

## 1. INTRODUÇÃO

---

O PROFLUXO (Projeto Fluxo dos Alunos do Ensino do Primeiro Grau), desenvolvido por Fletcher e Ribeiro (1988, 1989), ver também Klein (1991) e Klein e Ribeiro (1991), foi concebido para estimar os termos do modelo de fluxo dos alunos no Sistema Regular de Ensino através de dados demográficos. Para isso foi necessário desenvolver dois conceitos

---

<sup>1</sup> Este trabalho foi parcialmente financiado pela Fundação Ford no projeto "Indicadores de Fluxo Escolar e Estudos sobre o Mecanismo da Repetência".

\* Endereço para correspondência: Rua Lauro Müller, 455 - RJ - Brasil - e-mail - rube@lncc.br

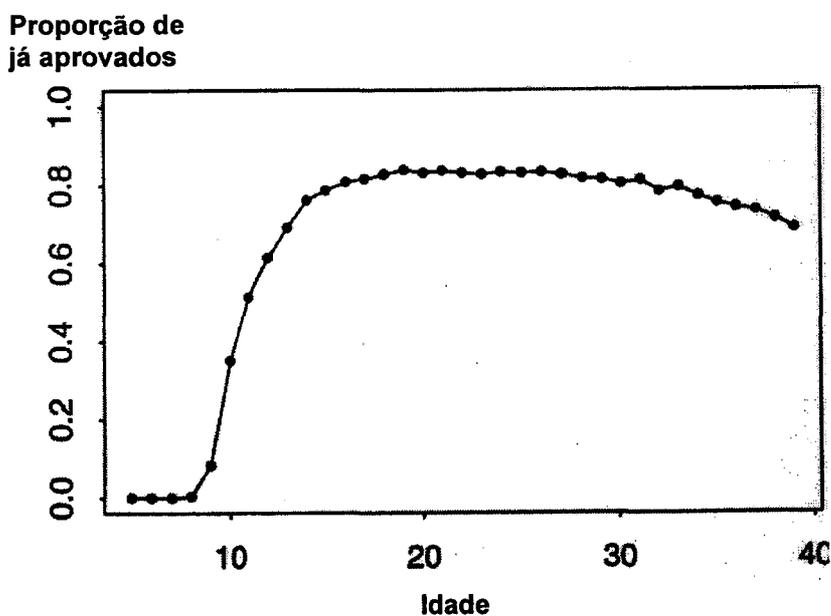
\*\* Departamento de Métodos Estatísticos- RJ- Brasil - e-mail - fmoura@dme.ufrj.br

que têm interesse próprio como indicadores educacionais de acesso à escola. Esses conceitos são: a taxa de participação e a taxa de conclusão em uma série.

Antes de definir esses termos precisamos do conceito de uma coorte de idade  $i$ , definida como sendo o conjunto dos indivíduos com idade  $i$ . A taxa de participação (conclusão) de uma coorte de idade  $i$  na série  $k$  é definida como sendo a proporção de indivíduos da coorte de idade  $i$  que eventualmente tem acesso (concluem) à série  $k$ .

Em um determinado ano, pode-se calcular a proporção de indivíduos da coorte de idade  $i$  que já tiveram acesso (concluíram) à série  $k$ . Essa proporção será igual à taxa desejada para as coortes mais velhas para as quais pode-se supor que ninguém mais terá acesso ou concluirá a série. A Figura 1 mostra a proporção estimada de indivíduos de cada coorte de idade  $i$ , no Brasil, no ano de 1990, que já concluíram a 3ª série do 1º grau. Pode-se ver que a partir de uma certa idade, essa proporção começa a cair, indicando que as taxas de conclusão dessas coortes já foram atingidas e que essas taxas vinham crescendo para as coortes mais novas.

Figura 1  
Proporção de já Aprovados, por Idade, para a 3ª série, no Brasil



A taxa de participação (conclusão) na série  $k$ , em um determinado ano, é definida como sendo o máximo das taxas de participação (conclusão) das coortes de idade na série  $k$  até aquele ano. Na Figura 1, o máximo da curva corresponde à taxa de conclusão da 3ª série do 1º grau no ano de 1990.

Os dados necessários para se estimar essas taxas podem ser obtidos no questionário de mão-de-obra da Pesquisa Nacional por Amostra de Domicílios - PNAD -, realizada anualmente pelo IBGE, através das seguintes perguntas feitas a cada indivíduo:

- i) idade;
- ii) se o indivíduo freqüenta a escola, qual é o grau e a série que freqüenta; e
- iii) se o indivíduo não freqüenta a escola, qual foi a última série e o grau da última série concluída com êxito.

Destas perguntas, obtém-se para cada ano estimativas para as seguintes quantidades:

$T_i$ : tamanho da coorte de idade  $i$ ;

$m_{i,k}$ : número de indivíduos de idade  $i$  matriculados na série  $k$ ; e

$d_{i,k}$ : número de indivíduos de idade  $i$  que não freqüentam a escola, mas para os quais a última série concluída com êxito foi a série  $k$ .

$M_{i,k} = m_{i,k} / T_i$ : proporção da coorte de idade  $i$  matriculada na série  $k$ ; e

$D_{i,k} = d_{i,k} / T_i$ : proporção da coorte de idade  $i$  que não freqüenta a escola, mas para quem a última série concluída com êxito foi a série  $k$ .

Supondo que se um aluno está matriculado ou já cursou uma determinada série, ele já cursou todas as séries anteriores, pode-se escrever as seguintes equações:

$$A_{i,k} = \sum_{j=k}^{12} D_{i,j} + \sum_{j=k+1}^{12} M_{i,j} = \frac{\sum_{j=k}^{12} d_{i,j} + \sum_{j=k+1}^{12} m_{i,j}}{T_i} \quad (1)$$

$$I_{i,k} = A_{i,k} + M_{i,k} + DN_{i,k} = \frac{\sum_{j=k}^{12} d_{i,j} + \sum_{j=k}^{12} m_{i,j}}{T_i} + DN_{i,k} \quad (2)$$

onde:

$j = 1, \dots, 8$  representa cada uma das oito séries do ensino fundamental;  $j = 9, 10, 11$  representa cada uma das três séries do ensino médio e  $j = 12$  representa a primeira série do ensino superior e qualquer série e grau superior ou igual a 12.

$A_{i,k}$ : é a proporção de indivíduos de idade  $i$  que já foram aprovados na série  $k$ ;

$I_{i,k}$ : é a proporção de indivíduos de idade  $i$  que já ingressaram na série  $k$ , tendo concluído a série ou não; e

$DN_{i,k}$ : proporção de indivíduos de idade  $i$  que não estão freqüentando a escola, mas que chegaram a freqüentar a série  $k$ , sem, no entanto, concluí-la.

As estimativas dos  $A_{i,k}$ 's e dos  $I_{i,k}$ 's nos fornecem, respectivamente, as taxas estimadas de conclusão e participação, na série  $k$ , de cada coorte de idade para as coortes

mais velhas. Os máximos nos dão as estimativas para as taxas de conclusão e participação na série  $k$ , respectivamente.

Para os que não freqüentam escola, a PNAD só coleta informação sobre a última série concluída com êxito, logo esta última parcela na equação (2),  $DN_{i,k}$  não pode ser estimada e é desconsiderada. Portanto os  $I_{i,k}$ 's e logo as taxas de participação na série  $k$  de cada coorte de idade são subestimadas. Porém, simulações do fluxo escolar utilizando-se taxas corrigidas de transição entre séries obtidas a partir dos Censos Educacionais (Klein (1995)) indicam que essa subestimação não é grande, talvez da ordem de 3% da coorte. Desta forma, as taxas de conclusão na série  $k$  de cada coorte de idade são, em princípio, bem estimadas.

Cabe ressaltar que as PNADs até 1990, por problemas conceituais e/ou de coleta, não distinguiam os indivíduos que estavam cursando ou que já tinham cursado o Ensino Supletivo Seriado, confundindo-os com os alunos que cursavam ou cursaram o Ensino Regular. Como conseqüência, as taxas de conclusão e participação incluem alunos que concluíram ou tiveram acesso à série através do Ensino Supletivo.

Na próxima seção, será dada sucintamente uma descrição do plano amostral das PNADs e fórmulas aproximadas para estimar o erro amostral das estimativas dos  $A_{i,k}$ 's e dos  $I_{i,k}$ 's. Finalmente, na última seção serão fornecidas estimativas dos  $A_{i,k}$ 's e dos  $I_{i,k}$ 's e de seus erros-padrão para o Brasil e para os Estados de São Paulo e Acre referentes ao ano de 1990.

## **2. ESTIMAÇÃO DOS INDICADORES EDUCACIONAIS ATRAVÉS DA PNAD**

### **2.1. Desenho Amostral da PNAD**

A Pesquisa Nacional por Amostra de Domicílios - PNAD -, de âmbito nacional e periodicidade anual, tem como objetivo obter informações socioeconômicas da população através de uma amostra probabilística de domicílios. O desenho amostral consiste em uma amostra estratificada de conglomerados com três estágios de seleção: municípios, setores e domicílios. Os estratos são compostos por municípios pertencentes, na medida do possível, a uma mesma microrregião geográfica.

Considerando-se a PNAD referente ao ano de 1990, a seleção das unidades primárias de amostragem (municípios), dentro de cada estrato, foi sistemática e com probabilidade proporcional à população total do município, segundo o Censo Demográfico 1980. Para alguns municípios, por razões técnicas, e para todos os municípios pertencentes às regiões metropolitanas, atribuíram-se probabilidades de seleção iguais a um, ou seja, esses municípios foram incluídos com certeza na amostra. Tais municípios foram denominados auto-representativos e cada um deles passou a constituir um estrato isolado.

A seleção das unidades secundárias de amostragem (os setores censitários) para cada unidade primária selecionada (município) foi sistemática e com probabilidade proporcional ao número de domicílios em cada setor, segundo o Censo Demográfico 1980.

Para os setores selecionados, é feita uma operação de listagem. A operação de listagem consiste no arrolamento de todos os domicílios e unidades dos setores selecionados, atualizando-se assim a informação cadastral para permitir a seleção de domicílios.

Finalmente, a seleção dos domicílios em cada setor é feita de forma sistemática e com intervalo de seleção calculado de tal forma que a amostra seja autoponderada (fração de amostragem constante) para cada uma das regiões metropolitanas e não-metropolitanas, ou para o conjunto das Unidades da Federação nos casos em que estas não possuam região metropolitana. Considerando-se a PNAD 1990, foram selecionados 698 municípios, 6 556 setores e investigados 87 540 domicílios.

## 2.2. Estimadores dos Indicadores e Respectivas Variâncias

Com o objetivo de analisar a precisão das estimativas em função do tamanho do nível geográfico pretendido, os indicadores definidos pelas equações (1) e (2) foram estimados em nível de Brasil como também em nível das Unidades da Federação de São Paulo e do Acre, na zona urbana<sup>2</sup>. O processo de estimação desses indicadores e das respectivas variâncias é descrito nesta seção.

Antes de proceder à estimação dos  $A_{i,k}$ 's e dos  $I_{i,k}$ 's, os numeradores e denominadores das equações (1) e (2) são totalizados, respectivamente, por domicílio e então, para cada indicador considerado, representa-se genericamente estes totais pelos respectivos valores de  $y$  e definidos abaixo.

Notação:

$H_{(d)}$ : número de estratos no domínio geográfico de interesse ( $d$ ): estado ou união de estados.

$I_h$ : número de municípios selecionados no  $h$ -ésimo estrato.

$P_{hi}$ : probabilidade de seleção do  $i$ -ésimo município do  $h$ -ésimo estrato.

$m_{hi}$ : número de setores selecionados no  $i$ -ésimo município do  $h$ -ésimo estrato.

$P_{hij}$ : probabilidade de seleção do  $j$ -ésimo setor do  $i$ -ésimo município do  $h$ -ésimo estrato.

$N_{hij}$ : número de domicílios particulares (ocupados, vagos, fechados e de uso ocasional) no  $j$ -ésimo setor do  $i$ -ésimo município do  $h$ -ésimo estrato.

<sup>2</sup> As PNADs só coletam informações na zona urbana de Unidades da Federação da Região Norte.

$n_{hij}$ : número de domicílios entrevistados no j-ésimo setor do i-ésimo município do h-ésimo estrato.

$y_{hijk}$ : valor de  $y$  para o k-ésimo domicílio entrevistado do j-ésimo setor do i-ésimo município do h-ésimo estrato.

Logo um estimador do total de  $y$  para um domínio geográfico (d), é dado por:

$$\hat{Y}_{(d)} = \sum_{h=1}^{H_{(d)}} \frac{1}{l_h} \sum_{i=1}^{l_h} \frac{1}{m_{hi} P_{hi}} \sum_{j=1}^{m_{hi}} \frac{N_{hij}}{P_{hij} n_{hij}} \sum_{k=1}^{n_{hij}} y_{hijk} \quad (3)$$

Na equação (3) considera-se que cada município auto-representativo forme um estrato e  $l_h=1$  e  $P_{hi}=1$  para todo h, tal que h seja um município auto-representativo.

Os estimadores dos  $A_{i,k}$ 's e dos  $l_{i,k}$ 's podem ser obtidos através das respectivas razões dos estimadores de totais fornecidos pela equação (3). Representando-se genericamente estas razões num domínio (d) de interesse por  $\hat{R}_{(d)}$ , tem-se:

$$\hat{R}_{(d)} = \frac{\hat{N}_{(d)}}{\hat{D}_{(d)}} \quad (4)$$

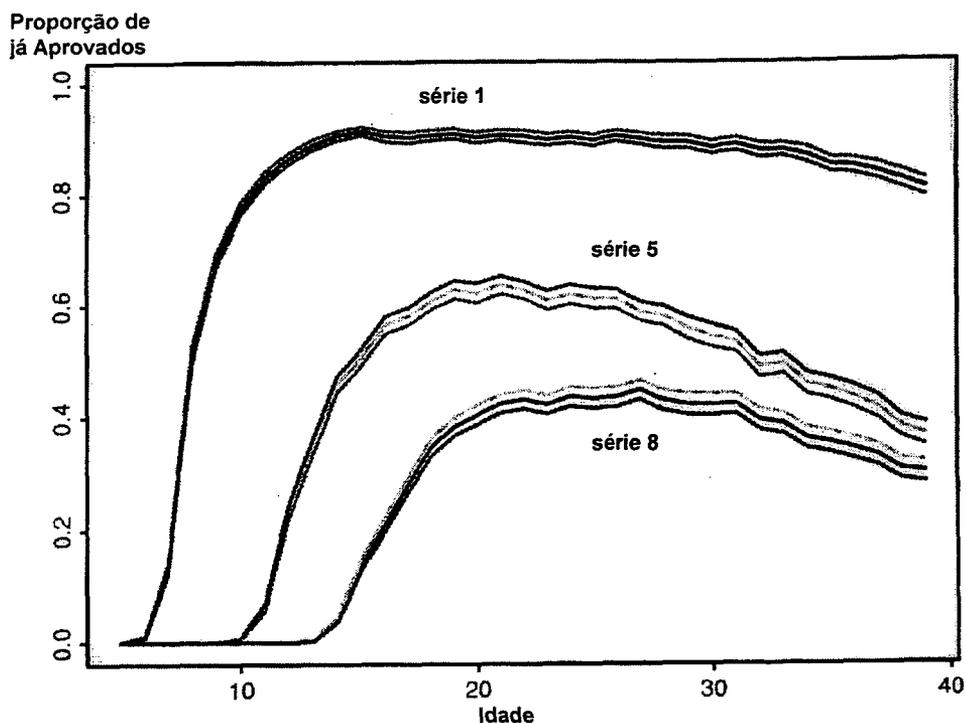
onde  $\hat{N}_{(d)}$  e  $\hat{D}_{(d)}$  representam genericamente os respectivos estimadores de totais para os numeradores e denominadores das equações (1) e (2).

Os estimadores de variâncias das razões representadas pela equação (4) estão descritos no apêndice.

### 3. RESULTADOS E CONCLUSÕES

A Figura 2 mostra os gráficos das taxas estimadas dos já aprovados, para o Brasil, por idade para a 1ª, 5ª e 8ª séries do Ensino Fundamental, com os respectivos intervalos de 95% de confiança. Na Tabela 1, no Apêndice B, são exibidas as estimativas das taxas dos já aprovados (iguais às taxas de conclusão para as coortes mais velhas) em nível de Brasil, para a 1ª, 5ª e 8ª séries do Ensino Fundamental e algumas idades de 5 a 39 anos, com os respectivos erros-padrão e coeficientes de variação. Pode-se ver que as curvas são crescentes no início, estabilizam em torno de um máximo, e depois decrescem.

Figura 2  
 Proporção de já Aprovados, com intervalo de confiança, por Idade,  
 para a 1<sup>a</sup>, 5<sup>a</sup> e 8<sup>a</sup> séries, no Brasil



Pode-se observar da Tabela 1 que as estimativas das taxas de já aprovados por idade são bastante precisas para todas as séries e idades, com exceção das idades iniciais e finais. Contudo a precisão das estimativas decresce com a série, conseqüência do menor número de alunos nas séries superiores. Em torno da taxa de conclusão (máximo das taxas dos já aprovados), tem-se um coeficiente de variação de aproximadamente 0,5% , 1,4% e 2,0% para a 1<sup>a</sup> , 5<sup>a</sup> e 8<sup>a</sup> séries, respectivamente.

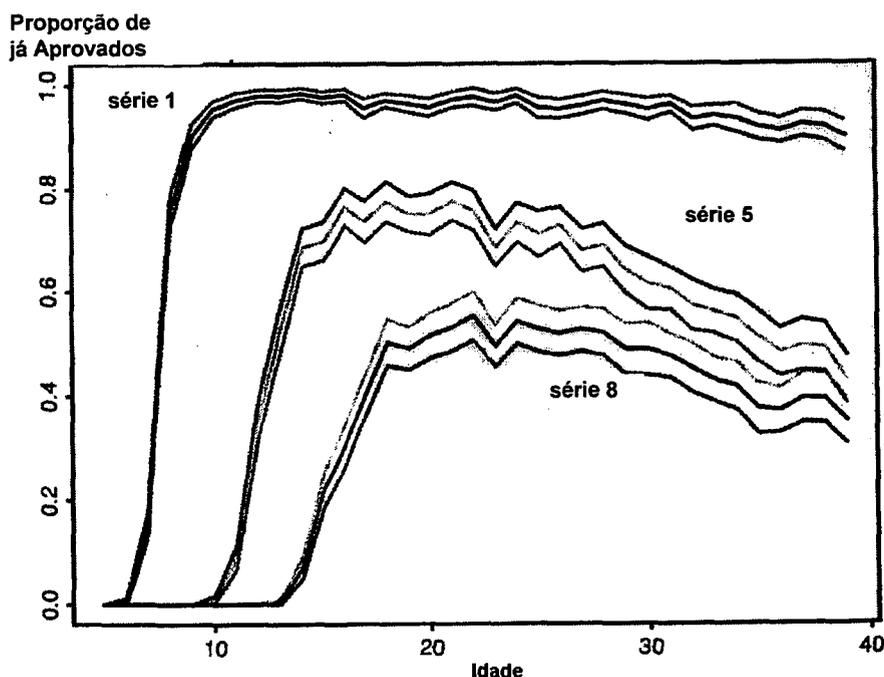
Estima-se que 51% da coorte de 8 anos (a idade desejada) e 77% da coorte de 10 anos já concluíram a 1<sup>a</sup> série. O máximo dessas taxas por idade (a taxa de conclusão da 1<sup>a</sup> série) é cerca de 91% e ocorre por volta da coorte de 15 anos. Em 1990, havia alunos dos 7 aos 15 anos terminando esta série, uma faixa muito ampla de idade, reflexo da alta repetência e de algum acesso tardio à série. Acima dos 15 anos, praticamente não existe ninguém mais terminando esta série. O fato da curva ser decrescente dessa idade em diante implica o fato de que as taxas de conclusão apresentam uma tendência de crescimento em função dos anos.

Na 5<sup>a</sup> série, o máximo da curva é de cerca de 63%, ocorrendo em torno de 19 anos de idade em 1989 e na 8<sup>a</sup> série é de cerca de 43%, ocorrendo por volta dos 22 anos. Novamente verifica-se que existem alunos em uma faixa larga de idade terminando cada uma das duas séries. A Figura 2 mostra ainda que o crescimento das taxas de conclusão é mais acelerado na 5<sup>a</sup> série do que na 8<sup>a</sup>.

Para se investigar o efeito do tamanho das amostras na precisão das estimativas consideraram-se o Estado de São Paulo e o Estado do Acre (zona urbana), que representam respectivamente cerca de 21,5% e 0,2% da população brasileira e cerca de 14% e 0,2% da amostra da PNAD 1990.

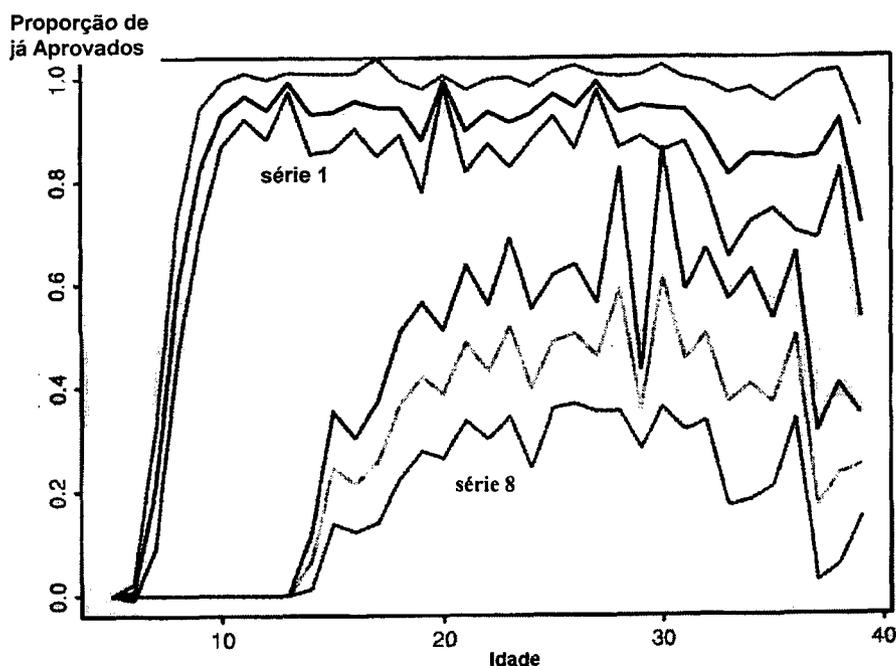
A Figura 3 mostra os gráficos das taxas estimadas dos já aprovados, para o Estado de São Paulo, por idade para a 1<sup>a</sup>, 5<sup>a</sup> e 8<sup>a</sup> séries do Ensino Fundamental, com os respectivos intervalos de 95% de confiança. A Tabela 2 do Apêndice B mostra as estimativas das taxas de já aprovados e seus respectivos erros-padrão e coeficientes de variação para o Estado de São Paulo, cujo tamanho da amostra ainda é relativamente grande. Pode-se ver que a precisão diminui, embora ainda seja aceitável. Os coeficientes de variação em torno das taxas de conclusão estão em torno de 0,6%, 2,6% e 4,3% para a 1<sup>a</sup>, 5<sup>a</sup> e 8<sup>a</sup> séries, respectivamente.

Figura 3  
Proporção de já Aprovados, com intervalos de confiança, por idade, para a 1<sup>a</sup>, 5<sup>a</sup> e 8<sup>a</sup> séries, no Estado de São Paulo



Para a zona urbana do Estado do Acre, as estimativas são bastante imprecisas. Inclusive nota-se que as estimativas dos erros-padrão são bastante instáveis (Figura 4 e Tabela 3 do Apêndice B.).

Figura 4  
Proporção de já Aprovados, com intervalos de confiança, por Idade,  
para a 1ª e 8ª séries, no Estado do Acre



Cabe ressaltar que para o Estado de São Paulo, as estimativas das taxas de conclusão nas diversas séries são maiores que no Brasil. Temos também que as idades onde os máximos das curvas ocorrem são menores em São Paulo e conseqüentemente a faixa de idade de conclusão de cada série é mais estreita.

Resultados análogos podem ser obtidos para as outras séries e para as taxas dos já ingressos (e conseqüentemente para as taxas de participação).

As taxas de participação e conclusão são indicadores importantes do acesso e conclusão de uma série por uma coorte. Os resultados acima nos indicam que para o Brasil, regiões e estados grandes, podemos utilizar as PNADs para obter esses indicadores com uma boa precisão.

Supondo-se estabilidade do sistema escolar, pode-se estimar a proporção da coorte de idade  $i$  que termina a série com esta idade pela diferença da taxa dos já aprovados com a idade  $i$  e dos já aprovados com a idade  $(i-1)$ . Trabalhando-se com duas PNADs consecutivas, digamos 1989 e 1990, poderia-se, em princípio, estimar sem a hipótese de estabilidade. A diferença entre a proporção dos já aprovados com a idade  $i$  em 1990 e a idade  $(i-1)$  em 1989 é a proporção da coorte de idade  $i$  que termina a série com a idade  $i$ . Pretende-se em um próximo trabalho estudar a precisão dessas estimativas.

## Apêndice A Estimadores de Variância

Os estimadores de variância das razões representadas pela equação (4) na seção 2.2 foram calculados através do método do "ultimate cluster" como descrito em Hansen, Hurwitz e Madow (1953). Embora os setores e municípios tenham sido selecionados sistematicamente, na estimação de variância admitiu-se que foram selecionados com reposição.

As variâncias dos totais  $\hat{Y}_{(d)}$  podem ser obtidas pela seguinte equação:

$$\hat{V}(\hat{Y}_{(d)}) = \hat{V}_1 + \hat{V}_2 \quad (\text{A.1})$$

onde:

$$\hat{V}_1 = \sum_{h \in \tilde{A}_{(d)}} \frac{1}{l_h(l_h - 1)} \sum_{i=1}^{l_h} \left( \frac{\hat{Y}_{hi}}{P_{hi}} - \hat{Y}_h \right)^2 = \sum_{h \in \tilde{A}_{(d)}} \frac{1}{l_h} s_{h,(Y,Y)} \quad (\text{A.2})$$

$$\hat{V}_2 = \sum_{hi \in A_{(d)}} \frac{1}{m_{hi}(m_{hi} - 1)} \sum_{j=1}^{m_{hi}} \left( \frac{\hat{Y}_{hij}}{P_{hij}} - \hat{Y}_{hi} \right)^2 = \sum_{hi \in A_{(d)}} \frac{1}{m_{hi}} s_{hi,(Y,Y)} \quad (\text{A.3})$$

$\tilde{A}_{(d)}$  estratos formados por municípios não-autoponderados pertencentes ao domínio de interesse (d)

$A_{(d)}$  municípios autoponderados pertencentes ao domínio de interesse (d)

$$\hat{Y}_h = \frac{1}{l_h} \sum_{i=1}^{l_h} \frac{\hat{Y}_{hi}}{P_{hi}} \quad \hat{Y}_{hi} = \frac{1}{m_{hi}} \sum_{i=1}^{m_{hi}} \frac{\hat{Y}_{hij}}{P_{hij}} \quad \hat{Y}_{hij} = \frac{N_{hij}}{n_{hij}} \sum_{k=1}^{n_{hij}} y_{hijk}$$

Considerando-se o estimador de razão dado pela expressão (4) na seção 2.2, tem-se o seguinte estimador de variância obtido pelo método do "ultimate cluster" (Hansen, M.H., Hurwitz, W.N. e Madow, W.G. (1953), Vol I, p. 419):

$$\begin{aligned} \hat{V}(\hat{R}_{(d)}) = & \frac{1}{(\hat{D}_{(d)})^2} \left[ \sum_{h \in \tilde{A}_{(d)}} \frac{1}{l_h} (s_{h,(N,N)} + \hat{R}_{(d)}^2 s_{h,(D,D)} - 2\hat{R}_{(d)} s_{h,(N,D)}) \right] \\ & + \frac{1}{(\hat{D}_{(d)})^2} \left[ \sum_{hi \in A_{(d)}} \frac{1}{m_{hi}} (s_{hi,(N,N)} + \hat{R}_{(d)}^2 s_{hi,(D,D)} - 2\hat{R}_{(d)} s_{hi,(D,N)}) \right] \end{aligned} \quad (\text{5})$$

onde  $s_{h,(Y,Y)}$  para  $Y=N, D$  vem da segunda igualdade da equação (6), e  $s_{h,(N,Y)}$  é definido

por:  $s_{h,(Y,X)} = \frac{1}{(l_h - 1)} \sum_{i=1}^{l_h} \left( \frac{\hat{Y}_{hi}}{P_{hi}} - \hat{Y}_h \right) \left( \frac{\hat{X}_{hi}}{P_{hi}} - \hat{X}_h \right)$  para  $Y=N$  e  $X=D$ . Analogamente define-se

$s_{i,(Y,X)}$  para  $Y=N, D$  e  $X=N, D$ .

## Apêndice B

Tabela 1 - Estimativas das Taxas dos já Aprovados por Idade e Respectivos Erros-Padrão e Coeficientes de Variação para a Primeira, Quinta e Oitava Séries do Ensino Fundamental. Brasil

Idade	Primeira Série			Quinta Série			Oitava Série		
	Estimat.	Erro-padrão	Coef. Var.	Estimat.	Erro-padrão	Coef. Var.	Estimat.	Erro-padrão	Coef. Var.
5	0.0003	0.0002	62.29	0.0000	0.0000	-	0.0000	0.0000	-
6	0.0086	0.0012	14.37	0.0000	0.0000	-	0.0000	0.0000	-
7	0.1314	0.0050	3.77	0.0000	0.0000	-	0.0000	0.0000	-
8	0.5116	0.0072	1.41	0.0000	0.0000	-	0.0000	0.0000	-
9	0.6834	0.0064	0.93	0.0002	0.0002	104.86	0.0000	0.0000	-
10	0.7745	0.0060	0.77	0.0050	0.0010	20.37	0.0000	0.0000	-
11	0.8317	0.0053	0.64	0.0635	0.0036	5.59	0.0000	0.0000	-
12	0.8673	0.0049	0.56	0.2352	0.0064	2.71	0.0001	0.0001	104.26
13	0.8918	0.0046	0.52	0.3501	0.0074	2.11	0.0037	0.0009	23.46
14	0.9079	0.0043	0.47	0.4598	0.0075	1.62	0.0417	0.0030	7.29
15	0.9163	0.0037	0.41	0.5069	0.0075	1.48	0.1397	0.0051	3.62
16	0.9063	0.0045	0.50	0.5645	0.0081	1.43	0.2103	0.0069	3.28
17	0.9043	0.0049	0.54	0.5809	0.0082	1.41	0.2819	0.0069	2.46
18	0.9076	0.0045	0.50	0.6078	0.0079	1.30	0.3470	0.0078	2.24
19	0.9106	0.0047	0.52	0.6285	0.0084	1.33	0.3835	0.0080	2.09
20	0.9036	0.0050	0.55	0.6220	0.0090	1.45	0.4038	0.0086	2.13
21	0.9084	0.0047	0.51	0.6368	0.0084	1.33	0.4236	0.0084	1.99
27	0.9015	0.0052	0.57	0.5890	0.0086	1.46	0.4483	0.0087	1.93
33	0.8809	0.0058	0.65	0.4949	0.0096	1.93	0.3880	0.0095	2.44
39	0.8113	0.0082	1.01	0.3673	0.0100	2.71	0.3011	0.0095	3.14

Tabela 2 - Estimativas das Taxas dos já Aprovados por Idade e Respectivos Erros-Padrão e Coeficientes de Variação para a Primeira, Quinta e Oitava Séries do Ensino Fundamental. São Paulo

Idade	Primeira Série			Quinta Série			Oitava Série		
	Estimat.	Erro-padrão	Coef. Var.	Estimat.	Erro-padrão	Coef. Var.	Estimat.	Erro-padrão	Coef. Var.
5	0.0000	0.0000	-	0.0000	0.0000	-	0.0000	0.0000	-
6	0.0061	0.0032	52.49	0.0000	0.0000	-	0.0000	0.0000	-
7	0.1545	0.0129	8.32	0.0000	0.0000	-	0.0000	0.0000	-
8	0.7576	0.0165	2.18	0.0000	0.0000	-	0.0000	0.0000	-
9	0.9020	0.0120	1.33	0.0000	0.0000	-	0.0000	0.0000	-
10	0.9536	0.0077	0.81	0.0085	0.0036	42.18	0.0000	0.0000	-
11	0.9702	0.0070	0.72	0.0896	0.0111	12.43	0.0000	0.0000	-
12	0.9782	0.0057	0.59	0.3556	0.0197	5.54	0.0000	0.0000	-
13	0.9785	0.0062	0.63	0.5284	0.0220	4.16	0.0049	0.0028	56.44
14	0.9834	0.0052	0.53	0.6847	0.0186	2.72	0.0705	0.0105	14.94
15	0.9750	0.0058	0.59	0.6964	0.0190	2.73	0.2212	0.0170	7.70
16	0.9808	0.0058	0.59	0.7640	0.0187	2.44	0.3035	0.0209	6.89
17	0.9548	0.0093	0.98	0.7347	0.0206	2.80	0.3998	0.0203	5.08
18	0.9696	0.0065	0.67	0.7738	0.0199	2.57	0.5018	0.0229	4.57
19	0.9640	0.0081	0.84	0.7495	0.0177	2.37	0.4901	0.0209	4.27
20	0.9561	0.0086	0.90	0.7490	0.0206	2.75	0.5144	0.0221	4.29
21	0.9715	0.0075	0.77	0.7749	0.0192	2.48	0.5316	0.0231	4.34
27	0.9585	0.0090	0.94	0.6787	0.0211	3.11	0.5266	0.0213	4.04
33	0.9396	0.0102	1.09	0.5600	0.0214	3.81	0.4290	0.0214	4.99
39	0.8994	0.0151	1.67	0.4294	0.0235	5.47	0.3498	0.0226	6.46

Tabela 3 - Estimativas das Taxas dos já Aprovados por Idade e Respectiveiros Erros-Padrão e Coeficientes de Variação para a Primeira, Quinta e Oitava Séries do Ensino Fundamental.  
Acre - Zona Urbana

Idade	Primeira Série			Quinta Série			Oitava Série		
	Estimat.	Erro-padrão	Coef. Var.	Estimat.	Erro-padrão	Coef. Var.	Estimat.	Erro-padrão	Coef. Var.
5	0.0000	0.0000	-	0.0000	0.0000	-	0.0000	0.0000	-
6	0.0070	0.0081	114.49	0.0000	0.0000	-	0.0000	0.0000	-
7	0.2074	0.0587	28.30	0.0000	0.0000	-	0.0000	0.0000	-
8	0.5982	0.0678	11.33	0.0000	0.0000	-	0.0000	0.0000	-
9	0.8212	0.0609	7.42	0.0000	0.0000	-	0.0000	0.0000	-
10	0.9300	0.0320	3.44	0.0000	0.0000	-	0.0000	0.0000	-
11	0.9656	0.0226	2.34	0.0470	0.0341	72.49	0.0000	0.0000	-
12	0.9398	0.0299	3.18	0.1693	0.0518	30.58	0.0000	0.0000	-
13	0.9915	0.0097	0.98	0.4617	0.0880	19.06	0.0000	0.0000	-
14	0.9308	0.0400	4.30	0.5708	0.0683	11.96	0.0643	0.0270	42.06
15	0.9343	0.0379	4.06	0.6970	0.0663	9.51	0.2450	0.0551	22.50
16	0.9561	0.0276	2.89	0.6496	0.0586	9.03	0.2119	0.0458	21.62
17	0.9429	0.0485	5.14	0.7303	0.0715	9.79	0.2538	0.0592	23.33
18	0.9420	0.0274	2.91	0.6849	0.0628	9.16	0.3653	0.0721	19.73
19	0.8790	0.0512	5.82	0.6210	0.0484	7.80	0.4196	0.0732	17.44
20	0.9948	0.0056	0.56	0.7834	0.0875	11.17	0.3865	0.0633	16.38
21	0.8977	0.0411	4.58	0.6669	0.0919	13.79	0.4860	0.0768	15.80
27	0.9918	0.0082	0.83	0.6235	0.0555	8.90	0.4581	0.0538	11.74
33	0.8108	0.0809	9.97	0.4557	0.1248	27.40	0.3689	0.1018	27.61
39	0.7175	0.0931	12.98	0.2757	0.0648	23.49	0.2469	0.0507	20.51

## REFERÊNCIAS

- COCHRAN, W.G. (1977). *Sampling Techniques*. John Wiley and Sons, New York.
- FLETCHER, P. R. E RIBEIRO, S. C. (1988). *Projeto Fluxo dos Alunos de Primeiro Grau*. - PROFLEXO. (mimeo).
- FLETCHER, P. R. E RIBEIRO, S. C. (1989). *Modeling Education System Performance with Demographic Data. An Introduction to the PROFLEXO Model*. (mimeo)
- HANSEN, M.H. , HURWITZ, W.N. E MADOW, W.G. (1953). *Sample Survey Methods and Theory*. John Wiley and Sons, New York, Vol. I e II.
- KLEIN, R. (1991). Relatório 1. *Descrição da Metodologia do PROFLEXO*. United Nations - Project Symbol: BRA/90/026 - M.O.D. Nº. 91-006.
- KLEIN, R. (1995). *Produção e Utilização de indicadores Educacionais (2ª Versão Preliminar)*. (mimeo)
- KLEIN, R. E RIBEIRO, S. C. (1991). *O Censo Educacional e o Modelo de Fluxo: O Problema da Repetência*. Revista Brasileira de Estatística, 52(197/198), pg. 5-45.
- Pesquisa Nacional Por Amostra de Domicílios (1990). Rio de Janeiro: IBGE, Vol. 14, Num 1.

## ABSTRACT

Fletcher e Ribeiro (1988) develop educational indicators, using data from one PNAD (National Survey by Household Sampling). This methodology was called Profluxo. The aim of this paper is to estimate and analyse the precision of the participation and conclusion rates of a cohort in a certain grade. It can be seen that the estimates at the Brazilian level have good precision. The precision of the estimates for a state with a "large" sample size as São Paulo and one with a "small" sample size as Urban Acre are also studied.

# MÉTODOS BAYESIANOS EM MODELOS DE CONFIABILIDADE DE SOFTWARE USANDO PROCESSOS DE POISSON NÃO-HOMOGÊNEOS

Karin Storani\*

Jorge Alberto Achcar\*\*

## RESUMO

Neste artigo, utilizamos métodos Bayesianos para fazer inferências dos modelos de confiabilidade de software que utilizam processos de Poisson não-homogêneos. Uma extensão desses modelos é proposta, considerando-se a distribuição gaussiana inversa na modelagem da função de valor médio dos processos de Poisson não-homogêneos. Além disso, exploramos o uso de algoritmos computacionais como o amostrador de Gibbs e Metropolis-Hastings, para obter os sumários *a posteriori*. Tendo em vista as suposições dos modelos de estatísticas de ordem, um método de diagnóstico do modelo, utilizando-se o amostrador de Gibbs é proposto. A metodologia desenvolvida neste artigo é exemplificada com um conjunto de dados introduzido por Jelinski e Moranda (1972).

**Palavras-Chave:** Processos de Poisson não-homogêneos, Distribuição Gaussiana Inversa, Gibbs Sampling e Metropolis-Hastings.

## 1. INTRODUÇÃO

---

Muitos modelos probabilísticos e técnicas estatísticas são utilizados por pesquisadores da área de computação para avaliar a performance de um software, o qual está sujeito a falhas devido à presença de erros.

---

\* Endereço para correspondência: ICEX - Des - UFMG - C.P. 702, 30123-970 Belo Horizonte - MG, Brasil.

\*\* ICMSC - USP - C.P. 668, 13560-970 - São Carlos, SP, Brasil.

A literatura em estatística e em engenharia de software apresenta vários modelos para modelar a confiabilidade de um software (ver por exemplo Mazuchi e Singpurwalla, 1988; ou Singpurwalla e Wilson, 1994).

Duas estratégias de modelagem muito utilizadas são:

- (i) Estratégia de tipo I: modela os tempos entre falhas sucessivas do software; e
- (ii) Estratégia de tipo II: modela o número de falhas do software até um determinado tempo (usual: tempo de CPU).

Os modelos de confiabilidade de software de estratégia de tipo II são classificados de acordo com a especificação da distribuição do número de falhas observadas  $M(t)$  no intervalo de tempo  $(0, t]$ . Os processos de Poisson homogêneo e não-homogêneo são os mais usados para modelar o número de falhas observadas  $M(t)$  no intervalo  $(0, t]$  (ver Musa, Iannino e Okumoto, 1987).

Neste artigo, damos ênfase à utilização de processos de Poisson não-homogêneo para modelar dados de confiabilidade de software, considerando formas diferentes para a função de valor médio  $m(t)$  ou equivalentemente para a função intensidade que é a derivada de  $m(t)$  com respeito a  $t$ .

Usando métodos Bayesianos de inferência estatística, obtemos sumários *a posteriori* de interesse a partir de simulação estocástica baseada em Cadeias de Markov com Monte Carlo - MCMC - com especial destaque aos algoritmos Gibbs Sampling e Metropolis-Hastings (ver por exemplo, Gelfand e Smith, 1990; ou Smith e Roberts, 1992).

## **2. ALGUNS MODELOS DE CONFIABILIDADE DE SOFTWARE COM PROCESSOS DE POISSON NÃO-HOMOGÊNEOS**

Os modelos de confiabilidade de software com processos de Poisson não-homogêneos podem ser classificados de acordo com o comportamento limite da função de valor médio  $m(t)$ : (1) modelos com número de falhas finito ( $\lim_{t \rightarrow \infty} m(t) < \infty$ ), denotados por PPNH-I; (2) modelos com número de falhas infinito (ou seja  $m(t) \rightarrow \infty$  com  $t \rightarrow \infty$ ), denotados por PPNH-II.

Recentemente, Kuo e Yang (1995) propuseram uma teoria unificada para modelos com números de falhas finito e infinito.

Um caso especial de processos de Poisson não-homogêneos com número de falhas finito (PPNH-I) é dado por,

$$m(t) = \theta F(t) \quad (1)$$

onde é suposto que existe um número desconhecido  $N$  de falhas no início da fase de testes de um software e modelamos os tempos de falhas como sendo as  $n$  primeiras estatísticas de ordem tomadas de  $N$  observações i.i.d com função densidade  $f$  e função

distribuição F. Casos especiais de F(t) em (1) levam a vários modelos usados em confiabilidade de software:

(i) Se  $F(t)=1-e^{-\beta t}$ , temos o processo de Goel e Okumoto (1979) (um modelo de estatísticas de ordem exponencial);

(ii) Se  $F(t)=1-e^{-\beta t^\alpha}$ , temos o processo de Goel (1983) (um modelo de estatísticas de ordem Weibull);

(iii) Se  $F(t)=1-(1+\beta t)e^{-\beta t}$ , temos o processo de Ohba-Yamada (Yamada et al., 1983);

(iv) Se  $F(t)=I_k(\beta t^\alpha)$  onde  $I_k(\cdot)$  é uma integral gama incompleta dada por

$$I_k(s) = \frac{1}{\Gamma(k)} \int_0^s x^{k-1} e^{-x} dx,$$

temos um modelo de estatísticas de ordem gama-generalizada (ver Achcar, Dey e Niverthy, 1996);

(v) Se  $F(t)=\Phi\left(\frac{\ln t - \mu}{\sigma}\right)$  onde  $\Phi(\cdot)$  é a função distribuição de uma variável aleatória com

distribuição normal padronizada  $N(0,1)$ , temos um modelo de estatísticas de ordem log-normal; e

(vi) Se  $F(t) = \Phi\left[\sqrt{\frac{\lambda}{t}}\left(\frac{t}{\mu} - 1\right)\right] + e^{2\lambda/\mu}\Phi\left[-\sqrt{\frac{\lambda}{t}}\left(1 + \frac{t}{\mu}\right)\right]$ , temos um modelo de estatísticas

de ordem gaussiana inversa (ver Achcar e Storani, 1996).

Considerando processos de Poisson não-homogêneos com número de falhas infinito (PPNH-II), um caso especial é dado por,

$$m(t)=-\ln(1-F(t)) \quad (2)$$

que se reduz ao processo de Musa e Okumoto (1984) com  $\lambda(t) = \alpha/(t + \beta)$ ; ou ao processo de Duane (1964) com  $\lambda(t) = \alpha\beta t^{\alpha-1}$  e ao processo de Cox e Lewis (1966) com  $\lambda(t) = \exp(\alpha + \beta t)$ .

### **3. INFERÊNCIA BAYESIANA PARA MODELOS PPNH-I**

Assumir que a função de valor médio  $m(t)$  é indexada pelos parâmetros desconhecidos  $\theta$  e  $\beta$ .

Dado o modelo de tempo truncado que testa até o tempo  $t$ , os tempos ordenados dos  $n$  tempos de falhas observados são denotados por  $x_1, x_2, \dots, x_n$ .

A função de verossimilhança é dada por

$$L(\theta, \beta \mid D_t) = \prod_{i=1}^n \lambda(x_i) \exp\{-m(t)\}. \quad (3)$$

onde  $D_t = \{n; x_1, \dots, x_n; t\}$  é o conjunto de dados (ver por exemplo, Cox e Lewis, 1966).

Para o modelo de falhas truncadas, uma expressão similar a (3) pode ser aplicada com  $t$  substituído por  $x_n$ .

Para inferência Bayesiana deste modelo, consideramos o uso dos algoritmos de Metropolis com Gibbs. Como a presença da expressão  $m(t) = \theta F(t)$  para modelos PPNH-I na função de verossimilhança (3) usualmente dificulta a especificação de uma forma conveniente para as densidades condicionais de  $\theta$  e  $\beta$  dado  $D_t$ , necessárias no algoritmo Gibbs sampling, consideramos a introdução de uma variável latente  $N' = N - n$  que tem uma distribuição de Poisson com parâmetro  $\theta[1 - F(t \mid \beta)]$  (ver Kuo e Yang, 1995).

Nesta forma, a distribuição *a posteriori* conjunta  $p(\theta, \beta \mid D_t)$  pode ser obtida da distribuição conjunta  $P(\theta, \beta, N' \mid D_t)$  por marginalização. A densidade *a posteriori* conjunta  $p(\theta, \beta, N' \mid D_t)$  é aproximada das amostras de Gibbs geradas das seguintes densidades condicionais:  $p(N' \mid \theta, \beta, D_t)$ ,  $p(\theta \mid N', \beta, D_t)$  e  $p(\beta \mid N', \theta, D_t)$ .

Como um caso especial, considerar a função de valor médio  $m(t)$  dada em (1) com  $F(t) = I_k(\beta t^\alpha)$  (distribuição gama generalizada) que corresponde à função intensidade

$$\lambda(t) = \frac{dm(t)}{dt} = \frac{1}{\Gamma(k)} \theta \beta^k \alpha t^{k-1} e^{-\beta t^\alpha} \quad (4)$$

de (3), a função de verossimilhança para  $\theta, \alpha, \beta$  e  $k$  é dada por,

$$L(\theta, \alpha, \beta, k) = \frac{\theta^n \alpha^n \beta^{kn}}{\{\Gamma(k)\}^n} \left\{ \prod_{i=1}^n x_i^{\alpha k - 1} \right\} \exp\left\{ -\beta \sum_{i=1}^n x_i^\alpha - \theta I_k(\beta t^\alpha) \right\}. \quad (5)$$

Considere as seguintes densidades *a priori* para  $N', \theta, \alpha, \beta$  e  $k$ :

(i)  $N' \sim P\{\theta[1 - I_k(\beta t^\alpha)]\}$ , para  $N' > 0$ ;

(ii)  $\theta \sim \Gamma(a,b)$ , onde  $a$  e  $b$  são conhecidos e  $\theta > 0$ ;

(iii)  $\beta \sim \Gamma(c,d)$ , onde  $c$  e  $d$  são conhecidos e  $\beta > 0$ ;

(iv)  $\alpha \sim \pi_1(\alpha)$ , onde  $\pi_1(\alpha)$  é uma densidade *a priori* para  $\alpha$  ( $\alpha > 0$ ); e (6)

(v)  $k \sim \pi_2(k)$ , onde  $\pi_2(k)$  é uma densidade *a priori* para  $k$  ( $k > 0$ ).

onde  $P(\lambda)$  denota a distribuição de Poisson com parâmetro  $\lambda$ ;  $\Gamma(a,b)$  denota a distribuição gama com média  $a/b$  e variância  $a/b^2$ . Além disso assumir independência entre os parâmetros  $\theta, \alpha, \beta$  e  $k$ .

A densidade *a posteriori* conjunta é dada por,

$$\pi(N', \alpha, \beta, k, \theta \mid D_t) \propto \frac{\theta^{N'+n+\alpha-1} \alpha^n \beta^{kn+c-1}}{N'! \{\Gamma(k)\}^n} \times \left\{ \prod_{i=1}^n x_i^{\alpha k - 1} \right\} \left\{ 1 - I_k(\beta t^\alpha) \right\}^{N'} e^{-(b+1)\theta - (d + \sum_{i=1}^n x_i^\alpha) \beta} \pi_1(\alpha) \pi_2(k), \quad (7)$$

As densidades marginais para o algoritmo Gibbs Sampling são dadas por:

(i)  $N' \mid \alpha, \beta, k, \theta, D_t \sim P[\theta(1 - I_k(\beta t^\alpha))]$ ;

(ii)  $\theta \mid N', \alpha, \beta, k, D_t \sim \Gamma[a + n + N', b + 1]$ ;

(iii)  $\pi(\beta \mid N', \alpha, k, \theta, D_t) \propto \beta^{kn+c-1} e^{-\beta \left[ \sum_{i=1}^n x_i^\alpha + d \right]} \left\{ 1 - I_k(\beta t^\alpha) \right\}^{N'}$ ;

(iv)  $\pi(\alpha \mid \theta, N', \beta, k, D_t) \propto \alpha^n \left\{ \prod_{i=1}^n x_i^{\alpha k - 1} \right\} e^{-\beta \sum_{i=1}^n x_i^\alpha} \left\{ 1 - I_k(\beta t^\alpha) \right\}^{N'} \pi_1(\alpha)$ ; (8)

(v)  $\pi(k \mid N', \alpha, \beta, \theta, D_t) \propto \frac{\beta^{kn}}{\{\Gamma(k)\}^n} \left\{ \prod_{i=1}^n x_i^{\alpha k - 1} \right\} \left\{ 1 - I_k(\beta t^\alpha) \right\}^{N'} \pi_2(k)$ .

Observar que as variáveis  $\alpha, \beta$  e  $k$  devem ser geradas usando o algoritmo de Metropolis-Hastings (ver Chib e Greenberg, 1995).

Também observar que o modelo de estatísticas de ordem gama generalizado é um supermodelo que engloba vários casos especiais usados em confiabilidade de software. Por exemplo, considerando-se  $\alpha = 1$  e  $k=1$ , temos um processo de Goel e Okumoto (1979); se  $k=1$  temos um processo de Goel (1983); se  $\alpha = 1$  temos um modelo de estatísticas de ordem gama e se  $k \rightarrow \infty$ , temos um modelo de estatísticas de ordem normal (ver Lawless, 1982).

Assim, podemos obter sumários *a posteriori* para os parâmetros de interesse considerando densidade *a priori* similares às dadas em (6) e amostras geradas pelo algoritmo Metropolis com Gibbs.

Considerando agora a função de valor médio  $m(t)$  dada em (1) com

$$F(t) = \Phi \left[ \sqrt{\frac{\lambda}{t}} \left( \frac{t}{\mu} - 1 \right) \right] + e^{2\lambda/\mu} \Phi \left[ -\sqrt{\frac{\lambda}{t}} \left( 1 + \frac{t}{\mu} \right) \right] \quad (9)$$

(distribuição gaussiana inversa), onde  $\Phi(\cdot)$  denota a função distribuição de uma variável aleatória com distribuição normal padronizada  $N(0,1)$ , que corresponde à função intensidade

$$\lambda(t) = \frac{\theta\sqrt{\lambda}}{\sqrt{2\pi}} t^{-3/2} \exp \left\{ -\frac{\lambda(t-\mu)^2}{2\mu^2 t} \right\}, \quad (10)$$

a função de verossimilhança (3) para  $\theta, \mu$  e  $\lambda$  é dada por,

$$L_{GI}(\theta, \mu, \lambda) = \frac{\theta^n \lambda^{n/2}}{(2\pi)^{n/2}} \left\{ \prod_{i=1}^n x_i^{-3/2} \right\} \exp \left\{ -\theta F(x_n) - \frac{\lambda}{2\mu^2} \sum_{i=1}^n \frac{(x_i - \mu)^2}{x_i} \right\}, \quad (11)$$

onde  $F(\cdot)$  é definida em (9).

Supor as seguintes densidades *a priori* informativas para  $N'$ ,  $\theta$ ,  $\mu$  e  $\lambda$  dadas por,

(i)  $N' \sim P[\theta[1-F(t)]];$

(ii)  $\mu \sim \Gamma(a_1, b_1); a_1, b_1$  conhecidos;

(iii)  $\lambda \sim \Gamma(a_2, b_2); a_2, b_2$  conhecidos; e (12)

(iv)  $\theta \sim \Gamma(a_3, b_3)$ ;  $a_3, b_3$  conhecidos.

Assumindo independência *a priori* entre os parâmetros, a densidade *a posteriori* conjunta é dada por,

$$\begin{aligned} \pi(N', \mu, \lambda, \theta \mid D_t) &\propto \theta^{N'+n+a_2-1} \lambda^{n/2+a_1-1} \mu^{a_1-1} \frac{\{1-F(x_n)\}^{N'}}{N!} \\ &\exp\left\{-\frac{\lambda}{2\mu^2} \sum_{i=1}^n \frac{(x_i - \mu)^2}{x_i} - b_1\mu - b_2\lambda - \theta(b_3 + 1)\right\}, \end{aligned} \quad (13)$$

As densidades condicionais de  $N'$ ,  $\theta$ ,  $\mu$  e  $\lambda$  para o algoritmo Gibbs Sampling são dadas por,

(i)  $N' \mid \mu, \lambda, \theta, D_t \sim P[\theta(1-F(x_n))]$ ;

(ii)  $\theta \mid \mu, \lambda, N', D_t \sim \Gamma[N' + n + a_3, b_3 + 1]$ ;

(iii)  $\pi(\mu \mid \lambda, \theta, N', D) \propto \mu^{a_1-1} e^{-b_1\mu} \Psi_1(\mu, \lambda)$ ,

$$\text{onde } \Psi_1(\mu, \lambda) = \exp\left\{-\frac{\lambda}{2\mu^2} \sum_{i=1}^n \frac{(x_i - \mu)^2}{x_i} + N' \log(1-F(x_n))\right\}; \quad (14)$$

e

(iv)  $\pi(\lambda \mid \theta, \mu, N', D) \propto \lambda^{a_2-1} e^{-b_2\lambda} \Psi_2(\mu, \lambda)$

$$\text{onde } \Psi_2(\mu, \lambda) = \exp\left\{N' \log(1-F(x_n)) + \frac{n}{2} \log(\lambda) - \frac{\lambda}{2\mu^2} \sum_{i=1}^n \frac{(x_i - \mu)^2}{x_i}\right\}.$$

Observar que devemos utilizar o algoritmo Metropolis-Hastings para gerar  $\mu$  e  $\lambda$ .

Outra possibilidade é considerar a função de valor médio  $m(t)$  dada em (1) com  $F(t) = \Phi\left(\frac{\ln t - \mu}{\sigma}\right)$  (função distribuição acumulada de uma distribuição log-normal). Neste caso, a função de verossimilhança para  $\theta$ ,  $\mu$  e  $\sigma$  é dada [de (3)] por,

$$\begin{aligned} L_{lg}(\theta, \mu, \sigma \mid D_t) &= \frac{\theta^n}{(2\pi)^{n/2} \sigma^n} \left\{ \prod_{i=1}^n 1/x_i \right\} \\ &\exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^n (\ln x_i - \mu)^2 - \theta \Phi\left(\frac{\ln t - \mu}{\sigma}\right)\right\}. \end{aligned} \quad (15)$$

Supor as seguintes densidades *a priori* para os parâmetros  $N'$ ,  $\theta$ ,  $\mu$  e  $\sigma$ :

$$(i) N' \sim P \left[ \theta \left( 1 - \Phi \left( \frac{\ln t - \mu}{\sigma} \right) \right) \right],$$

$$(ii) \theta \sim \Gamma[a_4, b_4], \text{ onde } a_4 \text{ e } b_4 \text{ são conhecidos;}$$

$$(iii) \mu \sim \pi_3(\mu), \text{ onde } \pi_3(\mu) \text{ é a densidade } a \text{ priori para } \mu; \text{ e} \quad (16)$$

$$(iv) \sigma \sim \pi_4(\sigma), \text{ onde } \pi_4(\sigma) \text{ é uma densidade } a \text{ priori para } \sigma.$$

Assumindo independência *a priori* para os parâmetros, a densidade *a posteriori* conjunta é dada por,

$$\begin{aligned} \pi(N', \theta, \mu, \sigma \mid D_t) &\propto \frac{1}{\sigma^n N'!} \theta^{n+N'+a_4-1} e^{-\theta(b_4+1)} \\ &\times \left\{ 1 - \Phi \left( \frac{\ln t - \mu}{\sigma} \right) \right\}^{N'} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n (\ln x_i - \mu)^2 \right\} \pi_3(\mu) \pi_4(\sigma). \end{aligned} \quad (17)$$

As densidades condicionais de  $N'$ ,  $\theta$ ,  $\mu$  e  $\sigma$  são dadas por,

$$(i) N' \mid \theta, \mu, \sigma, D_t \sim P \left[ \theta \left( 1 - \Phi \left( \frac{\ln t - \mu}{\sigma} \right) \right) \right];$$

$$(ii) \theta \mid N', \mu, \sigma, D_t \sim \Gamma[a+n+N', b+1]; \quad (18)$$

$$\begin{aligned} (iii) \mu \mid N', \theta, \sigma, D_t &\propto \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n (\ln x_i - \mu)^2 \right\} \\ &\times \left\{ 1 - \Phi \left( \frac{\ln t - \mu}{\sigma} \right) \right\}^{N'} \pi_3(\mu); \end{aligned}$$

e

$$\begin{aligned} (iv) \sigma \mid N', \mu, \theta, D_t &\propto \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n (\ln x_i - \mu)^2 \right\} \\ &\times \sigma^{-n} \left\{ 1 - \Phi \left( \frac{\ln t - \mu}{\sigma} \right) \right\}^{N'} \pi_4(\sigma). \end{aligned}$$

#### 4. DISCRIMINAÇÃO DE MODELOS DE CONFIABILIDADE DE SOFTWARE

Uma questão que surge é a seguinte “*Dentre os modelos que utilizam processos de Poisson, trata-se de saber qual é aquele que melhor se ajusta a um determinado conjunto de dados de confiabilidade de software?*”

Para determinarmos o melhor modelo, dois fatores devem ser considerados (Gelfand, Dey e Chang, 1992): adequabilidade e seleção do modelo.

Sob o enfoque Bayesiano, procura-se, neste artigo, responder a esta questão através dos seguintes critérios:

(i) Soma dos Erros Relativos;

(ii) Análise freqüencial de densidades preditivas condicionais ordenadas, denotada por PCPO (em inglês “*prequential conditional predictive ordinates*”) (Yang, 1994).

A técnica PCPO define o CPO<sup>1</sup>, de  $t_i$  condicionado aos dados passados dos tempos de falhas  $t_1, \dots, t_{i-1}$  conjuntamente. Essa técnica difere do CPO que utiliza os dados *cross-validated* onde o condicionamento é de todos os dados passados e futuros, exceto  $t_i$ . A análise freqüencial é considerada, pois o teste do software é usualmente conduzido em um tempo real. Para cada instante de falha, temos interesse em prever o instante da próxima falha. A regra CPO providencia uma taxa de explicação dos dados para o modelo em estudo.

Gelfand, Dey e Chang (1992) fizeram uma revisão da técnica CPO dos dados *cross-validated* na discriminação de modelos.

Dawid (1984) propõe, para a seleção dos modelos de confiabilidade de software, o critério da verossimilhança freqüencial. Singpurwalla e Soyer (1992) também utilizam a análise freqüencial para a comparação dos modelos onde a confiabilidade crescente de um software é modelada por um processo auto-regressivo de coeficiente aleatório.

Yang (1994) e Achcar, Dey e Niverthy (1996) consideram a técnica PCPO para discriminação dos modelos de confiabilidade de software que utilizam os processos de Poisson não-homogêneos.

Para verificação da adequabilidade do modelo, Littlewood, Abdel-Ghaly e Chan (1986) propõem o u-plot; ordenando os valores  $u_j = P(T_i < t_j)$  e representando graficamente os pontos  $(u_j, j/c)$ , onde  $c$  é o número total de falhas no software, observadas no intervalo  $(0, t]$ .

Algum vício na predição será destacado pelo desvio da linha unitária com inclinação na origem.

<sup>1</sup> CPO é a densidade preditiva condicional ordenada; para o tempo de falha futuro  $t_{i+1}$  é definido como

$$c_{i+1} = p(t_{i+1} / t_1, t_2, \dots, t_{i-1}, t_i, t_{i+2}, \dots, t_n).$$

Portanto, além dos procedimentos analíticos de discriminação pode-se também ter uma boa idéia do melhor modelo com processos de Poisson não-homogêneos, através das suposições básicas dos modelos de estatísticas de ordem PPNH-I: O número de erros  $N$  no início do teste do software tem distribuição de Poisson com média  $\theta$ . Então, a função de valor médio é dada por  $m(t) = \theta F(t)$ . Se essas suposições estiverem corretas, então  $m(t)/\theta = F(t)$  tem uma distribuição uniforme padronizada.

Assim, para cada tempo de falha, podemos considerar Q-Q plots empíricos de estimadores Monte Carlo de  $m(t)/\theta$  versus uma distribuição Uniforme (0,1) para cada tempo de falha (ver Achcar, Dey e Niverthy, 1996). Um distanciamento da distribuição uniforme indica inadequação do modelo.

É importante ressaltarmos que várias outras estratégias Bayesianas ou clássicas, propostas na literatura, podem ser utilizadas na discriminação dos modelos de confiabilidade de software. Além das estratégias Bayesianas de discriminação citadas anteriormente, acrescenta-se o Fator de Bayes (Jeffreys, 1939), Pseudo Fator de Bayes (Geisser e Eddy, 1979), etc.

Para os modelos PPNH-I, podemos considerar a técnica da verossimilhança preditiva condicional ordenada, denotada por PCPO. A técnica PCPO define o CPO de  $t_i$  condicionado aos dados passados dos tempos de falhas  $t_1, \dots, t_{i-1}$  conjuntamente.

A análise freqüencial é considerada, porque o teste do software é usualmente conduzido em um tempo real.

Nesta seção os dados são definidos como os instantes de falha do software, denotados por  $x_1, \dots, x_n$ . O conjunto de dados é escrito por  $D_{x_i} = \{x_1, \dots, x_i\}$ .

Admitindo-se os modelos de estatísticas de ordem denotados por PPNH-I, a verossimilhança da densidade preditiva condicional ordenada para uma observação futura  $x_{i+1}$  é definida por  $c_i = p(x_{i+1} / D_{x_i})$ , ou seja, é a densidade condicional estimada de  $X_{i+1}$  para o tempo futuro observado de falha  $x_{i+1}$ , condicionado aos dados  $\{x_1, \dots, x_i\}$ .

A seqüência  $\{X_i\}_{i \geq 1}$  é uma cadeia de Markov (ver Yang, 1994, pág. 52). Portanto,  $c_i = p(x_{i+1} \setminus D_{x_i})$  pode ser escrito como  $c_i = p(x_{i+1} \setminus x_i)$ .

Assim, a densidade preditiva condicional ordenada pode ser calculada por

$$\begin{aligned} p(x_{i+1} \setminus x_i) &= \int p(x_{i+1} \setminus \beta, D_{x_i}) p(\beta, D_{x_i}) d\beta \\ &= \int \lambda_{II}(x_{i+1}) \exp\{-m_{II}(x_{i+1}) + m_{II}(x_i)\} p(\beta, D_{x_i}) d\beta \end{aligned} \quad (19)$$

onde  $m_{II}(t) = -\ln(1-F(t))$  é a função de valor médio dos PPNH-II, definida em (2) e  $\lambda_{II}(t)$  é a função intensidade dada por  $\lambda_{II}(t) = \frac{dm_{II}(t)}{dt} = \frac{f(t)}{1-F(t)}$ .

Considerando-se o modelo de estatísticas de ordem gama generalizada com  $F(t)=I_k(\beta t^\alpha)$ , temos a função de valor médio (2) dada por,

$$m_{II}(t) = -\ln(1-I_k(\beta t^\alpha)) \text{ e } \lambda_{II}(t) = \alpha\beta^k t^{\alpha k-1} e^{-\beta t^\alpha} / \{\Gamma(k)[1-I_k(\beta t^\alpha)]\}.$$

Para o modelo de estatísticas de ordem gama generalizada, o PCPO (19) é dado por,

$$p(x_{i+1} \setminus x_i) = \iiint \frac{\alpha\beta^k x_{i+1}^{\alpha k-1} e^{-\beta x_{i+1}^\alpha}}{\Gamma(k)\{1-I_k(\beta x_i^\alpha)\}} \pi(\alpha, \beta, k \setminus D_{x_i}) d\alpha d\beta dk. \quad (20)$$

Utilizando-se o amostrador de Gibbs e Metropolis-Hastings (20) pode ser aproximado por um estimador Monte Carlo dado por:

$$\hat{p}(x_{i+1} \setminus x_i) = \frac{2}{RS} \sum_{s=1}^S \sum_{r=\frac{R}{2}+1}^R \frac{\alpha^{(r,s)k(r,s)} \beta^{(r,s)} x_{i+1}^{\alpha(r,s)k(r,s)-1} e^{-\beta^{(r,s)} x_{i+1}^{\alpha(r,s)}}}{\Gamma(k^{(r,s)}) \{1-I_k(\beta^{(r,s)} x_i^{\alpha(r,s)})\}}.$$

O PCPO dos modelos de estatísticas de ordem exponencial, gama e Weibull são facilmente obtidos de (20); por exemplo, considerando-se  $k=1$  em (20), temos o PCPO dos modelos de estatísticas de ordem Weibull dado por,

$$p(x_{i+1} \setminus x_i) = \iint \alpha\beta x_{i+1}^{\alpha-1} \exp\{-\beta x_{i+1}^\alpha + \beta x_i^\alpha\} \pi(\alpha, \beta \setminus D_{x_i}) d\alpha d\beta. \quad (21)$$

O PCPO (21) pode ser aproximado por um estimador Monte Carlo, utilizando-se as amostras geradas pelo amostrador de Gibbs e Metropolis-Hastings.

Considerando, agora, o modelo de estatísticas de ordem log-normal, a função de valor médio dos PPNH-II (2) é dada por  $m_{II}(t) = -\ln\left\{1 - \Phi\left(\frac{\ln t - \mu}{\sigma}\right)\right\}$  e a função intensidade

$$\lambda_{II}(t) = m'_{II}(t) = \frac{\exp\left\{-\frac{(\ln t - \mu)^2}{2\sigma^2}\right\}}{\sqrt{2\pi\sigma} \left\{1 - \Phi\left(\frac{\ln t - \mu}{\sigma}\right)\right\}}.$$

O PCPO (19), para o instante de falha futura  $x_{i+1}$ , dos modelos de estatísticas de ordem log-normal, é dado por,

$$p(x_{i+1} | x_i) = \iint \frac{\exp\{-\ln(x_{i+1} - \mu)^2 / 2\sigma^2\}}{\sqrt{2\pi\sigma x_{i+1}} \left\{1 - \Phi\left(\frac{\ln x_{i+1} - \mu}{\sigma}\right)\right\}} p(\mu, \sigma | D_{x_i}) d\mu d\sigma. \quad (22)$$

Admitindo-se o modelo de estatísticas de ordem gaussiana inversa, a função de valor médio (2) é dada por,

$$m_{ii}(t) = -\ln\left\{1 - \left[\Phi\left[\sqrt{\frac{\lambda}{t}}\left(\frac{t}{\mu} - 1\right)\right] + e^{2\lambda/\mu}\Phi\left[-\sqrt{\frac{\lambda}{t}}\left(1 + \frac{t}{\mu}\right)\right]\right]\right\} \text{ e a função}$$

$$\text{intensidade } \lambda_{ii}(t) = \frac{\frac{\sqrt{\lambda}}{\sqrt{2\pi}} t^{-3/2} \exp\left\{-\frac{\lambda(t-\mu)^2}{2\mu^2 t}\right\}}{1 - \left[\Phi\left[\sqrt{\frac{\lambda}{t}}\left(\frac{t}{\mu} - 1\right)\right] + e^{2\lambda/\mu}\Phi\left[-\sqrt{\frac{\lambda}{t}}\left(1 + \frac{t}{\mu}\right)\right]\right]}.$$

O PCPO (19) para o modelo de estatísticas de ordem gaussiana inversa é dado por,

$$p(x_{i+1} / x_i) = \iint \frac{\frac{\sqrt{\lambda}}{\sqrt{2\pi}} x_{i+1}^{-3/2} \exp\left\{-\frac{\lambda(x_{i+1} - \mu)^2}{2\mu^2 x_{i+1}}\right\}}{1 - \left[\Phi\left[\sqrt{\frac{\lambda}{x_n}}\left(\frac{x_n}{\mu} - 1\right)\right] + e^{2\lambda/\mu}\Phi\left[-\sqrt{\frac{\lambda}{x_n}}\left(1 + \frac{x_n}{\mu}\right)\right]\right]} \pi(\mu, \lambda | D_{x_i}) d\lambda d\mu$$

Os PCPOs dos modelos de estatísticas de ordem log-normal e gaussiana inversa podem ser aproximados por um estimador Monte Carlo, utilizando-se o amostrador de Gibbs e Metropolis-Hastings.

Para seleção dos modelos, podemos "plotar"  $c_i$  versus  $i$ , para  $i=2, \dots, n$ , para cada modelo. O maior dos  $c_i$  em média indica o melhor modelo.

O critério de verossimilhança preditivo Bayesiano seleciona o modelo, que maximiza a

verossimilhança  $c(l) = \prod_{i=1}^n c_i(l)$ , onde  $l$  indexa o  $l$ -ésimo modelo de confiabilidade. Neste caso,

( $l=1, \dots, 6$ ), ( $l=1$  modelo de estatísticas de ordem gama generalizada,  $l=2$  Weibull,  $l=3$  gama,  $l=4$  exponencial,  $l=5$  log-normal,  $l=6$  gaussiana inversa).

Neste trabalho, selecionamos também o modelo que minimiza a soma dos erros relativos, denotada por RE. A soma dos erros relativos para os modelos que utilizam processos de Poisson não-homogêneos é dada pela expressão,

$$RE = \frac{\sum_{i=1}^n (n_i - \hat{m}(x_i))^2}{\hat{m}(x_i)}, \quad (23)$$

onde  $n_i$  é o número de erros no software observados no  $i$ -ésimo intervalo de tempo e  $\hat{m}(x_i)$  é um estimador Bayesiano da função de valor médio.

Por exemplo, considere-se o modelo de estatísticas de ordem exponencial, com função de valor médio dada por,

$$m(t) = \theta(1 - e^{-\beta t}). \quad (24)$$

Considerando as amostras geradas pelo amostrador de Gibbs e Metropolis-Hastings, aproximamos a função de valor médio (24) por um estimador Monte Carlo dado por,

$$\hat{m}(x_i) = \frac{2}{RS} \sum_{r=1}^R \sum_{s=\frac{S}{2}+1}^S \theta^{(r,s)} (1 - e^{-\beta^{(r,s)} x_i}). \quad (25)$$

Para encontrar a soma dos erros relativos do modelo de estatísticas de ordem exponencial, considera-se (25) em (23).

Para calcular a soma dos erros relativos RE (23) para os modelos de estatísticas de ordem Weibull, gama, gama generalizada, log-normal e gaussiana inversa, aproximamos as funções de valor médio desses modelos, por um estimador Monte Carlo e, substituímos esses estimadores, em (23) para cada modelo.

## 5. UM EXEMPLO

Neste exemplo, vamos considerar o conjunto de dados da Tabela 1, os quais foram introduzidos por Jelinski e Moranda (1972), consistindo no número de dias entre as 26 falhas que ocorreram durante a fase de produção do software NTDS (em inglês *Naval Tactical Data System*).

Na Tabela 1, observamos que  $n=26$ ,  $\sum_{i=1}^n \ln x_i = 112.4776$  e  $x_n = x_{26} = 250$ .

Tabela 1-Número de dias entre as 26 falhas que ocorreram na fase de teste do software NTDS ( $t_i = X_i - X_{i-1}$  refere-se ao tempo entre falha)

i	$t_i$	$x_i$	i	$t_i$	$x_i$	i	$t_i$	$x_i$
1	9	9	11	1	71	21	11	116
2	12	21	12	6	77	22	33	149
3	11	32	13	1	78	23	7	156
4	4	36	14	9	87	24	91	247
5	7	43	15	4	91	25	2	249
6	2	45	16	1	92	26	1	250
7	5	50	17	3	95			
8	8	58	18	3	98			
9	5	63	19	6	104			
10	7	70	20	1	105			

Assumindo o modelo de estatísticas de ordem gama generalizada na modelagem dos dados da Tabela 1, e o modelo de falha truncado com  $t$  substituído por  $x_{26} = 250$ , consideramos de (6) as densidades *a priori*  $N' \sim P\{\theta[1 - I_k(\beta(250)^\alpha)]\}$ ,  $\theta \sim \Gamma(60,2)$ ,  $\beta \sim \Gamma(10,2)$  e densidades *a priori* não informativas para  $\alpha$  e  $k$  dadas por  $\pi_1(\alpha) \propto 1/\alpha, \alpha > 0$  e  $\pi_2(k) \propto 1/k, k > 0$ . Considerando as densidades *a posteriori* marginais para  $N', \theta, \beta, \alpha$  e  $k$  descritas em (8), geramos 10 cadeias de Gibbs, com 2000 iterações cada, e monitoramos a convergência do amostrador de Gibbs, utilizando o critério de Gelman e Rubin (1992) que admite a técnica de análise de variância para determinar o número de iterações necessárias para obtenção da convergência.

Em cada seqüência, desprezamos os primeiros 500 pontos gerados e selecionamos 100 pontos tomados de 15 em 15, obtendo-se uma amostra final de 1000 pontos amostrais. A Tabela 2 mostra os sumários *a posteriori* para os parâmetros  $N', \theta, \beta, \alpha$  e  $k$ . É interessante se observar que os estimadores de máxima verossimilhança para  $N', \theta, \beta, \alpha$  e  $k$  obtidos utilizando-se o software SAS são dados por:  $\hat{\theta} = 29.4301$ ,  $\hat{k} = 18.2010$ ,  $\hat{\alpha} = 0.2748$  e  $\hat{\beta} = 5.1279$ .

Usualmente, engenheiros de software consideram casos especiais do supermodelo gama generalizada para analisar os dados de confiabilidade de software.

Considerando  $k=1$  ( processo de Goel, 1983) e assumindo um modelo de falha truncado, admitimos as densidades *a priori*  $N' \sim P(\theta e^{-(250)^\alpha \beta})$ ,  $\theta \sim \Gamma(90,3)$ ,  $\beta \sim \Gamma(26,160)$  e uma densidade *a priori* não informativa para  $\pi_1(\alpha) \propto 1/\alpha, \alpha > 0$ . Na Tabela 3, temos os sumários *a posteriori* para os parâmetros  $N', \theta, \beta$  e  $\alpha$ , utilizando-se 1 000 amostras obtidas pelas 10 cadeias de Gibbs com 1 200 iterações cada.

Quando  $\alpha=1$  (modelo de estatísticas de ordem) e assumindo as densidades *a priori*  $N' \sim P(\theta(1 - I_k(250\beta)))$ ,  $\theta \sim \Gamma(90,3)$ ,  $\beta \sim \Gamma(12,680)$  e uma densidade *a priori* não informativa para  $k$  proporcional a  $1/k, k > 0$ , obtemos os sumários *a posteriori*, considerando 10 cadeias de Gibbs com 1 200 iterações cada das densidades condicionais.

Considerando-se  $k=1$  e  $\alpha=1$  (processo de Goel e Okumoto, 1979) e assumindo as densidades *a priori*  $N' \sim P(\theta e^{-(250)\beta})$ ,  $\theta \sim \Gamma(60,2)$ ,  $\beta \sim \Gamma(5,1000)$ , com  $\theta$  independente de  $\beta$ , temos, na Tabela 5, sumários *a posteriori* para os parâmetros  $N'$ ,  $\theta$  e  $\beta$  considerando as amostras de Gibbs  $S=1\ 000$  obtidas das 10 cadeias com 1 200 iterações cada das densidades *a posteriori* condicionais.

Na Tabela 6, mostramos os sumários *a posteriori* para os parâmetros  $N'$ ,  $\theta$ ,  $\mu$  e  $\sigma$  do modelo de estatísticas de ordem log-normal. Consideramos  $S=1\ 000$  amostras obtidas do algoritmo de Gibbs com Metropolis gerados utilizando as densidades condicionais descritas em (18) e assumindo as densidades *a priori*  $N' \sim P[\theta(1-\Phi(\frac{\ln 250 - \mu}{\sigma}))]$ ,  $\theta \sim \Gamma(36,1.2)$ , e densidades *a priori* não informativas para  $\mu$  e  $\sigma$  dadas por  $\pi_3(\mu) \propto \text{constante}$ ,  $-\infty < \mu < \infty$  e  $\pi_4(\sigma) \propto 1/\sigma$ ,  $\sigma > 0$ .

Os sumários da distribuição gaussiana inversa, apresentados na Tabela 7, foram obtidos considerando-se as densidades condicionais definidas em (14) e 10 cadeias de Gibbs com 1 200 iterações cada. Em cada seqüência, desprezamos os 200 valores iniciais e selecionamos 100 pontos tomados de 10 em 10, obtendo assim uma amostra final de 1 000 pontos. As densidades *a posteriori* marginais aproximadas para os parâmetros  $N'$ ,  $\theta$ ,  $\mu$  e  $\lambda$  pelo amostrador de Gibbs e Metropolis são apresentadas na Figura 1.

Tabela 2 - Sumário *a posteriori* do modelo de estatísticas de ordem gama generalizada

Parâmetros	Média	Mediana	Variância	Intervalo de Credibilidade Bayesiano 95%
$N'$	2.4511	2.0000	1.0254	(1; 4)
$\theta$	29.996	30.010	12.1848	(24.0194; 38.1233)
$\alpha$	0.3061	0.3084	0.00390	(0.2101; 0.4316)
$\beta$	4.0125	3.9879	1.2455	(2.9874; 6.1257)
$k$	16.976	16.9201	16.4042	(9.1443; 25.0124)

Tabela 3 - Sumário *a posteriori* do modelo de estatísticas de ordem Weibul

Parâmetros	Média	Mediana	Variância	Intervalo de Credibilidade Bayesiano 95%
$N'$	6.949	6.0000	9.879	(1; 16)
$\theta$	30.854	30.684	5.2542	(25.1755; 37.2316)
$\alpha$	0.3061	0.3084	0.00390	(0.0.3573; 0.6107)
$\beta$	0.11148	0.11010	0.04497	(0.0724; 0.1609)

Tabela 4 - Sumário *a posteriori* do modelo de estatísticas de ordem gama

Parâmetros	Média	Mediana	Variância	Intervalo de Credibilidade Bayesiano 95%
$N'$	2.1310	2.000	4.1360	(0; 8)
$\theta$	29.4676	29.3576	7.914500	(24.2259; 35.5030)
$k$	1.8750	1.8364	0.1491	(1.9997; 2.7292)
$\beta$	0.0171	0.0168	0.00184	(0.0098; 0.0266)

Tabela 5 - Sumário *a posteriori* do modelo de Goel e Okumoto (1979)

Parâmetros	Média	Mediana	Variância	Intervalo de Credibilidade Bayesiano 95%
$N'$	7.724	7.000	17.2070	(1; 17)
$\theta$	31.2174	31.0882	12.4283	(24.8528; 38.1791)
$\beta$	0.0059	0.0058	$2,428 \times 10^{-6}$	(0.0034; 0.0094)

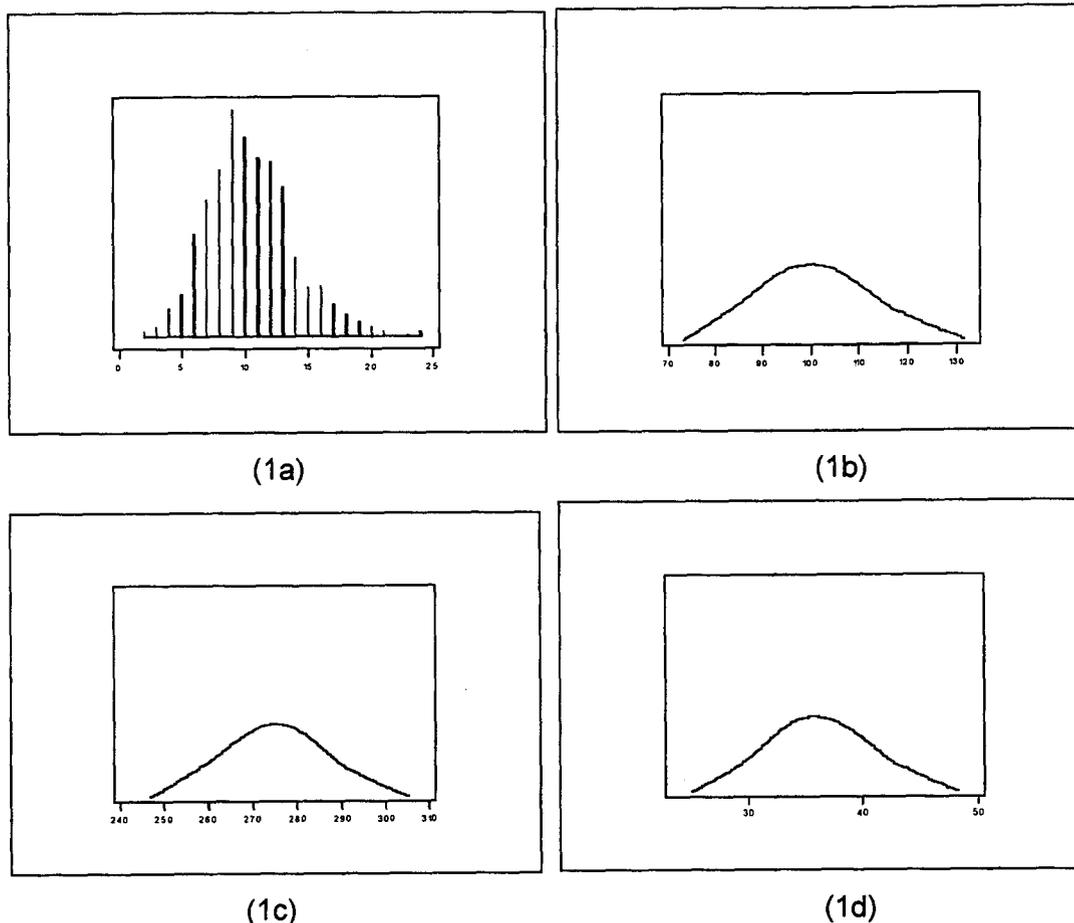
Tabela 6 - Sumário *a posteriori* do modelo de estatísticas de ordem log-norma

Parâmetros	Média	Mediana	Variância	Intervalo de Credibilidade Bayesiano 95%
$N'$	4.2001	2.000	24.812	(0; 21)
$\theta$	30.105	29.844	16.451	(22.511; 40.173)
$\mu$	4.590	4.580	0.1706	(4.13; 5.94)
$\sigma$	0.8541	0.7787	0.449	(0.567; 1.7012)

Tabela 7- Sumário *a posteriori* do modelo de estatísticas de ordem gaussiana inversa

Parâmetros	Média	Mediana	Variância	Intervalo de Credibilidade Bayesiano 95%
$N'$	10.4270	10	11.3821	(5.0, 18.0)
$\theta$	36.2004	36.1753	11.6491	(29.8630; 43.4157)
$\mu$	274.6688	274.6903	87.7225	(256.012; 293.322)
$\lambda$	100.1174	99.9262	84.8426	(82.971; 119.424)

Figura 1- Densidades *a posteriori* de  $N, \lambda, \mu$  e  $\theta$ , (1a), (1b), (1c) e (1d), respectivamente obtidas pelo amostrador de Gibbs e Metropolis-Hastings.



## 6. DISCRIMINAÇÃO BAYESIANA DOS MODELOS DE ESTATÍSTICAS DE ORDEM

A seguir, conduziremos a discriminação Bayesiana dos modelos de estatísticas de ordem gama, gama-generalizada, log-normal, exponencial, Weibull e gaussiana inversa, utilizando as técnicas de seleção de modelos apresentadas na seção 4.

Na Tabela 8, temos os estimadores Monte Carlo das funções de valor médio dos modelos de estatísticas de ordem gaussiana inversa, gama generalizada, exponencial Weibull, gama e log-normal com respeito à função de perda quadrática, utilizando os dados da Tabela 1 e as amostras geradas pelo amostrador de Gibbs e Metropolis-Hastings.

Tabela 8 - Estimadores Monte Carlo das funções de valor médio  $m(x_i)$  dos modelos de estatísticas de ordem gama generalizada, Weibull, gama, exponencial, log-normal e gaussiana inversa

i	$x_i$	$n_i$	Gama generalizada	Weibull	Gama	Exponencial	Log-normal	Gaussiana inversa
1	9	1	0.3847	8.4412	0.5877	1.6011	0.1416	0.0442
2	21	2	2.0812	11.7309	2.1222	3.4021	1.1853	1.4955
3	32	3	4.3854	13.6602	3.9554	5.3110	3.0754	3.9496
4	36	4	5.3102	14.2111	4.6694	5.8544	3.8987	4.8902
5	43	5	6.8978	15.1044	5.9728	6.8610	5.4101	6.4974
6	45	6	7.3999	15.3388	6.3544	7.1546	5.8612	6.9412
7	50	7	8.6423	15.8751	7.3012	7.8012	6.8777	8.0144
8	58	8	10.4689	16.6410	8.8199	8.8614	8.8012	9.6159
9	63	9	11.5575	17.0676	9.7751	9.4755	9.9122	10.5444
10	70	10	13.0123	17.6201	11.064	10.3319	11.255	11.7552
11	71	11	13.2117	17.6944	11.243	10.4421	11.4512	11.9201
12	77	12	14.3755	18.1321	12.441	11.1318	12.5914	12.8696
13	78	13	14.5507	18.1914	12.495	11.2449	12.787	13.0214
14	87	14	16.1164	18.7677	14.005	12.2513	14.3522	14.3118
15	91	15	16.7584	18.9998	14.559	12.7010	14.9978	14.9442
16	92	16	16.9130	19.0541	14.883	12.7431	15.1311	14.9737
17	95	17	17.3422	19.2331	15.335	13.0947	15.6012	15.3535
18	98	18	17.8001	19.4011	15.778	13.4119	16.0411	15.7212
19	104	19	18.6199	19.7107	16.744	13.8754	16.8841	16.4220
20	105	20	18.7422	19.7644	16.754	14.0844	17.0079	16.5346
21	116	21	20.0765	22.2999	18.288	15.1004	18.3666	17.6994
22	149	22	22.9722	21.5820	21.956	17.8999	21.4785	20.5412
23	156	23	23.4415	21.8456	22.601	18.2606	21.9803	21.0447
24	247	24	26.7045	24.0511	27.433	23.2911	25.8897	25.6768
25	249	25	26.8011	24.0877	27.511	23.4143	25.9214	25.7507
26	250	26	26.7644	24.1055	27.601	23.3999	25.9561	25.7873

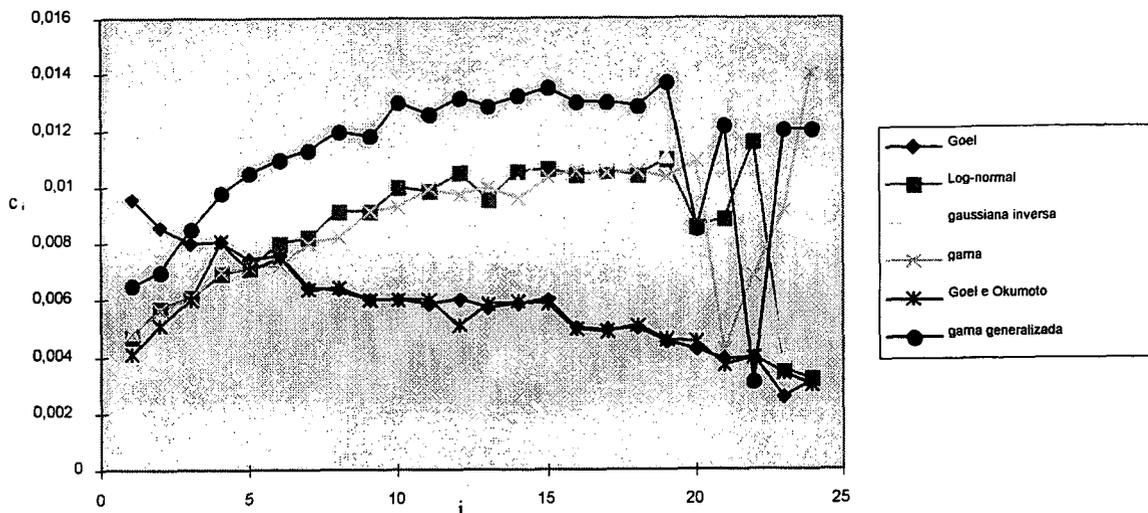
Podemos observar que as estimativas Bayesianas da função de valor médio para os modelos de estatísticas de ordem gama, gama-generalizada, log-normal, exponencial e gaussiana inversa são próximas. Notamos uma razoável discrepância das estimativas Bayesianas  $\hat{m}(x_i)$  para o modelo de estatísticas de ordem Weibull comparado aos demais modelos.

Na Tabela 9, temos os critérios de seleção, definidos na seção 4, para discriminar os modelos em estudos: a soma dos erros relativos e os PCPOs, definidos por  $c(l) = \prod_{i=1}^n c_i$ , onde  $c_i = p(x_{i+1}/D_{x_i})$  dos modelos de estatísticas de ordem Weibull, gama, gama generalizada, exponencial, log-normal e gaussiana inversa.

Tabela 9 - Soma dos erros relativos e  $C(l)$  dos modelos de estatísticas de ordem gama generalizada, gama, exponencial, Weibull, log-normal e gaussiana inversa

Modelos	$C(l)$	RE
Modelo de estatísticas de ordem gama genera	$9.56 \times 10^{-50}$	6.8110
Modelo de estatísticas de ordem gama	$1.33 \times 10^{-51}$	24.112
Modelo de estatísticas de ordem exponencial	$3.61 \times 10^{-57}$	225.414
Modelo de estatísticas de ordem Weibull	$9.91 \times 10^{-58}$	67.45
Modelo de estatísticas de ordem log-normal	$1.130 \times 10^{-51}$	8.1012
Modelo de estatísticas de ordem gaussiana inversa	$1.011 \times 10^{-51}$	25.467

Figura 2 - Valores de  $c_i$  versus  $i$  para os modelos de estatísticas de ordem gama, gama-generalizada, Weibull, log-normal, exponencial e gaussiana inversa



## 7. ADEQUABILIDADE DO MODELO

Sabemos que a construção dos modelos de estatísticas de ordem, denotados de PPNH-I, está baseada nas seguintes suposições:

- O número de erros  $N$  no início do teste do software tem distribuição de Poisson;
- a média do número de erros  $N$  é  $\theta$ .

Então, a função de valor médio é dada por  $m(x_i) = \theta F(x_i | \beta)$ . Se essas suposições estiverem corretas é fácil verificar que  $m(x_i) / \theta = F(x_i | \beta)$  tem uma distribuição Uniforme (0,1).

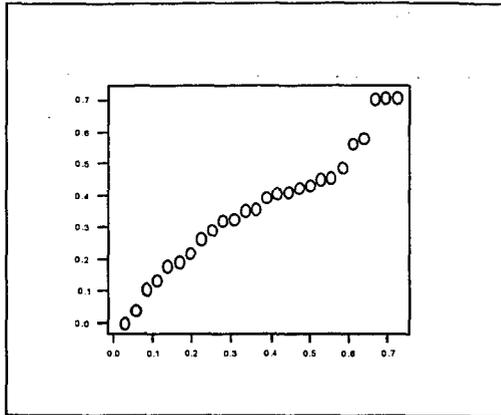
Assim para verificar se o modelo é adequado para os dados, ou seja, se a suposição é válida, fazemos um estudo para cada tempo de falha, construindo Q-Q plots empíricos dos estimadores Monte Carlo de  $m(x_i) / \theta$  *versus* Uniforme (0,1) para cada modelo.

A título de decidir quais dos seis modelos analisados melhor explicam o fenômeno de confiabilidade crescente do software em estudo, podemos concluir, através das técnicas utilizadas, que, para este conjunto de dados, o modelo de estatísticas de ordem gama generalizada mostrou-se superior aos demais principalmente se comparado com os modelos especificados a partir de 2 parâmetros (modelos: Weibull e exponencial).

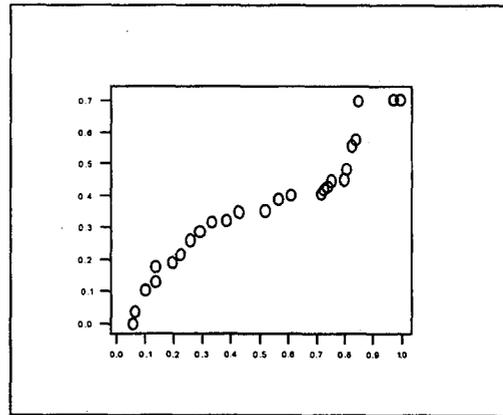
Fazendo a comparação entre modelos de mesma dimensão, concluímos que o modelo de estatísticas de ordem gaussiana inversa é superior ao modelo estatísticas de ordem log-normal. Quanto aos modelos especificados a partir de 2 parâmetros, podemos afirmar, através das técnicas utilizadas, que o modelo de estatísticas de ordem gama é melhor que os modelos de estatísticas de ordem Weibull e exponencial, para esse conjunto de dados [Tabela 1].

A adequação dos modelos é verificada na Figura 3. Observamos uma razoável linearidade dos pontos indicando uma boa adequabilidade dos modelos analisados para esse conjunto de dados.

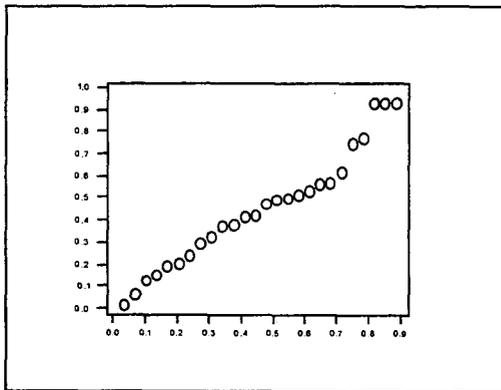
Figura 3 - Q-Q plots empíricos para  $m(X_j)/\theta$  versus a distribuição uniforme para os seis modelos de estatística de ordem



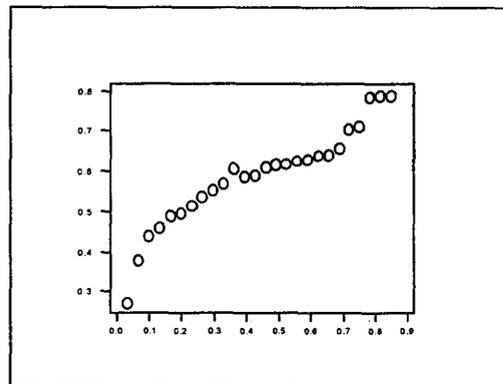
(gaussiana inversa)



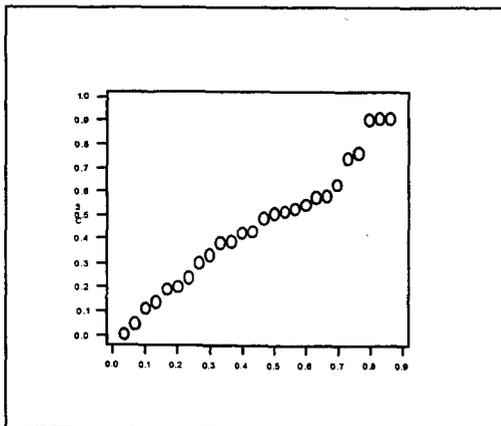
(Goel e Okumoto)



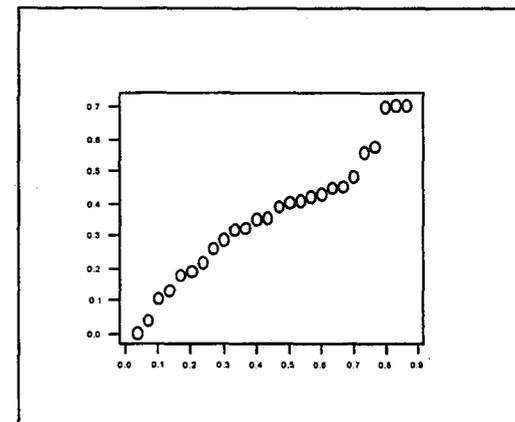
(gama)



(Goel)



(gama generalizada)



(log-normal)

## 8. CONCLUSÕES

O uso de técnicas Bayesianas para analisar dados de confiabilidade de software torna-se viável a partir de algoritmos de simulação como o amostrador de Gibbs e Metropolis-Hastings. O uso de métodos clássicos assintóticos para modelos de confiabilidade de software em processos de Poisson não-homogêneos, em geral, não é de simples obtenção e a precisão pode ser pequena.

Além disso, é possível a incorporação de opinião de especialistas em software na forma de distribuição *a priori*, o que torna a análise Bayesiana muito realista. Também, o procedimento Bayesiano apresenta técnicas para discriminar diferentes modelos de simples implementação a partir das amostras geradas pelo amostrador de Gibbs.

## REFERÊNCIAS

- ACHCAR, J.A., DEY, D. e NIVERTHY, M. (1996). "A Bayesian Approach Using nonhomogeneous Poisson Process for Software Reliability Models", Technical Report 96-07, a aparecer em *Frontiers of Reliability Theory* (ed. A. Basu).
- ACHCAR, J.A., STORANI, K. (1996). "Nonhomogeneous Poisson Processes Assuming a Inverse Gaussian Order Statistics Model for Software Reliability Data: a Bayesian Approach", *Notas do ICMS- USP*, N° 35.
- BOX, G.E.P.; TIAO, G.C (1973). "*Bayesian Inference in Statistical Analysis*". Addison-Wesley: Reading, Mass.
- CHIB, S. e GREENBERG, E. (1995). "Understanding the Metropolis-Hastings algorithm", *American Statistician*, 49, 4, 327-335.
- CHIKKARA, R.S.; FOLKS J.L. (1989) *The inverse Gaussian distribution*, New York: Marcel Dekker, Inc.
- COX, D.R. e LEWIS, P.A. (1966) *Statistical Analysis of the Series of Events*. London: Methuen.
- DAWID, A.P. (1984). "Statistics theory: the prequential approach". *Journal of the Royal Statistics Society, Ser. A*, 147, 278-292.
- DUANE, J.T. (1964). "Learning curve approach to reliability monitoring", *IEEE Transaction on Aerospace*, AS-2(2), 563-566.
- GEISSER, S.; EDDY, W.F. (1979). A Predictive Approach to Model Selection. *Journal American Stat. Assoc.*, 74, 153-160.
- GELFAND, A.E., DEY, D.K., e CHANG, H. (1992), "Model Determination Using Predictive Distribution with Implementation via Sampling-Basead Methods" (with discussion), *Bayesian Statistics 4*, eds. J. M. Bernardo, J.O.Berger, A. P.Dawid, e A.F.M.Smith, Oxford: Oxford University Press, 147-169.
- GELFAND, A.E.; SMITH, A.F.M. (1990). "Sampling Basead Approaches to Calculating Marginal Densities". *Journal of the Royal Statistics Society, Ser. A*, 147, 278-292.
- GELMAN, A. E., e RUBIN, D.B. (1992). "Inference from Iterative Simulation Using Multiple Sequence", *Statistical Science*, 7, 457-472.
- GOEL, A.L.; e OKUMOTO, K. (1979) "nTime-dependent error detection rate model for software reliability and other performance measures", *IEEE Transaction Reliability*, R-28, 206-211.
- GOEL, A.L. (1983) "A guidebook for software reliability assesment", Technical report, RADC- TR-83-176.
- JEFFREYS, H. (1939). *Theory of Probability*. Oxford University Press.
- KUO, L. e YANG, T.Y. (1995). "Bayesian Computation for Nonhomogeneous Poisson Process in (Software) Reliability". Technical Report N° .84-19, Statistics Departament, University of Connecticut.
- LAWLESS, J.F.(1982). *Statistical models and methods for lifetime data*, New York: John Wiley & Sons.

- LITTLEWOOD, B.; ABDEL-GHALY A.A. e CHAN, P.Y. (1986) "Evaluation of Competing Software Reliability Predictions," *IEEE Transactions on Software Engineering*, 12, 950-967.
- MAZZUCHI, T.A.; SINGPURWALLA, N.D. (1988) "Software reliability models", in Handbook of Statistical, vol.7, Eds Krishnaiah, P.r.; Rao, C.R., Elsevier Science Publishers B.V., 73-98.
- MAZZUCHI, T.A.; SOYER, R. (1988). "A Bayes empirical- Bayes model for software reliability", *IEEE Transactions on reliability*, R-37, 2, 248-254.
- MEINHOLD, R.J. e SINGPURWALLA, N.D. (1983). "Bayesian Analysis of a Comonly Used Model for Describing Software Failures", *The Statistician*, 32, 168-173.
- MUSA, J.D.; IANNINO, A.; OKUMOTO, K. (1987) *Software Reliability: Measurement, Prediction, Application*. New York:Wiley.
- MUSA, J.D.; OKUMOTO, K. (1984) "A Logarithm poisson execution time model for software reliability measurement". Proceedings of the 7th international conference on software engineering, Orlando, 230-237.
- SINGPURWALLA, N.D.; WILSON, S.P., (1994) "Software reliability modelling", *Intern. Statis Review*, 62, 3, 289-317.
- SMITH, A.F.M. e ROBERTS, G.O. (1992) Bayesian computation via Gibbs and related Markov chain Monte Carlo methods (with discussion). *J. Royal Statistical Society (Ser.B)* 55, 3-24.
- YAMADA, S.; OHBA, M. e OSAKI, S. (1983). "S-shaped reliability growth modeling for software error detection". *IEEE Transactions on Reliability*, 32, 475-478.
- YANG, T.Y. (1994) "Computational approaches to Bayesian inference for software reliability" P.h.D. Thesis, Department of Statistics, University of Connecticut, Storrs.

### ABSTRACT

In this paper, we use Bayesian methods to obtain inferences for software reliability models using nonhomogeneous Poisson processes. Different models are considered to model the mean value function, in special, the inverse Gaussian distribution. Posterior summaries of interest are obtained using Gibbs and Metropolis-Hastings sampling algorithms. We also propose a Bayesian model checking approach to discriminate the different models. The proposed methodology is illustrated considering a software reliability data set introduced by Jelinski and Moranda (1972).

# Alguns Modelos de Análise de Variância em Séries Temporais Utilizando Transformada de Fourier - Uma Aplicação

Chang Chiann \*

Clélia Maria de Castro Toloí \*

## RESUMO

Neste trabalho utilizamos a técnica da análise de variância para estudar as variações apresentadas em dados cujas medidas respostas são séries temporais. O nosso objetivo consiste na aplicação das técnicas convencionais à transformação de Fourier dos dados. Os modelos estudados incluem: o modelo com um sinal comum, com um fator e o modelo de planejamento. Serão considerados efeitos fixos e aleatórios. Uma aplicação a uma série real é dada.

## 1. INTRODUÇÃO

---

A motivação desse trabalho deve-se ao fato de que, muitas vezes, queremos utilizar uma análise de variância para estudar as variações apresentadas em dados cujas medidas respostas são séries temporais. A análise convencional não é adequada neste caso, uma vez que os dados são freqüentemente correlacionados, isto é, dependentes no tempo. O nosso objetivo é "adaptar" os dados, no sentido de amenizar a dependência entre eles e apresentar soluções para resolver nosso problema. Veremos que, uma vez feita a transformação dos dados, a análise utilizada será uma aplicação das técnicas convencionais da análise de variância.

---

\* Endereço para correspondência: Departamento de Estatística, IME-USP. C. Postal 66281, CEP. 05315-970, São Paulo.

Na seção 2 apresentamos uma parte do conjunto de dados que será analisado e o objetivo da análise. Na seção 3, baseado em Brillinger(1981), abordamos alguns conceitos fundamentais sobre a transformada de Fourier e periodograma. Na seção 4 apresentamos alguns modelos que são mais utilizados, incluindo efeitos fixo e aleatório. Finalmente, na seção 5, apresentamos a análise dos dados e na seção 6 as principais conclusões.

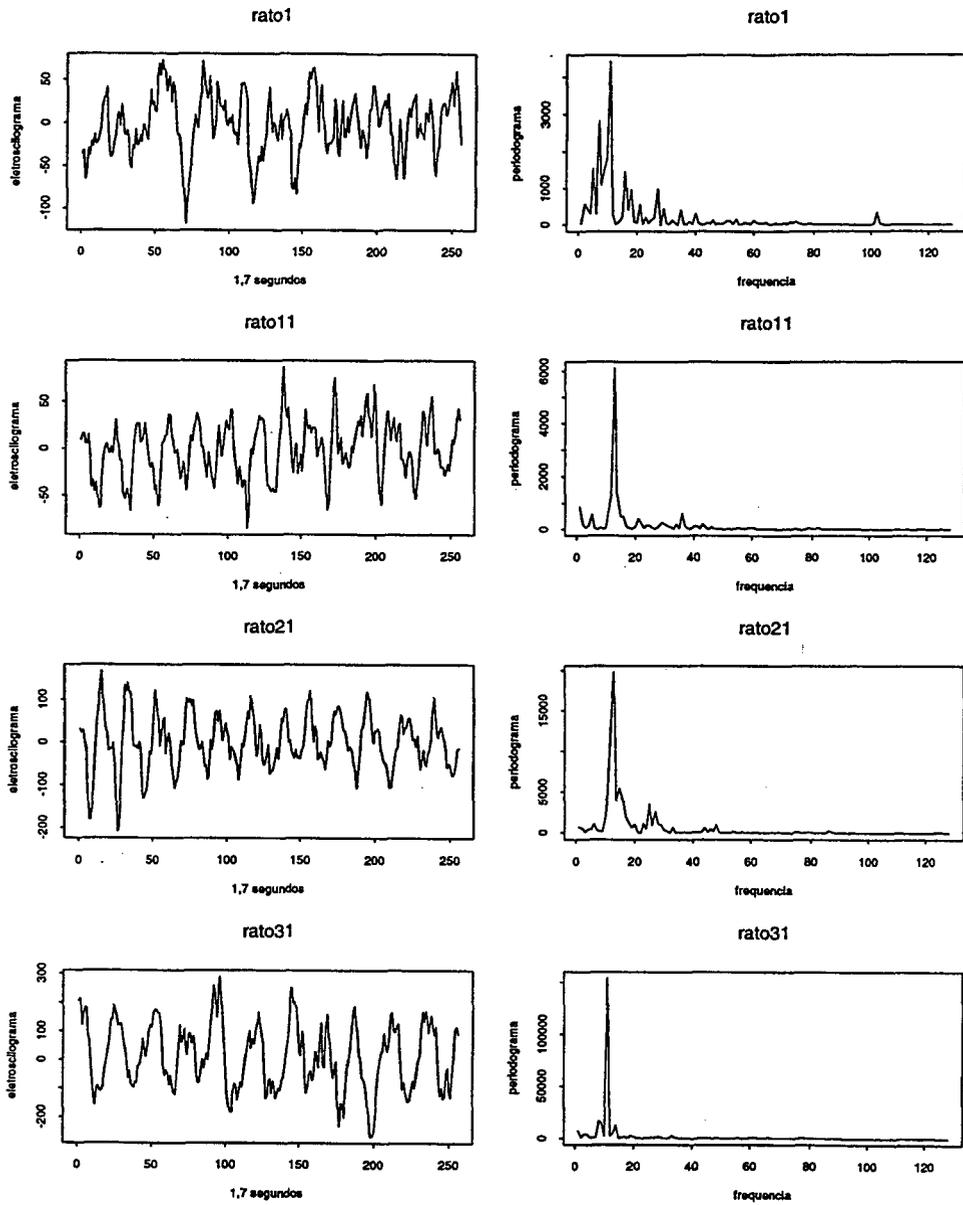
## **2. APRESENTAÇÃO DOS DADOS E OBJETIVO DA ANÁLISE**

---

As séries temporais analisadas neste trabalho referem-se aos eletroscilogramas de quarenta ratos normais de três a quatro meses de idade, de ambos os sexos, com pesos compreendidos entre duzentos e duzentos e cinquenta gramas. As observações foram coletadas, na Faculdade de Medicina - USP-, durante a atividade rítmica dos ratos, de acordo com um modelo de planejamento de experimento cruzado com dois fatores fixos: região do cérebro (posteriotálamo ou hipotálamo) e estado de sono (alerta ou sono paradoxal), que será apresentado na seção 5 com mais detalhes. Dez ratos são alocados em cada uma das combinações dos dois fatores.

Os dados coletados dos ratos 1, 11, 21 e 31 estão apresentados na Figura 1; cada série contém 256 pontos do eletroscilograma medidos em microvolts durante 1,7 segundo. Notamos que a maioria deles se distribuem na faixa de -200 a 200 microvolts, excetuando-se o rato 31 (do grupo hipotálamo-sono paradoxal) que apresenta alguns valores fora desse intervalo.

Figura 1 - Eletroscilogramas e periodogramas dos ratos 1, 11, 21 e 31.



O objetivo da análise será verificar a existência de efeito da região do cérebro e do estado de sono nas medidas dos eletroscilogramas.

A aplicação das técnicas convencionais de análise de variância não é adequada para o nosso conjunto de dados, pois as séries de eletroscilogramas de cada um dos ratos são altamente correlacionadas, o que viola a hipótese da independência entre as observações. Para amenizar essa dependência utilizaremos, inicialmente, uma ferramenta denominada transformada de Fourier (seção 3) e, em seguida, as técnicas da análise de variância adaptadas para séries temporais (seção 4). Voltaremos à análise dos dados na seção 5.

### 3. TRANSFORMADA DE FOURIER E PROPRIEDADES

Seja  $X(n), n = 0, \pm 1, \dots$ , um processo estacionário real. A transformada finita de Fourier dos valores  $X(0), X(1), \dots, X(T-1)$  de  $X(n)$  é dada por

$$d_x^{(T)}(\lambda) = \sum_{t=0}^{T-1} X(t) \exp(-i\lambda t), -\infty < \lambda < \infty, \quad (1)$$

e a transformada discreta de Fourier, por

$$d_x^{(T)}\left(\frac{2\pi s}{T}\right), s = 0, \pm 1, \dots, T-1. \quad (2)$$

A relação (1) pode ser invertida e a série  $X(t)$  pode ser escrita como

$$X(t) = (2\pi)^{-1} \int_0^{2\pi} \exp(-it\lambda) d_x^{(T)}(\lambda) d\lambda, \quad t = 0, \pm 1, \dots, T-1$$

ou

$$X(t) = T^{-1} \sum_{s=0}^{T-1} \exp\left(i\frac{2\pi st}{T}\right) d_x^{(T)}\left(\frac{2\pi s}{T}\right), \quad t = 0, \dots, T-1.$$

Para maiores detalhes, ver Brillinger(1981).

**Suposição 2.1:** Suponha  $X(t) = (X_1(t), \dots, X_r(t))'$  um processo estritamente estacionário satisfazendo  $\sum_{\mu_1, \dots, \mu_{k-1}}^{\infty} |C_x(\mu_1, \dots, \mu_{k-1})| < \infty$ , onde  $C_x(\mu_1, \dots, \mu_{k-1})$  é o cumulante de ordem  $k-1$ .

O *periodograma* de uma série temporal estacionária  $X(t), t = 0, \dots, T-1$  é definido por

$$I_{xx}^{(T)}(\lambda) = (2\pi T)^{-1} |d_x^{(T)}(\lambda)|^2, \lambda \neq 0 \pmod{2\pi}. \quad (3)$$

**Teorema 2.1:** Se  $X(t)$  satisfaz a Suposição 2.1,  $s_j(T)$  um inteiro com

$$\lambda_j(T) = \frac{2\pi s_j(T)}{T} \rightarrow \lambda_j, \quad T \rightarrow \infty, \quad j = 1, 2, \dots, J,$$

então para  $2\lambda_j(T)$ ,  $\lambda_j(T) \pm \lambda_k(T) \neq 0 \pmod{2\pi}$ ,  $1 \leq j < k \leq J$  e  $T = 1, 2, \dots$ , temos que  $I_{xx}^{(T)}(\lambda_j(T)) \xrightarrow{D} f_{xx}(\lambda_j) \chi^2_{2} / 2$  independentes.

Podemos notar que, embora o periodograma seja assintoticamente não viciado, ele não é consistente (Brillinger, 1981), o que leva à construção de estimadores espectrais com propriedades melhores, ou seja, para  $\lambda \neq 0 \pmod{\pi}$ ,

$$f_{xx}^{(T)}(\lambda) = (2m+1)^{-1} \sum_{j=-m}^m I_{xx}^{(T)}\left(\frac{2\pi[s(T)+j]}{T}\right), \quad (4)$$

que é uma média das ordenadas do periodograma em torno de  $\lambda$ , sendo denominado *periodograma suavizado*.

**Teorema 2.2:** Nas mesmas condições do Teorema 2.1, para

$$j = 1, \dots, J, f_{xx}^{(T)}(\lambda_j) \xrightarrow{D} f_{xx}(\lambda_j) \chi^2_{4m+2} / (4m+2) \text{ independentes se } \lambda_j \neq 0 \pmod{\pi}.$$

#### 4. ANÁLISE DE VARIÂNCIA USANDO A TRANSFORMADA DE FOURIER

Nesta seção, analisaremos as variações apresentadas nos dados cujas medidas respostas são séries temporais. As análises propostas podem ser vistas como extensões de procedimentos usados na ANOVA (Análise de Variância), quando cada ocorrência (ou série) de um planejamento conduz a uma variável real. Nosso objetivo consiste na aplicação de uma versão complexa de ANOVA aos componentes das transformadas discretas de Fourier das séries observadas. Serão considerados efeitos fixo e aleatório.

Uma abordagem comum da análise de variância é feita através de um modelo de regressão linear com a seguinte forma (ver Roy et al., 1971)

$$Y = Z \Theta X + \varepsilon \quad (5)$$

onde:

$Y_{(N \times T)}$  é a matriz das observações, onde as N linhas correspondem a cada um dos indivíduos

e T ao número de respostas de cada um deles;

$Z_{(N \times M)}$  é a matriz de planejamento entre os indivíduos;

$X_{(S \times T)}$  é a matriz de planejamento dentro de indivíduos;

$\Theta_{(M \times S)}$  é a matriz de coeficientes não observados (parâmetros); e

$\varepsilon_{(N \times T)}$  é uma matriz de erros estocásticos.

Geralmente, a matriz de planejamento Z é construída com 0's e 1's e Z e X são consideradas na maioria das vezes como fixas. A matriz  $\Theta$  é considerada fixa no modelo de

efeitos fixos e estocástica no modelo de efeitos aleatórios. Nosso objetivo é verificar se algum elemento de  $\Theta$  é zero, estimar os parâmetros de interesse e calcular os erros das estimativas.

O modelo (5) pode ser reescrito como

$$Y_n(t) = \sum_{j=1}^S \sum_{i=1}^M Z_{mi} \Theta_{ij} X_j(t) + \varepsilon_n(t), \quad n = 1, \dots, N, \quad t = 0, \dots, T-1, \quad (6)$$

onde  $Y_n(t)$  representa o elemento da  $n$ -ésima linha e  $t$ -ésima coluna de  $Y$ .

A análise de variância, quando temos como respostas um conjunto de  $N$  séries temporais estacionárias,  $\{Y_n(t), t = 0, \dots, T-1 \text{ e } n = 1, \dots, N\}$ , é realizada utilizando como ferramenta básica a transformada discreta de Fourier de cada uma delas,

$$d_{y_n}^{(T)}(k) = \sum_{t=0}^{T-1} Y_n(t) \exp\{-i2\pi kt / T\}, \quad k = 0, 1, \dots, \frac{[T-1]}{2}. \quad (7)$$

Pode-se demonstrar que, para  $k = 1, \dots, K, K \ll T, \frac{2\pi k}{T} \cong \lambda \text{ e } \frac{2\pi k}{T} \neq 0, \pm\pi, \dots,$

convergem em distribuição para  $N^c(0, 2\pi T f_{y_n y_n}(\lambda))$  independentes, onde  $f_{y_n y_n}(\lambda)$  é o espectro da série  $Y_n(t)$ . Portanto,

$$Y_n^{\wedge}(k) = (2\pi T)^{-1/2} d_{y_n}^{(T)}(k) \quad (8)$$

tem a seguinte representação:

$$Y_n^{\wedge}(k) = \xi_{nk} + O_{a.s.}(1), \quad (9)$$

para todo  $n$ , onde  $\xi_{n1}, \dots, \xi_{nK}$  são variáveis  $N^c(0, f_{y_n y_n}(\lambda))$  independentes e  $O_{a.s.}(1) \rightarrow 0$  quando  $T \rightarrow \infty$ .

Vimos na seção anterior que um estimador padronizado de  $f_{y_n y_n}(\lambda)$  denominado periodograma suavizado, é dado por

$$f_{y_n y_n}^{(T)}(\lambda) = K^{-1} \sum_{k=1}^K |Y_n^{\wedge}(k)|^2, \quad (10)$$

com a seguinte representação:

$$f_{y_n y_n}^{(T)}(\lambda) = f_{y_n y_n}(\lambda) \chi_{2K}^2 / 2K + O_{a.s.}(1). \quad (11)$$

Na prática, queremos que  $K$  seja grande de forma a estabilizar o estimador indicado por (11), mas, por outro lado, um valor muito grande de  $K$  aumentará muito o viés do estimador. Ver Skorokhod(1956), Hannan(1970) e Brillinger(1981). Veremos, agora, alguns casos particulares do modelo (6).

## 4.1 Modelo com um sinal comum

Um importante caso especial da família ANOVA é detectar um sinal comum  $\Omega(t)$ , possivelmente obscurecido pelo erro, em uma coleção de  $N$  séries temporais. Neste caso, podemos usar o modelo

$$Y_n(t) = \mu_n + \Omega(t) + \varepsilon_n(t), n = 1, \dots, N, t = 0, 1, \dots, T-1. \quad (12)$$

### 4.1.1 O modelo com sinal determinístico

Suponha que queremos detectar um comportamento padrão para o electrocorticograma medido no córtex frontal de ratos normais durante o estado do sono paradoxal. Seja  $Y_n(t), n = 1, \dots, N, t = 0, \dots, T-1$ , o electrocorticograma do  $n$ -ésimo rato no instante  $t$ . Utilizando o modelo (12) temos:

$\mu_n$  é a média do electrocorticograma do  $n$ -ésimo rato e é constante;

$\Omega(t)$  representa o sinal comum a todos os ratos e é determinístico; e

$\varepsilon_n(\cdot)$  são séries estacionárias, independentes com média zero e função de covariância  $c_{\varepsilon\varepsilon}(u)$  e espectro  $f_{\varepsilon\varepsilon}(\lambda)$ .

Segundo as suposições do modelo, temos:

$$EY_n(t) = \mu_n + \Omega(t) \text{ e } Cov\{Y_n(t+u), Y_n(t)\} = c_{\varepsilon\varepsilon}(u).$$

Testar a existência de um sinal comum é equivalente a testar a hipótese

$$H_{01}: \Omega(t) = 0, \quad (13)$$

que consiste em verificar se existe um sinal comum a todos os ratos.

Pela definição (8) e representação (9), o modelo (12) pode ser reescrito

$$Y_n^{\wedge}(k) = \Omega^{\wedge}(k) + \xi_{nk} + O_{a.s.}(1), \quad (14)$$

para  $k = 0, \dots, \frac{[T-1]}{2}$ , com  $\xi_{nk}$  variáveis  $N^c(0, f_{\varepsilon\varepsilon}(\lambda))$  independentes.

Para testar (13) consideraremos as seguintes somas de quadrados:

$$SQH_{01}(\lambda) = N \sum_{k=1}^K |Y^{\wedge}(k)|^2 \quad (15)$$

$$SQR_{01}(\lambda) = \sum_{n=1}^N \sum_{k=1}^K |Y_n^{\wedge}(k) - Y^{\wedge}(k)|^2, \quad (16)$$

onde:

$$Y^{\wedge}(k) = N^{-1} \sum_{n=1}^N Y_n^{\wedge}(k). \quad (17)$$

Utilizando a expressão (14) e a extensão complexa do teorema de Fisher-Cochran, (15) e (16) convergem quase certamente para

$$f_{\varepsilon\varepsilon}(\lambda)\chi_{2K}^2 \left( \frac{N \sum_k |\hat{\Omega}(k)|^2}{f_{\varepsilon\varepsilon}(\lambda)} \right) / 2 \text{ e } f_{\varepsilon\varepsilon}(\lambda)\chi_{2K(N-1)}^2 / 2, \quad (18)$$

respectivamente, com  $\chi_v^2$  independentes.

Utilizando (10), podemos calcular os valores esperados de (15) e (16)

$$E(SQH_{01}(\lambda)) = N K f_{\Omega\Omega}^{(T)}(\lambda) + K f_{\varepsilon\varepsilon}(\lambda),$$

$$E(SQR_{01}(\lambda)) = K(N-1) f_{\varepsilon\varepsilon}(\lambda).$$

O teste (13) é realizado utilizando a estatística

$$F_{01}(\lambda) = \frac{2SQH_{01}(\lambda) / 2K}{2SQR_{01}(\lambda) / 2K(N-1)}, \quad (19)$$

que tem uma distribuição assintótica F com parâmetro de não centralidade  $\frac{N \sum_k |\hat{\Omega}(k)|^2}{f_{\varepsilon\varepsilon}(\lambda)}$ .

Sob  $H_{01}$ , (19) tem uma distribuição  $F(2K, 2K(N-1))$  para  $0 \leq \lambda < \pi$ . Ver Brillinger(1980), Chiann(1993).

Os resultados obtidos estão resumidos na Tabela 1: ANOVAF (Análise de Variância usando Fourier).

Tabela 1: ANOVAF para teste (13) ou (23) na frequência  $\lambda$

fonte	quadrado médio = $2SQ / gl$ (QM)	estatística F	quadrado médio esperado (EQM)
$\Omega(t)$	$\frac{2N \sum_k  \hat{Y}_\cdot(k) ^2}{2K}$ (*)	(*) / (**)	$f_{\varepsilon\varepsilon}(\lambda) + N f_{\Omega\Omega}^{(T)}(\lambda)$
resíduo	$\frac{2 \sum_k \sum_{n=1}^N  Y_n^\wedge(k) - \hat{Y}_\cdot(k) ^2}{2K(N-1)}$ (**)		$f_{\varepsilon\varepsilon}(\lambda)$

Se quisermos usar os componentes espectrais não suavizados para testar (13) na frequência  $\lambda_k = \frac{2\pi k}{T}$ ,  $k = 0, 1, \dots, \frac{[T-1]}{2}$ , utilizaremos a estatística

$$F_{01}(\lambda_k) = \frac{N |Y_\cdot^\wedge(k)|^2}{\sum_{n=1}^N |Y_n^\wedge(k) - \hat{Y}_\cdot(k)|^2 / (N-1)} \quad (20)$$

que tem uma distribuição F com parâmetro de não centralidade  $\frac{N|\hat{\Omega}(k)|^2}{f_{\varepsilon\varepsilon}(\lambda_k)}$  e que sob  $H_{01}$ , tem uma distribuição  $F(2,2(N-1))$ . Ver Shumway(1988), Chiann(1993).

Pode-se demonstrar que o sinal  $\Omega(t)$  pode ser estimado por  $Y.(t) = \frac{\sum_N Y_n(t)}{N}$  com as seguintes propriedades

$$E(Y.(t)) = \mu + \Omega(t) \text{ e } Var(Y.(t)) = c_{\varepsilon\varepsilon}(0) / N.$$

#### 4.1.2 O modelo com sinal estocástico

Brillinger(1973) apresenta um exemplo onde o interesse é analisar a similaridade climática das estações americanas. Foram selecionadas aleatoriamente  $N = 5$  estações e utilizado o modelo (12),

$$Y_n(t) = \mu_n + \Omega(t) + \varepsilon_n(t),$$

onde:

$Y_n(t)$  temperatura mensal da n-ésima estação,  $n = 1, \dots, N$ ;

$\mu_n$  é a média da temperatura mensal da n-ésima estação e é constante;

$\Omega(\cdot)$  representa o componente climático comum a todas as estações americanas e é uma série temporal estacionária com média zero,  $Cov\{\Omega(t+u), \Omega(t)\} = c_{\Omega\Omega}(u)$  e espectro  $f_{\Omega\Omega}(\lambda), u = 0, \pm 1, \dots$ ;

$\varepsilon_n(\cdot)$  são séries temporais estacionárias, independentes com média zero,  $Cov\{\varepsilon_n(t+u), \varepsilon_n(t)\} = c_{\varepsilon\varepsilon}(u)$  e espectro  $f_{\varepsilon\varepsilon}(\lambda)$ ; e

$\Omega(\cdot)$  e  $\varepsilon_n(\cdot)$  são independentes.

Segundo as suposições do modelo, temos:

$$EY_n(t) = \mu_n,$$

$$Cov\{Y_n(t+u), Y_n(t)\} = c_{\Omega\Omega}(u) + c_{\varepsilon\varepsilon}(u),$$

Sob estas condições, a expressão (14) continua valendo, mas pode ser melhor representada por

$$\hat{Y}_n(k) = \gamma_k + \xi_{nk} + O_{a.s.}(1) \quad (21)$$

onde:

- (i)  $\gamma_k$  é uma variável  $N^c(0, f_{\Omega\Omega}(\lambda))$ ;
- (ii)  $\xi_{nk}$  são variáveis  $N^c(0, f_{\varepsilon\varepsilon}(\lambda))$  independentes; e
- (iii)  $\gamma_k$  e  $\xi_{nk}$  são independentes.

Neste caso, utilizando a expressão (21) e a extensão complexa do teorema de Fisher-Cochran, as somas de quadrados (15) e (16) convergem quase certamente para

$$(N f_{\Omega\Omega}(\lambda) + f_{\varepsilon\varepsilon}(\lambda)) \chi_{2K}^2 / 2 \text{ e } f_{\varepsilon\varepsilon}(\lambda) \chi_{2K(N-1)}^2 / 2, \quad (22)$$

respectivamente, com  $\chi_v^2$  independentes. Testar (13) é equivalente a testar

$$H_{02}: f_{\Omega\Omega}(\cdot) = 0, \quad (23)$$

que no exemplo significa verificar a existência de uma similaridade climática a todas as estações.

De acordo com (22), o teste (23) pode ser realizado utilizando a estatística (19) ou (20). Os resultados obtidos também estão resumidos na Tabela 1, onde  $E(\Omega(t)) = f_{\varepsilon\varepsilon}(\lambda) + N f_{\Omega\Omega}(\lambda)$ .

Um estimador de  $f_{\Omega\Omega}(\lambda)$ , na frequência  $\lambda$ , pode ser obtido por

$$f_{\Omega\Omega}^{(T)}(\lambda) = \frac{\sum_k |Y^{\wedge}(k)|^2}{K} - \frac{\sum_k \sum_n |Y_n^{\wedge}(k) - Y^{\wedge}(k)|^2}{N(N-1)K}, \quad (24)$$

com

$$E\{f_{\Omega\Omega}^{(T)}(\lambda)\} = f_{\Omega\Omega}(\lambda) \text{ e } \text{Var}\{f_{\Omega\Omega}^{(T)}(\lambda)\} = \frac{(f_{\Omega\Omega}(\lambda) + f_{\varepsilon\varepsilon}(\lambda) / N)^2}{K} + \frac{f_{\varepsilon\varepsilon}^2(\lambda)}{N^3 K}.$$

## 4.2 Modelo com um fator

Suponhamos que um experimento seja realizado de tal forma que M tratamentos diferentes são aplicados e cujos resultados são séries temporais estacionárias. Um modelo apropriado para esta situação é

$$Y_{lm}(t) = \mu_{lm} + \Omega(t) + \Gamma_m(t) + \varepsilon_{lm}(t), \quad (25)$$

$l = 1, \dots, L_m, m = 1, \dots, M, N = L_1 + \dots + L_M, t = 0, 1, \dots, T - 1$ , onde:

$L_m$ : número de séries submetidas ao m-ésimo tratamento;

$\Gamma_m(t)$ : efeito do m-ésimo tratamento; e

$\Omega(t)$ : sinal comum entre todas as séries observadas.

### 4.2.1 O modelo com fator fixo

Para exemplificar esse tipo de modelo, suponha que estamos interessados em verificar se existe diferença de comportamento entre os eletrocorticogramas medido no córtex frontal de ratos normais durante o estado de sono paradoxal e de alerta; também é de interesse verificar a existência de um comportamento comum para os dois estados. Consideremos  $Y_{lm}(t)$ , o eletrocorticograma do  $l$ -ésimo rato no  $m$ -ésimo estado e a expressão (25), onde:

$\mu_{lm}$  é a média do eletrocorticograma do  $l$ -ésimo rato no  $m$ -ésimo estado;

$\Omega(t)$  representa o sinal comum a todos os ratos e é determinístico;

$\Gamma_m(t)$  representa o sinal comum do  $m$ -ésimo estado e é determinístico com restrição

$$\sum_{m=1}^M L_m \Gamma_m(t) = 0; \text{ e}$$

$\varepsilon_{lm}(\cdot)$  são séries estacionárias, independentes, com média zero, função de covariância  $c_{\varepsilon\varepsilon}(u)$  e espectro  $f_{\varepsilon\varepsilon}(\lambda)$ .

Para este modelo, temos:

$$EY_{lm}(t) = \mu_{lm} + \Omega(t) + \Gamma_m(t),$$

$$\text{Cov}\{Y_{lm}(t+u), Y_{lm}(t)\} = c_{\varepsilon\varepsilon}(u)$$

$$\text{Cov}\{Y_{lm}(t+u), Y_{l'm'}(t)\} = 0, \quad (l', m') \neq (l, m).$$

Nosso problema é:

(1) detectar se existe diferença entre os efeitos dos tratamentos, testando

$$H_{03}: \Gamma_1(t) = \Gamma_2(t) = \dots = \Gamma_M(t) = 0; \text{ e} \quad (26)$$

(2) verificar a existência de similaridade entre todas as séries, testando

$$H_{04}: \Omega(t) = 0. \quad (27)$$

Se  $k$  é um inteiro com  $\frac{2\pi k}{T} \cong \lambda \neq 0$ , o modelo (25) pode ser escrito em termos da

transformada discreta de Fourier

$$\hat{Y}_{lm}(k) = \Omega^{\wedge}(k) + \Gamma_m^{\wedge}(k) + \xi_{lmk} + O_{a.s.}(1), \quad (28)$$

$l = 1, \dots, L_m; m = 1, \dots, M; k = 1, \dots, K$  e  $\xi_{lmk} \sim N^c(0, f_{\varepsilon\varepsilon}(\lambda))$  independentes.

Para fazer a análise consideraremos as seguintes somas de quadrados:

$$SQH_{04}(\lambda) = N \sum_k |\hat{Y}_{lm}(k)|^2, \quad (29)$$

$$SQH_{03}(\lambda) = \sum_k \sum_m L_m |Y_{..}^{\wedge}(k) - Y_{..}^{\wedge}(k)|^2, \quad (30)$$

e

$$SQR_{03}(\lambda) = \sum_k \sum_m \sum_l |Y_{lm}^{\wedge}(k) - Y_{..}^{\wedge}(k)|^2, \quad (31)$$

onde:

$$Y_{..}^{\wedge}(k) = L_m^{-1} \sum_l Y_{lm}^{\wedge}(k) \text{ e } Y_{..}^{\wedge}(k) = N^{-1} \sum_m L_m Y_{..}^{\wedge}(k).$$

Usando a expressão (28) e a extensão complexa do teorema de Fisher-Cochran, as expressões (29), (30) e (31) convergem quase certamente para

$$f_{\varepsilon\varepsilon}(\lambda) \chi_{2K}^2 \left( \frac{N \sum_k |\Omega^{\wedge}(k)|^2}{f_{\varepsilon\varepsilon}(\lambda)} \right) / 2,$$

$$f_{\varepsilon\varepsilon}(\lambda) \chi_{2K}^2 (M-1) \left( \frac{\sum_K \sum_m L_m |\Gamma_m^{\wedge}(k)|^2}{f_{\varepsilon\varepsilon}(\lambda)} \right) / 2 \text{ e } f_{\varepsilon\varepsilon}(\lambda) \chi_{2K(N-M)}^2 / 2,$$

respectivamente, com  $\chi_v^2$  independentes. Os resultados estão na Tabela 2.

O teste (26) pode ser realizado utilizando a estatística

$$F_{03}(\lambda) = \frac{2SQH_{03}(\lambda) / 2K(M-1)}{2SQR_{03}(\lambda) / 2K(N-M)}, \quad (32)$$

que tem uma distribuição assintótica F com parâmetro de não centralidade

$$\frac{\sum_K \sum_{m=1}^M L_m |\Gamma_m^{\wedge}(k)|^2}{f_{\varepsilon\varepsilon}(\lambda)}. \text{ Sob } H_{03}, (32) \text{ tem uma distribuição } F(2K(M-1), 2K(N-M))$$

para  $0 \leq \lambda < \pi$ .

O teste (27) pode ser realizado utilizando a estatística

$$F_{04}(\lambda) = \frac{2SQH_{04}(\lambda) / 2K}{2SQR_{03}(\lambda) / 2K(N-M)} \quad (33)$$

que também tem uma distribuição assintótica F com parâmetro de não centralidade

$$\frac{N \sum_k |\Omega^{\wedge}(k)|^2}{f_{\varepsilon\varepsilon}(\lambda)}. \text{ Sob } H_{04}, (33) \text{ tem uma distribuição } F(2K, 2K(N-M)).$$

Maiores informações, ver Brillinger(1973, 1980), Chiann(1993).

Se quisermos usar os componentes espectrais não suavizados para testar (26) e (27), na frequência  $\lambda_k = \frac{2\pi k}{T}$ ,  $k = 0, 1, \dots, \frac{[T-1]}{2}$ , utilizaremos as estatísticas (32) e (33) com  $K=1$ , que sob  $H_{03}$  e  $H_{04}$ , têm distribuições  $F(2(M-1), 2(N-M))$  e  $F(2, 2(N-M))$ , respectivamente.

O efeito do tratamento  $\Gamma_m^{\wedge}(k)$  pode ser estimado por

$$Y_m^{\wedge}(k) - Y_{..}^{\wedge}(k) \quad (34)$$

com as seguintes propriedades

$$E[Y_m^{\wedge}(k) - Y_{..}^{\wedge}(k)] = \Gamma_m^{\wedge}(k) \text{ e } Var[Y_m^{\wedge}(k) - Y_{..}^{\wedge}(k)] = \frac{L_m - 1}{L_m} f_{\varepsilon\varepsilon}(\lambda).$$

Tabela 2: ANOVA para os testes (26) ou (27) na frequência  $\lambda$

fonte	quadrado médio = $2SQ / gl$ (QM)	estatística F	quadrado médio esperado (EQM)
$\Omega(t)$	$\frac{2N \sum_k  Y_{..}^{\wedge}(k) ^2}{2K}$ (*)	(*) / (***)	$f_{\varepsilon\varepsilon}(\lambda) + N f_{\Omega\Omega}^T(\lambda)$
$\Gamma_m(t)$	$\frac{2 \sum_k \sum_{m=1}^M L_m  Y_m^{\wedge}(k) - Y_{..}^{\wedge}(k) ^2}{2K(M-1)}$ (**)	(**) / (***)	$f_{\varepsilon\varepsilon}(\lambda) + \frac{\sum_{m=1}^M L_m f_{\Gamma_m \Gamma_m}^T(\lambda)}{M-1}$
resíduo	$\frac{2 \sum_k \sum_{m=1}^M \sum_{l=1}^{Lm}  Y_{lm}^{\wedge}(k) - Y_m^{\wedge}(k) ^2}{2K(N-M)}$ (***)		$f_{\varepsilon\varepsilon}(\lambda)$

#### 4.2.2 O modelo com fator aleatório

Para exemplificar a utilização desse modelo, suponha que queremos verificar se existe diferença entre os componentes climáticos de estações continentais e, também, se existe um componente climático comum para todos os continentes. Para fazer a análise, selecionamos aleatoriamente dois continentes e  $L$  estações de cada um dos continentes, sendo  $Y_{lm}(t)$  as temperaturas mensais da  $l$ -ésima estação no  $m$ -ésimo continente. Este exemplo é citado por Brillinger(1973) e neste caso, utilizando o modelo (25) temos que

$\mu_{lm}$  é a média da temperatura da  $l$ -ésima estação no  $m$ -ésimo continente;

$\Omega(\cdot)$  representa o componente climático comum a todas as estações e é uma série estacionária com média zero, função de covariância  $c_{\Omega\Omega}(u)$  e espectro  $f_{\Omega\Omega}(\lambda)$ ;

$\Gamma_m(\cdot)$  representa o componente climático comum do  $m$ -ésimo continente e é uma série estacionária com média zero, função de covariância  $c_{\Gamma\Gamma}(u)$  e espectro comum  $f_{\Gamma\Gamma}(\lambda)$ ;

$\Omega(\cdot)$ ,  $\Gamma_m(\cdot)$  e  $\varepsilon_{lm}(\cdot)$  são independentes.

Sob as suposições do modelo, temos:

$$\begin{aligned} E(Y_{lm}(t)) &= \mu_{lm}, \\ \text{Cov}\{Y_{lm}(t+u), Y_{lm}(t)\} &= c_{\Omega\Omega}(u) + c_{\Gamma\Gamma}(u) + c_{\varepsilon\varepsilon}(u), \\ \text{Cov}\{Y_{lm}(t+u), Y_{l'm}(t)\} &= c_{\Omega\Omega}(u) + c_{\Gamma\Gamma}(u), \quad l \neq l', \\ \text{Cov}\{Y_{lm}(t+u), Y_{l'm'}(t)\} &= c_{\Omega\Omega}(u), \quad m \neq m'. \end{aligned}$$

Neste caso, a representação (28) pode ser reescrita

$$\hat{Y}_{lm}(k) = \gamma_k + v_{mk} + \xi_{lmk} + O_{a.s.}(1) \quad (35)$$

onde:

- (i)  $\gamma_k$  é uma variável  $N^c(0, f_{\Omega\Omega}(\lambda))$ ;
- (ii)  $v_{mk}$  são variáveis  $N^c(0, f_{\Gamma\Gamma}(\lambda))$  independentes;
- (iii)  $\xi_{lmk}$  são variáveis  $N^c(0, f_{\varepsilon\varepsilon}(\lambda))$  independentes; e
- (iv)  $\gamma_k$ ,  $v_{mk}$  e  $\xi_{lmk}$  são independentes.

No caso de modelo balanceado, ou seja,  $L_m = L$ , para  $m = 1, \dots, M$ , utilizando a expressão (35) e a extensão complexa do teorema de Fisher-Cochran, a soma de quadrados (31) continua valendo e as somas de quadrados (29) e (30) convergem quase certamente para

$$\left[ f_{\varepsilon\varepsilon}(\lambda) + Lf_{\Gamma\Gamma}(\lambda) + LMf_{\Omega\Omega}(\lambda) \right] \chi_{2K}^2 / 2 \text{ e } \left[ f_{\varepsilon\varepsilon}(\lambda) + Lf_{\Gamma\Gamma}(\lambda) \right] \chi_{2K(M-1)}^2 / 2, \quad (36)$$

respectivamente, com  $\chi_v^2$  independentes (Brillinger, 1973, 1980).

As hipóteses equivalentes a (26) e (27) são

$$H_{05}: f_{\Gamma\Gamma}(\lambda) = 0 \quad (37)$$

e

$$H_{06}: f_{\Omega\Omega}(\lambda) = 0, \quad (38)$$

respectivamente para  $0 \leq \lambda < \pi$ .

Neste caso, os resultados estão resumidos na Tabela 3.

A estatística para a realização do teste (37) tem, sob  $H_{05}$ , uma distribuição  $F(2K(M-1), 2K(N-M))$ . Se a hipótese  $H_{05}$  é verdadeira, ou seja,  $f_{\Gamma\Gamma}(\lambda) = 0$ , então o teste para (38) pode ser feito utilizando uma estatística  $F(2K, 2K(N-M))$ .

Pela Tabela 3, um estimador de  $f_{\Omega\Omega}(\lambda)$  pode ser obtido por

$$f_{\Omega\Omega}^{(T)}(\lambda) = \left\{ \frac{\sum_k |Y_{..}^{\wedge}(k)|^2}{K} - \frac{\sum_k \sum_{m=1}^M L_m |Y_m^{\wedge}(k) - Y_{..}^{\wedge}(k)|^2}{KN(M-1)} \right\},$$

com

$$E\{f_{\Omega\Omega}^{(T)}(\lambda)\} = f_{\Omega\Omega}(\lambda).$$

Além disso, também podemos obter um estimador de  $f_{\Gamma\Gamma}(\lambda)$ ,

$$f_{\Omega\Omega}^{(T)}(\lambda) = \left\{ \frac{\sum_k \sum_{m=1}^M L_m |Y_m^{\wedge}(k) - Y_{..}^{\wedge}(k)|^2}{K(N-L)} - \frac{\sum_k \sum_{m=1}^M \sum_{l=1}^{L_m} |Y_{lm}^{\wedge}(k) - Y_m^{\wedge}(k)|^2}{KN(L-1)} \right\},$$

com

$$E\{f_{\Gamma\Gamma}^{(T)}(\lambda)\} = f_{\Gamma\Gamma}(\lambda).$$

Tabela 3: ANOVA para os testes (37) ou (38) na frequência  $\lambda$

fonte	quadrado médio = $2SQ / gl$ (QM)	estatística F	quadrado médio esperado (EQM)
$\Omega(t)$	$\frac{2N \sum_k  Y_{..}^{\wedge}(k) ^2}{2K}$ (**)	(*) / (***)	$f_{\varepsilon\varepsilon}(\lambda) + L f_{\Gamma\Gamma}(\lambda) + LM f_{\Omega\Omega}(\lambda)$
$\Gamma_m(t)$	$\frac{2 \sum_k \sum_{m=1}^M L_m  Y_m^{\wedge}(k) - Y_{..}^{\wedge}(k) ^2}{2K(M-1)}$ (**)	(**) / (***)	$f_{\varepsilon\varepsilon}(\lambda) + L f_{\Gamma\Gamma}(\lambda)$
resíduo	$\frac{2 \sum_k \sum_{m=1}^M \sum_{l=1}^{L_m}  Y_{lm}^{\wedge}(k) - Y_m^{\wedge}(k) ^2}{2K(N-M)}$ (***)		$f_{\varepsilon\varepsilon}(\lambda)$

A análise do modelo não-balanceado é realizada utilizando a analogia feita na ANOVA usual; ver Scheffé(1959) ou Searle(1971).

### 4.3 Modelo linear geral com matriz de planejamento fixa (modelo de planejamento)

O modelo (5) pode ser reparametrizado nos seguintes modelos:

#### 4.3.1 O modelo com efeito fixo

Neste caso, para  $t = 0, 1, \dots, T - 1$ , consideraremos o modelo

$$Y(t) = \mu + \sum_u X(t-u)\theta(u) + \varepsilon(t) \tag{39}$$

que pode ser particionado da seguinte forma:

$$Y(t) = \mu + \sum_u X_1(t-u)\theta_1(u) + \sum_u X_2(t-u)\theta_2(u) + \varepsilon(t) \quad (40)$$

onde:

$Y(t)$  é um vetor observável, de dimensão  $N \times 1$ ;

$\mu$  é um vetor constante, de dimensão  $N \times 1$ ;

$X_1(t), X_2(t)$  são matrizes fixas e observáveis, de dimensão  $N \times P_1$  e  $N \times P_2$ ,

respectivamente, e satisfazendo  $\sum |X_1(t)| < \infty$  e  $\sum |X_2(t)| < \infty$ ;

$\theta_1(u), \theta_2(u)$  são vetores determinísticos, de dimensão  $P_1 \times 1$  e  $P_2 \times 1$ , respectivamente;

$\varepsilon(t)$  é um vetor de séries temporais estacionárias de dimensão  $N \times 1$ , com média zero, matriz espectral  $f_{\varepsilon\varepsilon}(\lambda)$  e satisfazendo Suposição 2.1.

Podemos notar que o modelo com um sinal comum, (12), é um caso particular de (40) com  $P_1 = 0, P_2 = 1, \theta_2(u) = \Omega(u)$  fixo para  $u = 0, \dots, T-1$  e  $X_2(t-u) = [1, \dots, 1]'$ .

No caso balanceado, o modelo com um fator, (25), também é um caso particular de (40) com  $P_1 = 1, \theta_1(u) = \Omega(u), P_2 = M, \theta_2(u) = (\Gamma_1(u), \dots, \Gamma_M(u))'$  fixo,  $X_1(t-u) = [1, \dots, 1]'$  e  $X_2(t-u)$  uma matriz  $(N \times M)$ , onde cada linha (série) tem o valor 1 na m-ésima coluna (série pertencente ao m-ésimo nível do fator) e zero nas demais colunas.

Além disso, existem dois importantes casos particulares de (40):

(1) O modelo de planejamento com dois fatores fixos cruzados,

$$Y_{lmq}(t) = \mu_{lmq} + \Omega(t) + \Gamma_m(t) + \Delta_q(t) + (\Gamma\Delta)_{mq}(t) + \varepsilon_{lmq}(t), \quad (41)$$

para  $l = 1, \dots, L, m = 1, \dots, M, q = 1, \dots, Q$  e  $N = LMQ$ , onde :

$L$ : número de séries submetidas ao m-ésimo nível do fator 1 e q-ésimo nível do fator 2;

$M$ : número de níveis do fator 1;

$Q$ : número de níveis do fator 2;

$\Gamma_m(t)$ : efeito do m-ésimo nível do fator 1, com restrição  $\sum_m \Gamma_m(t) = 0$ ;

$\Delta_q(t)$ : efeito do q-ésimo nível do fator 2, com restrição  $\sum_q \Delta_q(t) = 0$ ;

$(\Gamma\Delta)_{mq}(t)$ : interação entre m-ésimo nível do fator 1 e q-ésimo nível do fator 2, com restrições  $\sum_m (\Gamma\Delta)_{mq}(t) = 0$  e  $\sum_q (\Gamma\Delta)_{mq}(t) = 0$  ; e

$\Omega(t)$ : sinal comum entre todas as séries observadas.

Neste caso, nosso problema é:

(a) verificar a existência de interação entre os dois fatores, para isso, podemos testar, para todo  $t$ ,

$$H_{07}: (\Gamma\Delta)_{11}(t) = \dots = (\Gamma\Delta)_{M1}(t) = \dots = (\Gamma\Delta)_{12}(t) = \dots = (\Gamma\Delta)_{MQ}(t) = 0; \quad (42)$$

(b) detectar se existe diferença entre os efeitos dos níveis do fator 1, e

$$H_{08}: \Gamma_1(t) = \dots = \Gamma_M(t) = 0; \text{ e} \quad (43)$$

(c) detectar se existe diferença entre os efeitos dos níveis do fator 2,

$$H_{09}: \Delta_1(t) = \dots = \Delta_Q(t) = 0. \quad (44)$$

No caso deste modelo  $P_1 = 1 + M + Q$ ,  $\theta_1(u) = (\Omega(u), \Gamma_1(u), \dots, \Gamma_M(u),$

$\Delta_1(u), \dots, \Delta_Q(u))'$ ,  $P_2 = MQ$ ,  $\theta_2(u) = ((\Gamma\Delta)_{11}(u), \dots, (\Gamma\Delta)_{M1}(u),$

$(\Gamma\Delta)_{12}(u), \dots, (\Gamma\Delta)_{MQ}(u))'$ ,  $X_1(t-u)$  é uma matriz  $N \times [1 + M + Q]$ , onde cada linha (série) tem valor 1 na primeira coluna,  $(m+1)$ -ésima coluna (série pertence ao  $m$ -ésimo nível do fator 1) e  $(1 + M + q)$ -ésima coluna (série pertence ao  $q$ -ésimo nível do fator 2) e zeros nas demais colunas.  $X_2(t-u)$  é uma matriz  $(N \times MQ)$ , onde cada linha (série) tem valor 1 na  $mq$ -ésima coluna (série pertence ao  $m$ -ésimo nível do fator 1 e  $q$ -ésimo nível do fator 2) e zeros nas demais colunas.

(2) O modelo de planejamento com dois fatores fixos hierárquicos

$$Y_{lmq}(t) = \mu_{lmq} + \Omega(t) + \Gamma_m(t) + \Delta_{q(m)}(t) + \varepsilon_{lmq}(t), \quad (45)$$

para  $l = 1, \dots, L$ ,  $m = 1, \dots, M$ ,  $q = 1, \dots, Q$  e  $N = LMQ$ , onde:

$L$ : número de séries submetidas ao  $q$ -ésimo nível do fator 2 dentro do  $m$ -ésimo nível do fator 1;

$M$ : número de níveis do fator 1;

$Q$ : número de níveis do fator 2 dentro do fator 1;

$\Gamma_m(t)$ : efeito do  $m$ -ésimo nível do fator 1, com restrição  $\sum_m \Gamma_m(t) = 0$ ;

$\Delta_{q(m)}(t)$ : efeito do  $q$ -ésimo nível do fator 2 dentro do  $m$ -ésimo nível do fator 1, com restrição

$\sum_q \Delta_{q(m)}(t) = 0$  para todo  $m$ ; e

$\Omega(t)$ : sinal comum entre todas as séries observadas.

Neste caso, nosso problema é:

(a) detectar se existe diferença entre os efeitos dos níveis do fator 1,

$$H_{10}: \Gamma_1(t) = \dots = \Gamma_M(t) = 0; \quad (46)$$

(b) detectar se existe diferença entre os efeitos dos níveis do fator 2 dentro do  $m$ -ésimo nível do fator 1,

$$H_{11}: \Delta_{1(m)}(t) = \dots = \Delta_{Q(m)}(t) = 0, \text{ para } m = 1, \dots, M. \quad (47)$$

Aqui,  $P_1 = 1 + MQ$ ,  $\theta_1(u) = (\Omega(u), \Delta_{1(1)}(u), \dots, \Delta_{Q(1)}(u), \Delta_{1(2)}(u), \dots, \Delta_{Q(M)}(u))'$ ,

$P_2 = M$ ,  $\theta_2(u) = (\Gamma_1(u), \dots, \Gamma_M(u))'$ ,  $X_1(t-u)$  é uma matriz  $(N \times (1 + MQ))$ , onde cada linha (série) tem valor 1 na primeira coluna e na  $(1 + mq)$ -ésima coluna (série pertence ao  $q$ -ésimo nível dentro do  $m$ -ésimo nível do fator 1), as demais posições da linha fixada assumem o valor

zero.  $X_2(t-u)$  é uma matriz  $N \times M$ , onde cada linha (série) tem valor 1 na  $m$ -ésima coluna (série pertence ao  $m$ -ésimo nível do fator 1) e zero nas demais posições.

Para verificar se algum subconjunto dos parâmetros do modelo (40) está contribuindo significativamente para a resposta, testaremos a hipótese

$$H_{12} : \theta_2(u) = 0. \quad (48)$$

Para o modelo (12), este é essencialmente um problema de detectar um sinal comum (hipótese (13)), para o modelo (25), é de detectar a diferença entre os efeitos dos tratamentos (hipótese (26)), para o modelo (41), é de detectar a interação entre os dois fatores (hipótese (42)) e, para o modelo (45), é de detectar a diferença entre os efeitos dos níveis do fator 1 (hipótese (46)).

Um procedimento para fazer o teste (48) é comparar o modelo completo (40) com o modelo sob a hipótese  $H_{12}$  (modelo reduzido),

$$Y(t) = \sum_u X_1(t-u)\theta_1(u) + \varepsilon(t). \quad (49)$$

A representação de (40) em termos da transformada discreta de Fourier é

$$Y^{\wedge}(k) = X_1^{\wedge}(k)\theta_1^{\wedge}(k) + X_2^{\wedge}(k)\theta_2^{\wedge}(k) + \xi_k + O_{a.s.}(1) \quad (50)$$

$k = 1, \dots, K$ , onde  $\frac{2\pi k}{T} \equiv \lambda, \frac{2\pi k}{T} \neq 0, \pi$  com  $\xi_k \sim N^c(0, f_{\varepsilon\varepsilon})$  independentes.

Analogamente, (39) e (49) podem ser escritos como

$$Y^{\wedge}(k) = X^{\wedge}(k)\theta^{\wedge}(k) + \xi_k + O_{a.s.}(1) \quad (51)$$

e

$$Y^{\wedge}(k) = X_1^{\wedge}(k)\theta_1^{\wedge}(k) + \xi_k + O_{a.s.}(1), \quad (52)$$

respectivamente.

Para fazer a análise consideraremos as seguintes somas de quadrados

$$SQH_c(\lambda) = \sum_k \overline{Y^{\wedge}(k)}^T X^{\wedge}(k) (\overline{X^{\wedge}(k)}^T X^{\wedge}(k))^{-1} \overline{X^{\wedge}(k)}^T Y^{\wedge}(k), \quad (53)$$

$$SQH_{12}(\lambda) = \sum_k \overline{Y^{\wedge}(k)}^T X_1^{\wedge}(k) (\overline{X_1^{\wedge}(k)}^T X_1^{\wedge}(k))^{-1} \overline{X_1^{\wedge}(k)}^T Y^{\wedge}(k), \quad (54)$$

e

$$SQR_c(\lambda) = \sum_k \{ \overline{Y^{\wedge}(k)}^T Y^{\wedge}(k) - \overline{Y^{\wedge}(k)}^T X^{\wedge}(k) (\overline{X^{\wedge}(k)}^T X^{\wedge}(k))^{-1} \overline{X^{\wedge}(k)}^T Y^{\wedge}(k) \} \quad (55)$$

onde  $X^{\wedge}(k) = [X_1^{\wedge}(k) \quad X_2^{\wedge}(k)]$ .

Os resultados estão resumidos nas Tabelas 4 e 5, onde:

$$(i) \hat{e}(k) = \hat{Y}(k) - \hat{X}(k)\hat{\theta}(k);$$

$$(ii) \hat{A}(k) = \hat{X}(k)\hat{\theta}(k); \text{ e}$$

$$(iii) \hat{B}(k) = \hat{X}_1(k)\hat{\theta}_1(k).$$

O teste (48) pode ser realizado utilizando a estatística

$$F_{12}(\lambda_k) = \frac{(SQH_c(\lambda_k) - SQH_{12}(\lambda_k)) / 2P_2}{SQR_c(\lambda_k) / 2K(N - P)} \quad (56)$$

que, sob  $H_{12}$ , tem uma distribuição  $F(2K P_2, 2K(N - P))$ .

Para mais detalhes, ver Brillinger(1980) e Chiann(1993).

Se quisermos usar os componentes espectrais não suavizados para testar (48), na

freqüência  $\lambda_k = \frac{2\pi k}{T}, k = 0, 1, \dots, \frac{[T-1]}{2}$ , utilizaremos a estatística

$$F_{12}(\lambda) = \frac{(SQH_c(\lambda) - SQH_{12}(\lambda)) / 2KP_2}{SQR_c(\lambda) / 2K(N - P)} \quad (57)$$

que, sob  $H_{12}$ , tem uma distribuição  $F(2P_2, 2(N - P))$ , onde:

$$SQH_c(\lambda_k) = \overline{\hat{Y}(k)^T \hat{X}(k) (\hat{X}(k)^T \hat{X}(k))^{-1} \hat{X}(k)^T \hat{Y}(k)},$$

$$SQH_{12}(\lambda_k) = \overline{\hat{Y}_1(k)^T \hat{X}_1(k) (\hat{X}_1(k)^T \hat{X}_1(k))^{-1} \hat{X}_1(k)^T \hat{Y}(k)}$$

e

$$SQR_c(\lambda_k) = \overline{\hat{Y}(k)^T \hat{Y}(k) - \hat{Y}(k)^T \hat{X}(k) (\hat{X}(k)^T \hat{X}(k))^{-1} \hat{X}(k)^T \hat{Y}(k)}.$$

Tabela 4: ANOVA para o teste (48) na freqüência  $\lambda$

fonte	soma de quadrados (SQ)		
Modelo reduzido	$\sum_k \overline{\hat{Y}(k)^T \hat{X}_1(k) (\hat{X}_1(k)^T \hat{X}_1(k))^{-1} \hat{X}_1(k)^T \hat{Y}(k)}$ (*)		
Modelo só com $\theta_2(t)$	(**) - (*)		
Modelo completo	$\sum_k \overline{\hat{Y}(k)^T \hat{X}(k) (\hat{X}(k)^T \hat{X}(k))^{-1} \hat{X}(k)^T \hat{Y}(k)}$ (***)		
resíduo	(***) - (**)		
Total	$\sum_k \overline{\hat{Y}(k)^T \hat{Y}(k)}$ (***)		

fonte	grau de liberdade (gl)	quadrado médio = $\frac{2SQ}{g'}$ (QM)	estatística F
Modelo reduzido	$2KP_1$	$\sum_{n=1}^N f_{B_n B_n}^{(T)}(\lambda) / P_1$	
Modelo só com $\theta_2(t)$	$2KP_2$		$\frac{(**) - (*) / 2KP_2}{(***) - (**) / 2K(N - P)}$
Modelo completo	$2KP$	$\sum_{n=1}^N f_{A_n A_n}^{(T)}(\lambda) / P$	$\frac{(**) / 2KP}{(***) - (**) / 2K(N - P)}$
resíduo	$2K(N - P)$	$\sum_{n=1}^N f_{e_n e_n}^{(T)}(\lambda) / (N - P)$	
Total	$2KN$		

Um estimador de  $\hat{\theta}(k)$  é

$$\hat{\theta}(k) = (\overline{X^{\wedge}(k)}^T X^{\wedge}(k))^{-1} \overline{X^{\wedge}(k)}^T Y^{\wedge}(k), \quad (58)$$

se  $(\overline{X^{\wedge}(k)}^T X^{\wedge}(k))^{-1}$  existe, com as seguintes propriedades

$$E[\hat{\theta}(k)] = \hat{\theta}(k) \quad e \quad Var[\hat{\theta}(k)] = (\overline{X^{\wedge}(k)}^T X^{\wedge}(k))^{-1} f_{\varepsilon\varepsilon}(\lambda).$$

No caso  $\theta_2(u) = 0$ , um estimador de  $\hat{\theta}_1(k)$  pode ser

$$\hat{\theta}_1(k) = (\overline{X_1^{\wedge}(k)}^T X_1^{\wedge}(k))^{-1} \overline{X_1^{\wedge}(k)}^T Y^{\wedge}(k). \quad (59)$$

Para mais detalhes, ver Roy et al. (1971) e Brillinger(1980).

Tabela 5: Quadro médio esperado na frequência  $\lambda$  no caso  $\theta_2(u) = 0$

fonte	quadrado médio esperado (EQM)
Modelo reduzido	$f_{\varepsilon\varepsilon}(\lambda) + \frac{tr \left\{ f_{\theta_1\theta_1}^{(T)}(\lambda) \overline{X_1(\lambda)}^T X_1(\lambda) \right\}}{P_1}^a$
Modelo só com $\theta_2(t)$	$f_{\varepsilon\varepsilon}(\lambda)$
Modelo completo	$f_{\varepsilon\varepsilon}(\lambda) + \frac{tr \left\{ f_{\theta\theta}^{(T)}(\lambda) \overline{X(\lambda)}^T X(\lambda) \right\}}{P}$
Resíduo	$f_{\varepsilon\varepsilon}(\lambda)$

$$^a X(\lambda) = \sum_u \exp\{-i\lambda u\} X(u)$$

Um teste análogo ao (48) é testar a hipótese

$$H_{13}: \theta(u) = 0, \quad (60)$$

que pode ser feita da seguinte forma:

(1) utilizar os componentes espectrais suavizados,

$$F_{13}(\lambda) = \frac{SQH_c(\lambda) / 2KP}{SQR_c(\lambda) / 2K(N - P)} \quad (61)$$

que, sob  $H_{13}$ , tem uma distribuição  $F(2KP, 2K(N - P))$ ; e

(2) utilizar os componentes não suavizados, na frequência  $\lambda_k = \frac{2\pi k}{T}, k = 0, 1, \dots, \frac{[T-1]}{2}$ ,

$$F_{13}(\lambda_k) = \frac{SQH_c(\lambda_k) / 2P}{SQR_c(\lambda_k) / 2(N - P)} \quad (62)$$

que, sob  $H_{13}$ , tem uma distribuição  $F(2P, 2(N - P))$  (Shumway, 1988).

### 4.3.2 O modelo com efeito aleatório

Consideremos o modelo (40),

$$Y(t) = \mu + \sum_u X_1(t-u)\theta_1(u) + \sum_u X_2(t-u)\theta_2(u) + \varepsilon(t),$$

onde, agora  $\theta_1(u), \theta_2(u)$  são séries estacionárias, independentes com média zero, matrizes espectrais  $f_{\theta_1\theta_1}(\lambda)$  e  $f_{\theta_2\theta_2}(\lambda)$ , respectivamente, e satisfazendo Suposição 2.1.

Neste caso, a representação (50) pode ser reescrita

$$Y^{\wedge}(k) = X_1^{\wedge}(k)\Theta_{1k} + X_2^{\wedge}(k)\Theta_{2k} + \xi_k + O_{a.s.} \quad (63)$$

onde

- (i)  $\Theta_{1k}, \Theta_{2k}$ : variáveis  $N^c(0, f_{\theta_1\theta_1}(\lambda))$  e  $N^c(0, f_{\theta_2\theta_2}(\lambda))$  independentes;
- (ii)  $\xi_k$ : variáveis  $N^c(0, f_{\varepsilon\varepsilon}(\lambda))$  independentes; e
- (iii)  $\Theta_{1k}, \Theta_{2k}$  e  $\xi_k$  independentes.

A hipótese equivalente a (48) é

$$H_{14}: f_{\theta_2\theta_2}(\lambda) = 0, \quad 0 < \lambda < \pi. \quad (64)$$

Analogamente, a hipótese equivalente a (60) é

$$H_{15}: f_{\theta\theta}(\lambda) = 0, \quad 0 < \lambda < \pi. \quad (65)$$

Neste caso, os resultados apresentados na Tabela 4 continuam valendo, e os valores de quadrado médio esperado estão representados na Tabela 6.

Tabela 6: Quadro médio esperado para os testes (64) e (65) na  $\lambda$

fonte	quadrado médio esperado (EQM)
Modelo reduzido	(**) - (*)
Modelo só com $\theta_2(t)$	$f_{\varepsilon\varepsilon}(\lambda) + \frac{\text{tr} \left\{ f_{\theta_2\theta_2}^{(T)}(\lambda) \overline{X_2(\lambda)}^T \left[ I - X_1(\lambda) \overline{X_1(\lambda)}^T X_1(\lambda) \right]^{-1} \overline{X_1(\lambda)}^T \right\} X_2(\lambda)}{P_2}$ (*)
Modelo completo	$f_{\varepsilon\varepsilon}(\lambda) + \frac{\text{tr} \left\{ f_{\theta\theta}(\lambda) \overline{X(\lambda)}^T X(\lambda) \right\}}{P}$ (**)
Resíduo	$f_{\varepsilon\varepsilon}(\lambda)$

Podemos notar que o modelo com um sinal comum estocástico é um caso particular do (63) com  $P_1 = 0$ ,  $P_2 = 1$  e  $\Theta_{2k} = \gamma_k$ . No caso balanceado, o modelo com um fator aleatório, também é um caso particular do (63) com  $P_1 = 0$ ,  $\Theta_{1k} = \gamma_k$ ,  $P_2 = M$  e  $\Theta_{2k} = (v_{1k}, \dots, v_{Mk})'$ .

As análises feitas para o modelo com um sinal comum estocástico (21) e o modelo com um fator aleatório (35) são análogas às do modelo linear com matriz de planejamento fixa com efeito aleatório (63). No caso de modelo balanceado, as hipóteses (64) e (65) podem ser testadas da mesma maneira que para o modelo com efeito fixo, isto é, utilizando as estatísticas (56)-(57) e (61)-(62), respectivamente. No caso não balanceado as estatísticas envolvidas nos testes têm distribuições assintóticas mais complicadas. Para melhores referências ver Scheffé(1959), Searle(1971), Brillinger(1980, 1985) e Shumway(1988).

## 5. APLICAÇÕES

As observações dos eletroscilogramas dos quarenta ratos normais foram obtidas durante a atividade *rítmica* de acordo com um modelo de planejamento cruzado com dois fatores fixos,

$$Y_{lmq}(t) = \mu_{lmq} + \Omega(t) + \Gamma_m(t) + \Delta_q(t) + (\Gamma\Delta)_{mq}(t) + \varepsilon_{lmq}(t),$$

com os quarenta ratos subdivididos da seguinte forma:

		Região	
		Posteriorotálamo	Hipotálamo
Estado	Alerta	rato 01 ( $Y_{10,1,1}$ )	rato 21 ( $Y_{1,2,1}$ )
		∴ ( $PAR$ )	∴ ( $HAR$ )
	rato 10 ( $Y_{10,1,1}$ )	rato 30 ( $Y_{10,2,1}$ )	
	Sono paradoxal	rato 11 ( $Y_{1,1,2}$ )	rato 31 ( $Y_{1,2,2}$ )
∴ ( $PSR$ )		∴ ( $HSR$ )	
		rato 20 ( $Y_{10,1,2}$ )	rato 40 ( $Y_{10,2,2}$ )

onde:

$Y_{lmq}(t)$  é o eletroscilograma do  $l$ -ésimo rato, sujeito aos fatores região ( $m$ ) e estado ( $q$ ), no instante  $t$ ,  $l = 1, \dots, 10$ ,  $m = 1, 2$  e  $q = 1, 2$ ;

$\mu_{lmq}$  é um valor médio dos eletroscilogramas do  $l$ -ésimo rato, sujeito aos fatores região ( $m$ ) e estado ( $q$ );

$\Omega(t)$  é o sinal comum dos eletroscilogramas de todos os ratos, no instante  $t$ ;

$\Gamma_m(t)$  é o comportamento comum dos eletroscilogramas dos ratos na  $m$ -ésima região (posteriorotálamo ou hipotálamo), no instante  $t$ ,  $m = 1, 2$ ;

$\Delta_q(t)$  é o comportamento comum dos eletroscilogramas dos ratos durante o  $q$ -ésimo estado (alerta ou sono paradoxal), no instante  $t$ ,  $q = 1, 2$ ;

$(\Gamma\Delta)_{mq}(t)$  é a interação entre os fatores região e estado, no instante  $t$ ; e

$\varepsilon_{lmq}(t)$  é o erro aleatório.

Os eletroscilogramas de alguns ratos, juntamente com os respectivos periodogramas, estão apresentados na Figura 1. Analisando os periodogramas, notamos que existem valores altos em baixas frequências.

Na primeira etapa da análise, vamos verificar a existência de interação entre região e estado, ou seja, testar a hipótese

$$H_{01}: (\Gamma\Delta)_{11}(t) = (\Gamma\Delta)_{12}(t) = (\Gamma\Delta)_{21}(t) = (\Gamma\Delta)_{22}(t) = 0. \quad (66)$$

A análise é feita utilizando a Tabela 4 (ANOVAF). Desenvolvendo as fórmulas para um modelo com dois fatores cruzados fixos, com dois níveis cada, chegamos à ANOVAF dada pela Tabela 7.

Tabela 7: ANOVAF para o modelo com dois fatores cruzados fixos com dois níveis cada

fonte	soma de quadrados (SQ)	gl
$\Omega(t)$	$40 \sum_k \left  \hat{Y}_{\dots}(k) \right ^2$ (**)	2K
$\Gamma_m(t)$	$\sum_k \sum_{m=1}^2 20 \left  \hat{Y}_{m.}(k) - \hat{Y}_{\dots}(k) \right ^2$ (***)	2K
$\Delta_q(t)$	$\sum_k \sum_{q=1}^2 20 \left  \hat{Y}_{.q}(k) - \hat{Y}_{\dots}(k) \right ^2$ (****)	2K
$(\Gamma\Delta)_{mq}(t)$	$\sum_k \sum_{m=1}^2 \sum_{q=1}^2 10 \left  \hat{Y}_{mq}(k) - \hat{Y}_{m.}(k) - \hat{Y}_{.q}(k) + \hat{Y}_{\dots}(k) \right ^2$ (*)	2K
resíduo	$\sum_k \sum_{l=1}^{10} \sum_{m=1}^2 \sum_{q=1}^2 \left  \hat{Y}_{lmq}(k) - \hat{Y}_{mq}(k) \right ^2$ (*****)	72K
Total	$\sum_k \sum_{l=1}^{10} \sum_{m=1}^2 \sum_{q=1}^2 \left  \hat{Y}_{lmq}(k) \right ^2$	80K

fonte	quadrado médio = $\frac{2SQ}{gl}$ (QM)	estatística (F)	quadrado médio esperado (EQM)
$\Omega(t)$	$\frac{**}{2K}$	$\frac{** / 2K}{***** / 72K}$	$f_{\varepsilon\varepsilon} + 40f_{\Omega\Omega}^{(T)}(\lambda)$
$\Gamma_m(t)$	$\frac{***}{2K}$	$\frac{*** / 2K}{***** / 72K}$	$f_{\varepsilon\varepsilon} + 20f_{\Gamma\Gamma}^{(T)}(\lambda)$
$\Delta_q(t)$	$\frac{****}{2K}$	$\frac{**** / 2K}{***** / 72K}$	$f_{\varepsilon\varepsilon} + 20f_{\Delta\Delta}^{(T)}(\lambda)$
$(\Gamma\Delta)_{mq}(t)$	$\frac{*}{2K}$	$\frac{* / 2K}{***** / 72K}$	$f_{\varepsilon\varepsilon} + 10f_{(\Gamma\Delta)_{mq}}^{(T)}(\lambda)$
resíduo	$\frac{*****}{72K}$		$f_{\varepsilon\varepsilon}$

O teste (66) pode ser realizado utilizando a estatística que, sob

$$F_{01}(\lambda) = \frac{\sum_k \sum_{m=1}^2 \sum_{q=1}^2 10 |Y_{mq}^{\wedge}(k) - Y_{.m}^{\wedge}(k) + Y_{.q}^{\wedge}(k) + Y_{..}^{\wedge}(k)|^2 / 2K}{\sum_k \sum_{l=1}^{10} \sum_{m=1}^2 \sum_{q=1}^2 |Y_{lmq}^{\wedge}(k) - Y_{.mq}^{\wedge}(k)|^2 / 72K}, \quad (67)$$

que, sob  $H_{01}$ , tem uma distribuição  $F(2K, 72K)$ , para  $0 < \lambda < \pi$ .

A Figura 2 mostra os resultados da estatística (67), para  $K=1$ , que indica a existência, em nível de 0,05, de uma forte interação na frequência  $\frac{2\pi 57}{256}$  (período de 0,030 segundos). A

existência de interação significa que o comportamento do eletroscilograma dos ratos numa determinada região depende do estado em que eles se encontram, ou seja, o comportamento do eletroscilograma dos ratos na região hipotálamo durante o sono paradoxal é diferente daquele durante o estado de alerta. As retas horizontais indicam os níveis críticos 0,05 e 0,01. Cada ponto da figura representa o grau de interação entre estado e região naquela frequência; quanto maior o valor da ordenada em uma determinada frequência mais forte será a interação. As frequências, onde a interação é significativa, estão representadas na Tabela 8.

Figura 2 - Estatística (67) para verificar a interação entre estado e região.

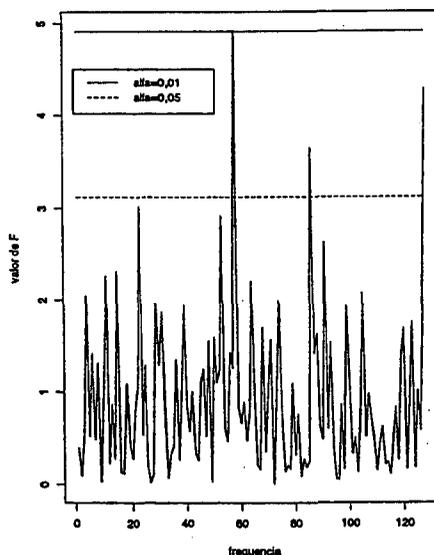


Tabela 8: Freqüências selecionadas pelo teste(67)

índice	freqüência	período	estatística F	nível descritivo
57	1,399	0,030	4,9271	0,0099
85	2,086	0,020	3,6567	0,0307
127	3,117	0,013	4,2871	0,0174

Uma vez detectada a existência de interação no modelo, verificaremos, numa **segunda etapa**, se existe diferença entre os sinais do eletroscilograma na região do posteriotálamo e do hipotálamo durante o estado de alerta (comparação entre as regiões no estado de alerta) e, também, durante o sono paradoxal (comparação entre as regiões no estado de sono paradoxal). A análise de variância para comparação entre as regiões durante cada um dos estados está representada na Tabela 9, onde:

$\Gamma_{m/i}(t)$  é o comportamento comum dos eletroscilogramas dos ratos na m-ésima região durante o estado  $i$ , no instante  $t$ ;  $i = 1,2$ ; e

$$f_{\Gamma_{m1}\Gamma_{m2}}^T(\lambda) = K^{-1} \sum_{k=1}^K |\Gamma_{m/i}^{\wedge}(k)|^2; i = 1,2.$$

Tabela 9: ANOVAF para comparação entre as regiões durante cada um dos estados

fonte	quadrado médio = $2SQ / gl$ (QM)	estatística F	quadrado médio esperado (EQM)
$\Gamma_{m/1}(t)$	$\frac{2 \sum_k \sum_{m=1}^2 10  Y_{m1}^{\wedge}(k) - Y_{.1}^{\wedge}(k) ^2}{2K}$ (*)	*/***	$f_{\varepsilon\varepsilon}(\lambda) + 10L f_{\Gamma_{m1}\Gamma_{m1}}^{(T)}(\lambda)$
$\Gamma_{m/2}(t)$	$\frac{2 \sum_k \sum_{m=1}^2 10  Y_{m2}^{\wedge}(k) - Y_{.2}^{\wedge}(k) ^2}{2K}$ (**)	**/**	$f_{\varepsilon\varepsilon}(\lambda) + 10L f_{\Gamma_{m2}\Gamma_{m2}}^{(T)}(\lambda)$
resíduo	$\frac{2 \sum_k \sum_{l=1}^{10} \sum_{m=1}^2 \sum_{q=1}^2  Y_{lmq}^{\wedge}(k) - Y_{.mq}^{\wedge}(k) ^2}{2K(N - M)}$ (***)		$f_{\varepsilon\varepsilon}$

Para fazer a comparação entre as regiões no estado de alerta, testamos

$$H_{02}: \Gamma_{1/1}(t) = \Gamma_{2/1}(t) = 0, \quad (68)$$

e no caso de comparações entre as regiões durante sono paradoxal,

$$H_{03}: \Gamma_{1/2}(t) = \Gamma_{2/2}(t) = 0. \quad (69)$$

O teste para (68) pode ser realizado utilizando a estatística

$$F_{02}(\lambda) = \frac{\sum_k \sum_{m=1}^2 10 |Y_{m1}^{\wedge}(k) - Y_{.1}^{\wedge}(k)|^2 / 2K}{\sum_k \sum_{l=1}^{10} \sum_{m=1}^2 \sum_{q=1}^2 |Y_{lmq}^{\wedge}(k) - Y_{.mq}^{\wedge}(k)|^2 / 72K}, \quad (70)$$

que, sob  $H_{02}$ , tem uma distribuição  $F(2K,72K)$ , para  $0 < \lambda < \pi$ .

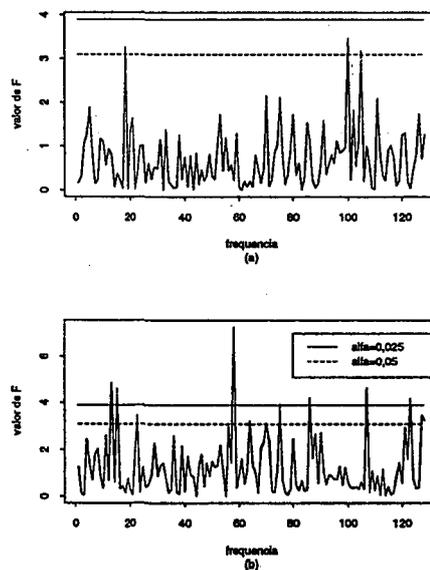
O teste para (69) também pode ser realizado utilizando a estatística

$$F_{03}(\lambda) = \frac{\sum_k \sum_{m=1}^2 10 |Y_{m2}^{\wedge}(k) - Y_{..2}^{\wedge}(k)|^2 / 2K}{\sum_k \sum_{l=1}^{10} \sum_{m=1}^2 \sum_{q=1}^2 |Y_{lmq}^{\wedge}(k) - Y_{mq}^{\wedge}(k)|^2 / 72K}, \quad (71)$$

que, sob  $H_{03}$ , tem uma distribuição  $F(2K,72K)$ , para  $0 < \lambda < \pi$ .

Os resultados, para  $K = 1$ , estão apresentados na Figura 3(a) (comparação entre as regiões no estado de alerta) e Figura 3(b) (comparação entre as regiões durante o estado de sono paradoxal). As linhas horizontais indicam os níveis críticos 0,05 e 0,025 de uma distribuição  $F(2,72)$ . Cada um dos pontos da Figura 3 representa o grau de diferença entre os periodogramas dos ratos na região do hipotálamo e do posteriotálamo, durante cada um dos estados, isto é, quanto maior o valor da ordenada em uma determinada frequência maior será a diferença dos periodogramas entre os ratos de diferentes regiões, durante cada um dos estados de sono. Os valores das estatísticas (70) e (71) nas frequências significantes estão nas Tabelas 10 e 11.

Figura 3 -Estatística (70) e (71) para comparação de sinais entre as regiões durante o estado de (a) alerta e (b) sono paradoxal, respectivamente.



Analisando os resultados da Tabela 10, vemos que existe uma diferença, entre as regiões do cérebro, no comportamento dos eletroscilogramas nas frequências  $\frac{2\pi 17}{256}$  (correspondente a um período de 0,1 segundos),  $\frac{2\pi 99}{256}$  (período de 0,017 segundo) e  $\frac{2\pi 104}{256}$  (período de 0,016 segundo), durante o estado de alerta.

Tabela 10: Frequências selecionadas pelo teste(70)  
(comparação entre as regiões do cérebro durante o estado de alerta)

índice	frequência	período	estatística F	nível descritivo
17	0,417	0,100	3,2668	0,0438
99	2,430	0,017	3,4666	0,0365
104	2,553	0,016	3,1872	0,0472

Durante o estado de sono paradoxal, Tabela 11, também existe uma diferença, entre as regiões do cérebro, no comportamento dos eletroscilogramas. Entretanto, essa diferença se manifesta de forma diferente do caso anterior, isto é, em um número maior e diferente de frequências.

Estes resultados são consistentes com a existência de interação entre estado e região.

Tabela 11: Frequências selecionadas pelo teste(71)  
(comparação entre as regiões do cérebro durante o estado de sono paradoxal)

índice	frequência	período	estatística F	nível descritivo
12	0,294	0,142	4,8712	0,0104
14	0,344	0,121	4,6262	0,0129
22	0,540	0,077	3,4545	0,0369
57	1,399	0,030	7,2849	0,0013
63	1,546	0,027	3,2488	0,0449
69	1,693	0,025	3,1496	0,0446
74	1,816	0,023	3,9623	0,0233
85	2,086	0,020	4,2199	0,0185
106	2,601	0,016	4,6447	0,0127
122	2,994	0,014	4,1909	0,0190
126	3,093	0,013	3,4759	0,0362
127	3,117	0,013	3,2475	0,0446

Numa **terceira etapa**, podemos construir os eletroscilogramas comuns para os ratos de cada um dos grupos: PAR, PSR, HAR e HSR.

Para esta análise utilizamos o modelo (12)

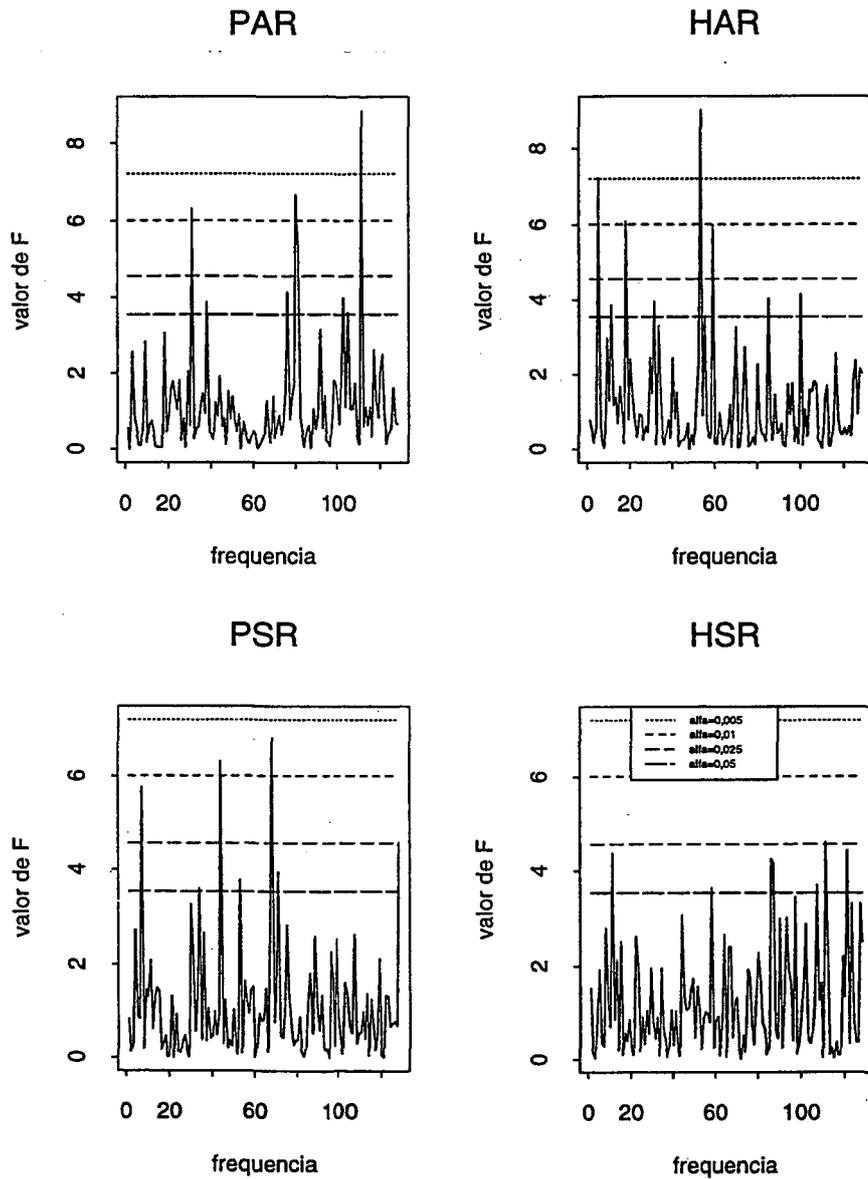
$$Y_n(t) = \mu_n + \Omega(t) + \varepsilon_n(t), \quad n = 1, \dots, 10,$$

e testamos a hipótese  $H_{04} : \Omega(t) = 0$ , através da estatística (20), para cada um dos grupos.

Os resultados estão na Figura 4, onde cada ponto representa o grau de similaridade entre os periodogramas dos ratos naquela frequência, isto é, para cada grupo, quanto maior o valor da ordenada em uma determinada frequência mais parecidos serão os periodogramas dos ratos naquela frequência. As retas horizontais representam os níveis críticos 0,05, 0,025, 0,01 e 0,005 de uma distribuição  $F(2,18)$ . Os valores das ordenadas superiores a essas retas corresponderão às frequências comuns presentes nos eletroscilogramas dos ratos do mesmo grupo. Notamos que, em nível de 0,05, temos oito frequências comuns nos grupos PAR e HAR, sete nos grupos PSR e HSR; em nível de 0,025, temos quatro frequências comuns nos

grupos PAR, HAR e PSR, e uma no grupo HSR; em nível de 0,01, temos três nos grupos PAR e HAR, duas no grupo PSR e nenhuma no grupo HSR enquanto em nível de 0,005, uma no grupo PAR, duas no grupo HAR e nenhuma nos grupos PSR e HSR.

Figura 4 - Estatística (20) para testar um sinal comum em cada um dos grupos.



A Tabela 12 apresenta as freqüências comuns, cujos níveis críticos são inferiores a 0,05, para cada um dos grupos. No grupo PAR, o maior pico é  $F(110) = 8,8296$ , significando que o componente correspondente a um período de 0,016 segundo é um forte componente comum. Analogamente, no grupo HAR é  $F(52) = 9,0022$  que corresponde a um período de 0,033 segundo, no grupo PSR é  $F(67) = 6,814$  correspondente a um período de 0,025 segundo, e no grupo HSR é  $F(110) = 4,615$  correspondente a um período de 0,016 segundo. Notamos também que os grupos PAR e HSR têm um forte componente comum em alta freqüência (período curto), enquanto os grupos HAR e PSR têm um forte componente comum em baixa freqüência (período longo).

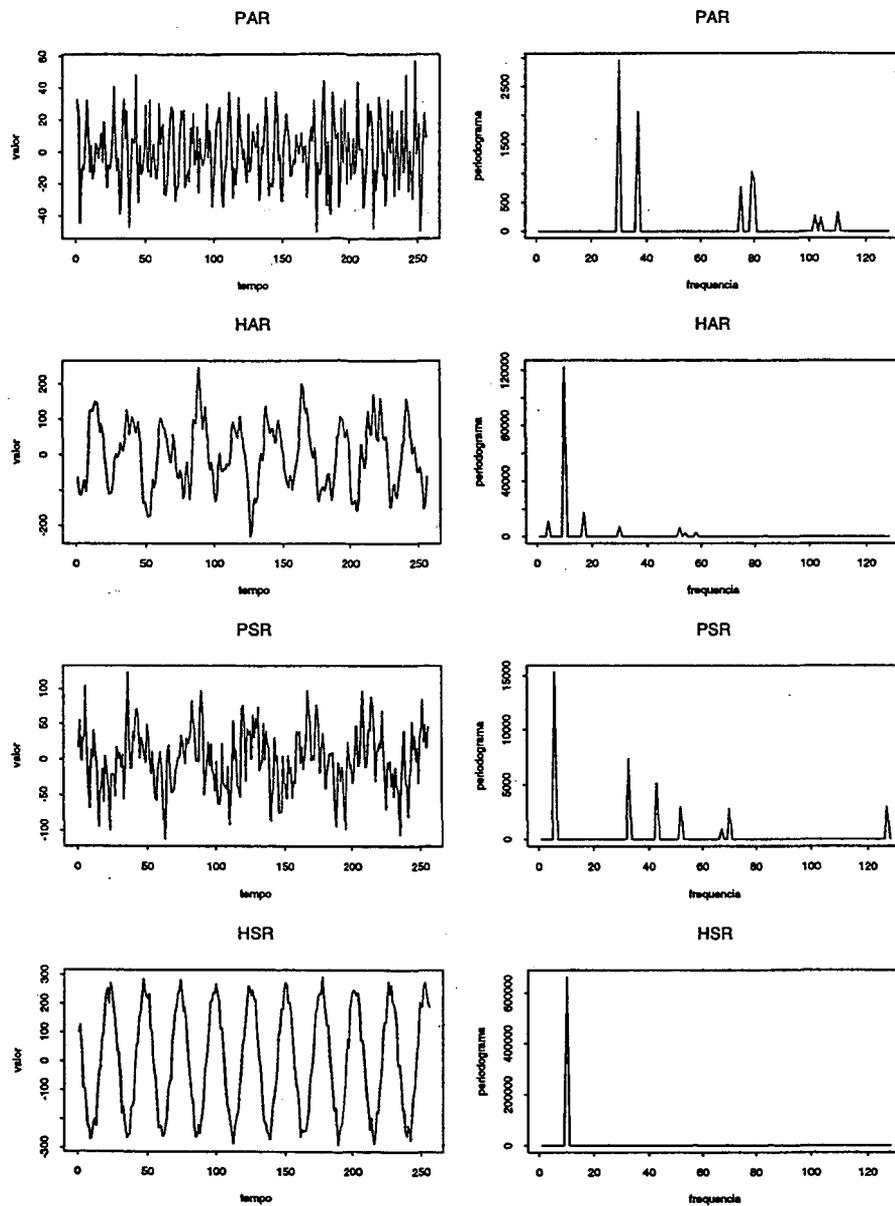
Com base na Tabela 12 construímos a “série comum” de cada um dos grupos, isto é, a série que só contém as freqüências comuns em cada um deles.

Os resultados e os periodogramas correspondentes a essas séries comuns estão na Figura 5. De uma forma geral, podemos notar que os maiores valores dos periodogramas estão localizados em baixas freqüências, estatisticamente isto significa que as baixas freqüências explicam a maior parte da variabilidade dos sinais comuns obtidos.

Tabela 12: Freqüências comuns

grupo	índice	freqüência	período	F	P
PAR	30	0,736	0,057	6,32006	0,0083
	37	0,908	0,046	3,87967	0,0397
	75	1,841	0,023	4,14205	0,0331
	79	1,939	0,022	6,67442	0,0068
	80	1,963	0,021	5,32180	0,0153
	102	2,503	0,017	3,98325	0,0370
	104	2,552	0,016	3,60251	0,0483
	110	2,699	0,015	8,82965	0,0021
HAR	4	0,098	0,427	7,24691	0,0049
	10	0,245	0,171	3,87470	0,0399
	17	0,417	0,100	6,09668	0,0095
	30	0,736	0,057	3,96384	0,0375
	52	1,276	0,056	9,00219	0,0020
	58	1,423	0,029	5,99777	0,0101
	84	2,061	0,020	4,04116	0,0355
	99	2,429	0,017	4,14768	0,0330
PSR	6	0,147	0,283	5,77492	0,0115
	33	0,810	0,052	3,60931	0,0481
	43	1,055	0,040	6,32910	0,0083
	52	1,276	0,033	3,79633	0,0421
	67	1,644	0,025	6,81400	0,0063
	70	1,718	0,024	3,95908	0,0376
	127	3,117	0,013	4,58574	0,0246
HSR	10	0,245	0,170	4,37215	0,0283
	57	1,399	0,030	3,64982	0,0467
	85	2,086	0,020	4,25211	0,0307
	86	2,110	0,020	4,16119	0,0327
	106	2,601	0,016	3,70882	0,0448
	110	2,699	0,015	4,61495	0,0241
	120	2,945	0,014	4,44101	0,0271

Figura 5 - Série comum e periodograma para cada um dos grupos ao nível de 0,05.



Nota: Para maiores detalhes, ver Chiann(1993).

## 6. CONCLUSÕES

As principais conclusões da análise realizada na seção anterior são:

a) Existe interação entre os fatores região do cérebro e tipo de sono, isto é, o comportamento do eletroscilograma dos ratos numa determinada região (posteriotálamo ou hipotálamo) depende do estado de sono em que eles se encontram (alerta ou sono paradoxal);

b) Existe diferença, entre as regiões do cérebro, no comportamento dos eletroscilogramas durante o sono paradoxal e, também, durante o estado de alerta;

c) Foi detectado um comportamento comum dos eletroscilogramas para os ratos pertencentes a cada um dos grupos. Os grupos PAR e HSR têm um forte componente periódico de 0,016 segundo comum a todos os ratos de cada um dos grupos. No grupo HAR, o comportamento periódico mais forte e comum a todos os ratos é de 0,033 segundo e no grupo PSR de 0,025 segundo; e

d) Os grupos PAR e HSR têm componentes comuns concentrados em altas freqüências, ou seja, de períodos curtos, enquanto que os grupos HAR e PSR apresentam componentes comuns em baixas freqüências, indicando períodos mais longos nos eletroscilogramas.

## REFERÊNCIAS

- BRILLINGER, D.R. (1973). The Analysis of Time Series Collected in an Experimental Design, in *Multivariate Analysis III*, ed. P.R. Krishnaiah, New York: Academic Press, pp. 241-256.
- BRILLINGER, D.R. (1980). Analysis of Variance Problems under Time Series Models, in *Handbook of statistics* (Vol. 1), ed. P.R. Krishnaiah, Amsterdam: North Holland, pp. 237-278.
- BRILLINGER, D.R. (1981). *Time Series: Data Analysis and Theory* (2nd ed.), San Francisco: Holden-Day.
- BRILLINGER, D.R. (1985). A Maximum Likelihood Approach to Frequency-Wavenumber Analysis, *IEEE Trans. on Acoustics, Speech and Signal Processing*, ASSP-33, pp. 1076-1085.
- CHIANN, C. (1993). *Análise de Variância em Séries Temporais*, Dissertação, Departamento de Estatística, USP.
- HANNAN, E.J. (1970). *Multiple Time Series*, New York: John Wiley.
- ROY, S.N., GNANADESIKAN, R. AND SRIVASTAVA, J.N. (1971). *Analysis and Design of Certain Quantitative Multiresponse Experiments*, Pergamon, Oxford.
- SCHEFFÉ, H. (1959). *The Analysis of Variance*, New York: John Wiley.
- SEARLE, S.R. (1971). *Linear Models*, New York: John Wiley.
- SHUMWAY, R.H. (1970). Applied Regression and Analysis of Variance for Stationary Time Series, *Journal of the American Statistical Association* 65, pp. 1527-1546.
- SHUMWAY, R.H. (1988). *Applied Statistical Time Series Analysis*, Englewood Cliffs, NJ: Prentice-Hall.
- SKOROKHOD, A.V. (1956). Limit Theorems for Stochastic Processes, *Theory Prob. Appl.* 1, pp. 261-290.

**ABSTRACT**

In this work the analysis of variance is employed to study the variation present in data having the character that the basic responses measured are time series. Our objective consists of the application of the conventional techniques to the Fourier transform of the data. The models covered in the paper include: case of a common signal, one-way classification and the design model. The cases of both fixed and random effects are considered. An application with a real series is given.

# Revista Brasileira de Estatística - RBEs

## Política Editorial

A Revista Brasileira de Estatística - RBEs -, objetiva promover a Estatística relevante para aplicação em questões sociais, interpretadas amplamente para incluir questões educacionais, de saúde, demográficas, econômicas, legais, de políticas públicas e de estatísticas oficiais, entre outras. A revista pretende apresentar artigos num formato que permita fácil assimilação pelos membros da comunidade científica em geral. Os artigos devem incluir aplicações práticas como assunto central. Essas aplicações deverão ter conteúdo estatístico substancial. As análises deverão ser exaustivas e bem apresentadas, mas o emprego de métodos estatísticos inovadores não é essencial para publicação.

Artigos contendo exposição de métodos são aceitáveis, desde que estes sejam relevantes para as áreas cobertas pela revista, auxiliem na compreensão do problema e contenham interpretação clara das expressões matemáticas apresentadas. A apresentação de aplicações ilustrativas envolvendo dados adequados é requerida. Tratamentos algébricos extensos devem ser evitados.

A RBEs tem periodicidade semestral e publicará também artigos escritos a convite e resenhas de livros, bem como artigos abordando os diversos aspectos de metodologias relevantes para órgãos produtores de estatísticas, incluindo:

- a) planejamento de pesquisas;
- b) avaliação e mensuração de erros em pesquisas;
- c) uso e combinação de fontes alternativas de informação e integração de dados;
- d) novos desenvolvimentos em metodologia de pesquisa;
- e) crítica e imputação de dados;
- f) amostragem e estimação;
- g) disseminação e confiabilidade de dados;
- h) análise de dados;
- i) análise de séries temporais;
- j) modelos e métodos demográficos; e
- k) modelos e métodos econométricos.

Todos os artigos submetidos serão avaliados quanto à qualidade e relevância por dois especialistas indicados pelo Comitê Editorial da RBEs. Os artigos submetidos deverão ser inéditos e não deverão ter sido, simultaneamente, submetidos a qualquer outro periódico nacional. O processo de avaliação é do tipo duplo cego, isto é, os artigos são avaliados sem identificação da autoria, e os comentários dos avaliadores também são repassados aos autores sem identificação.

## INSTRUÇÕES PARA SUBMISSÃO DE ARTIGOS À RBEs

Os artigos submetidos para publicação deverão ser remetidos em 3 vias (que não serão devolvidas) para:

Pedro Luis do Nascimento Silva  
Editor Responsável  
Revista Brasileira de Estatística - RBEs  
Av. República do Chile, 500 10º andar  
Rio de Janeiro – RJ – 20031-170  
Tel.: +55 - 21 - 514 0470  
Fax: +55 - 21 - 514 4785  
E-mail: pedrosilva@ibge.gov.br

Os artigos submetidos à RBEs não devem ter sido publicados ou estar sendo considerados para publicação em outros periódicos.

Para cada artigo publicado, serão fornecidas gratuitamente 20 separatas.

### Instruções para preparo de originais:

1. A primeira página do original (folha de rosto) deve conter o título do artigo, seguido do(s) nome(s) completo(s) do(s) autor(es), indicando-se para cada um a filiação e endereço para correspondência. Agradecimentos a colaboradores, a instituições e auxílios recebidos devem figurar também nesta página;
2. A segunda página do original deve conter resumos em português e em inglês (*Abstract*), destacando os pontos relevantes do artigo. Cada resumo deve ser datilografado seguindo o mesmo padrão do restante do texto, em um único parágrafo, sem fórmulas, com no máximo 150 palavras;
3. O artigo deve ser dividido em seções numeradas progressivamente, com títulos concisos e apropriados. Todas as seções e subseções devem ser numeradas e receber título apropriado.
4. A citação de referências no texto e a listagem final das referências devem ser feitas de acordo com as normas da ABNT;
5. As tabelas e gráficos devem ser precedidos de títulos que permitam perfeita identificação do conteúdo. Devem ser numeradas seqüencialmente (Tabela 1, Figura 3, etc.) e referidas nos locais de inserção pelos respectivos números. Quando houver tabelas e demonstrações extensas ou outros elementos de suporte, podem ser empregados apêndices. Os apêndices devem ter título e numeração, tais como as demais seções do trabalho;
6. Gráficos e Diagramas para publicação devem ser incluídos nos arquivos com os originais do artigo, sempre que possível. Quando isto não ocorrer, devem ser traçados em papel branco, com nitidez e boa qualidade, para permitir que a redução seja feita mantendo qualidade. Fotocópias não serão aceitas. É fundamental que não existam erros quer no desenho, quer nas legendas ou títulos; e
7. Serão preferidos originais processados pelo editor de texto Word for Windows.

Se o assunto é Brasil,  
procure o IBGE

<http://www.ibge.gov.br>

<http://www.ibge.org>

---

atendimento  
0800 21 81 81