

Abordagem Clássica e Bayesiana em Modelos
Auto-regressivos com transformações de Box
& Cox

Juana Rosa Lindo Samaniego

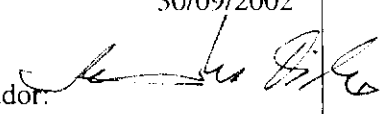
Orientador:

Prof. Dr. *Marinho Gomes de Andrade Filho*

*Dissertação apresentada ao Instituto de Ciências Matemáticas e de Computação -
ICMC-USP, como parte dos requisitos para obtenção do título de Mestre em
Ciências de Computação e Matemática Computacional.*

USP – São Carlos
Novembro/2002

“VERSÃO REVISADA APÓS A DEFESA”

Data da Defesa:	30/09/2002
Visto do Orientador:	

ÍNDICE

CAPÍTULO 1 : Introdução.	1
CAPÍTULO 2 : Abordagem Clássica e Bayesiana na estimação dos parâmetros em processos auto-regressivos.	4
2.1 Introdução.	4
2.2 Processo Auto-regressivo AR(p).	5
2.3 Priori não informativa de Jeffreys para o modelo AR(p).	7
2.4 Priori conjugada normal gama para o modelo AR(p).	10
2.5 Priori T-Gama para o modelo AR(p)	14
2.6 Verossimilhança exata vs Verossimilhança aproximada Processo Auto-regressivo AR(1).	16
2.6.1 Verossimilhança Exata.	16
2.6.2 Verossimilhança Aproximada.	17
2.6.3 Verossimilhança Exata vs. Verossimilhança Aproximada.	18
2.7 Priori não informativa de Jeffreys para o modelo AR(1).	18
2.8 Priori conjugada normal gama para o modelo AR(1).	21
2.9 Priori T-Gama par modelo AR(1).	24
2.10 Critérios de Seleção de modelos.	25
2.10.1 Critérios AIC e BIC.	25
2.10.2 Fator de Bayes.	26
2.11 Previsão k passos a frente.	27
2.11.1 Previsão usando MCMC.	29
CAPÍTULO 3 : Abordagem Clássica e Bayesiana na estimação do parâmetro de transformação λ em processos auto-regressivos de primeira ordem.	31
3.1 Introdução.	31
3.2 Tratamento Bayesiano nas transformações de Box & Cox no processo AR(1).	31
3.2.1 Priori no informativa para o modelo AR(1)	35
3.3 Alternativas para a obtenção da distribuição de (λZ) .	38
3.3.1 Mediante a condicional de $(\lambda Z, \phi)$.	38
3.3.2 Mediante a marginal de (λZ) .	39
3.3.3 Mediante a moda a posteriori - Análise Clássica e Bayesiana.	39
3.4 Transformação de Box e Cox para modelos AR(p).	41

3.4.1	Transformações de Box e Cox para séries de vazões.	44
3.4.2	Previsão em transformações de Box e Cox.	45
3.5	Tratamento clássico em outros tipos de transformação em processos Auto-regressivos.	50
3.5.1	Propriedades da função de Auto-correlação em séries transformadas: Processos Estacionários.	50
3.5.2	Propriedades da função de Auto-correlação em séries transformadas: Processos não Estacionários.	54
CAPÍTULO 4 : Aplicações.		58
4.1	Introdução.	58
4.2	Volume de gás oxigênio em pacientes com queimaduras severas.	58
4.2.1	Resultados obtidos usando priori não informativa.	59
4.2.2	Resultados obtidos usando priori Normal- Gama.	60
4.2.3	Conclusões.	62
4.3	Vazões médias mensais no Reservatório de Furnas - Série Original	62
4.3.1	Resultados obtidos usando priori não informativa.	64
4.3.2	Resultados obtidos usando priori Normal- Gama.	65
4.3.3	Resultados obtidos usando priori T- Gama.	65
4.3.4	Conclusões.	66
4.4	Vazões médias mensais no reservatório de Furnas - Série transformada.	66
4.4.1	Alternativa Bayesiana para encontrar a marginal de λ .	67
4.4.2	Estimadores clássicos para a série transformada	69
4.4.3	Conclusões.	70
CAPÍTULO 5 : Conclusões e Propostas futuras.		71
5.1	Introdução.	71
5.2	Conclusões.	71
5.3	Propostas.	73
ANEXO A :		74
Distribuição a priori de Jeffreys.		74
ANEXO B :		76
Propriedades dos Polinômios Hermite.		76
REFERÊNCIAS BIBLIOGRÁFICAS.		79

RESUMO

Atualmente as projeções de demanda e ganho tornam-se variáveis importantes no processo de tomada de decisões para investimentos envolvendo custo e capital, em pesquisa de mercado envolvendo produtos de consumo, em pesquisa de populações e em qualquer outro tipo de previsões que tenham a ver com ganhos ou demandas futuras, por exemplo o volume de água que é preciso para ser gerada determinada quantidade de energia consumido por uma população através de um sistema de operação e planejamento de um sistema hidroelétrico, etc.

Em resposta desse interesse muitos estudos examinaram a possibilidade de gerar previsões usando séries temporais, ajustando modelos mediante a metodologia de Box e Jenkins, porém estas séries sugeriam variabilidade maior em diferentes níveis, violando deste modo a suposição de variância constante na formulação dos modelos ARIMA. Nestas situações, é comum na prática, contemplar uma extensão destes modelos, assumindo que alguma transformação da série obedeça um modelo ARIMA, freqüentemente são usadas transformações de Box e Cox, porém as previsões destas séries transformadas afeta as interpretações em quanto à série original. Uma abordagem combinada de métodos clássicos e bayesianos é apresentada no tratamento destas transformações, os quais estimam junto com os parâmetros do modelo a potência desta transformação, apresentamos também uma alternativa para examinar a estrutura das auto-covariâncias através do Polinômio de Hermite. A pergunta que surge é, se a incorporação destas transformações resulta numa melhora nas previsões. No caso particular apresentamos resultados em processos auto-regressivos. É feita uma aplicação destes métodos em séries de vazões médias mensais no Reservatório de Furnas.

ABSTRACT

Nowadays the demand and gain projections become important variables in the process of making decisions for investments involving cost and capital, regarding the market research involving consuming products, the population research and any other forecast research which deals with the earnings or the future demands as an example, the water volume which is necessary to generate a determined amount of energy to be consumed by a population through the operation and planning system of a hydroelectric system and so on.

In order to answer this demand a lot of studies examined the possibility to generate forecasts by using the time series, and by adjusting the models used in the Box and Jenkins methodology, however, these series suggested a larger variability in different levels, and therefore violating the constant variance supposition in the ARIMA models formulation. Considering these situations, it is common in the practice to contemplate an extension of these models, assuming that some of these series transformation will follow the ARIMA model. Frequently the Box and Cox transformations are used; however, the forecasts of these transformed series affects the interpretation regarding the original series. An approach combining the classical and bayesian methods is introduced to the consideration of these transformation, which allows us to estimate, along with the parameters of the model, the power of this transformation. Also, we present an option to examine the structure of the auto-covariances through the Hermite polynomials. The question that arises is, if the incorporation of these transformations will result in an improvement in the forecasts. Considering this particular case we present results in the auto-regressive processes. An application of these methods is made in a regular flow series measured monthly at Reservoir of Furnas.



CAPÍTULO 1

INTRODUÇÃO

A metodologia descrita por Box & Jenkins (1970) para ajustar modelos ARIMA foi amplamente usada numa grande variedade de aplicações, embora a possibilidade de utilizar transformações de potência era muito brevemente descrito. O interesse difundido pelo uso de transformações na construção de modelos ARIMA foi estimulado pelos resultados de Chatfield e Prothero (1973), estes autores produziram modelos com transformação Logaritmo obtendo previsões de venda insatisfatórias, surgindo muitas discussões deste artigo, principalmente por Box e Jenkins (1976), que discutiram a possibilidade de obter um modelo com erros de variância homogênea ajustando modelos ARIMA não necessariamente à série original $\{Z_t\}$ mas sim para uma série derivada de um conjunto de membros de uma classe de transformações de potência analisadas por Box e Cox (1964) :

$$\begin{aligned} Z_t^{(\lambda)} &= (Z_t^\lambda - 1) / \lambda & \lambda \neq 0 \\ &= \log Z_t & \lambda = 0 \end{aligned} \quad (1.1)$$

O parâmetro de transformação λ pode ser estimado juntamente com os coeficientes do modelo ARIMA usando a função de máxima verossimilhança exata. Ansley e Newbold (1980) indicam que o uso da função de máxima verossimilhança exata produz melhores estimadores que o método de mínimos quadrados descrito por Box e Jenkins (1976) particularmente em séries com pequeno comportamento sazonal.

Alguns aspectos teóricos do ajuste de modelos ARIMA com transformações de potência foram discutidos por Granger e Newbold (1976), principalmente na aplicação da transformação inversa para as previsões. Em particular, estes autores mostraram que o problema de selecionar um modelo inicial para sua análise posterior mediante modelos ARIMA é complicado, pelo fato que a estrutura da autocorrelação da série transformada $\{Z_t^{(\lambda)}\}$, depende da escolha do parâmetro de transformação λ . Assim por exemplo, se a autocorrelação amostral dos dados originais são empregados para ajustar um modelo ARIMA específico, a escolha do modelo pode não ser adequada para descrever as propriedades lineares de $\{Z_t^{(\lambda)}\}$ para um λ apropriado. Um exemplo numérico em séries temporais macroeconômicas dado por Nelson e Granger (1979) mostra a importância prática deste ponto.

Recentemente, há um crescente interesse em métodos estatísticos bayesianos, a sua aplicação tem se difundido para diversas áreas como epidemiologia, bioestatística, engenharia, ciências da computação e séries temporais. Além disso, uma das áreas de mais rápido

crescimento tem sido os métodos de simulação Monte Carlo em Cadeias de Markov na pesquisa e análise de séries temporais, sendo mais comum em modelos não lineares. A abordagem bayesiana considera o uso de densidades de probabilidade para representar informações a priori sobre os parâmetros do modelo.

Box e Jenkins (1976) e Zellner (1971) foram uns dos primeiros a aplicar a abordagem bayesiana para a estimação dos parâmetros de processos auto-regressivos em modelos ARIMA, porém, a maioria da literatura bayesiana em séries temporais focalizou mais a sua atenção na estimação dos parâmetros de processos auto-regressivos como foi indicado por Broemeling e Shaarawy (1988). Harrison e Stevens (1976) apresentam previsões calculadas por métodos bayesianos em modelos lineares dinâmicos. Um desenvolvimento mais aprofundado nesta classe de modelos é encontrado em West e Harrison (1989). As previsões baseadas em modelos auto-regressivos assim como os critérios de seleção baseados na densidade preditiva é estudada por Schervish e Tsay (1988). Não foi encontrada nenhuma literatura que trate estes modelos com transformações de potência.

No Capítulo 2 é descrito principalmente o tratamento bayesiano dos processos auto-regressivos sem considerar as transformações de potência. É determinada a função de verossimilhança exata e aproximada dos dados, são calculados os estimadores de máxima verossimilhança dos parâmetros do modelo. Na abordagem bayesiana são especificadas três densidades a priori para os parâmetros do modelo, a priori não informativa de Jeffreys, a priori conjugada normal-gama e a priori t-gama, são construídas as densidade a posteriori e preditiva para estas prioris, trata-se particularmente com processos auto-regressivos de primeira ordem. É feita a seleção do modelo via métodos clássicos como AIC e BIC, bem como através de métodos bayesianos, via fator de Bayes. São estimados os parâmetros do modelo mediante técnicas bayesianas usando algoritmos MCMC.

No Capítulo 3 é feita uma análise análoga ao Capítulo 2 considerando agora um parâmetro a mais no processo auto-regressivo, o parâmetro de transformação λ . É considerada uma distribuição a priori não informativa de Jeffreys para os parâmetros do processo. Neste ponto existe a dificuldade de encontrar a distribuição a posteriori para λ , tendo esta uma expressão muito complexa. É apresentada uma alternativa que combina métodos clássicos e bayesianos na solução deste problema. É feita uma análise da estrutura das autocovariancias através do Polinômio de Hermite supondo que a série aceita uma classe geral de transformações.

No Capítulo 4 fazemos uma aplicação da teoria desenvolvida nos capítulos anteriores. Tratamos com uma série do volume de oxigênio desprendido por pacientes queimados, este volume mede o grau de estresse destes pacientes Esta primeira série não precisa de transformações, porém é um bom exemplo para ilustrar o Capítulo 2. Para ilustrar o Capítulo 3

tratamos com uma série de vazões médias mensais no reservatório de Furnas, estas vazões permitem estimar custos ligados ao processamento de energia elétrica, esta série apresenta o problema de variância instável em diferentes níveis da série. Ajustamos modelos para a série transformada e para a série original e fazemos uma comparação dos resultados, são aplicados os critérios para a seleção do modelo em cada caso e feita uma comparação para ambas as séries.

No Capítulo 5 temos as conclusões do trabalho e são propostos futuros desenvolvimentos para trabalhos posteriores, Propomos a possibilidade de trabalhar com outras prioris para a série transformada, como a priori Normal-Gama e a priori t-Gama, propomos também desenvolver técnicas de previsão segundo uma abordagem bayesiana para a série transformada e generalizar os resultados para processos auto-regressivos de ordem p . Tratamos também sobre a possibilidade de considerar erros não gaussianos nos processos auto-regressivos, por exemplo, erros com distribuição exponencial, gama e log-normal, baseados em distribuições condicionais.

CAPÍTULO 2

ABORDAGEM CLÁSSICA E BAYESIANA NA ESTIMAÇÃO DOS PARÂMETROS EM PROCESSOS AUTO-REGRESSIVOS.

2.1 INTRODUÇÃO :

Apresentamos neste capítulo, o desenvolvimento de métodos clássicos e bayesianas para processos auto-regressivos de ordem p , $AR(p)$, em particular para o processos auto-regressivos $AR(1)$. Em geral, esta abordagem considera o uso de densidades de probabilidade para representar informações a priori sobre os parâmetros do modelo. O objetivo deste capítulo é utilizar a abordagem bayesiana e técnicas de simulação de Monte Carlo em Cadeias de Markov (MCMC) para estimar parâmetros, selecionar modelos, fazer previsões e avaliar a influência do parâmetro de localização das densidades a priori na estimativa a posteriori dos parâmetros dos modelos $AR(p)$.

Em casos mais gerais como fazemos aqui, a influência dos parâmetros de localização das densidades a priori só pode ser avaliada usando técnicas de simulação MCMC. Os resultados em [Barreto e Andrade, 2000] mostram que o uso de densidades a priori t-student tornam a abordagem bayesiana mais robusta com relação aos parâmetros de locação das prioris. Apresentamos também vários critérios de seleção de modelos como o fator de Bayes, observando como tais critérios podem ser utilizados na análise de modelos de séries temporais.

É feita uma comparação das estimativas clássicas e bayesianas usando a verossimilhança exata e a verossimilhança aproximada, observando-se que pode ser usada a verossimilhança aproximada dos dados ao invés da verossimilhança exata tendo resultados similares, simplificando assim vários cálculos. Três alternativas para as densidades a priori dos parâmetros do modelo serão estudadas. Primeiro, adota-se uma densidade não informativa de Jeffreys para o modelo $AR(p)$. Em seguida, usa-se uma priori conjugada normal-gama. Neste caso, o estimador bayesiano para os parâmetros podem ser calculados explicitamente (Broemeling e Land, 1984). Na terceira alternativa, utiliza-se uma priori informativa t-Student, p -dimensional, para os parâmetros ϕ 's e uma densidade a priori gama para o inverso da variância dos resíduos $\tau = 1/\sigma^2$. Esta escolha resulta em uma densidade a posteriori não padrão, a qual só permite uma análise a posteriori por meio de simulação de algoritmos de Monte Carlo em cadeias de Markov

(MCMC). Aqui serão adotados os algoritmos Amostrador de Gibbs (Casella & George, 1992) e Metropolis-Hastings (Chib & Greenberg, 1995).

Os resultados desse capítulo mostram as vantagens e flexibilização do processo de inferência bayesiana pelo uso de técnicas de simulação MCMC.

2.2 PROCESSO AUTO-REGRESSIVO - AR(p) :

O modelo AR(p) é dado por: $Z_t = \phi_1 Z_{t-1} + \dots + \phi_p Z_{t-p} + a_t$.

Seja Z_t , $t = 1, 2, \dots, n$ dados observados da série e denotemos o vetor de parâmetros por:

Z_t = Variável aleatória observável

ϕ = Parâmetro não observável do processo

a_t = Ruído branco não observável, independentes com $V(a_t) = \sigma^2$

$\tau = \sigma^{-2}$ (precisão)

$$\phi \in R^p, \quad \phi = (\phi_1, \dots, \phi_p)^T$$

onde a_t é um ruído iid $N(0, \tau^{-1})$

usando a notação vetorial.

$$\underline{Z}_n = (Z_1, Z_2, \dots, Z_n)^T$$

$$\underline{Z}_p = (Z_1, Z_2, \dots, Z_p)^T$$

$$\underline{Z} = (Z_{p+1}, Z_{p+2}, \dots, Z_n)^T$$

A função de verossimilhança condicionada nos p-primeiros valores é dada por :

$$f(Z_{p+1}, \dots, Z_n) = f(Z_{p+1} | Z_p) f(Z_{p+2} | Z_{p+1}) \dots f(Z_n | Z_{n-1})$$

Então podemos escrever :

$$Z_{p+1} = \phi_1 Z_p + \dots + \phi_p Z_1 + a_{p+1}$$

$$Z_{p+2} = \phi_1 Z_{p+1} + \dots + \phi_p Z_2 + a_{p+2}$$

⋮

$$Z_n = \phi_1 Z_{n-1} + \dots + \phi_p Z_{n-p} + a_n$$

Sabemos que para estes modelos: $a_t = Z_t - \phi_1 Z_{t-1} - \dots - \phi_p Z_{t-p}$, $t=p+1, \dots, n$ é uma transformação com jacobiano unitário, temos:

$$f(Z_{p+1}, \dots, Z_n | Z_1, \dots, Z_p) = \prod_{t=p+1}^n \frac{\tau^{1/2}}{\sqrt{2\pi}} \exp\left(-\frac{\tau}{2} (Z_t - \phi_1 Z_{t-1} - \dots - \phi_p Z_{t-p})^2\right) \quad (2.1)$$

$$\Rightarrow f(\underline{Z} | \underline{\phi}, \tau) \propto \tau^{(n-p)/2} \exp\left\{-\frac{\tau}{2} \sum_{t=p+1}^n (Z_t - \phi_1 Z_{t-1} - \dots - \phi_p Z_{t-p})^2\right\} \quad (2.2)$$

Usando a notação :

$$\text{Seja: } \underline{Z} = \begin{pmatrix} Z_{p+1} \\ \vdots \\ \vdots \\ Z_n \end{pmatrix}, \quad \underline{X} = \begin{pmatrix} Z_p & \dots & \dots & Z_1 \\ Z_{p+1} & \dots & \dots & Z_2 \\ \vdots & \dots & \dots & \vdots \\ Z_{n-1} & \dots & \dots & Z_{n-p} \end{pmatrix}, \quad \underline{\phi} = \begin{pmatrix} \phi_1 \\ \vdots \\ \vdots \\ \phi_p \end{pmatrix}$$

$$\text{Assim temos que : } \sum_{t=p+1}^n (Z_t - \phi_1 Z_{t-1} - \dots - \phi_p Z_{t-p})^2 = (\underline{Z} - \underline{X} \underline{\phi})^T (\underline{Z} - \underline{X} \underline{\phi})$$

rescrevendo a função de verosimilhança na forma matricial :

$$f(\underline{Z} | \underline{\phi}, \tau) \propto \tau^{(n-p)/2} \exp\left\{-\frac{\tau}{2} (\underline{Z} - \underline{X} \underline{\phi})^T (\underline{Z} - \underline{X} \underline{\phi})\right\} \quad (2.3)$$

Podemos também representar está função por :

$$L(\underline{\phi}, \tau) \propto \tau^{(n-p)/2} \exp\left\{-\frac{\tau}{2} (\underline{Z} - \underline{X} \underline{\phi})^T (\underline{Z} - \underline{X} \underline{\phi})\right\} \quad (2.4)$$

Aplicando o logaritmo temos que :

$$l(\underline{\phi}, \tau) = \ln\left(L(\tau, \underline{\phi})\right) \propto \frac{n-p}{2} \ln \tau - \frac{\tau}{2} (\underline{Z} - \underline{X} \underline{\phi})^T (\underline{Z} - \underline{X} \underline{\phi}) \quad (2.5)$$

Derivando $l(\underline{\phi}, \tau)$ em relação aos parâmetros τ e $\underline{\phi}$ e igualando as expressões a 0 temos que :

$$\frac{\partial l(\underline{\phi}, \tau)}{\partial \underline{\phi}} = -\frac{\tau}{2} \left[\underline{X}^T (\underline{Z} - \underline{X} \underline{\phi}) - (\underline{Z} - \underline{X} \underline{\phi})^T \underline{X} \right]$$

$$\frac{\partial l(\underline{\phi}, \tau)}{\partial \tau} = \frac{n-p}{2} \frac{1}{\tau} - \frac{1}{2} \left(\underline{Z} - \underline{X} \underline{\phi} \right)' \left(\underline{Z} - \underline{X} \underline{\phi} \right) = 0$$

Portanto os estimadores de máxima verossimilhança são :

$$\hat{\underline{\phi}} = \left(\underline{X}' \underline{X} \right)^{-1} \left(\underline{X}' \underline{Z} \right) \quad (2.6)$$

$$\hat{\tau}^{-1} = \frac{1}{n-p} \left(\underline{Z} - \underline{X} \hat{\underline{\phi}} \right)' \left(\underline{Z} - \underline{X} \hat{\underline{\phi}} \right) \quad (2.7)$$

2.3 PRIORI NÃO INFORMATIVA DE JEFFREYS PARA O MODELO AR(p) :

Na teoria da Inferência bayesiana os parâmetros $\underline{\phi}$ (parâmetro qualquer) de um modelo são considerados variáveis aleatórias para os quais temos alguma informação a priori, resumida na densidade de probabilidade a priori $f_0(\underline{\phi})$.

No contexto da inferência bayesiana determina-se a densidade a posteriori dos parâmetros $\underline{\phi}$ combinando-se a função de verossimilhança com a densidade a priori $f_0(\underline{\phi})$ através do teorema de Bayes, ou seja :

$$f_1(\underline{\phi} | \underline{Z}) = \frac{L(\underline{\phi} | \underline{Z}) f_0(\underline{\phi})}{\int_{\mathcal{B}(\underline{\phi})} L(\underline{\phi} | \underline{Z}) f_0(\underline{\phi}) d\underline{\phi}}$$

A inferência bayesiana para os modelos AR(p) começa pela escolha da densidade a priori $f_0(\underline{\phi}, \tau)$.

Quando as informações sobre os parâmetros são insuficientes podemos adotar uma priori não informativa de Jeffreys. Supondo que $\underline{\phi}$ e τ são independentes e assumindo que $f_0(\underline{\phi} | \tau)$ tem distribuição uniforme (-1,1), podemos escrever a densidade conjunta a priori como :

$$f_0(\underline{\phi}, \tau) \propto 1/\tau, \quad \tau > 0 \quad (2.8)$$

Combinando a função de verossimilhança $L(\underline{\phi}, \tau)$ com a priori de Jeffreys $f_0(\underline{\phi}, \tau)$ temos que :

$$\begin{aligned} f(\underline{\phi}, \tau | Z) &\propto L(\underline{\phi}, \tau) f_0(\underline{\phi}, \tau) \\ &\propto \tau^{(n-p)/2} \exp\left\{-\frac{\tau}{2} \left(\underline{Z} - \underline{X}\underline{\phi}\right)^T \left(\underline{Z} - \underline{X}\underline{\phi}\right)\right\} \tau^{-1} \end{aligned}$$

Utilizando os resultados :

$$\left(\underline{Z} - \underline{X}\underline{\phi}\right)^T \left(\underline{Z} - \underline{X}\underline{\phi}\right) = \left(\underline{Z} - \underline{\hat{Z}}\right)^T \left(\underline{Z} - \underline{\hat{Z}}\right) + \left(\underline{\phi} - \underline{\hat{\phi}}\right)^T \underline{X}^T \underline{X} \left(\underline{\phi} - \underline{\hat{\phi}}\right)$$

onde : $\underline{\hat{Z}} = \underline{X}\underline{\hat{\phi}}$ e $\underline{\hat{\phi}} = (\underline{X}^T \underline{X})^{-1} (\underline{X}^T \underline{Z})$

Temos que a densidade posteriori para $\underline{\phi}$ e τ é :

$$f(\underline{\phi}, \tau | Z) \propto \tau^{\frac{n-p}{2}-1} \exp\left\{-\frac{\tau}{2} \left[\left(\underline{\phi} - \underline{\hat{\phi}}\right)^T (\underline{X}^T \underline{X}) \left(\underline{\phi} - \underline{\hat{\phi}}\right) + \left(\underline{Z} - \underline{\hat{Z}}\right)^T \left(\underline{Z} - \underline{\hat{Z}}\right) \right]\right\} \quad (2.9)$$

Agora vamos obter as densidades condicionais de $\underline{\phi}$ e τ .

Temos que :

$$f(\tau | Z) \propto \int_{\underline{\phi}} \tau^{\frac{n-p}{2}-1} \exp\left\{-\frac{\tau}{2} \left[\left(\underline{\phi} - \underline{\hat{\phi}}\right)^T (\underline{X}^T \underline{X}) \left(\underline{\phi} - \underline{\hat{\phi}}\right) + \left(\underline{Z} - \underline{\hat{Z}}\right)^T \left(\underline{Z} - \underline{\hat{Z}}\right) \right]\right\} d\underline{\phi}$$

$$\propto \tau^{\frac{n-p}{2}-1} \exp\left\{-\frac{\tau}{2} \left(\underline{Z}-\hat{\underline{Z}}\right)^T \left(\underline{Z}-\hat{\underline{Z}}\right)\right\} \int_{\underline{\phi}} \exp\left\{-\frac{\tau}{2} \left[\left(\underline{\phi}-\hat{\underline{\phi}}\right)^T \left(\underline{X}^T \underline{X}\right) \left(\underline{\phi}-\hat{\underline{\phi}}\right)\right]\right\} d\underline{\phi}$$

$$\propto \tau^{\frac{n(n-p)}{2}-1} \exp\left\{-\frac{\tau}{2} \left(\underline{Z}-\hat{\underline{Z}}\right)^T \left(\underline{Z}-\hat{\underline{Z}}\right)\right\}$$

pois : $\int_{\underline{\phi}} \exp\left\{-\frac{\tau}{2} \left[\left(\underline{\phi}-\hat{\underline{\phi}}\right)^T \left(\underline{X}^T \underline{X}\right) \left(\underline{\phi}-\hat{\underline{\phi}}\right)\right]\right\} d\underline{\phi} \propto \tau^{-1}$ é uma integral de uma distribuição Normal multivariada.

Portanto τ tem uma densidade a posteriori:

$$\text{Gama}\left(\frac{n-p}{2}, \left(\underline{Z}-\hat{\underline{Z}}\right)^T \left(\underline{Z}-\hat{\underline{Z}}\right) / 2\right)$$

Considerando o estimador bayesiano como a esperança a posteriori de τ , temos :

$$E(\tau | \underline{Z}) = (n-2p) \left[\left(\underline{Z}-\hat{\underline{Z}}\right)^T \left(\underline{Z}-\hat{\underline{Z}}\right) \right]^{-1}$$

$$\text{Temos também que: } \text{Var}(\tau | \underline{Z}) = 2(n-2p) \left[\left(\underline{Z}-\hat{\underline{Z}}\right)^T \left(\underline{Z}-\hat{\underline{Z}}\right) \right]^{-2}$$

Por outro lado temos que :

$$f(\underline{\phi} | \underline{Z}) \propto \int_{\tau} \tau^{\frac{n-p}{2}-1} \exp\left\{-\frac{\tau}{2} \left[\left(\underline{\phi}-\hat{\underline{\phi}}\right)^T \left(\underline{X}^T \underline{X}\right) \left(\underline{\phi}-\hat{\underline{\phi}}\right) + \left(\underline{Z}-\hat{\underline{Z}}\right)^T \left(\underline{Z}-\hat{\underline{Z}}\right)\right]\right\} d\tau$$

$$\propto \left[\left(\underline{\phi}-\hat{\underline{\phi}}\right)^T \left(\underline{X}^T \underline{X}\right) \left(\underline{\phi}-\hat{\underline{\phi}}\right) + \left(\underline{Z}-\hat{\underline{Z}}\right)^T \left(\underline{Z}-\hat{\underline{Z}}\right) \right]^{-\frac{n-p}{2}} \Gamma\left(\frac{n-p}{2}\right)$$

$$\propto \left[\left(\underline{\phi}-\hat{\underline{\phi}}\right)^T \left(\underline{X}^T \underline{X}\right) \left(\underline{\phi}-\hat{\underline{\phi}}\right) + \left(\underline{Z}-\hat{\underline{Z}}\right)^T \left(\underline{Z}-\hat{\underline{Z}}\right) \right]^{-\frac{n-p}{2}} \quad (2.10)$$

Após algumas simplificações algébricas podemos escrever a densidade a posteriori dos parâmetros $\underline{\phi}$ como :

$$f(\underline{\phi}) \propto \left[1 + \frac{(\underline{\phi} - \hat{\underline{\phi}})^T V (\underline{\phi} - \hat{\underline{\phi}})}{n-2p} \right]^{-\frac{n-p}{2}} = \left[1 + \frac{(\underline{\phi} - \hat{\underline{\phi}})^T V (\underline{\phi} - \hat{\underline{\phi}})}{v} \right]^{-\frac{v+p}{2}}$$

onde $V = (n-2p)(\underline{X}^T \underline{X})^{-1} \left[(\underline{Z} - \hat{\underline{Z}})^T (\underline{Z} - \hat{\underline{Z}}) \right]^{-1}$ e $v = n-2p$

Portanto $\underline{\phi}$ tem densidade a posteriori t-Student, com $v = (n-2p)$ graus de liberdade, vetor de localização $\hat{\underline{\phi}}$ e matriz de precisão V . O estimador bayesiano é o vetor esperado a posteriori de $\underline{\phi}$ ou seja: $E(\underline{\phi} | \underline{Z}) = \hat{\underline{\phi}}$.

Assim o estimador bayesiano para $\underline{\phi}$ coincide com o EMV.

Temos também que: $V(\underline{\phi} | \underline{Z}) = V^{-1} \left(\frac{v}{v-2} \right)$

2.4 PRIORI CONJUGADA NORMAL GAMA PARA O MODELO AR(p) :

A função de verossimilhança sugere o uso de uma densidade a priori conjugada Normal-Gama. Assim temos :

$$f_0(\underline{\phi}, \tau) = f_1(\underline{\phi} | \tau) f_2(\tau)$$

$$f_1(\underline{\phi} | \tau) \propto \tau^{\frac{p}{2}} \exp \left\{ -\frac{\tau}{2} (\underline{\phi} - \underline{\mu})^T P (\underline{\phi} - \underline{\mu}) \right\}$$

$$f_2(\tau) \propto \tau^{\alpha-1} \exp\{-\beta\tau\} \tag{2.11}$$

onde $\underline{\phi} \subset R^p$, $P \subset R^{p \times p}$ e α, β escalares são hiperparâmetros. Combinando a função de verossimilhança com a densidade a priori temos a densidade a posteriori dada por :

$$f(\underline{\phi}, \tau | \underline{Z}) \propto \tau^{\frac{(n-p+2\alpha)}{2}} \exp\{-\tau D\} \tau^{\frac{p}{2}} \exp\left\{-\frac{\tau}{2} \left(\underline{\phi} - \underline{\phi}_b\right)^T V \left(\underline{\phi} - \underline{\phi}_b\right)\right\} \quad (2.12)$$

onde :

$$\underline{\phi}_b = \left(\underline{X}^T \underline{X} + P\right)^{-1} \left(\underline{X}^T \underline{Z} + P \underline{\mu}\right)$$

$$V = \underline{X}^T \underline{X} + P$$

$$D = \beta + \frac{1}{2} \left[\left(\underline{Z}^T \underline{Z} + \underline{\mu}^T P \underline{\mu}\right) - \left(\underline{X}^T \underline{Z} + P \underline{\mu}\right)^T \left(\underline{X}^T \underline{X} + P\right)^{-1} \left(\underline{X}^T \underline{Z} + P \underline{\mu}\right) \right]$$

Devido a conjugação, temos que a densidade conjunta a posteriori é uma Normal-Gama. Assim podemos determinar a densidade a posteriori de τ , integrando com relação aos parâmetros $\underline{\phi}$. Procedendo desta forma, temos :

$$f(\tau) \propto \tau^{\frac{(n-p+2\alpha)}{2}-1} \exp\{-\tau D\} \int \tau^{\frac{p}{2}} \exp\left\{-\frac{\tau}{2} \left[\left(\underline{\phi} - \underline{\phi}_b\right)^T V \left(\underline{\phi} - \underline{\phi}_b\right)\right]\right\} d\underline{\phi}$$

$$f(\tau) \propto \tau^{\frac{(n-p+2\alpha)}{2}-1} \exp\{-\tau D\} \quad (2.13)$$

Temos portanto que o valor esperado e a variância a posteriori de τ são dados por:

$$E(\tau | \underline{Z}) = (n - p + 2\alpha)(2D)^{-1}$$

$$\text{Var}(\tau | \underline{Z}) = E(\tau | \underline{Z})(D)^{-1}$$

Da mesma forma, integrando com relação a τ , temos a densidade marginal a posteriori para os parâmetros $\underline{\phi}$ é dada por :

$$f(\underline{\phi} | \underline{Z}) \propto \int_0^\infty \tau^{\frac{(n+2\alpha)}{2}-1} \exp\left\{-\frac{\tau}{2} \left[\left(\underline{\phi} - \underline{\phi}_b\right)^T V \left(\underline{\phi} - \underline{\phi}_b\right) + 2D\right]\right\} d\tau$$

$$f(\underline{\phi}) \propto \left[2D + \left(\underline{\phi} - \underline{\phi}_b \right)^T V \left(\underline{\phi} - \underline{\phi}_b \right) \right]^{-\frac{n+2\alpha}{2}}$$

reagrupando os termos temos :

$$f(\underline{\phi}) \propto \left[1 + \frac{\left(\underline{\phi} - \underline{\phi}_b \right)^T W \left(\underline{\phi} - \underline{\phi}_b \right)}{n - p + 2\alpha} \right]^{-\frac{n+2\alpha}{2}} \quad (2.14)$$

$$\text{onde } W = (n - p + 2\alpha) V (2D)^{-1}$$

Portanto, vemos que $\underline{\phi}$ tem densidade a posteriori t-student, p-dimensional com $v = (n - p + 2\alpha)$ graus de liberdade, vetor de localização $\underline{\phi}_b$ e matriz de precisão W.

Então podemos estimar os parâmetros $\underline{\phi}$, pelo valor esperado da densidade acima e a matriz de covariância dos parâmetros pode ser obtida diretamente da matriz de precisão W. Assim temos :

$$E(\underline{\phi} | Z) = \left(\underline{X}^T \underline{X} + P \right)^{-1} \left(\underline{X}^T \underline{Z} + P \underline{\mu} \right) \quad (2.15)$$

$$\text{Var}(\underline{\phi} | Z) = \frac{v}{v - 2} W^{-1} = \frac{\left(\underline{X}^T \underline{X} + P \right)^{-1} (2D)}{(n - p + 2\alpha - 2)} \quad (2.16)$$

O valor esperado a posteriori, $\underline{\phi}_b$ é uma média ponderada do estimador de máxima verossimilhança $\hat{\underline{\phi}}$ e da média da densidade a priori, $\underline{\mu}$, ou seja :

$$\underline{\phi}_b = \left(\underline{X}^T \underline{X} + P \right)^{-1} \left(\underline{X}^T \underline{Z} + P \underline{\mu} \right) = \left(\underline{X}^T \underline{X} + P \right)^{-1} \left[\left(\underline{X}^T \underline{X} \right) \hat{\underline{\phi}} + P \underline{\mu} \right]$$

Apresentaremos os procedimentos necessários à aplicação de simulação MCMC para a análise bayesiana.

Algoritmo do Amostrador de Gibbs

O amostrador de Gibbs (Gelfand & Smith, 1990; Casella & George, 1992) é basicamente um esquema iterativo de amostragem de uma cadeia de Markov cujo o núcleo de transição é formado pela densidade conjunta a posteriori. Para o modelo com priori informativa normal-gama, tem-se que a posteriori conjunta é dada por (2.9) :

$$f(\underline{\phi}, \tau | \underline{Z}) \propto \tau^{(n+2\alpha)-1} \exp\left\{-\tau D - \frac{\tau}{2}(\underline{\phi} - \underline{\phi}_b)^T V(\underline{\phi} - \underline{\phi}_b)\right\}$$

As densidades condicionais são facilmente obtidas devido à conjugação:

$$f(\underline{\phi} | \tau, \underline{Z}) \sim N(\underline{\phi}_b, (\tau V)^{-1}) \quad (2.17)$$

$$f(\tau | \underline{\phi}, \underline{Z}) \sim \Gamma\left(\frac{n+2\alpha}{2}, \frac{1}{2}(\underline{\phi} - \underline{\phi}_b)^T V(\underline{\phi} - \underline{\phi}_b)\right) \quad (2.18)$$

onde, $V = \underline{X}^T \underline{X} + P$

$$D = \beta + \frac{1}{2} \left[\left(\underline{Z}^T \underline{Z} + \underline{\mu}^T P \underline{\mu} \right) - \left(\underline{X}^T \underline{Z} + P \underline{\mu} \right)^T \left(\underline{X}^T \underline{X} + P \right)^{-1} \left(\underline{X}^T \underline{Z} + P \underline{\mu} \right) \right]$$

As Equações (2.17) e (2.18) são usadas para gerar os valores dos parâmetros como segue:

Passo 1: Atribua valores iniciais aos parâmetros e inicie o contador de iterações $j=0$.

$$\underline{\phi}^{(0)} = \left(\phi_1^{(0)}, \dots, \phi_p^{(0)} \right) e \tau^{(0)}$$

Passo 2: Obtenha um novo valor $\underline{\phi}^{(j+1)}$ a partir da densidades condicional .

$$\underline{\phi}^{(j+1)} \sim f(\underline{\phi} | \tau^{(j)}, \underline{Z})$$

Passo 3: Obtenha um novo valor $\tau^{(j+1)}$ a partir da densidade condicional.

$$\tau^{(j+1)} \sim f(\tau | \underline{\phi}^{(j+1)}, \underline{Z})$$

Passo 4: Verifique se houve convergência. Caso contrário, faça $j=j+1$ e repita passo 2 .

Após um número suficientemente grande de iterações, as observações geradas $\left(\underset{\sim}{\phi}^{(j)}, \tau^{(j)} \right)$ convergirão para uma amostra aleatória da densidade conjunta $f\left(\underset{\sim}{\phi}, \tau \mid \underset{\sim}{Z}\right)$ a posteriori.

2.5 PRIORI T-GAMA PARA O MODELO AR(p)

Considera-se agora, para o vetor de parâmetros $\underset{\sim}{\phi}$, uma densidade a priori t-Student, p-dimensional, com k graus de liberdade, vetor de localização μ e matriz de precisão P . Para o parâmetro τ , atribui-se uma densidade a priori gama com hiperparâmetros α e β .

Desta forma, tem se:

$$f_0\left(\underset{\sim}{\phi}, \tau\right) = f_1\left(\underset{\sim}{\phi} \mid \tau\right) f_2(\tau) \quad (2.19)$$

onde:

$$f_1\left(\underset{\sim}{\phi}\right) \propto \left[1 + \frac{\left(\underset{\sim}{\phi} - \underset{\sim}{\mu}\right)^T P \left(\underset{\sim}{\phi} - \underset{\sim}{\mu}\right)}{k} \right]^{-\frac{(k+p)}{2}} \quad (2.20)$$

$$f_2(\tau) \propto \tau^{\alpha-1} \exp\{-\beta\tau\} \quad (2.21)$$

Os coeficientes $\underset{\sim}{\phi}$ podem ser considerados independentes com uma matriz de correlação dada por $\sum kP^{-1}/(k-2)$. A priori $f_0\left(\underset{\sim}{\phi}, \tau\right)$ resultante do produto (2.20) e (2.21) será doravante chamada de priori t-gama. Assumindo-se independência entre $\underset{\sim}{\phi}$ e τ , a análise a posteriori é feita multiplicando-se a verossimilhança em (2.3) com as densidades a priori em (2.20) e (2.21):

$$f\left(\underset{\sim}{\phi}, \tau \mid \underset{\sim}{Z}\right) \propto \tau^{\frac{n-p-2\alpha-1}{2}} \exp\left\{-\tau \left[\beta + \frac{B\left(\underset{\sim}{\phi}\right)}{2} \right]\right\} f_1\left(\underset{\sim}{\phi}\right) \quad (2.22)$$

onde $B\left(\underset{\sim}{\phi}\right) = \left(\underset{\sim}{\phi} - \hat{\underset{\sim}{\phi}}\right)^T \underset{\sim}{X}^T \underset{\sim}{X} \left(\underset{\sim}{\phi} - \hat{\underset{\sim}{\phi}}\right) + \left(\underset{\sim}{Z} - \hat{\underset{\sim}{Z}}\right)^T \left(\underset{\sim}{Z} - \hat{\underset{\sim}{Z}}\right)$. Percebe-se que a densidade conjunta a posteriori não tem uma forma padrão. Por tanto, a avaliação das densidades marginais a posteriori pode ser feita por meio de métodos de simulação estocástica do tipo

Monte Carlo em Cadeia de Markov, usando algoritmos como o Amostrador de Gibbs e Metropolis-Hastings. Além disso, não é possível chegar a uma equação semelhante à (2.14) que estabelece a relação entre $\hat{\phi}_b$, $\hat{\phi}$ e $\hat{\mu}$, portanto esta relação será analisada também via simulação MCMC.

Algoritmo do Metropolis – Hasting + Amostrador de Gibbs

Quando as distribuições condicionais não são facilmente identificadas, de forma tal que não podemos gerar amostras dessas distribuições diretamente, utiliza-se o algoritmo Metropolis-Hastings. Para este caso, a posteriori conjunta é dada por (2.22):

$$f(\hat{\phi}, \tau | \hat{Z}) \propto \tau^{\frac{n-p+2\alpha-1}{2}} \exp\left\{-\tau \left[\beta + \frac{B(\hat{\phi})}{2}\right]\right\} f_1(\hat{\phi})$$

onde:

$$B(\hat{\phi}) = (\hat{\phi} - \hat{\phi})^T X^T X (\hat{\phi} - \hat{\phi}) + (\hat{Z} - \hat{Z})^T (\hat{Z} - \hat{Z})$$

$$f_1(\hat{\phi}) \propto \left[1 + \frac{(\hat{\phi} - \hat{\mu})^T P (\hat{\phi} - \hat{\mu})}{k}\right]^{\frac{(k+p)}{2}}$$

As densidades condicionais a posteriori são dadas por:

$$f(\hat{\phi} | \tau, \hat{Z}) \propto \exp\left\{-\frac{\tau}{2} \left[(\hat{\phi} - \hat{\phi})^T (X^T X (\hat{\phi} - \hat{\phi})) \right]\right\} f_1(\hat{\phi}) = \Psi(\hat{\phi}) f_1(\hat{\phi}) \quad (2.23)$$

$$f(\tau | \hat{\phi}, \hat{Z}) \sim \Gamma\left(\frac{n-p+2\alpha}{2}, \beta + \frac{1}{2} B(\hat{\phi})\right) \quad (2.24)$$

Este algoritmo é implementado a seguir:

Passo 1: Atribua valores aos parâmetros e inicie o contador de iterações $j=0$.

$$\hat{\phi}^{(0)} = \left(\hat{\phi}_1^{(0)}, \dots, \hat{\phi}_p^{(0)}\right) e^{\tau^{(0)}}$$

Passo 2: Obtenha novos valores a partir da função de transição: $\phi_{\sim}^{(j+1)} \sim f_1\left(\phi_{\sim}^{(j)}\right)$

Passo 3: Calcule a probabilidade de aceitação do novo valor.

$$\alpha\left(\phi_{\sim}^{(j+1)}, \phi_{\sim}^{(j)}\right) = \begin{cases} \min\left\{1, \frac{\Psi\left(\phi_{\sim}^{(j+1)}\right)}{\Psi\left(\phi_{\sim}^{(j)}\right)}\right\}, & \text{se } \Psi\left(\phi_{\sim}^{(j)}\right) > 0 \\ 1, & \text{c.c.} \end{cases}$$

Passo 4: Gere μ de Uniforme (0,1) e faça $\phi_{\sim}^{(j+1)} = \phi_{\sim}^{(j)}$ se $\mu > \alpha\left(\phi_{\sim}^{(j+1)}, \phi_{\sim}^{(j)}\right)$

Passo 5: Gerar $\tau^{(j+1)} \sim f\left(\tau \mid \phi_{\sim}, Z_{\sim}\right)$ a partir da expressão

Passo 6: Verifique se houve convergência. Caso contrário, faça $j=j+1$ e repita passos 2-6

2.6 VEROSSIMILHANÇA EXATA VS APROXIMADA - PROCESSO AUTO-REGRESSIVO AR(1).

Considere a série temporal $Z_t, t = 1, \dots, n$ gerada por um processo AR(1) :

$$Z_t = \phi Z_{t-1} + a_t, \quad t=1,2,\dots,n \quad (2.25)$$

onde:

Z_t = Variável aleatória observável

ϕ = Parâmetro não observável do processo

a_t = Ruído branco não observável, independentes com $V(a_t) = \sigma^2$

$\tau = \sigma^{-2}$ (precisão)

2.6.1 Verossimilhança Exata :

A Verossimilhança exata para os dados é dada por :

$$f(\underline{Z} \mid \phi, \sigma^2) = \prod_{t=2}^n f(Z_t \mid Z_{t-1}) f(Z_1)$$

temos que :

$$f(\underline{Z} \mid \phi, \sigma^2) \propto \tau^{\frac{(n-1)}{2}} \exp\left\{-\frac{\tau}{2} \left[(1-\phi^2) Z_1^2 + \sum_{t=2}^n (Z_t - \phi Z_{t-1})^2 \right]\right\}$$

Aplicando logaritmo e derivando em relação aos parâmetros e igualando as expressões a 0 temos:

$$\hat{\phi}_{EMV} = \frac{\sum_{t=2}^n Z_t Z_{t-1}}{\sum_{t=2}^n Z_{t-1}^2 - Z_1^2} \quad \hat{\tau}^{-1} = \frac{1}{n-1} \sum_{t=2}^n \left(Z_t - \hat{\phi} Z_{t-1} \right)^2$$

Então a verossimilhança pode ser escrita como :

$$f(\underline{Z} | \phi, \tau) \propto \tau^n \sqrt{1 - \phi^2} \exp \left\{ -\frac{\tau}{2} \left[r + b_1 (\phi - \hat{\phi})^2 \right] \right\} \quad (2.26)$$

onde:

$$b_1 = \sum_{t=2}^n Z_{t-1}^2 - Z_1^2 \quad b_3 = \sum_{t=1}^n Z_t^2$$

$$b_2 = \sum_{t=2}^n Z_t Z_{t-1} \quad \hat{\phi}_{EMV} = \frac{b_1}{b_2} \quad r = b_3 - \frac{b_2^2}{b_1}$$

2.6.2 Verossimilhança Aproximada :

A Verossimilhança aproximada para os dados é dada por :

$$f(\underline{Z} | \phi, \sigma^2) = \prod_{t=2}^n f(Z_t | Z_{t-1})$$

$$f(\underline{Z} | \phi, \sigma^2) = \prod_{t=2}^n \frac{\tau^{1/2}}{\sqrt{2\pi}} \exp \left\{ -\frac{1}{2\sigma^2} (Z_t - \phi Z_{t-1})^2 \right\}$$

$$f(\underline{Z} | \phi, \sigma^2) = \tau^{\frac{(n-1)}{2}} \exp \left\{ -\frac{\tau}{2} \left[\sum_{t=2}^n (Z_t - \phi Z_{t-1})^2 \right] \right\}$$

Após algumas simplificações algébricas obtemos o EMV para ϕ e τ :

$$\hat{\phi}_{EMV} = \frac{\sum_{t=2}^n Z_t Z_{t-1}}{\sum_{t=2}^n Z_{t-1}^2} \quad \hat{\tau}^{-1} = \frac{1}{n-1} \sum_{t=2}^n \left(Z_t - \hat{\phi} Z_{t-1} \right)^2$$

2.6.3 Verossimilhança Exata vs. Verossimilhança Aproximada :

O objetivo desta seção é mostrar que pode ser usada a verossimilhança aproximada dos dados ao invés da verossimilhança exata tendo resultados similares simplificando assim vários cálculos. Na continuação são apresentados os gráficos da função de verossimilhança para ambos dos casos onde pode-se notar que estas no apresentam maiores diferenças.

Figura 1 : Verossimilhança exata

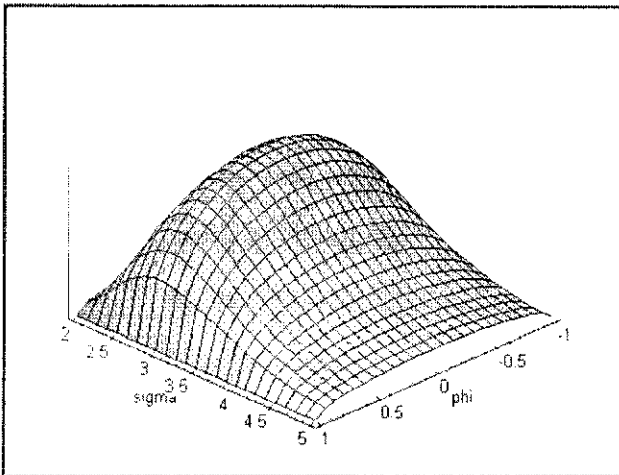
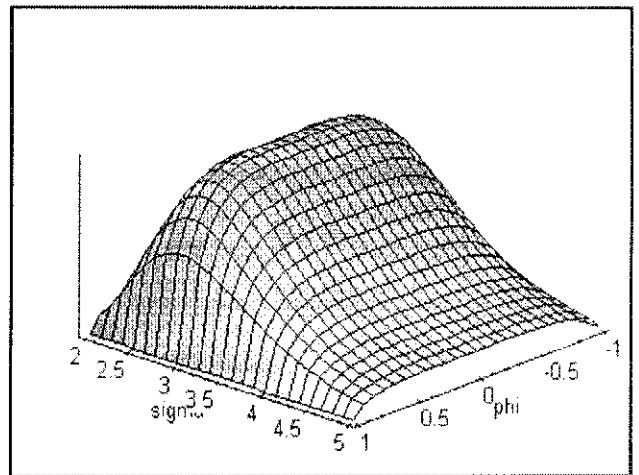


Figura 2 : Verossimilhança Aproximada



2.7 PRIORI NÃO INFORMATIVA DE JEFFREYS PARA O MODELO AR(1) :

Quando temos pouca ou nenhuma informação sobre os parâmetros podemos adotar a priori não informativa de Jeffreys :

$$J = E \left\{ \frac{\partial^2 \log f(Z/\phi, \tau)}{\partial \phi \partial \tau} \right\}$$

$$I = |J|^{-\frac{1}{2}} \quad \text{Priori de Jeffreys}$$

logo :

$$f(\phi, \tau) \propto \tau^{-1} \quad -1 < \phi < 1, \tau > 0$$

$$f(\phi, \tau) = f(\tau | \phi) f(\phi), \text{ onde } f(\phi) \text{ é assumido } U(-1,1)$$

Pelo teorema de Bayes a densidade a posteriori conjunta para (ϕ, τ) é dada por :

$$f(\phi, \tau | Z) \propto f(Z | \phi, \tau) f(\phi, \tau)$$

Então:

$$f(\phi, \tau | Z) \propto \tau^{\frac{n}{2}} \sqrt{1 - \phi^2} \exp\left\{-\frac{\tau}{2} \left[r + b_1 (\phi - \hat{\phi})^2 \right]\right\} \tau^{-1}$$

Por tanto:

$$f(\phi, \tau | Z) \propto \tau^{n-1} \sqrt{1 - \phi^2} \exp\left\{-\frac{\tau}{2} \left[r + b_1 (\phi - \hat{\phi})^2 \right]\right\} \quad (2.27)$$

Precisamos obter uma amostra para essa densidade.

A distribuição condicional de $(\tau | \phi)$ é encontrada fazendo que todo termo que dependa de ϕ seja constante, então :

$$f(\tau | Z, \phi) \propto \tau^{n-1} \exp\left\{-\frac{\tau}{2} \left[r + b_1 (\phi - \hat{\phi})^2 \right]\right\} \quad \tau > 0$$

Podemos identificar que $(\tau | Z, \phi)$ tem distribuição gama :

$$(\tau | Z, \phi) \sim \text{Gama} \left[\frac{n}{2}, \frac{r + b_1 (\phi - \hat{\phi})^2}{2} \right] \quad \tau > 0$$

Agora precisamos da densidade de $(\phi | Z)$, para isso integramos (2.27) com respeito a τ e obtemos :

$$f(\phi | Z) = \int_0^{\infty} f(\phi, \tau | Z) d\tau \propto \int_0^{\infty} \sqrt{1 - \phi^2} \tau^{n-1} \exp\left\{-\frac{\tau}{2} \left[r + b_1 (\phi - \hat{\phi})^2 \right]\right\} d\tau$$

$$= \sqrt{1-\phi^2} \left\{ \frac{2}{r+b_1(\phi-\hat{\phi})^2} \right\}^{\frac{n}{2}} \int_0^\infty \left\{ \frac{r+b_1(\phi-\hat{\phi})^2}{2} \right\}^{\frac{n}{2}} \tau^{\frac{n}{2}-1} \exp\left\{-\frac{\tau}{2}\left[r+b_1(\phi-\hat{\phi})^2\right]\right\} d\tau$$

$$= 1$$

$$\propto \sqrt{1-\phi^2} \left[r+b_1(\phi-\hat{\phi})^2 \right]^{\frac{n}{2}}, |\phi| < 1$$

$$\propto \sqrt{1-\phi^2} \left[1+\frac{b_1}{r}(\phi-\hat{\phi})^2 \right]^{\frac{n}{2}}, |\phi| < 1$$

Podemos reconhecer que o segundo fator da densidade marginal é proporcional a uma densidade t-Student :

$$\left[1+\frac{1}{v}\left(\frac{\phi-\mu}{\sigma}\right)^2 \right]^{-\frac{(v+1)}{2}} \propto \left[1+\frac{b_1}{r}(\phi-\hat{\phi})^2 \right]^{-\frac{n}{2}}$$

Assim :

$$\frac{v+1}{2} = \frac{n}{2} \quad v = n-1$$

$$\frac{1}{v\sigma^2} = \frac{b_1}{r} \quad \sigma^2 = \frac{r}{b_1 v}$$

então:

$$f(\phi | Z) \propto \sqrt{1-\phi^2} t_{n-1}\left(\hat{\phi}, \frac{r}{b_1 v}\right)$$

Para gerar uma amostra da distribuição conjunta $(\phi, \tau | Z)$, aplicamos o algoritmo de Metropolis-Hastings para $f(\phi | Z)$, gerando uma distribuição t-Student com (n-1) g.l e logo multiplicamos a esse valor gerado pelo desvio padrão e somamos a média :

$$t_{n-1, X} \sqrt{\frac{r}{b_1 V} + \hat{\phi}}$$

com este núcleo geramos uma amostra aleatória para ϕ .

Esses valores de ϕ geram valores para τ da:

$$Gama \left[\frac{n}{2}, \frac{r + b_1 (\hat{\phi} - \phi)^2}{2} \right]$$

2.8 PRIORI CONJUGADA NORMAL GAMA PARA O MODELO AR(1) :

A função de verossimilhança (2.26) sugere o uso de uma densidade a priori conjugada Normal-Gama, assim temos :

$$f(\phi, \tau) = f(\phi | \tau) f(\tau)$$

onde :

$$(\phi | \tau) \sim N \left(\hat{\phi}_0, \frac{1}{\tau} \right)$$

$$(\tau) \sim Gama(\alpha, \beta)$$

$\hat{\phi}_0$ é o estimador obtido pelo método dos momentos, ou seja $\hat{\phi}_0 = \rho_1$

Para obter os parâmetros da distribuição Gama achamos um estimador de $\tau_0 = \frac{1}{\sigma^2}$

temos que:

$$V(Z_t) = \sigma_z^2 = \frac{\sigma^2}{1 - \hat{\phi}_0^2} = \frac{\sigma^2}{1 - \rho_1^2} \quad \hat{\tau}_0 = \frac{1}{\sigma^2} = \frac{1}{(1 - \rho_1^2) \sigma_z^2}$$

Assim:

$$E(\tau) = \frac{\alpha}{\alpha + \beta} \approx \hat{\tau}_0 = \frac{1}{(1 - \rho_1^2) \sigma_z^2}$$

Da aqui podemos tirar os valores apropriados para os hiperparâmetros α e β
 Na continuação, temos :

$$f(\phi | \tau) \propto \tau^{\frac{1}{2}} \exp\left\{-\frac{\tau}{2}(\phi - \hat{\phi})^2\right\} \quad \text{e} \quad f(\tau) \propto \tau^{\alpha-1} \exp\{-\beta\tau\}$$

$$f(\tau) \propto \tau^{\alpha-1} \exp\{-\beta\tau\}$$

então:

$$f(\phi, \tau) = f(\phi | \tau)f(\tau)$$

$$= \tau^{\frac{2\alpha+1}{2}-1} \exp\left\{-\left(\beta\tau + \frac{\tau}{2}(\phi - \hat{\phi})^2\right)\right\}$$

Pelo teorema de Bayes, temos :

$$f(\phi, \tau | Z) \propto f(Z | \phi, \tau)f(\phi, \tau)$$

$$\propto \tau^n \sqrt{1-\phi^2} \exp\left\{-\frac{\tau}{2}\left(r + b_1(\phi - \hat{\phi})^2\right)\right\} \tau^{\frac{2\alpha+1}{2}-1} \exp\left\{-\beta\tau - \frac{\tau}{2}(\phi - \hat{\phi})^2\right\}$$

Achamos agora:

$$f(\tau | \phi, Z) \propto \tau^{\frac{n+2\alpha+1}{2}-1} \exp\left\{-\beta\tau - \tau \left[\frac{r + b_1(\phi - \hat{\phi})^2 + (\phi - \hat{\phi})^2}{2} \right]\right\}$$

$$\propto \tau^{\frac{n+2\alpha+1}{2}-1} \exp\left\{-\left[\beta + \frac{1}{2}\left\{r + (b_1 - 1)(\phi - \hat{\phi})^2\right\}\right]\tau\right\}$$

Assim:

$$(\tau | \phi, Z) \sim \text{Gama} \left[\frac{n+2\alpha+1}{2}, \beta + \frac{1}{2} \left\{ r + (b_1 + 1) (\phi - \hat{\phi})^2 \right\} \right]$$

B

Procedemos a encontrar a densidade marginal:

$$f(\phi | Z) = \int_0^{\infty} f(\phi, \tau | Z) d\tau$$

$$\propto \sqrt{1-\phi^2} \int_0^{\infty} \tau^{\frac{n+2\alpha+1}{2}-1} \exp\{-\beta\tau\} d\tau$$

$$\propto \sqrt{1-\phi^2} \left[\frac{1}{B} \right]^{\frac{n+2\alpha+1}{2}} \int_0^{\infty} B^{\frac{n+2\alpha+1}{2}} \tau^{\frac{n+2\alpha+1}{2}-1} \exp\{-\beta\tau\} d\tau$$

1

$$\propto \sqrt{1-\phi^2} \left[\frac{1}{2\beta + r + (b_1 + 1) (\phi - \hat{\phi})^2} \right]^{\frac{n+2\alpha+1}{2}}$$

$$\propto \sqrt{1-\phi^2} \left[(2\beta + r) + (b_1 + 1) (\phi - \hat{\phi})^2 \right]^{-\frac{n+2\alpha+1}{2}}$$

$$\propto \sqrt{1-\phi^2} \left[1 + \frac{(b_1 + 1)}{(2\beta + r)} (\phi - \hat{\phi})^2 \right]^{-\frac{n+2\alpha+1}{2}}$$

Reconhecemos que o segundo termo da densidade acima é proporcional a uma densidade t-Student, assim:

$$\left[1 + \frac{1}{v\sigma^2} (\phi - \hat{\phi})^2 \right]^{-\frac{(v-1)}{2}} \propto \left[1 + \frac{(b_1 + 1)}{(2\beta + r)} (\phi - \hat{\phi})^2 \right]^{-\frac{n+2\alpha+1}{2}}$$

Assim:

$$\frac{v+1}{2} = \frac{n+2\alpha+1}{2} \quad v = n+2\alpha$$

$$\frac{1}{v\sigma^2} = \frac{b_1+1}{2\beta+r} \quad \sigma^2 = \frac{2\beta+1}{(b_1+1)v}$$

então:

$$f(\phi | Z) = \sqrt{1-\phi^2} \cdot t_{n+2\alpha} \left(\hat{\phi}, \frac{2\beta+1}{(b_1+1)v} \right)$$

Para gerar uma a.a da distribuição conjunta $(\phi, \tau | Z)$, aplicamos o algoritmo de M-H para

$f(\phi | Z)$ fazendo:

$$t_{n+2\alpha} \cdot \sqrt{\frac{2\beta+1}{(b_1+1)v} + \hat{\phi}}$$

com este núcleo geramos uma amostra aleatória para ϕ .

Esses valores de ϕ geram valores para τ da:

$$\text{Gama} \left[\frac{n+2\alpha+1}{2}, \beta + \frac{1}{2} \left\{ r + (b_1+1) \left(\phi - \hat{\phi} \right)^2 \right\} \right]$$

2.9 PRIORI T-GAMA PARA O MODELO AR(1)

Infelizmente Priors conjugados não existem para todas as probabilidades. Em geral, para o caso não conjugado, o posterior e seu condicional não pode ser diretamente provado. Acontece, por exemplo, se nós escolhermos para ϕ uma densidade t-Student, p-dimensional com graus de liberdade, parâmetro μ e matriz de precisão P; e uma densidade de Gama para τ , i.e.,

$$f(\phi, \tau) = f(\phi | \tau) f(\tau) \quad (2.28)$$

De forma que:

$$f(\phi | \tau) \propto \left[1 + \frac{(\phi - \mu)P(\phi - \mu)}{K} \right]^{-\frac{(K+p)}{2}}$$

$$f(\tau) \propto \tau^{\alpha-1} \exp\{-\beta\tau\}$$

Em este caso a posteriori esta dado por:

$$f(\phi, \tau | Z) \propto \tau^{\frac{n-p-2\alpha-1}{2}} \exp\left\{-\tau \left[\beta + \frac{B(\phi)}{2} \right]\right\} f(\phi/\tau)$$

Em que:

$$B(\phi) = \left(\phi - \hat{\phi} \right)^T X^T X \left(\phi - \hat{\phi} \right) + \left(Z - \hat{Z} \right)^T \left(Z - \hat{Z} \right)$$

$$f(\phi | \tau) \propto \left[1 + \frac{(\phi - \mu)P(\phi - \mu)}{K} \right]^{-\frac{(K+p)}{2}}$$

Deste modo as densidades condicionais são dadas por:

$$f(\phi | \tau, Z) \propto \exp\left\{-\frac{\tau}{2} \left[\left(\phi - \hat{\phi} \right)^T \left(X^T X \right) \left(\phi - \hat{\phi} \right) \right]\right\} f_1(\phi) \quad (2.29)$$

$$= \psi(\phi).f(\phi | \tau)$$

$$f(\tau | \phi, Z) \sim \text{Gama}\left(\frac{n-p+2\alpha}{2}, \beta + \frac{1}{2} B(\phi)\right)$$

A densidade condicional $f(\phi | \tau, Z)$ em (2.29) não pode ser diretamente provado. Para superar esta dificuldade nós podemos usar o algoritmo Metropolis-Hastings.

2.10 CRITÉRIOS DE SELEÇÃO DE MODELOS :

Uma vez escolhidos vários modelos e calculados os seus parâmetros, o passo seguinte é escolher o modelo mais “parcimonioso”, ou seja, aquele que tem o menor número de parâmetros possível e menor variância dos resíduos. A seleção do modelo é de extrema importância pois a escolha de um valor “pequeno” demais para p pode levar a uma representação inadequada da série, enquanto que um valor “grande” pode levar a um grau de complexidade desnecessário.

2.10.1 Critérios AIC e BIC :

Dentre os critérios mais utilizados para seleção de modelos estão o critério de informação de Akaike (AIC) e o critério de informação bayesiano de Schwarz (BIC).

Estes critérios têm as seguintes formulações (Akaike, 1974; Schwarz, 1978) :

$$AIC = \ln \left(\hat{\sigma}_a^2 \right) + \frac{2M}{N}$$

$$BIC = \ln \left(\hat{\sigma}_a^2 \right) + \frac{M \ln(N)}{N}$$

Onde $M = p+1$ é o número de parâmetros e N é o tamanho da série em questão. Portanto, nota-se que as equações anteriores quantificam o chamado “critério da parcimônia”.

Escolhendo M e $\hat{\phi}$ para minimizar o AIC e o BIC, estamos obtendo um compromisso entre um ajuste acurado (pequeno $S(\hat{\phi})$) e conseqüentemente pequeno $\hat{\sigma}_a^2$ e parcimônia na parametrização (número adequado M de parâmetros). ■

2.10.2 Fator de Bayes :

Pode-se ainda determinar o modelo via fator de Bayes, definido como :

$$B_{ij} = \frac{f(Z | M_i)}{f(Z | M_j)}$$

onde $f(Z | M_i)$ é a função de verossimilhança marginal para todo o conjunto de dados Z definida como :

$$f(Z | M_p) = \int_{\phi} \int_{\tau} L(Z | \phi, \tau, M_p) f(\phi, \tau | M_p) d\phi d\tau$$

onde M_p é o modelo AR(p) sendo analisado, $f(\phi, \tau | M_p)$ é a densidade a priori e $L(Z | \phi, \tau, M_p)$ é a função de verossimilhança dada por (Box & Jenkins, 1976; Oliveira, 1998) :

$$L(Z | \phi, \tau, M_p) \propto \tau^{\frac{N-p}{2}} \exp \left\{ -\frac{\tau}{2} \sum_{t=p+1}^N (Z_t - \phi_1 Z_{t-1} - \dots - \phi_p Z_{t-p})^2 \right\}$$

A verossimilhança marginal pode ser estimada, usando as amostras geradas por simulação MCMC, da seguinte maneira (Raftery, 1996) :

$$\hat{f}(Z | M_p) = \frac{1}{m} \sum_{j=1}^m L(Z | \phi^{(j)}, \tau^{(j)}, M_p)$$

$$\hat{f}(Z | M_p) = \left[\frac{1}{m} \sum_{j=1}^m \frac{1}{L(Z | \phi^{(j)}, \tau^{(j)}, M_p)} \right]^{-1}$$

De acordo com o critério do fator de Bayes, prefere-se o modelo M_j ao modelo M_i se:

$$\hat{B}_{ji} = \frac{\hat{f}(Z | M_i)}{\hat{f}(Z | M_j)} < 1$$

2.11 PREVISÃO K PASSOS A FRENTE :

O ajuste de modelos AR(p) tem como um dos objetivos finais a predição de valores futuros para a série temporal em questão. Denota-se por Z_f os k valores da série a serem previstos e por Z_p as p últimas observações, isto é:

$$\begin{aligned} Z_f &= (Z_{(n+1)}, Z_{(n+2)}, \dots, Z_{(n+k)}) \\ Z_p &= (Z_{n-(p-1)}, Z_{n-(p-2)}, \dots, Z_n) \end{aligned}$$

Seja $f(\phi, \tau | Z)$ a densidade conjunta a posteriori para os parâmetros de interesse ϕ e

τ . A densidade preditiva pode ser calculada por:

$$f(Z_f | Z_p) = \int_{\phi} \int_{\tau} f(Z_f | \phi, \tau, Z_p) f(\phi, \tau | Z_p) d\tau d\phi$$

onde \int_{ϕ} corresponde a p integrais $\int_{\phi_1} \dots \int_{\phi_p}$. A previsão dos valores futuros Z_f pode ser obtida através da resolução da seguinte integral múltipla:

$$E(Z_f | Z) = \int Z_f f(Z_f | Z) dZ_f$$

$$E(Z_f | Z) = \int Z_f \left[\int \int f(Z_f | \phi, \tau, Z_p) f(\phi, \tau | Z_p) d\tau d\phi \right] dZ_f$$

Alterando a ordem das integrais tem-se:

$$E(Z_f | Z) = \int \int E(Z_f | \phi, \tau, Z_p) f(\phi, \tau | Z_p) d\tau d\phi \quad (2.30)$$

Usando o modelo $AR(p)$, tem-se:

$$Z_{n-k} = \phi_1 Z_{n+k-1} + \dots + \phi_p Z_{n-(p-k)} + a_{n+k}$$

Esta expressão na forma matricial fica:

$$A(\phi) = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ -\phi_1 & 1 & 0 & \dots & 0 \\ -\phi_2 & -\phi_1 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\phi_{k-1} & -\phi_{k-2} & -\phi_{k-3} & \dots & 1 \end{bmatrix}_{k \times k} \quad B(\phi) = \begin{bmatrix} -\phi_p & -\phi_{p-1} & -\phi_{p-2} & \dots & -\phi_1 \\ 0 & -\phi_p & 0 & \dots & -\phi_2 \\ 0 & 0 & -\phi_p & \dots & -\phi_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & -\phi_{k-3} & \dots & -\phi_k \end{bmatrix}_{k \times p}$$

Pode-se escrever, portanto:

$$A(\phi)Z_f + B(\phi)Z_p = a \Rightarrow Z_f = -A^{-1}(\phi)B(\phi)Z_p + A^{-1}(\phi)a$$

$N\left[E(Z_f | \phi, \tau, Z), \text{Var}(Z_f | \phi, \tau, Z)\right]$, em que a é um vetor de variáveis aleatórias i.i.d, $N(0, \tau^{-1})$. Tem-se que Z_f também é um vetor de variáveis aleatórias que obedece a uma distribuição

$$E(Z_f | \phi, \tau, Z) = -A^{-1}(\phi)B(\phi)Z_p \quad (2.31)$$

$$\text{Var}(Z_f | \phi, \tau, Z) = \tau^{-1} \left(A^{-1}(\phi) \right) \left(A^{-1}(\phi) \right)^T \quad (2.32)$$

Substituindo-se (2.31) em (2.30) pode-se calcular a previsão dos valores futuros Z_f resolvendo-se a seguinte integral:

$$E(Z_f | Z) = \int_{\underline{\phi}} \int_{\underline{\tau}} \left(-A^{-1}(\underline{\phi})B(\underline{\phi})Z_p \right) f(\underline{\phi}, \tau | Z_p) d\tau d\phi \quad (2.33)$$

A variância é calculada como:

$$Var(Z_f | Z) = E_{\phi, \tau} \left[Var(Z_f | \phi, \tau, Z) \right] + Var_{\phi, \tau} \left[E(Z_f | \phi, \tau, Z) \right] \quad (2.34)$$

onde:

$$\begin{aligned} E_{\phi, \tau} \left[Var(Z_f | \phi, \tau, Z) \right] &= \int_{\underline{\phi}} \int_{\underline{\tau}} Var(Z_f | \phi, \tau, Z) f(\underline{\phi}, \tau | Z) d\tau d\phi \quad (2.35) \\ &= \int_{\underline{\phi}} \int_{\underline{\tau}} \left[\tau^{-1} \left(A^{-1}(\underline{\phi}) \right) \left(A^{-1}(\underline{\phi}) \right)^T \right] f(\underline{\phi}, \tau | Z) d\tau d\phi \end{aligned}$$

Em geral, devido a complexidade das expressões (2.33) e (2.34) para $p \geq 2$, essas integrais só podem ser calculadas de forma aproximada. Uma alternativa bastante viável e que simplifica consideravelmente o tratamento do problema da predição consiste no uso de métodos de simulação de Monte Carlo.

2.11.1 Previsão usando MCMC

Retomando a Equação (2.33), afirma-se que o valor esperado $E(Z_f | Z)$ pode ser calculado usando algoritmos de simulação de Monte Carlo. Para isso, considera-se uma amostra $\phi^{(j)} = (\phi_1^{(j)}, \phi_2^{(j)}, \dots, \phi_p^{(j)})$, $j = 1, \dots, m$ gerada por simulação MCMC. Pode-se, portanto, calcular o estimador de Monte Carlo para $E(Z_f | Z)$ usando essa amostra através da seguinte expressão :

$$E(Z_f | Z) = \frac{1}{m} \sum_{j=1}^m -A^{-1}(\phi^{(j)})B(\phi^{(j)})Z_p \quad (2.36)$$

tomando-se agora a expressão (2.34), dada por :

$$Var(Z_f | Z) = E_{\phi, \tau} \left[Var(Z_f | \phi, \tau, Z) \right] + Var_{\phi, \tau} \left[E(Z_f | \phi, \tau, Z) \right] \quad (2.37)$$

Pode-se escrever cada termo da equação (2.37) usando a amostra gerada por simulação de Monte Carlo, ou seja :

$$\hat{E}_{\phi, \tau} \left[\text{Var} \left(Z_f \mid \phi, \tau, Z \right) \right] = \text{diag} \left(\frac{1}{m} \sum_{j=1}^m (\tau^{(j)})^{-1} \left[A^{-1} \left(\phi^{(j)} \right) \left[A^{-1} \left(\phi^{(j)} \right) \right]^T \right] \right) \quad (2.38)$$

e, usando (2.36),

$$\begin{aligned} \hat{\text{Var}}_{\phi, \tau} \left[E \left(Z_f \mid \phi, \tau, Z \right) \right] &= E \left(Z_f^2 \mid \phi, \tau, Z \right) + \left[E \left(Z_f \mid \phi, \tau, Z \right) \right]^2 \\ &= \frac{1}{m} \sum_{j=1}^m \left(A^{-1} \left(\phi^{(j)} \right) B \left(\phi^{(j)} \right) Z_p \right)^2 + \left(\frac{1}{m} \sum_{j=1}^m A^{-1} \left(\phi^{(j)} \right) B \left(\phi^{(j)} \right) Z_p \right)^2 \end{aligned} \quad (2.39)$$

onde $E \left(Z_f \mid \phi^{(j)}, \tau^{(j)}, Z \right) = -A^{-1} \left(\phi^{(j)} \right) B \left(\phi^{(j)} \right) Z_p, j=1, \dots, m.$

Assim, substituindo as Equações (2.38) e (2.39) na Equação (2.37), tem-se o estimador de Monte Carlo para $\text{Var} \left(Z_f \mid Z \right)$ obtido por simulação MCMC.

CAPITULO 3

ABORDAGEM CLÁSSICA E BAYESIANA NA ESTIMAÇÃO DO PARÂMETRO DE TRANSFORMAÇÃO λ EM PROCESSOS AUTO-REGRESSIVOS DE PRIMEIRA ORDEM.

3.1 INTRODUÇÃO :

Neste capítulo são desenvolvidos procedimentos de estimação em modelos auto-regressivos de primeira ordem AR(1) com erros gaussianos quando existe a possibilidade de incorporar uma transformação de potência nos dados com a finalidade de conseguir uma série estacionária.

Aqui observaremos que surgem problemas na geração de uma amostra para a distribuição a posteriori de λ dado que a sua expressão é muito complexa, assim é desenvolvido um método recursivo para estimar os parâmetros do modelo onde é feita uma combinação de métodos clássicos e bayesianos, usando os algoritmos de Newton-Rapshon e as técnicas de simulação Monte Carlo em Cadeias de Markov (MCMC).

3.2 TRATAMENTO BAYESIANO NAS TRANSFORMAÇÕES DE BOX E COX NO PROCESSO AR(1) :

Box & Jenkins (1970) discutiam a possibilidade de obter um modelo com erros de variância homogênea, ajustando modelos ARIMA não necessariamente para a série dada Z_t , porém, para uma série derivada de um conjunto de transformações de potência, da forma,

$$Z_t^{(\lambda)} = \begin{cases} Z_t^\lambda - 1, \lambda \neq 0 \\ \lambda \\ \log Z_t, \lambda = 0 \end{cases} \quad (3.1)$$

o uso dessa transformação exige que $Z_t > 0$ e $t \geq 0$, quando este não é o caso a transformação (3.1) pode ser modificada para :

$$Z_t^{(\lambda)} = \begin{cases} \frac{(Z_t + c)^\lambda - 1}{\lambda}, \lambda \neq 0 \\ \log(Z_t + c), \lambda = 0 \end{cases}$$

Assim definimos o modelo auto-regressivo de primeira ordem AR(1) como:

$$Z_t = a_t + \phi Z_{t-1} \text{ onde } a_t \sim N(0, \sigma^2)$$

Fazendo $\lambda \rightarrow 0$ em (3.1) a transformação de Box e Cox fica : $Z_t^{(\lambda)} = \ln Z_t$

A Verossimilhança do Modelo AR(1) para a série transformada fica :

$$\begin{aligned} f(Z^{(\lambda)} | \phi, \tau) &\propto \tau^{-n} \exp\left\{-\frac{\tau}{2} \left[\sum_{t=2}^n (Z_t^{(\lambda)} - \phi Z_{t-1}^{(\lambda)})^2 \right]\right\} \\ &\propto \tau^{-n} \exp\left\{-\frac{\tau}{2} \left[\sum_{t=2}^n (\ln Z_t - \phi \ln Z_{t-1})^2 \right]\right\} \end{aligned}$$

Agora:

$$Z_t^{(\lambda)} = a_t + \phi Z_{t-1}^{(\lambda)} \text{ modelo AR(1) para } Z_t^{(\lambda)}$$

sendo:

$$a_t = Z_t^{(\lambda)} - \phi Z_{t-1}^{(\lambda)}$$

$$a_t = \ln Z_t - \phi \ln Z_{t-1}$$

$$E(a_t) = E(\ln Z_t - \phi \ln Z_{t-1}) = 0$$

Assim para $\lambda \neq 0$, temos a função de verossimilhança exata para os dados transformados dada por :

$$f(Z^{(\lambda)} | \phi, \tau, \lambda) \propto \sqrt{1 - \phi^2} \tau^{-n} \exp\left\{-\frac{\tau}{2} \left[(1 - \phi^2) Z_1^{(\lambda)2} + \sum_{t=2}^n (Z_t^{(\lambda)} - \phi Z_{t-1}^{(\lambda)})^2 \right]\right\} \quad (3.2)$$

Achamos o EMV para ϕ :

$$f(\mathbf{Z}^{(\lambda)} | \phi, \tau, \lambda) \propto \tau^n \sqrt{1-\phi^2} \exp \left\{ -\frac{\tau}{2} \left[(1-\phi^2) \left(\frac{Z_1^\lambda - 1}{\lambda} \right)^2 + \sum_{i=2}^n \left(\frac{Z_i^\lambda - 1}{\lambda} - \phi \frac{Z_{i-1}^\lambda - 1}{\lambda} \right)^2 \right] \right\}$$

$$\propto \tau^n \sqrt{1-\phi^2} \exp \left\{ -\frac{\tau}{2\lambda^2} \left[(1-\phi^2) (Z_1^\lambda - 1)^2 + \sum_{i=2}^n \left[(Z_i^\lambda - 1) - \phi (Z_{i-1}^\lambda - 1) \right]^2 \right] \right\} \quad (3.3)$$

(*)

trabalhando em (3.3) temos :

$$\ln f(\mathbf{Z}^{(\lambda)} | \phi, \tau, \lambda) \propto -\frac{\tau}{2\lambda^2} \left[(1-\phi^2) (Z_1^\lambda - 1)^2 + \sum_{i=2}^n \left[(Z_i^\lambda - 1) - \phi (Z_{i-1}^\lambda - 1) \right]^2 \right] + \ln \left[\tau^{n/2} \sqrt{1-\phi^2} \right]$$

$$\frac{\partial \ln f(\mathbf{Z}^{(\lambda)} | \phi, \tau, \lambda)}{\partial \phi} \propto -\frac{\tau}{2\lambda^2} \left[-2\phi (Z_1^\lambda - 1)^2 - 2 \sum_{i=2}^n (Z_i^\lambda - 1) (Z_{i-1}^\lambda - 1) + 2\phi \sum_{i=2}^n (Z_{i-1}^\lambda - 1)^2 \right] - \frac{\phi}{1-\phi^2}$$

Podemos desconsiderar o termo final da equação anterior como é o caso na verossimilhança aproximada. Igualando a zero, obtemos :

$$\phi \left\{ \sum_{i=2}^n [Z_{i-1}^\lambda - 1]^2 - [Z_1^\lambda - 1]^2 \right\} = \sum_{i=2}^n [Z_i^\lambda - 1] [Z_{i-1}^\lambda - 1]$$

Logo o EMV para ϕ é :

$$\hat{\phi}_{EMV} = \frac{\sum_{i=2}^n [Z_i^\lambda - 1] [Z_{i-1}^\lambda - 1]}{\sum_{i=2}^n [Z_{i-1}^\lambda - 1]^2 - [Z_1^\lambda - 1]^2}$$

O EMV para λ é apresentado em (3.7).

Agora, precisamos de uma expressão mais reduzida da verossimilhança em (3.3), assim trabalhando em (*) temos que essa expressão pode ser escrita como

$$(1-\phi^2) (Z_1^\lambda - 1)^2 + \sum_{i=2}^n \left[(Z_i^\lambda - 1) - \phi (Z_{i-1}^\lambda - 1) \right]^2$$

$$= (Z_1^\lambda - 1)^2 - \phi^2 (Z_1^\lambda - 1)^2 + \sum_{i=2}^n \left[(Z_i^\lambda - 1)^2 - 2\phi (Z_i^\lambda - 1) (Z_{i-1}^\lambda - 1) + \phi^2 (Z_{i-1}^\lambda - 1)^2 \right]$$

$$= (Z_1^\lambda - 1)^2 - \phi^2 (Z_1^\lambda - 1)^2 + \sum_{t=2}^n (Z_t^\lambda - 1)^2 - 2\phi \sum_{t=2}^n (Z_t^\lambda - 1)(Z_{t-1}^\lambda - 1) + \phi^2 \sum_{t=2}^n (Z_{t-1}^\lambda - 1)^2$$

Agrupando convenientemente :

$$\begin{aligned} &= \sum_{t=1}^n (Z_t^\lambda - 1)^2 + \phi^2 \left[\sum_{t=2}^n (Z_{t-1}^\lambda - 1)^2 - (Z_1^\lambda - 1)^2 \right] - 2\phi \sum_{t=2}^n (Z_t^\lambda - 1)(Z_{t-1}^\lambda - 1) \\ &= \sum_{t=1}^n (Z_t^\lambda - 1) - \frac{\left[\sum_{t=2}^n (Z_t^\lambda - 1)(Z_{t-1}^\lambda - 1) \right]^2}{\sum_{t=2}^n (Z_{t-1}^\lambda - 1)^2 - (Z_1^\lambda - 1)^2} + \phi^2 \left[\sum_{t=2}^n (Z_{t-1}^\lambda - 1)^2 - (Z_1^\lambda - 1)^2 \right] \\ &\quad - 2\phi \sum_{t=2}^n (Z_t^\lambda - 1)(Z_{t-1}^\lambda - 1) + \frac{\sum_{t=2}^n (Z_t^\lambda - 1)(Z_{t-1}^\lambda - 1)}{\sum_{t=2}^n (Z_{t-1}^\lambda - 1) - (Z_1^\lambda - 1)^2} \\ &= \sum_{t=1}^n (Z_t^\lambda - 1) - \frac{\left[\sum_{t=2}^n (Z_t^\lambda - 1)(Z_{t-1}^\lambda - 1) \right]^2}{\sum_{t=2}^n (Z_{t-1}^\lambda - 1)^2 - (Z_1^\lambda - 1)^2} \\ &\quad + \frac{\left[\sum_{t=2}^n (Z_{t-1}^\lambda - 1) - (Z_1^\lambda - 1)^2 \right] \left[\phi^2 - 2\phi \frac{\sum_{t=2}^n (Z_t^\lambda - 1)(Z_{t-1}^\lambda - 1)}{\sum_{t=2}^n (Z_{t-1}^\lambda - 1)^2 - (Z_1^\lambda - 1)^2} + \frac{\left[\sum_{t=2}^n (Z_t^\lambda - 1)(Z_{t-1}^\lambda - 1) \right]^2}{\left[\sum_{t=2}^n (Z_{t-1}^\lambda - 1)^2 - (Z_1^\lambda - 1)^2 \right]} \right]}{\left[\sum_{t=2}^n (Z_{t-1}^\lambda - 1) - (Z_1^\lambda - 1)^2 \right]} \\ &\quad \left[\phi - \frac{\sum_{t=2}^n (Z_t^\lambda - 1)(Z_{t-1}^\lambda - 1)}{\sum_{t=2}^n (Z_{t-1}^\lambda - 1) - (Z_1^\lambda - 1)^2} \right]^2 \end{aligned}$$

Podemos assim denotar as expressões:

$$\begin{array}{c}
\frac{(b_2)^2}{\sum_{t=1}^n (z_t^\lambda - 1)^2} - \frac{\left[\sum_{t=2}^n (z_t^\lambda - 1)(z_{t-1}^\lambda - 1) \right]^2}{\sum_{t=2}^n (z_{t-1}^\lambda - 1)^2 - (z_1^\lambda - 1)^2} + \left[\sum_{t=2}^n (z_{t-1}^\lambda - 1)^2 - (z_1^\lambda - 1)^2 \right] \left[\phi - \frac{\sum_{t=2}^n (z_t^\lambda - 1)(z_{t-1}^\lambda - 1)}{\sum_{t=2}^n (z_{t-1}^\lambda - 1) - (z_1^\lambda - 1)^2} \right]^2 \\
\hline
\begin{array}{cccc}
b_3 & b_1 & b_1 & b_1 \\
\hline
& & & r
\end{array}
\end{array}$$

onde $\hat{\phi}_{EMV} = \frac{b_2}{b_1}$

Substituindo em (3.3), temos a expressão do verossimilhança exata:

$$L(Z^{(\lambda)} | \phi, \tau, \lambda) \propto \tau^n \sqrt{1 - \phi^2} \exp \left\{ -\frac{\tau}{2\lambda^2} \left[r + b_1 \left(\phi - \hat{\phi} \right)^2 \right] \right\}$$

3.2.1 PRIORI NÃO INFORMATIVA PARA O MODELO AR(1)

Precisamos agora ter uma expressão para a priori de (ϕ, τ, λ) ; dado que não temos nenhuma informação optamos pela priori :

$$f(\phi, \tau, \lambda) \propto \frac{1}{\tau}, \quad \tau > 0$$

$$f(\phi, \tau, \lambda) = f(\tau | \lambda, \phi) f(\lambda | \phi) f(\phi), \text{ onde } f(\lambda | \phi) \text{ e } f(\phi) \text{ são assumidos } U(-1,1)$$

Pelo teorema de Bayes, obtemos uma expressão para a distribuição a posteriori conjunta de $(\phi, \tau, \lambda | Z)$:

$$f(\phi, \tau, \lambda | Z) \propto \tau^{\frac{n}{2}-1} \sqrt{1-\phi^2} \exp\left\{-\frac{\tau}{2\lambda^2} \left[r + b_1 (\phi - \hat{\phi})^2 \right]\right\} \quad (3.4)$$

Precisamos obter agora uma amostra para essa distribuição.

Achamos a distribuição condicional de $(\tau | Z, \phi, \lambda)$, isto implica que as expressões que não dependam de τ passam a formar parte da constante de proporcionalidade.

Assim:

$$f(\tau | Z, \phi, \lambda) \propto \tau^{\frac{n}{2}-1} \exp\left\{-\frac{\tau}{2\lambda^2} \left[r + b_1 (\phi - \hat{\phi})^2 \right]\right\}$$

Temos que:

$$(\tau | Z, \phi, \lambda) \propto \text{Gama} \left[\frac{n}{2}, \frac{r + b_1 (\phi - \hat{\phi})^2}{2\lambda^2} \right] \quad \tau > 0$$

Obtemos agora uma expressão para a distribuição condicional de $(\phi, \lambda | Z)$, para este fim precisamos integrar (3.4) com respeito a τ :

$$\begin{aligned} f(\phi, \lambda | Z) &= \int_0^{\infty} f(\phi, \tau, \lambda | Z) d\tau \\ &\propto \int_0^{\infty} \tau^{\frac{n}{2}-1} \sqrt{1-\phi^2} \exp\left\{-\frac{\tau}{2\lambda^2} \left[r + b_1 (\phi - \hat{\phi})^2 \right]\right\} \left[\frac{r + b_1 (\phi - \hat{\phi})^2}{2\lambda^2} \right]^{\frac{n}{2}} \left[\frac{2\lambda^2}{r + b_1 (\phi - \hat{\phi})^2} \right]^{\frac{n}{2}} d\tau \\ &\propto \sqrt{1-\phi^2} \left[\frac{2\lambda^2}{r + b_1 (\phi - \hat{\phi})^2} \right]^{\frac{n}{2}} \int_0^{\infty} \tau^{\frac{n}{2}-1} \exp\left\{-\tau \left[\frac{r + b_1 (\phi - \hat{\phi})^2}{2\lambda^2} \right]\right\} \left[\frac{r + b_1 (\phi - \hat{\phi})^2}{2\lambda^2} \right]^{\frac{n}{2}} d\tau \\ &= 1 \end{aligned}$$

$$\propto \sqrt{1-\phi^2} \lambda^n \tau^{-\frac{n}{2}} \left[1 + \frac{b_1}{r} (\phi - \hat{\phi})^2 \right]^{-\frac{n}{2}} \quad (3.5)$$

A partir da aqui podemos obter uma expressão para a distribuição condicional de $(\phi | Z, \lambda)$

Assim :

$$f(\phi | Z, \lambda) \propto \sqrt{1-\phi^2} \left[1 + \frac{b_1}{r} (\phi - \hat{\phi})^2 \right]^{-\frac{n}{2}} \quad (3.6)$$

O segundo termo da expressão é proporcional á dist. t-student :

$$\left[1 + \frac{1}{v} \left(\frac{\phi - \hat{\phi}}{\sigma} \right)^2 \right]^{-\frac{(v+1)}{2}} \propto \left[1 + \frac{b_1}{r} (\phi - \hat{\phi})^2 \right]^{-\frac{n}{2}}$$

$$\rightarrow -\frac{(v+1)}{2} = -\frac{n}{2} \quad \rightarrow v = n-1$$

$$\frac{1}{v\sigma^2} = \frac{b_1}{r} \quad \rightarrow \sigma^2 = \frac{r}{b_1 v} \quad \text{r e } b_1 \text{ depende de } \lambda .$$

Com este segundo temos gerado valores para ϕ mediante o algoritmo de M-H, usando o núcleo :

$$t(v) x \sqrt{\frac{r}{b_1 v}} + \hat{\phi}$$

é a constante $\sqrt{1-\phi^2}$ para completar o algoritmo.

Na seguinte seção apresentamos algumas alternativas para se obter a distribuição de $(\lambda | Z)$ e poder gerar as amostras da posteriori $(\phi, \tau, \lambda | Z)$.

3.3 ALTERNATIVAS PARA A OBTENÇÃO DA DISTRIBUIÇÃO DE $(\lambda | Z)$:

Apresentamos 3 alternativas para a obtenção de $(\lambda | Z)$ e completar o algoritmo para obter (3.4)

3.3.1 Mediante a Condicional $(\lambda | Z, \hat{\phi})$:

De (3.5) podemos escrever :

$$f(\lambda | \hat{\phi}, Z) \propto \left[\frac{2\lambda^2}{r + b_1(\hat{\phi} - \phi)^2} \right]^{\frac{n}{2}}$$

$$\propto \lambda^n \left[r + b_1(\hat{\phi} - \phi)^2 \right]^{-\frac{n}{2}}$$

temos que r , b_1 e $\hat{\phi}$ depende de λ .

Assim :

$$f(\lambda | \hat{\phi}, Z) \propto \lambda^n \left\{ \sum_{i=1}^n (Z_i^\lambda - 1)^2 - \frac{\left[\sum_{i=2}^n (Z_{i-1}^\lambda)(Z_{i-1}^\lambda - 1) \right]^2}{\sum_{i=2}^n (Z_{i-1}^\lambda - 1)^2 - (Z_1^\lambda - 1)^2} \right\}$$

$$+ \left[\sum_{i=2}^n (Z_{i-1}^\lambda - 1)^2 - (Z_1^\lambda - 1)^2 \right] \left[\phi - \frac{\sum_{i=2}^n (Z_{i-1}^\lambda - 1)(Z_{i-1}^\lambda - 1)}{\sum_{i=2}^n (Z_{i-1}^\lambda - 1)^2 - (Z_1^\lambda - 1)^2} \right]^2$$

Desenvolvendo a expressão acima obtemos :

$$f(\lambda | \hat{\phi}, Z) \propto \lambda^n \left\{ (1 - \phi^2) (Z_1^\lambda - 1)^2 + \sum_{i=2}^n [(Z_i^\lambda - 1) - \phi(Z_{i-1}^\lambda - 1)]^2 \right\}$$

aqui não é impossível reconhecer alguma distribuição conhecida para λ o que dificulta a obtenção da amostra.

3.3.2 Mediante a Marginal $(\lambda | Z)$:

Aqui podemos tomar a conjunta de $(\phi, \lambda | Z)$ em (3.5) e integrar com respeito a ϕ , teremos :

$$f(\lambda | Z) = \int (\phi, \lambda | Z) d\phi$$

$$\propto r^{-n} \lambda^n \int_{-1}^1 \sqrt{1-\phi^2} \left[1 + \frac{b_1}{r} (\phi - \hat{\phi})^2 \right]^n d\phi$$

O cálculo dessa integral foi feito no programa MAPLE, sem obter resultados satisfatórios, resulta difícil para o programa a resolução dessa integral dado que a sua expressão é complexa.

De qualquer modo, tendo o valor dessa integral, temos os termos r , b_1 e $\hat{\phi}$ que dependem de λ , e como no caso da condicional de $(\lambda | \phi, Z)$ temos problemas na identificação de um núcleo para começar o algoritmo de Metropolis-Hastings.

Dados esses problemas, como terceira alternativa sugerimos o uso de uma metodologia que combina a análise clássica dos dados com a análise bayesiana, trabalhando com distribuições a posteriori para os parâmetros ϕ e τ e encontrando o EMV para λ , dados os parâmetros anteriores.

3.3.3 Mediante a moda a posteriori - Análise Clássica e Bayesiana :

Nas seções anteriores foram desenvolvidas distribuições a posteriori para :

$$(\tau | \phi, \lambda, Z) \sim \text{Gama} \left[\frac{n}{2}, \frac{r + b_1 (\phi - \hat{\phi})^2}{2\lambda^2} \right]$$

e para:

$$(\phi | \lambda, Z) \sim \sqrt{1 - \phi^2} \left\{ t_{(v)} \sqrt{\frac{r}{b_1 v}} + \hat{\phi} \right\}$$

$$\text{t-Student}_{(v)} \left\{ \hat{\phi}, \frac{r}{b_1 v} \right\} \quad \text{núcleo para o algoritmo M-H}$$

Tivemos problemas para obter a distribuição de $(\lambda | Z)$ a qual é necessária para gerar amostras das duas posteriores anteriores, como alternativa podemos tomar um valor inicial para λ e substituí-lo nas distribuições condicionais de τ e ϕ , assim geramos um novo valor para λ a partir de uma equação.

Para isto precisamos que esta equação combine os valores de ϕ, τ e λ .

A verossimilhança em (3.3) combina estes parâmetros,

$$f(Z | \phi, \tau, \lambda) \propto \exp \left\{ -\frac{\tau}{2\lambda^2} \left[(1 - \phi^2)(Z_1^\lambda - 1) + \sum_{i=2}^n [(Z_i^\lambda - 1) - \phi(Z_{i-1}^\lambda - 1)]^2 \right] \right\}$$

$$\ln f(Z | \phi, \tau, \lambda) \propto -\frac{\tau}{2\lambda^2} \left[(1 - \phi^2)(Z_1^\lambda - 1)^2 + \sum_{i=2}^n [(Z_i^\lambda - 1) - \phi(Z_{i-1}^\lambda - 1)]^2 \right]$$

$$\propto \frac{-\tau(1 - \phi^2)(Z_1^\lambda - 1)^2}{2\lambda^2} - \frac{\tau \sum_{i=2}^n [(Z_i^\lambda - 1) - \phi(Z_{i-1}^\lambda - 1)]^2}{2\lambda^2}$$

$$\frac{\partial \ln f(Z | \phi, \tau, \lambda)}{\partial \lambda} \propto \frac{-2\tau(1 - \phi^2)(Z_1^\lambda - 1)Z_1^\lambda \ln Z_1 - 4\lambda}{4\lambda^4} - \frac{2\tau \sum_{i=2}^n [(Z_i^\lambda - 1) - \phi(Z_{i-1}^\lambda - 1)] [Z_i^\lambda \ln Z_i - \phi Z_{i-1}^\lambda \ln Z_{i-1}] - 4\lambda}{4\lambda^4} = 0$$

assim:

$$-\frac{\tau}{2\lambda^4} \left[(1 - \phi^2)(Z_1^\lambda - 1)Z_1^\lambda \ln Z_1 + \sum_{i=2}^n [(Z_i^\lambda - 1) - \phi(Z_{i-1}^\lambda - 1)] [Z_i^\lambda \ln Z_i - \phi Z_{i-1}^\lambda \ln Z_{i-1}] \right] = 0$$

$\neq 0$

$$g(\lambda) = (1 - \phi^2)(Z_1^\lambda - 1)Z_1^\lambda \ln Z_1 + \sum_{i=2}^n [(Z_i^\lambda - 1) - \phi(Z_{i-1}^\lambda)] [Z_i^\lambda \ln Z_i - \phi Z_{i-1}^\lambda \ln Z_{i-1}] = 0$$

então:

$$\begin{aligned} g(\lambda) &= (1 - \phi^2)(Z_1^\lambda - 1)Z_1^\lambda \ln Z_1 + \sum_{i=2}^n [(Z_i^\lambda - 1 - \phi Z_{i-1}^\lambda + 1)(Z_i^\lambda \ln Z_i - \phi Z_{i-1}^\lambda \ln Z_{i-1})] \\ &= (1 - \phi^2)(Z_1^\lambda - 1)Z_1^\lambda \ln Z_1 + \sum_{i=2}^n Z_i^{2\lambda} \ln Z_i - \phi Z_i^\lambda Z_{i-1}^\lambda \ln Z_{i-1} - \phi Z_{i-1}^\lambda Z_i^\lambda \ln Z_i + \phi^2 Z_{i-1}^{2\lambda} \ln Z_{i-1} \\ &= (1 - \phi^2)(Z_1 - 1)Z_1^\lambda \ln Z_1 + \sum_{i=2}^n Z_i^{2\lambda} \ln Z_i - \phi \sum_{i=2}^n Z_i^\lambda Z_{i-1}^\lambda \ln Z_{i-1} - \phi \sum_{i=2}^n Z_{i-1}^\lambda Z_i^\lambda \ln Z_i + \phi^2 \sum_{i=2}^n Z_{i-1}^{2\lambda} \ln Z_{i-1} = 0 \end{aligned} \quad (3.7)$$

logo para um valor inicial de $\lambda = \lambda_0$ [obtido do ajuste $(\ln \bar{X}_i, \ln S_i^2)$] obtemos amostras para as posteriores de ϕ e τ , assim, de (3.7) temos um novo valor para λ , fazemos isto até λ convergir.

A equação $g(\lambda) = 0$. Pode ser solucionada mediante o algoritmo de Newton-Rapshon.

$$\lambda^{(m+1)} = \lambda^{(m)} - \frac{g(\lambda)}{g'(\lambda)}$$

Este algoritmo é interessante porque combina ambas abordagens clássico e bayesiano e permite comparar como se comporta o estimador para λ em ambas situações. Este procedimento é equivalente à construção da função de verossimilhança “profile” conhecida na abordagem clássica e que será desenvolvida na seguinte seção.

3.4 A TRANSFORMAÇÃO DE BOX E COX PARA MODELOS AR(p) :

Nesta seção vamos generalizar os resultados clássicos encontrados para modelos auto-regressivos de ordem p. Consideramos uma série temporal $\{Z_t, t = 1, 2, \dots, n\}$ onde $Z_t > 0$ para todo $t \geq 0$, o uso da transformação de Box e Cox para Z_t consiste em construir a série transformada dada por :

$$Z_t^{(\lambda)} = \begin{cases} \frac{Z_t^\lambda - 1}{\lambda}, \lambda \neq 0 \\ \ln Z_t, \lambda = 0 \end{cases}$$

Se $Z_t^{(\lambda)}$ pode ser modelada por um AR(p) então temos :

$$Z_t^{(\lambda)} = \sum_{j=1}^p \phi_j Z_{t-j}^{(\lambda)} + a_t$$

onde $\{a_t, t \geq 0\}$ é um ruído i.i.d. $N(0, \sigma^2)$.

As estimativas dos parâmetros $\phi = (\phi_1, \dots, \phi_p)$ e $\tau = 1/\sigma^2$ devem ser calculadas juntamente com o parâmetro de transformação λ . Para isso vamos considerar a função de verossimilhança para os dados reais $\{Z_t, t \geq 0\}$. Assim temos :

$$L(\phi, \lambda, \tau | Z) = \prod_{t=p+1}^n f(Z_t | Z)$$

onde $f(Z_t | Z) = g(Z_t^{(\lambda)} | Z^{(\lambda)}) \cdot |J_{Z_t, Z_t^{(\lambda)}}|$ sendo $g(Z_t^{(\lambda)} | Z^{(\lambda)})$ a função de verossimilhança dos dados transformados e $J_{Z_t, Z_t^{(\lambda)}}$ o Jacobiano da transformação de Box e Cox.

Sendo a_t i.i.d. $N(0, \sigma^2)$ temos que :

$$g(Z_t^{(\lambda)} | Z^{(\lambda)}) \propto \tau^{1/2} \exp\left\{-\frac{\tau}{2} \left(Z_t^{(\lambda)} - \sum_{j=1}^p \phi_j Z_{t-j}^{(\lambda)}\right)^2\right\} \quad (3.8)$$

Além disso $J_{Z_t, Z_t^{(\lambda)}} = Z_t^{\lambda-1} \left(J_{Z_t, Z_t^{(\lambda)}} = \frac{\partial Z_t^{(\lambda)}}{\partial Z_t} \right)$ então :

$$f(Z_t | Z) \propto \tau^{1/2} \exp\left\{-\frac{\tau}{2} \left(Z_t^{(\lambda)} - \sum_{j=1}^p \phi_j Z_{t-j}^{(\lambda)}\right)^2\right\} \cdot Z_t^{\lambda-1} \quad (3.9)$$

A função de verossimilhança para os dados reais $\{Z_t, t \geq 0\}$ é dado por :

$$L(\phi, \tau, \lambda | Z) \propto \prod_{t=p+1}^n \tau^{1/2} \exp\left\{-\frac{\tau}{2} \left(Z_t^{(\lambda)} - \sum_{j=1}^p \phi_j Z_{t-j}^{(\lambda)}\right)^2\right\} \cdot Z_t^{\lambda-1} \quad (3.10)$$

Usando a forma matricial :

$$\underline{Z}^{(\lambda)} = \begin{pmatrix} Z_{p+1}^{(\lambda)} \\ \vdots \\ Z_n^{(\lambda)} \end{pmatrix}_{(n-p) \times 1} \quad X = \begin{pmatrix} Z_p^{(\lambda)} & Z_{p-1}^{(\lambda)} & \dots & \dots & Z_1^{(\lambda)} \\ \vdots & \vdots & \dots & \dots & \vdots \\ Z_{n-1}^{(\lambda)} & Z_{n-2}^{(\lambda)} & \dots & \dots & Z_{n-p}^{(\lambda)} \end{pmatrix}_{(n-p) \times p}$$

temos :

$$l(\underline{\phi}, \tau, \lambda | \underline{Z}) \propto \tau^{n-p/2} \exp\left\{-\frac{\tau}{2} \left(\underline{Z}^{(\lambda)} - \underline{X}^{(\lambda)} \underline{\phi}\right)^T \left(\underline{Z}^{(\lambda)} - \underline{X}^{(\lambda)} \underline{\phi}\right)\right\} \prod_{i=p+1}^n Z_i^{\lambda-1} \quad (3.11)$$

tomando agora o logaritmo de $l(\underline{\phi}, \tau, \lambda | \underline{Z})$, temos :

$$l(\underline{\phi}, \tau, \lambda | \underline{Z}) \propto \frac{n-p}{2} \ln \tau - \frac{\tau}{2} \left(\underline{Z}^{(\lambda)} - \underline{X}^{(\lambda)} \underline{\phi}\right)^T \left(\underline{Z}^{(\lambda)} - \underline{X}^{(\lambda)} \underline{\phi}\right) + (\lambda-1) \sum_{i=p+1}^n \ln Z_i \quad (3.12)$$

para determinar o estimador de máxima verossimilhança para λ temos a função de verossimilhança "profile" (Box e Cox, 1964), obtida substituindo-se os estimadores $\hat{\underline{\phi}}$ e $\hat{\tau}$ em (3.12). Assim temos :

$$lp(\lambda | \hat{\underline{\phi}}, \hat{\tau}, \underline{Z}) \propto \frac{n-p}{2} \ln \hat{\tau} + (\lambda-1) \sum_{i=p+1}^n \ln Z_i \quad (3.13)$$

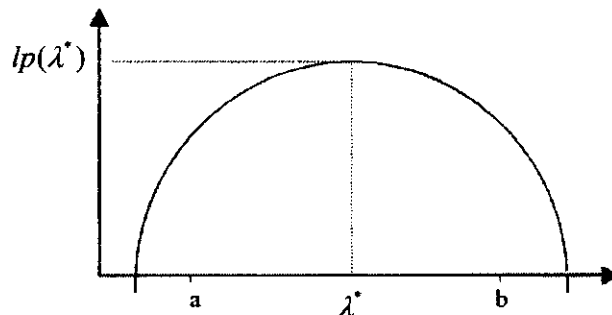
as estimativas $\hat{\underline{\phi}}$ e $\hat{\tau}$ são dada por :

$$\hat{\underline{\phi}} = \left(\underline{X}^{(\lambda)T} \underline{X}^{(\lambda)}\right)^{-1} \left(\underline{X}^{(\lambda)T} \underline{Z}^{(\lambda)}\right) \quad (3.14)$$

$$\hat{\tau}^{-1} = \frac{1}{n-p} \left(\underline{Z}^{(\lambda)} - \underline{X}^{(\lambda)} \hat{\underline{\phi}}\right)^T \left(\underline{Z}^{(\lambda)} - \underline{X}^{(\lambda)} \hat{\underline{\phi}}\right) \quad (3.15)$$

Assim, escolhendo-se $\lambda \in [a, b]$, fazemos variar λ neste intervalo e avaliamos este na função de verossimilhança "profile", graficamente teríamos :

Gráfico 1 : λ estimado que maximiza a função de verossimilhança "profile".



Então escolhe-se $\hat{\lambda}$ como o estimador de máxima verossimilhança "profile" para λ .

3.4.1 TRANSFORMAÇÕES DE BOX E COX PARA SERIES DE VAZÕES :

As séries de vazões são séries sazonais que usualmente apresentam comportamentos repetitivos por anos, tendo grande influência as precipitações pluviais e as mudanças de temperatura que acompanham as diferentes estações do ano. Assim, as transformações de Box e Cox são aplicadas a uma serie sazonal, onde temos :

$$q_{t(r,m)}^{(\lambda)} = \begin{cases} \frac{q_{t(r,m)}^{\lambda} - 1}{\lambda}, & \lambda \neq 0 \\ \ln q_{t(r,m)}, & \lambda = 0 \end{cases} \quad (3.16)$$

a serie $q_{t(r,m)}^{(\lambda)}$ então é padronizada para estimar a componente sazonal na média e a variância, ou seja :

$$Z_{t(r,m)} = \frac{q_{t(r,m)}^{(\lambda)} - \mu_m^{(\lambda)}}{\sigma_m^{(\lambda)}} \quad (3.17)$$

onde : $\mu_m^{(\lambda)} = \frac{1}{n} \sum_{r=1}^n q_{t(r,m)}^{(\lambda)}$, $m = 1, 2, \dots, 12$ e $\sigma_m^{(\lambda)2} = \frac{1}{n} \sum_{r=1}^n \left(q_{t(r,m)}^{(\lambda)} - \mu_m^{(\lambda)} \right)^2$

O modelo AR(p) para $Z_{t(r,m)}$ é dado por :

$$Z_{t(r,m)} = \sum_{j=1}^p \phi_j Z_{t(r,m)-j} + a_{t(r,m)} \quad (3.18)$$

A função de verossimilhança para os dados reais é escrita como :

$$L\left(\phi, \tau, \lambda \mid Z\right) \propto \prod_{t=p+1}^{n-s} \left[\tau^{1/2} \exp\left\{-\frac{\tau}{2} \left(Z_t - \sum_{j=1}^p \phi_j Z_{t-j} \right)^2\right\} \cdot J_{Z_t, q_t} \right] \quad (3.19)$$

neste caso o Jacobiano J_{Z_t, q_t} é dado por :

$$J_{Z_t, q_t} = \frac{\partial Z_t}{\partial q_t^{(\lambda)}} \cdot \frac{\partial q_t^{(\lambda)}}{\partial q_t} = \frac{1}{\sigma_m^{(\lambda)}} \cdot q_{t(r,m)}^{(\lambda)}$$

assim, temos :

$$l(\hat{\phi}, \hat{\tau}, \hat{\lambda} | Z) \propto \frac{n-p}{2} \ln \tau - \frac{\tau}{2} \left(Z - X \hat{\phi} \right)^T \left(Z - X \hat{\phi} \right) + \frac{n}{2} \sum_{m=1}^s \ln \sigma_m^{(\lambda)} + (\lambda - 1) \sum_{t=p+1}^{n-s} \ln q_{t(r,m)} \quad (3.20)$$

A função de verossimilhança "profile" é dada por :

$$lp(\hat{\lambda} | \hat{\phi}, \hat{\tau}, Z) \propto \frac{n-p}{2} \ln \tau + \frac{n}{2} \sum_{m=1}^s \ln \sigma_m^{(\lambda)} + (\lambda - 1) \sum_{t=p+1}^{n-s} \ln q_{t(r,m)} \quad (3.21)$$

3.4.2 PREVISÃO EM TRANSFORMAÇÕES DE BOX E COX :

O ajuste de modelos AR(p) com transformações de Box e Cox tem como um dos objetivos finais uma melhor predição de valores futuros para a série temporal em questão.

Seja :

$$q_{t(r,m)}^{(\lambda)} = \frac{q_{t(r,m)}^{\lambda} - 1}{\lambda} \quad (3.22)$$

com :

$$\mu_m^{(\lambda)} = \frac{1}{n} \cdot \sum_{\lambda=1}^n q_{t(r,m)}^{(\lambda)}, \quad m = 1, 2, \dots, s \quad \sigma_m^{(\lambda)^2} = \frac{1}{n-1} \cdot \sum_{\lambda=1}^n \left(q_{t(r,m)}^{(\lambda)} - \mu_m^{(\lambda)} \right)^2$$

e :

$$Z_{t(r,m)} = \frac{q_{t(r,m)}^{(\lambda)} - \mu_m^{(\lambda)}}{\sigma_m^{(\lambda)}} \quad (3.23)$$

Denotamos por :

$$Z_{t(r,m)} = T(q_{t(r,m)}) = \frac{q_{t(r,m)}^{\lambda} - 1}{\lambda} - \mu_m^{(\lambda)} \quad (3.24)$$

então :

$$q_{t(r,m)} = T^{-1}(Z_{t(r,m)}) = \left[\lambda (\sigma_m^{(\lambda)} Z_{t(r,m)} + \mu_m^{(\lambda)}) + 1 \right]^{1/\lambda} \quad (3.25)$$

A previsão da variável Z_t é :

$$\hat{Z}_{t(r,m)} = E[Z_{t(r,m)} | Z] \quad (3.26)$$

sendo a_t i.i.d. $N(0, \sigma_a^2)$ então podemos escrever :

$$Z_{t(r,m)} = \hat{Z}_{t(r,m)} + a_t \quad (3.27)$$

uma primeira aproximação para a previsão de $q_{t(r,m)}$ pode ser calculada por :

$$\hat{\eta}_{t(r,m)} = T^{-1}\left(\hat{Z}_{t(r,m)}\right) \quad (3.28)$$

$$\hat{\eta}_{t(r,m)} = \left[\lambda \left(\sigma_m^{(\lambda)} \hat{Z}_{t(r,m)} + \mu_m^{(\lambda)} \right) + 1 \right]^{1/\lambda}$$

Esta primeira aproximação não é a previsão MMSE e sim a previsão que minimiza a média do erro absoluto MMAE, ou seja $\hat{\eta}_{t(r,m)}$ é a mediana da função de densidade condicional de $q_{t(r,m)}$. Isso ocorre porque a_t tem distribuição simétrica e a média e a mediana da densidade condicional de $Z_{t(r,m)}$ coincidem e são dadas por $\hat{Z}_{t(r,m)}$. Assim sendo $q_{t(r,m)} = T^{-1}\left(Z_{t(r,m)}\right)$ é uma função crescente de $Z_{t(r,m)}$, então o percentil 50% de $q_{t(r,m)}$ é a transformação inversa do percentil 50% de $Z_{t(r,m)}$.

Para $\lambda \neq 1$ a função de densidade condicional de $q_{t(r,m)}$ é assimétrica, então $\hat{\eta}_{t(r,m)}$ não coincide com a média da densidade condicional de $q_{t(r,m)}$, denotando a média condicional por :

$$\hat{q}_{t(r,m)} = E\left[q_{t(r,m)} \mid q \right] \quad (3.29)$$

$\hat{q}_{t(r,m)}$ é a previsão MMSE para $q_{t(r,m)}$. Podemos deduzir $\hat{q}_{t(r,m)}$ em função de $\hat{Z}_{t(r,m)}$ adotando a transformação inversa dada por :

$$\begin{aligned} q_{t(r,m)} &= T^{-1}\left(Z_{t(r,m)}\right) \\ \hat{q}_{t(r,m)} &= T^{-1}\left(\hat{Z}_{t(r,m)} + a_t\right) \end{aligned} \quad (3.31)$$

então $\hat{q}_{t(r,m)}$ pode ser calculada usando a densidade de probabilidade de a_t , por :

$$\hat{q}_{t(r,m)} = \int_{-\infty}^{+\infty} T^{-1}\left(\hat{Z}_{t(r,m)} + a_t\right) f(a_t) da_t \quad (3.32)$$

$$\hat{q}_{t(r,m)} = \int_{-\infty}^{+\infty} \left[\lambda \left(\sigma_m^{(\lambda)} \left(\hat{Z}_{t(r,m)} + a_t \right) + \mu_m^{(\lambda)} \right) + 1 \right]^{1/\lambda} f(a_t) da_t$$

então :

$$\hat{q}_{t(r,m)} = \int_{-\infty}^{+\infty} \left[\lambda \left(\sigma_m^{(\lambda)} \hat{Z}_{t(r,m)} + \mu_m^{(\lambda)} \right) + 1 + \lambda \sigma_m^{(\lambda)} a_t \right]^{1/\lambda} f(a_t) da_t$$

$$\hat{q}_{t(r,m)} = \int_{-\infty}^{+\infty} \lambda \left(\sigma_m^{(\lambda)} \hat{Z}_{t(r,m)} + \mu_m^{(\lambda)} \right) + 1 \Big]^{\frac{1}{\lambda}} [1 + ra_t]^{\frac{1}{\lambda}} f(a_t) da_t, \quad (3.33)$$

onde :

$$r = \frac{\lambda \sigma_m^{(\lambda)}}{\left[\lambda \left(\sigma_m^{(\lambda)} \hat{Z}_{t(r,m)} + \mu_m^{(\lambda)} \right) + 1 \right]} = \frac{\sigma_m^{(\lambda)}}{\left(\sigma_m^{(\lambda)} \hat{Z}_{t(r,m)} + \mu_m^{(\lambda)} \right) + \frac{1}{\lambda}}$$

sendo a primeira aproximação para a previsão :

$$\hat{\eta}_{t(r,m)} = \left[\lambda \left(\sigma_m^{(\lambda)} \hat{Z}_{t(r,m)} + \mu_m^{(\lambda)} \right) + 1 \right]^{\frac{1}{\lambda}}$$

temos a previsão MMSE de $q_{t(r,m)}$ dada por :

$$\hat{q}_{t(r,m)} = \hat{\eta}_{t(r,m)} \cdot \int_{-\infty}^{+\infty} [1 + ra_t]^{\frac{1}{\lambda}} f(a_t) da_t, \quad (3.34)$$

denotando por $w_t = a_t / \sigma_a$, com $w_t \sim N(0,1)$, assim temos :

$$\hat{q}_{t(r,m)} = \hat{\eta}_{t(r,m)} \cdot \int_{-\infty}^{+\infty} [1 + uw_t]^{\frac{1}{\lambda}} f(w_t) dw_t, \quad (3.35)$$

onde :

$$u = \frac{\sigma_a \sigma_m^{(\lambda)}}{\left(\sigma_m^{(\lambda)} \hat{Z}_{t(r,m)} + \mu_m^{(\lambda)} \right) + \frac{1}{\lambda}}$$

calculamos uma aproximação para a integral em (3.35) :

$$G = \int_{-\infty}^{+\infty} [1 + uw_t]^{\frac{1}{\lambda}} f(w_t) dw_t,$$

Método de Monte Carlo :

A aproximação de G por Monte Carlo, consiste em gerar uma amostra de tamanho M (grande) da variável aleatória i.i.d. $w_t \sim N(0,1)$ que denotamos por $w_t^{(j)}$ para $j=1,2,\dots,M$, então avaliamos G por :

$$\hat{G} = \frac{1}{M} \cdot \sum_{j=1}^M [1 + uw_t^{(j)}]^{\frac{1}{\lambda}} \quad (3.36)$$

Método Numérico :

Para resolver a integral G numericamente, vamos rescrever essa integral como :

$$G = \int_{-\infty}^{+\infty} [1 + uw_t]^{1/\lambda} f(w_t) dw_t$$

Denotando por :

$$g(w_t) = \left[\frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{w_t^2}{2}\right\} \right] / \int_{-w_t}^{+w_t} f(w_t) dw_t \quad (3.37)$$

então $g(w_t)$ é uma densidade truncada para w_t , dessa forma temos :

$$G = \int_{-\infty}^{+\infty} [1 + uw_t]^{1/\lambda} g(w_t) dw_t \quad (3.38)$$

existem restrições para u impostas pela restrição de que $q_{t(r,m)} > 0$, essa condição na transformação (3.25) implica que :

$$q_{t(r,m)}^\lambda = \lambda q_{t(r,m)}^{(\lambda)} + 1$$

$$q_{t(r,m)}^\lambda = \lambda \left(\sigma_m^{(\lambda)} Z_{t(r,m)} + \mu_m^{(\lambda)} \right) + 1 > 0$$

substituindo $Z_{t(r,m)} = \hat{Z}_{t(r,m)} + \sigma_a w_t$, temos :

$$q_{t(r,m)}^\lambda = \lambda \left[\sigma_m^{(\lambda)} \left(\hat{Z}_{t(r,m)} + \sigma_a w_t \right) + \mu_m^{(\lambda)} \right] + 1 > 0 \quad (3.39)$$

então se $\lambda > 0$:

$$\sigma_m^{(\lambda)} \left(\hat{Z}_{t(r,m)} + \sigma_a w_t \right) + \mu_m^{(\lambda)} > -\frac{1}{\lambda}$$

$$\sigma_m^{(\lambda)} \left(\hat{Z}_{t(r,m)} + \sigma_a w_t \right) > -\mu_m^{(\lambda)} - \frac{1}{\lambda} \Rightarrow \sigma_m^{(\lambda)} \hat{Z}_{t(r,m)} + \sigma_m^{(\lambda)} \sigma_a w_t > -\mu_m^{(\lambda)} - \frac{1}{\lambda}$$

$$w_t > \frac{1}{\sigma_m^{(\lambda)} \sigma_a} \left[\sigma_m^{(\lambda)} \hat{Z}_{t(r,m)} + \mu_m^{(\lambda)} + \frac{1}{\lambda} \right] \Rightarrow w_t > \frac{1}{u}$$

Agora se $\lambda < 0$ temos de (3.39) que :

$$\sigma_m^{(\lambda)} \left(\hat{Z}_{t(r,m)} + \sigma_a w_t \right) + \mu_m^{(\lambda)} < -\frac{1}{\lambda}$$

que resulta em :

$$w_i < -\frac{1}{\sigma_m^{(\lambda)} \sigma_a} \left[\sigma_m^{(\lambda)} Z_{t(r,m)} + \mu_m^{(\lambda)} + \frac{1}{\lambda} \right] \Rightarrow w_i < -\frac{1}{u}$$

Considerando como ponto de truncamento da $N(0,1)$ $|w^*| \geq 4$ teremos $|w_i| \geq 4$, assim devemos ter: $|u_i| \leq \frac{1}{4}$.

Como aproximação numérica para G , vamos adotar o método de Gauss, onde :

$$G = \int_{-w^*}^{+w^*} (1 + uw_i)^{1/\lambda} g(w_i) dw_i \quad (3.40)$$

sendo :

$$I = \int_a^b \varphi(x) dx = \frac{b-a}{2} \cdot \sum_{i=1}^n A_i \varphi(x_i)$$

com :

$$x_i = \frac{b+a}{2} + \frac{b-a}{2} \cdot t_i, \quad i=1,2,\dots,n$$

por exemplo, para :

i	t_i	A_i
1, 8	± 0.96028986	0.10122854
2, 7	± 0.79666648	0.22238104
3, 6	± 0.52553242	0.31370664
4, 5	± 0.18343464	0.36268378

com $b = 4$, $a = -4$

temos : $x_i = 4t_i$, assim :

$$\varphi(x_i) = (1 + ux_i)^{1/\lambda} g(x_i)$$

e :

$$g(x_i) = \frac{\frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{x_i^2}{2}\right\}}{\Phi(b) - \Phi(a)}$$

3.5 TRATAMENTO CLÁSSICO EM OUTROS TIPOS DE TRANSFORMAÇÃO EM PROCESSOS AUTO-REGRESSIVOS.

Suponhamos agora a série z_t , mas estamos interessados nas previsões da série transformada :

$$Y_t = T(z_t) \quad (3.41)$$

onde $T(\cdot)$ é uma função bem comportada. Supondo que z_t é um processo gaussiano estacionário e que a função $T(\cdot)$ pode ser expandida numa série de polinômios Hermite, podemos derivar as propriedades da função de autocovariâncias de Y_t .

O uso do polinômio Hermite foi introduzido por Barrett e Lampard (1955) e foi considerado mais recentemente por Donelson e Maltz (1972) em processos de tempo contínuo. As propriedades mais importantes do polinômio Hermite são apresentadas no apêndice B e é discutida a expansão de algumas funções em series deste polinômio.

3.5.1 PROPRIEDADES DA FUNÇÃO DE AUTO-CORRELAÇÃO EM SERIES TRANSFORMADAS

PROCESSOS ESTACIONÁRIOS:

Seja z_t uma série de tempo estacionária com média μ , variância, σ^2 , e autocorrelação $corr(z_t, z_{t-\tau}) = \rho_\tau$. Com:

$$Z_t = \frac{z_t - \mu}{\sigma} \text{ para todo } t.$$

Os Z_t e $Z_{t-\tau}$ tem distribuição conjunta Normal bivariada com médias zero, variância unitária e correlação ρ_τ .

Consideremos agora uma transformação da forma $Y_t = T(Z_t)$, com expansão :

$$T(Z) = \sum_{j=0}^M \alpha_j H_j(Z),$$

temos que:

$$E(Y_{t-\tau}) = E\left\{ \sum_{j=0}^M \alpha_j H_j(Z_{t-\tau}) \right\} = \alpha_0 \quad (3.42)$$

pelo anexo (B.2), a série transformada tem média α_0 . Usando o anexo (B.4) :

$$E(Z_t Y_{t-\tau}) = E\left\{H_1(Z_t) \sum_{j=0}^M \alpha_j H_j(Z_{t-\tau})\right\} = \alpha_1 \rho_\tau \quad (3.43)$$

a covariância entre a série transformada e a série original é $COV(Z_t Y_{t-\tau}) = \alpha_1 \rho_\tau \sigma$.

Usando (3.42), a autocovariância da serie transformada é dada por:

$$C_{Y,\tau} = COV(Y_t Y_{t-\tau}) = E\left\{\sum_{j=1}^M \alpha_j H_j(Z_t) \sum_{i=1}^M \alpha_i H_i(Z_{t-\tau})\right\} = \sum_{j=1}^M \alpha_j^2 j! \rho_\tau^j \quad (3.44)$$

pelo anexo (B.1). Em particular, quando $\tau = 0$ temos:

$$Var(Y_t) = \sum_{j=1}^M \alpha_j^2 j!$$

Assim, a autocorrelação da série transformada é:

$$\rho_{Y,\tau} = \frac{\sum_{j=1}^M \alpha_j^2 j! \rho_\tau^j}{\sum_{j=1}^M \alpha_j^2 j!}$$

Consequentemente, se $\rho_\tau \neq 0$ para algum $\tau \neq 0$, segue que:

$$|\rho_{Y,\tau}| < |\rho_\tau| \text{ para } M > 1.$$

Deste modo, a série transformada é mais fechada para um processo de ruído branco que a série original.

Agora consideramos alguns exemplos específicos destas transformações.

Exemplo 1: Transformações Quadráticas

Considere primeiro a transformação:

$$Y_t = a + bz_t + cz_t^2 = a + b(\sigma Z_t + \mu) + c(\sigma Z_t + \mu)^2 = a + b\mu + c\mu^2 + (b + 2c\mu)\sigma Z_t + c\sigma^2 Z_t^2$$

Esta pode ser escrita como:

$$Y_t = T(Z_t) = \alpha_0 H_0(Z_t) + \alpha_1 H_1(Z_t) + \alpha_2 H_2(Z_t) = \alpha_0 + \alpha_1 Z_t + \alpha_2 (Z_t^2 - 1)$$

onde:

$$\alpha_0 = a + b\mu + c(\mu^2 + \sigma^2), \quad \alpha_1 = (b + 2c\mu)\sigma, \quad \alpha_2 = c\sigma^2 \quad (3.45)$$

isto segue de (3.44) que a autocovariância da serie transformada é:

$$C_{Y,t} = \alpha_1^2 \rho_\tau + 2\alpha_2^2 \rho_\tau^2 \quad (3.46)$$

onde α_1 e α_2 são dados em (3.45). Como tínhamos observado em (3.42), a media da serie transformada es α_0 .

Exemplo 2: Transformação Quadrática dos processos auto-regressivos AR(p)

Suponhamos agora, que a serie original z_t segue de um processo auto-regressivo AR(p). Então podemos escrever:

$$\rho_\tau = \sum_{j=1}^p a_j G_j^\tau, \quad |G_j| < 1, \quad j=1,2,\dots,p.$$

Deste modo:

$$\rho_\tau^2 = \sum_{i=1}^p a_i^2 G_i^{2\tau} + 2 \sum_{j < k} a_j a_k (G_j G_k)^\tau$$

segue de (3.46). Assim, em geral:

$$C_{Y,t} = \sum_{j=1}^p b_j F_j^\tau,$$

onde $p^* = p + \frac{1}{2}p(p+1) = \frac{1}{2}p(p+3)$. Deste modo, a série transformada terá a mesma estrutura de autocovariâncias que um processos AR(1). Isto pode ser usado em processos

ARMA $\left\{ \frac{1}{2}p(p+3), \frac{1}{2}p(p+1) \right\}$.

Exemplo 3: O quadrado de um processo AR(1)

Com os exemplos 1 e 2 podemos considerar o caso onde z_t é um processo auto-regressivo AR(1): $\hat{z}_t - \phi \hat{z}_{t-1} = a_t$, onde $\hat{z}_t = z_t - \mu$ e a_t é um ruído branco com variância σ_a^2 . Consequentemente $\rho_r = \phi^r$ e $\sigma^2 = \sigma_a^2 / (1 - \phi^2)$. Consideremos a transformação $Y_t = z_t^2$.

De (3.45) no exemplo 1, vemos que isto pode ser escrito como:

$$Y_t = T(Z_t) = \alpha_0 H_0(Z_t) + \alpha_1 H_1(Z_t) + \alpha_2 H_2(Z_t)$$

onde: $\alpha_0 = \mu^2 + \sigma^2$, $\alpha_1 = 2\mu\sigma$, $\alpha_2 = \sigma^2$. Consequentemente de (3.42), observamos que a série transformada tem média $\mu^2 + \sigma^2$ e de (3.46) que a estrutura de autocovariâncias é dada por:

$$C_{Y,t} = \alpha_1^2 \rho_t + 2\alpha_2 \rho_t^2 = \sigma^2 (4\mu^2 \phi^{2t} + 2\sigma^2 \phi^{2t}) \quad (3.47)$$

Então Y_t terá a mesma estrutura de autocovariância como $Y_{1,t} + Y_{2,t}$, onde $Y_{1,t} - \phi Y_{1,t-1} = e_{1,t}$, $Y_{2,t} - \phi^2 Y_{2,t-1} = e_{2,t}$; $e_{1,t}$ e $e_{2,t}$ são independentes do processo ruído branco com variâncias $4\mu^2 \sigma^2 (1 - \phi^2)$ e $2\sigma^4 (1 - \phi^4)$ respectivamente. Consequentemente Y_t tem as mesmas propriedades das autocovariâncias da série:

$$(1 - \phi B)(1 - \phi^2 B)Y_t = (1 - \Theta B)e_t.$$

Onde B é o operador usual utilizado em séries temporais. Como é bem conhecido, este modelo pode ser escrito como:

$$(1 - \phi B)(1 - \phi^2 B)Y_t = (1 - \Theta B)e_t.$$

Note que se a média μ é muito grande comparado com σ , então o comportamento da série transformada se aproxima de um processo auto-regressivo AR(1) por (3.47).

3.5.2 PROPRIEDADES DA FUNÇÃO DE AUTO-CORRELAÇÃO EM SÉRIES TRANSFORMADAS:

PROCESSOS NÃO ESTACIONÁRIOS.

Nesta seção examinamos a aplicação dos métodos da seção anterior para séries temporais que requerem de diferenças finitas para produzir estabilidade na série. Estes casos são comuns em séries econômicas.

Seja z_t uma série temporal tal que:

$$\nabla z_t = u_t \quad (3.48)$$

Onde u_t é uma série temporal gaussiana e estacionária com média zero e variância σ_u^2 . Consideramos (3.48) como tendo sido gerado por um modelo da forma.

$$z_t = \mu + \sum_{j=1}^t u_j \quad (3.49)$$

com $t=0$ é $z_0 = \mu$. Podemos escrever:

$$C_{t,t} = E\left(\sum_{j=1}^t u_j \sum_{i=1}^t u_i\right) \quad (3.50)$$

Então de (3.49) é (3.50) z_t tem distribuição normal com média μ e variância $C_{t,t}$. Agora fazendo $Z_t = (z_t - \mu) / \sqrt{C_{t,t}}$, temos que Z_t e $Z_{t-\tau}$ tem distribuição normal bivariada com medias zero, variâncias unitárias e correlação $\rho_{t,t} = C_{t,t-\tau} / \sqrt{(C_{t,t} C_{t-\tau,t-\tau})}$.

Consideramos agora uma transformação da forma:

$$Y_t = T(Z_t) = \sum_{j=0}^M \alpha_j^{(t)} H_j(Z_t) \quad (3.51)$$

Usando a seção previa temos:

$$E(Y_t) = \alpha_0^{(t)} \quad (3.52)$$

Assim, por um argumento exatamente análogo da seção anterior, temos:

$$\text{cov}(Y_t, Y_{t-\tau}) = \sum_{j=1}^M \alpha_j^{(t)} \alpha_j^{(t-\tau)} j! \rho_{t,t-\tau}^j$$

e conseqüentemente:

$$E(Y_t Y_{t-\tau}) = \sum_{j=1}^M \alpha_j^{(t)} \alpha_j^{(t-\tau)} j! \rho_{t,t-\tau}^j + \alpha_0^{(t)} \alpha_0^{(t-\tau)} \quad (3.53)$$

Agora, o nosso interesse estará em ∇Y_t , no lugar de Y_t . De (3.52) e (3.53) temos que:

$$E(\nabla Y_t) = \alpha_0^{(t)} - \alpha_0^{(t-1)} \quad (3.54)$$

é

$$COV(\nabla Y_t, \nabla Y_{t-\tau}) = \sum_{j=1}^M \alpha_j^{(t)} \alpha_j^{(t-\tau)} j! \rho_{t,t-\tau}^j + \sum_{j=1}^M \alpha_j^{(t-1)} \alpha_j^{(t-1-\tau)} j! \rho_{t-1,t-1-\tau}^j - \sum_{j=1}^M \alpha_j^{(t)} \alpha_j^{(t-1-\tau)} j! \rho_{t,t-1-\tau}^j - \sum_{j=1}^M \alpha_j^{(t-1)} \alpha_j^{(t-\tau)} j! \rho_{t-1,t-\tau}^j \quad (3.55)$$

Os resultados em (3.54) e (3.55) são completamente gerais, embora devemos notar que, as variâncias em lugar das esperadas em (3.55) são explosivas. Consideremos agora dois exemplos específicos.

Exemplo 1: Transformações Quadráticas

Examinando novamente a transformação $Y_t = a + bz_t + cz_t^2$. Esta pode ser escrita como $Y_t = T(Z_t) = \alpha_0^{(t)} + \alpha_1^{(t)} Z_t + \alpha_2^{(t)} (Z_t^2 - 1)$, onde :

$$\alpha_0^{(t)} = a + b\mu + c(\mu^2 + C_{t,t}), \quad \alpha_1^{(t)} = (b + 2c\mu)C_{t,t}^{\frac{1}{2}}, \quad \alpha_2^{(t)} = cC_{t,t} \quad (3.56)$$

De (3.54) temos que a média de ∇Y_t é dada por:

$$E(\nabla Y_t) = c(C_{t,t} - C_{t-1,t-1}) = cE\left(u_t^2 + 2u_t \sum_{j=1}^{t-1} u_j\right) \quad (3.57)$$

Notemos que para t grande e assumindo que a série começa em $t=1$ (em (3.49)), a $E(\nabla Y_t)$ é quase constante se u_t constitui um processo estacionário. Substituindo $M=2$ e (3.56) em (3.55), as autocovariâncias são obtidas como:

$$\text{cov}(\nabla Y_t, \nabla Y_{t-\tau}) = (b + 2c\mu)^2 (C_{t,t-\tau} + C_{t-1,t-\tau-1} - C_{t,t-\tau-1} - C_{t-1,t-\tau}) + 2c^2 (C_{t,t-\tau}^2 + C_{t-1,t-\tau-1}^2 - C_{t-1,t-\tau}^2) \quad (3.58)$$

Assim, em particular, se $Y_t = z_t^2$ e substituindo $b=0$ e $c=1$ em (3.57) e (3.58), temos :

$$E(\nabla Y_t) = E\left(u_t^2 + 2u_t \sum_{j=1}^{t-1} u_j\right) \quad (3.59)$$

e:

$$\text{cov}(\nabla Y_t, \nabla Y_{t-1}) = 4\mu^2 (C_{t,t-1} + C_{t-1,t-2} - C_{t,t-2} - C_{t-1,t-1}) + 2(C_{t,t-1}^2 + C_{t-1,t-2}^2 - C_{t,t-2}^2 - C_{t-1,t-1}^2) \quad (3.60)$$

Exemplo 2: O quadrado do processo AR(1) com uma diferencia finita.

Consideramos o caso onde u_t é gerada por um processo AR(1) : $u_t - \phi u_{t-1} = a_t$, então z_t é um processo ARIMA(1,1,0) e $\nabla z_t - \phi \nabla z_{t-1} = a_t$.

De (3.59) temos:

$$E(\nabla Y_t) = C_{t,t} - C_{t-1,t-1} = \sigma_u^2 + 2\sigma_u^2 \sum_{j=1}^{t-1} \phi^j = \sigma_u^2 \left\{ 1 + \frac{2\phi(1-\phi^{t-1})}{1-\phi} \right\}.$$

Agora, supondo que a série começou em um tempo p pequeno, e que t é grande e ϕ^t desprezível, temos que $E(\nabla Y_t) = C_{t,t} - C_{t-1,t-1} \cong \sigma_u^2(1+\phi)/(1-\phi)$. Então claramente $C_{t,t} - C_{t-\tau,t-\tau} \cong \tau\sigma_u^2(1+\phi)/(1-\phi)$. De (3.50) temos:

$$C_{t-j,t-j-\tau} = \frac{\sigma_u^2 \phi (1-\phi^{t-j-\tau}) (1-\phi^\tau)}{1-\phi^2} + C_{t,t} - \frac{(\tau+j)\sigma_u^2(1+\phi)}{1-\phi} \text{ se } j, \tau \geq 0,$$

ou, aproximadamente, para t grande,

$$C_{t-j,t-j-\tau} = \frac{\sigma_u^2 \phi}{1-\phi^2} (1-\phi^\tau) + C_{t,t} - \frac{(\tau+j)\sigma_u^2(1+\phi)}{1-\phi} \text{ se } j, \tau \geq 0.$$

de (3.50), $C_{t-j-t,t-j} = C_{t-j,t-j-t}$. Depois de algumas operações trabalhosas obtemos de (3.60) o seguinte resultado : $\text{cov}(\nabla Y_t, \nabla Y_{t-\tau}) = A(\tau) + B\phi^\tau + C\phi^{2\tau}$, onde $A(\tau) = 0$ para todo $\tau \neq 0$, $A(0) > 0$.

Deste modo ∇Y_t , tem a mesma estrutura de autocovariâncias que $\nabla Y_{1,t} + \nabla Y_{2,t} + \nabla Y_{3,t}$, onde:

$$\nabla Y_{1,t} - \phi \nabla Y_{1,t-1} = e_{1,t}, \quad \nabla Y_{2,t} - \phi^2 \nabla Y_{2,t-1} = e_{2,t}, \quad \nabla Y_{3,t} = e_{3,t},$$

onde $e_{1,t}$, $e_{2,t}$, $e_{3,t}$ são processos independentes ruído branco. Consequentemente ∇Y_t tem as propriedades das autocovariâncias do modelo:

$$(1 - \phi B)(1 - \phi^2 B)\nabla Y_t = (1 - \phi^2 B)e_{1,t} + (1 - \phi B)e_{2,t} + (1 - \phi B)(1 - \phi^2 B)e_{3,t}.$$

Este modelo pode ser escrito como:

$$(1 - \phi B)(1 - \phi^2 B)\nabla Y_t = (1 - \Theta_1 B - \Theta_2 B^2)e_t,$$

onde e_t é um ruído branco. Em geral, ∇Y_t tem as mesmas propriedades das autocovariâncias que um processo ARMA(2,2).

Como temos notado em alguns exemplos específicos, as autocovariâncias da série diferenciada não é geralmente invariante sobre transformações.

CAPÍTULO 4

APLICAÇÕES.

4.1 INTRODUÇÃO :

Neste capítulo apresentamos dois exemplos de séries, uma das quais precisa de transformação. A primeira delas trata com dados de energia desprendida por pacientes com queimaduras severas, esta série não precisa de transformação. O segundo conjunto de dados trata das vazões médias mensais no reservatório de Furnas, esta série precisa de transformação e daremos a esta especial atenção. Nestes duas séries a análise a posteriori é feita usando-se algoritmos de simulação em cadeias de Markov (MCMC).

4.2 VOLUME DE GÁS OXIGÊNIO EM PACIENTES COM QUEIMADURAS SEVERAS :

O primeiro exemplo corresponde à análise de dados de pacientes queimados, estes pacientes receberam queimaduras severas, a sua energia precisou ser monitorada de perto. As medidas são obtidas em períodos bases. Uma máscara metabólica analisa os gases inspirados e expirados e mede a quantidade de oxigênio e o dióxido de carbono, esta máscara metabólica inclui um equipamento que consiste num computador, um analisador de gás e um dispositivo de calibração interna. Baseado no volume de oxigênio e dióxido de carbono o equipamento pode calcular a despesa calórica do paciente. O método usado é chamado de calorimetria indireta. Apresentamos uma análise de 30 medidas sucessivas de volume de oxigênio tomadas em um paciente com queimaduras em 79% do seu corpo.

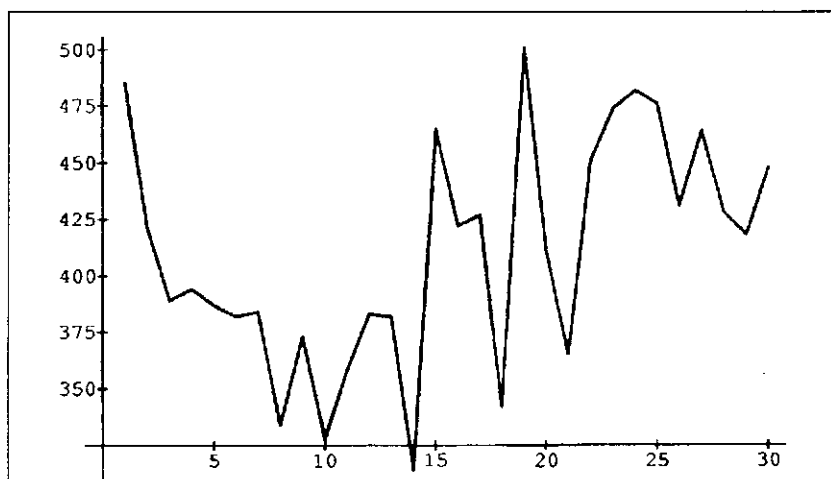


Figura 3 : Série temporal de volume de oxigênio em pacientes com queimaduras severas.

Um modelo auto-regressivo de primeira ordem é ajustado aos dados, são usadas prioris não informativa, normal-gama e t-gama. Foram geradas 5 cadeias independentes cada uma de tamanho 1000, as primeiras 400 observações são despreziadas (burn-in) e das 600 restantes são tomadas observações de 3 em 3, gerando uma amostra de tamanho 1000.

4.2.1 RESULTADOS OBTIDOS USANDO PRIORI NÃO INFORMATIVA :

HISTOGRAMAS DAS DENSIDADES MARGINAIS DE ϕ E τ :

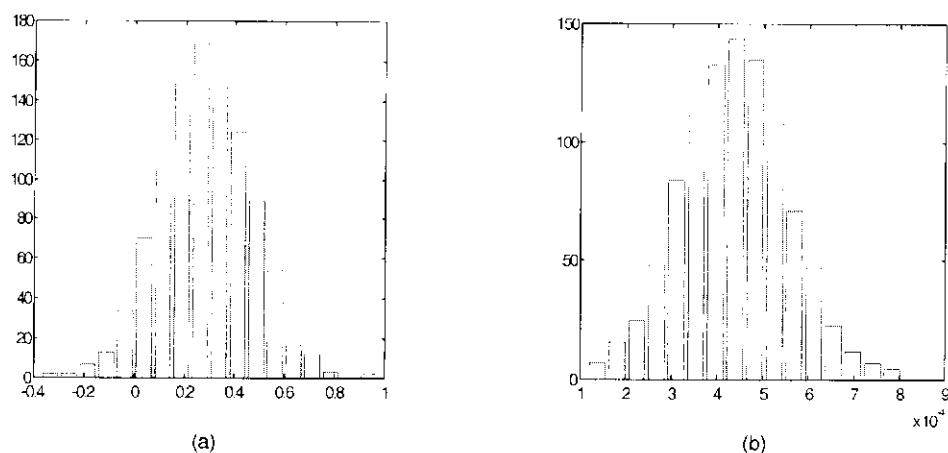


Figura 4 : Histograma de densidade marginal (a) para ϕ e (b) para τ .

PARÂMETROS ESTIMADOS :

ϕ :		τ :	
M	= 0.2786	M	= 4.3684e-004
S	= 0.1833	S	= 1.1576e-004
Linf	= -0.0795	Linf	= 2.4147e-004
Lsup	= 0.6026	Lsup	= 6.9699e-004

GRÁFICO DA DENSIDADE CONJUNTA DE ϕ E τ :

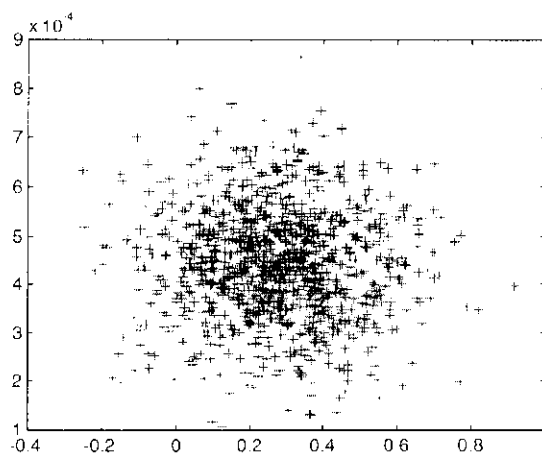


Figura 5 : Gráfico para a densidade conjunta de ϕ e τ .

GRÁFICO DE CONVERGÊNCIA PARA ϕ :

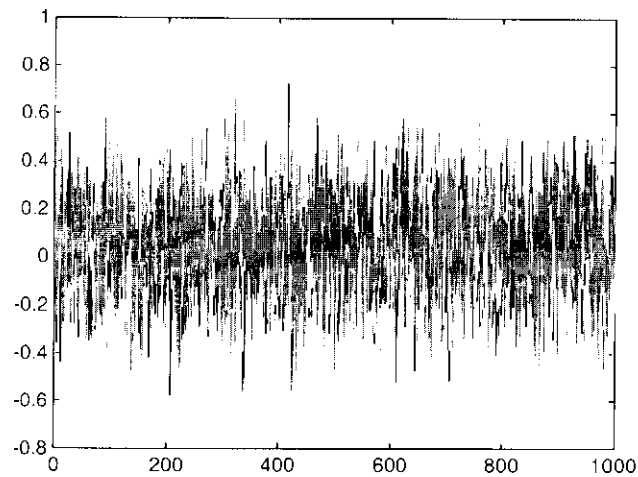


Figura 6 : Gráfico de convergência para as cadeias que geraram ϕ (Gelman Rubin = 1.0034).

4.2.2 RESULTADOS OBTIDOS USANDO PRIORI NORMAL-GAMA :

HISTOGRAMAS DAS DENSIDADES MARGINAIS DE ϕ E τ :

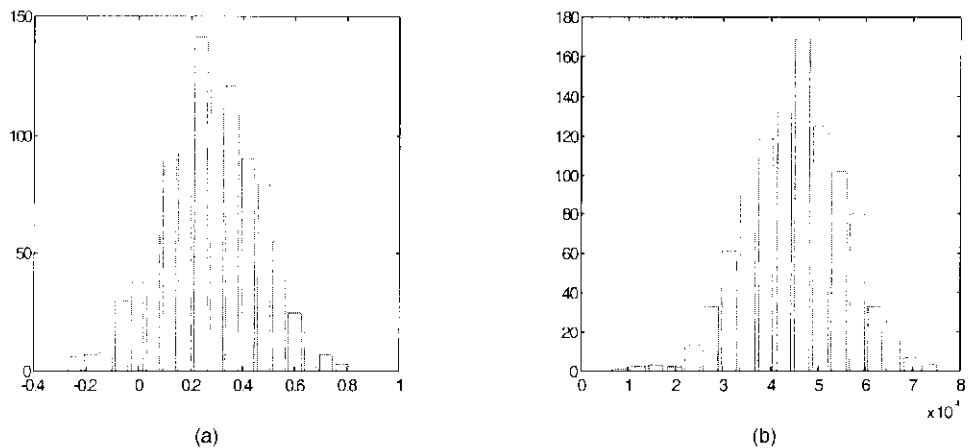


Figura 7 : Histograma de densidade marginal (a) para ϕ e (b) para τ .

PARÂMETROS ESTIMADOS :

ϕ :

M = 0.2724
S = 0.1796
Linf = -0.0764
Lsup = 0.6270

τ :

M = 4.5685e-004
S = 1.0289e-004
Linf = 2.7318e-004
Lsup = 6.7510e-004

GRÁFICO DA DENSIDADE CONJUNTA DE ϕ E τ :

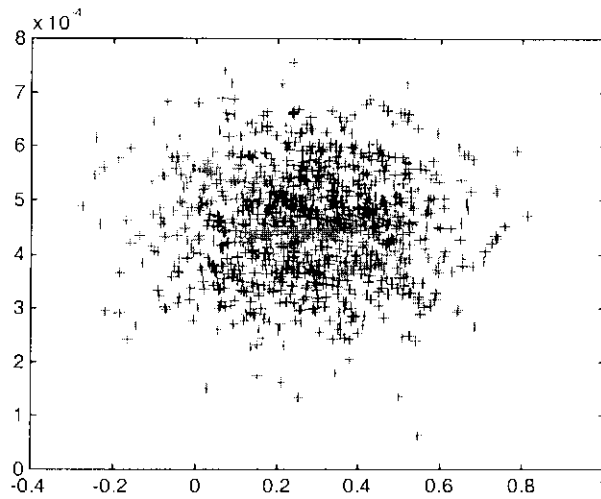


Figura 8 : Gráfico para a densidade conjunta de ϕ e τ .

GRÁFICO DE CONVERGÊNCIA PARA ϕ :

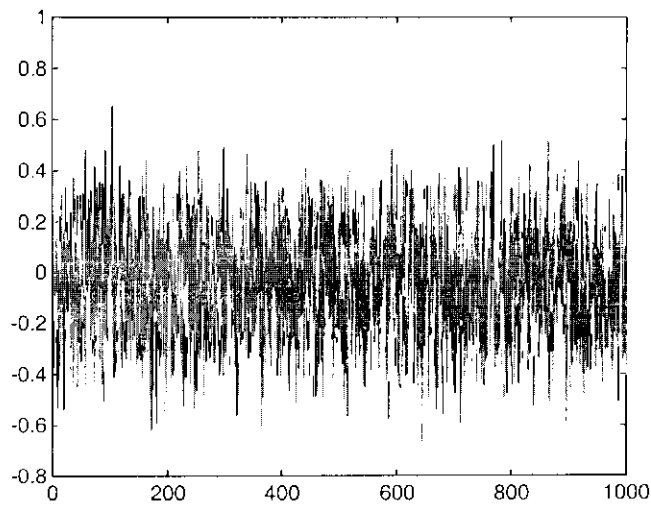


Figura 9 : Gráfico de convergência para as cadeias que geram ϕ (Gelman Rubin = 1.0008).

COMPARAÇÃO DOS RESULTADOS :

Tabela 1 : Comparação das estimativas para os parâmetros do modelo AR(1)
Volume de oxigênio em pacientes com queimaduras severas.

MÉTODO	ϕ_1	$Var(\phi_1)$	τ	$Var(\tau)$
E.M.V.	0,2803	0,1992	4,8461e-004	1,2020e-004
Não Informativa	0,2786	0,1873	4,3684e-004	1,1576e-004
Normal-Gama	0,2724	0,1796	4,5685e-004	1,0289e-004
T-Gama	0,2706	0,1802	4,6245e-004	1,0853e-004

4.2.3 CONCLUSÕES :

Estes resultados mostram pouca diferença nas estimativas, tendo uma variância ligeiramente maior nas estimativas clássicas, porém os resultados clássicos e bayesianos podem ser usados sem problemas. A FAC e FACP sugeriram um modelo AR(1) para esta série, assim não usaremos neste caso os critérios de seleção de modelos.

4.3 VAZÕES MÉDIAS MENSIS NO RESERVATÓRIO DE FURNAS - SÉRIE NORMALIZADA :

Este segundo exemplo corresponde as vazões médias mensais que chegam ao reservatório de Furnas no sudeste do Brasil. São analisados registros mensais de vazões num período de 30 anos. Estas vazões permitem estimar custos ligados ao processamento de energia elétrica durante determinados períodos de tempo que podem ser de meses ou de anos. Estes registros apresentam um comportamento sazonal anual e cíclico de aproximadamente 7 anos.

Assume-se para a série original o modelo $\log(X_{t(r,m)}) = \mu_m + \sigma_m Z_{t(r,m)}$ onde $m = 1, 2, \dots, 12$, representa o número de meses, $r = 1, 2, \dots, 30$, representa o número de anos, e $t(r,m) = (r-1)12 + m$, μ_m e σ_m são as médias e desvios-padrão estimados dos dados para cada mês do ano. Desta forma, a componente sazonal $Y_{t(r,m)} = \log(Z_{t(r,m)})$ é calculada como :

$$Z_{t(r,m)} = \frac{\log(X_{t(r,m)}) - \mu_m}{\sigma_m} \quad (4.1)$$

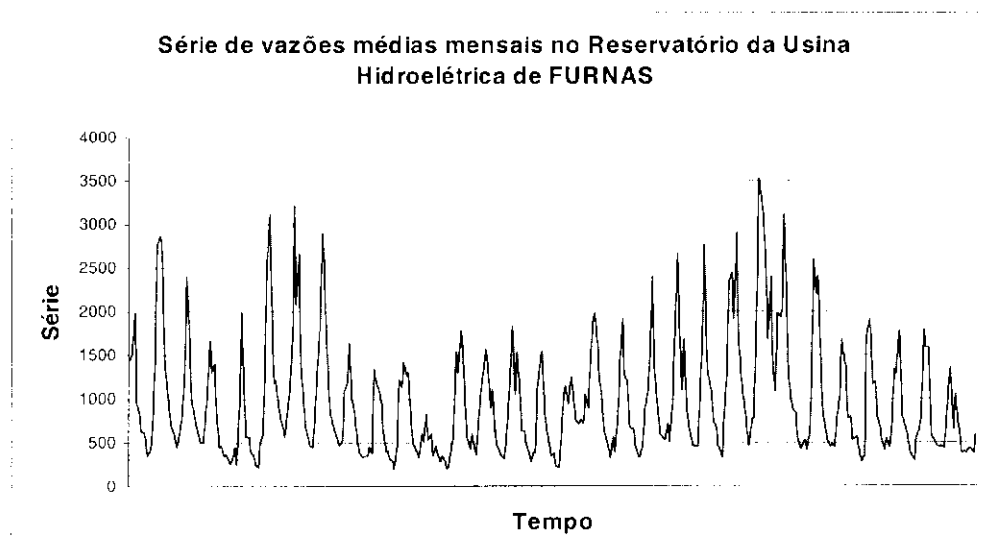


Figura 10 : Gráfico da série de vazões médias mensais no Reservatório da Usina Hidroelétrica de Furnas.

assim os dados originais são normalizados mediante (4.1). Para a série normalizada $Z_{(t, m)}$ ajusta-se um modelo AR(p): $Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + a_t$.

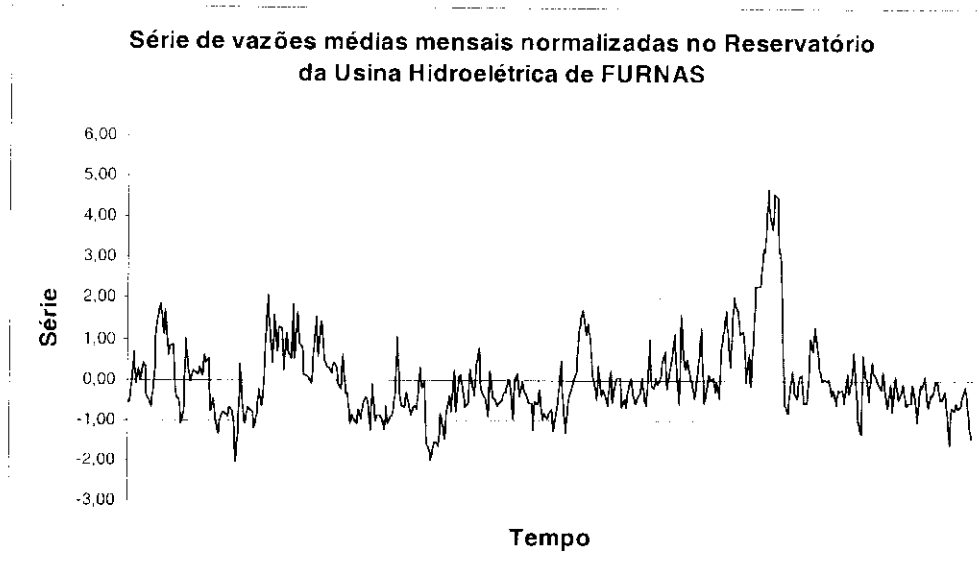


Figura 11 : Gráfico da série normalizada das vazões médias mensais na Usina Hidroelétrica de Furnas.

ESCOLHA DO MODELO :

A função de autocorrelação (FAC) e a função de autocorrelação parcial (FACP) para a série normalizada identificam um processo auto-regressivo de segunda ordem. A FAC apresenta um decaimento exponencial próprio dos processos auto-regressivos.

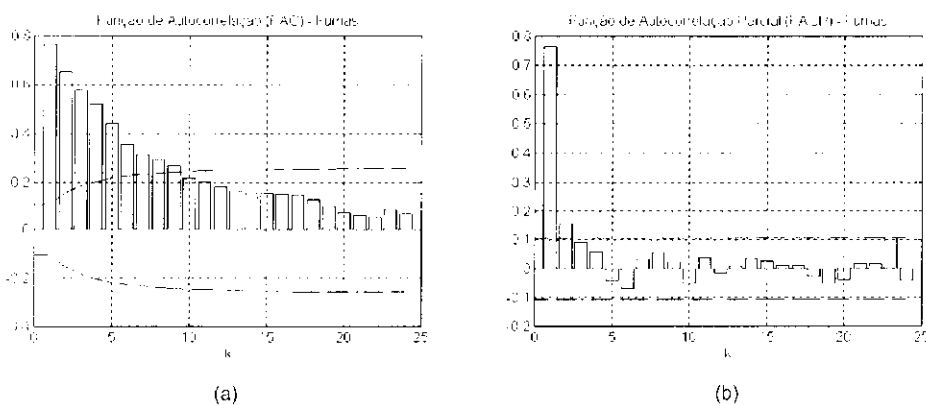


Figura 12 : (a) Função de Autocorrelação Simples e (b) Função de Autocorrelação Parcial.

As sugestões fornecidas pela análise gráfica da FAC e FACP acerca da ordem do modelo são confrontadas com técnicas numéricas. Apresentamos os critérios de seleção AIC, BIC e fator de Bayes.

Os critérios usados têm as seguintes formulações (Akaike, 1974; Schwarz, 1978; Raftery, 1996) :

$$AIC = \ln \left(\hat{\sigma}_a^2 \right) + \frac{2M}{N} \quad BIC = \ln \left(\hat{\sigma}_a^2 \right) + \frac{M \ln(N)}{N} \quad B_{ij} = \frac{f(Z|M_i)}{f(Z|M_j)}$$

CRITÉRIOS DE SELEÇÃO DE MODELOS :

Tabela 2 : Critérios de seleção de modelos - Série normalizada.

CRITÉRIOS	AR(1)	AR(2)	AR(3)
AIC	-1,0350	-1,0467	-1,0456
BIC	-1,0243	-1,0255	-1,0141
Fator de Bayes	0,0107	0,5581	0,1636

4.3.1 RESULTADOS OBTIDOS USANDO PRIORI NÃO INFORMATIVA :

HISTOGRAMAS DAS DENSIDADES MARGINAIS DE ϕ_1 , ϕ_2 E τ :

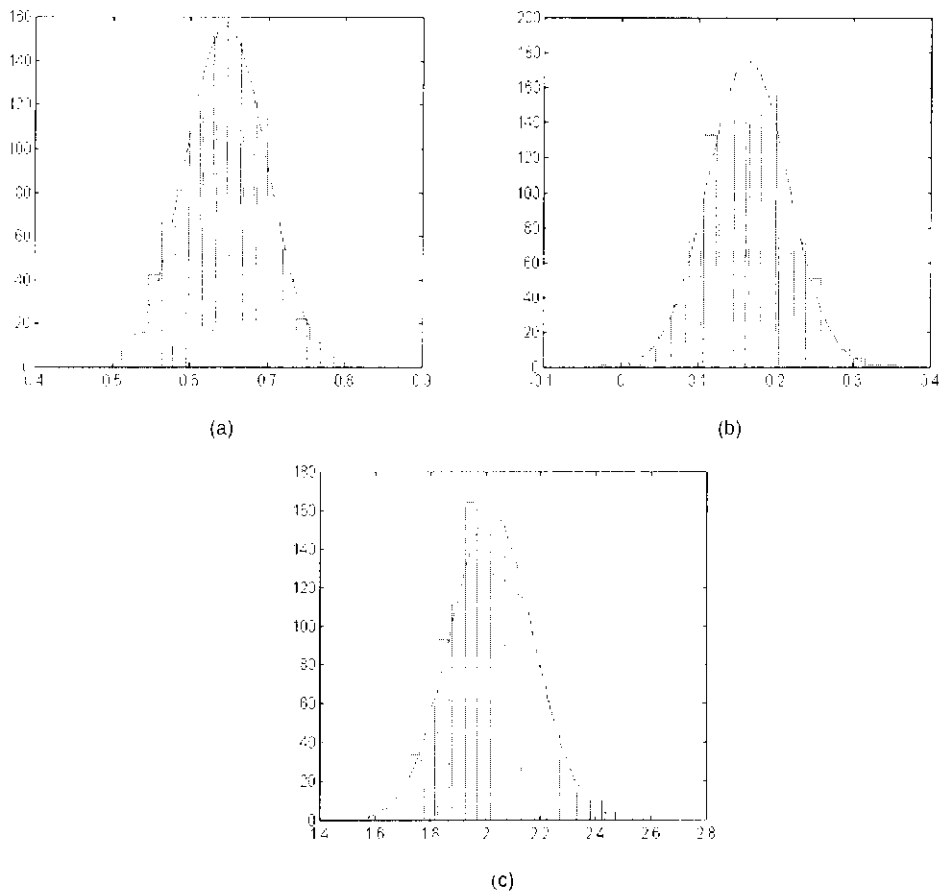


Figura 13 : Histograma de densidade marginal (a) para $\hat{\phi}_1$, (b) para $\hat{\phi}_2$ e (c) para τ .

4.3.2 RESULTADOS OBTIDOS USANDO PRIORI NORMAL-GAMA :

HISTOGRAMAS DAS DENSIDADES MARGINAIS DE $\hat{\phi}_1$, $\hat{\phi}_2$ E τ :

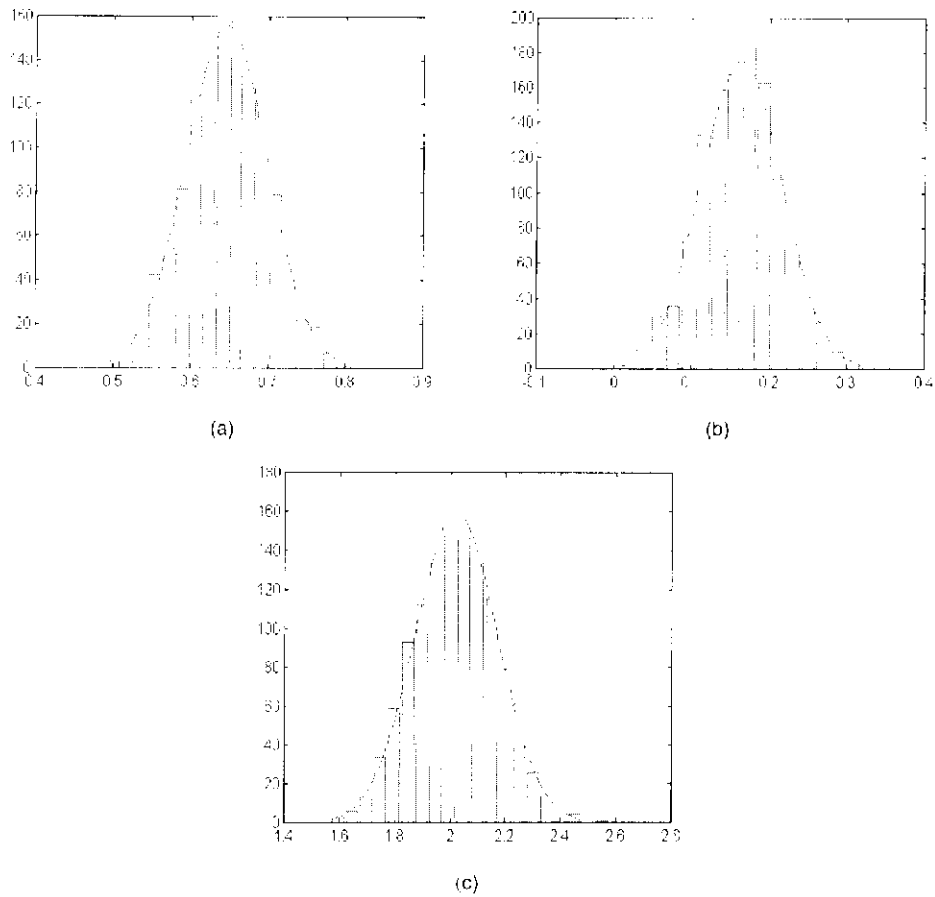
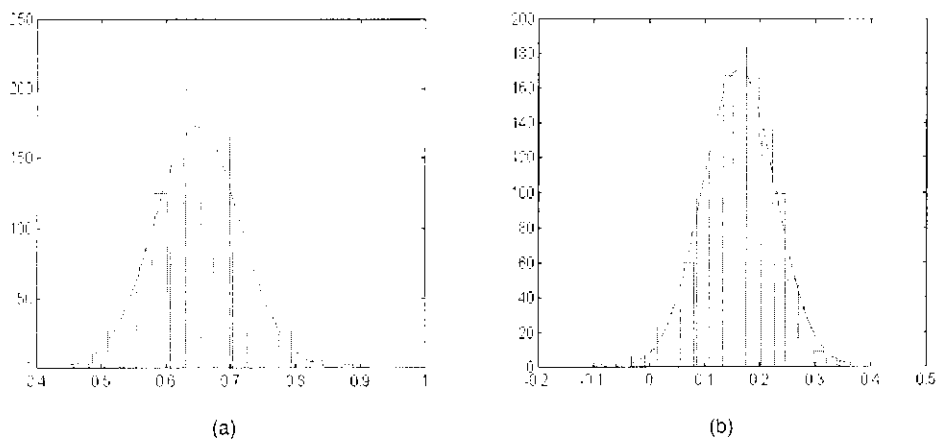
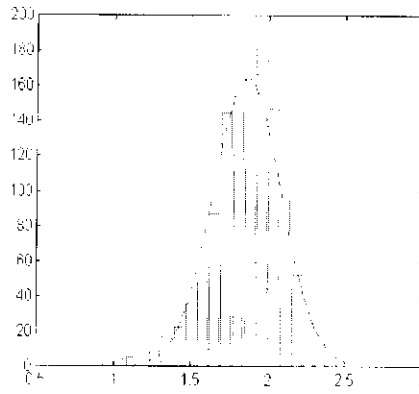


Figura 14 : Histograma de densidade marginal (a) para $\hat{\phi}_1$, (b) para $\hat{\phi}_2$ e (c) para τ .

4.3.3 RESULTADOS OBTIDOS USANDO PRIORI T-GAMA :

HISTOGRAMAS DAS DENSIDADES MARGINAIS DE $\hat{\phi}_1$, $\hat{\phi}_2$ E τ :





(c)

Figura 15 : Histograma de densidade marginal (a) para $\hat{\phi}_1$, (b) para $\hat{\phi}_2$ e (c) para τ .

COMPARAÇÃO DOS RESULTADOS :

Tabela 2 : Comparação das estimativas para os parâmetros do modelo AR(2).
Vazões médias mensais no Reservatório de Furnas - serie normalizada.

MÉTODO	ϕ_1	$Var(\phi_1)$	ϕ_2	$Var(\phi_2)$	τ	$Var(\tau)$
E.M.V.	0,6478	0,0027	0,1623	0,0027	2,0285	---
Não Informativa	0,6484	0,0029	0,1621	0,0029	2,0178	0,0222
Normal-Gama	0,6435	0,0026	0,1644	0,0028	1,9987	0,0217
T-Gama	0,6480	0,0025	0,1619	0,0027	1,9488	0,0317

4.3.4 CONCLUSÕES :

Esta série foi ajustada a um modelo auto-regressivo de segunda ordem usando diferentes prioris, o critério para a simulação das cadeias é o mesmo que para o primeiro exemplo. A análise gráfica é confirmada com os critérios de seleção do modelo, tendo como melhor ajuste o modelo AR(2), assim foram aplicados os métodos clássicos e bayesianos resultando em estimativas equivalentes.

4.4 VAZÕES MÉDIAS MENSAIS NO RESERVATÓRIO DE FURNAS - SÉRIE TRANSFORMADA :

Nesta seção analisamos a série de vazões medias mensais original aplicando uma transformação de potência aos dados, na tentativa de estabilizar a variância. Aqui usamos a alternativa que combina os métodos clássicos e bayesianos para estimar o parâmetro de transformação λ . Assumimos para estimar λ , que os dados podem ser ajustados a um processo auto-regressivo de primeira ordem, e logo após, com essa estimativa, ajustamos outras ordens para o processo auto-regressivo.

4.4.1 ALTERNATIVA BAYESIANA PARA ENCONTRAR A MARGINAL DE λ :

É complexo encontrar uma expressão fechada para a marginal de λ , ainda trabalhando com uma priori de expressão simples como a priori não informativa mostrada na seção 3.3.2 e ajustando os dados a um modelo auto-regressivo de primeira ordem. Assim tentou-se encontrar uma expressão para a distribuição condicional de λ dado ϕ a qual também resultou em uma expressão desconhecida.

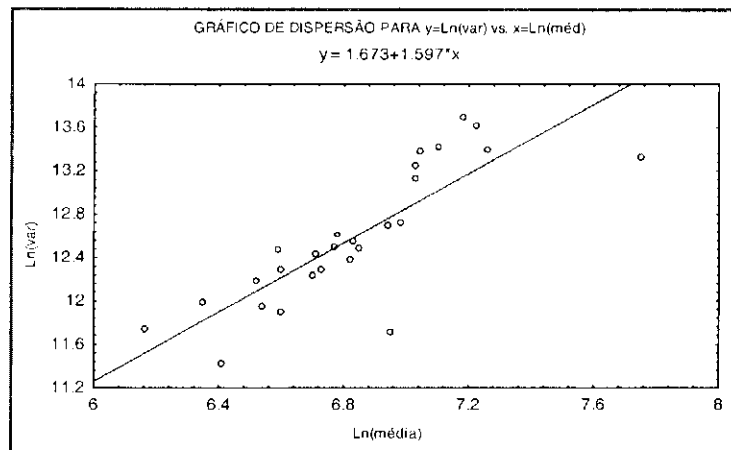
Dadas estas dificuldades foi proposto trabalhar com a moda a posteriori para λ , assim é tomado um valor inicial para o parâmetro e logo são gerados valores para ϕ e τ mediante as distribuições encontradas, esses valores são avaliados na função de verossimilhança e é obtido um novo valor para λ mediante o algoritmo de Newton-Rapshon. Estes passos serão repetidos até o valor de λ convergir. Um valor inicial para λ poderia ser tomado ajustando uma reta aos pontos:

$$\left(\ln \bar{x}_i, \ln S_i^2 \right) \quad i: 1, \dots, 30$$

sendo \bar{x}_i a média e S_i^2 a variância para cada ano da série, tomamos como valor inicial de λ

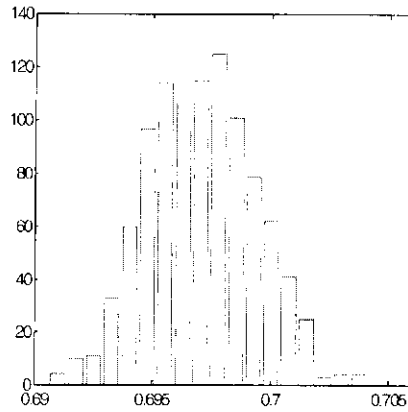
a: $\lambda^{(0)} = 1 - a$ - a inclinação da reta ajustada.

Valor inicial para λ :

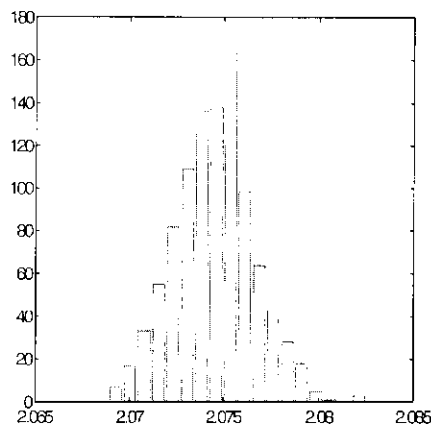


um novo valor para $\lambda^{(i)}$, este valor gera outra dupla $(\phi^{(i)}, \tau^{(i)})$, repetindo estes passos chegamos à convergência de $\hat{\lambda}$.

HISTOGRAMA DAS DENSIDADES MARGINAIS DE ϕ E τ :



(a)



(b)

Figura 17 : Histograma de densidade marginal (a) para ϕ e (b) para τ .

COMPARAÇÃO DOS RESULTADOS :

Tabela 3 : Comparação das estimativas para os parâmetros do modelo AR(1).
Vazões médias mensais no Reservatório de Furnas - série transformada.

MÉTODO	ϕ_1	$Var(\phi_1)$	τ	$Var(\tau)$	λ	$Var(\lambda)$
E.M.V.	0,7623	0,0019	2,3151	---	-0,0082	---
Não Informativa	0,6971	0,0022	2,0745	0,0021	-0,0076	---

Foi feito o ajuste usando somente uma priori não informativa para um modelo auto-regressivo de primeira ordem, porém é proposto trabalhar com outras distribuições a priori como a normal-gama e a t-gama.

4.4.2 ESTIMADORES CLÁSSICOS PARA A SÉRIE TRANSFORMADA :

Também foram calculados para este exemplo os estimadores clássicos para os parâmetros do modelo auto-regressivo, estes processos foram desenvolvidos na seção 3.4. O estimador do parâmetro de transformação λ é calculado via a função de verossimilhança "profile", a mesma idéia foi aplicada no cálculo da marginal de λ na abordagem bayesiana. Começamos fazendo a escolha do modelo segundo os critérios de seleção apresentados.

CRITÉRIOS DE SELEÇÃO DE MODELOS :

Tabela 4 : Critérios de seleção de modelos - Serie transformada.

CRITÉRIOS	AR(1)	AR(2)	AR(3)
AIC	-0,8334	-0,8630	-0,8630
BIC	-0,8198	-0,8425	-0,8356
Fator de Bayes	0,0314	0,7256	0,2713

ESTIMADOR PARA LAMBDA :

O modelo escolhido para a série transformada segundo estes critérios de seleção foi o modelo AR(2). Para este modelo foi estimado o parâmetro de transformação λ , a figura 18 mostra o valor de λ que maximiza a função de verossimilhança "profile".

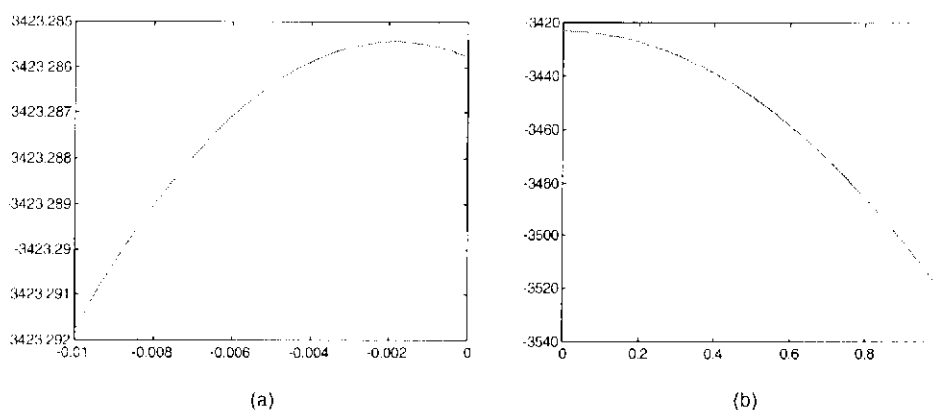


Figura 18 : Função de verossimilhança "profile" avaliada para uma grade de λ .
(a) para $\lambda \in [-0,01,0)$ e (b) para $\lambda \in (0,1]$

COMPARAÇÃO DOS RESULTADOS :

Tabela 5 : Comparação das estimativas clássicas para os parâmetros dos modelos auto-regressivos.
Vazões médias mensais no Reservatório de Furnas - serie transformada.

MODELO	ϕ_1	$Var(\phi_1)$	ϕ_2	$Var(\phi_2)$	ϕ_3	$Var(\phi_3)$	τ	λ
AR(1)	0,7623	0,0019	---	---	---	---	2,3151	-0,0082
AR(2)	0,6451	0,0018	0,1543	0,0019	---	---	2,3922	-0,0019
AR(3)	0,6342	0,0021	0,1117	0,0020	0,0667	0,0023	2,3994	-0,0100

PREVISÕES PARA A SÉRIE ORIGINAL :

As previsões são encontradas mediante a equação (3.35) da seção 3.4.2, o cálculo da integral é feito a través da integral de Gauss.

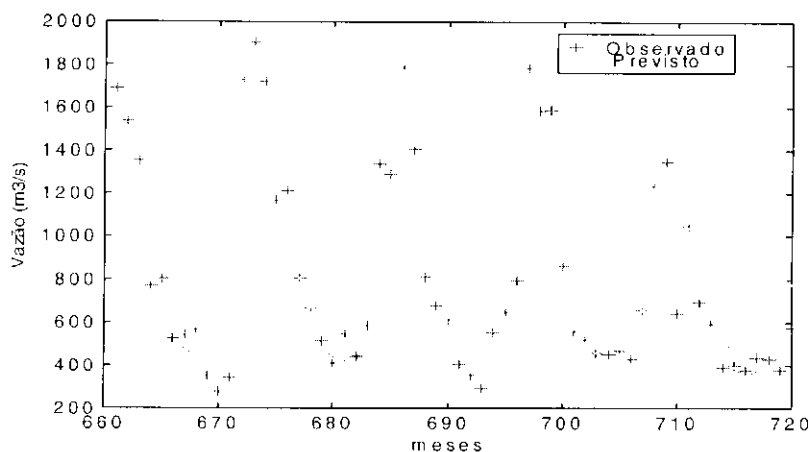


Figura 19 : Previsões para a série original. valores observados vs. valores previstos.

4.4.3 CONCLUSÕES :

Segundo os critérios de seleção, é escolhido o processo auto-regressivo de segunda ordem para a série transformada, as previsões reforçam esta escolha. Podemos concluir então, que a transformação de potência consegue estabilizar a serie, assim podem ser usados os modelos ARIMA em séries que apresentam este problema, particularmente nas séries de vazões e precipitações pluviais.

Tabela 6 : Critérios de seleção de modelos.

CRITÉRIOS	AR(2) normalizada	AR(2) transformada
AIC	-1,0467	-0.8630
BIC	-1,0255	-0.8425
Fator de Bayes	0,5581	0,7246

Nesta seção fizemos uso de um vasto conjunto de ferramentas numéricas que facilitam o cálculo das técnicas desenvolvidas nos capítulos anteriores, vemos a importância dos algoritmos de simulação MCMC e dos algoritmos numéricos no cálculo das integrais, assim a abordagem clássica e bayesiana podem se complementar em muitos casos como este.

CAPITULO 5

CONCLUSÕES E PROPOSTAS FUTURAS

5.1 INTRODUÇÃO :

Neste capítulo discutiremos as conclusões principais deste trabalho e algumas propostas para serem desenvolvidas em estudos posteriores. Em primeiro lugar propomos trabalhar com outras distribuições a priori para a série transformada e em segundo lugar fazemos uma proposta para trabalhar com modelos auto-regressivos considerando erros não gaussianos.

5.2 CONCLUSÕES :

Este trabalho trata da utilização, principalmente, de técnicas bayesianas em séries temporais ajustadas por modelos auto-regressivos de ordem p , $AR(p)$. De particular interesse foram os estudos sobre a estimação do parâmetro de transformação λ em modelos $AR(1)$ para estabilizar a série.

Apresentamos uma seqüência de procedimentos para análise de modelos auto-regressivos de ordem p , a través de técnicas de inferência clássica e bayesiana, além de simulação de Monte Carlo em Cadeias de Markov. Os procedimentos envolvidos são :

- Determinação da função de verossimilhança dos dados, especificação da densidade a priori dos parâmetros, e cálculo da densidade a posteriori via teorema de Bayes.
- Determinação da função de verossimilhança "profile" no caso dos dados transformados, cálculo dos estimadores clássicos e previsões para a série.
- Análise da série buscando observar propriedades particulares da mesma como tendências, comportamento sazonal, etc. e posterior determinação da FAC e FACP.
- Construção e análise da FAC e FACP da serie transformada através de técnicas clássicas e bayesianas, e através do Polinômio de Hermite.
- Seleção da ordem do modelo via métodos clássicos como AIC e BIC, bem como através de métodos bayesianos via fator de Bayes.
- Estimação dos parâmetros via métodos clássicos como o de máxima verossimilhança, e através de técnicas bayesianas via simulação Monte Carlo.

A metodologia utilizada na abordagem bayesiana para estimar o valor do parâmetro de transformação λ é equivalente ao uso da função de verossimilhança “profile” (Box e Cox, 1996) utilizada na abordagem clássica, porém os estimadores de ϕ e τ , substituídos na função de verossimilhança são os estimadores bayesianos. Podemos observar, ao longo do trabalho, a importância, na estimação dos parâmetros do modelo, dos algoritmos de simulação Monte Carlo em Cadeias de Markov (MCMC - Gibbs-Sampler e Metropolis-Hasting), e dos algoritmos numéricos, como Newton-Raphson e a Integral de Gauss.

No caso particular da série de vazões, obtivemos estimativas dos parâmetros para a série normalizada e a para a série transformada, porém, a metodologia desenvolvida para a série transformada na abordagem bayesiana, se limitou aos processos auto-regressivos de primeira ordem utilizando uma priori não informativa, consideramos a possibilidade que em trabalhos posteriores sejam desenvolvidos resultados para outras priors.

Nas seguintes tabelas temos um resumo dos resultados para ambas as séries, normalizada e transformada, observando que as estimativas clássicas e bayesianas são bastante próximas.

Tabela 7 : Comparação das estimativas para os parâmetros do modelo AR(2).
Vazões médias mensais no Reservatório de Furnas - série normalizada.

MÉTODO	ϕ_1	$Var(\phi_1)$	ϕ_2	$Var(\phi_2)$	τ	$Var(\tau)$
E.M.V.	0,6478	0,0027	0,1623	0,0027	2,0285	---
Não Informativa	0,6484	0,0029	0,1621	0,0029	2,0178	0,0222
Normal Gama	0,6435	0,0026	0,1644	0,0028	1,9987	0,0217
T Gama	0,6480	0,0025	0,1619	0,0027	1,9488	0,0317

Tabela 8 : Comparação das estimativas para os parâmetros do modelo AR(1).
Vazões médias mensais no Reservatório de Furnas - série transformada.

MÉTODO	ϕ_1	$Var(\phi_1)$	τ	$Var(\tau)$	λ	$Var(\lambda)$
E.M.V.	0,7623	0,0019	2,3151	---	-0,0082	---
Não Informativa	0,6971	0,0022	2,0745	0,0021	-0,0076	---

Tabela 9 : Comparação das estimativas clássicas para os parâmetros dos modelos auto-regressivos.
Vazões médias mensais no Reservatório de Furnas - série transformada.

MODELO	ϕ_1	$Var(\phi_1)$	ϕ_2	$Var(\phi_2)$	ϕ_3	$Var(\phi_3)$	τ	λ
AR(1)	0,7623	0,0019	---	---	---	---	2,3151	-0,0082
AR(2)	0,6451	0,0018	0,1543	0,0019	---	---	2,3922	-0,0019
AR(3)	0,6342	0,0021	0,1117	0,0020	0,0667	0,0023	2,3994	-0,0100

Na seguinte tabela fazemos uma comparação dos critérios de seleção utilizados para a escolha dos modelos, tendo que a série transformada apresenta um melhor desempenho em comparação com a série normalizada, assim, podem ser usados os modelos ARIMA em séries que apresentam este problema, particularmente nas séries de vazões e precipitações pluviais.

Tabela 10 : Critérios de seleção de modelos.

CRITÉRIOS	AR(2) normalizada	AR(2) transformada
AIC	-1,0467	-0.8630
BIC	-1,0255	-0.8425
Fator de Bayes	0,5581	0,7246

5.3 PROPOSTAS :

Em primeiro lugar observamos que para fins de comparação devemos usar outras distribuições a priori para a série transformada, como a normal-gama e a t-gama, e comparar estes resultados com os modelos análogos na série original. Podemos generalizar os resultados para processos auto-regressivos de ordem p e desenvolver técnicas de previsão segundo uma abordagem bayesiana para a série transformada.

Em segundo lugar sugerimos trabalhar com modelos auto-regressivos com erros não gaussianos, dado que a estruturas dos erros na realidade não é normal. Propomos neste ponto trabalhar com erros com distribuição condicional Exponencial, Gama e log-Normal.

ANEXO A :

DISTRIBUIÇÃO A PRIORI DE JEFFREYS

Uma aproximação que é muitas vezes usadas para definir distribuições a priori não informativas foi introduzida por Jeffreys. A continuação é dada uma versão multiparamétrica desta regra.

Seja a distribuição de \mathbf{x} dependendo de m parâmetros p , sobre certas condições de regularidade e para amostras suficientemente grandes, a função de verossimilhança de p é aproximada por uma distribuição normal. A log verossimilhança é aproximadamente quadrática,

$$l(p/x) = \log l(p/x) \approx L\left(\hat{p}/x\right) - \frac{n}{2} \left(p - \hat{p}\right)' D_{\hat{p}} \left(p - \hat{p}\right), \quad (\text{A.1})$$

onde \hat{p} é o vetor de estimadores de máxima verossimilhança de p e $-nD_{\hat{p}}$ é a matriz $m \times m$

de segundas derivadas avaliadas em \hat{p} , isto é :

$$D_{\hat{p}} = \left\{ -\frac{1}{n} \frac{\partial^2 L}{\partial p_i \partial p_j} \right\}_{\hat{p}}, \quad i, j = 1, \dots, m$$

Em geral, $D_{\hat{p}}$ depende de \mathbf{x} . Para n grande este é aproximadamente fechado para :

$$D_{\hat{p}} \approx \frac{1}{n} f_n(\hat{p}), \quad (\text{A.2})$$

a qual é função só de \hat{p} . Especificamente $f_n(\hat{p})$ é a função,

$$f_n(\hat{p}) = E \left\{ -\frac{\partial^2 L}{\partial p_i \partial p_j} \right\}, \quad (\text{A.3})$$

onde a esperança é tomada com respeito à distribuição dos dados $L(x/p)$. Em outras palavras $f_n(\hat{p})$ é a matriz de informação associada com a amostra \mathbf{x} .

Agora, pensemos em uma transformação para \hat{p} seja $\hat{\phi}(\hat{p})$, poderíamos calcular a função $f_n(\hat{\phi})$ sobre uma verossimilhança aproximadamente transladada. Isto em geral não é

possível, então tentaremos uma transformação ϕ que garanta que o conteúdo da região de verossimilhança aproximada para $\hat{\phi}$,

$$\left(\phi - \hat{\phi}\right) f_n(\hat{\phi}) \left(\phi - \hat{\phi}\right) < const., \quad (A.4)$$

fique constante para diferentes $\hat{\phi}$. Desde que a raiz quadrada do determinante, $|f_n(\hat{\phi})|^{1/2}$ mede o volume da região de verossimilhança, o requerimento acima é equivalente a indagar por uma transformação para a qual $|f_n(\hat{\phi})|$ seja independente de $\hat{\phi}$. Para encontrar essa transformação devemos ter,

$$f_n(\phi) = A f_n(p) A', \quad (A.5)$$

onde A é uma matriz $m \times m$ de derivadas parciais,

$$A = \begin{bmatrix} \frac{\partial(p_1, \dots, p_m)}{\partial(\phi_1, \dots, \phi_m)} \end{bmatrix}.$$

Assim,

$$|f_n(\phi)| = |A|^2 |f_n(p)|, \quad (A.6)$$

de aqui o requerimento acima será satisfeito se,

$$|A| = \frac{\partial(p_1, \dots, p_m)}{\partial(\phi_1, \dots, \phi_m)} \propto |f_n(p)|^{-1/2}, \quad (A.7)$$

e a priori não informativa aproximada será localmente uniforme em ϕ . A correspondente não informação em p é então,

$$\Pi(p) = \Pi(\phi) \left| \frac{\partial(\phi_1, \dots, \phi_m)}{\partial(p_1, \dots, p_m)} \right|;$$

que é,

$$\Pi(p) \propto |f_n(p)|^{1/2} \quad (A.8)$$

logo temos a seguinte regra :

Regra de Jeffreys para modelos multiparamétricos : A distribuição a priori para um conjunto de parâmetros é proporcional à raiz quadrada do determinante da matriz de informação. Esta regra deve ser aplicada com precaução, especialmente quando os parâmetros de locação e escala acontecem simultaneamente.

ANEXO B :

PROPRIEDADES DOS POLINÔMIOS HERMITE

O Sistema de Polinômios Hermite $H_n(x)$ é definido em termos da distribuição Normal standar como :

$$H_n(x) = \exp(x^2/2) \left(\frac{-d}{dx} \right)^n \exp(-x^2/2) \text{ ou } H_n(x) = (-1)^n \Phi^{(n)}(x) / \Phi(x),$$

onde Φ é uma p.d.f. Normal Standar. Explicitamente, podemos escrever :

$$H_n(x) = n! \sum_{m=0}^{\lfloor n/2 \rfloor} (-1)^m \{2^m m! (n-2m)!\}^{-1} x^{n-2m},$$

Onde $[N]$ é o maior inteiro menor ou igual a N . Deste modo, temos :

$$H_0(x) = 1, H_1(x) = x, H_2(x) = x^2 - 1, H_3(x) = x^3 - 3x, H_4(x) = x^4 - 6x^2 + 3,$$

$$H_5(x) = x^5 - 10x^3 + 15x$$

e assim por diante.

Definimos operadores de esperança E_0 e E como:

$$E_0\{\psi(x)\} = (2\pi)^{-1/2} \int_{-\infty}^{\infty} \psi(x) \exp(-x^2/2) dx,$$

$$E\{\psi(x)\} = (2\pi\sigma^2)^{-1/2} \int_{-\infty}^{\infty} \psi(x) \exp\left[-(x-\mu)^2 / 2\sigma^2\right] dx.$$

O Polinômio Hermite constitui um sistema ortogonal com respeito à p.d.f. Normal standar, assim :

$$E_0\{H_n(x)H_k(x)\} = \begin{cases} 0, n \neq k, \\ n!, n = k, \end{cases} \quad (\text{B.1})$$

desde que $H_0(x) = 1$, temos que:

$$E_0\{H_n(x)\} = 0, n > 0 \quad (\text{B.2})$$

Os Polinômios obedecem a uma fórmula recursiva $H_{n+1}(x) - xH_n(x) + nH_{n-1}(x) = 0$, e tem uma função geradora dada por:

$$\exp(tx - t^2/2) = \sum_{n=0}^{\infty} H_n(x) t^n / n! \quad (\text{B.3})$$

Um resultado particularmente importante é que a p.d.f de uma Normal bivariada com a mesma variância pode ser escrita como:

$$\begin{aligned} f(x_1, x_2) &= \frac{(1-\rho^2)^{-1/2}}{2\pi\sigma^2} \exp\left\{-\frac{1}{2} \frac{x_1^2 + x_2^2 - 2\rho x_1 x_2}{\sigma^2(1-\rho^2)}\right\} \\ &= \frac{1}{2\pi\sigma^2} \exp\left\{-\frac{1}{2\sigma^2}(x_1^2 + x_2^2)\right\} \sum_{n=0}^{\infty} \frac{\rho^n H_n(x_1/\sigma) H_n(x_2/\sigma)}{n!}, \end{aligned}$$

a fórmula dada por Mehler. Então temos que se X e Y tem distribuição conjunta Normal bivariada com médias zero, variâncias unitárias e coeficiente de correlação ρ ,

$$E\{H_n(X)/Y = y\} = \rho^n H_n(y)$$

e

$$E\{H_n(X)H_k(Y)\} = \begin{cases} 0, n \neq k, \\ \rho^n n!, n = k. \end{cases} \quad (\text{B.4})$$

Um resultado útil é a fórmula de adição :

$$H_n(Ax + By) = \sum_{k=0}^n C_k A^k B^{n-k} H_k(x) H_{n-k}(y), \quad \text{de } A^2 + B^2 = 1 \quad (\text{B.5})$$

Um método geral para se obter a expansão do polinômio Hermite para uma função $T(x)$ é notando que se:

$$\sum_{j=0}^{\infty} \alpha_j H_j(x) = T(x)$$

onde, $\alpha_n = E_0\{(d/dx)^n T(x)\}/n!$. Se $T(x)$ e alguma função onde $E_0\{T(x)\} < \infty$, então sempre existe uma expansão assintótica da forma $T(x) = \lim S_N(x)$ quando $N \rightarrow \infty$, onde

$$S_N(x) = \sum_{j=0}^N \alpha_j H_j(x)$$

tal que o $\lim E_0 \left[\{T(x) - S_N(x)\}^2 \right] = 0$ quando $N \rightarrow \infty$. Algumas expansões particulares são:

$$|x|^\lambda = \alpha_0 + \alpha_0 \sum_{m=1}^{\infty} \left\{ \prod_{v=0}^{m-1} (\lambda - 2v) \right\} H_{2m}(x) / (2m)!, \lambda + 1 > 0,$$

onde:

$$\alpha_0 = 2^{\lambda/2} \Gamma \left\{ \frac{1}{2}(\lambda + 1) \right\} / \Gamma \left(\frac{1}{2} \right),$$

e:

$$|x|^{\lambda+1} \operatorname{sign} x = (\lambda + 1) \alpha_0 \left\{ H_1(x) + \sum_{m=1}^{\infty} \prod_{v=0}^{m-1} (\lambda - 2v) H_{2m+1}(x) / (2m + 1)! \right\}, \lambda + 2 > 0.$$

A expansão de $\exp(tx)$ é obtida da função geradora em (B.3).

Para mais detalhes sobre as propriedades dos polinômios Hermite e indicações das provas de alguns destes resultados podem ser encontrados em Erdélyi et al. (1953, Vol. 2 Capítulo 10), embora a notação seja diferente.

REFERÊNCIAS BIBLIOGRÁFICAS

- Akaike, H. (1974). New look at statistical model identification, *IEEE Transactions Automatic Control*, AC-19, 716-723.
- Ansley, C. F. and Newbold (1980), "Finite Sample Properties of Estimators for Autoregressive Moving Average Models," *Journal of Econometrics* (June 1980), pp. 159-183.
- Barreto, G. de A. & Andrade, M. G. (2000). Bayesian Inference and Markov Chain Monte Carlo Methods Applied to Streamflow Forecasting. 6^a International Conference on Probabilistic Methods Applied to Power System Funchal, Madeira - Portugal, pp. 1-8.
- Barrett, J. F. and Lampard, D. G. (1955). An expansion for some second order probability distributions and its application to noise problems. *I.R.E. Trans. PGIT, IT-1*, 10-15.
- Bates, J.M. & Granger, C.W.J. (1969). The combination of forecasting. *Oper. Res.* Q20, 451-468.
- Box, G. E. P. and Jenkins, G. M. (1970), *Time Series Analysis, Forecasting and Control*, San Francisco: Holden Day.
- Box, G.E.P. & Cox, D.R. (1964), An analysis of transformations, *Journal of Royal Statistical Society B*, 26, 211-243.
- Box, G. E. P. & Jenkins, G. M. (1976) : *Time series analysis. Forecasting and control*. San Francisco. Prentice Hall.
- Brillinger, D. R. (1980). Analysis of variance and problems under time series. *Handbook of Statistics*, Vol. I, pp. 237-277.
- Broemeling, L. D. & Land, M. (1984). On forecasting with univariate autoregressive processes. *Comm. Statist.-Theor. Meth.*, vol. 13, no. 11, pp. 1305-1320.
- Broemeling, L. D. & Shaarawy, S. (1988). Time Series : A Bayesian Analysis in the time domain. In : Spall, J., *Bayesian Analysis of Time Series and Dynamic Models*, pp 1-21. New York, Marcel Dekker.
- Broemeling, L. D. & Cook, P. (1993). Bayesian estimation of the mean of an autoregressive process. In: *Journal of Applied Statistics*, Vol. 20, N^o. 1, pp. 25-39.
- Brooks, S. P. & Roberts, G. O. (1999). Convergence assessment techniques for Markov chain Monte Carlo. *Statistics and Computing*, vol. 8, pp. 319-335.
- Casella, G. & George, E. J. (1992). Explaining the Gibbs sampler. *The American Statistician*, vol. 46, no. 3, pp. 167-174.
- Chatfield, C. & Prothero, D.L. (1973), Box-Jenkins seasonal forecasting: problems in a case study, *Journal of Royal Statistical Society A*, 136, 295-336.
- Chib, S. & Greenberg, E. (1995). Understanding the Metropolis-Hastings algorithm. *The American Statistician*, vol. 49, no. 4, pp. 327-335.
- Donelson, J. and Maltz, F. (1972). A comparison of linear versus non-linear prediction for polynomial functions of the Ornstein-Uhlenbeck process. *J. Appl.*, 9, 725-744.
- Erdélyi, A. et al. (1953). *Higher Transcendental Functions*. Bateman Manuscript Project. New York: McGraw-Hill.

- Gamerman, D. (1997): Markov Chain Monte Carlo. Stochastic Simulation for Bayesian Inference. CRC Press.
- Gelman, A. E. & Rubin, D. (1992). Inference from iterative simulation using multiple sequences, *Statistical Science*, vol. 7, pp. 457-511.
- Gelfand, A. E. & Smith, A.F.M. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association*, vol. 85, 398-409.
- Granger, C.W.J. & Newbold, P. (1976), Forecasting transformed series, *Journal of Royal Statistical Society B*, 38, 189-203.
- Harrison, P. J. & Stevens, C. F. (1976) Bayesian forecasting, *Journal of the Royal Statistical Society, Series B*, 38, pp. 205-247.
- Hopwood, W ; Mckeown, J. (1984): Time Series Forecasting Models Involving Power Transformation. In : *Journal of Forecasting*, Vol. 3, pp. 57-61.
- Lyle D. Broemeling & Peyton Cook (1993). Bayesian estimation of the mean of a autoregressive process, *Journal of Applied Statistics*, Vol. 20, Nro 1, 25-39.
- Nelson, H. L. & Granger, C. W. J. (1979), Experience with using the Box-Cox transformation when forecasting economic time series, *Journal of Econometrics*, 10, 190-204.
- Oliveira, J. R. T. (1998). Uso do amostrador de Gibbs e Metropolis-Hastings em análise bayesiana de modelos AR(p). *Dissertação de Mestrado*, ICMC, Universidade de São Paulo, Campus São Carlos.
- Pericchi, L. R. & Smith, A. F. M. (1992). Exact and approximate posterior moments for a normal location parameter, *Journal of Royal Statistical Society*, vol. 54, no. 3, pp. 793-804.
- Raftery, A. E. (1996). Hypothesis testing and model selection. In: *Markov Chain Monte Carlo in Practice*, W. R. Gilks, E. Richardson e D. J. Spiegelhalter (Ed.), London: Chapman & Hall.
- Schervish, M. J. ; Tsay, R.S. (1988) Bayesian Modeling and Forecasting in Autoregressive Models, *Bayesian Analysis of Time Series and Dynamic Models*, Ed by James L. Spall, CAP 2, pp 23-52.
- Sriram, T. N. (1987). Sequential estimation of the mean of a first-order stationary autoregressive process. *Annals of Statistics*, 15(3), pp. 1079-1090.
- Schwarz, G.(1978), Estimating the dimension of a model, the *Annals of Statistics*, Vol 6, Nro 2, pp. 461-464.
- Tierney, L. (1994). Markov chains for exploring posterior distributions. *Annals of Statistics*, vol. 22, pp. 1701-1762.
- West, M. & Harrison, J. (1989) *Bayesian Forecasting and Dynamic Models* (New York, Springer-Verlag)
- Whittle, P. (1951). *Hypothesis Testing in Time Series Analysis*. Uppsala, Almqvist & Wiksell.
- Zellner, A. (1971). *An Introduction to Bayesian Inference in Econometrics*. New York, Wiley.