

1

Introdução

Séries temporais para dados de contagem são registros da frequência relativa de ocorrência de determinados eventos em sucessivos intervalos de tempo, geralmente equiespaçados. Como exemplos temos: o número mensal de sinistros de fogo associados a uma seguradora, o número diário de homicídios por arma de fogo em determinada região ou cidade, o número diário de internações hospitalares em uma cidade em consequência de problemas cardio-respiratórios etc.

Como pode-se observar, dados de contagem são necessariamente números inteiros não negativos, o que elimina a possibilidade do processo gerador de dados ser Gaussiano, e assim torna inadequada, em princípio, a utilização de modelos ARIMA (Box-Jenkins). Os modelos apropriados para dados de contagem devem utilizar distribuições de probabilidade cujo domínio seja formado apenas pelos números inteiros não-negativos como é o caso das distribuições Poisson, hipergeométrica, binomial e binomial negativa.

Abaixo, para fins de comparação, apresentamos o gráfico de 3 séries temporais e respectivos histogramas, geradas por processos dependentes com as distribuições normal($\mu=1, \sigma^2=1$), Poisson($\lambda=1$), e Binomial Negativa ($p=0.5, k=1$).

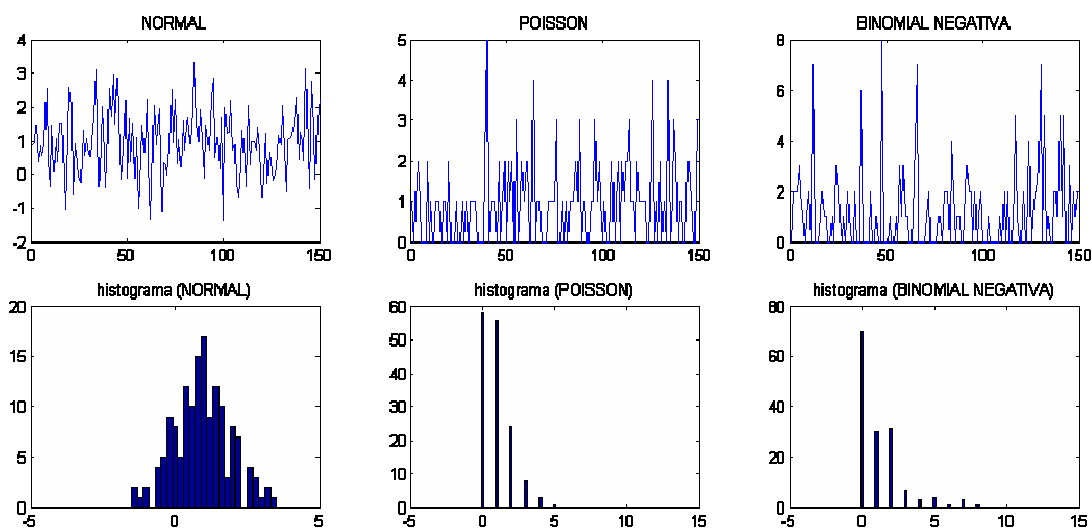


Figura 1: Gráficos comparativos relativos às distribuições Normal, Poisson, e Binomial Negativa.

É possível notar claras diferenças entre os casos apresentados. A inexistência de valores negativos bem como a presença apenas de valores inteiros nos casos Poisson e binomial negativa são os aspectos mais marcantes a serem observados. Na tabela abaixo apresentamos algumas estatísticas descritivas das séries acima simuladas. Como era de se esperar, a série de distribuição Binomial Negativa apresenta superdispersão em relação às outras.

Tabela 1: Estatísticas descritivas das séries geradas.

	<i>média</i>	<i>variância</i>	<i>desvio padrão</i>	<i>curtose</i>	<i>simetria</i>
normal	1.0416	1.0027	1.0013	3.7436	-0.4956
Poisson	0.9867	1.0736	1.0362	3.0211	0.8619
binomial negativa	1.0133	1.9596	1.3998	5.7903	1.6996

O presente trabalho tem como objetivo principal investigar por meio de simulação Monte Carlo algumas propriedades estatísticas das especificações Poisson e Binomial Negativa dos modelos GARMA (*Generalized Autoregressive Moving Average Models*). Esta classe de modelo foi proposta por Benjamin, Rigby e Stasinopoulos (2003) [1] em uma extensão do trabalho originalmente apresentado por Li (1994) [13], Zeger e Qaqish (1988) [23] e Zeger e Qaqish desenvolveram modelos autoregressivos para a família exponencial e Li alterou este modelo para que incorporasse também termos médias móveis.

As propriedades estatísticas dos modelos GARMA que serão foco da nossa investigação computacional serão as condições de estacionariedade dos modelos GARMA e metodologias para a identificação da ordem (p,q) dos polinômios AR e MA que definem o modelo. Na parte de estacionariedade foi realizada uma varredura no espaço paramétrico dos modelos GARMA (p,q) , $p= 0,1,2$ e $q = 0,1,2$, e através de um teste chi-quadrado foi investigada a estacionariedade da densidade marginal implicada por estes modelos. Na identificação foi investigada a adequação de alguns critérios de informação - Akaike, Schwarz e Hanna-Quinn - utilizados na identificação da ordem correta de modelos de séries temporais. Também foi sistematizado as formas da FAC e FACP obtidas para diversos processos GARMA (p,q) , observando a existência, tal como em processos ARMA gaussianos, de padrões definidos para processos GARMA puros (AR ou MA). Por fim, serão realizadas aplicações do modelo a séries reais, comparando os resultados com os ajustes produzidos por um modelo ARMA gaussiano.

No capítulo 2 são apresentadas as principais propriedades estatísticas dos modelos GARMA para dados de contagem com distribuições condicionais Poisson e Negativa Binomial. O capítulo 3 é dedicado à investigação, por simulação Monte Carlo, das propriedades de estacionariedade e da adequação dos critérios de identificação para os modelos GARMA. No capítulo 4 apresentamos aplicações dos modelos a duas séries reais, o número mensal de casos de poliomielite nos EUA e o número diário de casos de enfarto registrados no estado do Rio de Janeiro. Finalmente no capítulo 5 apresentamos as conclusões da investigação e sugestões para futuros estudos.