

# Modelar a sobre-dispersão com processos de médias móveis de valor inteiro

**Isabel Silva**

*Faculdade de Engenharia, Universidade do Porto & CIDMA, ims@fe.up.pt*

**Maria Eduarda Silva**

*Faculdade de Economia, Universidade do Porto & CIDMA, mesilva@fep.up.pt*

**Cristina Torres**

*IPP-ISCAP, cmptorres@gmail.com*

**Palavras-chave:** Distribuição binomial negativa, Modelos BINMA, Operação *thinning*, Séries temporais de contagem, Sobre-dispersão.

**Resumo:** Séries temporais de contagens, caracterizadas por apresentar valores baixos, assimetrias, valores nulos e sobre-dispersão, podem ser encontradas nas mais variadas áreas de aplicação. Uma das abordagens propostas para analisar estes tipos de dados consiste na utilização de modelos baseados numa operação aleatória chamada *thinning*, que conjugada com inovações discretas, permite preservar a natureza discreta das contagens. Desta forma surge a família de modelos auto-regressivos e de médias móveis de valor inteiro, INARMA, que tem sido amplamente estudada na literatura, especialmente no que toca a modelos auto-regressivos (ver Scottó *et al.* [3]). Por outro lado, os modelos de média móvel de valor inteiro têm atraído menos atenção, devido sobretudo à impossibilidade de escrever de forma explícita a verosimilhança (Torres [4]).

Neste trabalho, propõe-se e caracteriza-se um novo modelo de médias móveis de valor inteiro de primeira ordem, baseado na operação *thinning* binomial negativa, definida por Ristić *et al.* [2], e denotado por NB-INMA(1). Quando as inovações seguem uma distribuição binomial negativa (NB), este modelo apresenta distribuição marginal NB, pelo que pode ser utilizado em situações em que os dados apresentem sobre-dispersão. Adicionalmente, este modelo é estendido para o contexto bivariado, obtendo-se o modelo denotado por BINMA\_NB-II(1,1), baseado na mesma operação *thinning* e em que o processo das inovações segue uma distribuição NB bivariada do tipo II, proposta por Edwards e Gurland [1].

O método dos momentos generalizado (GMM) é usado na estimação dos parâmetros desconhecidos dos modelos propostos e os resultados de um estudo de simulação que pretende investigar o desempenho do método mostram que, em geral, as estimativas são consistentes e simétricas.

Finalmente, os modelos propostos são ajustados a conjuntos de dados reais e a qualidade do ajuste é avaliada.

**Agradecimentos:** Este trabalho foi parcialmente financiado pela Fundação Portuguesa para a Ciência e Tecnologia (FCT), através do CIDMA - Centro de Investigação e Desenvolvimento em Matemática e Aplicações, no âmbito do projeto UID / MAT / 04106/2013.

## **Bibliografia**

- [1] Edwards, C.B., Gurland, J. A class of distributions applicable to accidents. *Journal of the American Statistical Association*, 56, 503–517, 1961.
- [2] Ristić, M.M., Bakouch, H.S., Nastić, A.S. A new geometric first-order integer-valued autoregressive (NGINAR (1)) process. *Journal of Statistical Planning and Inference*, 139, 2218–2226, 2009.
- [3] Scotto, M.G., Weiß, C.H., Gouveia, S. Thinning-based models in the analysis of integer-valued time series: a review, *Statistical Modelling*, 15, 590–618, 2015.
- [4] Torres, C. *Contributions to the analysis of multivariate count data*. PhD Thesis, Universidade do Porto, Portugal, 2016.