Statistical Inference for Integer-Valued Autoregressive (INAR(p)) Processes

Tamás Szabó

Let $(\varepsilon_k)_{k\in\mathbb{N}}$ be an independent and identically distributed (i.i.d.) sequence of non-negative integer-valued random variables, and let $\alpha_1, \dots, \alpha_p \in [0,1]$. An INAR($p$) time series model with coefficients $\alpha_1, \dots, \alpha_p$ and innovations $(\varepsilon_k)_{k\in\mathbb{N}}$ is a stochastic process $(X_n)_{n\geq -p+1}$ given by

$$X_k = \sum_{j=1}^{X_{k-1}} \xi_{k,1,j} + \cdots + \sum_{j=1}^{X_{k-p}} \xi_{k,p,j} + \varepsilon_k \,, \qquad k \in \mathbb{N} \,,$$

where for all $k \in \mathbb{N}$ and $i \in \{1, \dots, p\}$, $(\xi_{k,i,j})_{j\in\mathbb{N}}$ is a sequence of i.i.d. Bernoulli random variables with mean α_i such that these sequences are mutually independent and independent of the sequence $(\varepsilon_k)_{k\in\mathbb{N}}$, and $X_0, X_{-1}, \dots, X_{-p+1}$ are non-negative integer-valued random variables independent of the sequences $(\xi_{k,i,j})_{j\in\mathbb{N}}$, $k \in \mathbb{N}$, $i \in \{1, \dots, p\}$, and $(\varepsilon_k)_{k\in\mathbb{N}}$.

The INAR(p) model can be written in another way using the binomial thinning operator $\alpha\circ$ (due to Steutel and van Harn) which we recall now. Let X be a non-negative integer-valued random variable. Let $(\xi_j)_{j\in\mathbb{N}}$ be a sequence of i.i.d. Bernoulli random variables with mean $\alpha \in [0,1]$. We assume that the sequence $(\xi_j)_{j\in\mathbb{N}}$ is independent of X. The non-negative integer-valued random variable $\alpha \circ X$ is defined by

$$\alpha \circ X := \begin{cases} \sum_{j=1}^{X} \xi_j \,, & \text{if } X > 0, \\ 0 \,, & \text{if } X = 0. \end{cases}$$

The INAR(p) model takes the form

$$X_k = \alpha_1 \circ X_{k-1} + \cdots + \alpha_p \circ X_{k-p} + \varepsilon_k \,, \qquad k \in \mathbb{N} \,.$$

Change in the coefficients or in the mean or variance of the innovations of an INAR(p) process over time is a sign of disturbance that is important to detect. The methods presented in this talk can test for change in any one of these quantities separately, or in any collection of them. They are available in forms that make one-sided tests possible, furthermore, they can be used to test for a temporary change. The test statistics are based on an analogue of the efficient score vector. The large sample properties of the change-point estimator are also explored.