Time Series Models for Deer Populations using Time Dependent Reproduction Rates

Daniela Gandini
University of Kaiserslautern
e-mail: gandini@mathematik.uni-kl.de
Abstract

June 3, 1998

General remarks

Linear models like autoregression or ARMA-processes have primarily been designed for real-valued time series. They are, however, no adequate tools for integer-valued processes as e. g. the modelling of populations require. A flexible parametric class of parametric models for time series with values in $\mathbf{N}_0$ are the integer-valued autoregressive (INAR) processes. In the multivariate case, these models and estimation procedures are useful for fitting time series of vectors of counts. Their main advantage is that their autocorrelation matrices are the same as for the AR(1)-process $Y(t) = A \cdot Y(t-1) + \varepsilon(t)$. Therefore, INAR-processes retain some of the properties of the familiar autoregressions while allowing for the discreteness of the data.

Notations

$$p \circ U = \sum_{j=1}^{U} Y_j$$

where U is a random variable with values in $\mathbf{N}_0$, $0 \leq p \leq 1$ and $Y_1, Y_2, \ldots$ are i.i.d. Bernoulli variables, independent of U, with

$$p = pr(Y_j = 1)$$

In general, for $M \geq 1$ let A be an $M \times M$-matrix with entries a_{ij} satisfying $0 \leq a_{ij} \leq 1$ for $i, j = 1, \ldots, M$. Then, for a random vector X with values in $\mathbf{N}_0^M$, $A \circ X$ is defined as the $\mathbf{N}_0^M$-valued random vector with i-th component

$$(A \circ X)_i = \sum_{j=1}^{M} a_{ij} \circ X_j, \qquad i = 1, \ldots, M$$

where we assume independence of the counting series of all $a_{ij} \circ X_i$, $i, j = 1, \ldots, M$.

Special case

An important aspect of wildlife management is the estimation and forcasting of the actual size and structure of the game population. An effective tool to approximately determine the size of a region's deer population is the use of mortality statistics. Here, annually the age and sex of every member of the population that was hunted or that died from another cause are registered. From these data, the number of animals of a given age and sex, being part of the population in a given year, can be reconstructed.

In this case, the mortality statistics were collected over a span of 21 years in a nearly closed area in the Rhine valley, such that emigration and immigration give rise to only small errors. The calculations will be restricted to the does only, as they are responsible for the reproduction of the population. Using the reconstructed data and assuming stationarity of the population and independent and identical survival of the single animals, it is possible to use the following M-variate INAR(1)-process

$$X(t) = A \circ X(t-1) + \varepsilon(t).$$

Here, $\varepsilon_i(t)$, $i = 1, \ldots, M$ are included to cope for the small error mentioned above, and $M = 3$ is the fixed number of classes considered (separated by age). The first class includes only the newborn deer of age zero to one; in the second class the animals are of the age between one and two years, whereas the third class consideres all the rest, i.e. older than two years. The coefficient matrix then takes the following form:

$$A = \begin{pmatrix} 0 & \gamma & \delta \\ \alpha_1 & 0 & 0 \\ 0 & \alpha_2 & \beta \end{pmatrix}$$

where α_1, α_2, and β are the probabilities of the deer to survive till the following year and γ and δ are the reproduction rates of the classes two and three. Furthermore, we assume that the deer belonging to class one do not yet contribute to the reproduction.

If all coefficients are assumed to be constant, their maximum likelihood estimates can be calculated.

On the other hand, the reproduction rate seems to decrease as the size of the population increases. This important feature would be neglected assuming the coefficients to be constant. The new task now is to find function estimates for γ and δ depending on the density of the population.

To avoid further assumptions, nonparametric estimation procedures for γ and δ using the Nadaraya-Watson-estimator, can be applied. On the basis of these estimations, we can then fit a function and estimate its parameters. The remaining parameters α_1, α_2, and β, which are still constants, are easily approximated by least squares estimates.

References

[1] Franke, J., and Seligmann, Th. (1993): *Conditional maximum-likelihood estimates for INAR(1)-processes and their applcation to modelling epileptic seizure counts.* In: Developments in Time Series Analysis, ed. T. Subba Rao, Chapman & Hall, London.

[2] Franke, J., and Subba Rao, T. (1993): *Multivariate first-order integer-valued autoregressions.* Berichte der Arbeitsgruppe Technomathematik, Bericht Nr.95, Universität Kaiserslautern.

[3] Schmitt, A. (1995): *Modellierung der Populationsentwicklung bei Rehen mit Hilfe ganzzahliger autoregressiver Prozesse.* Diploma thesis, Universität Kaiserslautern.