

Szeregi czasowe dla danych dyskretnych

Szeregi czasowe

- Szeregi czasowe to narzędzie analizy danych uporządkowanych w czasie.
- Standardowe modele zakładają dane ciągłe.
- W praktyce często spotykamy dane dyskretne: liczby osób, zdarzeń, przypadków.

1.1.2 Proces autoregresji (AR)

Proces autoregresji rzędu p , to proces następującej postaci

$$Y_t = f(y_{t-1}, y_{t-2}, \dots, y_{t-p}, Z_t). \quad (1.7)$$

$$Y_t = \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + Z_t = \phi_0 + \sum_{i=1}^p \phi_i Y_{t-i} + Z_t, \quad (1.8)$$

gdzie $\phi_1, \phi_2, \dots, \phi_p$ to parametry modelu, p to wielkość opóźnienia, natomiast $Z \sim WN(0, \sigma^2)$ to biały szum

Model INAR(1)

Proces $\{X_t\}_{t \in \mathbb{Z}}$ nazywamy procesem INAR(1) z parametrem $\alpha \in [0, 1)$ i napływem $\{\epsilon_t\}$, jeśli spełnia równanie:

$$X_t = \alpha \circ X_{t-1} + \epsilon_t,$$

gdzie:

- $\{\epsilon_t\}$ jest ciągiem niezależnych i identycznie rozłożonych zmiennych losowych na $\mathbb{Z}_+$,
- ϵ_t jest niezależne od X_{t-1} dla każdego t .

Operator przerzedzenia

Niech X będzie zmienną losową przyjmującą wartości w zbiorze liczb całkowitych nieujemnych $\mathbb{Z}_+$. Dla ustalonego $\alpha \in [0, 1]$ definiujemy operator przerzedzania $\circ$ w następujący sposób:

$$\alpha \circ X = \sum_{i=1}^X Y_i,$$

gdzie Y_i są niezależnymi zmiennymi losowymi typu Bernoulliego o parametrach $P(Y_i = 1) = \alpha$, $P(Y_i = 0) = 1 - \alpha$.

Interpretacja modelu

$$X_t = \alpha \circ X_{t-1} + \epsilon_t,$$

The INAR(1) model defined by (2.1) simply states that the components of the process at time t , X_t , are (i) the survivors of the elements of the process at time $t - 1$, X_{t-1} , each with probability of survival α and (ii) elements which entered the system in the interval $(t - 1, t]$ as innovation term (ϵ_t) .

As an illustration of how the sample path of the INAR(1) process differs from that of the standard AR(1) process, a simulated sample path for each process was generated. For comparison of the two sample paths, the two simulated processes are taken to have the same mean and the same variance. Figure 1 presents the simulated sample path based on a sample of size 100 observations generated from model (2.1), in which Y_i is taken to be Bernoulli random variables with $\alpha = 0.5$ and ε_i is taken to be Poisson with parameter $\lambda = 1.0$. Figure 2 shows the corresponding sample path of the standard AR(1) model, in which ε_t is taken to be normally distributed with mean λ and variance $\lambda(1 + \alpha)$.

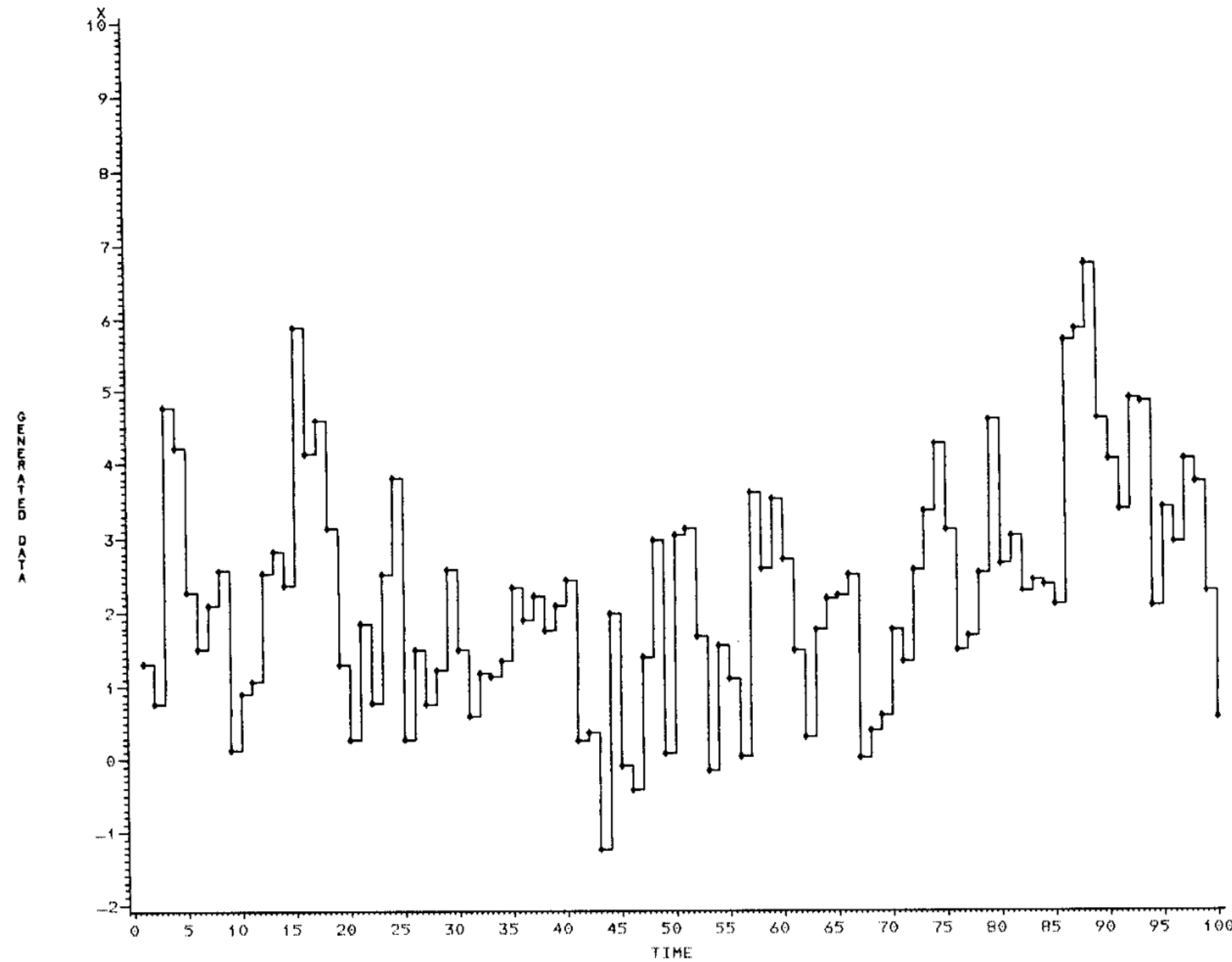


FIGURE 2. Continuous time series generated from normal (2, 2).

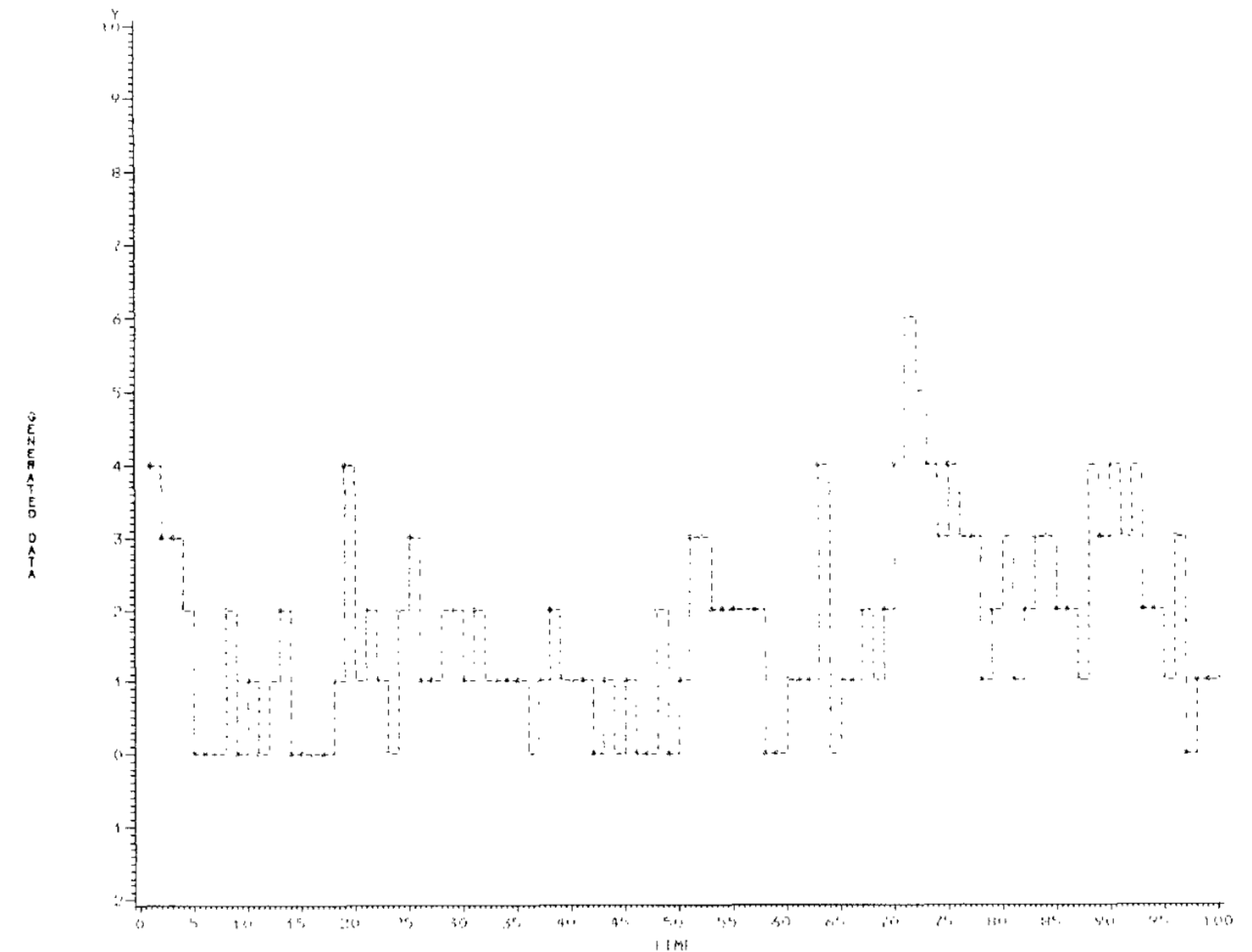


FIGURE 1. Integer valued time series generated from Bernoulli (0.5) and Poisson (1.0).

Model INAR(p)

Definition 2.1 allows the specification of the integer autoregressive process of order p . Suppose that $\alpha_i \in [0, 1)$ for $i = 1, \dots, p$ and let $\{\epsilon_t, t \in \mathbb{N}\}$ be a sequence of independent identically distributed nonnegative integer-valued random variables with $E(\epsilon_t) = \mu$ and $\text{var}(\epsilon_t) = \sigma_\epsilon^2$. Then, the process

$$X_t = \sum_{i=1}^p \alpha_i \circ X_{t-i} + \epsilon_t, \quad (1)$$

is called an integer autoregressive process of order p , denoted INAR(p). The Bernoulli variables used to define the random variable $\alpha_1 \circ X_{t-1}$ are independent of those used to define $\alpha_2 \circ X_{t-2}$, and so forth, so the INAR(p) process has the classical AR(p) correlation structure (Du and Li 1991). A unique stationary and ergodic solution of (1) exists if $\sum_{i=1}^p \alpha_i < 1$, and then $E(X_t) = \mu / (1 - \sum_{i=1}^p \alpha_i)$.