Mikailalsys Journal of Mathematics and Statistics

e-ISSN : 3030-816X
p-ISSN : 3030-8399

Index : Harvard, Boston, Sydney University, Dimensions, Lens, Scilit, Semantic, Google, etc

https://doi.org/10.58578/MJMS.v3i2.5379

A Study on Mixture Poisson Autoregressive (P) Model

Yaska Mutah[1], S. S. Abdulkadir[2], A. A. Akinrefon[3], E. Torsen[4]
[1]Federal Polytechnic, Mubi, Adamawa State, Nigeria
[2,3,4]Modibbo Adama University, Yola, Adamawa State, Nigeria
yaskamutah@gmail.com

Article Info:

Submitted:	Revised:	Accepted:	Published:
Mar 15, 2025	Mar 30, 2025	Apr 11, 2025	Apr 16, 2025

Abstract

This research presents and assesses novel models for time series count data called the Mixture Poisson Autoregressive (MPAR) model addresses difficulties of discreteness, overdispersion, and serial correlation. A completely parametric technique was used, and a marginal distribution for the counts was defined. The parameters of the model were estimated using the Expectation Maximization method, through extensive Monte Carlo simulations, the stability of the estimates of the MPAR was evaluated and the results clearly revealed that the model was stable as the estimated parameters were converging to the values of the true parameters as the sample gets larger. Also, the results from the simulations revealed that the MPAR model outperform other count data models thus, Poisson distribution, Poisson Autoregressive (PAR) and Poisson Exponentially Weighted Moving Average (PEWMA).

Keywords: Mixture; Poisson; Autoregressive

INTRODUCTION

In the study of count time series models, the use of the lagged values of the dependent variable (y_{t-1}) are fundamental in predicting the present outcome (y_t). Time series models with this feature are called autoregressive (AR) models. Just like other statistical methods, the AR method was built on some assumptions, time dependent, Stationarity and uncorrelated error terms. The AR models only allow for unimodal marginal and conditional densities, which cannot capture heterogeneity, multimodality, or skewness evident in data collected overtime (Davide, 2021). Yang *et al* (2019) pointed out that count time series data exhibit three distinctive features; overdispersion, zero-inflation, and temporal correlation, these three characteristics of time series count data pose a serious challenge in developing a model framework. To address these challenge(s), Wong & Li (2000) introduced the technique of mixture model based on the Gaussian function (GF) in relation AR time series. This study observed the literatures above and uses Mixture Poisson Autoregressive model where the Poisson distribution evaluates the rate of occurrence in a particular component, AR was used because of the time dependent Univariate data and mixture model as a result of various risk factors that influence the data collected over time. A Mixture Poisson Autoregressive (MPAR) model is an extension of the traditional Poisson Autoregressive (PAR) model, which combines the features and parameters of both mixture models and Poisson autoregressive processes. It is particularly useful when analyzing time series count data that exhibit temporal dependencies and heterogeneity, this implies the data can be divided into distinct components with different underlying patterns.

Let C_t be a continuous observed random variable (at time t), and let G_t be the information set up to time t. $C_t \sim GF-AR(p)$ if $C_t \mid G_{t-1}$ has a density of the form;

$$g\left(c_t \middle| G_{t-1};\tau\right)=\sum_{i=1}^{k}\pi_i\phi\left(c_t;\alpha_{i0}+\sum_{j=1}^{p}\alpha_{ij}c_{t-j},\sigma_i^2\right) \tag{1}$$

where $\pi_i>0,\sum_{i=1}^{k}\pi_i=1,\lambda_i=\left(\alpha_{i0},\ldots,\alpha_{ip}\right)',\sigma_i^2>0\ \forall\, i=1,2,\ldots,k,\tau=\left(\pi_1,\alpha_1,\sigma_i^2,\ldots,\pi_k,\alpha_k,\sigma_k^2\right)'$ is the parameter (to be estimated) and $\phi\left(c_t;\mu,\sigma^2\right)$ is the density of the normal distribution with mean and variance $\left(\mu,\sigma^2\right)$.

However, since the assumptions was that the variable C_t is a continuous random variable, model expressed in equation (1) is not suitable for count-generated time series data. To model count data, a discrete density function is essential. One among the prominent density functions employed for count data is the Poisson distribution. This is because of equality of mean and variance which is very rare to obtain in practice. Poisson exponentially weighted moving average model (PEWMA) was proposed by Brandt et al, (2000) to capture the dynamics by estimating a time-dependent discounted average of the mean of the event count process, this model is not suitable for cyclical and short memories processes that are mean reverting. The Poisson Autoregressive (PAR) proposed by Brandt and Williams, (2001) tends to solve the challenge of mean reversion, cyclical events, and short memories, however, this model cannot capture events of heterogeneity, multimodal, or skewness in count-based time series. Due to one or more weakness of the Poisson AR model, this research proposes a Mixture Poisson Autoregressive (MPAR) Model on the basis of equation (1) for count-based time series.

METHODS

Autoregressive Processes AR (p)

The autoregressive model by Grunwald *et al* (1993) is meant to explain the present value of the series, X_t, by a function of past values, $C_{t-1,}C_{t-2}, \dots, C_{t-p}$. An autoregressive process of order p is written as;

$$C_t = \phi_1 C_{t-1} + \phi_2 C_{t-2} + \dots + \phi_i C_{t-p} + e_t \equiv \sum_{i=1}^{p} \left(\phi_i C_{t-p} + e_t \right) \tag{2}$$

where (e_t) is white noise, $\emptyset_i$ coefficient of autoregressive, p is the order of the autoregressive.

Stationarity of Linear Autoregressive AR (p) processes

Using iterated expectations of an AR (p) process according to Brandt and William (2000) the equation (3.3) generates a stationary mean m_t

$$E\left[E\left(C_t \mid C_{t-1}\right)\right] = E\left[\sum_{i=1}^{p} \rho_i C_{t-1} + \lambda\right] \tag{3}$$

$$E\left[C_t\right] = \sum_{i=1}^{p} \rho_i E\left[C_{t-1}\right] + \lambda \tag{4}$$

It can be shown that equation (4) converges to a geometric series for ρ such that $\lambda = \left(1 - \sum_{i=1}^{p} \rho_i\right)\mu$. Since $E(c_0) = \mu$ then equation (4) can be expressed as

$$E[C_t|C_{t-1}] = \sum_{i=1}^{p} \rho_i C_{t-1} + \left(1 - \sum_{i=1}^{p} \rho_i\right)\mu$$

which is a stationary linear AR (*p*) process and has two major parts;

i. Measurement equation $\Pr(c_t|m_t) = \frac{e^{-m_t} m_t^{c_t}}{c_t!}$ is a Poisson distribution used in measuring the conditional mean of the homogeneous Markov process and

ii. The transition equation, is $m_t = \sum_{i=1}^{p} \rho_i c_{t-1} + \left(1 - \sum_{i=1}^{p} \rho_i\right)\mu$

Mixture of Poisson Autoregressive (MPAR) model

In this section a Mixture Poisson Autoregressive (MPAR) model of order *p* for analyzing time series count data that exhibit heterogeneity. As explained by Wood *et al* (2011), Brandt & William (2001) and Grunwald *et al* (1993), a typical Mixture model is of the form as:

$$g(c) = \sum_{i=1}^{k} \pi_i g(c; m_t) \tag{5}$$

where $g(c)$ is a function with c variables, $c_1, c_2, \ldots, c_n$ are generated from Poisson process

π_k = Mixing proportion of a give *kth* component

k =Number of components,

m_t = Parameter of the mixture model which include π, ρ *and* λ

Brandt & William (2001) the state variable of the marginal Poisson distribution evolves according to a stationary AR(p) process with autoregressive parameters $\rho_i, i = 1,2, \ldots, p$ is given by;

$$m_t = \sum_{i=1}^{p} \rho_i c_{t-1} + \left(1 - \sum_{i=1}^{p} \rho_i\right)\mu \tag{6}$$

Considering equations (5) and (6), Mixture Poisson Autoregressive (p) model can be written as:

$$g(c) = \sum_{j=1}^{k} \pi_j \left[\sum_{i=1}^{p} \rho_{ij} c_{t-1} + \left(1 - \sum_{i=1}^{p} \rho_{ij}\right)\mu_j\right] \tag{7}$$

Equation (7) can be simplified and be written as:

$$g(c)=\sum_{j=1}^{k}\sum_{i=1}^{p}\pi_j\left[\rho_{ij}c_{t-1}+\left(1-\sum_{i=1}^{p}\rho_{ij}\right)\mu_j\right] \tag{8}$$

where;

π_j= mixing weight (probabilities), $\sum_{j=1}^{k}\pi_j = 1$and π_j is nonnegative

k= number of components in the mixture distribution

ρ_{ij}= coefficient of autoregressive parameters for the k^{th} components

$p =$ total number of lags

c_t= current observed sample data

c_{t-i}= immediate past observed sample data

μ_j= mean of a give kth component

Expectation–Maximization (EM) Algorithm

The Expectation-Maximization Algorithm is the steps involves in finding the parameters of the model where it comprises of E-steps and M-steps. The E-step and M-step are alternate repeatedly, wherein their subsequent executions, where the initial fit parameters ω replaced by the current $\omega^{(v)}$, say $\omega^{(v-1)}$ at the *k-th* stage. The process stops when it converges, say when the difference between the current parameters and the previous or initial parameters is 0.0001 or $L(\omega^{(v+1)}) \geq \omega^{(v)}$

Steps Involves in Fitting the Model

Mixture of Poisson Autoregressive (PAR) using a time series count data, with Poisson distribution as the Measurement equation, and autoregressive model as transition model. According to Ghojogh *et al,* (2019), likelihood for a mixture model can be determined as:

$$L(\omega)=f\left(y_1,y_2,\ldots,y_{t-i};\rho_1,\rho_2,\ldots,\rho_k\right)=\prod_{i=1}^{n}f\left(y_i;\rho_1,\rho_2,\ldots,\rho_k\right)$$

$$L(\omega)=\prod_{i=1}^{n}\sum_{j=1}^{k}\pi_k f\left(y_i;\rho_k\right)L\left(\rho_1,\rho_2,\ldots,\rho_k\right)=\sum_{i=1}^{n}\log\sum_{j=1}^{k}\pi_i f\left(y_i;\rho_k\right)$$

the assumption is that $y_1, \dots, y_n$ are dependent

Obtaining the log likelihood is to determine which component does the variable or sample data belongs to. Let $Z_{i,k}$ being the unobserved variable assuming two components

$Z_{i,k} = \begin{cases} 1 & \text{if } y_i \text{ belongs to } k, \\ 0 & \text{otherwise.} \end{cases}$ and its corresponding probabilities stands as: $\begin{cases} P(Z_{i,k} = 1) = \pi_k \\ P(Z_{i,k} = 0) = 1 - \pi_k \end{cases}$

Therefore, the log-likelihood can be written as:

$$L(\omega) = \begin{cases} \sum_{i=1}^{n} \log\left[\pi_1 f(y_i; \rho_1, \lambda_1)\right] & \text{if } Z_{i,1} = 1 \text{ and } Z_{i,k} = 0, \\ \sum_{i=1}^{n} \log\left[\pi_2 f(y_i; \rho_2, \lambda_2)\right] & \text{if } Z_{i,2} = 1 \text{ and } Z_{i,k} = 0, \\ \quad\quad \mathrm{M} & \\ \sum_{i=1}^{n} \log\left[\pi_k f(y_i; \rho_k, \lambda_k)\right] & \text{if } Z_{i,k} = 1 \text{ and } Z_{i,k} = 0. \end{cases}$$

where $\omega = \pi, \rho, \lambda$

The above expression can be written as:

$$1\left(\rho_1, \dots, \rho_k\right) = \sum_{i=1}^{n}\left[\sum_{j=1}^{p} Z_{ij} \log\left(\pi_i f\left(c_i; \rho_p\right)\right)\right] \quad (9)$$

The Z_{ij} here is the incomplete or unobserved datum because it is not known whether it is $Z_{ij} = 0$ or $Z_{ij} = 1$ for a specific component. Therefore, using the EM algorithm, the algorithm try to estimate it by its expectation.

The E-step in EM-algorithm

$$Q(\omega) = \sum_{i=1}^{k}\left[\sum_{j=1}^{p}\left(E[Z_{i,k} \mid y_{t-1}, \lambda, \rho] \cdot \log\left(\pi_k f(y_{t-1}, \lambda, \rho)\right)\right)\right] \quad (10)$$

The $Z_{i,k}$ is either 0 or 1; therefore:

$$E[Z_{i,k} \mid y_{t-1}, \lambda, \rho] = 0 \cdot P(Z_{i,k} = 0 \mid y_{t-1}, \lambda, \rho) + 1 \cdot P(Z_{i,k} = 1 \mid y_{t-1}, \lambda, \rho) + 1 \cdot P(Z_{i,k} = 1 \mid y_{t-1}, \lambda, \rho)$$

Using Bayes rule,

$$P(Z_i = 1 \mid y_{t-1}, \lambda, \rho) = \frac{P(y_{t-1}, \lambda, \rho, Z_i = 1)}{P(y_{t-1}, \lambda, \rho)} P(Z_i = 1 \mid y_{t-1}, \lambda, \rho) = \frac{P(y_{t-1}, \lambda, \rho \mid Z_i = 1) P(Z_i = 1)}{\sum_{j=0}^{k} P(y_{t-1}, \lambda, \rho \mid Z_i = j) P(Z_i = j)} \quad (11)$$

Assuming for k component the marginal probability in the denominator is:

$$P(y_{t-1},\lambda,\rho)=\pi_1 f_1(y_{t-1};\rho_1)+\pi_2 f_2(y_{t-1};\rho_2)+\ldots+\pi_k f_k(y_{t-p};\rho_k)$$

Thus assume

$$\hat{\upsilon}_{ik}^{v}=\frac{\hat{\pi}^{v-1} f_1(y_{t-1},\lambda,\rho)}{\hat{\pi}_1^{v-1} f_1(y_{t-1},\lambda,\rho)+\hat{\pi}_2 f_2(y_{t-1},\lambda,\rho)+\ldots+\hat{\pi}_k f_k(y_{t-1},\lambda,\rho)} \tag{12}$$

where $\hat{\upsilon}_i=E[Z_i \mid y_{t-1},\lambda,\rho]$ is called responsibility of y_i (Friedman *et al,* 2009).

$$Q(\omega)=\sum_{i=1}^{n}\left[\hat{\upsilon}_1\log\left[\pi_1 f_1(y_{t-1},\lambda,\rho)\right]+\hat{\upsilon}_2\log\left[\pi_2 f_2(y_{t-1},\lambda,\rho)\right]+\ldots+\hat{\upsilon}_k\log\left[\pi_k f_k(y_{t-p},\lambda,\rho)\right]\right] \tag{13}$$

The M-Step

The M-Step in a mixture model, the algorithm computes new parameter values that maximize the expected log-likelihood,

$$\hat{\varnothing}_k,\hat{\pi}_k=\arg\max_{\varnothing_1,\pi_k}\left[Q(\rho_1,\ldots,\rho_p,\pi_1,\ldots,\pi_k)\right] \tag{14}$$

$$iif\ \sum_{k=1}^{K}\pi_k=1$$

The equation (14) is a problem that can be solve through optimization as follows:

Using Lagrange multiplier by introducing a new variable α known as Lagrange multiplier

$$\begin{aligned} L(\rho_1,\ldots,\rho_k,\pi_1,\pi_2,\ldots,\pi_K,\alpha)&=Q(\rho_1,\ldots,\rho_k,\pi_1,\pi_2,\ldots,\pi_K)-\alpha\left(\sum_{k=1}^{K}\pi_k-1\right)\\ &=\sum_{i=1}^{n}\sum_{k=1}^{K}\left[\hat{\upsilon}_{i,k}\log\pi_k+\hat{\upsilon}_{i,k}\log f_k(x;\rho_k)\right]-\alpha\left(\sum_{k=1}^{K}\pi_k-1\right)\end{aligned} \tag{15}$$

$$\frac{\partial L}{\partial\rho_k}=\sum_{i=1}^{n}\left[\frac{\hat{\upsilon}_{i,k}}{f_k(c_i;\rho_k)}\frac{\partial f_k(c_i;\rho_k)}{\partial\rho_k}\right]=0 \tag{16}$$

$$\frac{\partial L}{\partial\rho_k}=\sum_{i=1}^{n}\frac{\hat{\upsilon}_{i,k}}{\pi_k}-\alpha=0 \;\Rightarrow\; \pi_k=\frac{1}{\alpha}\sum_{i=1}^{n}\hat{\upsilon}_{i,k};\ \frac{\partial L}{\partial\alpha}=\sum_{k=1}^{K}\left(\pi_k-1\right)=0 \;\Rightarrow\; \sum_{k=1}^{K}\pi_k=1$$

$$\therefore\sum_{k=1}^{K}\frac{1}{\alpha}\sum_{i=1}^{n}\hat{\upsilon}_{i,k}=1 \;\Rightarrow\; \alpha=\sum_{i=1}^{n}\sum_{k=1}^{K}\hat{\upsilon}_{i,k}$$

$$\therefore \hat{\pi}_k = \frac{\sum_{i=1}^{n} \hat{\upsilon}_{i,k}}{\sum_{i=1}^{n} \sum_{k=1}^{K} \hat{\upsilon}_{i,k}} \qquad (17)$$

The value of ω, say $\omega^{(v)}$, that maximizes $Q(\omega, \omega^{(v)}) = L_c(\omega)$ from E-step with Z_{ij} replaced with υ_{ij}.Using equations (16) and (17) we obtain the new values of the estimations $\hat{\rho}_k, \hat{\lambda}_k$ *and* $\hat{\pi}_k$ in every iteration

Estimating the number of components

To estimate the number of required components in a given set of data, the study uses model estimation method of Akaike Information Criterion (AIC). Various expected number of components will be tested and the number of component with the least AIC is assumed to be the best fit for the set of count tine series data. The expected number of components is > 1 , the heterogeneous nature of the data is expected to have a minimum of two risk factors that influence the occurrence of the time series count data. The AIC equation is:

$$AIC = 2k - 2In(\hat{L}) \qquad (19)$$

Where $\hat{L}$is the likelihood function of the MPAR model

k number of components in the MPAR model.

Hartigan's Dip Test Statistic for Unimodality

The Dip test statistic according to Hartigan & Hartigan (1985), is a statistical method used to determine whether a dataset is unimodal or multimodal. The test statistic, known as the Dip statistic, measures as if the p-value is less than 0.05 the data set is said to be multimodal else unimodal

The Dip statistic, denoted as *D*, is calculated as the maximum difference between the empirical distribution function and the unimodal distribution function at any point in the sample. This statistic quantifies the departure from unimodality in the data

Test Statistic: The test statistic is called the "dip" statistic, denoted as

$$D = \max \mid F(x) - U(x)| \qquad (18)$$

where: F(x) is the empirical distribution function U(x) is the unimodal distribution function that minimizes the maximum difference, Dip test has a hypothesis which suggested

Null Hypothesis (H_0): The data comes from a unimodal distribution.

Alternative Hypothesis (H_1): The data comes from a non-unimodal distribution, meaning it is at least bimodal. The test evaluates the significance of the Dip statistic to determine whether to reject the null hypothesis in favour of the alternative hypothesis.

RESULTS

Numerical Monte Carlo Simulations

To test the validity of the Mixture Poisson Autoregressive Model, this study simulate data that exhibit Count Time series behavior and applied it using the developed R code. Mean Square Error (MSE) and Root Mean Square Error was adopted as statistical measures whereas the sample size increases the MSE and RMSE tends toward zero, this indicate the validity and best performance of the model.

Simulation Studies

The simulations result indicates and estimate the parameters of the MPAR model this include $\rho, \pi \; and \; \lambda$ with MSE and RMSE as measures for performance.

Table 1: Parameter Estimates for 3 components of MPAR model using Simulation Studies.

Initial Parameters ($\lambda_1 = 2, \lambda_2 = 4$ and $\lambda_3 = 6, \rho_1 = 0.8, \; \rho_2 = 0.6 \; and \; \rho_3 = 0.4$), mixing weights ($\pi_1 = 0.6, \pi_2 = 0.3 \; and \; \pi_3 = 0.1$)

Sample size	Comp. 1	Comp. 2	Comp. 3	MSE	RMSE
	$\lambda_1 = 2.11$	$\lambda_2 = 4.20$	$\lambda_3 = 6.06$		
150	$\rho_1 = 0.91$	$\rho_2 = 0.63$	$\rho_3 = 0.55$	177.67	13.85
	$\pi_1 = 0.37$	$\pi_2 = 0.30$	$\pi_3 = 0.33$		
	$\lambda_1 = 2.07$	$\lambda_2 = 3.99$	$\lambda_3 = 6.06$		
250	$\rho_1 = 0.93$	$\rho_2 = 0.60$	$\rho_3 = 0.50$	164.67	12.83
	$\pi_1 = 0.32$	$\pi_2 = 0.36$	$\pi_3 = 0.32$		
	$\lambda_1 = 1.94$	$\lambda_2 = 4.15$	$\lambda_3 = 5.92$		
300	$\rho_1 = 0.88$	$\rho_2 = 0.47$	$\rho_3 = 0.36$	74.17	8.61
	$\pi_1 = 0.35$	$\pi_2 = 0.33$	$\pi_3 = 0.32$		
	$\lambda_1 = 1.98$	$\lambda_2 = 4.09$	$\lambda_3 = 4.99$		
400	$\rho_1 = 0.73$	$\rho_2 = 0.57$	$\rho_3 = 0.51$	71.67	8.46
	$\pi_1 = 0.35$	$\pi_2 = 0.34$	$\pi_3 = 0.31$		
	$\lambda_1 = 2.09$	$\lambda_2 = 4.01$	$\lambda_3 = 5.89$		
500	$\rho_1 = 0.67$	$\rho_2 = 0.57$	$\rho_3 = 0.33$	57.83	7.60
	$\pi_1 = 0.34$	$\pi_2 = 0.34$	$\pi_3 = 0.32$		
700	$\lambda_1 = 1.92$	$\lambda_2 = 4.12$	$\lambda_3 = 6.04$	43.00	6.56

	$\rho_1 = 0.89$	$\rho_2 = 0.57$	$\rho_3 = 0.55$		
	$\pi_1 = 0.34$	$\pi_2 = 0.35$	$\pi_3 = 0.31$		
	$\lambda_1 = 2.10$	$\lambda_2 = 3.94$	$\lambda_3 = 5.99$		
800	$\rho_1 = 0.72$	$\rho_2 = 0.40$	$\rho_3 = 0.49$	36.67	6.06
	$\pi_1 = 0.35$	$\pi_2 = 0.31$	$\pi_3 = 0.34$		
	$\lambda_1 = 1.80$	$\lambda_2 = 4.20$	$\lambda_3 = 6.33$		
1000	$\rho_1 = 0.82$	$\rho_2 = 0.68$	$\rho_3 = 0.52$	9.50	3.08
	$\pi_1 = 0.37$	$\pi_2 = 0.33$	$\pi_3 = 0.30$		

Table 1 presents the results of estimated parameters for a three-component system using varying sample sizes, ranging from 150 to 1000. The model is characterized by three parameters for each component: λ (lambda), representing the mean of a Poisson distribution; ϱ (rho), an autoregressive coefficient; and π (pi), the mixing weight of each component. Mean Squared Error (MSE) and Root Mean Squared Error (RMSE) indicates the model performance. The most striking observation in the results of the simulations is the serial decrease in MSE estimate as sample size increases. This is evident from the decrease in both MSE and RMSE values. The MSE drops from 177.67 with 150 sample size to 9.50 with 1000 sample size, while the RMSE decreases from 13.85 to 3.08 over the same range. This trend revealed that as the sample size increases, the parameter estimates converge to the value of the true parameter.

Comparison analysis of MPAR and PD, PAR, PEWMA models

Table 2 presents the results of a comparison of different models for count data, specifically the Poisson distribution, Poisson Autoregressive (PAR) Model, Poisson Exponential Weighted Moving Average (PEWMA) Model, and Mixture of Poisson Autoregressive (MPAR) Model, using simulated data. The comparison is based on various sample sizes and the number of components in the MPAR model, with the Akaike Information Criterion (AIC)

Table 2: Comparison between Poisson distribution, PAR, PEWMA and MPAR models

Sample size	Models	Components	AIC
	Poisson	-	5016.056
	PAR	-	3143.371
	PEWMA	-	2854.731
60	MPAR	2	2155.718*
		3	2508.745
		4	2738.141
		5	2856.061
100	Poisson	-	4401.2845

	PAR	-	2251.4697
	PEWMA	-	2228.3785
	MPAR	2	2017.321*
		3	2200.6996
		4	2219.0513
		5	2228.4849
150	Poisson	-	3467.095
	PAR	-	1750.4899
	PEWMA	-	1742.7928
	MPAR	2	1672.4403*
		3	1733.5665
		4	1739.6838
		5	1742.8283
250	Poisson	-	3138.683
	PAR	-	1670.019
	PEWMA	-	1869.711
	MPAR	2	1266.897*
		3	1369.342
		4	1469.587
		5	1545.713
350	Poisson	-	3136.004
	PAR	-	1665.34
	PEWMA	-	1862.032
	MPAR	2	1364.218
		3	1266.663*
		4	1466.908
		5	1543.034
500	Poisson	-	3130.827
	PAR	-	1660.163
	PEWMA	-	1856.855
	MPAR	2	1289.041
		3	1161.486*
		4	1461.731
		5	1537.857
1000	Poisson	-	3121.73
	PAR	-	1651.066
	PEWMA	-	1847.758
	MPAR	2	1349.944
		3	1152.389*
		4	1452.634
		5	1528.76

From the results presented, the AIC was used to compare the goodness of fit of the models, with lower values indicating a better fit, as the sample size increases, the AIC values across all the models decrease. This is because larger sample sizes provide more information, which can lead to better model fits and thus lower AIC values. For the MPAR model, as the number of components increases from two to five, the AIC values generally decrease, suggesting that the model with two components is the most parsimonious and appropriate for the data. The Poisson distribution, being the simplest, has the highest AIC value across all sample sizes, indicating that it is the least suitable for the data. The PAR and PEWMA models perform better than the Poisson distribution but MPAR model still outperformed the three models. The AIC values for each model decrease as the sample size increases, with the MPAR model with two components consistently providing the best fit across all sample sizes. The MPAR model with two components is the most effective in modelling the simulated count time series data across different sample sizes, as evidenced by the lowest AIC values. This suggests that the MPAR model captures the underlying data-generating process more accurately than the other models considered. Since the MPAR model outperformed other count time series models, meanwhile in the preceding section, the MPAR model will be used to analyse real-life data.

CONCLUSION

This study developed a new model called a Mixture Poisson Autoregressive (MPAR). The MPAR model derived in this study proves to be a powerful and flexible tool for analyzing count-based time series data. The parameters of the model were estimated via the expectation maximization (EM) method. Montecarlo simulations were conducted to determine the stability of the model with respect to the sample size. The results from the simulations demonstrated that the MPAR model was stable, i.e., as the sample size tends to infinity, the root mean square error tends to zero (0). The results further revealed that the MPAR model outperform other count data AR models by capturing both the temporal dependencies and the heterogeneity.

REFERENCES

Brandt P. T. and Williams J. T. (2001). A Linear Poisson Autoregressive Model: The Poisson AR(p) Model. *Political Analysis 9*(2), 164-184.

Brandt, P. T., & Williams, J. T. (2001). A linear Poisson autoregressive model: The Poisson AR (p) model. *Political Analysis*, *9*(2), 164-184.

Davide, R. (2021). *Mixture autoregressive models with applications to heteroskedastic time series*. The University of Manchester (United Kingdom).

Ghojogh, B., Ghojogh, A., Crowley, M., & Karray, F. (2019). Fitting a mixture distribution to data: tutorial. *arXiv preprint arXiv:1901.06708*.

Grunwald, G. K., Raftery, A. E., & Guttorp, P. (1993). Time series of continuous proportions. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *55*(1), 103-116.

Hartigan, J. A., & Hartigan, P. M. (1985). The dip test of unimodality. *The annals of Statistics*, 70-84.

Wong, C. S., & W. K. Li. 2000. On a mixture autoregressive model. Journal of Royal Statistical Society B 62:95–115.

Wood, S., Rosen, O. and Kohn, R. (2011). Bayesian mixtures of autoregressive models, *Journal of Computational and Graphical Statistics 20*(1), 174–195.

Yang, T.; Gallagher, C.M.; McMahan, C.S. (2019). A robust regression methodology via M-estimation. Commun. Stat. Theory Methods, 48, 1092–1107.