پانزدهمین سمینار احتمال و فرآیندهای تصادفی
۸ و ۹ شهریور ۱۴۰۴
دانشگاه کردستان
15th Seminar On Probability and Stochastic Processes
August 30-31, 2025
University of Kurdistan
انجمن آمار ایران

فرایند INAR(1)-PJ : یک جایگزین جدید فرایند INAR(1) پواسونی

مهدی راسخی [۱]

[۱]گروه آمار، دانشکده علوم ریاضی و آمار، دانشگاه ملایر، ملایر، ایران

چکیده: در این مقاله، یک مدل خودبازگشتی مرتبه اول با مقادیر صحیح نامنفی معرفی می‌شود که نوفه فرآیند از یک توزیع یک پارامتری جدید پیروی می‌کند. برای برآورد پارامترهای این فرآیند، دو روش مختلف برآوردیابی در نظر گرفته شده است: حداکثر درستنمایی شرطی و روش یول-والکر. همچنین در یک مطالعه شبیه‌سازی عملکرد و کارایی این روش‌ها ارزیابی شده است. علاوه بر این، کاربرد فرآیند پیشنهادی را از طریق تحلیل یک مجموعه داده واقعی ارائه شده و در این قسمت برتری آن نسبت به مدل INAR(1) پواسونی نشان داده شده است.

واژه‌های کلیدی: توزیع احتمال جوچز، داده‌های شمارشی، بیش‌پراکنش، فرایند خودبازگشتی عدد صحیح مرتبه اول .

کد موضوع‌بندی ریاضی (۲۰۲۰): ۶۲M۱۰، ۶۲F۱۰، ۶۲E۱۵.

---

۱ مقدمه

داده‌های سری‌های زمانی با مقادیر اعداد صحیح نامنفی در حوزه‌های گوناگونی همچون علوم اجتماعی، اپیدمیولوژی، بیمه، مهندسی صنایع و بسیاری دیگر از زمینه‌ها به طور گسترده‌ای دیده می‌شوند. برای مدل‌سازی بهینه این نوع داده‌ها که دارای وابستگی زمانی نیز هستند، مدل خودبازگشتی عدد صحیح مرتبه اول ((INAR(1)) پیشنهاد شده است که مکنزی (۱۹۸۶) و الزید و الاش (۱۹۸۷) طور مستقل به معرفی آن پرداختند. ویژگی مشترکی که اغلب در این نوع سری‌های زمانی دیده می‌شود، بیش‌پراکنش است؛ یعنی واریانس داده‌ها بیشتر از میانگین آنهاست. با توجه به اینکه توزیع پواسون اغلب برای مدل‌سازی داده‌های شمارشی با بیش‌پراکنش مناسب نیست، برخی از پژوهشگران مدل‌های INAR(1) مبتنی بر توزیع‌های نوفه انعطاف‌پذیرتری را پیشنهاد کرده‌اند. از جمله این مدل‌ها می‌توان به موارد زیر اشاره کرد: مدل INAR(1) گسسته هندسی توسط جازی و همکاران (۲۰۱۲) ، فرایند INAR(1) مبتنی بر خانواده توزیع‌های کاتس که توسط لی و کیم (۲۰۱۷) معرفی شد. در راستای پیشرفت‌های اخیر در مدل‌سازی داده‌های شمارشی، در این مقاله توزیع پواسون-جوچز

[۱] سخنران، m.rasekhi@malayeru.ac.ir

(PJ) معرفی شده است: یک توزیع گسسته جدید با یک پارامتر که از ترکیب توزیع‌های پواسون و جوچز (امبگو و اکبیری (۲۰۲۲)) حاصل شده است. توزیع PJ به ویژه در مدل‌بندی داده‌ها با ویژگی بیش‌پراکنش، کارا است و دارای فرم‌های بسته و ساده‌ای برای تابع جرم احتمال و گشتاورهای آن می‌باشد. با استفاده از این توزیع به عنوان مکانیزم نوفه، یک فرآیند خودبازگشتی عدد صحیح مرتبه اول را توسعه داده و بصورت INAR(1)-PJ نامگذاری شده است. فرآیند پیشنهادی روی یک مجموعه داده واقعی اعمال شده است تا کاربردپذیری و انعطاف‌پذیری آن نشان داده شود. همچنین، مطالعات شبیه‌سازی برای ارزیابی دقت و کارایی روش‌های برآوردیابی پارامترها در چارچوب پیشنهادی انجام شده است. مقاله به شرح زیر سازمان‌دهی گردیده است: در بخش ۲ ، توزیع PJ معرفی و ویژگی های آماری آن به طور کامل بررسی شده است. در بخش ۳ ، یک فرایند خودبازگشتی عدد صحیح مبتنی بر نوفه‌های PJ پیشنهاد شده و خواص ساختاری مهم آن ارائه شده است. در بخش ۴ ، دو روش برآوردیابی برای پارامترهای مدل INAR-PJ مورد بحث قرار می‌گیرد: روش حداکثر درستنمایی شرطی و روش یول-والکر. در بخش ۵ ، نتایج یک مطالعه شبیه‌سازی به منظور ارزیابی دقت و کارایی این روش‌های برآوردیابی گزارش شده‌اند. در بخش ۶ ، فرآیند پیشنهادی روی یک مجموعه داده واقعی اعمال شده است تا کاربرد عملی آنها نشان داده شود.

۲ توزیع پواسن-جوچز

تابع جرم احتمال (PMF) توزیع جوچز به صورت زیر تعریف می‌شود:

$$f(x;\theta)=\frac{\theta^{۴}}{\theta^{۳}+\theta^{۲}+۶}(۱+x+x^{۳})e^{-\theta x}\quad x>۰\,, \tag{۱.۲}$$

که $\theta>۰$. تابع توزیع تجمعی (CDF) متناظر با توزیع جوچز به صورت زیر بیان می‌شود:

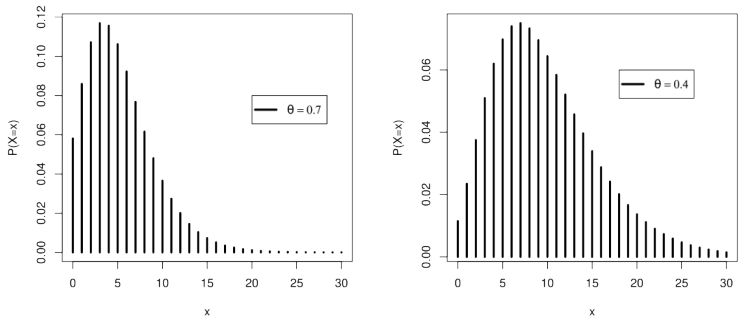
$$F(x;\theta)=۱-\left(۱+\frac{\theta x(\theta^{۲}+\theta^{۲}x^{۲}+۳\theta x+۶)}{\theta^{۳}+\theta^{۲}+۶}\right)e^{-\theta x}\quad x\geq ۰.$$

در ادامه، یک توزیع پواسون ترکیبی جدید یک پارامتری را تعریف می‌کنیم که از ترکیب توزیع پواسون با توزیع جوچز حاصل می‌شود. فرض کنید متغیر تصادفی λ دارای توزیع جوچز با پارامتر θ باشد که در برابری (۱.۲) تعریف شده است. همچنین فرض کنید که توزیع شرطی X به شرط λ ، یک توزیع پواسون با میانگین λ باشد. در این صورت، توزیع حاشیه‌ای X که از حذف λ به دست می‌آید، مربوط به توزیع پواسون-جوچز (PJ) است و دارای تابع جرم احتمال (PMF) به صورت زیر می‌باشد:

$$p(x;\theta)=\left(\frac{\theta^{۴}}{\theta^{۳}+\theta^{۲}+۶}\right)\left(\frac{۱}{(\theta+۱)^{x+۱}}+\frac{(x+۱)}{(\theta+۱)^{x+۲}}+\frac{(x+۳)(x+۲)(x+۱)}{(\theta+۱)^{x+۴}}\right),\quad x=۰,۱,۲,\ldots \tag{۲.۲}$$

فرض کنید X نشان‌دهنده متغیر تصادفی با تابع چگالی (۲.۲) باشد؛ در این صورت می‌نویسیم $X\sim PJ(\theta)$. شکل ۱ انواع مختلفی از شکل‌های تابع جرم احتمال (PMF) توزیع PJ را تحت مقادیر مختلف پارامتر نشان می‌دهد. دو گشتاور اول توزیع PJ به‌ترتیب به فرم زیر هستند:

$$\begin{aligned}E(X)&=\left(\tfrac{\theta^{۴}}{\theta^{۳}+\theta^{۲}+۶}\right)\left\{\sum_{x=۰}^{\infty}\tfrac{x}{(\theta+۱)^{x+۱}}+\sum_{x=۰}^{\infty}\tfrac{x(x+۱)}{(\theta+۱)^{x+۲}}+\sum_{x=۰}^{\infty}\tfrac{x(x+۳)(x+۲)(x+۱)}{(\theta+۱)^{x+۴}}\right\}\\&=\left(\tfrac{\theta^{۴}}{\theta^{۳}+\theta^{۲}+۶}\right)\left\{\tfrac{۱}{\theta^{۲}}+\tfrac{۲}{\theta^{۳}}+\tfrac{۲۴}{\theta^{۵}}\right\}=\tfrac{\theta^{۳}+۲\theta^{۲}+۲۴}{\theta(\theta^{۳}+\theta^{۲}+۶)},\end{aligned} \tag{۳.۲}$$



شکل ۱: نمودار تابع جرم احتمال توزیع PJ تحت مقادیر مختلف پارامتر

$$E(X^{۲}) = \left(\frac{\theta^{۴}}{\theta^{۳}+\theta^{۲}+۶}\right)\left\{\sum_{x=۰}^{\infty}\frac{x^{۲}}{(\theta+۱)^{x+۱}} + \sum_{x=۰}^{\infty}\frac{x^{۲}(x+۱)}{(\theta+۱)^{x+۲}} + \sum_{x=۰}^{\infty}\frac{x^{۲}(x+۳)(x+۲)(x+۱)}{(\theta+۱)^{x+۴}}\right\}$$
$$= \left(\frac{\theta^{۴}}{\theta^{۳}+\theta^{۲}+۶}\right)\left\{\frac{\theta+۲}{\theta^{۳}} + \frac{۲(\theta+۳)}{\theta^{۴}} + \frac{۲۴(\theta+۵)}{\theta^{۶}}\right\} = \frac{\theta^{۴}+۴\theta^{۳}+۶\theta^{۲}+۲۴\theta+۱۲۰}{\theta^{۲}(\theta^{۳}+\theta^{۲}+۶)}. \quad (۴.۲)$$

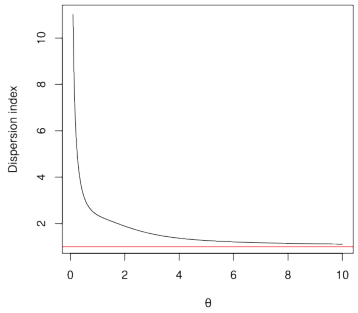
بر اساس روابط (۳.۲) و (۴.۲) واریانس توزیع PJ عبارتست از

$$V(X) = \frac{\theta^{۷}+۴\theta^{۶}+۶\theta^{۵}+۳۲\theta^{۴}+۱۲۰\theta^{۳}+۶۰\theta^{۲}+۱۴۴\theta+۱۴۴}{\theta^{۲}(\theta^{۳}+\theta^{۲}+۶)^{۲}}.$$

همچنین شاخص پراکنش به فرم زیر حاصل می شود.

$$DI(X) = \frac{\theta^{۷}+۴\theta^{۶}+۶\theta^{۵}+۳۲\theta^{۴}+۱۲۰\theta^{۳}+۶۰\theta^{۲}+۱۴۴\theta+۱۴۴}{\theta(\theta^{۳}+\theta^{۲}+۶)(\theta^{۳}+۲\theta^{۲}+۲۴)}.$$

شکل ۲ نشان می‌دهد که برای تمامی مقادیر پارامتر θ ، شاخص پراکنشی بزرگ‌تر از ۱ است. این رفتار، کارایی این توزیع در مدل‌سازی داده‌های شمارشی با بیش‌پراکنش را تأیید می‌کند.



شکل ۲: نمودار شاخص پراکنش توزیع PJ تحت مقادیر مختلف پارامتر

الگوریتم زیر می‌تواند برای تولید نمونه تصادفی از توزیع PJ با بهره بردن از روش معکوس تابع توزیع و بر اساس توزیع جوچز استفاده شود: (۱) تولید U_i، $i = ۱, ۲, ..., n$ جاییکه $Ui \sim U(۰, ۱)$.

(۲) یافتن ریشه برابری زیر بر اساس λ_i به کمک دستور uniroot در نرم افزار R.

$$\ln\left(۱+\frac{\theta\lambda_i(\theta^۲+\theta^۲{\lambda_i}^۲+۳\theta\lambda_i+۶)}{\theta^۳+\theta^۲+۶}\right)-\theta\lambda_i-\ln(۱-U_i)=۰$$

(۳) تولید X_i از توزیع $Poisson(\lambda_i)$.

۳ فرایند INAR(1) مبتنی بر نوفه‌های PJ

فرض کنید $\{X_t;t\in\mathbb{Z}\}$ نشان دهنده فرآیند INAR(1) بر اساس تعریف زیر باشد.

$$X_t=p\circ X_{t-۱}+\varepsilon_t,\quad t\in\mathbb{Z}, \tag{۱.۳}$$

که در آن $۰\le\alpha<۱$ و $\{\varepsilon_t;t\in\mathbb{Z}\}$ دنباله ای از متغیرهای تصادفی هم توزیع و مستقل با مقادیر صحیح نامنفی و میانگین $E(\varepsilon_t)=\mu_\varepsilon$ و واریانس $V(\varepsilon_t)=\sigma^۲_\varepsilon$ است. فرض کنید نوفه $\{\varepsilon_t;t\in\mathbb{Z}\}$ مستقل از مشاهدات گذشته $X_{t-k},\ k\ge۱$ است. عملگر $\circ$ نشان دهنده عملگر نازل کردن دوجمله ای است که به فرم زیر تعریف می‌شود:

$$p\circ X_{t-۱}:=\sum_{j=۱}^{X_{t-۱}}W_j$$

که در آن $\{W_j;j\ge۱\}$ نشان دهنده دنباله ای از متغیرهای تصادفی همتوزیع و مستقل با توزیع برنولی با احتمال موفقیت p است. اگر $۰\le p<۱$ این فرآیند ایستا و درغیر اینصورت فرآیند ناایستا است. برای تشخیص ساختار حرکت در فرآیند، احتمال انتقال یک مرحله ای به شرح زیر است:

$$P(X_t=k\,|X_{t-۱}=l)=\sum_{i=۱}^{\min(k,l)}\binom{l}{i}p^i(۱-p)^{l-i}P(\varepsilon_t=k-i),\ \ k,l\ge۰$$

که در آن $۰<p<۱$. برای ایجاد یک فرآیند تصادفی INAR(1) جدید، دنباله متغیرهای تصادفی $\{\varepsilon_t;t\in\mathbb{Z}\}$ که از توزیع PJ پیروی می‌کند را به عنوان متغیر های نوفه فرآیند در برابری (۱.۳) لحاظ می‌شود. درنتیجه احتمال انتقال یک مرحله فرآیند تصادفی جدید INAR-PJ(1) به فرم زیر حاصل می‌شود.

$$\begin{aligned}P(X_t=k\,|X_{t-۱}=l)=\left(\tfrac{\theta^۴}{\theta^۳+\theta^۲+۶}\right)\sum_{i=۱}^{\min(k,l)}\binom{l}{i}p^i(۱-p)^{l-i}\left(\tfrac{۱}{(\theta+۱)^{k-i+۱}}\right.\\ \left.+\tfrac{(k-i+۱)}{(\theta+۱)^{k-i+۲}}+\tfrac{(k-i+۳)(k-i+۲)(k-i+۱)}{(\theta+۱)^{k-i+۴}}\right).\end{aligned} \tag{۲.۳}$$

از این پس، فرایند توصیف شده در معادله (۲.۳) فرآیند مربوطه را با INAR-PJ(1) نامگذاری میکنیم. بر اساس روش ویب (۲۰۱۸) میانگین، واریانس و شاخص پراکنش فرایند INAR-PJ(1) به‌ترتیب به صورت زیر بیان می‌شوند:

$$\mu_X=\frac{\theta^۳+۲\theta^۲+۲۴}{\theta(\theta^۳+\theta^۲+۶)(۱-p)},$$

$$\sigma^۲_X=\frac{\left(\begin{array}{l}(p+۱)\theta^۷+(۳p+۴)\theta^۶+(۲p+۶)\theta^۵+(۳۰p+۳۲)\theta^۴\\+(۳۶p+۱۲۰)\theta^۳+۶۰\theta^۲+(۱۴۴p+۱۴۴)\theta+۱۴۴\end{array}\right)}{(۱-p^۲)\theta^۲(\theta^۳+\theta^۲+۶)^۲},$$

$$DI_X = \frac{\frac{\theta^{۷}+۴\theta^{۶}+۶\theta^{۵}+۳۲\theta^{۴}+۱۲۰\theta^{۳}+۶۰\theta^{۲}+۱۴۴\theta+۱۴۴}{\theta(\theta^{۳}+\theta^{۲}+۶)(\theta^{۳}+۲\theta^{۲}+۲۴)} + p}{۱+p}.$$

تحت فرضیات مدل، بر اساس روش ارائه شده در الاش و الزید (۱۹۸۸) ، میانگین و واریانس شرطی فرآیند به ترتیب به شرح زیر حاصل می‌شود:

$$E(X_t\,|X_{t-۱}) = pX_{t-۱} + \frac{\theta^{۳}+۲\theta^{۲}+۲۴}{\theta(\theta^{۳}+\theta^{۲}+۶)},$$

و

$$V(X_t\,|X_{t-۱}) = p(۱-p)X_{t-۱} + \left(\frac{\left(\begin{array}{l} \theta^{۷}+۴\theta^{۶}+۶\theta^{۵}+۳۲\theta^{۴} \\ +۱۲۰\theta^{۳}+۶۰\theta^{۲}+۱۴۴\theta+۱۴۴ \end{array} \right)}{\theta^{۲}(\theta^{۳}+\theta^{۲}+۶)^{۲}} \right).$$

۴ برآوردیابی پارامترهای فرآیند

۱.۴ روش ماکسیمم درستنمایی شرطی

برای انجام روش ماکسیمم درستنمایی شرطی تحت فرآیند INAR-PJ(1)، تابع درستنمایی شرطی زیر تعریف شده است.

$$\begin{aligned} \ell(p,\theta) &= \sum_{t=۲}^{T} \ln\left(P(X_t = k\,|X_{t-۱} = l)\right) \\ &= \sum_{t=۲}^{T} \ln\left(\left(\frac{\theta^{۴}}{\theta^{۳}+\theta^{۲}+۶}\right)^{\min(X_t,X_{t-۱})} \sum_{i=۱} \left(\begin{array}{c} X_{t-۱} \\ i \end{array} \right) p^{i}(۱-p)^{X_{t-۱}-i} \left(\frac{۱}{(\theta+۱)^{X_t-i+۱}} \right.\right. \\ &\quad \left.\left. + \frac{(X_t-i+۱)}{(\theta+۱)^{X_t-i+۲}} + \frac{(X_t-i+۳)(X_t-i+۲)(X_t-i+۱)}{(\theta+۱)^{X_t-i+۴}} \right) \right). \end{aligned} \tag{۱.۴}$$

برآوردگر ماکسیمم درستنمایی شرطی (CML) فرآیند INAR-PJ(1) ، که با این $\hat{p}_{CLS}, \hat{\theta}_{CLS}$ نشان داده می‌شود، با ماکسیمم کردن تابع درستنمایی شرطی برابری (۱.۴) حاصل می‌شود. برآوردیابی پارامترها را به کمک تابع optim در نرم‌افزار R انجام می‌دهیم. برای مدیریت دامنه پارامترهای p و θ از الگوریتم L-BFGS-B استفاده می‌شود. ماتریس اطلاع مشاهده شده در پارامترهای برآورد شده را با استفاده از خروجی ماتریس هِسیان حاصل از بهینه‌سازی محاسبه می‌شود. خروجی حاصل امکان محاسبه خطاهای استاندارد تقریبی برای برآوردگرها را فراهم می‌نماید. تحت شرایط نظم ، برآوردگرهای CML دارای سازگاری، توزیع نرمال مجانبی و کارایی هستند. جو (۱۹۹۷)

۲.۴ روش یول-والکر

روش برآوردیابی یول-والکر براساس برابری توابع خودکواریانس و میانگین فرآیند با حالت نمونه‌ای آنها و برآورد پارامترها با حل دستگاه معادلات ایجاد شده حاصل می‌شود. برای فرآیند INAR(1) تابع خود همبستگی نمونه ای (ACF) عبارتست از $\rho_X(h) = p^h$ که

$h \geq ۱$ و $p \in (۰, ۱)$ عملگر نازل فرآیند است. بنابراین برآورد یول-والکر p عبارتست از

$$\hat{p}_{YW} = \frac{\sum_{t=۲}^{T} (X_t - \bar{X})(X_{t-۱} - \bar{X})}{\sum_{t=۲}^{T} (X_t - \bar{X})^۲}.$$

برآوردگر یول-والکر $\hat{\theta}_{YW}$ از برابر قرار دادن میانگین نظری فرایند INAR-PJ(1) با مقدار نمونه‌ای آن به دست می‌آید. معادله برآوردیابی با قرار دادن تفاضل این دو کمیت برابر با صفر تشکیل شده و سپس $\hat{\theta}_{YW}$ از حل این معادله نسبت به θ حاصل می‌شود.

$$\left((۱ - \hat{p}_{YW})\bar{X}\right)\theta^۴ + \left((۱ - \hat{p}_{YW})\bar{X} - ۱\right)\theta^۳ - ۲\theta^۲ + \left(۶(۱ - \hat{p}_{YW})\bar{X}\right)\theta - ۲۴ = ۰ \qquad (۲.۴)$$

برابری (۲.۴) به صورت عددی به کمک تابع uniroot در نرم افزار R بدست می‌آید.

۵ مطالعات شبیه سازی

یک مطالعه شبیه‌سازی به منظور ارزیابی رفتار برآوردگرهای CML و YW بر مبنای حجم متفاوت نمونه ای برای پارامترهای مدل INAR-PJ(1) انجام شده و این شبیه سازی بر اساس ۱۰۰۰ تکرار برای حجم های نمونه $n = ۱۰۰, ۳۰۰, ۵۰۰$ اجرا شده است. نمونه‌های تصادفی با توزیع PJ با استفاده از الگوریتم معرفی شده در بخش ۱ تولید شده‌اند. یک مجموعه انتخابی از مقادیر پارامتر برای انجام شبیه‌سازی عبارت است از $p = ۰/۵, \theta = ۱$. عملکرد برآوردگرهای CML و YW با استفاده از سه معیار انحراف (Bias) ، میانگین مربعات خطا (MSE) و میانگین خطای نسبی (MRE) ارزیابی می‌شود. این معیارها به صورت زیر تعریف می‌شوند:

$$Bias = \sum_{j=۱}^{N} \frac{\hat{\gamma}_{i,j} - \gamma_i}{N}, \quad MSE = \sum_{j=۱}^{N} \frac{(\hat{\gamma}_{i,j} - \gamma_i)^۲}{N}, \quad MRE = \sum_{j=۱}^{N} \frac{\hat{\gamma}_{i,j}/\gamma_i}{N},$$

برای $i = ۱, ۲$ و $\gamma = (p, \theta)$.

مقادیر محاسبه شده برای Bias و MSE که در جدول ۱ گزارش شده‌اند، به طور کلی نزدیک به صفر هستند و این موضوع میزان اطمینان هر دو روش برآوردیابی را در حجم‌های محدود تأیید می‌کند. با افزایش حجم نمونه، میانگین خطاهای نسبی (MREs) به سمت یک همگرا شده، که نشان از بهبود دقت برآوردیابی با افزایش حجم داده‌ها دارد. این نتایج نشان می‌دهند که روش CML از لحاظ هر دو معیار Bias و MSE در تمامی حجم‌های نمونه‌ای، عملکرد بهتری نسبت به روش YW دارد. علاوه بر این، مقادیر MRE محاسبه شده نیز برتری روش CML را تأیید می‌کنند، زیرا این مقادیر به طور مداوم به مقدار ایده‌آل یک نزدیک‌تر هستند.

۶ کاربرد فرآیند پیشنهادی

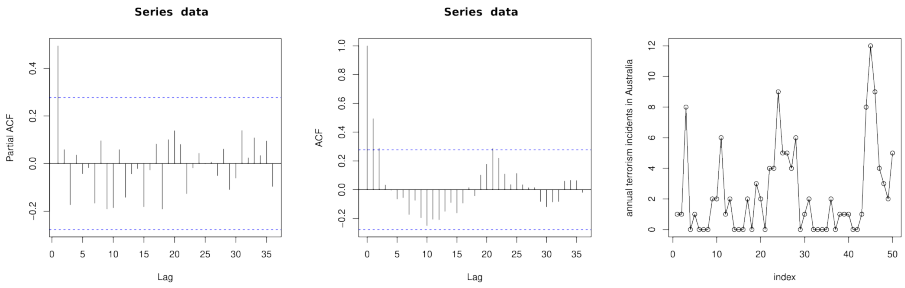
در این بخش، یک مجموعه داده واقعی برای بررسی کاربرد عملی فرایند پیشنهادی INAR-PJ(1) در مقایسه با مدل استاندارد INAR(1) پواسونی (INAR-P(1)) مورد تحلیل قرار می‌گیرد. برای مقایسه مدل‌های رقیب، از معیار AIC و معیار BIC استفاده می‌کنیم، به‌طوری که مقادیر کمتر نشان‌دهنده برازش بهتر مدل به داده‌ها است. ما یک سری زمانی سالانه از وقایع تروریستی در استرالیا که از سال ۱۹۷۰ تا ۲۰۲۰ ادامه داشته است، مورد تحلیل قرار می‌دهیم. داده‌ها از بسته Ecdat در نرم‌افزار R به دست آمده‌اند. آمار توصیفی در این مجموعه داده عبارتست از میانگین ۲/۴، واریانس ۸/۴۸ و شاخص پراکنش ۳/۵۳ ، که به‌خوبی رفتار بیش‌پراکنش را در داده‌های شمارشی نشان

جدول ۱: نتایج شبیه سازی برای فرآیند INAR-PJ(1) : اولین مجموعه انتخابی پارامترها

$\theta = ۱$		$p = ۰/۵$			
YW	CML	YW	CML		n
−۰/۰۲۴۶	۰/۰۰۵۸	−۰/۰۳۱۶	−۰/۰۰۳۲	Bias	
۰/۰۱۸۸	۰/۰۰۹۳	۰/۰۰۸۵	۰/۰۰۲۳	MSE	۱۰۰
۰/۹۴۹۱	۰/۹۷۹۰	۰/۹۱۱۲	۰/۹۶۶۲	MRE	
−۰/۰۰۶۹	۰/۰۰۲۴	−۰/۰۱۰۳	−۰/۰۰۱۵	Bias	
۰/۰۰۶۳	۰/۰۰۳۵	۰/۰۰۲۶	۰/۰۰۰۹	MSE	۳۰۰
۰/۹۸۲۷	۰/۹۹۱۰	۰/۹۶۸۷	۰/۹۸۵۲	MRE	
−۰/۰۰۲۷	۰/۰۰۱۹	−۰/۰۰۴۹	−۰/۰۰۰۲	Bias	
۰/۰۰۴۰	۰/۰۰۱۹	۰/۰۰۱۵	۰/۰۰۵	MSE	۵۰۰
۰/۹۸۶۲	۰/۹۹۰۹	۰/۹۷۹۰	۰/۹۸۸۴	MRE	

می‌دهند و لزوم استفاده از مدل‌هایی قوی تر از مدل استاندارد INAR(1) پواسونی را مطرح می‌کند. روش ارائه شده در ویب و شویر (۱۹۸۶) یک آزمون برای وجود بیش‌پراکنش در فرایندهای خودبازگشتی با مقادیر عدد صحیح نامنفی ارائه دادند. با اعمال این روش روی مجموعه داده حاضر، آماره آزمونی برابر با ۱۱/۳۷ و مقدار p آزمون کمتر از ۰/۰۰۱ به دست می‌آید که منجر به رد فرضیه صفر هم‌پراکنش می‌شود. این نتیجه برجسته می‌کند که توزیع نوفه مدل باید قادر به مدل‌سازی تغییرات بیشتر از میانگین باشد؛ بنابراین، ممکن است مدل مبتنی بر پواسون برای این داده‌ها مناسب نباشند.

شکل ۳ نمودار سری زمانی تعداد سالانه حملات تروریستی در استرالیا را به همراه توابع خودهمبستگی (ACF) و خودهمبستگی جزئی (PACF) برآورد شده را نشان می‌دهد. تابع PACF پس از گام اول قطع می‌شود، که نشان می‌دهد ساختار خودبازگشتی مرتبه اول به خوبی وابستگی دنباله‌ای داده‌ها را مدل‌سازی می‌کند. بنابراین، مدلی از نوع AR(1) برای مدل‌سازی این مجموعه داده مناسب است.



شکل ۳: نمودار سری زمانی، تابع خودهمبستگی (ACF) و تابع خودهمبستگی جزئی (PACF) برای تعداد سالانه حملات تروریستی در استرالیا

در جدول ۲ ، پارامترهای برآوردشده مدل‌ها، خطاهای استاندارد مرتبط به آنها و معیارهای انتخاب مدل ارائه شده است. مدل INAR-PJ(1) از نظر هر دو معیار AIC و BIC مقادیر پایین‌تری را نسبت به مدل INAR-P(1) ارائه می‌دهد. این نتایج حاکی از عملکرد بهتر مدل پیشنهادی در توصیف داده‌های مربوط به حملات تروریستی سالانه در استرالیا دارد.

جدول ۲: برآورد ماکسیمم درستنمایی شرطی پارامترهای مدل، خطای استاندارد، AIC، BIC و آماره های مرتبط با تعداد سالانه حملات تروریستی در استرالیا

مدل	پارامتر	برآورد	خطای استاندارد	AIC	BIC	μ_X	σ_X^2	DI_X
INAR-PJ(۱)	p	۰/۲۳۷	۰/۰۱۱۰	۲۰۷/۸۴۹	۲۱۱/۶۳۳	۲/۳۶	۴/۴۵	۱/۸۸
	θ	۱/۵۱۴	۰/۰۳۱					
INAR-P(۱)	p	۰/۳۳۱	۰/۰۰۵	۲۳۹/۳۳۹	۲۴۳/۱۲۳	۲/۴۶	۲/۴۶	۱
	λ	۱/۶۵۰	۰/۰۵۱					
تجربی						۲/۴	۸/۴۸	۳/۵۳

مراجع

Al-Osh, M and Alzaid, A (1987). First-order integer-valued autoregressive (INAR (1)) process. *J Time Ser Anal*,**8(3)**: 261-275.

Alzaid, A, Al-Osh, M (1988). First-order integer-valued autoregressive (INAR (1)) process: distributional and regression properties. *Stat Neerland*,**42(1)**: 53-61.

Echebiri, U.V., Mbegbu, J.I. (2022). Juchez Probability Distribution: Properties and Applications. *Asian Journal of Probability and Statistics*,**20(2)**: 56-71.

Jazi, M., Jones, G. and Lai, C.D. (2012). Integer valued AR (1) with geometric innovations. *J. Iran. Stat. Soc. (JIRSS)*,**11**: 173-190.

Joe, H (1997). Multivariate models and multivariate dependence concepts. *Chapman and Hall, London*.

Kim, H. and Lee, S. (2017). On first-order integer-valued autoregressive process with Katz family innovations. *J Stat Comput Simul*,**87**: 546-562.

McKenzie, E (1986). Autoregressive moving-average processes with negative binomial and geometric marginal distrbutions.*Adv Appl Probabx*, **18**: 679-705.

Schweer, S, WeiB, CH (2014). Compound Poisson INAR (1) processes: stochastic properties and testing for overdispersion. *Comput Stat Data Anal*,**77**: 267-284.

WeiB, C.H.(2018). An introduction to discrete-valued time series. *Wiley*.

The INAR(1)-PJ Process: A New Alternative to the Poisson INAR(1) Process

Mahdi Rasekhi[1]

[1]Department of Statistics, Faculty of Mathematical Sciences and Statistics,

Malayer University, Malayer, Iran

Abstract: In this paper, we introduce a first-order non-negative integer-valued autoregressive model, in which the innovation process follows a newly proposed one-parameter distribution. To estimate the unknown parameters of the model, two estimation approaches are considered: conditional maximum likelihood and Yule-Walker methods. A simulation study is conducted to evaluate the performance and efficiency of these estimation techniques. Furthermore, the practical applicability of the proposed model is demonstrated through the analysis of a real-world data sets, showing that it outperforms the traditional Poisson INAR(1) model.

Keywords: Juchez probability distribution, count data, over-dispersion, INAR(1) process.

Mathematics Subject Classification (2020): ۶۲M۱۰، ۶۲F۱۰، ۶۲E۱۵.
