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

مدل خودبازگشتی صحیح مقدار مرتبه اول فصلی با توزیع حاشیه‌ای دلاپورت

مریم شالباف [۱] ، غلامعلی پرهام [۲]

[۱]گروه آمار و ریاضی، واحد شوشتر، دانشگاه آزاد اسلامی، شوشتر، ایران

[۲]گروه آمار، دانشکده علوم ریاضی و کامپیوتر، دانشگاه شهید چمران اهواز

چکیده: در سری‌های زمانی شمارشی اغلب پدیده بیش پراکنشی مشاهده می‌شود. در این مقاله، یک فرایند خودبازگشتی صحیح مقدار مرتبه اول با ساختار فصلی معرفی می‌شود. توزیع حاشیه‌ای مدل توزیع دلاپورت می‌باشد. بر این اساس توزیع نوآوری‌ها پیچشی از توزیع پواسون با تعداد α تا توزیع هندسی تعدیل‌یافته در صفر خواهد شد. برخی از خواص مدل نمایش داده می‌شود. برای برآورد پارامترها از روش‌های یول–والکر، حداقل مربعات شرطی و حداکثر درستنمایی شرطی استفاده می‌شود. با انجام یک شبیه‌سازی عملکرد این برآوردگرها مورد ارزیابی قرار می‌گیرد. در نهایت این مدل به یک مجموعه داده واقعی برازش داده می‌شود.

واژه‌های کلیدی: توزیع دلاپورت، مدل خودبازگشتی صحیح مقدار مرتبه اول فصلی، α تا توزیع هندسی تعدیل‌یافته، عملگر رقیق‌کننده دوجمله‌ای

---

۱ مقدمه

داده‌های سری زمانی با ویژگی‌های فصلی را می‌توان در زمینه‌های مختلفی مانند علوم آماری، بهداشت، اقتصاد، محیط زیست و غیره یافت. آن‌ها عمدتاً یک الگوی فصلی با دوره‌هایی را نشان می‌دهند که پس از یک بازه زمانی منظم تکرار می‌شوند. کوچکترین دوره زمانی برای این رویداد، دوره فصلی نامیده می‌شود. عوامل متعددی مانند آب‌وهوا و ویژگی‌های ذاتی می‌توانند باعث ایجاد ساختارهای فصلی شوند. مدل‌های سری زمانی گسسته مقدار (INAR) معمولاً براساس عملگر رقیق‌کننده دوجمله‌ای که توسط استیوتل و ون هارن (۱۹۷۹) به‌صورت $\rho \circ X = \sum_{i=1}^{X} Y_i$ تعریف شده است، ساخته می‌شوند، به‌طوری که $\{Y_t\}_{t\in\mathbb{Z}}$ یک دنباله از متغیرهای تصادفی مستقل و هم‌توزیع برنولی با احتمال موفقیت $\rho \in [0, 1]$ می‌باشد و X یک متغیر تصادفی صحیح مقدار غیرمنفی و مستقل از Y می‌باشد. ملکنزی (۱۹۸۵) و آلاوش و آلزاید (۱۹۸۷) مدل INAR(1) را به‌صورت زیر معرفی کردند:

$$X_t = \rho \circ X_{t-1} + \varepsilon_t \qquad (1.1)$$

---

[۱] سخنران، Maryamshal@iau.ac.ir

که در آن $0 \leq \rho < 1$ و ε_t دنباله‌ای از متغیرهای تصادفی با مقادیر صحیح غیرمنفی ناهمبسته با میانگین μ_ε و واریانس متناهی σ_ε^2 است. اگر $\varphi_X(s)$ و $\varphi_\varepsilon(s)$ به ترتیب نشان‌دهنده تابع مولد احتمال $\{X_t\}_{t\in\mathbb{Z}}$ و $\{\varepsilon_t\}_{t\in\mathbb{Z}}$ باشند، آنگاه توزیع حاشیه‌ای $\{X_t\}_{t\in\mathbb{Z}}$ را می‌توان به‌صورت $\varphi_X(s) = \varphi_X(1-\rho+\rho s)\varphi_\varepsilon(s)$ نوشت.

متداول‌ترین فرایند برای مدل‌بندی این نوع از داده‌ها، مدل پواسون خودبازگشتی گسسته مقدار مرتبه اول (PINAR(1)) است بیش پراکنشی (یعنی حالتی که واریانس از میانگین بزرگتر است) پدیده‌ای است که معمولا در سری‌های شمارشی مشاهده می‌شود. چون توزیع پواسون هم پراکنش است (یعنی میانگین با واریانس برابر است) فرایند PINAR برای مدل‌سازی داده‌های با بیش پراکنشی مناسب نیست. یک رویکرد معمول برای برخورد با بیش پراکنشی در داده‌های شمارشی استفاده از توزیع‌های پواسون آمیخته است که توسط بارتو–سوآزا (۲۰۱۹) بررسی شده است. توزیع‌های پواسون آمیخته بوسیله معرفی یک اثر تصادفی پنهان روی میانگین توزیع پواسون به‌دست می‌آیند. مدل خود بازگشتی صحیح مقدار با توزیع حاشیه‌ای دلاپورت توسط شالباف و همکاران (۲۰۲۲) بررسی شده است. در سال‌های اخیر بر روی مدل‌های خود بازگشتی صحیح مقدار مطالعات زیادی صورت گرفته است ولی فرایندهای فصلی با توزیع حاشیه‌ای گسسته به‌ندرت بررسی و معرفی شده‌اند. بورگینکون و همکاران (۲۰۱۶) فصلی بودن را براساس مدل PINAR(1) بررسی کرده‌اند. فرایند $INAR(1)_s$ فصلی مرتبه اول با دوره فصلی s ساختار ساده‌ای دارد و برای برآورد کردن پارامترهای کمتری دارد اما برای مدل‌سازی سری‌های زمانی فصلی با بیش پراکنشی مناسب نیست. تیان و همکاران (۲۰۱۸) یک فرایند INAR(1) فصلی را براساس عملگر رقیق‌کننده دوجمله‌ای منفی و با توزیع حاشیه‌ای هندسی معرفی کردند. در این مقاله ما در جهت کمک به تحلیل سری‌های زمانی گسسته مقدار، یک فرایند خودبازگشتی صحیح مقدار مرتبه اول فصلی با توزیع حاشیه‌ای دلاپورت $(DELINAR(1)_s)$ را معرفی می‌کنیم. این مدل برای مدل‌سازی سری‌های زمانی فصلی با مقدار صحیح غیرمنفی و با بیش پراکنشی مناسب است. بر این اساس مدل در بخش ۲ معرفی شده است. همچنین برخی از ویژگی‌های مدل بیان شده است. در بخش ۳، روش‌های برآورد پارامترهای مدل مورد بحث قرار می‌گیرد. در بخش ۴ برخی از نتایج پیش‌بینی ارائه شده است. بخش ۵ برخی از نتایج شبیه‌سازی را برای روش‌های برآورد نشان می‌دهد. در بخش ۶، مدل بر روی یک مجموعه داده واقعی اعمال می‌شود. در نهایت در بخش ۷ نتایج بیان شده است.

۲ فرایند INAR(1) فصلی با توزیع حاشیه‌ای دلاپورت

در این بخش، ویژگی‌های ساختاری این فرایند، مانند توزیع‌های حاشیه‌ای و نوآوری، میانگین و واریانس این توزیع‌ها، تابع خودکوواریانس، امید شرطی و واریانس شرطی متغیر تصادفی حاشیه‌ای و احتمالات انتقال را بررسی می‌کنیم.

۱.۲ توزیع دلاپورت

متغیر تصادفی X دارای توزیع دلاپورت با پارامترهای λ، α و β است و آن را با نماد $DEL(\lambda,\alpha,\beta)$ نشان می‌دهند اگر تابع مولد احتمال آن به‌صورت $G_X(S) = e^{-\lambda(1-s)}\left(\frac{1}{1+\beta(1-s)}\right)^\alpha$ باشد، به‌طوری‌که $|s| \leq 1$، $\lambda > 0$ و $\alpha,\beta > 0$ (دلاپورت، ۱۹۵۹).
با دوبار مشتق گرفتن از تابع مولد احتمال توزیع دلاپورت می‌توان نشان داد که:

$$\mu = E(X_t) = \lambda + \alpha\beta \quad , \quad \sigma^2 = Var(X_t) = \lambda + \alpha\beta(1+\beta)$$

با توجه به اینکه واریانس بزرگ‌تر از میانگین است، می‌توان گفت این توزیع بیش پراکنشی دارد.

۲.۲ فرایند خودبازگشتی صحیح مقدار مرتبه اول فصلی با توزیع حاشیه‌ای دلاپورت $DELINAR(۱)_s$

فرض کنید $N_۰$ نشان‌دهنده مجموعه مقادیر صحیح غیرمنفی باشد. همچنین فرض کنید X یک متغیر تصادفی با مقادیر صحیح غیرمنفی و $\rho \in [۰, ۱]$ باشد. با استفاده از عملگر رقیق‌کننده دوجمله‌ای، مدل دلاپوریت فصلی مرتبه اول به‌صورت

$$X_t = \rho \circ X_{t-s} + \varepsilon_t \tag{۱.۲}$$

تعریف می‌شود، به‌طوری که $\rho \in [۰, ۱]$ است و $\{\varepsilon_t\}$ دنباله نوآوری از متغیرهای تصادفی با مقادیر صحیح غیرمنفی و مستقل از مقادیر گذشته $\{X_t\}$ می‌باشند و $s \in \mathbb{N}$ دوره فصلی را نشان می‌دهد. متغیرهای برنولی در عملگرهای رقیق‌کننده مستقل از هم و مستقل از نوآوری $\{\varepsilon_t\}$ می‌باشند. این مدل را با نماد $DELINAR(۱)_s$ نشان می‌دهیم.

قضیه ۱.۲. متغیر تصادفی ε_t را می‌توان به صورت $\varepsilon_t = Y_۱ + Y_۲$ نشان داد که در آن $Y_۱$ دارای توزیع پواسون با پارامتر $\lambda(۱-\rho)$ و $Y_۲$ پیچش α تا توزیع هندسی تعدیل‌یافته در صفر است. بنابراین امید ریاضی و واریانس ε_t به‌صورت $\mu_\varepsilon := E(\varepsilon_t) = (۱-\rho)\mu_X$ و $\sigma^۲_\varepsilon := Var(\varepsilon_t) = (۱-\rho^۲)\sigma^۲_X - \rho(۱-\rho)\mu_X$ می‌باشند. همچنین احتمال شرطی X_t بشرط $X_{t-۱}$ به‌صورت زیر به‌دست می‌آید.

$$p_{ij} = P(X_t = j | X_{t-۱} = i) = P(\rho \circ X_{t-۱} + \varepsilon_t = x_t | X_{t-۱} = x_{t-۱}).$$

اکنون فرض کنید

$$B^\rho_{X_{t-۱}} := \rho \circ X_{t-۱} | X_{t-۱} \sim Binomial(X_{t-۱}, \rho)$$

همچنین ε_t مستقل از $X_{t-۱}$ می‌باشد و دارای ساختار پیچشی از دو متغیر تصادفی $P_{(\lambda,\rho)} \sim Po(\lambda(۱-\rho))$ و پیچش α تا توزیع هندسی تعدیل‌یافته در صفر با پارامترهای $p = \frac{۱}{۱+\beta}$ و $p_۰$ می‌باشد، که $p_۰ = \frac{(۱+\rho\beta)}{۱+\beta}$ میزان جرم احتمال در صفر می‌باشد. بنابراین احتمال شرطی را می‌توان به‌صورت زیر نوشت:

$$\begin{aligned} P(X_t = x_t | X_{t-۱} = x_{t-۱}) &= P(B^\rho_{X_{t-۱}} + P_{(\lambda,\rho)} + \alpha FZMG(p, p_۰) = x_t) \\ &= \sum_{s=۰}^{x_t} P(\alpha FZMG = s) \sum_{d=۰}^{\min\{x_t - s, x_{t-۱}\}} P(B^\rho_{X_{t-۱}} = d) P(P_{(\lambda,\rho)} = x_t - s - d), \end{aligned} \tag{۲.۲}$$

۳ برآورد پارامترها

فرض کنید $X_n, \ldots, X_۲, X_۱$ یک نمونه از فرایند $DELINAR(۱)_s$ باشند. در این فرایند ما ۴ پارامتر داریم، فرض می‌کنیم پارامتر α معلوم است پس ۳ پارامتر دیگر باید برآورد شوند. در ادامه روش‌های یول والکر، حداقل مربعات شرطی و حداکثر درستنمایی شرطی و برخی از نتایج تحلیلی و مجانبی برآوردگرها را معرفی و توضیح می‌دهیم.

۱.۳ برآورد یول والکر

فرض کنید $\gamma(k) = \sum_{t=۲}^{n-k}(X_t - \bar{X})(X_{t+k} - \bar{X})$ و $۰ \leq k < n$ تابع خودکواریانس نمونه‌ای باشد. چون $\rho = \frac{\gamma(۱)}{\gamma(۰)}$ و $E(X_t) = \lambda + \alpha\beta$ و $Var(X_t) = \lambda + \alpha\beta(۱+\beta)$، برآوردهای یول والکر پارامترها به‌صورت زیر می‌باشد:

$$\hat{\rho}_{YW} = \frac{\sum_{t=s+۱}^{n}(X_t - \bar{X})(X_{t-s} - \bar{X})}{\sum_{t=۱}^{n}(X_t - \bar{X})^۲} \quad , \quad \hat{\beta}_{YW} = \sqrt{\frac{S^۲ - \bar{X}}{\alpha}} \quad , \quad \hat{\lambda}_{YW} = \bar{X} - \alpha\sqrt{\frac{S^۲ - \bar{X}}{\alpha}},$$

به‌طوری‌که $\bar{X}$ و $S^۲$ بترتیب نشان‌دهنده میانگین نمونه‌ای و واریانس نمونه‌ای می‌باشد.

۲.۳ برآورد حداقل مربعات شرطی

روش حداقل مربعات شرطی را در دو مرحله اجرا می‌کنیم (کلیمکو و نلست، ۱۹۸۷). در گام اول با نوشتن معادله حداقل مربعات شرطی برای پارامترهای ρ و μ و به کمک مشتق گیری جزئی نسبت به پارامترهای ρ و μ از معادله و کمینه کردن آن داریم:

$$\hat{\mu}_{cls} = \frac{\sum_{t=s+۱}^{n} X_t - \hat{\rho}_{cls}\sum_{t=s+۱}^{n} X_{t-s}}{(n-s)(۱-\hat{\rho}_{cls})}, \quad \hat{\rho}_{cls} = \frac{(n-s)\sum_{t=s+۱}^{n} X_t X_{t-s} - \sum_{t=s+۱}^{n} X_t \sum_{t=s+۱}^{n} X_{t-s}}{(n-s)\sum_{t=s+۱}^{n} X_{t-s}^۲ - (\sum_{t=s+۱}^{n} X_{t-s})^۲}.$$

در گام دوم، با نوشتن معادله حداقل مربعات شرطی برای پارامتر $\sigma^۲$ و کمینه کردن آن نسبت به $\sigma^۲$ داریم:

$$\hat{\sigma}_{cls}^۲ = \frac{\sum_{t=s+۱}^{n}(X_t - \hat{\rho}_{cls}X_{t-s} - (۱-\hat{\rho}_{cls})\hat{\mu}_{cls})^۲ - \hat{\rho}_{cls}(۱-\hat{\rho}_{cls})\sum_{t=s+۱}^{n}(X_{t-s} - \hat{\mu}_{cls})}{(n-s)(۱-\hat{\rho}_{cls}^۲)}$$

سرانجام با حل دستگاه و فرض معلوم بودن α، برآورد کمترین مربعات شرطی برای پارامترهای λ و β به‌صورت زیر به‌دست می‌آید:

$$\hat{\lambda}_{cls} = \hat{\mu}_{cls} - \alpha\hat{\beta}_{cls} \quad , \quad \hat{\beta}_{cls} = \sqrt{\frac{\hat{\sigma}_{cls}^۲ - \hat{\mu}_{cls}}{\alpha}}.$$

۳.۳ برآورد حداکثر درستنمایی شرطی

برآوردگرهای حداکثر درستنمایی شرطی با بیشینه کردن تابع لگاریتم درستنمایی به‌دست می‌آیند. تابع لگاریتم درستنمایی شرطی به‌صورت $CL(\rho, \lambda, \beta) = \sum_{t=s+۱}^{n} \log P(X_t = j|X_{t-s} = i)$ می‌باشد، به‌طوری‌که $P(X_t = j|X_{t-s} = i)$ در (۲.۲) تعریف شده است. مقادیر برآورد حداکثر درستنمایی شرطی با بیشینه کردن تابع فوق بدست می‌آید. چون $CL(\rho, \lambda, \beta)$ یک تابع غیرخطی است، برآوردهای حداکثر درستنمایی شرطی باید به کمک روش‌های عددی محاسبه شوند.

۴ پیش‌بینی

در اینجا هدف، محاسبه پیش‌بینی مقدار $X_{n+h}(h \in \mathbb{N})$ براساس اطلاعات تا زمان n ام می‌باشد. توزیع X_{n+h} براساس تعریف مدل $DELINAR(۱)_s$ را می‌توان به‌صورت زیر بیان کرد:

$$X_{n+h} \stackrel{d}{=} \rho^q \circ X_{n+h-qs} + \sum_{j=۰}^{q-۱} \rho^j \circ \varepsilon_{n+h-js}$$

به‌طوری‌که $q = \lceil h/s \rceil$ و $\lceil x \rceil = \min\{n \in \mathbb{N} | x \leq n\}$. در واقع $\lceil x \rceil$ نشان‌دهنده کوچک‌ترین عدد طبیعی بزرگ‌تر (یا مساوی) x می‌باشد. از عبارت بالا مشاهده می‌شود که توزیع پیش‌بینی h گام جلوتر شکل بسیار پیچیده‌ای دارد و بنابراین ما پیش‌بینی h گام

جلوتر را به کمک امید شرطی زیر بدست می‌آوریم:

$$\hat{X}_{n+h}|X_n = E(X_{n+h}|X_n) = \rho^q(X_{n+h-qs} - \mu_x) + \mu_x.$$

به‌طوری که $\mu_x = \lambda + \alpha\beta$.

۵ شبیه‌سازی

در این قسمت عملکرد برآوردهای یول والکر، حداقل مربعات شرطی و حداکثر درستنمایی شرطی برای پارامترهای مدل برای اندازه‌های نمونه مختلف با دوره فصلی ۱۲ در جداول ۱ تا ۳ ارائه شده است. اریبی‌ها و میانگین توان دوم خطاها ۱۰۰۰ بار تکرار و محاسبه شده‌اند. این جداول نشان می‌دهد که اریبی و خطای استاندارد برآورد پارامترها با افزایش حجم نمونه برای همه موارد کاهش می‌یابد. همان‌طور که از جداول مشاهده می‌شود، روش حداقل مربعات شرطی و روش یول والکر، رفتارهای مشابهی را نشان می‌دهند. با توجه به جداول، برآوردگرهای حداکثر درستنمایی شرطی مقادیر کم‌تری از اریبی (مقادیر قدرمطلق) و میانگین مربعات خطا در مقایسه با برآوردگرهای یول والکر و حداقل مربعات شرطی را نشان می‌دهند.

جدول ۱: میزان اریبی و میانگین مربعات خطا (در داخل پرانتز) برآوردگرها برای $(\rho, \beta, \lambda) = (۰/۲, ۳, ۱/۵)$

$\hat{\lambda}_{CML}$	$\hat{\lambda}_{cls}$	$\hat{\lambda}_{YW}$	$\hat{\beta}_{CML}$	$\hat{\beta}_{cls}$	$\hat{\beta}_{YW}$	$\hat{\rho}_{CML}$	$\hat{\rho}_{cls}$	$\hat{\rho}_{YW}$	n
۰/۰۱۷۶	۰/۰۴۹۲	۰/۰۰۷۵	−۰/۰۱۵۹	−۰/۰۳۲۰	−۰/۰۱۰۶	۰/۰۰۰۸	−۰/۰۰۲۳	−۰/۰۱۴۲	۲۰۰
(۰/۱۷۰۸)	(۰/۲۸۰۴)	(۰/۲۶۸۵)	(۰/۰۷۶۶)	(۰/۱۰۱۳)	(۰/۰۹۵۶)	(۰/۰۰۲۷)	(۰/۰۰۵۰)	(۰/۰۰۴۷)	
۰/۰۰۶۰	۰/۰۱۵۵	−۰/۰۰۰۳	−۰/۰۰۸۲	−۰/۰۱۳۱	−۰/۰۰۵۱	۰/۰۰۰۰	−۰/۰۰۱۳	−۰/۰۰۷۴	۴۰۰
(۰/۰۸۷۰)	(۰/۱۴۲۴)	(۰/۱۳۴۴)	(۰/۰۳۷۴)	(۰/۰۴۹۸)	(۰/۰۴۷۴)	(۰/۰۰۱۳)	(۰/۰۰۲۶)	(۰/۰۰۲۵)	
۰/۰۱۱۹	۰/۰۱۹۶	۰/۰۰۹۵	−۰/۰۰۶۷	−۰/۰۱۰۵	−۰/۰۰۵۷	−۰/۰۰۱۶	−۰/۰۰۲۹	−۰/۰۰۵۹	۸۰۰
(۰/۰۴۶۵)	(۰/۰۷۷۴)	(۰/۰۷۵۹)	(۰/۰۱۸۹)	(۰/۰۲۶۸)	(۰/۰۲۶۲)	(۰/۰۰۰۷)	(۰/۰۰۱۳)	(۰/۰۰۱۳)	

جدول ۲: میزان اریبی و میانگین مربعات خطا (در داخل پرانتز) برآوردگرها برای $(\rho, \beta, \lambda) = (۰/۵, ۲, ۱)$

$\hat{\lambda}_{CML}$	$\hat{\lambda}_{cls}$	$\hat{\lambda}_{YW}$	$\hat{\beta}_{CML}$	$\hat{\beta}_{cls}$	$\hat{\beta}_{YW}$	$\hat{\rho}_{CML}$	$\hat{\rho}_{cls}$	$\hat{\rho}_{YW}$	n
۰/۰۲۷۲	۰/۰۹۵۳	۰/۰۴۶۰	−۰/۰۱۱۸	−۰/۰۴۶۱	−۰/۰۲۳۵	−۰/۰۰۲۹	−۰/۰۰۸۶	−۰/۰۳۸۵	۲۰۰
(۰/۱۵۵۵)	(۰/۲۷۷۹)	(۰/۲۴۵۹)	(۰/۰۶۶۰)	(۰/۰۹۰۸)	(۰/۰۸۰۸)	(۰/۰۰۱۹)	(۰/۰۰۴۴)	(۰/۰۰۵۴)	
۰/۰۲۵۴	۰/۰۶۴۵	۰/۰۴۵۳	−۰/۰۱۶۳	−۰/۰۳۶۷	−۰/۰۲۷۸	−۰/۰۰۳۴	−۰/۰۰۸۱	−۰/۰۲۳۱	۴۰۰
(۰/۰۷۲۵)	(۰/۱۳۵۳)	(۰/۱۲۷۸)	(۰/۰۳۱۴)	(۰/۰۴۵۱)	(۰/۰۴۲۶)	(۰/۰۰۰۸)	(۰/۰۰۲۲)	(۰/۰۰۲۶)	
۰/۰۰۱۷	۰/۰۲۳۳	۰/۰۱۴۲	−۰/۰۰۶۸	−۰/۰۱۷۷	−۰/۰۱۳۰	−۰/۰۰۱۵	−۰/۰۰۳۷	−۰/۰۱۰۹	۸۰۰
(۰/۰۳۳۲)	(۰/۰۶۰۸)	(۰/۰۵۹۶)	(۰/۰۱۴۷)	(۰/۰۲۱۴)	(۰/۰۲۰۹)	(۰/۰۰۰۴)	(۰/۰۰۱۱)	(۰/۰۰۱۲)	

جدول ۳: میزان اریبی و میانگین مربعات خطا (در داخل پرانتز) برآوردگرها برای $(\rho, \beta, \lambda) = (۰/۸, ۱/۵, ۴)$

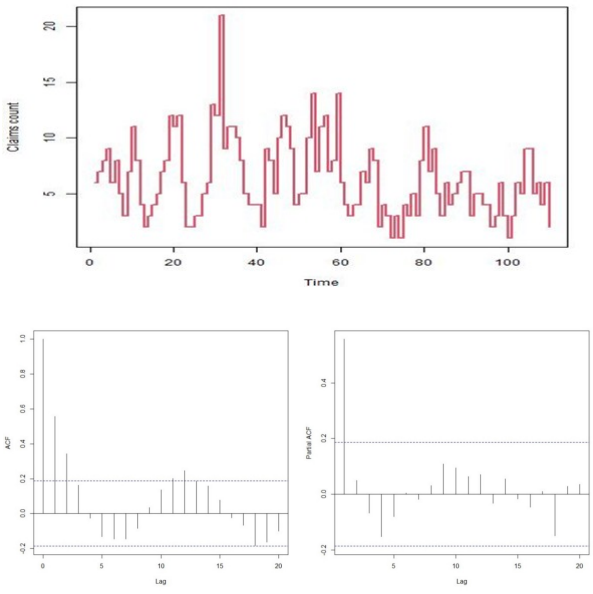
$\hat{\lambda}_{CML}$	$\hat{\lambda}_{cls}$	$\hat{\lambda}_{YW}$	$\hat{\beta}_{CML}$	$\hat{\beta}_{cls}$	$\hat{\beta}_{YW}$	$\hat{\rho}_{CML}$	$\hat{\rho}_{cls}$	$\hat{\rho}_{YW}$	n
۰/۰۴۶۱	۰/۲۶۰۲	۰/۱۵۳۲	−۰/۰۱۵۹	−۰/۱۲۴۱	−۰/۰۶۵۴	−۰/۰۰۳۰	−۰/۰۱۰۳	−۰/۰۵۸۵	۲۰۰
(۰/۴۴۸۸)	(۱/۲۰۰۵)	(۰/۸۱۵۱)	(۰/۱۰۸۸)	(۰/۲۶۷۶)	(۰/۱۹۳۶)	(۰/۰۰۰۵)	(۰/۰۰۲۲)	(۰/۰۰۵۷)	
۰/۰۱۸۰	۰/۱۷۶۸	۰/۱۱۲۰	−۰/۰۱۱۱	−۰/۰۸۸۸	−۰/۰۵۹۴	−۰/۰۰۲۱	−۰/۰۰۸۱	−۰/۰۳۲۶	۴۰۰
(۰/۲۴۱۶)	(۰/۵۶۴۵)	(۰/۴۹۴۰)	(۰/۰۵۵۵)	(۰/۱۳۳۴)	(۰/۱۱۶۵)	(۰/۰۰۰۲)	(۰/۰۰۱۰)	(۰/۰۰۲۱)	
۰/۰۳۱۸	۰/۱۱۳۰	۰/۰۹۰۷	−۰/۰۱۰۰	−۰/۰۵۰۲	−۰/۰۴۰۲	−۰/۰۰۰۹	−۰/۰۰۴۶	−۰/۰۱۶۹	۸۰۰
(۰/۱۱۳۰)	(۰/۲۶۷۳)	(۰/۲۴۸۶)	(۰/۰۲۶۹)	(۰/۰۶۷۶)	(۰/۰۶۲۵)	(۰/۰۰۰۱)	(۰/۰۰۰۵)	(۰/۰۰۰۸)	

۶ داده واقعی

در این بخش برای مدل‌سازی و پیش‌بینی و تجزیه و تحلیل از داده‌های سری زمانی شمارشی که از انجمن جبران خسارت کارگران (WCB) بریتیش کلمبیا و کانادا جمع‌آوری شده است، استفاده می‌کنیم. داده‌ها شامل ۱۲۰ مشاهده است که از ژانویه ۱۹۸۵ شروع شده و در دسامبر ۱۹۹۴ به پایان می‌رسد. میانگین نمونه ۶/۱۳ و واریانس نمونه ۱۱/۸۰ است. از این‌رو، به‌نظر می‌رسد داده‌ها بیش‌پراکنش داشته می‌باشند.

شکل ۱ نمودار سری زمانی داده‌ها، تابع خودهمبستگی (ACF) و تابع خودهمبستگی جزئی (PACF) را نشان می‌دهد. نمودار سری زمانی نشان می‌دهد که سری در میانگین ایستاست. الگوی کاهش هندسی با دوره فصلی ۱۲ را می‌توان در نمودار ACF مشاهده کرد و نشان می‌دهد که یک سری زمانی با رفتارهای همبستگی سریالی وجود دارد. فصلی بودن و بیش پراکنشی این مجموعه داده‌ها آشکار است، این ویژگی‌ها باعث شد که ما مدل $DELINAR(1)_{12}$ را به‌عنوان مدل برازش شده انتخاب کنیم. مجموعه داده‌ها به دو بخش تقسیم می‌شود. ۱۱۰ مشاهده اول برای مدل‌سازی سری استفاده می‌شوند و ۱۰ مشاهده باقیمانده برای اهداف پیش‌بینی در نظر گرفته شده است. همچنین برای مقایسه مدل‌های $INAR(1)_{12}$، $DELINAR(1)$ و $PINAR(1)$ را نیز در نظر می‌گیریم.

همان‌طور که قبلا اشاره شد پارامتر α را ثابت و معلوم در نظر گرفتیم و با توجه به این‌که α مقادیر صحیح غیرمنفی را می‌گیرد، بنابراین مدل را با مقادیر ۱، ۲ و ۳ برای α با رویکرد ماکزیمم درستنمایی شرطی برازش کردیم تا بهترین را با توجه به معیار اطلاعاتی آکاییک (AIC) و معیار اطلاعات بیزی (BIC) انتخاب کنیم. با توجه به تجزیه و تحلیل حساسیت پارامتر α، به نظر می‌رسد که معیار اطلاع آکاییک با افزایش مقدار α کمی افزایش می‌یابد. بنابراین ما برای مدل $\alpha = ۱$ را در نظر می‌گیریم.



شکل ۱: نمودار سری زمانی و تابع خودهمبستگی و تابع خودهمبستگی جزیی داده‌های واقعی از سال ۱۹۸۵ تا ۱۹۹۴

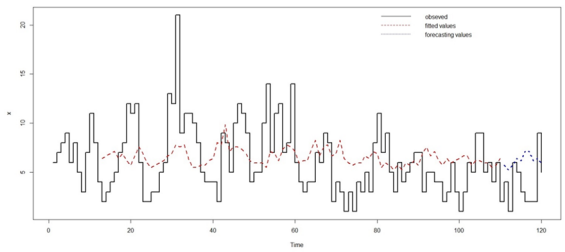
جدول ۴ برآوردهای حداقل مربعات و حداکثر درستنمایی شرطی (با میانگین مربعات خطا در پرانتز) را برای چهار مدل برازش شده و معیار اطلاع آکاییک (AIC) و معیار اطلاع بیزی (BIC) را نشان می‌دهد. همان‌طور که می‌بینیم، AIC و BIC مربوط به مدل $DELINAR(1)_{12}$ در مقایسه با مدل‌های دیگر کمترین است. نتایج بدست آمده در جدول ۴ نشان می‌دهد که مدل خودبازگشتی

با مقادیر صحیح غیرمنفی مرتبه اول دلاپوریت با ساختار فصلی بهتر از سایر مدل‌های خودبازگشتی با مقادیر صحیح غیرمنفی مرتبه اول فصلی و غیرفصلی می‌باشد.

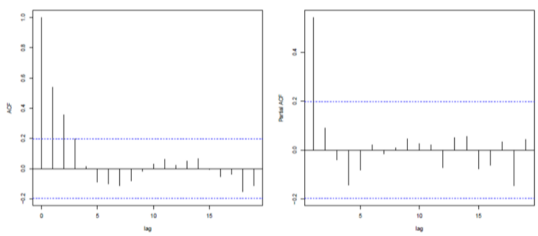
در شکل ۲ مقادیر مشاهده شده، مقادیر برازش و مقادیر پیش‌بینی به ترتیب با خطوط مشکی، قرمز و آبی نشان داده شده است. شکل ۳ نیز نشان دهنده تابع خودهمبستگی و تابع خودهمبستگی جزیی باقیمانده‌هاست و همان‌طور که مشاهده می‌شود همچنان همبستگی سریالی در باقیمانده‌ها دیده می‌شود.

جدول ۴: براورد پارامترها (میانگین مربعات خطا در پرانتز) و مقادیر AIC و BIC

مدل	پارامتر	برآورد CML	برآورد CLS	AIC	BIC
$PINAR(1)$	ρ	۰٫۵۷۰۵(۰٫۰۰۰۵)	۰٫۵۶۵۱(۰٫۰۰۷۵)	۵۴۵٫۸۰	۵۵۱٫۲۰
	λ	(۰٫۰۱۸۸)۲٫۷۰۵۹	(۰٫۳۱۹۹)۲٫۷۳۵۶		
$INAR(1)_{12}$	ρ	(۰٫۰۰۰۴)۰٫۲۴۵۹	(۰٫۰۰۹۸)۰٫۲۶۶۷	۵۳۲٫۰۹	۵۳۷٫۴۹
	λ	(۰٫۰۴۱۱)۴٫۶۷۶۹	(۰٫۴۲۰۱)۴٫۵۳۸۹		
$DELINAR(1)$	ρ	(۰٫۰۰۱۶)۰٫۴۴۳۶	(۰٫۰۰۷۰)۰٫۵۶۵۱	۵۲۷٫۰۸	۵۳۵٫۱۸
	β	(۰٫۰۰۷۷)۱٫۲۷۶۸	(۰٫۱۰۰۰)۱٫۶۸۳۱		
	λ	(۰٫۱۱۴۸)۳٫۶۹۴۵	(۰٫۳۲۲۳)۲٫۹۲۵۳		
$DELINAR(1)_{12}$	ρ	(۰٫۰۱۰۴)۰٫۲۲۸۵	(۰٫۰۱۰۰)۰٫۲۶۶۷	۵۰۰٫۲۱	۵۰۸٫۳۱
	β	(۰٫۲۰۳۲)۲٫۷۳۴۵	(۰٫۳۱۹۷)۲٫۵۰۸۹		
	λ	(۰٫۴۸۱۶)۳٫۵۵۷۵	(۰٫۲۹۱۱)۳٫۶۸۱۱		



شکل ۲: نمودار سری زمانی مقادیر مشاهده شده، برازش شده و پیش‌بینی مربوط به داده‌های واقعی از سال ۱۹۸۵ تا ۱۹۹۴



شکل ۳: تابع خودهمبستگی و تابع خودهمبستگی جزیی باقیمانده‌ها

۷ نتیجه‌گیری

در این مقاله یک فرایند خودبازگشتی مرتبه اول فصلی با توزیع حاشیه‌ای دلاپورت معرفی گردید که توزیع نوآوری مدل به‌صورت پیچشی از توزیع پواسون و α تا توزیع هندسی تعدیل‌یافته در صفر می‌باشد. در ادامه ویژگی‌های اصلی مدل مانند میانگین واریانس و تابع خودهمبستگی مورد بررسی قرارداده شد. برای برآورد پارامترها از روش‌های یول–والکر، حداقل مربعات شرطی و حداکثر درستنمایی شرطی استفاده شده است. همچنین به پیش‌بینی مدل پرداخته شده است. با انجام یک شبیه‌سازی در حجم نمونه‌های متناهی عملکرد این برآوردگرها مورد ارزیابی قرار گرفت. در نهایت این مدل به یک مجموعه داده واقعی، برازش داده شد. برای مقایسه مدل‌های $PINAR(۱)$، $INAR(۱)_{۱۲}$ و $DELINAR(۱)$ نیز در نظر گرفته شد و نتیجه نشان داد که براساس معیار AIC و BIC، مدل پیشنهادی ما در مقایسه با سایر مدل‌های INAR بهتر است. به منظور تحقیقات بیشتر، می‌توان مدل را به مرتبه‌های بیشتر گسترش داد. همچنین، طبق ACF باقیمانده‌ها که در شکل ۳ نشان داده شده است، همچنان همبستگی سریالی در باقیمانده‌ها وجود داشت بنابراین می‌توان یک مدل جدید معرفی کرد که بتواند همبستگی فصلی و سریالی را کاهش دهد.

مراجع

Al-Osh MA, Alzaid AA. (1987) First-order interger-valued autoregressive (INAR(1)) process, *J Time Ser Anal.*, **8**(3), 261–275.

Barreto-Souza W. (2019), Mixed Poisson INAR(1) processes, *Stat Pap.*, **60**(3), 2119-2139.

Bourguignon M, Vasconcellos KL, Reisen VA, Ispány M (2016), A Poisson INAR(1) process with a seasonal structure, *J Stat Comput Simul*, **86**(2), 373–387.

Delaporte P. (1959), Quelques problémes de statistiquemathématique posés par Ìassurance automobile et le bonus non sinistre, *Bulletin Trimestriel de l'lnstitut des Actuaires Franą uis.*, **227**, 87–102.

Klimko, L.A. and Nelson, P.I., (1978), On conditional least squares estimation for stochastic processes, *The Annals of statistics*, 629-642.

McKenzie E. (1985), Some simple models for discrete variate time series, *J Am Water Resour Assoc.*, **21**(4): 645–650.

Steutel FW, Van Harn K. (1979), Discrete analogues of self-decomposability and stability, *Ann Probab*; **7**(5), 893–899.

Shalbaf, M., Parham, G. and Chinipardaz, R., (2022), Binomial Thinning Integer-Valued AR (1) with Poisson-α Fold Zero Modified Geometric Innovations, *Journal of Sciences, Islamic Republic of Iran*, **33**(1), 55-63.

Tian S, Wang D, Cui S. (2020), A seasonal geometric INAR process based on negative binomial thinning operator, *Statistical Papers*, **61**(6), 2561–2581.

A seasonal Integer-Valued AR(1) model with Delaporte marginal distribution

Maryam Shalbaf[1], **Gholam Ali Parham**[2]

[1]Department of Statistics and Mathematics, Sho.C., Islamic Azad University, Shoushtar, Iran

[2]Department of Statistics, Faculty of Mathematical Sciences and Computer, Shahid Chamran University of Ahvaz

---

Abstract: Real-count data time series often show the phenomenon of the overdispersion. In this paper, we introduce the first-order integer-valued autoregressive process with a seasonal structure. The univariate marginal distribution is derived from the Delaporte distribution and the innovations are convolution of Poisson with α-fold zero modified geometric distribution, based on binomial thinning operator. Some properties of the model are derived. The methods of Yule-Walker, conditional least squares and conditional maximum likelihood are used for estimating the parameters. The Monte Carlo experiment is conducted to evaluate the performances of these estimators in finite samples. At the end this model is illustrated using a real data set.

Keywords: Delaporte distribution, seasonal INAR(1) model, α-fold zero modified geometric, binomial thining.
