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

یک کلاس از مدل‌بندی سری‌زمانی دوخطی گسسته‌مقدار با کاربردهای آن

رعنا بامدادی [1] **و مهرناز محمدپور**

گروه آمار، دانشگاه مازندران

چکیده: در این مطالعه، یک مدل تعدیل‌یافته دوخطی گسسته‌مقدار مرتبه دو برای تحلیل سری‌های زمانی غیرخطی معرفی می‌شود. این مدل از انعطاف‌پذیری بالایی برخوردار است به‌گونه‌ای که ورودی‌های جدید می‌توانند به‌طور مستقل رفتار کنند. مدل‌بندی ارائه‌شده برای فرایندهای بیش‌پراکنده به‌کار می‌رود که بر یک چارچوب غیرخطی تکیه دارند. ویژگی‌های اصلی مدل پیشنهادی به‌طور نظری بررسی شده‌است. همچنین، به‌منظور استفاده عملی از مدل، چند روش برای برآورد پارامترها ارائه و به‌منظور ارزیابی عملکرد این روش‌ها شبیه‌سازی انجام شده‌است. در پایان، برای نشان دادن کاربرد مدل، به تحلیل داده‌ها پرداخته شده است.

واژه‌های کلیدی: پیش‌بینی، مدل تعدیل‌یافته دوخطی گسسته‌مقدار، تقریب نقطه زینی، روش ویتل.

کد موضوع‌بندی ریاضی (۲۰۲۰): 62P10، 62M10، 60G25.

۱ مقدمه

در سال‌های اخیر، مدل‌بندی سری‌های زمانی شمارشی به دلیل کاربردهای وسیع آن مورد توجه قرار گرفته و توسعه‌ی چشمگیری یافته است. این نوع مدل‌بندی در بسیاری از حوزه‌های علمی، از جمله بیمه، اقتصاد، داروسازی، اپیدمیولوژی، سیستم‌های صف، ارتباطات، هواشناسی، مخابرات و غیره کاربردهای گسترده‌ای دارند و ضرورت دارد که برازش مناسب بر روی این داده‌ها صورت گیرد.

برای مدل‌سازی سری‌های زمانی شمارشی، بکارگیری عملگرهای رقیق‌ساز به‌عنوان روشی مؤثر شناخته شده‌اند. عملگرهای رقیق‌ساز مختلفی توسط ویب (۲۰۰۸) و اسکات و همکاران (۲۰۱۵) معرفی شده‌اند. علاوه بر این، ژانگ و همکاران (۲۰۱۰) به منظور مدل‌سازی داده‌های سری زمانی گسسته با مقادیر صحیح با تابع خودهمبستگی منفی، عملگری جدید با عنوان "عملگر رقیق‌ساز علامت" معرفی کردند.

با اینکه در حوزه مدل‌سازی سری‌های زمانی شمارشی پیشرفت‌های زیادی حاصل شده، ولی هنوز مدل‌های غیرخطی در این زمینه کمتر مورد توجه و بررسی قرار گرفته‌اند. در یکی از تلاش‌های برجسته در این حوزه، دوقان و همکاران (۲۰۰۶) مدل‌های دوخطی را با

[1] سخنران، bamdadirana@gmail.com

استفاده از عملگر رقیق‌ساز گسترش دادند. یک رویکرد جایگزین برای این مدل‌ها توسط بامدادی و همکاران (۲۰۲۴) ارائه شده است. آن‌ها مدل دوقان و همکاران (۲۰۰۶) را اصلاح کرده و مدل دوخطی را معرفی کردند که انعطاف‌پذیری بیشتری برای ورودی‌های جدید ارائه می‌دهد. این مدل از این مزیت برخوردار است که افراد جدید می‌توانند به‌طور مستقل رفتار کنند. در این مقاله، مدل فوق با مرتبه ۲ برای فرایندهایی که تابع خودهمبستگی جزیی آن در تاخیر ۲ به‌طور معنی‌داری از صفر اختلاف دارد بررسی می‌شود.

۲ معرفی مدل و برخی ویژگی‌های آن

فرآیند $\{X_t\}_{t\in\mathbb{Z}}$، را بر اساس عملگر رقیق‌ساز سری توانی تعمیم‌یافته علامت به‌صورت زیر در نظر بگیرید

$$X_t = a \circledast X_{t-۲} + (a \circledast X_{t-۲})(b \circledast \varepsilon_{t-۱}) + \varepsilon_t, \tag{۱.۲}$$

که در آن $a, b \in [-۱, ۱]$ و $\{\varepsilon_t\}$ دنباله‌ای از متغیرهای تصادفی گسسته i.i.d. با میانگین μ_ε و واریانس $\sigma^۲_\varepsilon$ هستند که از $\{X_s\}_{s<t}$ مستقل می‌باشند. همچنین، $\{W_i\}$ و $\{\tilde{W}_i\}$ که در عملگرهای $a\circledast$ و $b\circledast$ به‌کار رفته‌اند، دارای توزیع سری توانی با میانگین‌های $|a|$ و $|b|$ و واریانس‌های α و β هستند. این مدل به عنوان مدل تعدیل‌یافته دوخطی گسسته‌مقدار مبتنی بر عملگر رقیق‌ساز از مرتبه دو یا به اختصار MINBL$(۲, ۰, ۲, ۱)$ شناخته می‌شود.

گزاره ۱.۲. اگر $(|a| + |ab|\mu_{|\varepsilon|})^۲ + |ab|^۲\sigma^۲_\varepsilon + |a|^۲\beta\mu_{|\varepsilon|} < ۱$، آنگاه فرآیند $\{X_t\}_{t\in\mathbb{Z}}$ در (۱.۲) ایستا می باشد.

گزاره ۲.۲. فرض کنید $\{X_t\}_{t\in\mathbb{Z}}$ فرآیند $MINBL(۲, ۰, ۲, ۱)$ معرفی شده در (۱.۲) باشد که شرایط گزاره ۱.۲ برقرار باشد. در این‌صورت، امیدریاضی و تابع خودکواریانس فرآیند به‌صورت زیر می‌باشد

$$\mathrm{E}(X_t) = \frac{\mu_\varepsilon}{۱ - (a + ab\mu_\varepsilon)}, \tag{۲.۲}$$

و

$$\gamma_X(۱) = \frac{ab\sigma^۲_\varepsilon \mathrm{E}(X_t)}{۱ - (a + ab\mu_\varepsilon)}, \gamma_X(۲k) = (a + ab\mu_\varepsilon)^k \gamma_X(۰), \gamma_X(۲k + ۱) = (a + ab\mu_\varepsilon)^k \gamma_X(۱), k \geq ۱.$$

گزاره ۳.۲. میانگین شرطی فرآیند $\{X_t\}$ برای $k-$گام جلوتر به‌صورت زیر است

$$\begin{aligned}
\mathrm{E}(X_{t+۱}|t) &= (a + ab\varepsilon_t)X_{t-۱} + \mu_\varepsilon, \\
\mathrm{E}(X_{t+۲}|t) &= (a + ab\mu_\varepsilon)X_t + \mu_\varepsilon, \\
\mathrm{E}(X_{t+k}|t) &= (a + ab\mu_\varepsilon)^{\frac{k}{۲}} X_t + \mu_\varepsilon \sum_{i=۱}^{\frac{k}{۲}} (a + ab\mu_\varepsilon)^{i-۱}, \quad k = ۲h,
\end{aligned} \tag{۳.۲}$$

$$\mathrm{E}(X_{t+k}|t) = (a + ab\mu_\varepsilon)^{[\frac{k}{۲}]} \mathrm{E}(X_{t+۱}|t) + \mu_\varepsilon \sum_{i=۱}^{[\frac{k}{۲}]} (a + ab\mu_\varepsilon)^{i-۱}, \quad k = ۲h + ۱. \tag{۴.۲}$$

گزاره ۴.۲. واریانس شرطی یک گام به جلو فرآیند $\{X_t\}$ به‌صورت زیر می‌باشد

$$Var(X_{t+۱}|t) = a^۲\beta|\varepsilon_t|X^۲_{t-۱} + (\alpha + ۲\alpha ab\varepsilon_t + \alpha b^۲\varepsilon^۲_t + \alpha\beta|\varepsilon_t|)|X_{t-۱}| + \sigma^۲_\varepsilon.$$

۳ برآوردیابی و شبیه‌سازی

در این بخش، به برآورد پارامترهای مدل MINBL(۲, ۰, ۲, ۱) با استفاده از چهار روش مختلف می‌پردازیم. برای ارزیابی کارایی این روش‌ها، از داده‌های شبیه‌سازی‌شده استفاده شده و نتایج حاصل از هر روش با یکدیگر مقایسه می‌شوند. در روند برآورد، توزیع $\{\varepsilon_t\}$ پواسن با پارامتر μ_ε در نظر گرفته شده است.

۱.۳ روش یول‌واکر

با به‌کارگیری گزاره ۲.۲، امید ریاضی و تابع خودکواریانس مدل تحت فرض توزیع پواسن خطاها به‌صورت زیر بیان می‌شوند

$$\mathrm{E}(X_t) = \frac{\mu_\varepsilon}{1-(a+ab\mu_\varepsilon)}, \gamma_X(1) = \frac{ab\mu_\varepsilon \mathrm{E}(X_t)}{1-(a+ab\mu_\varepsilon)},$$
$$\gamma_X(2k) = (a+ab\mu_\varepsilon)^k \gamma_X(0), \gamma_X(2k+1) = (a+ab\mu_\varepsilon)^k \gamma_X(1).$$

فرض کنید $A = a + ab\mu_\varepsilon$ و $B = ab\mu_\varepsilon$. با برابر قرار دادن گشتاورهای مرتبه اول و توابع خودکواریانس فرآیند با میانگین نمونه‌ای و تابع خودکواریانس نمونه‌ای برآورد پارامترها به‌صورت زیر به دست می‌آیند

$$\hat{a}_{YW} = \hat{A} - \hat{B}, \ \hat{b}_{YW} = \frac{\hat{B}}{\hat{a}_{YW}\hat{\mu}_{YW}}, \ \hat{\mu}_{YW} = \bar{X}(1-\hat{A}),$$

که در آن $\hat{A} = \frac{\gamma_X(2)}{\gamma_X(0)}$ و $\hat{B} = \frac{\gamma_X(1)(1-\hat{A})}{\bar{X}}$.

۲.۳ روش کمترین مربعات شرطی

برآوردگرهای کمترین مربعات شرطی پارامترهای $\Theta = (a, b, \mu_\varepsilon)$، از طریق مینیمم‌سازی تابع هدف زیر به دست می‌آیند

$$Q(\Theta) = \sum_{t=3}^{n} (X_t - \mathrm{E}(X_t|t-1))^2 = \sum_{t=3}^{n} (X_t - (a+ab\varepsilon_{t-1})X_{t-2} - \mu_\varepsilon)^2,$$

که در آن $\varepsilon_t = X_t - aX_{t-2} - abX_{t-2}\varepsilon_{t-1}$. مینیمم کردن تابع $Q(\Theta)$ با استفاده از الگوریتم نیوتن–رافسون انجام می‌شود.

۳.۳ روش ویتل

در این بخش، روشی برای برآورد پارامترها در حوزه فرکانس با استفاده از معیار ویتل معرفی می‌شود. در بسیاری از موارد، محاسبه تابع چگالی طیفی، به‌ویژه برای فرآیندهای پیچیده‌تر، ساده‌تر و سریع‌تر از محاسبه تابع احتمال دقیق است. برآوردگرهای ویتل برای پارامترهای $\Theta = (a, b, \mu_\varepsilon)$، با مینیمم‌سازی تابع زیر به دست می‌آیند

$$\hat{l}_n(\Theta) = \frac{1}{n} \sum_{k=1}^{[\frac{n}{2}]} (\log f_X(\omega_k) + \frac{I_n(\omega_k)}{f_X(\omega_k)}),$$

که در آن

$$f_X(\omega) = \frac{1}{2\pi} \frac{(1-(a+ab\mu_\varepsilon))(\gamma_X(0)(1+(a+ab\mu_\varepsilon)) + 2\gamma_X(1)\cos\omega)}{1+(a+ab\mu_\varepsilon)^2 - 2(a+ab\mu_\varepsilon)\cos 2\omega}$$

تابع چگالی طیفی فرآیند MINBL(۲, ۰, ۲, ۱) در فرکانس فوریه $\omega_k = \frac{2\pi k}{n}$ و $I_n(.)$ دوره‌نگار فرآیند می‌باشند.

۴.۳ روش درستنمایی ماکزیمم نقطه زینی

روش درستنمایی ماکزیمم به دلیل پیچیدگی‌های مرتبط با عملگر رقیق‌ساز در محاسبه تابع درستنمایی

$$L_n(\Theta) = \sum_{t=1}^{n} \log f_{X_t|t-1}(x_t), \tag{۱.۳}$$

چالش‌برانگیز است. در ادامه با استفاده از تقریب نقطه زینی به تقریب تابع درستنمایی می‌پردازیم. لگاریتم تابع مولد تجمعی شرطی X_t، را به‌صورت $K_t(u) = \log \mathrm{E}(e^{uX_t}|t-1)$ در نظر بگیرید. تقریب نقطه زینی برای تابع چگالی شرطی (۱.۳) به‌صورت زیر می‌باشد

$$\tilde{f}_{X_t|t}(x_t) = (2\pi K_t''(\tilde{u}_t))^{-\frac{1}{2}} \exp\left\{K_t(\tilde{u}_t) - \tilde{u}_t x_t\right\}, \tag{۲.۳}$$

به‌طوری‌که مقدار u که در معادله نقطه زینی $K_t'(u) = x_t$ صدق می‌کند، با $\tilde{u}_t$ نشان داده می‌شود و K_t' و K_t'' به ترتیب مشتقات اول و دوم K_t نسبت به u هستند. محاسبه K_t مدل (۱.۲) در (۲.۳) به راحتی امکان‌پذیر نیست و با بکارگیری بسط تیلور مرتبه اول و دوم $K_t(u)$ در $u = 0$ و اندکی محاسبات می‌توان نشان داد تقریب نقطه زینی تابع چگالی شرطی به‌صورت زیر می‌باشد

$$\tilde{f}_{X_t|t-1}(X_t) = (2\pi\sigma_t^2(\Theta))^{-\frac{1}{2}} \exp\left\{\frac{-(x_t - \mu_t(\Theta))^2}{2\sigma_t^2(\Theta)}\right\}.$$

و لذا براورد درستنمایی ماکزیمم نقطه‌زینی پارامترها از طریق ماکزیمم کردن تابع زیر بدست می‌آید

$$\tilde{L}_n(\Theta) := \sum_{t=1}^{n} \log \tilde{f}_{X_t|t-1}(x_t) = -\sum_{t=1}^{n} \frac{1}{2} \log(2\pi\sigma_t^2(\Theta)) - \sum_{t=1}^{n} \frac{(x_t - \mu_t(\Theta))^2}{2\sigma_t^2(\Theta)}.$$

که در آن μ_t و $\sigma_t^2(\Theta)$ به ترتیب میانگین شرطی و واریانس شرطی X_t هستند.

۵.۳ شبیه‌سازی

در این بخش، با استفاده از شبیه‌سازی به بررسی عملکرد چهار روش براوردیابی می‌پردازیم. مطالعات شبیه‌سازی برای پارامترهای مختلف ارائه شده‌است. نتایج شبیه‌سازی بر اساس معیارهای Bias و MSE در جدول ۱ ارائه شده است. همان‌طور که در جدول ۱ مشاهده می‌شود، هر چهار روش برآورد از نظر اریبی و میانگین مربعات خطا عملکرد رضایت‌بخشی دارند و برآوردهای ویتل بهترین عملکرد را دارند.

جدول ۱: اریبی و میانگین مربعات خطا برآوردگرهای YW، W، CLS و SPML.

n		YW			W			CLS			SPML		
		$\hat{a}$	$\hat{b}$	$\hat{\mu}$	$\hat{a}$	$\hat{b}$	$\hat{\mu}$	$\hat{a}$	$\hat{b}$	$\hat{\mu}$	$\hat{a}$	$\hat{b}$	$\hat{\mu}$
		$(a,b,\mu)=(-۰/۱, ۰/۹, ۷)$											
۱۰۰	Bias	۰/۳۱۱۳۹۵	۰/۷۶۰۷۱۶	−۰/۲۹۱۸۹	−۰/۰۰۰۹۹۹۹	۰/۰۰۰۶۹۴	$۸/۹۲۱۳۴۲e-۰۵$	۰/۰۰۰۶۱۶۴	۰/۰۰۰۲۳۵۶	−۰/۰۰۰۳۱۴۸	۰/۰۰۰۱۰۴	۰/۰۰۱۳۳۷	$-۱/۲۹۰۵۱۲e-۰۷$
	MSE	۰/۱۲۳۹۳۶	۰/۵۹۷۸۱۴	۰/۱۱۲۵۸۹	۰/۰۰۰۹۹۹	$۴/۸۲۶۲۰۹e-۰۶$	$۷/۹۵۹۰۳۴e-۰۸$	۰/۰۰۰۳۷۹	۰/۰۰۰۱۰۸	$۹/۹۱۶۱۳۳e-۰۵$	$۱/۲۴۰۱۵۸e-۰۷$	$۹/۳۰۸۹۰۹e-۰۶$	$۹/۰۵۵۲۷۶e-۱۴$
۳۰۰	Bias	۰/۲۶۵۱۰۷	۰/۷۲۹۶۸۸	−۰/۳۰۰۹۲۷	−۰/۰۰۰۹۹۹۴	۰/۰۰۰۰۶۶۹	$۸/۵۹۱۲۳۲e-۰۵$	$۱/۶۵۴۶۹۵e-۰۵$	۰/۰۰۰۰۷۳۰	$-۲/۷۴۴۶۶۹e-۰۸$	$-۴/۹۲۷۴۴۵e-۰۵$	−۰/۰۰۰۹۰۷	$-۱/۱۵۲۴۶۶e-۰۷$
	MSE	۰/۰۹۱۳۵۷	۰/۵۶۰۹۷۷	۰/۱۱۲۶۹۳	۰/۰۰۰۹۹۸	$۴/۴۷۸۴۱۳e-۰۶$	$۷/۳۸۰۹۲۶e-۰۸$	$۱/۳۸۶۵۵۸e-۰۸$	$۲/۱۴۳۰۶۴e-۰۶$	$۱/۵۳۸۹۴۸e-۱۴$	$۲/۱۰۳۲۲۶e-۰۸$	$۷/۶۵۴۵۶۳e-۰۶$	$۹/۵۸۰۱۴۴e-۱۴$
۵۰۰	Bias	۰/۲۱۴۶۲۴	۰/۶۲۳۹۶۰	−۰/۳۵۸۲۸۵	−۰/۰۰۰۹۸۹۶	۰/۰۰۰۰۶۵۹	$۸/۴۶۲۰۹۳e-۰۵$	$۱/۰۶۸۱۶۷e-۰۵$	۰/۰۰۰۰۲۷۳	$۳/۶۱۷۰۴۶e-۰۹$	$۱/۷۹۹۶۲۹e-۰۵$	$۵/۶۵۳۸۹۵e-۰۵$	$۲/۸۷۸۰۹۴e-۰۹$
	MSE	۰/۰۷۸۰۹۶	۰/۴۳۸۳۰۲	۰/۱۷۵۶۹۵	$۴/۳۴۵۶۰۷e-۰۶$	$۷/۱۶۰۷۰۱e-۰۸$	$۷/۱۶۰۷۰۱e-۱۰$	$۷/۹۷۱۹۵۴e-۰۷$	$۱/۷۷۰۱۸۹e-۰۷$	$۳/۰۷۹۱۰۳e-۱۶$	$۳/۹۴۳۰۳۱e-۰۹$	$۳/۷۳۸۴۱۴e-۰۸$	$۱/۴۷۶۳۹۹e-۱۶$
		$(a,b,\mu)=(۰/۲, -۰/۵, ۸)$											
۱۰۰	Bias	−۰/۲۰۹۷۷۵	−۰/۱۴۲۶۴۱	−۰/۹۴۷۵۴۴	−۰/۰۰۷۷۰۵	۰/۰۱۹۶۴۳	۰/۰۰۰۴۶۴	−۰/۰۰۲۷۳۶	−۰/۰۰۰۶۸۰۱	$۳/۴۷۶۷۶۶e-۰۵$	−۰/۰۱۷۰۹۵	−۰/۰۲۵۹۴۴	−۰/۰۰۰۲۰۴
	MSE	۰/۱۰۲۰۵۶	۰/۰۵۷۴۶۷	۰/۸۶۹۸۸۴	۰/۰۰۰۳۸۵۸	۰/۰۰۰۵۹۳	$۲/۱۵۴۴۳۰e-۰۶$	$۶/۰۹۸۶۷۳e-۰۵$	۰/۰۰۰۷۳۴	$۱/۲۱۰۱۶۱e-۰۸$	۰/۰۰۰۲۲۳۰	۰/۰۰۵۰۵۶	$۳/۱۲۸۴۳۳e-۰۷$
۳۰۰	Bias	−۰/۱۵۷۹۲۳	−۰/۱۵۵۱۷۳	−۰/۹۴۶۵۶۰	−۰/۰۰۷۴۴۳	۰/۰۱۹۴۹۳	۰/۰۰۰۴۴۵	۰/۰۰۱۰۰۶	۰/۰۰۰۲۱۱۴	$۱/۰۲۴۰۷۷e-۰۷$	−۰/۰۰۰۹۲۵	−۰/۰۰۲۵۴۶	$-۵/۸۲۰۴۸۶e-۰۷$
	MSE	۰/۰۷۵۳۱۰	۰/۰۳۲۳۹۱	۰/۰۹۴۴۷۳۴	۰/۰۰۰۳۸۰۰	۰/۰۰۰۵۵۴	$۱/۸۸۱۳۴۲e-۰۶$	$۶/۸۶۱۹۶۹e-۰۶$	$۵/۲۲۴۶۷۰e-۰۵$	$۹/۶۵۸۹۷۲e-۱۴$	$۶/۷۳۳۶۴۷e-۰۶$	$۵/۲۴۱۰۰۸e-۰۵$	$۱/۹۳۵۰۹۴e-۱۲$
۵۰۰	Bias	−۰/۱۴۰۱۱۹	−۰/۱۴۸۷۵۸	−۰/۹۳۲۴۱۵	−۰/۰۰۷۳۶۶	۰/۰۱۹۲۵۴	۰/۰۰۰۴۴۰	$۱/۸۲۶۸۲۱e-۰۵$	$۹/۶۰۵۱۳۴e-۰۶$	$۵/۶۲۸۶۶۷e-۱۰$	$۴/۶۲۹۳۱۲e-۰۵$	$۶/۷۵۷۴۰۲e-۰۵$	$۲/۷۳۶۵۴۹e-۰۸$
	MSE	۰/۰۴۰۷۲۹	۰/۰۳۳۵۵۸	۰/۰۹۰۲۸۹۴	۰/۰۰۰۳۷۰۷	$۱/۸۳۷۳۰۸e-۰۶$	$۱/۸۳۷۳۰۸e-۰۶$	$۲/۶۶۹۳۶۲e-۰۸$	$۹/۴۳۳۶۶۳e-۰۸$	$۱/۵۲۳۸۸۰e-۱۷$	$۱/۲۷۵۳۳۶e-۰۷$	$۴/۷۹۱۲۶۴e-۰۷$	$۱/۷۷۵۸۷۵e-۱۴$
		$(a,b,\mu)=(-۰/۲, -۰/۷, ۴)$											
۱۰۰	Bias	۰/۱۸۹۰۰۰	−۰/۱۰۲۲۴۷	۰/۶۷۳۰۴۱	−۰/۰۰۸۱۲۸	−۰/۰۰۳۹۹۷	۰/۰۰۰۷۰۱	۰/۰۰۱۰۷۶	۰/۰۰۰۵۹۳	۰/۰۰۲۵۷۰	$-۵/۰۵۰۱۹۸e-۰۵$	۰/۰۰۰۲۹۰	$۱/۴۷۲۸۵۰e-۰۵$
	MSE	۰/۰۸۴۱۲۱	۰/۰۳۲۶۲۹	۰/۵۲۱۹۷۷	۰/۰۰۰۱۲۰۷	$۸/۸۵۱۸۷۰e-۰۶$	$۸/۸۵۱۸۷۰e-۰۶$	۰/۰۰۰۵۳۶	۰/۰۰۰۰۹۹۶	۰/۰۰۰۲۷۶	$۷/۵۶۱۳۸۹e-۰۶$	۰/۰۰۰۱۰۰	$۲/۴۳۴۸۴۹e-۰۸$
۳۰۰	Bias	۰/۱۰۵۸۷۹	۰/۰۲۰۰۴۵	۰/۶۵۱۹۱۸	−۰/۰۰۰۲۳۲	−۰/۰۰۰۱۸۷	$۳/۲۷۰۲۸۵e-۰۵$	۰/۰۰۰۱۹۲	−۰/۰۰۰۷۰۴	۰/۰۰۰۰۹۹۶	$-۲/۱۳۷۰۳۷e-۰۵$	$۳/۶۰۵۱۹۰e-۰۵$	$۸/۳۶۱۲۶۶e-۰۸$
	MSE	۰/۰۳۸۲۸۹	۰/۰۲۸۶۴۲	۰/۴۷۳۷۱۳	۰/۰۰۰۱۸۳	$۵/۰۷۱۲۶۰e-۰۵$	$۱/۵۸۶۸۳۲e-۰۶$	$۶/۰۹۵۳۹۷e-۰۶$	$۶۸۱۴۰۲۸e-۰۵$	$۹/۹۳۱۵۴۵e-۰۵$	$۱/۳۵۹۰۲۸e-۰۶$	$۱/۹۵۶۸۸۲e-۰۵$	$۱/۹۹۷۳۰۳e-۱۱$
۵۰۰	Bias	۰/۰۸۸۲۳۷	۰/۰۳۳۸۷۴	۰/۶۲۰۹۳۳	−۰/۰۰۰۵۹۲	−۰/۰۰۰۲۸۴	$-۵/۰۴۶۴۴۰e-۰۵$	۰/۰۰۰۲۰۹	−۰/۰۰۰۶۸۵	۰/۰۰۰۷۱۶	$-۲/۵۹۴۴۵۸e-۰۵$	$۸/۰۴۹۴۴۷e-۰۵$	$۱/۵۳۲۶۵۸e-۰۷$
	MSE	۰/۰۲۸۷۰۸	۰/۰۲۸۹۲۲	۰/۰۴۱۶۳۲۹	$۳/۵۰۸۷۱۶e-۰۵$	$۸/۰۸۶۰۳۲e-۰۶$	$۲/۵۴۶۶۵۶e-۰۷$	$۴/۲۶۱۴۵۲e-۰۶$	$۴/۶۸۰۴۱۶e-۰۵$	$۵/۱۴۰۱۰۶e-۰۵$	$۳/۶۱۷۸۹۴e-۰۷$	$۴/۷۵۶۸۰۰e-۰۶$	$۹/۳۲۵۵۶۱e-۱۳$
		$(a,b,\mu)=(۰/۳, ۰/۴, ۷)$											
۱۰۰	Bias	۰/۰۸۲۷۵۰	−۰/۰۸۲۶۴۷	۰/۱۲۷۴۶۳	−۰/۰۰۰۷۰۹	−۰/۰۰۱۱۶۶	$-۶/۵۴۶۱۵۴e-۰۶$	−۰/۰۵۶۶۴۲	۰/۰۱۸۶۰۵	−۰/۰۲۸۳۰۰	−۰/۴۵۴۱۷۳	−۰/۱۴۳۹۲۹	−۰/۰۲۱۳۶۲
	MSE	۰/۰۱۲۸۳۵	۰/۰۱۸۳۴۱	۰/۰۲۵۰۲۶	$۵/۰۲۷۷۳۹e-۰۶$	$۱/۳۵۹۶۰۱e-۰۵$	$۴/۲۸۵۲۱۴e-۱۰$	۰/۰۰۰۶۹۳۱	۰/۰۰۰۸۱۵	۰/۰۰۰۳۰۱۱	۰/۰۲۱۹۴۵۱	۰/۰۲۱۲۲۶	۰/۰۰۴۸۱۱
۳۰۰	Bias	۰/۰۱۸۷۲۰	−۰/۰۰۸۵۱۴	۰/۱۰۴۷۴۶	−۰/۰۰۰۴۶۸	۰/۰۰۰۰۷۶۹	$-۴/۳۷۷۹۲۷e-۰۶$	−۰/۰۰۲۶۵۰۲	۰/۰۰۰۲۵۰۲	−۰/۰۱۱۲۴۱	−۰/۴۱۲۱۸۴	−۰/۱۳۳۹۰۱	−۰/۰۰۵۸۱۰
	MSE	۰/۰۰۱۷۱۹	۰/۰۰۰۸۲۱	۰/۰۱۴۶۷۴	$۲/۱۹۵۶۶۰e-۰۶$	$۵/۹۱۵۶۴۷e-۰۶$	$۱/۹۱۶۶۲۵e-۱۰$	۰/۰۰۱۷۵۷	۰/۰۰۰۳۷۶	۰/۰۰۲۵۴۵	۰/۰۱۷۳۰۳۵	۰/۰۱۸۱۹۷	۰/۰۰۲۹۰۳
۵۰۰	Bias	۰/۰۱۸۳۷۴	−۰/۰۱۲۸۵۵	۰/۰۸۶۵۲۲	−۰/۰۰۰۲۹۲	−۰/۰۰۰۴۸۰	$-۲/۷۴۴۱۶۸e-۰۶$	−۰/۰۲۶۹۷۱	۰/۰۱۷۲۳۰	−۰/۰۲۴۵۵	−۰/۳۸۹۵۴۰	−۰/۱۲۷۵۰۰	۰/۰۰۶۴۴۵
	MSE	۰/۰۰۱۳۵۱	۰/۰۰۰۷۰۹	۰/۰۱۳۰۱۲	$۸/۵۷۵۵۲۹e-۰۷$	$۲/۳۰۸۲۵۲e-۰۶$	$۷/۵۳۰۴۶۱e-۱۱$	۰/۰۰۱۲۲۸	۰/۰۰۰۳۶۷	۰/۰۰۱۶۱۸	۰/۱۵۷۷۷۹	۰/۰۱۶۴۴۳	۰/۰۰۳۴۴۰

جدول ۲: برآورد پارامترها، RMSE، MAE ، FMSE، و FMAE داده‌های مربوط به اموال سرقت شده.

مدل	برآورد ویتل	RMSE	MAE	FMSE	FMAE
P-MINBL(۲, ۰, ۲, ۱)	$\hat{a} = ۰/۰۳۱$ $\hat{b} = ۱/۹۷۲$ $\hat{\mu} = ۰/۷۶۵$	۱/۰۵۰۷۷۵	۰/۸۰۳۹۰۱۸	۰/۷۰۰۲۷۴۷	۰/۵۲۰۲۹۰۷
PL-MINBL(۲, ۰, ۲, ۱)	$\hat{a} = ۰/۰۶۰$ $\hat{b} = ۰/۳۷۶$ $\hat{\theta} = ۱/۷۷۶$	۱/۰۴۹۱۸۳	۰/۷۹۹۵۴۶۶	۰/۶۸۹۷۵۱	۰/۵۱۳۱۴۱۴
NG-MINBL(۲, ۰, ۲, ۱)	$\hat{a} = ۰/۰۵۱$ $\hat{b} = ۰/۶۷۲$ $\hat{\mu} = ۰/۷۶۵$	۱/۰۴۹۳۵۱	۰/۸۰۰۵۲۱	۰/۶۹۳۳۱۸۲	۰/۵۱۶۱۹۴۴
G-MINBL(۲, ۰, ۲, ۱)	$\hat{a} = ۰/۰۵۲$ $\hat{b} = ۰/۶۷۲$ $\hat{p} = ۰/۵۷۳$	۱/۰۵۰۵۶۹	۰/۸۰۱۶۲۲۹	۰/۷۰۵۴۰۵۶	۰/۵۳۰۲۱۶۶

۴ تحلیل داده‌های واقعی

در این بخش، به کاربرد مدل MINBL(۲, ۰, ۲, ۱) پرداخته می‌شود. داده‌ها مربوط به تعداد گزارش‌های ماهانه پلیس در مورد خرید و فروش اموال سرقت‌شده در گانِدا [۱]، استرالیا است. این اطلاعات به‌صورت ماهانه از ژانویه ۱۹۹۵ تا دسامبر ۲۰۲۴ جمع‌آوری شده است و داده‌ها از سایت http://www.bocsar.nsw.gov.au، گرفته شده است. در نمودار ۱ مسیر نمونه‌ای، ACF و PACF برای داده‌ها نمایش داده شده است. در ادامه، ایستایی داده‌ها با استفاده از آزمون دیکی–فولر افزایشی تأیید شده است که $p-$مقدارهای آن برابر با ۰/۰۱ بوده است. این نتیجه بیان می‌کند که داده‌ها ایستا هستند. همچنین، آزمون غیرخطی کینان که $p-$مقدارهای ۰/۰۰۰ را نشان می‌دهد، غیرخطی بودن داده‌ها را تأیید می‌کند. از سوی دیگر، داده فوق با شاخص‌های پراکندگی $\hat{I}_X = ۱/۶۵۰$ بیش‌پراکنده می‌باشند. مدل MINBL(2,0,2,1) به‌عنوان مدلی مناسب برای برازش این مجموعه‌داده‌ها معرفی می‌شود، همان‌طور که در نمودار ۱ مشاهده می‌شود. در ادامه، پارامترهای مدل با در نظر گرفتن توزیع‌های مختلف از جمله پواسن، پواسن–لیندلی، هندسی و هندسی اصلاح‌شده برآورد شده‌اند. جدول ۲ شامل برآوردهای ویتل به همراه مقادیر ریشه میانگین مربعات خطا (RMSE)، میانگین قدرمطلق خطا (MAE) ، میانگین مربعات خطای پیش‌بینی (FMSE) و میانگین قدرمطلق خطای پیش‌بینی (FMAE) هستند. برای محاسبه FMSE، و FMAE، از پیش‌بینی درون‌نمونه‌ای ۱۰–گام به جلو استفاده شده است. نتایج جدول ۲ نشان می‌دهند که مدل PL-MINBL مقادیر کمتری از RMSE، MAE ، FMSE و FMAE را به‌دست می‌آورد، که حاکی از عملکرد قوی این مدل برای داده است.

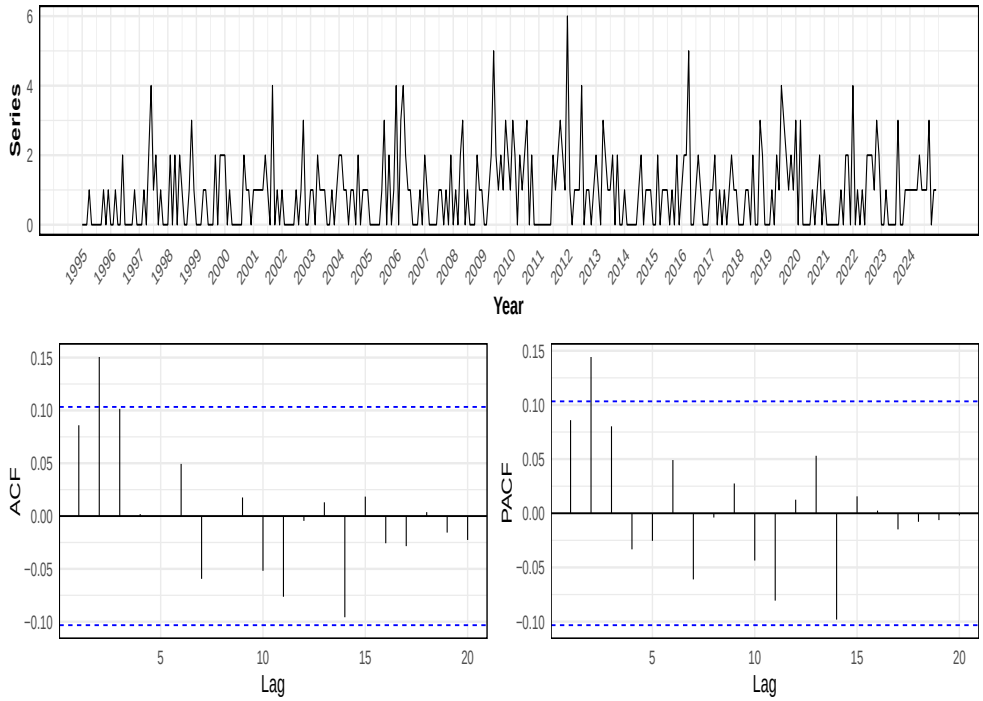
۱.۴ پیش‌بینی

در این بخش به پیش‌بینی مقادیر آینده می‌پردازیم. همان‌طور که قبلا اشاره شد، پیش‌بینی کلاسیک $k-$گام به جلو از طریق امیدریاضی شرطی (۳.۲) و (۴.۲) به‌صورت زیر محاسبه می‌شود

$$\hat{X}_{t+۱} = (a + ab\varepsilon_t)X_{t-۱} + \mu_\varepsilon, \hat{X}_{t+k} = (a + ab\mu_\varepsilon)X_{t+k-۲} + \mu_\varepsilon.$$

که در عمل پارامترهای a، b و μ_ε با برآوردهای مربوطه جایگزین می‌گردد.

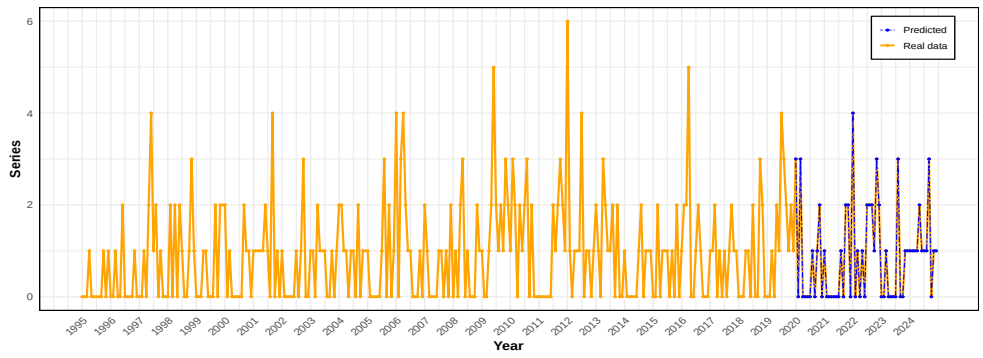
[1] Gunnedah



شکل ۱: مسیر نمونه‌ای، تابع خودهمبستگی و تابع خودهمبستگی جزئی داده‌های مربوط به خرید و فروش اموال سرقت شده.

پیش‌بینی بوت‌استرپ

در این قسمت از تکنیک بوت‌استرپ اصلاح‌شده برای پیش‌بینی استفاده می‌کنیم. این روش به‌طور خاص برای مدل MINBL(۲, ۰, ۲, ۱) با تنظیمات مشخص اعمال شده و در الگوریتم ۱ به تفصیل شرح داده شده است. برای ارزیابی قابلیت پیش‌بینی این روش در طول زمان، از رویکرد بوت‌استرپ برای پیش‌بینی پنج سال پایانی از داده‌های موجود استفاده شد. نتایج نشان می‌دهند که مقادیر پیش‌بینی‌شده شباهت زیادی به سری‌های اصلی دارند و این موضوع تأیید می‌کند که مدل MINBL(۲, ۰, ۲, ۱) برای پیش‌بینی‌های دقیق در مجموعه داده مؤثر است.



شکل ۲: مقادیر واقعی در مقابل مقادیر پیش‌بینی‌شده به روش بوت‌استرپ مربوط به داده‌های اموال سرقت شده در پنج سال گذشته.

الگوریتم ۱

گام اول: پارامترهای a، b و μ_ε را با استفاده از روش ویتل برآورد کنید.

گام دوم: باقیمانده‌های فرآیند $\hat{\varepsilon}_t = X_t - \hat{a}X_{t-۲} - \hat{a}\hat{b}X_{t-۲}\hat{\varepsilon}_{t-۱}$ را برای $t = ۳, \ldots, n$ تعیین کنید.

گام سوم: توزیع تجربی از باقیمانده‌های تعدیل شده $\tilde{\varepsilon}_t = [\hat{\varepsilon}_t]$ زمانی‌که $\hat{\varepsilon}_t > ۰$، در غیر اینصورت $\tilde{\varepsilon}_t = ۰$.

گام چهارم: داده‌های بوت‌استرپ X_t^b برای $b = ۱, \ldots, B$ به‌صورت زیر شبیه‌سازی شده است.

$$X_t^b = \hat{a} \circledast X_{t-۲}^b + (\hat{a} \circledast X_{t-۲}^b)(\hat{b} \circledast \tilde{\varepsilon}_{t-۱}^b) + \tilde{\varepsilon}_t^b,$$

که در آن $\tilde{\varepsilon}_t^b$ برای $t = ۱, ۲, \ldots, n$ در گام سوم شبیه‌سازی شده است.

گام پنجم: بر اساس داده های شبیه‌سازی شده X_t^b پارامترهای $\hat{a}^b$، $\hat{b}^b$ و $\hat{\mu}_\varepsilon^b$ مشابه گام اول محاسبه می‌شود.

گام ششم: پارامترهای a، b و μ_ε برای B تکرار بر اساس میانگین‌های نمونه‌ای به‌صورت زیر براورد شود:

$$\hat{a}^* = \frac{\sum_{b=۱}^{B} \hat{a}^b}{B}, \ \hat{b}^* = \frac{\sum_{b=۱}^{B} \hat{b}^b}{B}, \ \hat{\mu}_\varepsilon^* = \frac{\sum_{b=۱}^{B} \hat{\mu}_\varepsilon^b}{B}.$$

گام هفتم: پیش‌بینی $k-$گام به جلو بر اساس معادلات بازگشتی زیر محاسبه می‌گردد

$$X_{t+k}^b = \hat{a}^* \circledast X_{t+k-۲}^b + (\hat{a}^* \circledast X_{t+k-۲}^b)(\hat{b}^* \circledast \tilde{\varepsilon}_{t+k-۱}^b) + \tilde{\varepsilon}_{t+k}^b, \ k \geq ۱.$$

مراجع

Bamdadi, R., Mohammadpour, M., and Ramezani, S. (2024). A bilinear modeling in counts time series with applications, *Communications in Nonlinear Science and Numerical Simulation*, **139**, 108282.

Doukhan, P., Latour, A., and Oraichi, D. (2006). A Simple Integer-Valued Bilinear Time Series Model, *Advances in Applied Probability*, **38**, 559–578.

Scotto, M. G., Weiß, C. H., and Gouveia, S. (2015). Thinning-based models in the analysis of integer-valued time series: a review, *Statistical Modelling*, **15**, 590–618.

Weiß, C. H. (2008). Thinning operations for modeling time series of counts—a survey, *Advances in Statistical Analysis*, **92**, 319–341.

Zhang, H., Wang, D., and Zhu, D. (2010). Inference for INAR(p) processes with signed generalized power series thinning operator, *Journal of Statistical Planning and Inference*, **140**, 667–683.

A class of integer-valued bilinear time series model with applications

Rana Bamdadi and Mehrnaz Mohammadpour

Department of Statistics, University of Mazandaran

Abstract: This study presents a specific class of integer-valued bilinear model of order 2 for nonlinear time series. These models possess the flexibility to incorporate the new individuals into the model in a way that allows them to behave independently. The modeling methodology applies to processes characterized by overdispersion, relying on a non-linear framework. The main characteristics of the proposed model are established. For computational purposes, some approaches for parameter estimation are carried out. To evaluate the performance of them, simulation experiments are conducted. To illustrate the modeling strategy, a social science cases is analyzed.

Keywords: Forecasting, Integer-valued bilinear process, Saddlepoint approximation, Whittle method.

Mathematics Subject Classification (2020): 62P10، 62M10، 60G25.
