Statistical Sciences Extension
Vol. 2, No. 1, Winter 2017, (P 67-73)

گسترهٔ علوم آماری
سال دوم، شماره اول، پاییز و زمستان 1395، (ص 67-73)

بررسی و کاربرد دو مدل سری‌های زمانی صحیح مقدار

رسول روزگار[*1]، **مرضیه نیک‌آیین**[2]

1. استادیار، گروه آمار، دانشگاه یزد
2. کارشناسی ارشد، گروه آمار، دانشگاه یزد

تاریخ دریافت: 1395/10/03 **تاریخ پذیرش**: 1395/12/17

Review and Application of Two Integer – Valued Time Series Models

R. Roozegar[*1], **M. Nikaeen**[2]

1. Assistant Professor, Department of Statistics, Yazd University
2. M.Sc, Department of Statistics, Yazd University

Received: 2016/12/23 **Accepted:** 2017/03/07

چکیده

در این مقاله ابتدا به معرفی مدل‌های دلاپورت و لودرز-فورمل نوع اول پرداخته می‌شود. سپس این مدل‌ها به‌عنوان توزیع‌های حاشیه‌ای ایستا برای فرایندهای خودبازگشتی صحیح مقدار مرتبه اول در نظر گرفته شده و مورد بحث قرار می‌گیرند. در ادامه فرایندهای خودبازگشتی صحیح مقدار مراتب بالاتر مورد بررسی واقع می‌شوند. ویژگی‌های مختلفی از قبیل رفتار رگرسیونی، زمان معکوس‌پذیری و... برای مدل‌های خودبازگشتی صحیح مقدار مرتبه اول با توزیع‌های حاشیه‌ای ذکر شده مورد بررسی قرارگرفته و مطالعات شبیه‌سازی برای مطالعه مسیرهای نمونه‌ای این مدل‌ها صورت گرفته است. در نهایت نیز پارامترهای مدل، برآورد شده، به کمک یک سری داده واقعی این دو مدل مقایسه می‌شوند.

واژگان کلیدی

توزیع دلاپورت، توزیع لودرز-فورمل نوع اول، فرایندهای خودبازگشتی صحیح مقدار مرتبه اول.

Abstract

This paper first introduces Delaporte and first type Luders Formel models. Then these models are considered as stationary marginal distributions for first-order integer – valued autoregressive processes. After that, the higher order integer - valued autoregressive processes are considered. Various features of these models such as regressive behavior, time reversible and etc. are studied. Also Simulation studies for study of sample paths in these models were done. Finally, the model parameters were estimated and the two models are compared via a real data set.

Keywords

Delaporte Distribution, Luders Formel Distribution, First Order Integer – Value Autoregressive Processes.

*** نویسنده مسئول**: رسول روزگار

Corresponding Author: rroozegar@yazd.ac.ir

مقدمه

در مدل‌بندی داده‌های آماری گاهی با سری‌های زمانی شامل اعداد صحیح برخورد صحیح برخورد می‌کنیم که حفظ ماهیت صحیح بودن داده‌ها مهم است. بسیاری از متغیرهای سودمند (نامنفی) صحیح - مقدار هستند.

از جمله مواردی که در سری‌های زمانی گسسته مقدار نامنفی که با آن مواجه می‌شویم تعداد وقایع یا تعداد اشیا است. تعداد وقایع می‌تواند بروز تصادفات، قربانیان جرم و جنایت، شناسایی خطاها، بیماری‌ها، انتقال پیام‌ها و غیره باشد.

این مدل‌ها در مطالعات مربوط به برنامه‌ریزی شهری، مدل‌سازی ریسک، رشد حشرات و غیره کاربرد دارند.

سری‌های زمانی صحیح مقدار در بسیاری از مفاهیم معمولا به‌عنوان شمارشگر پیشامدها در فواصل زمانی مشخص مورد مطالعه قرار می‌گیرند. ثابت شده که این مدل‌ها به‌ویژه در فرایند پواسون بسیار مفیدند اما در بسیاری از موارد که سری‌های زمانی هم‌پراکنش نیست (میانگین و واریانس برابر نیستند) کارایی لازم را ندارند. داده‌های شمارا در بیولوژیک کاربردهای متعددی دارند و اغلب می‌تواند توسط یک توزیع پواسون، که در آن میانگین و واریانس یکسان هستند مدل‌سازی شود.

با این حال، در شرایطی که یک همبستگی مثبت در وقوع حوادث وجود دارد، تغییر مشاهده شده به طور قابل توجهی بیشتر از میانگین است و تعمیمی از مدل پواسون مناسب‌تر است. برای مدل‌سازی داده شمارا با بیش‌پراکندگی یک مدل مشهور و در دسترس توزیع دوجمله‌ای منفی است. در این مقاله دو توزیع گسسته که تعمیم‌هایی از توزیع‌های گسسته رایج هستند و برای داده‌های شمارا مناسب هستند را معرفی می‌کنیم و ویژگی‌هایی از این توزیع‌ها را در قالب قضیه بیان می‌کنیم.

در ادامه مدل‌های خودبازگشتی صحیح-مقدار با این توزیع‌های گسسته به‌عنوان توزیع‌های حاشیه‌ای فرایند را بیان کرده و توزیع دنباله نوآورهای این مدل‌ها را با استفاده از خواص تابع مولد احتمال و قضایای مختلف، برای تولید فرایند به دست آورده‌ایم. در آخر نیز با استفاده از شبیه‌سازی به بررسی ویژگی‌های این مدل‌ها پرداخته و در نهایت الگوی مورد مطالعه برای داده‌های واقعی خاصی که از نوع سری‌های زمانی شمارشی هستند به کار برده می‌شود.

توزیع لودرز-فورمل نوع اول

توزیع لودرز-فورمل نوع اول از پیچش توزیع‌های پواسون نوع یک، دو و سه به دست می‌آید. این توزیع از متغیرهایی به ترتیب با پارامترهای

$$\lambda_1 = \frac{\beta}{(1+\eta)}, \lambda_2 = \frac{\beta\eta}{2(1+\eta)^2},$$
$$\lambda_3 = \frac{\beta\eta^3}{3(1+\eta)^3}, \lambda_k = \frac{\beta\eta k}{\mathrm{k}(1+\eta)^{\mathrm{k}}}, \ldots.$$

به دست می‌آید. درنتیجه تابع مولد احتمال برای این پیچش به‌صورت زیر است:

$$G(s) = exp\left[\frac{\beta}{\eta}\sum_{k\geq 1}\left(\frac{\eta}{1+\eta}\right)^k \frac{s^k-1}{k}\right]$$
$$= [1+\eta-\eta s]^{\frac{-\beta}{\eta}}.$$

تابع جرم احتمال توزیع لودرز-فورمل نوع اول به صورت زیر است:

$$P(x) = P(X=x)$$
$$= \frac{\prod_{r=0}^{x-1}(\beta+r\eta)}{\mathrm{x}!}(1+\eta)^{-\left(\frac{\beta}{\eta}+\mathrm{x}\right)}$$
$$, \beta > 0, \eta \in (0,1).$$

توزیع لودرز-فورمل اول را با $LF(\beta,\eta)$، نشان می‌دهیم.

قضیه1: توزیع لودرز-فورمل نوع اول،گسسته خود تجزیه‌پذیر است.

فرایند خودبازگشتی صحیح مقدار مرتبه اول با توزیع حاشیه‌ای لودرز-فورمل نوع اول ($\boldsymbol{LFINAR(1)}$)

تعریف1. (فرایند خودبازگشتی صحیح مقدار مرتبه اول $((INAR(1))$ دنباله $\{X_n|n\geq 1\}$ از متغیرهای تصادفی صحیح مقدار نامنفی را در نظر بگیرید،

فرایند X_n یک فرایند خودبازگشتی صحیح مقدار مرتبه اول گفته می‌شود اگر برای هر $n\geq 1$، در رابطهٔ بازگشتی زیر صدق کند.

$$X_n = \gamma * X_{n-1} + \varepsilon_n. \qquad (1)$$

به سادگی می‌توان نشان داد تابع مولد احتمال $\gamma * X_{n-1}$، $G_X(1-\gamma+\gamma s)$ است. $\gamma \in (0,1)$ ضریب فرایند، $\gamma * X_{n-1}$ و ε_n مستقل‌اند و $\{\varepsilon_n\}$ دنباله نوآور است.

تحت فرض ایستای اکید، عبارت 1 را می‌توان با جملاتی از تابع مولد احتمال به صورت زیر نوشت:

$$G_X(s) = G_X(1-\gamma+\gamma s)G_\epsilon(s),$$
$$-1 \le s \le 1, \qquad \gamma \in (0,1).$$

با توجه به اینکه

$A_X(s) = G_X(1-s) = E[(1-s)^X]$ است، رابطه با جملاتی از تابع مولد احتمال متناوب به صورت زیر بازنویسی می‌شود:

$$A_X(s) = A_X(\gamma s)A_\epsilon(s), \forall\gamma \in (0,1).$$

فرایند خودبازگشتی (1) را در نظر بگیرید. هدف به دست آوردن توزیع $\{\varepsilon_n\}$ است، به طوری که $\{X_n\}$ دارای توزیع لودرز-فورمل نوع اول به‌عنوان توزیع حاشیه‌ای فرایند ایستا باشد.

$A_X(.)$ و $A_\varepsilon(.)$ را به ترتیب تابع مولد احتمال متناوب متغیر تصادفی X_n و ε_n در نظر بگیرید. تحت فرض ایستایی فرایند از به دست می‌آوریم که:

$$\begin{aligned} A_\epsilon(s) &= \frac{A_X(s)}{A_X(\rho s)} \\ &= \left[\frac{1+\rho\eta s}{1+\eta s}\right]^{\frac{\beta}{\eta}} \\ &= \left[\rho + (1-\rho)\frac{1}{1+\eta s}\right]^{\frac{\beta}{\eta}}. \end{aligned}$$

بنابراین فرایند خودبازگشتی که دنباله نوآورهای $\{\varepsilon_n\}$ آن دارای تابع مولد احتمال فوق باشد را فرایند لودرز-فورمل نوع اول گوییم. ضریب خودهمبستگی مرتبه اول فرایند و $\{\varepsilon_n\}$ نوآور است.

گشتاورهای فرایندهای LFINAR(1)

به راحتی می‌توان با توجه به روابط زیر:

$$G_\epsilon(s) = A_\epsilon(1-s),$$
$$G'(1) = \sum_r rP(X=r) = E(X),$$
$$G''(1) = \sum_r r(r-1)P(X=r) = E\big(X(X-1)\big).$$

امید، واریانس، ضریب چولگی، ضریب کشیدگی متغیر تصادفی ε_n را به ترتیب به دست آورد:

$$E(\epsilon_n) = \beta(1-\rho),$$
$$Var(\epsilon_n) = \beta(1-\rho)[\eta+\rho\eta+1],$$
$$\tau = -\frac{(2\eta^2\rho^2+3\rho\eta+2\rho\eta^2+3\eta+2\eta^2+1)^2}{\beta(\eta+\rho\eta+1)^2(\rho^2\eta+\rho-\eta-1)}.$$
$$\xi = \frac{(-6\eta^2\rho^2-6\eta^2+3\beta\eta\rho^2-6\eta\rho-3\beta\eta-6\eta+3\beta\rho-3\beta-1)}{\beta(\rho^2\eta+\rho-\eta-1)}.$$

تذکر 1. اگر X_0 دارای توزیع دلخواه باشد، فرایند خودبازگشتی فرایند مارکوف ایستای مجانبی با توزیع حاشیه‌ای لودرز-فورمل نوع اول است.

برهان. فرایند (1) را می‌توان به صورت زیر نوشت:

$$\begin{aligned} X_n &= \rho * X_{n-1} + \epsilon_n \\ &= \rho^n * X_0 + \sum_{k=0}^{n-1}\rho^k * \epsilon_{n-k}. \end{aligned}$$

عبارت فوق را برحسب تابع مولد احتمال متناوب می‌نویسیم:

$$\begin{aligned} A_{X_n}(s) &= A_{X_0}(\rho^n s)\prod_{k=0}^{n-1}A_{\epsilon_n}(\rho^k) \\ &= A_{X_0}(\rho^n s)\prod_{k=0}^{n-1}\left[\frac{1+\rho^{k+1}\eta s}{1+\rho^k\eta s}\right]^{\frac{\beta}{\eta}} \\ &\to \left[\frac{1}{1+\eta s}\right]^{\frac{\beta}{\eta}} \qquad n\to\infty. \end{aligned}$$

که این نشان‌دهنده این است که اگر X_0 دارای توزیع دلخواه باشد، فرایند مارکوف ایستای مجانبی با توزیع حاشیه‌ای لودرز-فورمل نوع اول است.
چون تابع مولد احتمال یکتاست با استفاده از تبدیل معکوس توزیع توأم (X_{n-1}, X_n) را می‌توان به دست آورد.

رفتار رگرسیونی و ویژگی‌های مسیر نمونه‌ای:

معادله رگرسیونی پیشرو برای فرایند مطرح شده در بخش (1) به صورت زیر است

$$\begin{aligned} &E(X_n|X_{n-1} = x) \\ &= E(\rho * X_{n-1} + \epsilon_n|X_{n-1} = x) \\ &= \rho x + E(\epsilon_n|X_{n-1} = x) \\ &= \rho x + (1-\rho)\beta. \end{aligned}$$

و واریانس شرطی به صورت زیر به دست می‌آید

$$\begin{aligned} &Var(X_n|X_{n-1} = x) \\ &= Var(\rho * X_{n-1} + \epsilon_n|X_{n-1} = x) \\ &= \rho(1-\rho)x. \\ &\quad +\beta(1-\rho)[\eta + \rho\eta + 1]. \end{aligned}$$

در مسیر پسرو از تابع مولد احتمال توأم X_n و X_{n+1} شروع کرده و نتایج به دست آمده توسط (لاورانس[1]، 1987) را مرحله به مرحله انجام می‌دهیم. با فرض ایستایی تابع مولد احتمال توأم X_n و X_{n+1} به صورت زیر حاصل می‌شود:

$$G_{X_n,X_{n+1}}(s_1, s_2) = \frac{G_X(s_2)G_X(s_1(1-\rho+\rho s_2))}{G_X(\rho s_2)}.$$

با مشتق‌گیری از عبارت فوق نسبت به s_1 و قرار دادن $s_1 = 1$ و $s_2 = s$ داریم:

$$\begin{aligned} G_{X_n,X_{n+1}}(s_1, s_2) &= E({s_1}^{X_n} {s_2}^{\rho * X_n + \epsilon_{n+1}}) \\ &= G_{\epsilon_{n+1}}(s_2) G_{X_n}(s_1) \\ &\quad G_{X_n}(1-\rho+\rho s_2) \\ &= G_{\epsilon_{n+1}}(s_2) E({s_1}^{X_{n+1}}) \end{aligned}$$

$$\begin{aligned} . \quad &E((1-\rho+\rho s_2)^{X_{n+1}}) \\ &= G_{\epsilon_{n+1}}(s_2)\, E(s_1(1-\rho+\rho s_2)^{X_n}) \\ &= G_{\epsilon_{n+1}}(s_2)\, G_{X_n}(s_1(1-\rho+\rho s_2)). \\ &= \left[\frac{1+\rho\eta(1-s_2)}{1+\eta(1-s_2)}\right]^{\frac{\beta}{\eta}} \end{aligned}$$

$$.\left[\frac{1}{1+\eta(1-[s_1(1-\rho+\rho s_2)])}\right]^{\frac{\beta}{\eta}}.$$

$$\begin{aligned} E(X_n\, s^{X_{n+1}}) &= \frac{G_X(s) G'_X(1-\rho+\rho s)}{G_X(\rho s)} \\ &= G_\epsilon(s) G'_X(1-\rho+\rho s) \\ &= \beta(1-\rho+\rho s) \\ &\quad \left[\frac{1+\rho\eta(1-s)}{1+\eta(1-s)}\right]^{\frac{\beta}{\eta}} \\ &. (1+\eta)^{\frac{\beta}{\eta}-1}. \qquad (1) \end{aligned}$$

با توجه به اینکه

$$\begin{aligned} &E(X_n s^{X_{n+1}}) \\ &= \sum_{m=0}^{\infty} s^m\, E(X_n|X_{n+1} = m) P(X_{n+1} = m). \end{aligned}$$

می‌توان $E(X_n|X_{n+1} = m)$ را از ضریب s^m در عبارت (2) به دست آورد.

درنتیجه چون

$$E(X_n|X_{n+1}) \neq E(X_n|X_{n-1}),$$

فرایند زمان معکوس‌پذیر نیست.

توزیع دلاپورت

متغیر تصادفی X روی Z_+ دارای توزیع دلاپورت است و با $Del(\lambda, p, k)$ نشان می‌دهیم اگر دارای تابع مولد احتمال زیر باشد:

$$\begin{aligned} G_X(s) &= e^{-\lambda(1-s)} \left(\frac{1}{1+\frac{q}{p}(1-s)}\right)^k, \\ &-1 \leq s \leq 1, 0 < p < 1, k > 0, \lambda > 0. \end{aligned}$$

تابع جرم احتمال متناظر با عبارت فوق به صورت زیر است:

---

1. Lawrance

$$f(x) = P(X = x)$$
$$= \frac{e^{-\lambda}\lambda^x p^k}{(k-1)!\,x!}\sum_{j=0}^{x}\binom{x}{j}\left(\frac{1-p}{\lambda}\right)^j$$
$$, x = 0,1,2, \ldots.$$

وقتی که $\lambda > 0$، $K \geq 1$، $0 < p < 1$ و $q = 1 - p$ است. متغیر تصادفی دلاپورت X را می‌توان به صورت $X = U + V$ نوشت که U و V متغیرهای تصادفی دوجمله‌ای منفی و پواسون مستقل هستند.

با مشتق گرفتن از تابع مولد احتمال توزیع دلاپورت به سادگی می‌توان نشان داد که

$$E(X) = \lambda + \frac{kq}{p},$$
$$Var(X) = \lambda + \frac{kq}{p^2},$$

ضریب چولگی

$$\tau = \frac{[\lambda p^3 + kq(1+q)]^2}{[p^2\lambda + kq]^3},$$

و ضریب کشیدگی

$$\xi = \frac{3p^4\lambda^2 + p^4\lambda + kq[3kq + 4q + q^2 + 6\lambda p^2 + 1]}{[p^2\lambda + kq]^2}.$$

فرایند INAR(1) با توزیع حاشیه‌ای دلاپورت

مدل خودبازگشتی (1) را در نظر بگیرید. هدف به دست آوردن توزیع $\{\varepsilon_n\}$ است، به طوری که $\{X_n\}$ دارای توزیع دلاپورت به‌عنوان توزیع حاشیه‌ای فرایند ایستا باشد. قضیه زیر نتایج یافته شده در این زمینه را ارائه می‌دهد.

قضیه 2. فرض کنید $\{X_n\}$ یک فرایند خودبازگشتی ایستا با ساختار داده شده در (1) باشد، که دنباله $\{\varepsilon_n\}$ متغیرهای تصادفی مستقل و هم‌توزیع هستند. یک شرط لازم و کافی که $\{X_n\}$ یک فرایند مارکوف ایستای اکید با توزیع‌های حاشیه‌ای دلاپورت $Del(\lambda, p, k)$ باشد این است که تابع مولد احتمال $\{\varepsilon_n\}$ به صورت زیر باشد:

$$A_\epsilon(s) = e^{-\lambda(1-\rho)s}\left\{\rho + (1-\rho)\frac{1}{1+\frac{q}{p}s}\right\}^k,$$

به شرط اینکه ε_n و X_0 هم‌توزیع باشند.

برهان. شرط لازم را به صورت زیر اثبات می‌کنیم

$$A_\epsilon(s) = \frac{A_X(s)}{A_X(\rho s)},$$

تابع مولد احتمال متناوب توزیع دلاپورت را جایگذاری می‌کنیم

$$A_\epsilon(s) = \left\{\frac{e^{-\lambda s}\left(1 + \frac{q}{p}s\right)^{-k}}{e^{-\lambda(\rho s)}\left(1 + \frac{q}{p}(\rho s)\right)^{-k}}\right\}$$
$$= e^{-\lambda(1-\rho)s}\left\{\frac{\rho\left(1 + \frac{q}{p}s\right) - \rho + 1}{1 + \frac{q}{p}s}\right\}^k$$
$$= \left\{\rho + (1-\rho)\frac{1}{1 + \frac{q}{p}s}\right\}^k.$$

این فرمول بیانگر آن است که $\{\varepsilon_n\}$ دارای ساختاری پیچشی $\varepsilon_n = X_1 + X_2$ است، که $\lambda \sim p(\lambda(1-\rho))$ و X_2 پیچش k تایی از توزیع هندسی با پارامتر p است که

$$X_2 = \begin{cases} 0 & \text{با احتمال } \rho \\ \mathrm{Geo}(p) & \text{با احتمال } 1-\rho \end{cases}.$$

شرط کافی را با روش استقراء ریاضی به صورت زیر اثبات می‌کنیم، فرض می‌کنیم X_n دارای توزیع $Del(\lambda, p, k)$ باشد، در نتیجه خواهیم داشت:

$$A_{X_n}(s) = A_{\epsilon_n}(s)A_{X_{n-2}}(\rho s)$$
$$= e^{-\lambda(1-\rho)s}\left\{\rho + (1-\rho)\frac{1}{1 + \frac{q}{p}s}\right\}^k$$
$$.\left\{e^{-\lambda(\rho s)}\left(1 + \frac{q}{p}(\rho s)\right)^{-k}\right\} = e^{-\lambda s}\left(1 + \frac{q}{p}s\right)^{-k}.$$

که تابع مولد احتمال توزیع دلاپورت است. بقیه را به‌طور مشابه می‌توان نتیجه گرفت. فرایند خودبازگشتی صحیح-مقدار مرتبه اول با توزیع حاشیه‌ای دلاپورت را با $DelAR(1)$ نشان می‌دهیم.

قضیه 3. اگر X_0 توزیع دلخواه داشته باشد، آنگاه فرایند معرفی شده در بخش (1.2) ، مارکوف ایستای مجانبی با توزیع حاشیه‌ای دلاپورت است.

رفتار رگرسیونی فرایند DelAR(1)

امید ریاضی و واریانس شرطی مشابه حالت قبل به صورت زیر به دست می‌آید:

$$\begin{aligned}
&E(X_n|X_{n-1} = x)\\
&= E(\rho * X_{n-1} + \epsilon_n|X_{n-1} = x)\\
&= \rho x + E(\epsilon_n|X_{n-1} = x)\\
&= \rho x + (1-\rho)\left(\lambda + k\frac{q}{p}\right),
\end{aligned}$$

$$\begin{aligned}
&Var(X_n|X_{n-1} = x)\\
&= Var(\rho * X_{n-1} + \epsilon_n|X_{n-1} = x)\\
&= \rho(1-\rho)x + (1-\rho^2)\left(\lambda + k\frac{q}{p^2}\right)\\
&+ \rho(1-\rho)\left(\lambda + k\frac{q}{p^2}\right),
\end{aligned}$$

از این رو در اینجا با محاسبه امید ریاضی شرطی پسرو، به این نتیجه خواهیم رسید که

$$E(X_n|X_{n+1}) \neq E(X_n|X_{n-1}),$$

بنابراین فرایند زمان معکوس‌پذیر نیست.

توسعه به فرایندهای مرتبه بالاتر

فرض می‌کنیم یک فرایند خودبازگشتی صحیح مقدار مرتبـه k دارای ساختار احتمال زیر باشد،

$$X_n = \rho_i * X_{n-i} + \epsilon_n, \quad \text{با احتمال } p_i$$

که $\rho_i > 0$، $p_i \leq 1$، $i = 1,2,\dots p$ و $\sum_{i=1}^{k} p_i = 1$.

مدل فوق بر حسب جملات تابع مولد احتمال متناوب به صورت زیر نوشته می‌شود:

$$A_{X_n}(s) = A_{\epsilon_n}(s)\left[\sum_{i=1}^{k} p_i A_{X_{n-1}}(\rho_i s)\right]$$

تحت فرض ایستای اکید عبارت بالا را به صورت زیر می‌نویسیم

$$A_X(s) = A_\epsilon(s)\left[\sum_{i=1}^{k} p_i A_X(\rho_i s)\right].$$

بنابراین

$$A_\epsilon(s)=\frac{A_X(s)}{\sum_{i=1}^{k} p_i A_X(\rho_i s)}$$

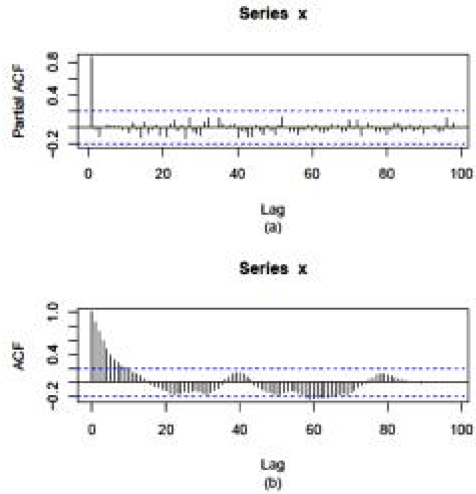
برای توزیع‌های حاشیه‌ای دلخواه، دنباله نوآور فرایند دارای تابع مولد احتمال متناوب فوق است.

مطالعات شبیه‌سازی و به کارگیری داده‌های واقعی برای مقایسه مدل‌ها

در این بخش آنچه را که به‌صورت تئوری ارائه شده، با استفاده از شبیه‌سازی بررسی می‌کنیم. با توجه به مدل‌های ارائه شده به بررسی و تجزیه و تحلیل این مدل‌ها می‌پردازیم. هم‌چنین با استفاده از داده‌های واقعی این دو مدل را مقایسه می‌کنیم.

ابتدا فرایند خودبازگشتی صحیح مقدار مرتبه اول لودرز-فورمل نوع اول و دلاپورت را با نرم‌افزار R شبیه‌سازی کرده و همچنین نمودار سری‌های زمانی، خودهمبستگی و خودهمبستگی جزئی این دو مدل را رسم کرده‌ایم. در نمودار 1و2 شاهد خودهمبستگی جزئی جملات داخل سری هستیم که جز لگ اول که تجاوز از محدوده مشخصه را داریم با سرعت خوبی به صفر همگرا است. نمودار خودهمبستگی نیز به صورت یک تنزل نمایی یا موج سینوسی به سمت صفر میل می‌کند.

با توجه به عدم وجود روند در نمودارهای سری‌های زمانی 3و4 ایستایی فرایند مشهود است. در ادامه نتایج حاصل از تحلیل تعداد واردات اداره کل گمرک استان یزد از فروردین‌ماه 1394 تا شهریورماه 1394 را مورد مطالعه قرار داده و برای پیش‌بینی ماه‌های آینده از مدل‌های دلاپورت و لودرز-فورمل استفاده کرده‌ایم.

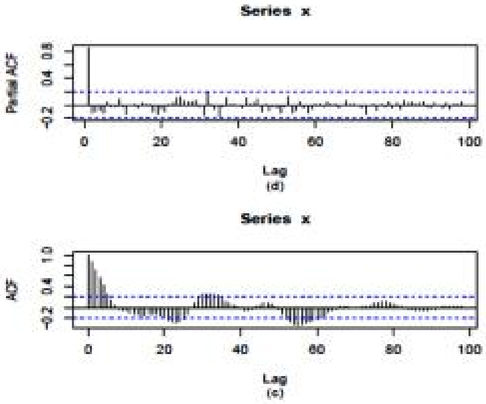


نمودار 1. نمودار خودهمبستگی و خودهمبستگی جزئی مدل $LFINAR(1)$ با پارامترهای $(\eta,\beta) = (0.5,2)$ و $\rho = 0.3$

ابتدا پارامترهای این دو مدل را با دو روش برآورد کردیم و سپس با مقایسه متوسط مربع خطاها در جدول 3 نتیجه می‌گیریم مدل $DelAR(1)$ مناسب‌تر است.

جدول 1. برآورد پارامترهای مدل $DelAR(1)$ بر اساس داده‌های واقعی

پارامترها	CLS	YW
λ	1/351	1/372
p	0/704	0/701
k	11/733	11/683
ρ	0/422	0/432



نمودار2. نمودار خودهمبستگی و خودهمبستگی جزئی مدل $DelAR(1)$ با پارامترهای $(\lambda,p,k) = (3, 0.8, 12)$ و $\rho = 0.125$

جدول 2. برآورد پارامترهای مدل $\boldsymbol{LFINAR(1)}$ بر اساس داده‌های واقعی

پارامترها	CLS	YW
β	6.961	6.012
η	0.536	0.501
ρ	0.404	0.466

جدول 3. مقایسه خطای مربع دو مدل

	CLS	YW
$DelAR(1)$	2389	2403
$LFINAR(1)$	2496	2512

منابع

[1] Al-Osh, M.A. and Alzaid, A.A. (1987). First-order integer-valued autoregressive (INAR (1)) process. Journal of Time Series Analysis, **8**(3), 261-275..

[2] Delaporte, P. (1959). Quelques probl´emes de statistique math´ematique pos´es par`I assurance automobile et le bonus non sinistre. Bulletin Trimestriel de l'Institut desActuaires Franc¸ais, **227**, 87-102.

[3] Lawrance, A.J. (1978). Some Autoregressive models for point processes. In Colloquia Mathematica Societatis Janos Bolyai, **24**, 257-275.

[4] Luders, R. (1934). Die Statistik der seltenen Ereignisse. Biometrika, **26**, 108-128.

[5] Jose, K. K., and Thomas, M.M. (2011). Generalized Laplacian distributions and autoregressive processes. Communications in Statistics-Theory and Methods, **40** (23), 4263-4277.