

آزمون استقلال سری زمانی مبتنی بر معیار واگرایی توان

عماد اشتری نژاد^۱، یدالله واقعی^۱، غلامرضا محتشمی برزادران^۲، حمیدرضا نیلی ثانی^۱، هادی علیزاده

نوقابی^۱

گروه آمار، دانشگاه بیرجند

گروه آمار، دانشگاه فردوسی مشهد

چکیده: در تحلیل سری‌های زمانی، بهتر است قبل از هرگونه تحلیلی، وابستگی داده‌ها مورد بررسی قرار گیرد. زیرا اگر داده‌ها از یکدیگر مستقل باشند، برازش مدل‌های متداول سری زمانی که مبتنی بر اصولی چون مانایی و وابستگی داده‌های زمانی است، اعتباری نخواهد داشت. ملاک واگرایی توان در سال‌های اخیر، اغلب برای آزمون نیکویی برازش مورد استفاده قرار گرفته است. در این مقاله با تشکیل بردارهای مجاور m تایی و استفاده از نمادسازی جایگشت، آزمونی مبتنی بر ملاک واگرایی توان برای بررسی استقلال سری‌های زمانی معرفی می‌شود که به پارامتر کنترل کننده نوع آزمون بستگی دارند. پس از بدست آوردن توزیع حدی آماره آزمون، با استفاده از یک مطالعه شبیه‌سازی، خطای نوع اول و توان آزمون برای برخی از حالت‌های خاص پارامتر کنترل کننده نوع آزمون بدست می‌آید. به وسیله نتایج شبیه‌سازی نشان داده می‌شود که برای حجم نمونه نسبتاً بزرگ به ازای تمامی مقادیر پارامتر کنترل کننده نوع آزمون خطای نوع اول آزمون به سطح اسمی آن نزدیک می‌شود و آزمون‌های خی-دو اصلاح شده، نسبت درستی‌ی اصلاح شده و فریمن-توکی بیشترین توان را دارند.

واژه‌های کلیدی: آزمون استقلال، سری زمانی، واگرایی توان، متغیرهای تصادفی m -وابسته.

^۱ آدرس الکترونیک مسئول مقاله: داله واقعی، ywaghei@birjand.ac.ir

^۲ کد موضوع بندی ریاضی (۲۰۱۰): 62G10, 68Q30, 37M10

۱ مقدمه

سری زمانی مشاهده‌هایی در طول زمان هستند. قبل از هر چیز بایست از وابسته یا مستقل بودن این داده‌ها مطمئن شد، زیرا در غیر این صورت مدل‌های سری زمانی فاقد اعتبار خواهند بود. آزمون‌های استقلال سری زمانی می‌تواند به عنوان یک آزمون تشخیصی در مدل‌سازی مورد استفاده قرار گیرد. کاربرد مهم‌تر این آزمون‌ها در تشخیص استقلال جملات خطای مدل‌های سری‌زمانی است. در این آزمون‌ها از هریک از متغیرهای تصادفی $\{Z_t, t = 1, 2, 3, \dots, n\}$ تنها یک مشاهده در دسترس خواهد بود، از این رو برآورد تابع چگالی Z_t ‌ها امکان پذیر نخواهد بود. در این راستا برای انجام آزمون می‌توان بردارهای m -بعدی را از این متغیرهای تصادفی به صورت

$$\mathbf{Z}_t(m) = (Z_t, \dots, Z_{t+m-1})^T \quad ; t \in \{1, \dots, n - m + 1\},$$

تشکیل داد و بر اساس این بردارها فرضیه مستقل و هم توزیع بودن متغیرها را بررسی نمود. طیف گسترده‌ای از آزمون‌های ناپارامتری به منظور انجام آزمون استقلال وجود دارد که از این جمله می‌توان به آزمون‌های مبتنی بر رتبه^۱ و گردش^۲ همانند بارتلز (۱۹۸۲)، آزمون‌های مبتنی بر اندازه‌های واگرایی همانند قودی و همکاران (۲۰۰۱)، آزمون‌هایی بر مبنای تابع مشخصه همانند هانگ (۲۰۰۰) و آزمون‌هایی بر مبنای همبستگی انتگرال همانند آزمون^۳ BDS (بروک و همکاران، ۱۹۹۶) اشاره کرد. ماتیلا-گارسیا و مارین (۲۰۰۸) توانستند آزمون‌های استفاده از نسبت درست‌نمایی و مبتنی بر آنتروپی جایگشت ارائه دهند که تحت تبدیلات یکنوا پایا است. سنسوی و همکاران (۲۰۱۵) کارایی بازار سهام سنتی و اسلامی را به وسیله این آزمون مورد بررسی قرار دادند. کانواس و همکاران (۲۰۱۳) سه آزمون، ماتیلا-گارسیا و مارین (۲۰۰۸)، آمیگو و همکاران (۲۰۰۷) و کانواس و گیلان (۲۰۰۹) را مقایسه کرده و به این نتیجه رسیدند که آزمون مبتنی بر آنتروپی جایگشت توان بالاتری از دو آزمون دیگر دارد. هرچند آزمون مبتنی بر آنتروپی جایگشت مورد توجه پژوهشگران بوده، اما در برابر آزمون‌هایی همانند BDS (بروک و همکاران، ۱۹۹۶) توان بالایی ندارد. این موضوع در پژوهش‌های کانواس و همکاران (۲۰۱۳) و ماتیلا-گارسیا و همکاران (۲۰۱۴) به وضوح دیده شده است. اما برتری آزمون مبتنی بر آنتروپی جایگشت بر آزمون BDS، داشتن برخی ویژگی‌های مطلوب همانند پایایی است. اشتری‌نژاد و همکاران (۲۰۱۸) به این نتیجه رسیدند که وجود همپوشانی در بردارهای $\mathbf{Z}_t(m)$ ، دقت آزمون

¹Rank

²RUN

³Brock, Dechert, Scheinkman (BDS)

مبتنی بر آنتروپی جایگشت را کاهش می‌دهد. در این راستا نسخه‌های اصلاح شده این آزمون را ارائه دادند که نسبت به آزمون ماتیللا-گارسیا و مارین (۲۰۰۸) از دقت و توان بالاتری داشت. بنابراین با توسعه نسخه اصلاح شده آزمون مبتنی بر آنتروپی جایگشت، می‌توان علاوه بر در اختیار داشتن شرایط مطلوب این آزمون، از مزایای دیگری همانند توان بالاتر نیز بهره برد. کرسی و رید (۱۹۸۴) برای اولین بار معیار واگرایی توان را معرفی نموده و از آن به منظور انجام آزمون فرضیه $p_1 = p_1^0, \dots, p_M = p_M^0$: H_0 در توزیع چندجمله‌ای استفاده کردند. آماره آزمون که بر اساس اختلاف بین برآورد p_i ها و p_i^0 ها بنا نهاده شده، از رابطه

$$T_n^\lambda(\hat{\mathbf{P}}, \mathbf{P}^0) = \frac{\gamma n}{\lambda(\lambda + 1)} \sum_{i=1}^M \hat{p}_i \left(\left(\frac{\hat{p}_i}{p_i^0} \right)^\lambda - 1 \right), \quad (1)$$

بدست می‌آید که λ یک عدد حقیقی معلوم و مخالف ۰ و ۱- بوده است. در این پژوهش بر مبنای معیار واگرایی توان و احتمال رخداد جایگشت‌ها، رده‌ای از آزمون‌های استقلال سری زمانی را ارائه شده است. این آزمون‌ها ویژگی‌های آزمون مبتنی آنتروپی جایگشت را دارند، زیرا بر مبنای احتمال رخداد جایگشت‌ها ارائه شده است. ملاک واگرایی توان رده‌ای از آزمون‌ها را می‌سازد که حالت‌های خاص آن توسط ماتیللا-گارسیا (۲۰۰۷) و ماتیللا-گارسیا و مارین (۲۰۰۸) به ترتیب بر اساس آماره خی-دو و اطلاع کولبک-لیبلر بررسی و با سایر آزمون‌های استقلال سری زمانی مقایسه شده‌اند. به منظور بسط این آزمون‌ها و مقایسه رده آزمون‌های مبتنی بر واگرایی توان، آزمون جدیدی مبتنی بر این معیار معرفی شده است. با در نظر گرفتن وجود هم‌پوشانی در بردارهای $\mathbf{Z}_t(m)$ ، توزیع حدی آماره آزمون بدست آمده و حالت‌های خاص آزمون تشریح شده است. به ازای حجم نمونه متفاوت، برآورد خطای نوع اول، سرعت همگرایی و توان آزمون برای دامنه وسیعی از λ مورد بررسی قرار گرفته است. آزمون معرفی شده برای داده‌های بازده قیمت روزانه طلا به کار رفته است. در انتها نیز به کمک نتایج بدست آمده از شبیه‌سازی و مثال کاربردی در مورد این گروه از آزمون‌ها نتیجه‌گیری شده است.

۲ مفاهیم اساسی آزمون

فرض کنید $\pi_i = (i_1, \dots, i_m)$ یک چندتایی مرتب از اعداد ۰، ۱، ۲، ...، $m-1$ و مجموعه $\Gamma = \{\pi_1, \pi_2, \dots, \pi_m\}$ شامل تمامی نمادهای جایگشت باشد. نمادسازی جایگشت به معنی اعمال تابع $f: \mathbb{R}^m \rightarrow \Gamma$ بر روی بردارهای $\mathbf{Z}_t(m)$ است به گونه‌ای که به هریک از بردارهای $\mathbf{Z}_t(m)$ بر حسب ترتیب (بزرگتر یا کوچکتر) عناصر آن، یک نماد از مجموعه Γ را منتسب می‌کند. این بدان معناست که با

مرتب کردن عناصر $\mathbf{Z}_t(m)$ برقراری رابطه $Z_{t+i_1} \leq Z_{t+i_2} \leq \dots \leq Z_{t+i_m}$ نشانه‌ی انتساب نماد $i_{s-1} < i_s$ آنگاه $Z_{t+i_{s-1}} = Z_{t+i_s}$ در حالت خاص اگر $Z_{t+i_{s-1}} = Z_{t+i_s}$ است. برای روشن شدن موضوع فرض کنید که $m = 3$ و $n = 7$ باشد. بنابراین مجموعه Γ شامل تمامی جایگشت‌های 3 تایی، به صورت

$$\Gamma = \{(0, 1, 2), (0, 2, 1), (1, 0, 2), (1, 2, 0), (2, 0, 1), (2, 1, 0)\},$$

خواهد بود. اگر $\{z_1 = 1, z_2 = 6, z_3 = 6, z_4 = 7, z_5 = 2, z_6 = 8, z_7 = 0\}$ باشد بردارهای 3 تایی $\mathbf{z}_1(3) = (1, 6, 6)$ ، $\mathbf{z}_2(3) = (6, 6, 7)$ ، $\mathbf{z}_3(3) = (6, 7, 2)$ ، $\mathbf{z}_4(3) = (7, 2, 8)$ و $\mathbf{z}_5(3) = (2, 8, 0)$ را تشکیل داده که ترتیب قرارگیری آن‌ها به صورت $(0, 1, 2)$ است. تحت نمادسازی جایگشت، ابتدا بردارهای 3 تایی به صورت نازولی مرتب می‌شوند. سپس ترتیب قرارگیری بردارها که بعد از مرتب سازی از $(0, 1, 2)$ به یکی از اعضای Γ تغییر کرده را به عنوان نماد هر آن‌ها در نظر گرفته می‌شوند. به طور مثال، از آن جا که $z_6 < z_4 < z_5$ مرتب شده‌ی بردار $\mathbf{z}_4(3)$ برابر $(2, 7, 8)$ و ترتیب قرارگیری آن $(1, 0, 2)$ است. بنابراین نتایج نمادسازی به صورت $f(\mathbf{z}_1(3)) = (0, 1, 2)$ ، $f(\mathbf{z}_2(3)) = (0, 1, 2)$ ، $f(\mathbf{z}_3(3)) = (1, 0, 2)$ ، $f(\mathbf{z}_4(3)) = (2, 0, 1)$ و $f(\mathbf{z}_5(3)) = (2, 0, 1)$ خواهد بود. برای نماد π_i متغیرهای برنولی $W_{\pi_i, t}$ را می‌توان به صورت

$$W_{\pi_i, t} = \begin{cases} 1 & f(\mathbf{Z}_t(m)) = \pi_i \\ 0 & f(\mathbf{Z}_t(m)) \neq \pi_i \end{cases}, \quad t = 1, \dots, K \quad i = 1, \dots, m!,$$

تعریف نمود که در آن $K = n - m + 1$ و با فرض هم توزیع بودن Z_t ها متغیرهای $W_{\pi_i, t}$ دارای توزیع برنولی با احتمال موفقیت p_{π_i} هستند. همچنین p_{π_i} ها در رابطه

$$\sum_{i=1}^{m!} p_{\pi_i} = 1,$$

صدق می‌کنند. با استفاده از رابطه $\hat{p}_{\pi_j} = \frac{1}{K} \sum_{t=1}^K W_{\pi_j, t}$ برآورد $\mathbf{P} = (p_{\pi_1}, \dots, p_{\pi_{m!}})^T$ را می‌توان به

صورت

$$\hat{\mathbf{P}} = \left(\frac{1}{K} \sum_{t=1}^K W_{\pi_1, t}, \dots, \frac{1}{K} \sum_{t=1}^K W_{\pi_{m!}, t} \right)^T,$$

محاسبه نمود (اشتری‌نژاد و همکاران، ۲۰۱۸). حال در این پژوهش آزمون فرضیه

$$H_0: \{Z_t, t \in \mathbb{Z}\},$$

مورد بررسی قرار خواهد گرفت.

بردار $\mathbf{P} = (p_{\pi_1}, \dots, p_{\pi_{m!}})^T$ شامل احتمال رخداد هریک از جایگشت‌های $\pi_1, \dots, \pi_{m!}$ است. از این رو، می‌توان به وسیله $T_n^\lambda(\hat{\mathbf{P}}, \mathbf{P}^\circ)$ با مقایسه بین برآورد آن یعنی $\hat{\mathbf{P}}$ و مقدار آن تحت فرضیه H_0 یعنی $\mathbf{P}^\circ = (\frac{1}{m!}, \dots, \frac{1}{m!})^T$ ، فرضیه مستقل و هم توزیع بودن $\{Z_t, t \in \mathbb{Z}\}$ را آزمون نمود. در این شرایط آماره $T_n^\lambda(\hat{\mathbf{P}}, \mathbf{P}^\circ)$ را برای آزمون استقلال سری زمانی به صورت

$$T_n^\lambda(\hat{\mathbf{P}}, \mathbf{P}^\circ) = \frac{2K}{\lambda(\lambda+1)} \sum_{i=1}^{m!} \hat{p}_{\pi_i} ((m! \hat{p}_{\pi_i})^\lambda - 1), \quad (2)$$

بازنویسی می‌شود که در آن $K = n - m + 1$. بنابراین فاصله گرفتن $\hat{\mathbf{P}}$ از \mathbf{P}° نشان از وابستگی $\{Z_t, t \in \mathbb{Z}\}$ داشته و برای مقادیر بزرگ $T_n^\lambda(\hat{\mathbf{P}}, \mathbf{P}^\circ)$ فرضیه H_0 رد خواهد شد.

۳ توزیع حدی آماره آزمون

در این بخش توزیع حدی آماره آزمون بیان و اثبات می‌شود. برای برهان توزیع حدی آماره آزمون تحت فرضیه H_0 ، باید وابستگی $\mathbf{Z}_t(m)$ ها مدنظر قرار گیرد.

توجه کنید که حتی اگر Z_t ها از یکدیگر مستقل باشند (تحت فرضیه H_0)، بردارهای $\mathbf{Z}_t(m)$ به دلیل وجود عناصر مشترک، در برخی موارد وابسته خواهند بود. به عبارت دقیق‌تر بردارهای $\mathbf{Z}_t(m)$ را $(m-1)$ -وابسته گویند (گات، ۲۰۰۶). لازم به ذکر است متغیرهای $f(\mathbf{Z}_t(m))$ که وسیله تبدیل جایگشت بر روی $\mathbf{Z}_t(m)$ بدست آمده‌اند نیز در صورت وابستگی $\mathbf{Z}_t(m)$ ها وابسته خواهند بود به طور مثال، در حالتی خاص، برای $m = 2$ داریم:

$$\begin{aligned} P(f(\mathbf{Z}_t(2)), f(\mathbf{Z}_{t+1}(2))) &= P(Z_t < Z_{t+1}, Z_{t+1} < Z_{t+2}) \\ &\neq P(Z_t < Z_{t+1})P(Z_{t+1} < Z_{t+2}). \end{aligned}$$

قضیه ۱. فرض کنید m ثابت، $\mathbf{A} = \text{diag}(m!, \dots, m!) = m! \mathbf{I}_{m!}$

$$\Sigma^\circ = \frac{1}{(m!)^2} (m! (\mathbf{I}_{m!} + \sum_{l=1}^{m-1} \mathbf{Q}^{(l)} + \sum_{l=1}^{m-1} \mathbf{Q}^{(l)T}) - (2m-1) \mathbf{1}_{m!} \mathbf{1}_{m!}^T),$$

و ν_i ها مقادیر ویژه ماتریس $\mathbf{A}\Sigma^\circ$ باشند. اگر $\{Z_1, \dots, Z_n\}$ مستقل و هم توزیع باشند، آنگاه رابطه

$$T_n^\lambda(\hat{\mathbf{P}}, \mathbf{P}^\circ) = \frac{\Upsilon K}{\lambda(\lambda+1)} \sum_{i=1}^{m!} \hat{p}_{\pi_i} ((m! \hat{p}_{\pi_i})^\lambda - 1) \xrightarrow[n \rightarrow \infty]{d} \sum_{i=1}^r \nu_i Y_i^\lambda. \quad (3)$$

برقرار است که در آن متغیرهای تصادفی Y_1, \dots, Y_r دارای توزیع نرمال استاندارد هستند.

برهان. برای بدست آوردن توزیع مجانبی آماره آزمون (تحت فرضیه H_0)، کفایت توزیع مجانبی بسط تیلور آن بدست آورده شود.

برای این منظور فرض کنید $\phi(x) = \frac{x^{\lambda+1} - x - \lambda(x-1)}{\lambda(\lambda+1)}$ و $\psi(x_1, \dots, x_{m!}) = \sum_{i=1}^{m!} p_{\pi_i}^\circ \phi(\frac{x_i}{p_{\pi_i}^\circ})$

بسط تیلور تابع ψ در $\hat{\mathbf{P}}$ حول \mathbf{P} تا مرتبه دوم از رابطه

$$\begin{aligned} \psi(\hat{p}_{\pi_1}, \dots, \hat{p}_{\pi_{m!}}) &= \psi(p_{\pi_1}, \dots, p_{\pi_{m!}}) \\ &+ \sum_{i=1}^{m!-1} \frac{\partial \psi(p_{\pi_1}, \dots, p_{\pi_{m!}})}{\partial p_{\pi_i}} (\hat{p}_{\pi_i} - p_{\pi_i}) \\ &+ \frac{1}{2} \sum_{i,j=1}^{m!-1} \frac{\partial^2 \psi(p_{\pi_1}, \dots, p_{\pi_{m!}})}{\partial p_{\pi_i} \partial p_{\pi_j}} (\hat{p}_{\pi_i} - p_{\pi_i})(\hat{p}_{\pi_j} - p_{\pi_j}) \\ &+ o(O_p(n^{-1})), \end{aligned} \quad (4)$$

بدست می‌آید که در آن اگر c مقدار ثابت باشد $\frac{O(n)}{n} \xrightarrow[n \rightarrow \infty]{} 0$ ، $\frac{O(n)}{n} \xrightarrow[n \rightarrow \infty]{} c$ و افزودن زیرنویس p

به آن‌ها به معنی همگرایی در احتمال خواهد بود. حال تحت فرضیه H_0 مقدار $\mathbf{P} = (p_{\pi_1}, \dots, p_{\pi_{m!}})$ برابر

$$\mathbf{P}^\circ = (\frac{1}{m!}, \dots, \frac{1}{m!})$$

بوده و بنابراین $\psi(p_{\pi_1}^\circ, \dots, p_{\pi_{m!}}^\circ) = 0$

با توجه به اینکه $\sum_{i=1}^{m!} p_{\pi_i} = 1$ داریم:

$$\begin{aligned} \psi(p_{\pi_1}, \dots, p_{\pi_{m!}}) &= \sum_{i=1}^{m!-1} p_{\pi_i}^\circ \phi(\frac{p_{\pi_i}}{p_{\pi_i}^\circ}) + p_{\pi_{m!}}^\circ \phi(\frac{p_{\pi_{m!}}}{p_{\pi_{m!}}^\circ}) \\ &= \sum_{i=1}^{m!-1} p_{\pi_i}^\circ \phi(\frac{p_{\pi_i}}{p_{\pi_i}^\circ}) + p_{\pi_{m!}}^\circ \phi(\frac{1 - \sum_{i=1}^{m!-1} p_{\pi_i}}{p_{\pi_{m!}}^\circ}). \end{aligned} \quad (5)$$

همچنین تحت فرضیه H_0 ، برای هر i ($i = 1, \dots, m!$) $p_{\pi_i} = p_{\pi_i}^\circ$ و بنابراین $\frac{p_{\pi_i}}{p_{\pi_i}^\circ} = 1$ از این رو

می‌توان نوشت:

$$\begin{aligned} \left(\frac{\partial \psi(p_{\pi_1}, \dots, p_{\pi_{m!}})}{\partial p_{\pi_i}} \right)_{\mathbf{P}=\mathbf{P}^\circ} &= \left(\frac{\partial (\sum_{i=1}^{m!-1} p_{\pi_i}^\circ \phi(\frac{p_{\pi_i}}{p_{\pi_i}^\circ}) + p_{\pi_{m!}}^\circ \phi(\frac{p_{\pi_{m!}}}{p_{\pi_{m!}}^\circ}))}{\partial p_{\pi_i}} \right)_{\mathbf{P}=\mathbf{P}^\circ} \\ &= \left(\phi'(\frac{p_{\pi_i}}{p_{\pi_i}^\circ}) - \phi'(\frac{p_{\pi_{m!}}}{p_{\pi_{m!}}^\circ}) \right)_{\mathbf{P}=\mathbf{P}^\circ} = 0. \end{aligned} \quad (6)$$

با توجه به برابری‌های $\phi'(\frac{p_{\pi_i}}{p_{\pi_m!}}) - \phi'(\frac{p_{\pi_m!}}{p_{\pi_m!}})$ و $\frac{\partial \psi(p_{\pi_1}, \dots, p_{\pi_m!})}{\partial p_{\pi_i}} = 1 - \sum_{i=1}^{m-1} p_{\pi_i}$ داریم:

$$\left(\frac{\partial^2 \psi(p_{\pi_1}, \dots, p_{\pi_m!})}{\partial p_{\pi_i} \partial p_{\pi_j}} \right)_{\mathbf{P}=\mathbf{P}^\circ} = \left(\frac{1}{p_{\pi_m!}^\circ} \phi''\left(\frac{p_{\pi_m!}}{p_{\pi_m!}^\circ}\right) \right)_{\mathbf{P}=\mathbf{P}^\circ} = m!, \quad i \neq j, \quad (۷)$$

$$\left(\frac{\partial^2 \psi(p_{\pi_1}, \dots, p_{\pi_m!})}{\partial p_{\pi_i} \partial p_{\pi_j}} \right)_{\mathbf{P}=\mathbf{P}^\circ} = \left(\frac{1}{p_{\pi_i}^\circ} \phi''\left(\frac{p_{\pi_i}}{p_{\pi_i}^\circ}\right) + \frac{1}{p_{\pi_m!}^\circ} \phi''\left(\frac{p_{\pi_m!}}{p_{\pi_m!}^\circ}\right) \right)_{\mathbf{P}=\mathbf{P}^\circ} = 2m! \quad i = j. \quad (۸)$$

بنابراین جمله اول و دوم رابطه (۴) برابر صفر خواهد شد. با توجه به روابطه (۷) و (۸) نیز جمله سوم رابطه (۴) به صورت

$$\begin{aligned} & \frac{m!}{2} \left[\sum_{i,j=1, i \neq j}^{m-1} \left(\hat{p}_{\pi_i} - \frac{1}{m!} \right) \left(\hat{p}_{\pi_j} - \frac{1}{m!} \right) + 2 \sum_{i=1}^{m-1} \left(\hat{p}_{\pi_i} - \frac{1}{m!} \right)^2 \right] \\ &= \frac{m!}{2} \left[\sum_{i,j=1}^{m-1} \left(\hat{p}_{\pi_i} - \frac{1}{m!} \right) \left(\hat{p}_{\pi_j} - \frac{1}{m!} \right) + \sum_{i=1}^{m-1} \left(\hat{p}_{\pi_i} - \frac{1}{m!} \right)^2 - \left(\hat{p}_{\pi_m} - \frac{1}{m!} \right)^2 \right] \\ &= \frac{m!}{2} \sum_{i=1}^{m-1} \left(\hat{p}_{\pi_i} - \frac{1}{m!} \right)^2, \end{aligned} \quad (۹)$$

محاسبه می‌شود. از این رو، با جایگذاری رابطه (۹) در رابطه (۴) و ضرب $2K$ در دو طرف برابری (۴) خواهیم داشت:

$$T_n^\lambda(\hat{\mathbf{P}}, \mathbf{P}^\circ) = Km! \sum_{i=1}^{m-1} \left(\hat{p}_{\pi_i} - \frac{1}{m!} \right)^2 + 2Ko(O_p(n^{-1})). \quad (۱۰)$$

همانطور که ملاحظه می‌شود تفاوت $T_n^\lambda(\hat{\mathbf{P}}, \mathbf{P}^\circ)$ و $Km! \sum_{i=1}^{m-1} \left(\hat{p}_{\pi_i} - \frac{1}{m!} \right)^2$ عبارت $2Ko(O_p(n^{-1}))$ است. با توجه به $K = n - m + 1$ ، می‌توان نشان داد:

$$2Ko(O_p(n^{-1})) \xrightarrow[n \rightarrow \infty]{p} 0.$$

از این رو، با توجه به قضیه اسلاتسکی (گات، ۲۰۰۶) توزیع حدی $T_n^\lambda(\hat{\mathbf{P}}, \mathbf{P}^\circ)$ و $Km! \sum_{i=1}^{m-1} \left(\hat{p}_{\pi_i} - \frac{1}{m!} \right)^2$ یکسان خواهد بود. در این راستا، برای بدست آوردن توزیع حدی آماره آزمون کافی است توزیع حدی جمله اول سمت راست برابری (۱۰) بدست آورده شود. حال با توجه به السینگر (۲۰۱۰، قضیه ۱) و با در نظر گرفتن توزیع حدی $\mathbf{P}^\circ = \left(\frac{1}{m!}, \dots, \frac{1}{m!} \right)$

$$\sqrt{K}(\hat{\mathbf{P}} - \mathbf{P}^\circ) \xrightarrow[n \rightarrow \infty]{d} N(0, \Sigma^\circ),$$

حاصل خواهد شد. بنابراین اگر $\mathbf{A} = \text{diag}(m!, \dots, m!)$ باشد با استناد به **دیک و گانست (۱۹۸۵)**، فرع (۲.۱) و قضیه نگاشت پیوسته **(گات، ۲۰۰۶)** رابطه

$$K(\hat{\mathbf{P}} - \mathbf{P}^\circ)^T \mathbf{A}(\hat{\mathbf{P}} - \mathbf{P}^\circ) = Km! \sum_{i=1}^{m!} \left(\hat{p}_{\pi_i} - \frac{1}{m!}\right)^2 \xrightarrow[n \rightarrow \infty]{d} \sum_{i=1}^r \lambda_i Z_i^2,$$

برقرار خواهد بود. بنابراین با توجه به یکسان بودن توزیع مجانبی $T_n^\lambda(\hat{\mathbf{P}}, \mathbf{P}^\circ)$ و $Km! \sum_{i=1}^{m!} \left(\hat{p}_{\pi_i} - \frac{1}{m!}\right)^2$ اثبات کامل است.

توزیع حدی $T_n^\lambda(\hat{\mathbf{P}}, \mathbf{P}^\circ)$ که در قضیه ۱ بدست آمد به اختصار مجموع موزون خی-دو نامیده می‌شود. به دلیل پیچیده بودن این توزیع فرم بسته برای چگالی آن وجود ندارد. چندک‌های این توزیع را می‌توان بر مبنای روش عددی ارائه شده توسط **لیو و همکاران (۲۰۰۹)** و با استفاده بسته *CompQuadForm* در نرم‌افزار *R* بدست آورد.

۴ مطالعات شبیه‌سازی

همانطور که اشاره شد، آماره آزمون به دلیل وابستگی به پارامتر دلخواه λ یک خانواده از آزمون‌های مختلف را شامل می‌شود. این موضوع برای نیکویی برازش توسط **پاردو (۲۰۰۶)**، صفحه ۱۱۳) مورد بررسی قرار گرفته است. اگر $p_{\pi_i}^\circ = \frac{1}{m!}$ و $\hat{p}_{\pi_i} = \frac{1}{K} \sum_{t=1}^K W_{\pi_i, t}$ ، حالت خاص $T_n^\lambda(\hat{\mathbf{P}}, \mathbf{P}^\circ)$ به صورت آماره‌های زیر خواهند بود:

$$1- \text{آماره خی-دو اصلاح شده: } T_n^{-2}(\hat{\mathbf{P}}, \mathbf{P}^\circ) = K \sum_{i=1}^{m!} \frac{(p_{\pi_i}^\circ - \hat{p}_{\pi_i})^2}{\hat{p}_{\pi_i}}$$

$$2- \text{نسبت درستنمایی اصلاح شده: } T_n^{-1}(\hat{\mathbf{P}}, \mathbf{P}^\circ) = 2K \sum_{i=1}^{m!} p_{\pi_i}^\circ \log\left(\frac{\hat{p}_{\pi_i}}{p_{\pi_i}^\circ}\right)$$

$$3- \text{فریمین-توکی: } T_n^{-\frac{1}{2}}(\hat{\mathbf{P}}, \mathbf{P}^\circ) = \lambda K \left(1 - \sum_{i=1}^{m!} \sqrt{\hat{p}_{\pi_i} p_{\pi_i}^\circ}\right)$$

$$4- \text{نسبت درستنمایی (ماتیلدا-گارسیا و مارین، ۲۰۰۸): } T_n^0(\hat{\mathbf{P}}, \mathbf{P}^\circ) = 2K \sum_{i=1}^{m!} \hat{p}_{\pi_i} \log\left(\frac{\hat{p}_{\pi_i}}{p_{\pi_i}^\circ}\right)$$

$$5- \text{آماره کرسی-رید: } T_n^{\frac{1}{2}}(\hat{\mathbf{P}}, \mathbf{P}^\circ) = \frac{1}{2} K \left(\sum_{i=1}^{m!} \hat{p}_{\pi_i} \left(\frac{\hat{p}_{\pi_i}}{p_{\pi_i}^\circ}\right)^{\frac{1}{2}} - 1\right)$$

$$6- \text{آماره خی-دو (ماتیلدا-گارسیا، ۲۰۰۷): } T_n^1(\hat{\mathbf{P}}, \mathbf{P}^\circ) = K \sum_{i=1}^{m!} \frac{(\hat{p}_{\pi_i} - p_{\pi_i}^\circ)^2}{p_{\pi_i}^\circ}$$

در ادامه به مقایسه توان و دقت (حفظ سطح اسمی خطای نوع اول) هر یک از حالت‌های خاص پرداخته

و در فرآیند شبیه‌سازی، سطح خطای نوع اول $\alpha = 0.05$ در نظر گرفته می‌شود.

سرعت همگرایی به سطح اسمی خطا

در قضیه ۱ اثبات شد که برای مقدار ثابت m تحت فرضیه H_0 آماره $T_n^\lambda(\hat{\mathbf{P}}, \mathbf{P}^\circ)$ به ازای هر λ دارای توزیع حدی مجموع موزون χ^2 -دو خواهد بود. سوالاتی که مطرح می‌شود این است که آیا مقدار λ سرعت همگرایی به توزیع حدی را تحت تاثیر قرار می‌دهد. این موضوع را می‌توان با مقایسه خطای اسمی و برآورد شده مورد بررسی قرار داد. برای بررسی این موضوع پس از شبیه‌سازی داده‌های مستقل از توزیع نرمال استاندارد با حجم نمونه متفاوت ($n \geq 30$)، برای $m = 3$ به ازای λ های متفاوت مقدار آماره $T_n^\lambda(\hat{\mathbf{P}}, \mathbf{P}^\circ)$ محاسبه شد. بعد از تکرار این محاسبات برای $N = 2000$ بار برآورد خطای نوع اول از رابطه

$$\hat{\alpha}_n^\lambda = \frac{\#\{T_{n,0}^\lambda(\hat{\mathbf{P}}, \mathbf{P}^\circ) > Q_{1-\alpha}\}}{N}, \quad (11)$$

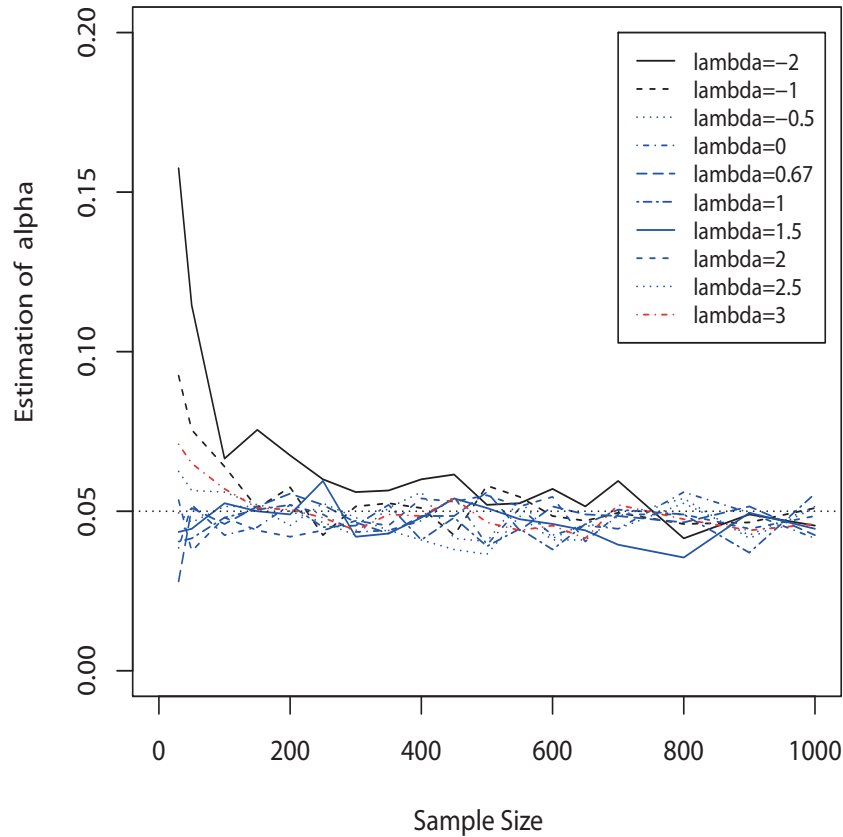
بدست می‌آید که در آن مقدار $T_{n,0}^\lambda(\hat{\mathbf{P}}, \mathbf{P}^\circ)$ مقدار $T_n^\lambda(\hat{\mathbf{P}}, \mathbf{P}^\circ)$ تحت فرضیه H_0 (داده شبیه سازی شده مستقل و هم توزیع) و $Q_{1-\alpha}$ چندک مرتبه $1 - \alpha$ توزیع مجموع موزون χ^2 -دو است. مقادیر بدست آمده برای $\alpha = 0.05$ در شکل ۱ رسم شده است. همان‌طور که ملاحظه می‌شود به ازای $\lambda = -2$ در حجم نمونه کم، $\hat{\alpha}_n^{-2}$ اختلاف زیادی با 0.05 دارد و با افزایش حجم نمونه این اختلاف بسیار کم خواهد شد. برای سایر مقادیر λ نیز شرایط مشابه بوده و افزایش حجم نمونه سبب نزدیکی مقادیر $\hat{\alpha}_n^\lambda$ به 0.05 شده است. این موضوع نشان دهنده سرعت همگرایی متفاوت به ازای مقادیر مختلف λ است.

تذکر ۱. محاسبات بالا برای مقادیر بیش‌تری از λ انجام شد که به دلیل کوتاه‌نویسی از ذکر آنها خودداری می‌شود. نتایج نشان می‌دهد برای طیف گسترده‌تری از مقادیر λ نیز سطح اسمی خطا با افزایش حجم نمونه به مقدار 0.05 نزدیک می‌شود.

میزان دقت آزمون

به منظور بررسی جامع‌تر درباره وابستگی خطای نوع اول به λ ، نمونه‌هایی مستقل با حجم 50 ، 240 و 580 به تفکیک از توزیع‌های $N(0, 1)$ ، $U(0, 1)$ ، χ^2_4 و t_4 شبیه‌سازی نموده و با تکرار $N = 2000$ بار مقدار $\hat{\alpha}_n^\lambda$ از رابطه (۱۱) محاسبه شد که نتایج در جدول ۱ آمده است. به توصیه پژوهشگران از جمله ماتیلدا-گارسیا (۲۰۰۷) و ماتیلدا-گارسیا و مارین (۲۰۰۸) مقدار m باید در نابرابری $n \leq 5m!$ صدق کند. در این پژوهش مقدار m را برای حجم نمونه 50 برابر ۳ و برای 240 و 580 برابر ۴ در نظر گرفته شده است. همان‌طور که

ملاحظه می‌شود برای $n = 580$ به ازای تمامی مقادیر λ برآورد خطای نوع اول به 5% بسیار نزدیک شده است.



شکل ۱: برآورد خطای نوع اول به ازای حجم نمونه متفاوت و برای $m = 3$.

تذکر ۲. میزان دقت آزمون برای مقادیر دیگر m نیز مورد بررسی قرار گرفت. بررسی‌ها نشان داد که برای انتخاب مقدار m باید توان آزمون را نیز در نظر گرفت. به طور مثال در تمامی آزمون‌ها، برای مقدار $m = 2$ ، در هر حجم نمونه، سطح اسمی خطا کاملاً حفظ می‌شود. اما ممکن است به ازای مقدار $m = 3$ با اندکی انحراف از سطح اسمی خطا، توان بسیار بالاتری از $m = 2$ نمایان شود. از این رو، باید با نگاهی واقع‌گرایانه از طرفی سطح اسمی خطا حفظ شده و از طرفی آزمونی با توان بالا انتخاب شود. بنابراین پس از بررسی تاثیر m بر توان آزمون، مقدار بهینه آن تعیین خواهد شد.

توان آزمون

هرچند در جدول ۱ میزان انحراف برآورد خطای نوع اول از مقدار اسمی مورد بررسی قرار گرفت، اما آزمون به عنوان آزمون برتر انتخاب خواهیم شد که علاوه بر میزان نزدیکی خطای آن به ۰/۰۵ توان بالایی را هم داشته باشد. از این رو، در گام بعدی توان آزمون به ازای مقادیر مختلف λ مقایسه خواهد شد. تحت فرضیه H_1 نوع وابستگی داده‌ها مشخص نیست. بنابراین برای مقایسه توان ۸ مدل

$$M1 : Z_t = \sqrt{h_t} \varepsilon_t, \quad h_t = (1 + 0.8 Z_{t-1}^2),$$

$$M2 : Z_t = 0.8 \varepsilon_{t-1} Z_{t-1} + \varepsilon_t,$$

$$M3 : Z_t = Z_{t-1} + \varepsilon_t,$$

$$M4 : Z_t = 0.3 Z_{t-1} + \varepsilon_t,$$

$$M5 : Z_t = \text{sign}(Z_{t-1}) + 0.43 \varepsilon_t,$$

$$M6 : Z_t = 0.8 \sqrt{|Z_{t-1}|} + \varepsilon_t,$$

$$M7 : Z_t = \varepsilon_t + 0.8 \varepsilon_{t-1}^2 + 0.8 \varepsilon_{t-2}^2 + 0.8 \varepsilon_{t-3}^2,$$

$$M8 : Z_t = \varepsilon_t + 0.8 \varepsilon_{t-1}^2,$$

استفاده شد که در آن ε_t ها دارای توزیع نرمال استاندارد هستند. به منظور مقایسه توان آزمون پس از شبیه‌سازی به ازای حجم نمونه $(n = 50, 240, 580)$ ، مدل وابستگی و مقدار λ معلوم آماره $T_n^\lambda(\hat{\mathbf{P}}, \mathbf{P}^\circ)$ محاسبه شد. با تکرار این موضوع برای $N = 2000$ بار، میانگین تعداد دفعاتی که فرضیه H_0 رد می‌شود برآورد توان آزمون برای آن مدل خواهد بود. در واقع اگر $T_{1,n}^\lambda(\hat{\mathbf{P}}, \mathbf{P}^\circ)$ مقدار $T_n^\lambda(\hat{\mathbf{P}}, \mathbf{P}^\circ)$ تحت فرضیه H_1 (داده شبیه‌سازی شده‌ی وابسته) و $Q_{1-\alpha}$ چندک مرتبه $1 - \alpha$ توزیع مجموع موزون خی-دو باشد، برآورد توان از رابطه

$$\hat{\beta}_n^\lambda = \frac{\#\{T_{1,n}^\lambda(\hat{\mathbf{P}}, \mathbf{P}^\circ) > Q_{1-\alpha}\}}{N}, \quad (12)$$

بدست می‌آید.

به ازای حجم نمونه ۵۰، ۲۴۰ و ۵۸۰ توان آزمون به تفکیک مدل‌های $M1$ تا $M8$ محاسبه شده و نتایج در جدول ۱ آمده است. همان‌طور که انتظار می‌رفت، با افزایش حجم نمونه در هر یک از مدل‌ها (با λ ثابت)

جدول ۱: برآورد خطای نوع اول و توان برای چند حالت خاص λ و حجم نمونه

برآورد توان											برآورد خطای نوع اول				نوع آزمون
M/A	M/V	M/E	M/D	M/C	M/T	M/R	M/I	$N(0,1)$	$U(0,1)$	χ^2	t	n			
۰/۱۳۳	۰/۲۵	۰/۱۱۳	۰/۰۸۱	۰/۱۵۱	۰/۶۶۳	۰/۱۲۲	۰/۰۸۲	۰/۱۱۲	۰/۰۹۵	۰/۱۰۲	۰/۱۰۴	۵۰	خی-دو		
۰/۳۲۵	۰/۶۸۳	۰/۲۳۲	۰/۰۵۹	۰/۲۲۴	۰/۹۹۵	۰/۳۶۴	۰/۰۷۹	۰/۱۷۹	۰/۱۷۹	۰/۱۷۹	۰/۱۸۸	۲۴۰	اصلاح شده		
۰/۸۷۶	۱/۰۰۰	۰/۷۵۹	۰/۱۰۶	۰/۸۱۱	۱/۰۰۰	۰/۹۹۵	۰/۳۳۶	۰/۰۷۹	۰/۰۸۰	۰/۰۹۴	۰/۰۸۷	۵۸۰	($\lambda = -2$)		
۰/۱۷۸	۰/۱۹۰	۰/۰۹۹	۰/۰۷۳	۰/۰۹۵	۰/۵۵۰	۰/۱۰۲	۰/۰۷۰	۰/۰۷۳	۰/۰۷۴	۰/۰۶۷	۰/۰۸۰	۵۰	نسبت درست‌سازی		
۰/۶۷۸	۰/۸۵۷	۰/۳۷۳	۰/۰۹۵	۰/۴۰۴	۱/۰۰۰	۰/۸۵۰	۰/۱۸۲	۰/۰۸۷	۰/۰۸۵	۰/۰۸۷	۰/۰۹۳	۲۴۰	اصلاح شده		
۰/۸۵۹	۱/۰۰۰	۰/۷۱۵	۰/۰۷۳	۰/۷۶۸	۱/۰۰۰	۰/۹۹۲	۰/۲۸۶	۰/۰۶۲	۰/۰۵۹	۰/۰۶۶	۰/۰۶۲	۵۸۰	($\lambda \approx -1$)		
۰/۱۱۲	۰/۱۷۹	۰/۰۹۳	۰/۰۶۴	۰/۰۸۴	۰/۵۲۸	۰/۱۱۶	۰/۰۵۸	۰/۰۶۰	۰/۰۵۹	۰/۰۶۸	۰/۰۶۷	۵۰	فرینس-توکی		
۰/۵۹۰	۰/۸۴۴	۰/۳۲۳	۰/۰۷۰	۰/۳۶۲	۱/۰۰۰	۰/۶۷۹	۰/۱۶۱	۰/۰۵۹	۰/۰۶۵	۰/۰۷۱	۰/۰۷۰	۲۴۰	($\lambda = -\frac{1}{2}$)		
۰/۸۵۳	۱/۰۰۰	۰/۷۰۰	۰/۰۶۹	۰/۷۶۱	۱/۰۰۰	۰/۹۹۳	۰/۲۶۷	۰/۰۵۱	۰/۰۵۷	۰/۰۴۷	۰/۰۵۴	۵۸۰			
۰/۰۹۸	۰/۱۶۸	۰/۰۸۳	۰/۰۵۳	۰/۰۶۶	۰/۵۲۸	۰/۱۰۴	۰/۰۳۹	۰/۰۵۸	۰/۰۴۵	۰/۰۴۸	۰/۰۴۸	۵۰	نسبت درست‌سازی		
۰/۵۲۸	۰/۸۴۹	۰/۲۹۴	۰/۰۵۱	۰/۳۱۸	۱/۰۰۰	۰/۵۹۸	۰/۱۱۵	۰/۰۴۳	۰/۰۴۹	۰/۰۵۰	۰/۰۵۴	۲۴۰	($\lambda \approx 0$)		
۰/۸۵۱	۱/۰۰۰	۰/۶۹۵	۰/۰۵۸	۰/۷۳۷	۱/۰۰۰	۰/۹۸۷	۰/۲۵۲	۰/۰۴۸	۰/۰۵۲	۰/۰۴۵	۰/۰۴۴	۵۸۰			
۰/۱۱۲	۰/۱۷۸	۰/۰۷۲	۰/۰۳۶	۰/۰۷۵	۰/۵۷۱	۰/۰۸۹	۰/۰۵۰	۰/۰۴۲	۰/۰۴۹	۰/۰۴۸	۰/۰۵۰	۵۰	گرسی-ریبند		
۰/۴۹۷	۰/۹۴۲	۰/۲۷۵	۰/۰۵۰	۰/۳۲۳	۱/۰۰۰	۰/۵۵۵	۰/۱۱۳	۰/۰۳۹	۰/۰۳۶	۰/۰۳۸	۰/۰۴۰	۲۴۰	($\lambda = \frac{1}{2}$)		
۰/۸۵۰	۱/۰۰۰	۰/۷۰۲	۰/۰۵۵	۰/۷۴۱	۱/۰۰۰	۰/۹۸۶	۰/۲۳۷	۰/۰۳۹	۰/۰۳۴	۰/۰۴۴	۰/۰۴۵	۵۸۰			
۰/۰۹۵	۰/۱۷۰	۰/۰۸۷	۰/۰۴۳	۰/۰۷۴	۰/۵۹۳	۰/۰۸۸	۰/۰۵۲	۰/۰۵۱	۰/۰۳۹	۰/۰۳۹	۰/۰۴۸	۵۰	خی-دو		
۰/۴۸۳	۰/۹۴۳	۰/۲۷۶	۰/۰۴۹	۰/۳۲۹	۱/۰۰۰	۰/۵۵۵	۰/۱۱۲	۰/۰۴۶	۰/۰۴۴	۰/۰۴۲	۰/۰۳۵	۲۴۰	($\lambda = 1$)		
۰/۹۴۱	۱/۰۰۰	۰/۷۰۱	۰/۰۶۲	۰/۷۳۸	۱/۰۰۰	۰/۹۸۱	۰/۲۳۹	۰/۰۴۵	۰/۰۴۲	۰/۰۴۵	۰/۰۵۶	۵۸۰			

توان آزمون افزایش می‌یابد. نکته قابل توجه در جدول ۱ برتری توان $\lambda = -2$ است. در صورتی که توان آزمون را به ازای مقدار $\lambda = -2$ (به دلیل عدم حفظ سطح اسمی خطای نوع اول) در نظر نگیریم، در جدول ۱ به ازای حجم نمونه 50° و 24° مقدار $\lambda = -1$ توان بیشتری نسبت به سایر مقادیر λ داشته و می‌توان این آزمون را برای این دو حجم نمونه انتخاب نمود. اما برای حجم نمونه بالا ($n = 580^\circ$) با توجه جدول ۱ مقدار $\lambda = -2$ توصیه می‌شود. شاید احساس شود انتخاب $\lambda = -2$ برای حجم نمونه بالا کم اهمیت باشد زیرا با افزایش حجم نمونه توان برای هر آزمونی به مقدار ۱ نزدیک می‌شود. اما تفاوت آزمون‌ها به ازای مدل‌های وابستگی پیچیده بوده که توان بالایی را حتی در حجم نمونه 580° نشان نخواهند داد.

تذکر ۳. توان آزمون برای طیف گسترده‌تری از λ نیز مورد بررسی قرار گرفت. نتایج نشان داد که هر چند برای برخی از مقادیر λ آزمون توان بالایی دارد، اما در این آزمون‌ها سطح اسمی خطا حفظ نخواهد شد.

تذکر ۴. شبیه‌سازی مجدد به ازای طیف گسترده‌تری از n انجام شد. نتایج نشان داد که می‌توان مقدار بهینه m را برابر $\max\{m, 5m! \leq n\}$ در نظر گرفت. همچنین پس از تعیین مقدار بهینه m (به ازای حجم نمونه معلوم) و محاسبه $5(m^*)! - 5(m^* + 1)!$ به وسیله جدول ۲، مقدار بهینه λ می‌شود.

جدول ۲: مقدار پیشنهادی λ به ازای هر حجم نمونه.

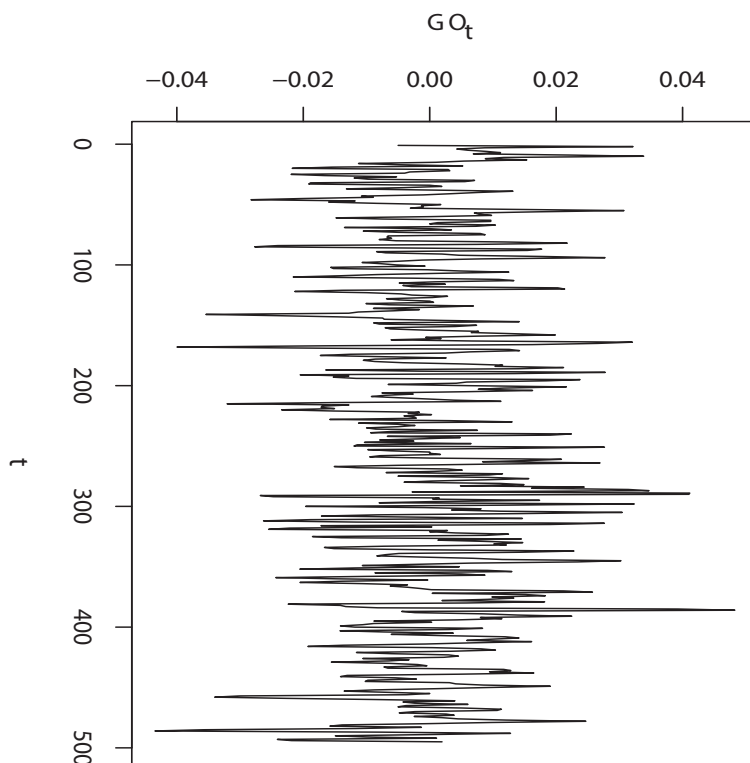
$5m^*! \leq n \leq 5m^*! + \frac{d}{4}$	$5m^*! + \frac{d}{4} < n \leq 5m^*! + \frac{3d}{4}$	$5m^*! + \frac{3d}{4} < n < 5(m^* + 1)!$
$\lambda = -\frac{1}{4}$	$\lambda = -1$	$\lambda = -2$

۵ مثال کاربردی

تغییرات قیمت طلا در جهان سهم عمده‌ای در اقتصاد ملی ایران دارد. اما ساختار پیچیده این تغییرات پیش‌بینی را برای این متغیر دشوار نموده است. برای بررسی این موضوع، بازده قیمت طلا (GO_t) در ۴۹۵ روز کاری پیاپی (دربازه‌ی زمانی ۲۰۱۵/۱/۱ تا ۲۰۱۶/۱۱/۲۵) از رابطه

$$GO_t = \ln\left(\frac{G_t}{G_{t-1}}\right)$$

محاسبه شد که در آن G_t قیمت روزانه هر انس طلا بر حسب دلار امریکا است. شکل ۲ سری زمانی بازده قیمت طلا را به تصویر می‌کشد.



شکل ۲: شاخص بازده قیمت طلا.

با توجه به جدول ۲، مقدار $\lambda = -2$ و $m = 4$ برای آزمون مبتنی بر واگرایی توان انتخاب شد. پس از انجام آزمون مقدار احتمال برابر $10^{-46} \times 36\%$ بدست آمده که حاکی از رد فرضیه H_0 است. با توجه به مشاهده ناهمسانی واریانس در داده‌ها، مدل از رده مدل‌های $GARCH$ در نظر گرفته شد. پس از مقایسه مدل‌ها مشخص شد، مدل $ARMA(1, 2) - GARCH(1, 1)$ با کمترین مقدار AIC بهترین برازش را داشته است. نتایج برازش این مدل به صورت

$$GO_t = -6\% \times 10^{-5} + 0.343GO_{t-1} + 0.04\epsilon_{t-1} - 0.337\epsilon_{t-2} + \hat{h}_t,$$

$$\hat{h}_t = \sqrt{2.4 \times 10^{-6} + 0.21\hat{h}_{t-1}\epsilon_{t-1}^2 + 0.857\hat{h}_{t-1}}$$

است. پس از برازش مدل، آزمون استقلال مبتنی بر واگرایی توان (با پارامترهای $\lambda = -2$ و $m = 4$) برای

مانده‌های مدل انجام شد. در این آزمون مقدار احتمال برابر $0/714$ بدست آمد که نشان دهنده‌ی مستقل و هم توزیع بودن مانده‌های مدل برازش شده است. از این رو، این مدل را می‌توان به عنوان مدلی معتبر برای برازش داده‌های مورد مطالعه استفاده نمود.

۶ بحث و نتیجه‌گیری

در این پژوهش رده جدیدی از آزمون‌های استقلال سری زمانی مبتنی بر اندازه و آگرایی توان ارائه شد. این آزمون‌ها پیش از آن که برای بررسی استقلال سری زمانی کاربرد داشته باشد، برای بررسی استقلال مانده‌های مدل‌های سری زمانی مفید است. ابتدا بردارهای m -بعدی شامل m نقطه مجاور در هر زمان تشکیل شد. سپس بر اساس این واقعیت که احتمال رخداد هر جایگشت در بردار m -بعدی تحت فرضیه H_0 (مستقل و هم توزیع بودن) برابر $\frac{1}{m!}$ می‌باشد آماره آزمون $T_n^\lambda(\hat{P}, P^0)$ تعریف شد. لازم به ذکر است که بردارهای m -بعدی تشکیل شده دارای هم پوشانی بوده و با در نظر گرفتن این موضوع توزیع حدی آماره آزمون را محاسبه شد. هر چند حالت خاص این آزمون توسط ماتریلا-گارسیا (۲۰۰۷) و ماتریلا-گارسیا و مارین (۲۰۰۸) با سایر آزمون‌های استقلال سری زمانی مقایسه شده است، اما در آینده تحقیق می‌توان این رده از آزمون‌ها را به طور گسترده‌تری با سایر آزمون‌های استقلال سری زمانی مقایسه نمود. در یک مثال کاربردی، ابتدا پیش‌بینی‌پذیری سری زمانی بازده قیمت طلا مورد آزمون قرار گرفت. پس از رد فرضیه H_0 ، مدل $ARMA(1, 2) - GARCH(1, 1)$ به این داده‌ها برازش داده شد. سپس مانده‌های مدل برازش شده مورد آزمون قرار گرفت که نتایج حاکی از مستقل و هم توزیع بودن مانده‌ها بود.

تقدیر و تشکر

نویسندگان کمال تشکر را از داوران و ویراستار مجله که با پیشنهادات ارزنده خود باعث بهبود مقاله شدند دارند.

مراجع

- Amigó, J. M., Zambrano, S. and Sanjuán, M. A. (2007), True and False Forbidden Patterns in Deterministic and Random Dynamics, *EPL (Europhysics Letters)*, **79**, 50001-p1-p5.
- Ashtari Nezhad, E., Waghei, Y., Mohtashami Borzadaran, G. R., Nilli Sani, H. R. and Alizadeh Noughabi, H. (2018), The Modified Permutation Entropy-Based Independence Test of Time Series, *Communications in Statistics: Simulation and Computation*, Doi:10.1080/03610918.2018.1469761.
- Bartels, R. (1982), The Rank Version of Von Neumann's Ratio Test for Randomness, *Journal of the American Statistical Association*, **77**, 40-46.
- Broock, W. A., Scheinkman, J. A., Dechert, W. D. and LeBaron, B. (1996), A Test for Independence Based on the Correlation Dimension, *Econometric reviews*, **15**, 197-235.
- Cánovas, J. S. and Guillamón, A. (2009), Permutations and Time Series Analysis, *Chaos: An Interdisciplinary Journal of Nonlinear Science*, **19**, 043103.
- Cánovas, J. S., Guillamón, A. and Vera, S. (2013), Testing for Independence: Permutation Based Tests vs. BDS Test, *The European Physical Journal Special Topics*, **222**, 275-284.
- Cressie, N. and Read, T. R. C. (1984), Multinomial Goodness-of-Fit Tests, *Journal of the Royal Statistical Society, Series B*, **46**, 440-464.
- Dik, J. J. and de Gunst, M. C. M. (1985), The Distribution of General Quadratic Forms in Normal Variables, *Statistica Neerlandica*, **39**, 14-26.

- Elsinger, H. (2010), Independence Tests Based on Symbolic Dynamics (No. 165), *Oesterreichische Nationalbank*,
- Gut, A. (2006), *Probability: A Graduate Course*, Springer Science & Business Media.
- Ghoudi, K., Kulperger, R. J. and Rémillard, B. (2001), A Nonparametric Test of Serial Independence for Time Series and Residuals, *Journal of Multivariate Analysis*, **79**, 191-218.
- Liu, H., Tang, Y. and Zhang, H. H. (2009), A New Chi-square Approximation to the Distribution of Non-negative Definite Quadratic Forms in Non-central Normal Variables, *Computational Statistics and Data Analysis*, **53**, 853-856.
- Hong, Y. (2000), Generalized Spectral Tests for Serial Dependence, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **62**, 557-574.
- Matilla-García, M. (2007), A Non-Parametric Test for Independence Based on Symbolic Dynamics, *Journal of Economic Dynamics and Control*, **31**, 3889-3903.
- Matilla-García, M. and Marín, M. R. (2008), A Non-parametric Independence Test Using Permutation Entropy, *Journal of Econometrics*, **144**, 139-155.
- Matilla-García, M., Marín, M. R., Dore, M. I. and Ojeda, R. B. (2014), Nonparametric Correlation Integral-Based Tests for Linear and Nonlinear Stochastic Processes, *Decisions in Economics and Finance*, **37**, 181-193.
- Pardo, L. (2006), *Statistical Inference Based on Divergence Measures*, Chapman and Hall/CRC, Taylor & Francis Group, Boca Raton, FL.
- Sensoy, A., Aras, G. and Hacıhasanoglu, E. (2015), Predictability Dynamics of Islamic and Conventional Equity Markets, *The North American Journal of Economics and Finance*, **31**, 222-248.

Independence Test of Time Series Based on Power-Divergence

Ashtarinezhad, E.¹, Waghei, Y.¹, Mohtashami Borzadaran, G. R.², Nilli Sani, H. R.¹ Alizadeh Noughabi, H.¹

¹Department of Statistics, University of Birjand, Birjand, Iran.

²Department of Statistics, Faculty of Mathematical Sciences, Ferdowsi University of Mashhad, Mashhad, Iran.

Abstract:

Before analyzing a time series data, it is better to verify the dependency of the data, because if the data be independent, the fitting of the time series model is not efficient. In recent years, the power divergence statistics used for the goodness of fit test. In this paper, we introduce an independence test of time series via power divergence which depends on the parameter λ . We obtain asymptotic distribution of the test statistic. Also using a simulation study, we estimate the error type I and test power for some λ and n . Our simulation study shows that for extremely large sample sizes, the estimated error type I converges to the nominal α , for any λ . Furthermore, the modified chi-square, modified likelihood ratio, and Freeman-Tukey test have the most power.

Keywords: Independence Test, Time Series, Power-divergence, m-dependence Random Variable.

Mathematics Subject Classification (2010): 37M10, 68Q30, 62G10.