



مدلسازی و پیش‌بینی نوسان تحقق‌یافته با در نظر گرفتن پرش در بورس اوراق بهادار تهران

سعید فلاح‌پور^۱
وحید مطهری‌نیا^۲

تاریخ دریافت: ۹۴/۱۲/۲۴ تاریخ پذیرش: ۹۵/۰۳/۱۹

چکیده

در سال‌های اخیر بازارهای مالی با نوسانات زیادی مواجه شده و عدم اطمینان ناشی از این نوسان‌ها، نگرانی‌هایی را در سرمایه‌گذاران ایجاد نموده است. از این رو مدل‌سازی نوسان و پیش‌بینی آن در مسائل مختلف تحقیقی و عملی مالی، مورد توجه قرار گرفته است. در این راستا، امکان دسترسی به داده‌های پرفراوانی، عرصه‌ی جدیدی برای مدل‌سازی نوسان و پیش‌بینی بازده دارایی‌های مالی ایجاد نموده است. در این پژوهش، مدل‌سازی نوسان با استفاده از داده‌های پرفراوانی و با کمک مدل‌های خانواده‌ی HAR-RV، انجام شده و اثر اضافه نمودن جزء پرش در کارایی پیش‌بینی نوسان شاخص کل بورس اوراق بهادار تهران مورد بررسی قرار گرفته است. نتایج مدل‌سازی و پیش‌بینی حاکی از این است که خطای پیش‌بینی با اضافه نمودن جزء پرش به مدل کاهش یافته و همچنین مجزا نمودن اجزای پرش و پیوسته نوسان تحقق‌یافته نیز در بهبود کارایی پیش‌بینی تاثیرگذار است.

واژه‌های کلیدی: نوسان تحقق یافته، مدل‌سازی نوسان، HAR-RV، شناسایی پرش، بورس اوراق بهادار تهران.

۱- استادیار گروه مدیریت مالی و بیمه، دانشکده مدیریت دانشگاه تهران، تهران، ایران Sfallahpour@gmail.com
۲- کارشناسی ارشد مهندسی مالی، دانشکده مدیریت دانشگاه تهران، تهران، ایران Vahid.motahari@gmail.com

۱- مقدمه

کلمه "نوسان" در مباحث مالی به تغییر قیمت در یک بازه زمانی مشخص اشاره می‌نماید. با این تعریف، معمولاً انحراف معیار استاندارد، به عنوان نوسان در نظر گرفته می‌شود. اگر چه در واقع نوسان و انحراف معیار استاندارد به طور کلی یکی نیستند. در مباحث مالی، نوسان بازده از اهمیت زیادی برخوردار است. از این نوسان در قیمت‌گذاری دارایی‌ها، تصمیمات مربوط به پرتفولیو و مدیریت ریسک از طریق محاسبات مربوط به ارزش در معرض خطر^۱ استفاده می‌شود. دو نوع اصلی از نوسان، نوسان ضمنی و نوسان تحقق‌یافته (نوسان تاریخی) است. نوسان ضمنی اغلب در بحث قیمت‌گذاری اختیار مورد استفاده قرار می‌گیرد. نوسان ضمنی از قیمت بازاری اختیار و بر اساس مدل قیمت‌گذاری اختیار به دست آمده و به عنوان نظر بازار نسبت به نوسان آینده دارایی در نظر گرفته می‌شود. نوسان ضمنی در بیشتر منابع با نماد σ نشان داده می‌شود. در سمت دیگر، نوسان تحقق‌یافته آنچه را که در گذشته اتفاق افتاده اندازه‌گیری می‌کند. این نوع نوسان از حاصل جمع توان دوم بازده‌های لگاریتمی با توجه به یک فرکانس خاص محاسبه می‌شود. با توجه به این تعریف، اگر فرکانس بالاتری استفاده شود، می‌توان به اطلاعات بیشتری دست یافت. افزایش فرکانس نمونه‌گیری به معنای استفاده از داده‌های میان روزانه و ورود به حوزه داده‌های پرفراوانی است.

پیش‌بینی نوسان‌ها و نااطمینانی‌های موجود در سرمایه‌گذاری و مدیریت ریسک با استفاده از آن، بخش مهمی از فرایند سرمایه‌گذاری است. پیش‌بینی نوسان، با استفاده از مدل‌سازی ریاضی انجام می‌شود. پیشرفت‌های اخیر صورت گرفته در زمینه اقتصادسنجی مالی، منجر به ارایه مدل‌هایی کمی گردیده است که با استفاده از آنها می‌توان خط مشی سرمایه‌گذاران را در ارتباط با ریسک و بازده مورد انتظار آنها و همچنین نوسان موجود در بازار، تشریح نمود. امروزه تحلیلگران مالی با استفاده از مدل‌های سری زمانی اقتصادسنجی به مدلسازی و تبیین رفتار بازده نوسان در بازار سهام می‌پردازند. دو ویژگی اساسی سری‌های زمانی مالی، یعنی وجود دنباله‌های پهن توزیع داده‌ها و نوسانات خوشه‌ای در آنها را می‌توان با استفاده از مدل‌های خانواده GARCH مدل‌سازی نمود. انگل^۲ (۱۹۸۲) برای اولین بار مدلی (ARCH) را ارائه نمود که بر اساس آن توانست ناهمسانی واریانس شرطی را مدلسازی نماید. بولرسلو^۳ (۱۹۸۶) با بسط مدل انگل، مدل ناهمسانی واریانس شرطی اتورگرسیو تعمیم‌یافته (GARCH) را معرفی نمود. پس از آن مدل‌های گوناگونی بر پایه مدل اولیه انگل (ARCH) توسط محققین مختلف ارائه گردیدند که هر کدام به دنبال رفع یکی از حقایق کلی داده‌های مالی و رسیدن به کمترین خطا در مدل‌سازی بوده‌اند.

در جستجوی یافتن بهترین مدل، اندرسن و همکارانش^۴ (۱۹۹۷، ۲۰۰۲) دریافتند که انتشار اخبار مهم در رابطه با یک دارایی، مانند اخبار اقتصاد کلان و یا گزارش سود و زیان، باعث ایجاد یک پرش^۵ گسسته در قیمت به جهت بازنگری سریع در ارزش دارایی می‌شود. از این رو می‌توان مدل‌های موجود را به گونه‌ای اصلاح نمود که بتوان با استفاده از آن، اثر این پرش‌ها را جدا نموده و مدلی پیشنهاد نمود که بر اساس اجزای پیوسته سری داده‌ها و اجزای پرش آن، قابلیت پیش‌بینی بهتری را ایجاد نماید.

۲- مبانی نظری و مروری بر پیشینه پژوهش

۱-۲- نوسان انباشته و نوسان تحقق یافته

برای نشان دادن مفهوم نوسان انباشته برای روز t ، $(\sigma_t^{(d)})$ با فرض یک فرایند براونی هندسی برای قیمت، فرایند نوسان تصادفی زیر در نظر گرفته می‌شود.

$$ds(t) = \mu(t)dt + \sigma(t)dW(t) \quad (1-2)$$

در این رابطه $s(t)$ لگاریتم قیمت‌های لحظه‌ای، $W(t)$ یک حرکت براونی استاندارد، $\mu(t)$ جمله رانش با تغییرات محدود و $\sigma(t)$ یک فرایند تصادفی مستقل از $dW(t)$ است. به زبان دیگر، $\mu(t)$ معمولاً نرخ بازده لگاریتم قیمت سهم و $\sigma(t)$ نوسان آنی (یا انحراف معیار استاندارد) است. به ترتیب قبل، $\sigma(t)dW(t)$ یک مارتینگل نسبی است و $ds(t)$ از یک فرایند نیمه مارتینگلی خاص تبعیت می‌کند. (پراتر^۶، ۱۹۹۰)

برای این فرایند، نوسان انباشته برای روز t ام، برابر انتگرال نوسانات لحظه‌ای در بازه زمانی یک روز $(t-1d, t)$ است. نماد $1d$ در اینجا به معنی یک روز کامل است.

$$\sigma_t^{2(d)} = \int_{t-1d}^t \sigma^2(w)dw \quad (2-2)$$

مرتون^۷ (۱۹۸۰) نشان داد که نوسان انباشته یک فرایند حرکت براونی را می‌توان با دقت خوبی از مجموع مربع بازده‌های میان روزانه محاسبه نمود. در سال‌های بعد نیز اندرسن، بولرسلو و دیبولد^۸ (۲۰۰۱) با استفاده از تئوری تغییرات درجه دوم، این موضوع را برای دسته خاصی (نوسان محدود) از فرایندهای نیمه مارتینگل اثبات نمودند. در واقع با وجود این شروط، مجموع مربع بازده‌های میان‌روزانه به سمت نوسان انباشته قیمت‌ها میل می‌کند. این تخمین زنده ناپارامتریک، نوسان تحقق‌یافته نامیده شده است. تعریف دقیق این تخمین زنده به شکل زیر است:

$$RV_t = \sum_{j=1}^M r_{t-j\Delta}^2 \quad (3-2)$$

که در آن $\Delta = 1d/M$ برابر تواتر میان روزانه و $p(t-j\Delta) - p(t-(j+1)\Delta)$ بازده میان روزانه می‌باشند. توجه شود که تعریف نوسان تحقق یافته (همانند سایر تعاریف موجود برای نوسان تاریخی) دارنده دو پارامتر زمان است: تواتر میان روزانه بازده و افق زمانی محاسبه (روزانه، هفتگی و ماهانه). برای محاسبه نوسان تحقق یافته هفتگی، به سادگی از روش زیر استفاده می‌شود:

$$RV_t^{(w)} = \frac{1}{5} (RV_{t-1d}^{(d)} + RV_{t-2d}^{(d)} + \dots + RV_{t-5d}^{(d)}) \quad (4-2)$$

جایی که $1w=5d$ نشان دهنده بازه زمانی یک هفته است. در عمل از افق‌های زمانی روزانه، هفتگی و ماهانه استفاده می‌شود.

۲-۲- پیشینه تحقیق

مدل اتورگرسیو ناهمگون نوسان تحقق یافته^۹ اولین بار توسط کرسی^{۱۰} (۲۰۰۴) معرفی شد. ایده اصلی این مدل از تئوری ناهمگونی بازار که توسط مولر و همکارانش^{۱۱} در سال ۱۹۹۳ ارائه شده، گرفته شده است و وجود ناهمگونی در معامله‌گران را تشخیص می‌دهد. تئوری ناهمگونی بازار تلاش می‌کند تا وجود همبستگی زیاد مقدار نوسان و میزان و نوع مشارکت‌کنندگان در بازار را توجیه نماید. در یک بازار همگون، جایی که خصوصیات تمامی مشارکت‌کنندگان در آن یکسان باشد؛ هر چه تعداد اشخاص بیشتری وجود داشته باشد، قیمت‌ها سریعتر به ارزش بازار واقعی که تمامی افراد با آن موافقت، میل می‌کند. بنابراین نوسان باید همبستگی منفی با میزان حضور در بازار و میزان فعالیت در آن داشته باشد. در مقابل در یک بازار ناهمگون، بازیگران متفاوت برای قیمت‌های مختلف مد نظر خود اقدام نموده و تلاش دارند که معاملات خود را در شرایط مختلف بازار به انجام رسانند. از این رو، این بازیگران نوسان را به وجود می‌آورند. (کرسی، ۲۰۰۴)

کرسی نوسان جزئی را به شکل نوسانی که از یک جزء مشخص بازار ناشی می‌شود، تعریف نمود و مدلی معرفی کرد که از مجموع این اجزا حاصل می‌شود. او با استفاده از داده‌های پرفراوانی مربوط به نرخ مبادله‌ی جفت ارز USD/CHF، مدل HAR-RV را با مدل‌های GARCH(1,1)، مدل ریسک متریک مورگان (RM)، AR(1)، AR(3) و مدل ARFIMA(5,0.401,0) مقایسه نمود. کرسی اثبات کرد که مدل HAR-RV بر سایر مدل‌ها غلبه نموده و مدل‌سازی رفتار حافظه‌ی طولانی نوسان را با استفاده از روشی ساده و با صرفه، با موفقیت به انجام می‌رساند. (کرسی، ۲۰۰۴)

اندرسن و همکارانش (۲۰۰۷)، با استفاده از سنجه‌ی bi-power variation و استفاده از روش تست پرش بارنداف، نیلسن و شفرد^{۱۲} (۲۰۰۴)، با توسعه مدل HAR-RV کرسی، برای اولین بار مدل HAR-RV-CJ را معرفی نمودند. در این مدل، ایده اصلی این بود که انتشار اخبار مهم در رابطه با یک دارایی، مانند اخبار اقتصاد کلان و یا گزارش سود و زیان، باعث ایجاد یک پرش گسسته در قیمت به جهت بازنگری سریع در ارزش دارایی می‌شود. (اندرسن و بولرسلو، ۱۹۹۷؛ اندرسن و همکارانش، ۲۰۰۲) از این رو می‌توان مدل معرفی شده را به گونه‌ای گسترش داده تا بتوان اثر این پرش‌ها را جدا نموده و مدلی پیشنهاد نمود که بر اساس اجزای پیوسته سری داده‌ها و اجزای پرش آن، قابلیت پیش‌بینی بهتری را ایجاد نماید. آنها با در نظر گرفتن جزء پرش و جزء پیوستگی نوسان تحقق یافته، مدل خود را روی داده‌های نرخ مبادله‌ی \$/DM، شاخص S&P 500 و نرخ بازده اوراق خزانه ۳۰ ساله آمریکا به کار گرفتند. نویسندگان به این نتیجه رسیدند که جزء پرش واریانس تحقق یافته که نویز را تشکیل می‌دهد، هیچ کمکی به پیش‌بینی نوسان تحقق یافته‌ی آینده نمی‌کند و تنها جزء پیوسته‌ی آن دارای قدرت پیش‌بینی است. آنها همچنین به این نتیجه رسیدند که بیشتر پرش‌های قابل توجه در ارتباط با انتشار اخبار مربوط به اقتصاد کلان هستند.

مکلیر و مدیراس^{۱۳} (۲۰۰۸) نیز مدل HAR-RV را گسترش دادند. آنها مدلی جدید معرفی نمودند که غیر خطی بودن و حافظه‌ی طولانی را با هم ترکیب می‌کند. این دسته‌ی جدید به نام مدل اتورگرسیو

ناهمگون انتقال هموار چند رژیمه^{۱۴} شناخته می‌شود. برتری اصلی مدل HARST این است که می‌تواند به روشی آسان، وابستگی داده‌ها در بلندمدت به یکدیگر و همچنین تغییر رژیم‌ها را مدل نماید. آنها با موفقیت نشان دادند زمانی که HAR-RV و HARST با یکدیگر ترکیب شوند، کارایی پیش‌بینی افزایش می‌یابد و زمانی که این مدل با مدل‌های دیگر (GARCH, EGARCH, GJR-GARCH) مقایسه می‌شود، کارایی هر دو مدل HAR و HARST بسیار بهتر است.

چانگ، هوانگ و تسنگ^{۱۵} (۲۰۰۸)، برای تحقق روی حجم معاملات، مدل HAR-RV را به کار گرفتند و یک مدل جدید پیش‌بینی به نام HAR-RV-cum-Volume ایجاد نمودند. بحث اختلاط تنوری توزیع‌ها^{۱۶} زیر بنای تنوری تحلیل ارتباط بین حجم و نوسان را ارائه می‌نماید. آنها با استفاده از داده‌های پرفراوانی سهام عادی IBM، نتیجه گرفتند که با رگرسیون یکسان، مدل جدید، کارایی پیش‌بینی نوسان تحقق‌یافته آینده را بهبود می‌بخشد.

سلیک و ارگن^{۱۷} (۲۰۱۴) با استفاده از داده‌های پرفراوانی شاخص آتی ISE-30 در بورس مشتقات ترکیه، عملکرد مدل‌های GARCH, HAR-RV, HAR-RV-J, HAR-RV-CJ, MIDAS و RGARCH را با هم مقایسه نمودند. یافته‌های تحقیق، بر عملکرد بهتر مدل‌های پیش‌بینی پرفراوانی نسبت به مدل GARCH تاکید می‌کند. در میان مدل‌های پرفراوانی نیز دو مدل MIDAS و HAR-RV-CJ بهتر از سایر مدل‌ها به پیش‌بینی نوسان می‌پردازند.

در رابطه با مدل‌های HAR-RV و در نظر گرفتن پرش در داده‌ها، تاکنون پژوهشی در منابع فارسی انجام نشده است.

۲-۳- تغییرات درجه دوم (QV)^{۱۸}

در ریاضیات، از تغییرات درجه دوم برای تحلیل یک فرایند تصادفی استفاده می‌شود. این نوع تغییر برای مارتینگل‌ها و نیمه مارتینگل‌ها کاربرد بیشتری دارد. در واقع تغییرات درجه دوم گونه‌ای از تغییر یک فرایند است. اگر فرض شود که X_t یک فرایند تصادفی روی دامنه اعداد حقیقی است که روی فضای احتمال (Ω, \mathcal{F}, P) تعریف شده و شمارنده t در اعداد حقیقی مثبت تغییر می‌کند، تغییرات مرتبه دوم این فرایند، $[X]_t$ از طریق رابطه زیر محاسبه می‌شود:

$$[X]_t = \lim_{n \rightarrow \infty} \sum_{k=1}^n (X_{t_k} - X_{t_{k-1}})^2 \quad (۵-۲)$$

به طوری که $t_0 = 0 < t_1 < \dots < t_n = t$.

۲-۴- فرایند وینر همراه با پرش

در حال حاضر هیچ تئوری اقتصادی وجود ندارد که تاکید نماید قیمت از یک فرایند پیوسته کامل تبعیت می‌نماید. در مشاهدات کاربردی نیز تغییرات سریع در قیمت (پرش) وجود دارد. بارندورف، نیلسون و

شفر (۲۰۰۵) برای در نظر گرفتن پرش در قیمت، یک فرایند پرش محدود (در یک محدوده زمانی ثابت، تعداد محدودی پرش وجود داشته باشد) به مدل حرکت براونی نیمه مارتینگل اضافه نمودند.

$$s(t) = \int_0^t \mu(u) du + \int_0^t \sigma(u) dW(u) + \sum_{j=1}^{N(t)} \kappa(t_j) \quad (6-2)$$

فرایند میانگین $\mu(t)$ پیوسته و دارای نوسان محدود بوده و $N(t)$ تعداد پرش‌ها با شدت $\lambda(t)$ و سائز $\kappa(s_j)$ است به طوریکه $\kappa(t) = s(t) - s(t-1)$ نشان‌دهنده سائز پرش مربوطه در فرایند لگاریتمی قیمت می‌باشد.

با پیروی از اندرسن و همکارانش (۲۰۰۲) ابتدا $S(t)$ به شکل یک فرایند لگاریتم قیمت تک متغیره روی فضای احتمال کامل (Ω, \mathcal{F}, P) تعریف می‌شود. فرایند قیمت $S(t)$ به شکل پیوسته در بازه زمانی $[0, T]$ نمو می‌کند و T یک عدد طبیعی محدود است. مجموعه اطلاعات در دسترس نیز در هر لحظه با \mathcal{F}_t نشان داده می‌شود. داریم: $(\mathcal{F}_t)_{t \in [0, T]} \subseteq \mathcal{F}$. این مجموعه اطلاعات در هر لحظه شامل تاریخچه کامل قیمت دارایی و سایر متغیرهای مرتبط است. به پیروی از باردورف، نیلسون و شفر (۲۰۰۵) فرض می‌شود $S(t)$ از یک فرایند وینر نیمه مارتینگل به همراه پرش تبعیت می‌کند.

بولرسلو و همکارانش (۲۰۰۴) بیان نمودند که تئوری تغییرات کوادراتیک این اجازه را می‌دهد که نوسان قیمت به دو جزء پیوسته و جزء پرش تجزیه شود. همچنین اثبات نمودند که برای یک فرایند وینر نیمه مارتینگلی $S(t)$ همراه با پرش، تغییرات کوادراتیک به شکل زیر بدست می‌آید:

$$[S]_t = \int_0^t \sigma^2(u) d(u) + \sum_{j=1}^{N(t)} \kappa^2(t_j) \quad (7-2)$$

که در آن $\int_0^t \sigma^2(s) d(s)$ نوسان انباشته و $\sum_{j=1}^{N(t)} \kappa^2(t_j)$ مجموع مربعات پرش است. در واقع اگر t_- را مقدار کمتر و بسیار نزدیک به t تعریف نماییم: $\kappa(s) = X_t - X_{t_-}$ برابر سائز پرش در مقدار فرایند در زمان t است.

۲-۵- جداسازی جزء پرش

کامت و رینالت^{۱۹} (۱۹۹۸)، اندرسن و بولرسلو (۱۹۹۸)، ABDL (۲۰۰۱) و باردورف، نیلسون و شفر (۲۰۰۲) بیان نموده‌اند که واریانس تحقق‌یافته مستقیماً از تئوری تغییرات درجه دوم تبعیت می‌کند. واریانس تحقق‌یافته با احتمال یکنواخت در صورت میل نمودن M به بی‌نهایت، به QV میل می‌کند.

$$RV_{t, t+1}^M \rightarrow \int_t^{t+1} \sigma^2(u) du + \sum_{t < s < t+1} \kappa^2(s) \quad (8-2)$$

سپس باردورف، نیلسون و شفر (۲۰۰۲) دو فرایند کاربردی مربوط به واریانس تحقق‌یافته به نام‌های RPV^{20} و BPV^{21} را معرفی نمودند.

$$RPV_{t,t+1}^M = \sum_{i=1}^M |R_{t,i}^M| \quad (9-2)$$

$$BPV_{t,t+1}^M = \mu_1^{-2} \sum_{i=2}^M |R_{t,i}^M| |R_{t,i-1}^M| \quad (10-2)$$

$$\mu_a = E(|Z|^a), Z \sim N(0,1), a > 0, E(|Z|) = \sqrt{2/\pi} \quad (11-2)$$

RPV که ارزش تحقق‌یافته مطلق (RAV) نیز نامیده می‌شود از مجموع قدرمطلق بازده تمامی M بازده میان‌روزانه به دست می‌آید. BPV نیز از مجموع ضرب هر بازده میان‌روزانه در بازده قبلی آن و به شکلی تعدیل شده به دست می‌آید.

باردورف، نیلسون و شفر (۲۰۰۲) اثبات نمودند که با میل M به بی‌نهایت داریم:

$$\lim_{M \rightarrow \infty} RPV_{t,t+1}^M \rightarrow \int_t^{t+1} \sigma(s) ds = \sigma_{t,t+1} \quad (12-2)$$

$$\lim_{M \rightarrow \infty} BPV_{t,t+1}^M \rightarrow \int_t^{t+1} \sigma^2(s) ds = \sigma_{t,t+1}^2 \quad (13-2)$$

بنابراین BPV تخمینی نارایب از واریانس انباشته است. حال با ادغام معادلات قبل، داریم:

$$RV_{t,t+1}^M \rightarrow \int_t^{t+1} \sigma^2(s) ds + \sum_{t < s < t+1} \kappa^2(s) \quad (14-2)$$

این معادله نشان می‌دهد که واریانس تحقق‌یافته سنج‌های از کل واریانس محقق شده در قیمت‌ها است که شامل بخش ناپیوسته پرش‌ها در قیمت نیز می‌شود. با فرض $M \rightarrow \infty$ ، با توجه به معادلات (۱۲-۲)، (۱۳-۲) و (۱۴-۲) رابطه زیر به دست می‌آید:

$$RV_{t,t+1}^M - BPV_{t,t+1}^M \rightarrow \sum_{t < s < t+1} \kappa^2(s) \quad (15-2)$$

پس می‌توان جزء پرش را از تفاضل RV و BPV محاسبه نمود. با تبعیت از پیشنهاد باردورف، نیلسون و شفر (۲۰۰۴) جزء پرش را می‌توان به شکل زیر تخمین زد:

$$J_{t,t+1}^M = \max[RV_{t,t+1}^M - BPV_{t,t+1}^M, 0] \quad (16-2)$$

این رابطه برای اطمینان از نامنفی بودن تمامی پرش‌های میان‌روزانه مورد استفاده قرار می‌گیرد.

به پیروی از باردورف، نیلسون و شفرد (۲۰۰۴)، معادله زیر نیز که از تست پرش BPV باردورف اقتباس شده، تعریف می‌شود:

$$TQ_t^M = M\mu_{4/3}^{-3} \sum_{j=3}^M (|R_{t,i}^M| |R_{t,i-1}^M| |R_{t,i-2}^M|)^{4/3} \quad (17-2)$$

این معادله به نام TQ شناخته شده و در آن داریم:

$$E(|Z|^{4/3}) = 2^{2/3} \Gamma\left(\frac{7}{6}\right) \Gamma\left(\frac{1}{2}\right) \quad (18-2)$$

همچنین باردورف، نیلسون و شفرد اثبات نمودند که اگر M به بی‌نهایت میل کند TQ به $\sigma^4_{t,t+1}$ میل می‌نماید.

$$\lim_{M \rightarrow \infty} TQ_{t,t+1}^M \rightarrow \int_t^{t+1} \sigma^4(s) ds = \sigma^4_{t,t+1} \quad (19-2)$$

با توضیحات گفته شده و بر اساس نتیجه‌گیری باردورف، نیلسون و شفرد (۲۰۰۴) و اندرسن و همکاران (۲۰۰۷)، اجزای پیوسته و پرش‌های با اهمیت با استفاده از رابطه‌های زیر تشخیص داده می‌شوند:

$$J_{t,t+1}^\alpha = I(Z_{t,t+1} > \Phi_\alpha)(RV_{t,t+1} - BPV_{t,t+1}) \quad (20-2)$$

$$C_{t,t+1}^\alpha = I(Z_{t,t+1} \leq \Phi_\alpha)RV_{t,t+1} + I(Z_{t,t+1} > \Phi_\alpha)BPV_{t,t+1} \quad (21-2)$$

$$Z_{t,t+1} = \sqrt{M} \frac{(RV_{t,t+1} - BPV_{t,t+1})/RV_{t,t+1}}{\sqrt{(\mu_1^{-4} + 2\mu_1^{-2} - 5) \max(1, \frac{TQ_{t,t+1}}{BPV_{t,t+1}^2})}} \quad (22-2)$$

$I_A(X)$ یک تابع صفر و یک با حالات روبرو است:

$$I_A(X) = \begin{cases} 1; X \in A \\ 0; X \notin A \end{cases} \quad (23-2)$$

$Z_{t,t+1}$ توسط هوانگ و تاچن^{۲۲} (۲۰۰۵) یافت شد. این تابع، بسیار به توزیع گوسی استاندارد نزدیک است. در اکثر مقالات از $Z_{t,t+1}$ به منظور تشخیص پرش استفاده می‌شود. در این رابطه α سطح اطمینان و Φ_α مقدار بحرانی است که براساس توزیع گوسی استاندارد $N(0,1)$ محاسبه می‌شود. در سطح اطمینان $\alpha = 0.999$ مقدار Φ_α برابر ۳٫۰۹ می‌شود. بنابراین اگر $Z_{t,t+1} > 3.09$ باشد، یک پرش اتفاق افتاده و از این طریق تشخیص داده می‌شود. به راحتی مشخص است که با کاهش سطح اطمینان، تعداد پرش‌های بیشتری تشخیص داده می‌شود.

۲-۶- ارزیابی پیش‌بینی

به منظور ارزیابی پیش‌بینی انجام شده، از معیار $MAPE^{23}$ استفاده می‌شود. این معیار به شکل زیر مورد محاسبه قرار می‌گیرد.

$$MAPE = 100 \sum_{t=T+1}^{T+h} \left| \frac{\hat{\sigma}_t - \sigma_t}{\sigma_t} \right| / h \quad (2-24)$$

توجه شود که $\hat{\sigma}_t$ مقدار پیش‌بینی شده نوسان تحقق یافته و σ_t مقدار واقعی مشاهده شده می‌باشند. H نیز تعداد نمونه پیش‌بینی شده است. از معیار MAPE به دلیل بدون مقیاس بودن و ماهیت درصدی آن استفاده شده است. در محاسبات مربوط به مدل‌ها، تفاوت در نوع داده‌ها و بازه تغییرات آنها، استفاده از معیارهایی نظیر MSE را با مشکل مواجه می‌سازد.

۳- روش‌شناسی پژوهش

از دیدگاه طبقه‌بندی تحقیقات بر اساس هدف، این پژوهش در دسته پژوهش‌های کاربردی قرار می‌گیرد. همچنین با توجه به روش پژوهش و نوع داده‌ها، پژوهش حاضر از نوع تحقیقات تاریخی و کمی محسوب می‌شود. به منظور بررسی مدل‌های مدنظر و آزمون فرضیه‌های تحقیق، از جامعه شاخص‌های محاسبه شده در بورس تهران، به عنوان نمونه شاخص TEPIX به صورت قضاوتی و بر این اساس که نشان‌دهنده میانگین تغییرات بازار می‌باشد، انتخاب شده و میزان عملکرد مدل‌های پیشنهادی در پیش‌بینی نوسان تحقق‌یافته شاخص کل بورس اوراق بهادار تهران مورد بررسی قرار می‌گیرد.

جهت آزمون نمودن فرضیه‌های تحقیق، نیاز به مقایسه دو مدل پیش‌بینی نوسان وجود دارد. بدین منظور می‌بایست از آزمون‌های زوجی توسعه داده شده استفاده نمود. با نرمال نبودن زوج‌ها و وجود همبستگی در آنها، در این تحقیق از آزمون غیرپارامتری ویلکاکسون²⁴ استفاده شده است.

داده‌های مورد استفاده به منظور تخمین مدل‌ها شامل مقادیر شاخص کل بورس تهران (TEPIX) به مدت ۵ سال (از ابتدای سال ۱۳۸۸ تا انتهای سال ۱۳۹۲) یا ۱۲۰۷ روز در تواتر ۱۵ دقیقه‌ای (۱۵۹۲۷ داده) از ساعت ۹ تا ۱۲ برای داده‌های تا تاریخ ۱۳۹۲/۰۸/۲۸ و از ساعت ۹ تا ۱۲:۳۰ برای سایر داده‌ها است. همچنین داده‌های مورد استفاده به منظور تعیین خطا و بررسی کارایی مدل‌ها شامل مقادیر شاخص کل بورس تهران (TEPIX) به مدت شش ماه (از ابتدای سال ۱۳۹۳ تا انتهای شهریور ماه این سال) یا ۱۲۱ روز در تواتر ۱۵ دقیقه‌ای (۱۶۶۰ داده) از ساعت ۹ تا ۱۲:۳۰ است. کلیه داده‌ها با کمک نرم‌افزار ره‌آورد نوین استخراج شده است.

۴- فرضیه‌های پژوهش

- بر اساس مسائل پیش‌گفته فرضیه‌های زیر برای تحقیق مطرح شده است:
- (۱) در نظر گرفتن پرش در مدل‌سازی نوسان تحقق‌یافته، پیش‌بینی دقیق‌تر نوسان تحقق‌یافته را ممکن می‌سازد.
 - (۲) تجزیه نوسان تحقق‌یافته به دو جزء پرش و پیوستگی، پیش‌بینی دقیق‌تر نوسان تحقق‌یافته را ممکن می‌سازد.

۵- مدل‌های پژوهش

۵-۱- مدل HAR-RV

مدل HAR-RV توسط کرسی (۲۰۰۴) معرفی شد. ایده‌ی اصلی مدل از این مطلب ناشی شده که یک شرکت دارای سهامداران با افق‌های زمانی متفاوت، مانند کوتاه‌مدت، میان‌مدت و بلندمدت است که در مواجهه با شرایط یکسان، به طرق متفاوت واکنش نشان می‌دهند و باعث ایجاد اجزاء مختلف نوسان می‌شوند. این مساله مرتبط با تئوری ناهمگونی بازار است که در سال ۱۹۹۳ توسط مولر و همکارانش ارائه شده است. کرسی نوسان جزئی را به شکل نوسانی که از یک جزء مشخص بازار ناشی می‌شود، تعریف نمود و مدلی معرفی کرد که از مجموع این اجزا حاصل می‌شود. نام این مدل به تناسب نام تئوری استفاده شده، مدل اتورگرسو ناهمگون^{۲۵} است.

مدل HAR-RV به شکل زیر بیان می‌شود:

$$RV_{t,t+1} = \beta_0 + \beta_D RV_{t-1,t} + \beta_W RV_{t-5,t} + \beta_M RV_{t-22,t} + \varepsilon_{t+1} \quad (۱-۵)$$

$RV_t, t = 1, 2, 3, \dots, T$ ، $RV_{t-5,t}$ و $RV_{t-22,t}$ به ترتیب نشان دهنده‌ی نوسان تحقق‌یافته‌ی روزانه، هفتگی و ماهانه هستند. $RV_{t,t+h}$ نشان دهنده‌ی نوسان بین t و $t+h$ است و داریم:

$$RV_{t,t+h} = h^{-1}(RV_{t,t+1} + RV_{t+1,t+2} + \dots + RV_{t+h-1,t+h}) \quad (۲-۵)$$

در این مدل اجزای ایجاد نوسان به شکل نوسانات روزانه، هفتگی و ماهانه‌ی قبلی در نظر گرفته شده‌اند که از مجموع آنها نوسان روز بعد تخمین زده می‌شود.

۵-۲- مدل HAR-RV-J

این مدل توسط اندرسون و همکارانش (۲۰۰۳) با وارد نمودن جزء پرش به مدل HAR-RV ایجاد شده است. در واقع در این مدل، یک متغیر توضیحی به مدل قبل اضافه شده که آن جزء پرش است. مدل HAR-RV-J به شکل زیر نوشته می‌شود:

$$RV_{t,t+1} = \beta_0 + \beta_D RV_{t-1,t} + \beta_W RV_{t-5,t} + \beta_M RV_{t-22,t} + \beta_j j_t + \varepsilon_{t,t+1} \quad (۳-۵)$$

در استفاده عملی از مدل‌های پیش‌بینی نوسان، اغلب به جای واریانس، نیاز به پیش‌بینی انحراف معیار وجود دارد. مدل HAR-RV-J در فرم انحراف استاندارد به شکل زیر نوشته می‌شود:

$$RV_{t,t+1}^{\frac{1}{2}} = \beta_0 + \beta_D RV_{t-1,t}^{\frac{1}{2}} + \beta_W RV_{t-5,t}^{\frac{1}{2}} + \beta_M RV_{t-22,t}^{\frac{1}{2}} + \beta_j j_t^{1/2} + \varepsilon_{t,t+1} \quad (۴-۵)$$

ABDL (۲۰۰۳) برای مدل نمودن واریانس از لگاریتم نوسان تحقق‌یافته استفاده نموده است. دلیل استفاده از این داده‌ها نزدیک‌تر بودن آنها به توزیع نرمال و دقیق‌تر بودن پیش‌بینی انجام شده است. به پیروی از ABDL، فرم لگاریتمی مدل HAR-RV-J به شکل زیر نوشته می‌شود:

$$\log(RV_{t,t+1}) = \beta_0 + \beta_D \log(RV_{t-1,t}) + \beta_W \log(RV_{t-5,t}) + \beta_M \log(RV_{t-22,t}) + \beta_j \log(j_t + 1) + \varepsilon_{t,t+1} \quad (۵-۵)$$

۵-۳ مدل HAR-RV-CJ

اندرسون و همکارانش (۲۰۰۷) مدل HAR-RV-J را با وارد نمودن جزء پیوستگی توسعه دادند و مدل HAR-RV-CJ را ایجاد کردند. اندرسون بیان کرد که می‌توان نوسان تحقق‌یافته را که در سمت راست رابطه‌ی رگرسیون وجود دارد، به دو جزء پرش و جزء مسیر پیوسته‌ی نمونه، تجزیه نمود. طبق این فرض هر مولفه‌ی $RV_{t-1,t}$ ، به دو مولفه‌ی پرش (J) و پیوستگی (C) تجزیه می‌شوند. به گونه‌ای که در صورت وجود پرش طی دوره‌ی مورد محاسبه، مقدار C برابر BPV و در غیر این صورت برابر RV طی آن دوره خواهد بود. این مدل به شکل زیر است:

$$RV_{t,t+1} = \beta_0 + \beta_{CD} C_{t-1,t} + \beta_{CW} C_{t-5,t} + \beta_{CM} C_{t-22,t} + \beta_{jD} j_{t-1,t} + \beta_{jW} j_{t-5,t} + \beta_{jM} j_{t-22,t} + \varepsilon_{t,t+1} \quad (۶-۵)$$

فرم انحراف استاندارد و لگاریتمی مدل HAR-RV-CJ به شکل زیر نوشته می‌شود:

(۷-۵)

$$RV_{t,t+1}^{1/2} = \beta_0 + \beta_{CD}C_{t-1,t}^{1/2} + \beta_{CW}C_{t-5,t}^{1/2} + \beta_{CM}C_{t-22,t}^{1/2} + \beta_{JD}j_{t-1,t}^{1/2} + \beta_{JW}j_{t-5,t}^{1/2} + \beta_{JM}j_{t-22,t}^{1/2} + \varepsilon_{t,t+1}$$

(۸-۵)

$$\log(RV_{t,t+1}) = \beta_0 + \beta_{CD}\log(C_{t-1,t}) + \beta_{CW}\log(C_{t-5,t}) + \beta_{CM}\log(C_{t-22,t}) + \beta_{JD}\log(j_{t-1,t} + 1) + \beta_{JW}\log(j_{t-5,t} + 1) + \beta_{JM}\log(j_{t-22,t} + 1) + \varepsilon_{t,t+1}$$

۶- نتایج پژوهش

در جدول زیر ویژگی‌های آماری فرایندهای RV، C و J که با پردازش داده‌های خام حاصل شده و به منظور مدل‌سازی مورد استفاده قرار گرفته‌اند، آورده شده است.

جدول ۶-۱- تحلیل آماری داده‌های ۱۵ دقیقه‌ای TEPIX

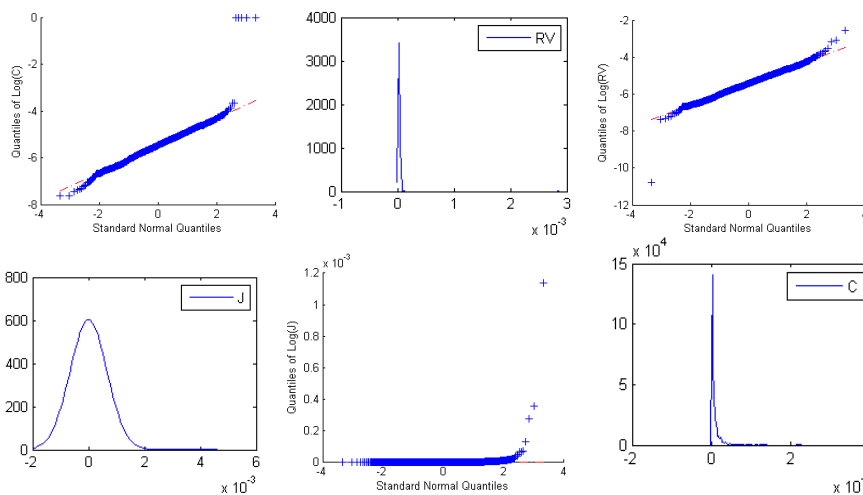
Log(J)	J ^{1/2}	J	Log(C)	C ^{1/2}	C	Log(RV)	RV ^{1/2}	RV	معیار
۰,۰۰۰	۰,۰۰۰	۰,۰۰۰	-۵,۴۷	۰,۰۰۲	۰,۰۰۰	-۵,۴۴	۰,۰۰۲	۰,۰۰۰	میانگین
۰,۰۰۰	۰,۰۰۲	۰,۰۰۰	۰,۶۹۱	۰,۰۰۲	۰,۰۰۰	۰,۶۲۴	۰,۰۰۳	۰,۰۰۰	انحراف معیار
۲۸,۲	۱۴,۴	۲۸,۲	۱,۸۶	۲,۲۸	۷,۳۷	-۰,۴۰۷	۹,۱۲	۲۷,۵	چولگی
۸۷۱	۲۸۲	۸۷۱	۱۸,۱	۱۲,۴	۸۴,۷	۷,۵۶	۱۵۲	۸۴۵	کشیدگی
۰,۰۰۰	۰,۰۰۰	۰,۰۰۰	-۷,۶۴	۰,۰۰۰	۰,۰۰۰	-۱۰,۸	۰,۰۰۰	۰,۰۰۰	کمینه
۰,۰۰۱	۰,۰۵۱	۰,۰۰۳	۰,۰۰۰	۰,۰۱۵	۰,۰۰۰	-۲,۵۵	۰,۰۵۳	۰,۰۰۳	بیشینه
۳۷,۸۹۵	۳,۹۳۰	۳۷,۹۳۷	۱۲	۵	۳۴۶	۱	۱,۱۳۶	۳۵,۶۵۳	جارك برا (X 10 ³)
۰,۱۲۷	۲,۱۰۳	۰,۱۲۶	۳۲۸	۳۷۱	۲۰۷	۴۷۳	۱۱۹	۱۱,۱	لیونگ باکس

(۱) این جدول ویژگی‌های آماری سه فرایند نوسان تحقق یافته (RV)، جزء پیوسته نوسان تحقق یافته (C) و جزء پرش نوسان تحقق یافته (J) را در سه حالت عادی، جذر گرفته شده و لگاریتمی نشان می‌دهد.

(۲) آماره جارك برا (J-B) از توزیع مربع کای با دو درجه آزادی برخوردار بوده و به منظور بررسی این موضوع که داده‌ها از توزیعی نرمال با میانگین و انحراف معیار نامعلوم برخوردار هستند یا خیر، به کار گرفته شده است. در سطح اطمینان ۵٪، مقدار آماره کمتر از ۵,۹۹۱ نشان از نرمال بودن داده‌ها دارد.

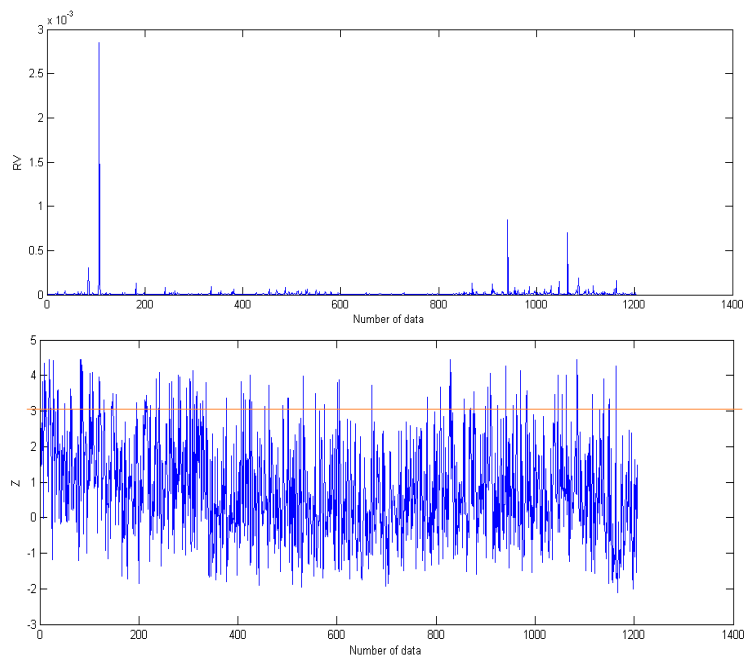
۳) آماره لیونگ باکس وجود همبستگی پیاپی در داده‌ها را تا ۱۰ وقفه آزمون می‌کند. این آماره دارای توزیع مربع کای با ۱۰ درجه آزادی است و مقدار آماره کمتر از ۱۸,۳۰۷ در سطح اطمینان ۵٪، نشان از عدم وجود همبستگی پیاپی در داده‌ها است.

در سه ستون اول مشخصات آماری فرایندهای نوسان تحقیق‌یافته، جذر نوسان تحقیق‌یافته و لگاریتم آن آورده شده است. با مشاهده مقادیر کشیدگی و چولگی، توزیع RV در تواتر ۱۵ دقیقه‌ای، یک توزیع چوله سمت راست با کشیدگی بسیار بیشتر از یک توزیع نرمال است. آماره جاکر برا نیز عدم تطبیق فرایند بر توزیع نرمال را تایید می‌کند. همچنین مقدار آماره لیونگ باکس نشان دهنده عدم وجود همبستگی پیاپی در داده‌ها است. مقادیر موجود در ستون‌های جذری و لگاریتمی مشخص‌کننده این نکته است که داده‌های جذری نسبت به داده‌های اصلی و داده‌های لگاریتمی نسبت به داده‌های جذری به توزیع نرمال نزدیک‌ترند. در سه ستون دوم جدول فوق، با مشاهده مشخصات آماری جزء پیوسته نوسان تحقیق‌یافته (C) مشخص می‌شود که برای هر کدام از داده‌های اصلی، جذری و لگاریتمی، توزیع فرایند در مقایسه با مقادیر متناظر در سه ستون اول، به توزیع نرمال نزدیک‌تر است. در نمودارهای زیر چگالی توزیع نوسان تحقیق‌یافته، جزء پیوسته نوسان و جزء پرش و همچنین نمودارهای کوانتایل‌های توزیع لگاریتم فرایندهای پیش‌گفته در مقایسه با کوانتایل‌های توزیع نرمال آورده شده که موید گفته‌های قبل می‌باشد.

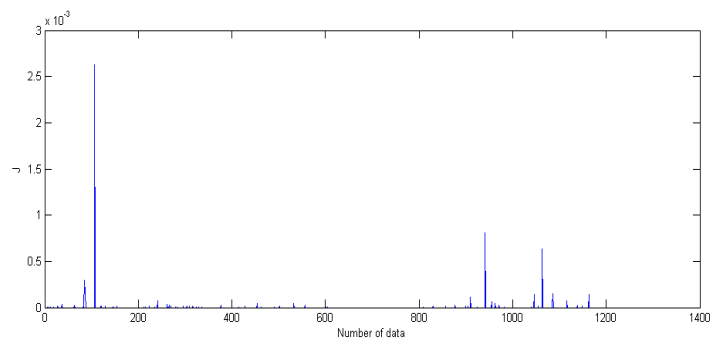


نمودار ۶-۱- چگالی توزیع RV، C و J در سمت چپ و مقایسه کوانتایل‌های توزیع لگاریتم‌ها با کوانتایل‌های توزیع نرمال در سمت راست

نمودارهای زیر، نوسان تحقق یافته (RV)، شناساگر اعلام وجود پرش در داده‌ها در سطح ۹۹,۹٪ اطمینان (Z) و مقدار پرش (J) را برای شاخص کل بورس تهران نشان می‌دهد.

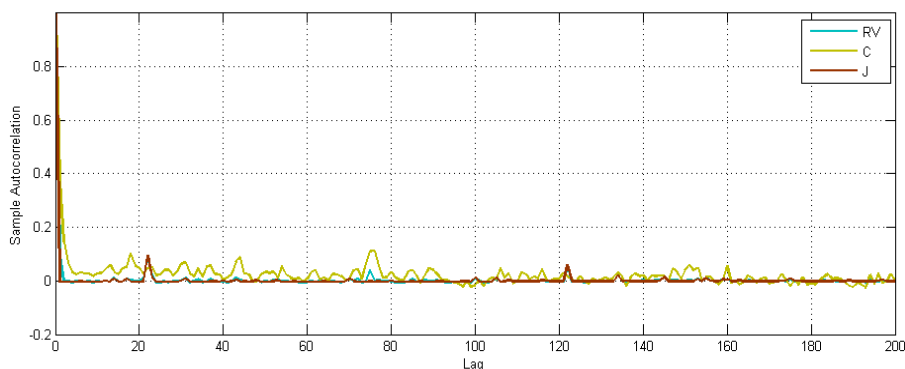


نمودار ۶-۲- نوسان تحقق یافته (RV) و شناساگر اعلام وجود پرش در داده‌ها در سطح ۹۹,۹٪ اطمینان (Z) به ترتیب در نمودارهای اول و دوم



نمودار ۶-۳- شناسایی وقوع پرش و مقدار آن (J)

با مقایسه مقادیر خودهمبستگی که در نمودار زیر نشان داده شده، مشخص است که در سه سری بررسی شده، سری جزء پیوسته نوسان تحقق‌یافته (C) بیشترین خودهمبستگی را داراست.



نمودار ۶-۴- خودهمبستگی سری‌های RV، C و J تا ۲۰۰ وقفه

با در نظر داشتن توضیحات پیش‌گفته، نتایج مدلسازی و تخمین ضرایب با استفاده از روش OLS در جدول (۲-۶) آورده شده است.

جدول ۶-۲- تخمین ضرایب مدل‌ها

HAR RV CJ			HAR RV J			HAR RV			پارامتر	افق زمانی رگرسیون
۳	۲	۱	۳	۲	۱	۳	۲	۱		
-۰٫۹۵۸۷۵	۰٫۰۰۰۵۹	۰٫۰۰۰۰۰	-۱٫۰۶۹۲	۰٫۰۰۰۶۳	-۰٫۰۰۰۰۱	-۱٫۰۶۳۷	۰٫۰۰۰۰۸۳	۰٫۰۰۰۰۱	cons	روزانه
-۰٫۱۶۱۵۵	-۰٫۴۷۹۳۷	۳٫۲۴۴۳	-۰٫۱۵۵۲۰	-۰٫۵۰۸۹۴	۳٫۰۳۱۴	-۰٫۱۵۸۴۱	-۰٫۲۲۰۸۲	۰٫۰۹۹۶۷	b1	
-۰٫۰۸۵۲۷	-۰٫۰۰۰۴۵۰	-۰٫۵۳۲۴۳	۰٫۱۰۸۰۹	-۰٫۱۱۶۲۹	-۰٫۱۴۳۹۷	۰٫۱۰۸۴۵	-۰٫۰۷۹۹۸	-۰٫۰۴۲۶۹	b2	
-۰٫۵۷۱۸۱	۰٫۳۵۱۴۱	-۱٫۱۸۷۲	۰٫۵۳۹۷۸	-۰٫۴۱۱۰۷	-۰٫۰۰۰۶۲۳	-۰٫۵۳۷۱۴	-۰٫۵۲۴۶۷	۰٫۱۰۱۱۳	b3	
۸۸۲٫۵۰	۰٫۰۰۰۳۲	-۰٫۱۹۵۳۷	۳۱۹٫۹۱	-۰٫۳۹۳۱۷	-۳٫۲۰۳۸				b4	
-۱٫۹۶۵٫۹	-۰٫۱۸۳۷۸	-۰٫۰۷۹۱۹							b5	
۱۰۰۱۵٫۸	۰٫۳۰۷۳۸	۰٫۲۶۹۳۰							b6	
-۰٫۱۸۴۵۵	۰٫۱۳۳۱۶	۰٫۲۰۷۷۸	۰٫۱۷۴۹۹	۰٫۱۳۳۸۱	۰٫۲۰۲۶۹	۰٫۱۷۴۸۲	۰٫۰۹۰۷۶	۰٫۰۰۹۵۸	R square	
-۱٫۲۴۵۳	۰٫۰۰۰۹۱	۰٫۰۰۰۰۱	-۱٫۴۱۰۵	۰٫۰۰۰۹۵	۰٫۰۰۰۰۱	-۱٫۴۱۸۶	۰٫۰۰۱۰۳	۰٫۰۰۰۰۱	cons	هفتگی
۰٫۱۰۲۵۵	۰٫۱۵۳۶۸	۰٫۶۹۹۰۲	۰٫۰۸۵۲۸	۰٫۱۷۱۳۰	۰٫۶۵۴۳۳	۰٫۰۸۰۵۷	۰٫۰۵۸۰۹	۰٫۰۲۰۴۶	b1	
-۰٫۰۸۰۷۱	-۰٫۰۴۵۲۳	-۰٫۱۳۷۱۹	۰٫۰۲۳۰۸	-۰٫۱۰۷۸۱	-۰٫۰۶۳۶۰	۰٫۰۲۲۵۵	-۰٫۰۹۳۵۰	-۰٫۰۴۱۷۰	b2	
۰٫۷۴۴۰۶	۰٫۵۵۵۸۹	-۰٫۱۷۶۶۱	۰٫۶۳۱۴۶	۰٫۵۷۹۱۰	۰٫۰۹۴۱۷	۰٫۶۳۵۳۴	۰٫۶۳۳۸۰	۰٫۱۱۷۴۷	b3	
۱٫۵۳۷۰	-۰٫۰۱۷۱۲	-۰٫۰۴۳۱۸	-۳۲۲٫۴۹	-۰٫۱۵۴۳۱	-۰٫۶۹۲۶۸				b4	
-۳۹۴٫۵۵	-۰٫۱۳۱۶۸	-۰٫۰۵۱۴۸							b5	
۱۲۶٫۶۰	۰٫۴۰۰۷۸	۰٫۱۵۶۹۰							b6	
-۰٫۳۴۰۱۳	۰٫۱۵۸۷۷	۰٫۰۴۴۱۴	۰٫۳۱۹۰۹	۰٫۱۵۹۶۵	۰٫۰۴۳۰۹	۰٫۳۱۸۲۹	۰٫۱۳۸۹۵	۰٫۰۰۳۸۸	R square	

HAR RV CJ			HAR RV J			HAR RV			پارامتر	افق زمانی رگرسیون
۳	۲	۱	۳	۲	۱	۳	۲	۱		
-۱,۴۷۶۹	۰,۰۰۱۰۲	۰,۰۰۰۰۱	-۱,۷۸۳۱	۰,۰۰۱۰۸	۰,۰۰۰۰۱	-۱,۷۸۵۹	۰,۰۰۱۱۱	۰,۰۰۰۰۱	cons	
۰,۰۲۹۶۰	۰,۰۴۶۶۵	۰,۱۶۳۵۹	۰,۰۳۳۷۵	۰,۰۵۱۹۱	۰,۱۴۴۹۴	۰,۰۳۲۰۸	۰,۰۱۶۴۸	۰,۰۰۴۱۲	b1	
۰,۰۱۱۹۲	-۰,۰۱۴۷۲	-۰,۱۹۷۷۷	۰,۰۵۳۱۷	-۰,۰۱۰۱۷	-۰,۰۰۷۸۳	۰,۰۵۲۹۸	-۰,۰۰۵۶۷	-۰,۰۰۲۹۷	b2	
۰,۶۸۱۳۰	۰,۵۸۲۲۷	۰,۴۳۲۴۵	۰,۵۸۴۰۳	۰,۵۳۴۱۶	۰,۰۸۳۶۵	۰,۵۸۵۴۰	۰,۵۴۸۱۸	۰,۰۸۸۸۴	b3	
۴۶,۳۴۴	-۰,۰۰۵۷۵	-۰,۰۱۰۶۱	-۱۱۳,۵۳	-۰,۰۴۸۱۷	-۰,۱۵۳۸۷				b4	
۱۱۸,۳۳	-۰,۰۰۳۱۶	۰,۰۲۶۴۸							b5	
-۱۰,۰۱۶,۵	۰,۲۸۷۱۲	۰,۰۱۰۶۵							b6	
۰,۵۱۴۲۸	۰,۳۱۹۷۶	۰,۰۲۰۲۷	۰,۴۵۷۳۶	۰,۳۱۹۹۹	۰,۰۱۶۸۲	۰,۴۵۷۱۹	۰,۳۱۴۶۸	۰,۰۰۸۳۲	R square	

- (۱) این جدول تخمین OLS را برای پیش‌بینی نوسان با استفاده از سه مدل HAR-RV-J, HAR-RV و HAR-RV-CJ برای افق‌های پیش‌بینی روزانه، هفتگی و ماهانه و سه نوع داده عادی، جذر گرفته شده و لگاریتمی در تواتر میان‌روزانه ۱۵ دقیقه‌ای نشان می‌دهد.
- (۲) برای هر مدل عدد ۱ به معنی استفاده از داده‌های عادی، عدد ۲ به معنی استفاده از داده‌های جذر گرفته شده و عدد ۳ به معنی استفاده از داده‌های لگاریتمی است.
- (۳) رنگ‌های سبز، آبی و نارنجی معنی‌داری پارامترهای تخمینی را در سطوح ۱٪، ۵٪ و ۱۰٪ نشان می‌دهد.
- (۴) مقادیر R^2 برای هر مدل با رنگ خاکستری مشخص شده است.

- با دقت در نتایج مدل‌سازی با استفاده از سری‌های محاسبه شده با تواتر ۱۵ دقیقه، مشاهده می‌شود که:
- با اضافه کردن جزء پرش به مدل، R^2 مدل اندکی افزایش می‌یابد. اندرسن و همکاران (۲۰۰۷) با اضافه نمودن جزء پرش شاهد افزایش چشمگیر سطح انطباق تخمین بر داده‌ها بوده‌اند.
 - استفاده از داده‌های لگاریتمی باعث کسب سطح انطباق بیشتر مدل با داده‌ها می‌شود.
 - مقادیر ثابت در تمامی مدل‌ها به صفر نزدیک است.
 - ضرایب J در اکثر موارد منفی هستند. این موضوع نشان می‌دهد که همانند نتیجه‌گیری اندرسن، در صورتی که فرض شود نوسان تحقق‌یافته تماماً ناشی از پرش است، سری پرش‌ها قابلیت استفاده برای پیش‌بینی آینده را ندارد. همچنین پارامترهای مربوط به پرش در بیشتر موارد معنی‌دار نیستند که این موضوع نیز موید نتیجه پیش‌گفته می‌باشد.
 - در بین تمامی ۲۷ مدل تخمین زده شده در تواتر ۱۵ دقیقه‌ای، مدل HAR-RV-CJ با افق پیش‌بینی ماهانه و با داده‌های لگاریتمی، بیشترین R^2 را داراست. این نشان می‌دهد که مدل خطی در افق ماهانه، دارای بهترین تطابق با داده‌ها است.
 - در تمامی مدل‌ها سطح معنی‌داری پارامترهای تخمین زده شده برای داده‌های جذر گرفته شده بیشتر از داده‌های عادی و برای داده‌های لگاریتمی بیشتر از داده‌های جذر گرفته شده است.

در ادامه تحقیق با استفاده از مدل‌های تخمینی، نوسان تحقق یافته برای شش ماهه آینده با افق‌های روزانه، هفتگی و ماهانه پیش‌بینی شده و خطای آن با استفاده از معیار MAPE محاسبه شده است. جدول (۳-۶) معیار MAPE به منظور اندازه‌گیری خطا برای تمامی مدل‌های تخمین زده شده در افق زمانی روزانه، هفتگی و ماهانه را نشان می‌دهد.

جدول ۳-۶- معیار MAPE محاسبه شده برای سه مدل پیش‌بینی

HAR-RV-CJ	HAR-RV-J	HAR-RV	ردیف
۷,۰۵۲	۹,۹۳۰	۱۴,۶۷۸	۱
۵,۲۰۰۸۴	۴,۶۷۹۴	۶,۰۲۶۸	۲
۳,۳۸۴۴	۳,۵۰۰۰	۳,۶۸۰۸	۳
۰,۹۱۹۲۰	۰,۹۲۶۲۸	۰,۹۷۲۴۳	۴
۰,۵۹۲۷۷	۰,۵۹۷۳۰	۰,۶۰۵۰۰	۵
۰,۴۵۴۸۱	۰,۴۶۲۶۲	۰,۴۶۳۹۳	۶
۰,۰۷۸۶۱	۰,۰۷۸۱۰	۰,۰۷۸۰۹	۷
۰,۰۴۹۶۶	۰,۰۴۹۲۵	۰,۰۴۹۲۸	۸
۰,۰۴۰۶۴	۰,۰۴۰۸۸	۰,۰۴۰۸۶	۹

(۱) این جدول معیار MAPE برای اندازه‌گیری خطای پیش‌بینی با استفاده از سه مدل HAR-RV، HAR-RV-CJ و RV-J را نشان می‌دهد.

(۲) داده‌های مشخص شده با خط زیر، کمترین مقدار MAPE در هر سطر را دارا می‌باشند.

با دقت در نتایج به نظر می‌رسد که برای تمامی نمونه‌های پیش‌بینی، کارایی مدل‌های HAR-RV-J و HAR-RV-CJ بهتر از مدل HAR-RV می‌باشد. برای آزمون معناداری تفاوت مشاهده شده در نتایج سه مدل، بار دیگر آزمون ویلکاکسون به کار گرفته می‌شود. نتایج آزمون در جدول (۴-۶) نشان داده شده است.

جدول ۴-۶- نتایج آزمون ویلکاکسون برای تعیین مدل با خطای پیش‌بینی کمتر

P-Value	مقدار آماره	مدل دوم	مدل اول	ردیف
۰,۰۲۱	-۲,۳۱۰	HAR-RV-J	HAR-RV	۱
۰,۰۳۸	-۲,۰۷۳	HAR-RV-CJ	HAR-RV	۲
۰,۲۶	-۱,۱۲۵	HAR-RV-CJ	HAR-RV-J	۳

در آزمون‌های اول و دوم، کمتر بودن مقادیر P-Value محاسبه شده از سطح اطمینان ۵٪، نشان می‌دهد که خطای پیش‌بینی (MAPE) دو مدل HAR-RV-J و HAR-RV-CJ به شکل معناداری با خطای پیش‌بینی مدل HAR-RV اختلاف دارد. در واقع دو آزمون اول، به ترتیب موید فرضیه‌های اول و دوم تحقیق بوده که بیان می‌کند اضافه نمودن جزء پرش به مدل و تجزیه نوسان تحقق‌یافته به دو جزء پیوسته و پرش، باعث کاهش خطا و پیش‌بینی دقیق‌تر نوسان تحقق‌یافته می‌شود. آزمون سوم نیز بیان می‌کند که تجزیه نوسان تحقق‌یافته به دو جزء پیوسته و پرش (مدل HAR-RV-CJ)، نسبت به مدل HAR-RV-J از لحاظ آماری تاثیر معنی‌داری در عملکرد پیش‌بینی ندارد.

۷- نتیجه‌گیری و بحث

در این پژوهش به معرفی یکی از روش‌های مدل‌سازی نوسان تحقق‌یافته پرداخته شد. در این روش با در نظر گرفتن یک فرایند وینر همراه با پرش برای نوسان و استفاده از نتایج تحقیقات کرسی (۲۰۰۴)، بارندورف، نیلسون، شفرد (۲۰۰۴) و اندرسن و همکارانش (۲۰۰۷)، مدل‌های HAR-RV، HAR-RV-J و HAR-RV-CJ به منظور مدل‌سازی و پیش‌بینی نوسان تحقق‌یافته معرفی شدند. سپس با به کارگیری داده‌های میان‌روزانه شاخص کل بورس اوراق بهادار تهران در تواتر ۱۵ دقیقه‌ای، پارامترهای مدل‌ها با استفاده از روش OLS تخمین زده شد و با استفاده از مدل تخمینی، نوسان تحقق‌یافته برای دوره‌ای شش ماهه و در افق‌های روزانه، هفتگی و ماهانه مورد پیش‌بینی قرار گرفت. در ادامه با استفاده از معیار MAPE خطای پیش‌بینی محاسبه شده و با دقت در نتایج و انجام آزمون ویلکاکسون، درستی فرضیه‌های تحقیق بررسی شد.

نتایج تحقیق عبارت است از (الف) در تواتر ۱۵ دقیقه‌ای، هر دو مدل HAR-RV-J و HAR-RV-CJ نسبت به مدل HAR-RV خطای پیش‌بینی کمتری داشته و دارای کارایی بهتری هستند. در واقع اضافه نمودن جزء پرش و یا تجزیه نوسان تحقق‌یافته، باعث افزایش کارایی نسبت به مدل اولیه می‌شود. این نتایج با نتایج کسب شده توسط اندرسن و همکارانش (۲۰۰۳ و ۲۰۰۷) منطبق است. (ب) مدل‌هایی که با سری‌های جذری تخمین زده شده‌اند، نسبت به مدل‌های عادی و مدل‌هایی که با سری‌های لگاریتمی تخمین زده شده‌اند نسبت به مدل‌های جذری از سطح انطباق (R^2) بیشتری برخوردار هستند. (ج) در تمامی مدل‌ها سطح معنی‌داری پارامترهای تخمین زده شده برای داده‌های جذر گرفته شده بیشتر از داده‌های عادی و برای داده‌های لگاریتمی بیشتر از داده‌های جذر گرفته شده است. (د) با اضافه کردن جزء پرش به مدل، R^2 مدل اندکی افزایش می‌یابد. اندرسن و همکاران (۲۰۰۳) با اضافه نمودن جزء پرش شاهد افزایش چشمگیر سطح انطباق تخمین بر داده‌ها بوده‌اند.

فهرست منابع

- * Andersen, T.G. and T.Bollerslev (1997). Heterogeneous information arrivals and return volatility dynamics: uncovering the long-run in high frequency returns. *Juornal of finance*, Vol.52, 975-1005.
- * Andersen, T.G., T. Bollerslev, F.X. Diebold and H. Ebens (2001). The Distribution of Realized Stock Return Volatility. *Journal of Financial Economics*, 61, 43-76.
- * Andersen, T.G., Benzoni, Luca, Lund, Jesper (2002). Estimating jump-diffusions for equity returns. *J. Financ.* 57, 1239–1284.
- * Andersen, Torben G., Bollerslev, Tim, Diebold, Francis X., (2003a). Some like it smooth and some like it rough: untangling continuous and jump components in measuring modelling and forecasting asset return volatility. *CFS Working Paper*, No: 2003/35.
- * Andersen, Torben G., Bollerslev, Tim, Diebold, Francis X., Labys, Paul, (2003b). Modeling and forecasting realized volatility. *Econometrica* 71, 579–625.
- * Andersen, T.G., T. Bollerslev, F.X. Diebold and P. Labys (2001). The Distribution of Realized Exchange Rate Volatility. *Journal of the American Statistical Association*, 96, 42-55.
- * Andersen, T.G., T. Bollerslev, F.X. Diebold and P. Labys (2003). Modeling and Forecasting Realized Volatility. *Econometrica*, 71, 579-625.
- * Andersen, T.G., T. Bollerslev and F.X. Diebold (2005). Parametric and Non-Parametric Volatility Measurement. In *Handbook of Financial Econometrics* (L.P Hansen and Y. Aït-Sahalia, Eds.). Elsevier Science, New York, forthcoming.
- * Andersen, Torben G., Tim, Bollerslev, Diebold, Francis X., (2007). Roughing it up: including jump components in the measurement, modeling and forecasting of return volatility. *Rev. Econ. Stat.* 89 (4), 701–720.
- * Barndorff-Nielsen, O.E. and N. Shephard (2002a), “Econometric Analysis of Realised Volatility and its use in Estimating Stochastic Volatility Models,” *Journal of the Royal Statistical Society*, 64, 253-280.
- * Barndorff-Nielsen, O.E. and N. Shephard (2002b), “Estimating Quadratic Variation Using Realized Variance,” *Journal of Applied Econometrics*, 17, 457-478.
- * Barndorff-Nielsen, O.E. and N. Shephard (2004a), “Power and Bipower Variation with Stochastic Volatility and Jumps,” *Journal of Financial Econometrics*, 2, 1-37.
- * Barndorff-Nielsen, O.E. and N. Shephard (2004b), “How Accurate is the Asymptotic Approximation to the Distribution of Realised Volatility,” in *Identification and Inference for Econometric Models. A Festschrift in Honour of T.J. Rothenberg* (D. Andrews, J. Powell, P.A. Ruud, and J. Stock, Eds.). Cambridge, UK: Cambridge University Press.
- * Bollerslev, T. (1986). Generalized Autoregressive Conditonal Heteroskedasticity. *Journal of Econometrics* 31 (1986) 307-327. North-Holland.
- * Çelik, Sibel and Ergin, Hüseyin (2014). Volatility forecasting using high frequency data: Evidence from stock markets. *Economic Modelling*, 36, 176–190.
- * Chung, H-M, Huang, C-S, and Tseng, T-C. (2008). The performance of the heterogeneous autoregressive model in the forecasting of the realized volatility of individual stocks. Working paper, National Chiao Tung University.
- * Chung, H-M, Huang, C-S, and Tseng, T-C. (2008). Modeling and forecasting of realized volatility based on high frequency data: evidence from Taiwan. *International Research Journal of Finance and Economics*, Vol.22, 178-191.
- * Comte, F. and Renault, E. (1998). Long memory in ontinuous-time stochastic0020 volatility models. *Mathematical Finance*, Vol.8, 291-323.

- * Corsi, Fulvio (2004). A Simple Long Memory Model of Realized Volatility. (Manuscript) University of Southern Switzerland.
- * Engle, Robert F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, Vol. 50, No. 4. (Jul., 1982), pp. 987-1007.
- * Huang, X. and Tauchen, G. (2005). The relative contribution of jumps to total price. Variance Manuscript, Duke University.
- * McAleer, M. and Medeiros, M. C. (2008). A multiple regime smooth transition Heterogeneous Autoregressive model for long memory and asymmetries. *Journal of Econometrics*, Vol.147, 104-119.
- * Merton, R. C. (1980). On estimating the expected return on the market: An exploratory investigation. *Journal of Financial Economics*, Vol.8, 323-361.
- * Müller, U., Dacorogna, M., Dave, R., Pictet, O., Olsen, R. and Weizsacker, J. (1997). Volatilities of different time resolutions - analyzing the dynamics of market components. *Journal of Empirical Finance*. Vol.4, 213-239.
- * Protter, P. (1990). *Stochastic integration and differential equations: A new approach*. Berlin Heidelberg, Germany: Springer-Verlag.

یادداشت‌ها

1. Value at risk
2. Engle
3. Bollerslev
4. Andersen and Bollerslev 'Andersen et al, 2002b
5. Jump
6. Protter
7. Merton
8. Andersen, Bollerslev and Diebold
9. Heterogeneous Autoregressive of Realized Volatility Model
10. Corsi
11. Muller et al
12. Barndorff, Nielsen and Shephard
13. McAleer and Medeiros
14. HARST
15. Chung, Huang and Tseng
16. Mixture of distributions hypothesis
17. Çelik and Ergin
18. Quadratic variation
19. Comte and Renault
20. Realized power variation
21. Realized bi-power variation
22. Huang and Tauchen
23. Mean Absolute Percent Error
24. Wilcoxon
25. Heterogeneous Autoregressive