



فصلنامه دانش سرمایه‌گذاری
سال اول / شماره دوم / تابستان ۱۳۹۱

مقایسه کارآمدی مدل‌های ARIMA و ARFIMA برای مدل‌سازی و پیش‌بینی شاخص قیمت تهران (TEPIX)

حبیب‌اله سالارزهی

عضو هیئت علمی دانشگاه تهران

منصور کاشی

دانش آموخته‌ی کارشناسی ارشد مدیریت بازرگانی - مالی دانشگاه سیستان و بلوچستان (مسئول مکاتبات)

mansoor.kashi@gmail.com

سیدحسن حسینی

دانش آموخته‌ی کارشناسی ارشد مدیریت بازرگانی - مالی دانشگاه سیستان و بلوچستان

محمد دنیایی

کارشناس ارشد رشته مدیریت بازرگانی - مالی، عضو باشگاه پژوهشگران جوان، دانشگاه آزاد اسلامی واحد زاهدان

تاریخ دریافت: ۹۱/۲/۳۰ تاریخ پذیرش: ۹۱/۴/۳۰

چکیده

این مقاله به بررسی عملکرد پیش‌بینی مدل‌های ARIMA و ARFIMA با استفاده از داده‌های روزانه بازده شاخص کل سهام تهران در بازه زمانی ۱۳۸۰/۰۹/۰۴ تا ۱۳۹۰/۰۹/۰۹ می‌پردازد. در این راستا جهت تخمین پارامتر d و دیگر پارامترها، از روش NLS در بسته نرم‌افزار Oxmetric/pcgive استفاده شد و پس از مقایسه نتایج مدل‌های تحقیق؛ مدل ARFIMA بر اساس معیار AIC مدلی برتر در مدل‌سازی TEPIX مشخص گردید. همچنین از میان برآوردهای پیش‌بینی، روش‌های پیش‌بینی ساده را برای تخمین پیش‌بینی آزمون می‌کنیم. از مقایسه دقت پیش‌بینی مدل‌های مذکور توسط معیارهای پیش‌بینی مانند MAPFE و RMSFE و فواصل اطمینانی که ارزش‌های واقعی در آن جای گرفته‌اند، می‌توان استنباط کرد که اولاً، تفاوت عملکرد بهتر پیش‌بینی مدل حافظه بلند مدت ARFIMA نسبت به مدل ARIMA بسیار جزئی است و ثانیاً، ناکارآمدی مدل ARFIMA در پیش‌بینی بازار سرمایه تهران کاملاً مشهود است.

واژه‌های کلیدی: پیش‌بینی، بازده، خود رگرسیون میانگین متحرک انباشته، خود رگرسیون میانگین متحرک انباشته جزئی.

۱- مقدمه

نگرش کلاسیک بر حسب مدل سازی سری های زمانی اقتصادی بنا بر کاربرد نگرش باکس جنکیس^۱ به ایستایی و غیر ایستایی سری های زمانی بستگی دارد. اگر چنین سری هایی خصوصیت حافظه بلند مدت را نشان دهند ارزش پیش بینی بر اساس مدل های خود رگرسیون میانگین متحرک (ARMA)^۲ و خود رگرسیون میانگین متحرک انباشته (ARIMA)^۳ معتبر نخواهد بود (Man, 2003). در صورت وجود حافظه بلند مدت در سری های زمانی، بین مشاهدات سری در فواصل بسیار زیاد، جدا و دور از هم وابستگی معنی داری وجود خواهد داشت که نمایانگر این موضوع است که مشاهدات مستقل از هم نبوده، همبستگی بین آنها وجود دارد و مشاهدات گذشته به پیش بینی داده ها کمک خواهند کرد. با این توصیفات وجود حافظه بلند مدت در بازارهای مالی شکل ضعیف کارایی بازار را نقض کرده و آنگاه تغییرات در بازار سرمایه تصادفی نبوده و قابل پیش بینی خواهد بود (Beran, 1995). همچنین در مدل خطی قیمت گذاری تردید ایجاد می کند و بیانگر آن است که در قیمت گذاری بایستی از مدل های غیر خطی استفاده کرد (محمودی، محمدی و چیت‌سازان، ۱۳۸۹). مدل حافظه بلند مدت را اولین بار هارست^۴ (۱۹۵۱) بر خلاف سایر محققین آب شناسی^۵ که جریان رود را فرایندی تصادفی می دانستند بیان نمود و تحلیل دامنه استاندارد شده (R/S)^۶ را طراحی کرد. مندلیبروت و والیس^۷ (۱۹۶۹) برای یافتن ماندگاری^۸ در سری های زمانی مالی R/S را براساس نگرش گرافیکی ابداعی^۹ توسعه دادند. در اوایل دهه ی ۱۹۸۰ گرنجر و جویکس^{۱۰} نگرشی جایگزین را برای مدل سازی حافظه بلند مدت با ایجاد مدل خود رگرسیون میانگین متحرک انباشته جزئی (ARFIMA)^{۱۱} ارائه دادند و از آن جایی که مدل ARFIMA بین فرایند حافظه کوتاه مدت^{۱۲} و فرایند حافظه ی بلند^{۱۳} در سری زمانی مالی تمایز قائل می شود مزیت ممتازی را نسبت به تجزیه و تحلیل کلاسیک R/S که گرایش بسیار زیادی به پذیرش فرض صفر حافظه بلند مدت با وجود فرایند حافظه کوتاه مدت را دارد ایجاد می کند. بطور دقیق تر با وجود ماندگاری در فرایند های ایستا، شوک ها با نرخ کاهش هیپربولیک^{۱۴}، کندتر (با میل بیشتر) از نرخ نمایی فرایند ایستا و معکوس پذیر^{۱۵} یا ARIMA(p,0,q) محو می شوند (Geweke, Porter, 1983). بنابراین مدل ARFIMA در موقعیت میانی بین مدل ARMA و ARIMA قرار می گیرد با نظر به اینکه، ویژگی آن دو مدل را نیز در بر دارد.

در دو دهه گذشته مطالعات اقتصاد مالی زیادی هم از نظر چشم انداز تئوریک و هم از نظر چشم انداز تجربی در حوزه حافظه بلند مدت انجام پذیرفته است. از نظر چشم انداز تئوریک علاوه بر مطالعات مندلیبروت و والیس (۱۹۶۹)، گرنجر و جویکس (۱۹۸۰) به مطالعاتی چون هاسکینگ^{۱۶}

(۱۹۸۱)، کوپک و پورتر-هوداک^{۱۷} (۱۹۸۳)، لو^{۱۸} (۱۹۹۱)، سوول^{۱۹} (۱۹۹۲)، چونگ و دایبولد^{۲۰} (۱۹۹۴)، رابینسون^{۲۱} (۱۹۹۵)، انگل و اسمیت^{۲۲} (۱۹۹۹)، دایبولد و اینو^{۲۳} (۲۰۰۱) و دیتمن و گرنجر^{۲۴} (۲۰۰۲) در زمینه‌ی اقتصاد مالی می‌توان اشاره داشت که بیشتر تمرکز آن‌ها بر تست و تخمین چنین مدل‌هایی بود. در حوزه تئوریک‌ی اخیراً، دورنیک و اوومس^{۲۵} (۲۰۰۳) رویه پایه راست نمایی^{۲۶} را بهبود بخشیدند و سپس آن‌ها در سال ۲۰۰۴ با تست این روش بهبود یافته، تورم را در انگلیس و امریکا مدل‌سازی کرده و نتیجه موفقیت‌آمیزی را بدست آوردند (Dittman, Granger, 2002). از نظر چشم انداز تجربی هم به مطالعاتی چون دایبولد و رودبوسج^{۲۷} (۱۹۹۱)، هاسلر و والترس^{۲۸} (۱۹۹۵)، هیونگ و فرانسس^{۲۹} (۲۰۰۱)، چیو و ذیوت^{۳۰} (۲۰۰۲) می‌توان اشاره کرد. از جمله مطالعاتی که به بررسی مقایسه عملکرد پیش‌بینی‌های ARIMA و ARFIMA پرداخته‌اند، مطالعه‌ی مان^{۳۱} (۲۰۰۳) می‌باشد که کارایی پیش‌بینی مدل ARIMA را در کوتاه مدت مناسب تر از مدل ARFIMA دریافت و برای پیش‌بینی‌های بلند مدت مدل ARFIMA نتیجه بهتری در پی داشت. بهاردواج و سوانسون^{۳۲} (۲۰۰۴) دریافتند که مدل ARFIMA در پیش‌بینی‌های بلند مدت بهتر از مدل ARIMA عمل می‌کند. شیتو و یایا^{۳۳} (۲۰۰۹) با استفاده از روش یک مرحله ایی به تخمین پارامتر کسری d و مقایسه آن با مدل ARIMA پرداخته است که نتایج، عملکرد بهتر پیش‌بینی مدل ARFIMA را تایید کردند. عرفانی و صمیمی (۲۰۰۹)، با استفاده از روش دو مرحله ایی مدل ARFIMA را برآورد کردند و نتایج پیش‌بینی، از عملکرد بهتر مدل ARFIMA در مقابل مدل ARIMA به شدت حمایت کردند. مشیری و مروت (۱۳۸۵)، چندین مدل را برای پیش‌بینی بازده شاخص کل سهام تخمین زد که در این میان مدل ARFIMA در بین سایر مدل‌های اقتصاد سنجی مانند ARMA و GARCH توانست پیش‌بینی بهتری انجام دهد اما در کل مدل شبکه عصبی پیش‌بینی بهتری ارائه داد. در مطالعه شعراپی و ثنایی اعلم (۱۳۸۹)، مدل ARMA عملکرد بهتری برای پیش‌بینی بازده یک روز بعد شاخص نسبت به مدل‌هایی که حافظه بلند مدت را در نظر می‌گیرند نشان داد، اما در پیش‌بینی بازده شاخص برای دوره‌های هفتگی، ماهانه، فصلی و شش ماهه، مدل واریانس شرطی FIGARCH که حافظه بلند مدت واریانس شرطی را در نظر می‌گیرد، همواره پیش‌بینی‌های دقیق‌تری ارائه کرده است.

در این مقاله به بررسی عملکرد پیش‌بینی مدل‌های ARIMA و ARFIMA با استفاده از بازده داده‌های شاخص قیمت تهران (TEPIX)^{۳۴} خواهیم پرداخت و برای دقت مقایسه پیش‌بینی مدل‌های مذکور از طریق معیارهای ریشه میانگین توان دوم خطاهای پیش‌بینی (RMSFE)^{۳۵} و درصد میانگین قدر مطلق خطاهای پیش‌بینی (MAPFE)^{۳۶} به تحلیل نتایج می‌پردازیم. برای ادامه، ساختار مقاله به این شرح خواهد بود، بخش دوم مرور تئوری و روش‌ها را پوشش می‌دهد، در

بخش سوم به تحلیل اولیه داده‌ها خواهیم پرداخت، بخش چهارم را به برازش مدل‌ها، نتایج مدل و پیش‌بینی‌های صورت‌گرفته‌ی مدل‌ها اختصاص می‌دهیم و نتیجه‌گیری کلی را در بخش پنجم ارائه می‌نمائیم. تجزیه و تحلیل‌ها و نمودارها در این مطالعه بر اساس برنامه Oxmetric/pcgive (Diebold & Inoue, 2004) و Eviews 6 انجام پذیرفته است.

۲- مبانی نظری تحقیق و مروری بر تئوری و روش‌ها

۲-۱- مدل ARIMA

مدل‌های ARMA و ARIMA، مدل‌های خطی-تصادفی سری‌های زمانی در اقتصاد سنجی می‌باشند که برای سری‌های زمانی مانا و نامانا کاربرد دارند. از جمله روش‌های تبدیل سری نامانا به مانا استفاده از روش تفاضل‌گیری است. برای پیش‌بینی با مدل ARIMA(p,d,q) از متدولوژی باکس-جنکیس استفاده می‌شود که دارای چهار مرحله شناسایی، برآورد، آزمون کنترل تشخیصی و پیش‌بینی می‌باشند. برای مطالعه روش‌ها و رویه‌ها و توصیفات مربوط به این گونه مدل‌ها، می‌توان نوشتجات باکس-جنکیس^{۳۷} (۱۹۷۶)، بریلینگر^{۳۸} (۱۹۸۱)، چات‌فیلد^{۳۹} (۱۹۷۵)، را بررسی کرد.

۲-۲- مدل ARFIMA

طبق مطالعه اوومس و دورنیک^{۴۰} [۱۶ و ۱۷]، مدل ARFIMA(p,d,q) بصورت رابطه زیر تعریف می‌شود:

$$\begin{aligned} \Phi(L)(1-L)^d(y_t - m_t) &= \Theta(L)\varepsilon_t \\ \Phi(L) &= (1 - \phi_1 L - \dots - \phi_p L^p) \\ \Theta(L) &= (1 + \theta_1 L + \dots + \theta_q L^q) \\ t &= 1, 2, \dots, T \end{aligned} \quad (1)$$

که θ پارامتر مدل میانگین، ϕ پارامتر مدل خود همبسته، q مرتبه مربوط به مدل میانگین متحرک، p مرتبه مربوط به مدل خود همبسته، L عملگر وقفه، y_t مقادیر مشاهده شده و m_t میانگین سری زمانی می‌باشد. جمله خطای خالص ε_t متغیری تصادفی با توزیع نرمال با میانگین صفر و واریانس واحد فرض شده است در حالی که پارامتر d عددی واقعی^{۴۱} است به طوری که $(1-L)^d$ نمایانگر پارامتر تفاضل‌گیری جزئی مدل می‌باشد که با استفاده از یک سری دو جمله‌ای بصورت زیر قابل تجزیه است:

$$(1-L)^d = \sum_{k=0}^{\infty} \binom{d}{k} (-L)^k = 1 - dL + \frac{d(d-1)}{2!} L^2 - \dots \quad (2)$$

همچنین فرض شده است $z_t = (y_t - m_t)$ ، ریشه‌های $\Phi(L) = 0$ و $\Theta(L) = 0$ خارج از دایره واحد می‌باشند و ریشه‌های مشترکی ندارند. بعلاوه قسمت ARMA مدل ARFIMA(p,d,q) ایستا و معکوس پذیر است. همانطور که پارامترهای p و q انعطاف بیشتری به خصوصیات کوتاه مدت مدل سازی می‌دهد، پارامتر d رفتار z_t را تعیین می‌کند و در این حالت خصوصیات سری زمانی z_t وابستگی زیادی نسبت به پارامتر d از خود بروز می‌دهد. اگر $-0.5 < d < 0.5$ باشد، فرایند z_t دارای حالت ایستایی کواریانس است و بدین ترتیب $0 < d < 0.5$ دلالت بر فرایند حافظه بلند مدت دارد و $0.5 < d < 1$ حالت برگشت به میانگین را بدون در نظر گرفتن ایستایی کواریانس برای z_t نمایان می‌سازد. زمانی که $0 < d < 0.5$ باشد در وقفه‌های بالا کاهش هیپربولیکی در تابع اتوکواریانس آشکار می‌شود. چنانچه γ_k نشان دهنده ی تابع اتوکواریانس^{۴۲} باشد کاهش هیپربولیکی در معادله ۳ به خوبی نمایان می‌شود به شرطی که $k \rightarrow \infty$ و $c > 0$ باشد.

$$\gamma_k \sim ck^{2d-1} \quad (3)$$

به ازای $-0.5 < d < 0$ مجموعه قدر مطلق خود همبستگی فرایند طبق معادله ۴ به یک مقدار ثابت میل می‌کند و ویژگی حافظه کوتاه مدت را برای سری زمانی z_t نمایان می‌سازد. همه خود همبستگی‌های این فرایند بجز وقفه ی صفر، منفی و کاهش هیپربولیکی به صفر دارند. در این حالت ARFIMA(0,d,0) ویژگی حافظه میانه^{۴۳} را آشکار می‌سازد (شعراپی و ثنایی اعلم، ۱۳۸۹).

$$\lim_{n \rightarrow \infty} \sum_{j=-n}^n |\rho_j| \quad (4)$$

اوداکی^{۴۴} (۱۹۹۳) معکوس پذیری را در محتوای خطای مربع میانگین پیش بینی کننده خطی بواسطه ی همگرایی مرتبه متناهی AR بر حسب تغییرات واریانس مطرح کرد و نشان داد که $d > 1$ شرط کافی برای معکوس پذیری فرایند z_t می‌باشد.

۲-۳- تخمین پارامتر d: نگرش پارامتریک مدل ARFIMA

تکنیک‌های تخمین پارامتر d مانند تکنیک جوک و پورتر - هوداک (GPH)^{۴۵}، تحلیل نوسانات روند زدایی (DFA)^{۴۶}، تخمین زن موجک^{۴۷} و تکنیک‌هایی از این نوع، شامل رویه تخمین دو

مرحله ایی می باشند. این نوع تکنیک ها در گام اول پارامتر تفاضل گیری را تخمین می زنند و در گام بعدی پارامتر تفاضل گیری تخمین زده به یک سری زمانی که احتمالاً از مدل ARMA تبعیت می کند، منتقل می شود. از محدودیت چنین مدل هایی این است که اطلاعات را در فرکانس های پایین بکار می برند بطوریکه هنگام تخمین پارامتر d خصوصیات کوتاه مدت سری را بررسی نمی کنند. برای رفع محدودیت این نوع تکنیک ها، بررسی سری های زمانی را تحت فرایند ARFIMA که توسط گرنجر و جوینکس (۱۹۸۰) و هاسکینگ (۱۹۸۱) معرفی شده اند و شامل رویه تخمین یک مرحله ایی هستند انجام می دهیم. این نوع فرایند پویای بلند مدت را بواسطه ی پارامتر انباشته کسری d بررسی می کند در حالی که اجزای سنتی AR و MA پویایی های کوتاه مدت سری زمانی را در بر می گیرند. لازم به ذکر است که در این نوع فرایند، تخمین d ، AR و MA به طور همزمان صورت می گیرد. در مباحث اقتصادسنجی از جمله روش های برآورد مدل ARFIMA، برآورد حداکثر راست نمایی دقیق (EML)^{۴۸} و حداقل مربعات خطی (NLS)^{۴۹} می باشند که به ترتیب از سوی سوول^{۵۰} (۱۹۸۷) و روشی که بران^{۵۱} (۱۹۹۵) بر اساس آن، تخمین زن ماکزیمم تقریبی راست نمایی را بر پایه ی مینیمم مجموعه مربعات باقیمانده های پیش بینی ساده، توسعه داده شده است، مطرح گشته اند. بران، کارایی مجانبی و به هنجاری نتایج تخمین زننده های پارامترهای (d, ϕ, θ) را اثبات نمود و محققانی چون بوردیگه و اوپکل^{۵۲} (۱۹۹۳) و چانگ و بایلی^{۵۳} (۱۹۹۳) با توجه به نتایج تخمین مونت کارلو^{۵۴} این روش برآورد را (که ما آن را روش NLS می نامیم) یک تخمین زن مناسب برای مدل های ARFIMA بیان کردند. تکنیک NLS بر خلاف روش EML احتیاجی به پیش فرض مانایی ($d < 0.5$) نخواهد داشت و همچنین محاسبه ی آسان تر آن نسبت به تکنیک EML در برآورد، روش NLS را تکنیکی دقیق تر از تکنیک EML معرفی کرده است. لگاریتم تقریبی راست نمایی NLS به صورت زیر تعریف می شود:

$$\log L_A(d, \phi, \theta, \beta) = c - \frac{1}{2} \log \frac{1}{T-K} \sum_{t=2}^T \tilde{e}_t^2 \quad (5)$$

که \tilde{e}_t^2 خطاهای پیش بینی یک گام به جلو است که بواسطه ی پیش بینی های ساده، تعریف تقریبی شده است (Dittman & Granger, 2002).

۲-۴- تشخیص مدل

برای چک کردن کفایت مدل های ARMA و ARFIMA، مدل های تحت بررسی آزمون ضریب لاگرانژ (LM)^{۵۵}، آزمون نرمال بودن باقیمانده ها و آزمون عدم همبستگی بین باقیمانده

ها، قرار خواهند گرفت. در صورتی که باقیمانده‌ها حاوی الگوی سیستماتیک و دیگر ویژگی‌های نامناسب برای برازش مدل باشند باید پارامترهای مختلف AR و MA را برای رفع چنین ویژگی‌هایی تخمین زد و مرتبه مناسب مدل، توسط معیار اکائیک (AIC) Δ^F برای انتخاب الگوی برتر تشخیص داده می‌شود. برای انجام پیش‌بینی، هم ارزش‌های درون نمونه و هم برون نمونه‌ی مدل‌های ARMA و ARFIMA محاسبه خواهند شد و تکنیک مورد استفاده پیش‌بینی در این مقاله روش‌های پیش‌بینی ساده خواهد بود. این روش به طور برگشتی مقادیر $\tilde{z}_{T+1}, \tilde{z}_{T+2}, \dots$ را با استفاده از AR پیش‌بینی می‌کند. واریانس مربوطه \tilde{z}_{T+H} با استفاده از ضریب MA محاسبه می‌شود بطوریکه:

$$\text{var}(\tilde{z}_{T+H}) = \sigma_\varepsilon^2 \left(1 + \sum_{j=1}^{H-1} \psi_j^2 \right) \quad (6)$$

در این مورد ارزش‌های پیش‌نمونه صفر تعیین می‌شوند و در صورتی که مشاهدات نامحدود شناخته شوند، پیش‌بینی‌ها بهینه خواهند بود. لازم به ذکر است که روش‌های پیش‌بینی بهینه مجموعه اطلاعات نمونه محدود را در نظر می‌گیرند در حالی که روش‌های پیش‌بینی ساده، چنین ویژگی را بکار نمی‌برند (Diebold & Rudebush, 1989). ارزیابی ارزش‌های پیش‌بینی با استفاده از معیارهای مانند RMSFE و MAPFE که بصورت معادلات ۷ و ۸ تعریف می‌شوند، انجام می‌شوند.

$$\text{MAPFE} = \frac{100}{h+1} \sum_{t=s}^{h+s} \left| \frac{\hat{X}_t - X_t}{\hat{X}_t} \right| \quad (7)$$

$$\text{RMSFE} = \sqrt{\frac{1}{h+1} \sum_{t=s}^{h+s} (\hat{X}_t - X_t)^2} \quad (8)$$

که $t = s, 1+s, \dots, h+s$ و ارزش‌های واقعی و پیش‌بینی شده به ترتیب با X_t و \hat{X}_t برابرند. ارزش‌های کوچک تر RMSFE و MAPFE، عملکرد پیش‌بینی بهتری را برای مدلی که چنین ویژگی را دارا باشد در پی خواهد داشت.

۳- توصیف داده‌ها

داده‌هایی که در این تخمین برای مدل‌سازی در بورس به کار برده شده است شاخص قیمت بورس تهران برای دوره زمانی ۱۳۸۰/۰۹/۰۴ تا ۱۳۹۰/۰۹/۰۹ که شامل ۲۳۵۷ مشاهده می‌باشد. در

این قسمت برای محاسبه ی بازده r_t از بازده ی لگاریتمی که در معادله ۹ تعریف شده است، استفاده می شود.

$$r_t = \log\left(\frac{p_t}{p_{t-1}}\right) \times 100 \quad (9)$$

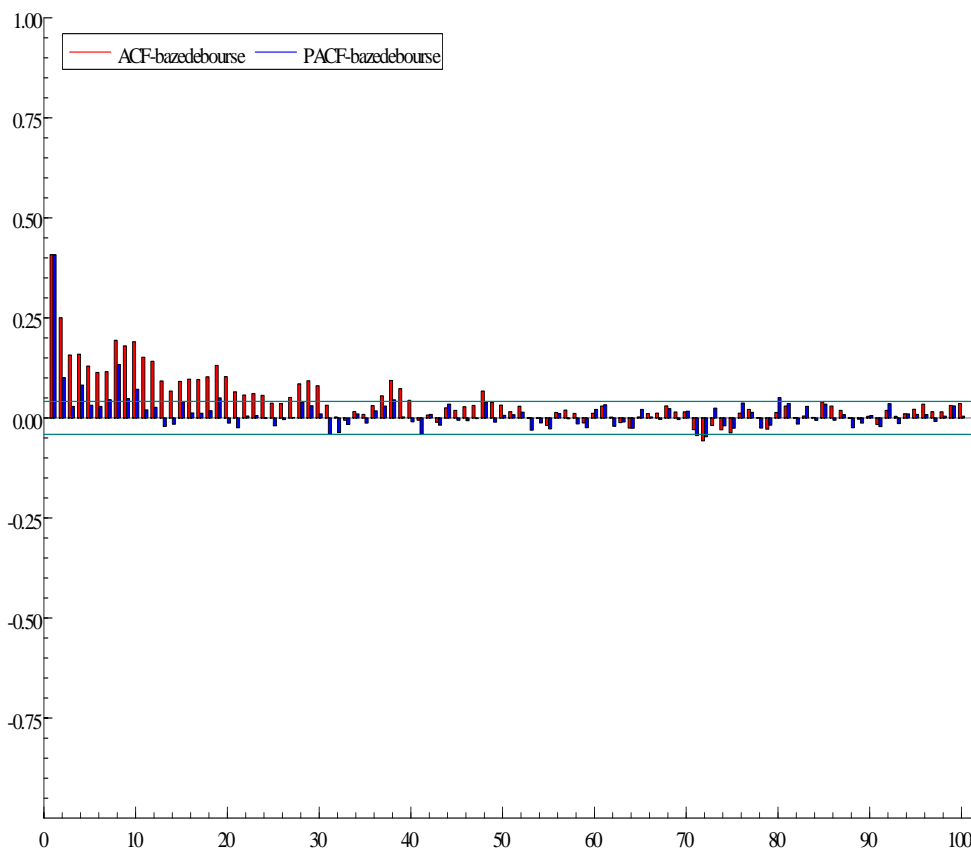
که p_t مقدار شاخص قیمت بورس در زمان t است. در تجزیه و تحلیل مقدماتی طبق جدول شماره ۱ توزیع بازده ها دارای کشیدگی مثبت است و همچنین داده ها درجه اضافی کشیدگی را نشان می دهند. چنین کشیدگی و چولگی ویژگی رایج در بازده دارایی ها می باشد. مقایسه انحراف معیار در برابر میانگین نشان می دهد که این متغیر در طول دوره بررسی از تلاطم زیادی برخوردار بوده است. آماره جاک-برا^{۵۷} برای بازده لگاریتمی برابر $۱۷۰۶۳/۱۶$ است که نشان دهنده ی رد شدن فرض نرمال بودن توزیع بازده ها است. آزمون ایستایی که توسط آماره فیلیپس-پرون^{۵۸} (pp) برای داده های مربوط به بازده شاخص کل بورس محاسبه شده است برابر $۳۹/۲۹۴۴۳$ است و از آن جایی که قدر مطلق مقدار بحرانی مربوط از سطح معنی داری $۰/۰۵$ بزرگ تر است، فرضیه صفر وجود ریشه واحد رد می شود. آزمون دیگر ایستایی که برای آزمون فرضیه صفر مانایی بر اساس آماره کوویت کووسکی-فیلیپس-اشمیت-شین^{۵۹} (kpss) صورت گرفته نشان می دهد که در سطح $۰/۰۵$ مقدار آماره برابر $۰/۲۷۰۵۰۸$ است و از آن جایی که در درون ناحیه بحرانی قرار می گیرد منجر به رد شدن فرضیه صفر مانایی آزمون می شود.

جدول شماره ۱: ویژگی های سری زمانی بازده TEPIX

آماره	TEPIX
تعداد مشاهدات	۲۳۵۷
میانگین	۰/۰۸۲۴۳۴
ماکزیمم	۵/۲۶۰۸۴۸
مینیمم	-۵/۴۵۰۲۹۶
انحراف معیار	۰/۵۹۸۳۹۸
چولگی	۰/۳۳۸۰۳۸
کشیدگی	۱۶/۱۶۷۴
کشیدگی اضافی	۱۳/۱۶۸
جاک-برا	۱۷۰۶۳/۱۶
PP	-۳۹/۲۹۴۴۳
KPSS	۰/۲۷۰۵۰۸

طبق مطالعات بایلی و همکاران^{۶۰} (۱۹۹۶)، رد شدن هر دو آماره pp و kpss حاکی از این است که فرایند نه توسط I(0) و نه توسط I(1) توصیف می‌شود، بنابراین چنین فرایندی ممکن است توسط فرایند کسری جزئی بهتر توصیف شود. همچنین شکل شماره ۱ تابع خود همبستگی TEPIX را نشان می‌دهد که کاهش خود همبستگی هایش در حالی که وقفه‌ها افزایش می‌یابد خیلی کند می‌باشد و تا حد زیادی از ویژگی تابع هیپربولیک پیروی می‌کند که بنابر ادبیات موجود نشانگر وجود حافظه بلند مدت خواهد بود.

شکل شماره ۱: توابع خود همبستگی و خود همبستگی جزئی



منبع: یافته‌های تحقیق

۴- برازش مدل‌ها، نتایج مدل و پیش‌بینی‌ها

۴-۱- برازش مدل‌ها و نتایج آن

برای یافتن مرتبه مناسب و برای لحاظ کردن محدودیت‌هایی که در تخمین چنین مدل‌هایی پیش خواهد آمد، برآوردها را در حالت مختلف برای $p < 6$ و $q < 6$ انجام می‌دهیم و سپس بر اساس الویت اعمال محدودیت ساختاری مدل‌ها و نتایج حاصل از کمترین معیار AIC بهترین مرتبه برای تعیین p و q را استخراج خواهیم نمود. بررسی شکل شماره ۱ نشان می‌دهد که مرتبه p و q برای مدل‌های ARIMA و ARFIMA می‌تواند بالاتر از یک باشد. با برآورد چندین مدل ARIMA و با کاربرد معیار AIC بهترین مدل مطلوب، ARMA(3,5) شده است. تخمین پارامترها، خطاهای استاندارد و p -values در جدول ۲ نمایان است.

جدول شماره ۲: تخمین مدل ARMA برای بازده TEPIX

	Estimates	Std. Errors	t-value	Prob.
Constant	0/088	0/039	2/25	0/024
AR(1)	0/844	0/078	10/07	0
AR(2)	0/952	0/032	29/0	0
AR(3)	-0/815	0/073	-11/2	0
MA(1)	-0/518	0/080	-6/40	0
MA(2)	-1/060	0/043	-24/6	0
MA(3)	0/416	0/069	6/00	0
MA(4)	0/140	0/029	4/75	0
MA(5)	0/085	0/023	3/68	0/001
Log-likelihood	-1618/030		Mean(residuals)	-0.00042
AIC	1/5095		Std.Devn(residuals)	0/51232
Mean(daily returns)	0/0846		Skewness	-0/40406
var(daily returns)	0/3297		Excess kurtosis	20/520

منبع: یافته‌های تحقیق

از تخمین مدل ARMA بالا، همه پارامترها در سطح ۰/۰۵ معنی‌دار می‌باشند و طبق جدول شماره ۳، علی‌رغم برآورد کردن مدل‌های ARMA بسیار تا رسیدن به مدلی که در موارد سنجش برازش مانند آزمون LM، آزمون نرمال بودن باقیمانده‌ها و آزمون عدم همبستگی بین باقیمانده‌ها، برازش خوبی ارائه دهد، نتیجه‌ای حاصل نشد (فرض صفر همه‌ی آزمون‌ها رد شد) و بر اساس کمترین مقدار معیار AIC بین سایر مدل‌های ARIMA، مدل ARMA(3,5) استخراج شد.

برای انتخاب درست مرتبه‌های مدل ARFIMA، جستجوی کلی اجزای سازنده مدل‌ها با کاربرد معیار AIC نشان از این دارد که بهترین مدل ARFIMA(5,0/380,4) می‌باشد. تخمین پارامترها، خطاهای استاندارد و p-values مدل ARFIMA(5,0/380,4) در جدول شماره ۴ ارائه شده است.

جدول شماره ۳: برازش مدل ARMA

Normality test: Chi ² (2)	**4918.2 [0.0000]
ARCH 1-1 test: F(5,2145)	**33.111 [0.0000]
Portmanteau(46): Chi ² (38)	*60.134 [0.0126]
* معناداری در سطح ۰/۰۵ ** معناداری در سطح ۰/۰۱	

منبع: یافته‌های تحقیق

جدول شماره ۴: تخمین مدل ARFIMA برای بازده TEPIX

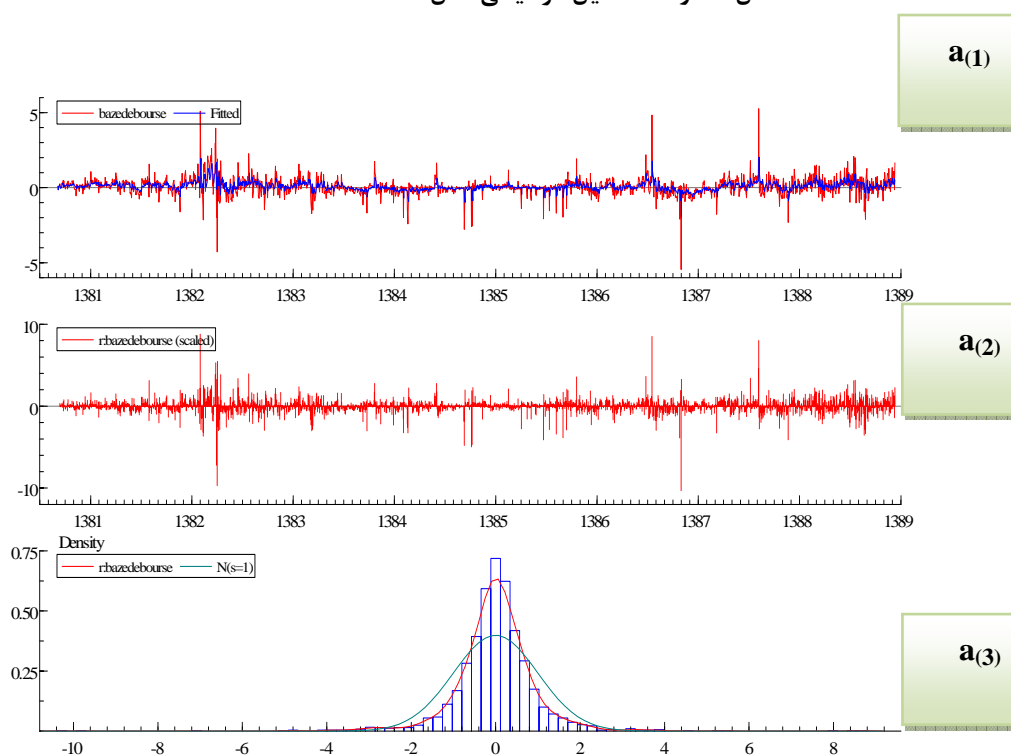
	Estimates	Std. Errors	t-value	Prob.
d	0/380	0/071	5/35	0
AR(1)	0/951	0/211	4/51	0
AR(2)	0/250	0/094	2/65	0/008
AR(3)	-1/099	0/153	-7/17	0
AR(4)	0/429	0/159	2/69	0/007
AR(5)	-0/066	0/032	-2/06	0/040
MA(1)	-1/006	0/234	-4/33	0
MA(2)	-0/271	0/119	-2/27	0/023
MA(3)	1/117	0/157	7/10	0
MA(4)	-0/440	0/184	-2/38	0/017
Log-likelihood	-1611/443		Mean(residuals)	-0/0011
AIC	1/5052		Std.Devn(residuals)	0/5107
Mean(daily returns)	0/08462		Skewness	0/3656
var(daily returns)	0/3297		Excess kurtosis	19/927

منبع: یافته‌های تحقیق

با در نظر گرفتن پارامتر d برآورده شده و قرار گرفتن آن بین $0 < d < 0.5$ و معنی داری آن در جدول شماره ۴، حافظه بلند مدت بودن بازده شاخص کل سهم تهران تایید می‌شود و همچنین نبود خود همبستگی بین باقیمانده‌ها برازش خوب مدل را تایید می‌کند. نسبت واریانس باقیمانده‌ها ی دو مدل عبارت است از $\frac{\sigma_{ARFIMA}^2}{\sigma_{ARIMA}^2} = 0.99 < 1$ که نشان دهنده ی برآورد بهتر مدل ARFIMA می‌باشد و همچنین کمتر بودن معیار AIC مدل ARFIMA، تصویر مدل سازی بهتری

را نسبت به مدل ARMA نشان می‌دهد. با توجه به این امر شکل شماره ۲ با هدف تحلیل گرافیکی مدل ARFIMA تهیه شده است. شکل a_1 ارزش برازش شده و واقعی را نشان می‌دهد، شکل a_2 نشان دهنده ی باقیمانده های مدل است و شکل a_3 بیانگر تابع چگالی باقیمانده ها به انضمام تابع چگالی نرمال می باشد.

شکل شماره ۲: تحلیل گرافیکی مدل ARFIMA



منبع: یافته های تحقیق

آنچه از شکل شماره ۲ نمایان است کفایت برازش مدل با توجه به شکل a_2 از لحاظ برازش نرمال بودن باقیمانده ها تایید نمی شود و همان طور که در جدول شماره ۵ نمایان است، با تست آزمون LM بین باقیمانده ها، اثراتی از ناهمسانی واریانس نمایان است و از لحاظ آزمون عدم خود همبستگی بین باقیمانده ها، مدل ARFIMA(5,0/380,4) فرضیه صفر این آزمون را تایید می کند.

جدول شماره ۵: برازش مدل ARFIMA

Normality test: Chi ² (2)	4786.3 [0.0000]**
ARCH 1-1 test: F(5,2145)	36.595 [0.0000]**
Portmanteau(46): Chi ² (38)	47.877 [0.1223]
**معناداری در سطح ۰/۰۱ * معناداری در سطح ۰/۰۵	

منبع: یافته‌های تحقیق

۲-۴- پیش‌بینی

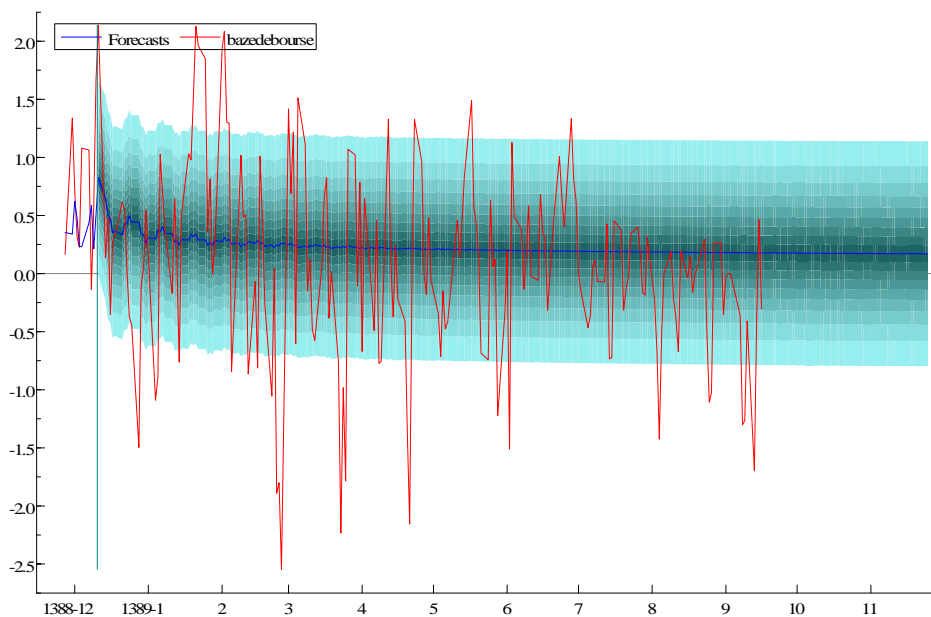
بعد از مشخص شدن پارامترهای ARMA و ARFIMA با ۲۳۵۷ داده که در این میان، ۲۰۰ داده برای دوره پیش‌بینی نمونه در نظر گرفته شده بودند و برای h گام پیش‌بینی جلو، ۲۵۰ گام را تعیین نمودیم که مشخصاً ۵۰ داده برای پیش‌بینی برون نمونه در نظر گرفته خواهد شد. از مقایسه ارزش‌های پیش‌بینی درون نمونه با مشاهدات واقعی برای مدل‌های ARIMA و ARFIMA، مدل ARFIMA بر پایه معیارهای RMSFE و MAPFE (همانطوری که در جدول شماره ۶ نشان داده شده است)، تفاوت عملکرد بسیار جزئی بهتری را در مقایسه با مدل ARMA از خود بروز می‌دهد و بر این اساس برآورد پیش‌بینی خارج از نمونه توسط مدل ARFIMA صورت گرفته است. نتایج برآورد ارزش‌های پیش‌بینی درون نمونه و خارج از نمونه مدل ARFIMA به صورت گرافیکی در شکل شماره ۳ نمایان است. بر اساس شکل شماره ۳ با وجود عملکرد بهتر بسیار جزئی مدل ARFIMA نسبت به مدل دیگر، پیش‌بینی مدل ARFIMA عملکرد خوبی به قطعیت از خود نشان نمی‌دهد زیرا اکثر ارزش‌های واقعی آن خارج از فاصله اطمینان ۰/۹۵ قرار دارند (فاصله اطمینان در شکل شماره ۳ به طور سایه نشان داده شده است). این امر به علت بحران‌ها و فراز و نشیب‌هایی بسیار زیادی است که بازار سرمایه تهران در این دوره زمانی مورد تحقیق متحمل آن شده است، می‌باشد.

جدول شماره ۶: ارزیابی عملکرد پیش‌بینی مدل‌ها

MODEL	RMSE	MAPE	Mean(Error)	SD
ARMA	0/80797	581/50	-/070097	-0/80493
ARFIMA	0/82063	269/72	-0/17804	0/80108

منبع: یافته‌های تحقیق

شکل شماره ۳: پیش بینی بازده TEPIX با مدل ARFIMA



منبع: یافته های تحقیق

۵- نتیجه گیری و بحث

این مطالعه نرخ بازده شاخص سهام کل تهران را از بازده زمانی ۱۳۸۰/۰۹/۰۴ تا ۱۳۹۰/۰۹/۰۹ پوشش داده است. با رد شدن فرض صفر آزمون های PP و KPSS و کاهش هیپربولیکی تابع خود همبستگی، احتمال اینکه که فرایند توسط فرایند کسری جزئی بهتر توصیف شود، امکان بالایی می یابد و با تخمین مدل ها با تکنیک NLS که ویژگی هایی چون نادیده گرفتن فرض مانایی سری های زمانی و همچنین روش محاسبه آسان تر آن، وجود حافظه بلند مدت در بازده TEPIX اثبات شد و می توانیم این استنباط را کنیم که بین داده ها در بازه زمانی مشخص شده خود همبستگی معناداری وجود دارد، بنابراین فرض گام تصادفی نفی می شود. با این وجود بازار سرمایه قابل پیش بینی است و فرض شکل ضعیف کارایی بازار نقض خواهد شد. نسبت واریانس باقیمانده های مدل ARFIMA بر ARMA و همچنین AIC کمتر مدل ARFIMA، این مدل را مدلی برتر در مدل سازی نمایان می سازد و نتایج پیش بینی، نشان دهنده ی این موضوع است که ارزش های پیش

بینی مدل ARFIMA با توجه به معیار های ارزیابی پیش بینی با تفاوت عملکرد بسیار جزئی، پیش بینی مناسب تری نسبت به مدل ساده ARMA ارائه داده است و قرار نگرفتن اکثر ارزش های واقعی در فواصل اطمینان ۰/۹۵، از عملکرد خوب پیش بینی مدل ARFIMA به وضوح حمایت نمی کند. تجزیه و تحلیل و نتایج مدل ها ذکر این نکته را ضروری می سازد که قبل از انجام هر گونه برآورد مدل، تحلیل لازم روی سری های زمانی صورت بگیرد. این امر خصوصیات نهانی سری های زمانی را آشکار می کند و کمک شایانی در انتخاب درست مدل مناسب به تحلیل گر خواهد کرد. در پایان، نتایج این تحقیق را می توان دقیقاً در راستای تحقیق شیتو و یایا (۲۰۰۹) دانست. از مطالعات دیگر که با روش یک مرحله ایی، مدل ARFIMA را تخمین زده است، مطالعه ی کوریتا (۲۰۱۰) می باشد که بر خلاف نتایج ما، از عملکرد بسیار خوب پیش بینی مدل ARFIMA در برابر مدل ARIMA دلالت می کند.

فهرست منابع

- ۱) شعراپی، سعید و محسن ثنائی اعلم (۱۳۸۹)، بررسی وجود حافظه بلندمدت در بورس اوراق بهادار تهران و ارزیابی مدل‌هایی که حافظه بلندمدت را در نظر میگیرند، مجله پژوهش های حسابداری مالی، شماره ۶، صص ۱۸۶-۱۷۳.
- ۲) محمودی، وحید. محمدی، شاپور و هستی چیت سازان (۱۳۸۹)، بررسی حافظه بلند مدت در بازار های جهانی نفت، فصل نامه تحقیق مدل سازی اقتصادی، شماره ۱، صص ۴۸-۲۹.
- ۳) مشیری، سعید و حبیب مروت (۱۳۸۵)، پیش بینی شاخص کل بازدهی سهام تهران با استفاده از مدل های خطی و غیر خطی، فصلنامه پژوهش‌های بازرگانی، شماره ۴۱، صص ۲۷۵-۲۴۵.
- 4) Baillie R.T. Long Memory Processes and Fractional Integration in Econometrics, Journal of Econometrics, , 1996. 73(1): 5-59.
- 5) Baillie R.T., Bollerslev T. and Mikkelsen H.O. Fractionally integrated generalized autoregressive conditional heteroskedasticity. Journal of Econometrics. 1996. 74(1): 3-30
- 6) Beran J. Maximum likelihood estimation of the differencing parameter for invertible short and long memory autoregressive integrated moving average models. Journal of the Royal Statistical Society. 1995. 57(4): 659-672.
- 7) Beveridge S. and Oickle C. Estimating fractionally integrated time series models, Economics Letters. 1993. 43(2): 137-142.

- 8) Bhardwaj G. and Swanson N. R. An Empirical Investigation of the Usefulness of ARFIMA Models for Predicting Macroeconomic and Financial Time Series. *Journal of Econometrics*. 2004: 539-578.
- 9) Bilel T. M. and Nadhem S. Long Memory in Stock Returns: Evidence of G7 Stocks Markets. *Research Journal of International Studies*. 2009. 9: 36-46
- 10) Box G. E. P. and Jenkins G. M. *Time Series Analysis: Forecasting and Control*. Revised Edition. Holden-Day. 1976.
- 11) Brillinger. *Time series Data Analysis and Theory Expanded Edition*. McGraw-Hill. Inc. New York. 1981
- 12) Chatfield C. *The Analysis of Time Series. An introduction*. New York, Chapman and Hall. 1975
- 13) Cheung Y.W. and Diebold F.X. On maximum likelihood estimation of the difference parameter of fractionally integrated noise with unknown mean. *Journal of Econometrics*. 1994. 62(2): 301-316.
- 14) Chio K. Zivot E. Long memory and structural changes in the forward discount: an empirical investigation. Working Paper. University of Washington. 2002
- 15) Chung C. F. and Baillie R.T. Small sample bias in conditional sum of squares estimators of fractionally integrated ARMA models. *Empirical Economics*. 1993. 18(4): 791-806
- 16) Diebold F. and Inoue A. Long memory and regime switching. *Journal of Econometrics*. 2001. 105(1): 131-159.
- 17) Diebold F.X. and Rudebusch G.D. Long memory and persistence in aggregate output. *Journal of Monetary Economics*. 1989. 24(7): 189-209.
- 18) Dittman I. and Granger C.W.J. Properties of nonlinear transformations of fractionally integrated processes. *Journal of Econometrics*. 2002. 110(2): 113-133.
- 19) Doornik J.A. and Hendry D.F. *Econometric Modelling: PcGive*. Timberlake Consultants Ltd. London. 2007.
- 20) Doornik J.A. and Ooms M. Computational aspects of maximum likelihood estimation of autoregressive fractionally integrated moving average models. *Computational Statistics & Data Analysis*. 2003. vol. 42(3): 333-348.
- 21) Doornik J.A. and Ooms M. Inference and forecasting for ARFIMA models, with an application to US and UK inflation. *Studies in Nonlinear Dynamics and Econometrics*. 2004. 8(2). Article 14.
- 22) Engle R.F. and Smith A.D. Stochastic permanent breaks. *Review of Economics and Statistics*. 1999. 81(4): 553-574.
- 23) Erfani A. and Samimi A.J. Long Memory Forecasting of Stock Price Index Using a Fractionally Differenced Arma Model. *Journal of Applied Sciences Research*. 2009. 5(10): 1721-1731
- 24) Geweke J. and Porter-Hudak S. The Test and Application of Long memory Time Series Models. *Journal of Time Series Analysis*. 1983. 4(4): 221-238.
- 25) Granger C. W. J. and Joyeux R. An Introduction to Long Memory Time Series Models and Fractional Differencing. *Journal of Time Series Analysis*. 1980. 1(1): 15-29.

- 26) Hassler U. and Wolters J. Long memory in inflation rates: international evidence. *Journal of Business and Economic Statistics*. 1995. 13(1): 37-45.
- 27) Hosking J. R. M. Fractional Differencing. *Biometrika*. 1981. 68(1): 165-176
- 28) Hurst H. E. Long-Term Storage Capacity of Reservoirs. *Transactions of the American Society of Civil Engineers*, 1951. 116(1): 770-799.
- 29) Hyung N. and Franses P.H. Structural breaks and long-memory in US inflation rates: do they matter for forecasting?. Working Paper. Erasmus University. 2001
- 30) Lo A.W. Long-Term Memory in Stock Market Volatility. *Journal of Econometrica*. 1991. 59(5): 1279-1313.
- 31) Mandelbrot B. B. and Wallis J. R. Computer Experiments with Fractional Gaussian Noises. Parts 1, 2, 3. *Water Resources Research*. 1969. 5(1): 967-988.
- 32) Man, K.S. Long memory time series and short term forecasts, *International Journal of Forecasting*. 2003. 19(3): 477-491.
- 33) Odaki M. On the invertibility of fractionally differenced ARIMA processes. *Biometrika*. 1993. 80(3): 703-709.
- 34) Robinson P. Log-periodogram regression of time series with long range dependence. *The Annals of Statistics*. 1995. 23(3): 1048-1072.
- 35) Shittu O. I. and Yaya D.S. Measuring Forecast Performance of ARMA and ARFIMA Models: An Application to US Dollar/UK Pound Foreign Exchange Rate. *European Journal of Scientific Research*. 2009. 32(2): 167-176
- 36) Sowell F. Maximum likelihood estimation of fractionally integrated time series models. Discussion Paper Department of Economics. Duke University. 1987.
- 37) Sowell, F. Maximum Likelihood Test of Stationary Univariate Fractionally Integrated Time Series Models. *Journal of Econometrics*. 1992. 53(1): 165-188

یادداشت‌ها

- | | |
|---|---------------------------|
| 1. Box – Jenkins approach | 16. Hosking |
| 2. Autoregressive moving average | 17. Geweke & Porter-Hodak |
| 3. Autoregressive integrated moving average | 18. Lo |
| 4. Hurst | 19. Sowell |
| 5. Hydrologist | 20. Cheung & Diebold |
| 6. Rescaled range analysis | 21. Robinson |
| 7. Mandelbrot & Wallis | 22. Engle & Smith |
| 8. Persistence | 23. Diebold & Inoue |
| 9. Heuristic | 24. Dittman & Garanger |
| 10. Granger & Joyeux | 25. Doornik & Ooms |
| 11. Autoregressive fractionally integrated moving average | 26. likelihood based |
| 12. Short memory process | 27. Diebold & Rudebusch |
| 13. Long memory process | 28. Hassler & Wolter |
| 14. Hyperbolic | 29. Hyung & Franses |
| 15. Stationary & invertible | 30. Chio & Zivot |
| | 31. Man |
| | 32. Bhardwaj & Swanson |

33. Shittu & Yaya
34. Tehran Exchange price index
35. Root mean Square forecast Error
36. Mean Absolute Percentage Forecast Error
37. Box & Jenkins
38. Brillinger
39. Chattfield
40. Doornik & Ooms
41. Real number
42. Autocovariance Function
43. Intermediate memory
44. Odaki
45. Geweke & Porter-Hudak
46. Detrended fluctuation analysis
47. Wavelet
48. Exact maximum likelihood
49. Nonlinear least squares
50. Sowell
51. Beran
52. Beveridge & Oickle
53. Chung & Baillie
54. Monte Carlo
55. Lagrange multiplier
56. Akaike information criterion
57. Jarque-Bera
58. Phillips & Perron
59. Kwiatkowski, Philips, Schmidt, Shin
60. Baillie & et al