تحلیل مانایی سری زمانی شاخص قیمت مصرف‌کننده در ایران و ارائه یک مدل ARIMA برای پیش‌بینی آن

صمد کاظمی (نویسنده مسئول)
دانشجوی کارشناسی ارشد دانشگاه علم و صنعت ایران
sa_kazemi@ind.iust.ac.ir

پوریا سوری
دانشجوی کارشناسی ارشد دانشگاه علم و صنعت ایران
P_souri@ind.iust.ac.ir

مهدی غضنفری
استاد دانشگاه علم وصنعت ایران
mehdi@iust.ac.ir

میرسامان پیشوایی
استادیار دانشگاه علم وصنعت ایران
pishvaee@iust.ac.ir

چکیده

در این پژوهش برای تعیین مانایی یا غیر مانایی داده‌های سری زمانی شاخص قیمت مصرف‌کننده در ایران از سال ۱۳۵۹ تا ۱۳۹۱ به قیمت‌های ثابت سال ۱۳۸۳، پس از کشف روند تغییرات داده‌های سری زمانی و تعیین مرتبه خود رگرسیو، میانگین متحرک و میانگین متحرک خودرگرسیو سری زمانی، آزمون ریشه واحد را روی داده‌های سری زمانی انجام می‌دهیم و عدم مانایی سری مشخص می‌شود. در ادامه با انجام آزمون دیکی‌فولر تکمیل‌شده، مشخص می‌شود که سری زمانی تنها یک ریشه واحد دارد. پس از حصول اطمینان از داشتن تنها یک ریشه واحد، مدل فرایند خودرگرسیون انباشته با میانگین متحرک را برای این سری زمانی برآورد می‌کنیم. سپس کفایت مدل را با آزمون پورتمن-تیو بررسی کرده و تصادفی خالص بودن سری زمانی مشخص می‌شود. در نهایت با استفاده از مدل به‌دست‌آمده، مقادیر آینده داده‌های سری زمانی را در دو فاصله اطمینان ۸۰ و ۹۵ درصد از سال ۱۳۹۲ تا ۱۴۰۱ پیش‌بینی می‌کنیم.

کلید واژه:

سری زمانی، آزمون ریشه واحد، مانایی، مدل خودرگرسیو انباشته با میانگین متحرک، پیش‌بینی

مقدمه

سری زمانی در فرایندهای کسب‌وکار، اقتصاد، محیط زیست، پزشکی و دیگر حوزه‌های علمی استفاده می‌شود تا الگوهایی همچون روند، نوسانات فصلی، دوره‌های نامنظم و تغییرات ناگهانی در سطح یا پراکندگی را مشخص سازد. هدف از مطالعه سری‌های زمانی اغلب، برون‌یابی الگوی پویای موجود در داده‌هاست تا بتوان مشاهدات آینده را پیش‌بینی کرد، اثرات تداخلات خارجی شناخته‌شده را تخمین زد و تداخلات نامحسوس را شناسایی کرد. وابستگی ترتیبی در داده‌های سری زمانی را می‌توان با یک مدل پویای خطی تقریب زد. دو مدل برای این کار توسعه یافته‌اند: مدل‌های خودرگرسیو با میانگین متحرک[1] (آرما) که برای شناسایی سری‌های زمانی مانا و غیر مانا به کار می‌رود و مدل‌های خودرگرسیو با میانگین متحرک انباشته[2] (آریما) که برای انجام آزمون ریشه واحد به‌کار می‌روند [۱].

شاخص قیمت مصرف‌کننده[3] به‌عنوان یکی از مؤثرترین شاخص‌ها برای نشان‌دادن وضعیت فعلی تورم در اقتصاد مطرح شده است. این شاخص ابزاری برای سنجش تغییرات قیمت گروه ثابتی از کالاها و خدمات است که در سبد خرید روزانه مردم جای دارد. یکی از موارد کاربرد این شاخص، محاسبه درصد تغییرات قیمت از یک سال به سال بعد است که نرخ تورم نامیده می‌شود و برای ارزیابی موفقیت یا شکست سیاست‌های اقتصادی دولت به‌کار می‌رود. یکی دیگر از کاربردهای آن، افزایش پرداخت‌های درآمدی است [۲].

با توجه به این موضوع که افزایش شاخص قیمت مصرف‌کننده یا تورم، قدرت خرید ارز ملی را کاهش می‌دهد و تاثیر به‌سزایی بر تصمیم‌گیری‌های اقتصادی دولت دارد، در این پژوهش قصد داریم با بررسی داده‌های سری زمانی هزینه مصرفی خانوار در ایران از سال ۱۳۵۹ تا ۱۳۹۱ به قیمت‌های ثابت سال ۱۳۸۳، ایستایی یا مانایی این سری زمانی و داشتن ریشه واحد آن را مشخص کنیم. ادامه مقاله بر بخشهای زیر استوار شده است:

در بخش پیشینه پژوهش، به بررسی چندین مطالعه مرتبط در حوزه‌های گوناگون اقتصادی، زیست‌محیطی، انرژی و قیمت مسکن می‌-پردازیم که از مدل‌های آریما یا سایر مدل‌های ترکیبی برای پیش‌بینی روند تغییرات یک‌سری داده‌های سری زمانی معین استفاده شده است. همان‌طور که پیشتر گفته شد سری زمانی در فرایندهای کسب‌وکار، اقتصاد، محیط زیست، پزشکی و دیگر حوزه‌های علمی استفاده می‌شود اما تاکنون در ایران، علیرغم این که از سری زمانی و مدل‌های آماری چندمتغیره برای پیش‌بینی داده‌های مختلفی اعم از قیمت گوشت و مرغ، میزان صادرات یک کالای خاص، میزان تقاضای آب شرب، میزان تقاضای سفر ریلی و ... استفاده شده است، تحلیلی روی داده‌های سری زمانی شاخص قیمت مصرف‌کننده از لحاظ آزمون داشتن ریشه واحد و پیش‌بینی مقادیر آینده آن صورت نگرفته است از طرفی تعیین مانایی یا عدم مانایی سری‌های زمانی در بررسی‌های مربوط به شاخص قیمت مصرف‌کننده تاکنون انجام نشده است که در این پژوهش، قصد داریم این شکاف تحقیقاتی را پر و پس از بررسی سری زمانی داده‌های شاخص قیمت مصرف‌کننده در ایران و تعیین مانایی یا عدم مانایی آن، بهترین مدل برازش‌داده‌شده را برای پیش‌بینی مقادیر آینده این شاخص ارائه کنیم.

در بخش مبانی نظری و روش تحقیق، روند داده‌های سری زمانی شاخص قیمت مصرف‌کننده و مرتبه فرایندهای میانگین متحرک، خودرگرسیو و ترکیب آنها را تعیین می‌کنیم و سپس با اجرای گامهای مختلف، مانایی داده‌ها را با آزمون ریشه واحد مورد بررسی قرار می‌دهیم. پس از آن که با انجام آزمون ریشه واحد چندگانه مشخص شد که سری زمانی فوق تنها یک ریشه واحد دارد، مدل آریمای مناسب را برآورد کرده و پس از حصول اطمینان از کفایت آن با استفاده از آزمون پورتمن-تیو، از آن برای پیش‌بینی مقادیر آینده این شاخص برای بازه زمانی ۱۳۹۲ تا ۱۴۰۱ در دو فاصله اطمینان ۸۰ و ۹۰ درصد استفاده می‌کنیم. تمامی گامهای فوق و نیز برآورد مدل مناسب آریما در نرم‌افزار آماری R-Project انجام شده‌است. در نهایت با توجه به نمودار پیش‌بینی و دقت مناسب مدل برآورد شده نتیجه می‌گیریم که برآورد نقطه‌ای سری زمانی شاخص قیمت مصرف‌کننده از سال ۱۳۹۲ تا ۱۴۰۱ روند صعودی دارد. سیاست‌گذاران بخش اقتصاد کشور می‌توانند از نتایج به‌دست آمده برای سیاست‌گذاری‌ها و اخذ تصمیمات اقتصادی مرتبط استفاده کنند.

سازمانها در راه تحقق ارزشهای اساسی و رسیدن به چشم انداز و اهداف خود دچار چالشهای جدی می گردند این چالشها در صورتیکه بدرستی پیش بینی و بر اساس آن چاره اندیشی نشده باشد میتواند حرکت سازمان را به انحراف کشیده و از اهداف دور سازد (نیلسون ، ۲۰۰۸) [۴]. لذا می بایست در یک برنامه دقیق بطور منظم به تعریف ارکان جهت ساز سازمان پرداخته و در حرکت بعدی با توجه به مدلهای آینده پژوهی به شناسایی آینده های ممکن اقدام و سپس مولفه های دخیل در ایجاد تهدیدها ، فرصتها ، قوت و ضعفهای سازمانی کنار هم قرار گرفته و نحوه صحیح حرکت با توجه به عدم اطمینان موجود مورد بررسی قرار گیرد. نکته مهم که در اکثر سازمان به نظر کم اهمیت دیده شده و همین امر باعث شکست در راه رسیدن به اهداف گردیده است عدم وجود مدل ارزیابی ، کنترل و هدایت می باشد چه بسا سازمان دارای برنامه ریزی مناسب برای رسیدن به اهداف باشد اما برنامه کنترل مناسبی برای خود طراحی نکرده باشد. لذا مسئله ما در این تحقیق ، طراحی الگوی ارزیابی عملکرد واحدهای صنعتی است که برای تحقق ارزشها و اهداف سازمانی خود با در نظر گرفتن عدم اطمینان و ریسک ناشی از مخاطرات محیطی و با توجه به رویکرد اقتصاد مقاومتی ، با اتخاذ روش برنامه ریزی پابرچا و تلفیق آن با چارچوب کارت امتیازی متوازن توسعه یافته پاسخگویی ارزیابی و کنترل سازمانها در شرایط پیچیده باشد.

سازمان ها اهمیت ارزیابی با ثبات و بدون تناقض را تشخیص داده و لذا سیستم های ارزیابی عملکرد متنوعی را در طی سالیان پیش، مورد استفاده قرارداده اند . این امر برگرفته از جهانی شدن کسب وکار ، تاثیر محیط خارجی برعملکرد فرآیندهای سازمان می باشد بصورتی که کشورها میتوانند با ایجاد قوانین عادلانه و یا ظالمانه در اقتصاد کشورهای دیگر تاثیر گذار باشند . لذا جهت مقابله با چنین تهدیداتی نیاز به یک مدل جامع می باشد که همه عناصر یک سازمان را درجهت چشم انداز ملی هدایت کند و عملکرد سازمان را در زمان حال و آینده مورد بررسی قرار دهد .این مدل جامع می تواند برگرفته از ابعاد توسعه یافته کارت امتیازی متوازن باشد که با تکیه بر مدل سازی منعطف می تواند رفتار آتی سازمان را با سناریوهای گوناگون از اهداف استراتژیک مختلف شبیه سازی کند و استراتژی های بهتری را جهت اجرا انتخاب نماید.

آنچه که در این مقاله ارائه می گردد تدوین مدلی است که براساس خبرگی و مطالعات قبلی به توسعه مدل کارت امتیازی متوازن براساس سیاستهای اقتصاد مقاومتی پرداخته و سپس شاخصهای ارزیابی عملکرد سازمانهای صنعتی بر اساس این ابعاد چیدمان نموده و دراخر با تدوین سناریوهای پیش رو به جایگذاری شاخصها در ابعاد کلی سازمان پرداخته است.

۱. پیشینه پژوهش

یکی از حوزه‌هایی که از مدل‌های آریما برای پیش‌بینی استفاده می‌شود، حوزه انرژی شامل آب، برق و سوختهای فسیلی است که از آن جمله می‌توان به کارهای ادیگر و همکاران[٥](٢٠٠٦)، ادیگر و آکار[٦](٢٠٠٧)، تابش و همکاران[٧](٢٠٠٤)، وانگ و همکاران[٨](٢٠١٢)، عبدالعال و القرنی[٩](١٩٩٧)، براک و صادق[١٠](٢٠١٦) و یوآن و همکاران[١١](٢٠١٦) اشاره کرد.

در حوزه پیش‌بینی داده‌های سری زمانی انرژی برق، عبدالعال و القرنی (١٩٩٧) از مدل آریما برای مدل‌سازی و پیش‌بینی مصرف داخلی ماهانه برق در استان شرقی عربستان سعودی استفاده می‌کنند. آنها برای ساخت مدل آریما از داده‌های ٥ سال متوالی استفاده کرده و مقادیر سال ششم را پیش‌بینی می‌کنند [٣]. همچنین وانگ و همکاران (٢٠١٢)، در مطالعه خود برای افزایش دقت مدل فصلی آریما در پیش‌بینی تقاضای برق چین و کمک به تصمیم‌گیران حوزه انرژی برای اخذ سیاست‌های مناسب، مدل‌های اصلاح باقیمانده[١٢] را ارائه می‌دهند. نتایج پژوهش آنها نشان می‌دهد که دقت پیش‌بینی مدل‌های اصلاح باقیمانده از مدل آریما بیشتر بوده و مدل ترکیبی استفاده‌شده بهترین عملکرد را در بین مدل‌های معرفی‌شده دارد [٤]. در حوزه پیش‌بینی داده‌های سری زمانی سوختهای فسیلی، ادیگر و همکاران (٢٠٠٦) با استفاده از مدل‌های آریما و آریمای فصلی[١٣] برای هر کدام از انواع سوختهای فسیلی در ترکیه از سال ١٩٥٠ تا ٢٠٠٣، مدل پیش‌بینی متفاوتی ارائه می‌دهند. در نهایت با توجه به پیش‌بینی‌های انجام شده و تعیین روند نزولی یا صعودی‌بودن نرخ رشد سوختهای فسیلی، سیاست‌هایی را برای تصمیم‌گیران حوزه انرژی ترکیه پیشنهاد می‌کنند [٥]. ادیگر و آکار (٢٠٠٧) با استفاده از دو مدل آریما و آریمای فصلی تقاضای انرژی را در ترکیه از سال ٢٠٠٥ تا ٢٠٢٠ پیش‌بینی می‌کنند. پس از پیش‌بینی، صعودی یا نزولی‌بودن نرخ رشد هر یک از منابع انرژی را در سالهای آتی تعیین کرده و بر مبنای نتایج به‌دست‌آمده، سیاست‌هایی را برای تصمیم‌گیران حوزه انرژی ترکیه پیشنهاد می‌کنند. [٦]

در حوزه پیش‌بینی داده‌های سری زمانی آب، تابش و همکاران (٢٠٠٤) به‌دلیل وجود همبستگی زیاد بین مقادیر مصرف آب در روزهای متوالی و نوسانات موجود در مصرف روزانه آن، دو مدل پیش‌بینی مصرف روزانه آب برای شهر تهران شامل مدل سنتی و مدل پیشرفته‌تر آریما تهیه کرده و نتایج آنها را با هم مقایسه می‌کنند که این نتایج نشانگر دقت بیشتر روش سنتی در شناسایی و پیش‌بینی الگوی مصرف

آب شهر تهران در ماههای گرم سال و دقت بیشتر روش آریما در ماههای سرد سال است [۷]. در برخی مطالعات از مدل‌های ترکیبی آریما برای پیش‌بینی مصرف انرژی استفاده می‌شود. مطالعات انجام‌شده توسط براک و صادق و یوآن و همکاران در این حوزه جای می‌گیرد. براک و صادق از الگوریتم ترکیبی ARIMA-ANFIS برای پیش‌بینی مصرف انرژی در ایران استفاده می‌کنند. نتایج پژوهش آنها نشان می‌دهد که مدل ترکیبی مذکور از مدل‌های آریما و ANFIS جداگانه دقت و عملکرد بهتری دارد [۸]. یوآن و همکاران (٢٠١٦)، با استفاده از مدل آریما و مدل خاکستری (GM(1,1)) مصرف انرژی را در چین پیش‌بینی کرده و کارایی این دو مدل را با هم مقایسه می‌کنند. همچنین آنها یک مدل ترکیبی را با استفاده از این دو مدل می‌سازند که خطای کمتری در پیش‌بینی نسبت به مدل‌های جداگانه دارد. آنها از این سه مدل برای پیش‌بینی مصرف انرژی چین از سال ٢٠١٤ تا سال ٢٠٢٠ استفاده می‌کنند [۹].

در حوزه زیست‌محیطی می‌توان به کارهای انجام‌شده توسط یوسف و همکاران[4] (٢٠١٤)، نارایانان و همکاران[5] (٢٠١٣)، جیا و همکاران[6] (٢٠١٠) و تانژا و همکاران[7] (٢٠١٦) اشاره کرد. نارایانان و همکاران (٢٠١٣)، به تحلیل روند و پیش‌بینی میزان بارش باران‌های پیش از موسم در غرب هندوستان می‌پردازند. آنها از مدل مان کندال[8] برای یافتن روند بارش باران در ماههای مارس، آوریل و می (پیش از موسم) استفاده می‌کنند [١٠]. جیا و همکاران (٢٠١٠) با توجه به این که تحلیل‌های کمّی دقیق پیرامون روند توسعه اثرات زیست-محیطی[9] هنوز هم کمیاب است، از یک مدل آریما برای تقویت توانایی و دقت پیش‌بینی این شاخص استفاده می‌کنند. نتایج آنها نشان می‌دهد که مدل آریما برای پیش‌بینی و شبیه‌سازی این دو شاخص توانمند است [١١]. یوسف و همکاران (٢٠١٤)، در مطالعه خود به پیش‌بینی گاز متان منتشرشده از فضولات دام زنده در مالزی در بازه زمانی ١٩٨٠ تا ٢٠٠٨ با استفاده از یک مدل آریما می‌پردازند [١٢]. تانژا و همکاران (٢٠١٦)، از مدل آریما برای شبیه‌سازی میانگین ماهانه عمق نوری اسپری[10] استفاده می‌کنند که در بازه زمانی ٢٠٠٤ تا ٢٠١٤ به‌دست آمده است [١٣].

در حوزه‌های اقتصادی مختلفی از جمله بورس اوراق بهادار، قیمت مسکن، شاخص قیمت مصرف‌کننده و تولید ناخالص داخلی مطالعاتی انجام شده است که می‌توان به کارهای انجام‌شده توسط کرباسی یزدی و همکاران[11] (٢٠١٣)، جادویسیوس و هوستون[12] (٢٠١٥)، دو و همکاران[13] (٢٠١٤)، بابو و ردی[14] (٢٠١٤)، ووسلر[15] (٢٠١٥)، هپسن و وطنسور[16] (٢٠١٢) و استیونسون[17] (٢٠٠٧) اشاره کرد.

کرباسی یزدی و همکاران (٢٠١٣)، پدیده بازگشت به میانگین را در بورس اوراق بهادار تهران در بازه زمانی ١٣٨٠ تا ١٣٨٩ مورد بررسی قرار می‌دهند. آنها با استفاده از آزمون ریشه واحد (دیکی فولر تعمیم‌یافته) وجود پدیده بازگشت به میانگین در شاخص کل قیمت سهام، شاخص قیمت و بازده نقدی سهام و شاخص پنجاه شرکت برتر را بررسی می‌کنند [١٤]. جادویسیوس و هوستون (٢٠١٥)، در مطالعه خود از مدل آریما برای پیش‌بینی شاخص قیمت مسکن در لیتوانی استفاده می‌کنند. آنها ٢٠ مدل آریمای مختلف از ARIMA(١،٠،٠) تا ARIMA(٤،٠،٤) توسعه می‌دهند [١٥]. دو و همکاران (٢٠١٤)، یک الگوریتم تقسیم و غلبه[28] برای پیش‌بینی شاخص هزینه قیمت مصرف‌کننده با استفاده از سه مدل آریما، خاکستری[29] شبکه عصبی مصنوعی پس‌انتشار[30] ارائه می‌دهند [١٦]. بابو و ردی (٢٠١٤)، یک مدل ترکیبی آریما و شبکه عصبی مصنوعی را با استفاده از فیلتر میانگین متحرک برای پیش‌بینی داده‌های سری زمانی معرفی می‌کنند. آنها ماهیت تغییرپذیری را با استفاده از فیلتر میانگین متحرک بررسی کرده و سپس یک مدل ARIMA و شبکه عصبی مصنوعی را برای داده‌های آزمایشی قیمت سهام بورس توسعه می‌دهند [١٧]. ووسلر در مطالعه خود رویکرد کاملا بیزینی[31] برای آزمون ریشه واحد[32] با چندین شکست ساختاری برای داده‌های مربوط به نرخ بیکاری سالانه ١٧ کشور عضو سازمان همکاری‌های اقتصادی و توسعه[33] معرفی می‌کند تا بتواند به این سوال پاسخ دهد که آیا بعد از شوک به بازار نیروی کار، ثبات مشاهده می‌شود یا خیر [١٨]. هپسن و وطنسور (٢٠١١)، در مطالعه خود به‌دنبال پیش‌بینی روندهای آینده در بازار مسکن دوبی با استفاده از مدل آریما هستند. آنها از داده‌های

سری زمانی ماهانه «شاخص قیمت املاک مسکونی دوبی»[۳۴] استفاده کرده و مقادیر آن را در بازه زمانی ژانویه تا دسامبر ۲۰۱۱ پیش‌بینی می‌کنند [۱۹]. استیونسون (۲۰۰۷)، در مطالعه خود به مقایسه قابلیت‌های مدل‌های مختلف آریما در پیش‌بینی می‌پردازد و برای بررسی موارد فوق، از داده‌های مربوط به اجاره املاک در بریتانیا به‌عنوان یک مطالعه موردی استفاده می‌کند [۲۰].

یکی دیگر از کاربردهای مدل‌های پیش‌بینی سری زمانی، پیش‌بینی میزان تقاضای سفر است. در این زمینه تیموری و همکاران[۳۵](۲۰۰۳) با استفاده از مدل SARIMA[۲]، تقاضای مسافرت ریلی در محور خراسان شبکه راه‌آهن ج.ا.ا. را پیش‌بینی می‌کنند [۲۱].

مدل‌های آریما در پیش‌بینی قابلیت اطمینان سیستم‌های قابل تعمیر و ارزیابی عملکرد مدل فعالیت فرایندهای کسب‌وکار نیز به‌کار می‌روند. هو و ژی[۳۶](۱۹۹۸) در مطالعه خود از مدل آریما برای پیش‌بینی قابلیت اطمینان سیستم‌های قابل تعمیر استفاده می‌کنند. بدین‌منظور آنها یک مثال از خرابی‌های یک سیستم مکانیکی زده و مقایسه‌ای بین مدل آریما و مدل دوآن انجام می‌دهند [۲۲]. لام و همکاران[۳۷](۲۰۰۹)، به‌دنبال ارزیابی عملکرد مدل فعالیت فرایندهای کسب‌وکار با استفاده از تحلیل تداخل سری‌های زمانی آریما هستند. آنها برای پیاده‌سازی مدل آریما از داده‌های شبیه‌سازی‌شده مربوط به فرایند خرید لوازم خانگی یک کارخانه که شامل ۲۰ مورد است، استفاده می‌کنند [۲۳].

مدل‌های آریما در حوزه زنجیره تامین نیز به کار می‌روند که می‌توان به مطالعات بابایی و همکاران[۳۸](۲۰۱۳) و گیلبرت و چاپاتانانان[۳۹] (۲۰۰۶) اشاره کرد. بابایی و همکاران (۲۰۰۶)، در پژوهش خود زنجیره تامین دو مرحله‌ای را در نظر می‌گیرند که توسط یک سازنده و خرده‌فروش ایجاد شده است. در این زنجیره، خرده‌فروش با فرایند تقاضایی مواجه است که از مدل ARIMA(۰،۱،۱) پیروی می‌کند. با توجه به مقدار خطای حداقل مربعات و مقایسه نتایج پیش‌بینی شده با مقادیر واقعی، کارایی بالای این مدل آریما مشخص می‌شود [۲۴]. گیلبرت و چاپاتانانان (۲۰۰۶)، از مدل‌های آریمای زنجیره تامین استفاده می‌کنند تا روش‌هایی را برای برنامه‌ریزی تولید ارائه دهند که بتوند اثر شلاقی را کاهش دهد [۲۵].

همان‌طور که از بررسی مقالات مطالعه‌شده پیداست، تاکنون از مدل‌های آریما برای پیش‌بینی سری زمانی شاخص قیمت مصرف‌کننده در ایران استفاده نشده است. همچنین بررسی مانایی سری‌های زمانی در مقالات تنها منحصر به موضوع پدیده بازگشت به میانگین در بورس اوراق بهادار تهران بوده است و این پژوهش با بررسی مانایی سری زمانی شاخص قیمت مصرف‌کننده از طریق آزمون ریشه واحد و سپس پیش‌بینی مقادیر آینده آن با مدل آریما سعی در پرکردن این شکاف تحقیقاتی دارد.

۲. مبانی نظری

یک سری زمانی تصادفی در صورتی ماناست که مشخصات توزیع احتمال آن به زمان بستگی نداشته باشد. به‌عبارت دیگر میانگین و واریانس آن طی زمان ثابت باقی بماند و همبستگی بین متغیر و متغیر باوقفه تنها بستگی به طول وقفه داشته باشد. برای استفاده از روش حداقل مربعات معمولی[۴۰] سری زمانی باید مانا باشد چرا که استفاده از این روش در مورد سری‌های زمانی غیرمانا منجر به نتایج ساختگی[۴۱] می‌شود. یکی از آزمون‌های نسبتا جدیدی که برای بررسی مانایی یا عدم مانایی یک سری زمانی به‌کار می‌رود، آزمون ریشه واحد است.

مدل‌های خودرگرسیو با میانگین متحرک انباشته شکل تعمیم‌یافته مدل‌های ساده خودرگرسیو می‌باشند که در آن از سه وسیله برای مدل‌سازی خودهمبستگی در عبارات خطا استفاده می‌شود [۲۲], [۲۷], [۲۶]:

- مدل خودرگرسیو مرتبه p، AR(p)
- مدل میانگین متحرک، MA(q)

- ARMA(p,q)

در انتخاب یک مدل آریما از سه ابزار سری‌های زمانی به نامهای تفاضل‌گیری[۲]، تابع خودهمبستگی و تابع خودهمبستگی جزئی استفاده می‌شود. یکی از کاربردهای تابع خودهمبستگی، بررسی داده‌های سری زمانی جهت داشتن یکی از مشخصه‌های مدل آریما است که عبارتند از: ۱- تابع خودهمبستگی بعد از چند وقفه صفر می‌شود و ۲- یا به‌طور نمایی کاهش می‌یابد. با افزایش وقفه k، ضریب خودهمبستگی پس از چند وقفه به صفر می‌رسد، در غیر این‌صورت به‌سرعت کاهش می‌یابد. اگر نمودار تابع خودهمبستگی هیچ کدام از این دو مشخصه یک مدل آریما را نشان ندهد، ممکن است تفاضل‌گیری داده‌های سری زمانی لازم شود. سری $y_t - y_{t-1}$ تفاضل اول سری زمانی اولیه نامیده می‌شود و به‌طور خلاصه به‌صورت $\Delta y_t = y_t - y_{t-1}$ نشان داده می‌شود. اگر تفاضل‌گیری اول برای ایجاد تابع خودهمبستگی که بعد از چند وقفه به صفر برسد یا به‌صورت نمایی کاهش یابد، کافی نباشد؛ ممکن است تفاضل‌گیری مجدد سری زمانی (تفاضل‌گیری داده‌های تفاضل‌گیری شده) لازم شود. البته اگر تفاضل‌گیری لازم شود معمولا به بیش از یک یا دو بار تفاضل‌گیری برای رسیدن به تابع خودهمبستگی با ویژگی‌های مذکور نیاز نخواهد بود. تعداد دفعاتی که سری زمانی باید تفاضل‌گیری شود با حرف d نشان داده و مرتبه تفاضل‌گیری[۳] نامیده می‌شود. سومین ابزار برای تشخیص یک مدل آریمای مناسب، تابع خودهمبستگی جزئی است. این تابع مشابه تابع خودهمبستگی است با این تفاوت که در آن، مقادیر خودهمبستگی جزئی در مقابل مقادیر وقفه نشان داده می‌شوند. ضرایب خودهمبستگی جزئی، سهم اضافه‌شدن یک وقفه بیشتر را نشان می‌دهد و اثرات وقفه‌های کوچک‌تر را در نظر می‌گیرد. جدول (۱ زیر مشخصه‌های توابع خودهمبستگی و خودهمبستگی جزئی را در مدل‌های آریما نشان می‌دهد [۲۸].

جدول (۱): مشخصه‌های توابع ACF و PACF در مدل‌های ARIMA

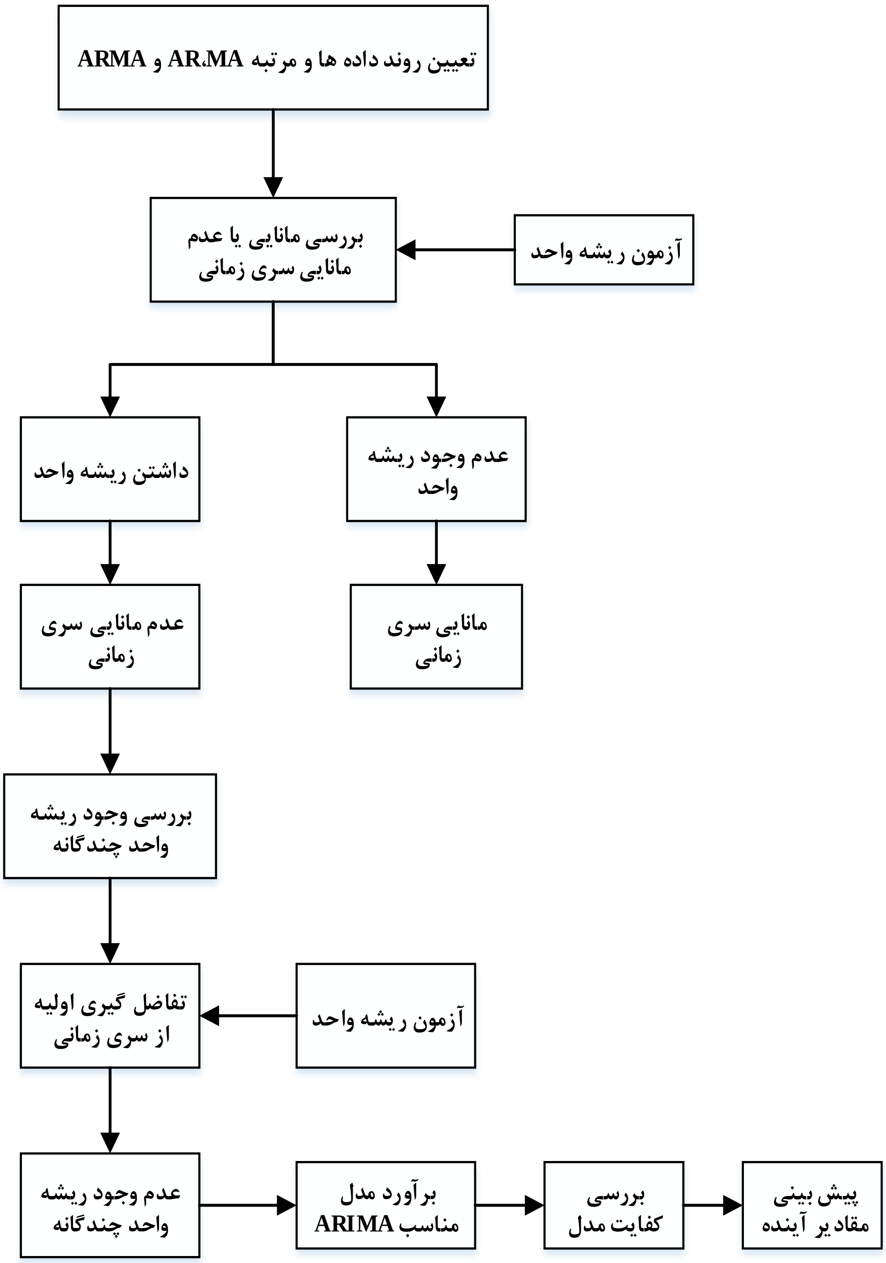
تابع خودهمبستگی جزئی (PACF)	تابع خودهمبستگی (ACF)	نوع مدل
بعد از p وقفه صفر می‌شود	به‌طور نمایی کاهش می‌یابد	AR(p)
به‌طور نمایی کاهش می‌یابد	بعد از q وقفه صفر می‌شود	MA(q)
به‌طور نمایی کاهش می‌یابد	به‌طور نمایی کاهش می‌یابد	ARMA(p,q)

به‌طور کلی اگر بخواهیم از یک مدل برای پیش‌بینی استفاده کنیم، باید ویژگی‌های این مدل در طی زمان ثابت باشد. بنابراین دلیل نیاز به داده‌های ایستا آن است که هر مدلی که از این داده‌ها به‌دست آید، باثبات بوده و می‌تواند مبنای معتبری برای پیش‌بینی محسوب شود. مراحل ساخت یک مدل ARIMA و استفاده از آن برای پیش‌بینی به‌صورت زیر است [۲۸], [۳]:

- مرحله تشخیص: تعیین مقادیر p، q و d با استفاده از نمودارهای خودهمبستگی و خودهمبستگی جزئی.
- مرحله تخمین پارامترهای مدل: استفاده از روش حداقل مربعات یا استفاده از روش‌های غیر خطی در صورت غیر خطی‌بودن مدل نسبت به پارامترها.
- کنترل تشخیصی[۴]: در صورتی‌که مدل برازش‌داده‌شده، انطباق خوبی ایجاد کرده باشد، مقادیر باقیمانده‌ها اختلال (نوفه) سفید می‌شوند.
- پیش‌بینی برای سالهای آتی

۳. روش تحقیق

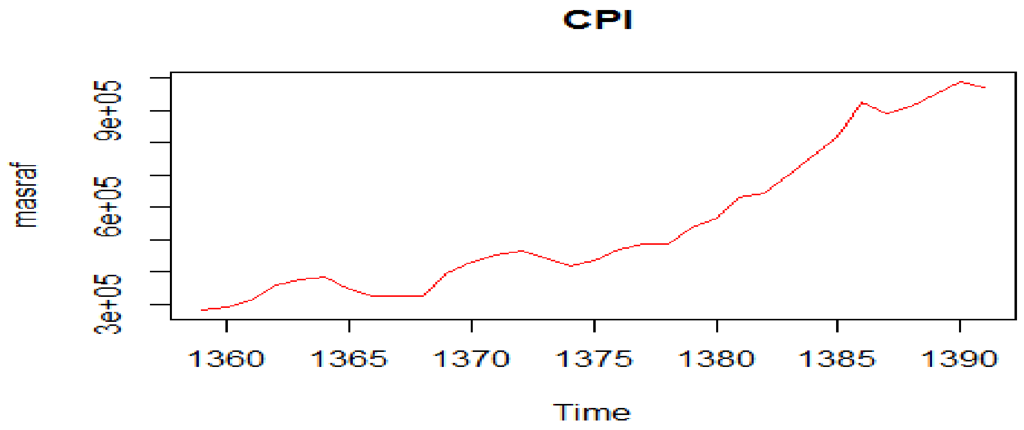
در نمودار زیر، گامهای کلی انجام تحقیق نشان داده شده است.



نمودار (۱): شمای کلی گامهای انجام تحقیق

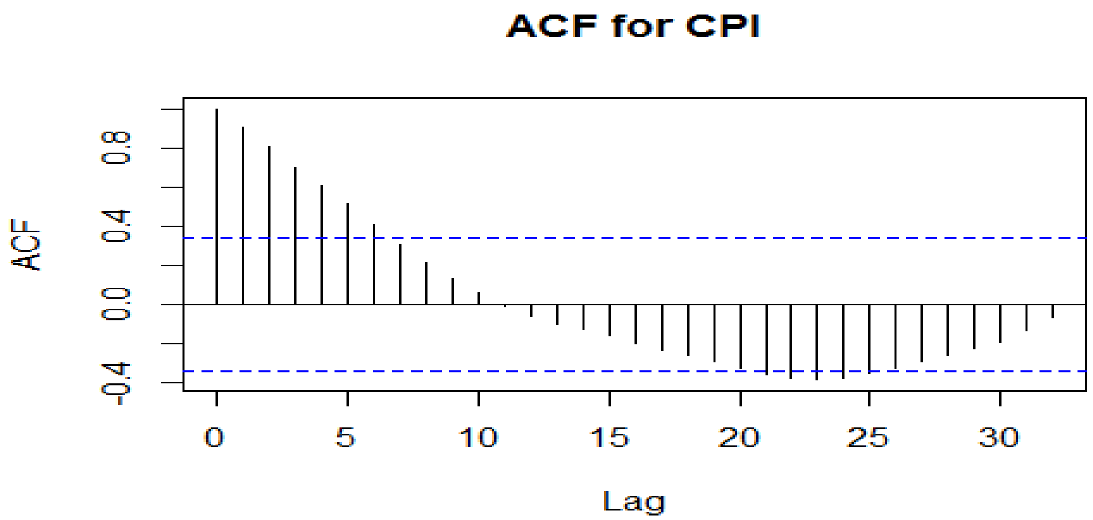
۳.۱. روند داده‌ها و مرتبه AR(p)، MA(q) و ARMA(p,q)

روند داده‌های هزینه مصرفی خانوار به قیمت ثابت سال ۱۳۸۳ در نمودار (۲)نشان داده شده است .[۲۹]



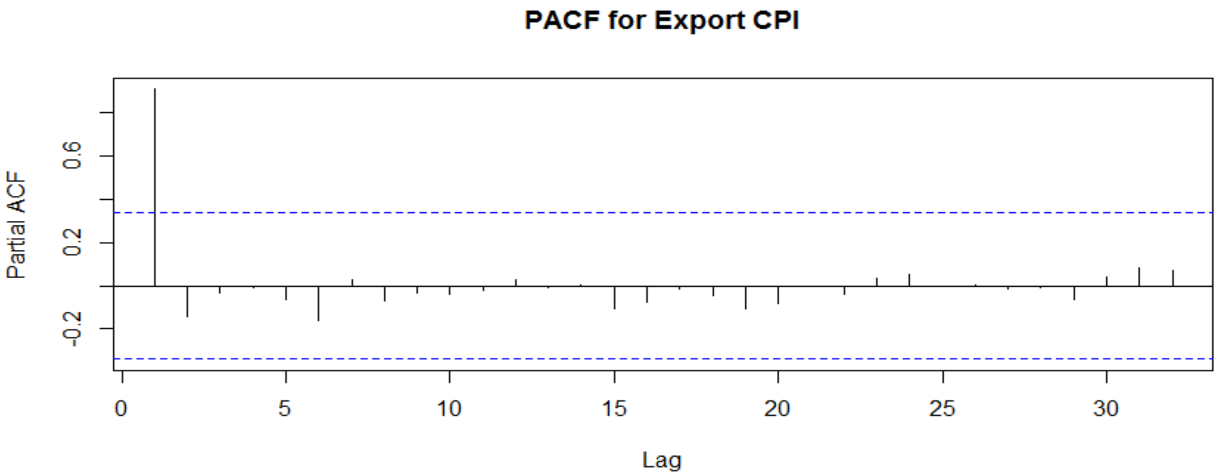
نمودار (۲): روند سری زمانی شاخص قیمت مصرف‌کننده به قیمت ثابت سال ۱۳۸۳

برای یک سری زمانی MA(q) تابع خودهمبستگی آن بعد از q دوره به صفر می‌گراید و این یعنی مقدار متغیر در دوره فعلی صرفا تابعی از q دوره قبل متغیر است و لذا جایی که ACF به صفر می‌رسد مرتبه MA(q) تعیین می‌شود. این نکته در نمودار (۳زیر نشان داده شده است.



نمودار (۳): تابع خودهمبستگی شاخص قیمت مصرف‌کننده به قیمت ثابت ۱۳۸۳

برای هر سری زمانی بعد از p دوره PACF آن به صفر می‌رسد و بنابراین جایی که تابع خودهمبستگی جزئی به صفر می‌گراید، مرتبه AR(p) تعیین می‌شود. این نکته در نمودار زیر نشان داده شده است.



نمودار (۴): تابع خودهمبستگی جزئی شاخص قیمت مصرف‌کننده به قیمت ثابت ۱۳۸۳

با توجه به این نمودارها مرتبه خودرگرسیو برابر $p = 1$ است و مرتبه MA با توجه به روند مشخص در نمودار خودهمبستگی ظاهرا برابر صفر است، اما ما با توجه به معیار اطلاعاتی آکائیک٦ادامه آن را برابر دو در نظر می‌گیریم.

```
Series: x
ARIMA(1,0,0) with non-zero mean

Coefficients:
         ar1  intercept
      0.9894   392570.2
s.e.  0.0140   166697.3

sigma^2 estimated as 671620549:  log likelihood=-360.96
AIC=727.92   AICc=728.81   BIC=732.22

Series: x
ARIMA(1,0,1) with non-zero mean

Coefficients:
         ar1     ma1  intercept
      0.9862  0.1583   393357.2
s.e.  0.0182  0.1511   162744.2

sigma^2 estimated as 646868189:  log likelihood=-360.4
AIC=728.81   AICc=730.34   BIC=734.54
```

```
Series: x
ARIMA(1,0,2) with non-zero mean

Coefficients:
         ar1      ma1     ma2  intercept
      0.9846  -0.2789  0.6509   382791.8
s.e.  0.0201   0.1745  0.1884   165102.0

sigma^2 estimated as 5.44e+08:  log likelihood=-358.38
AIC=726.75   AICc=729.15   BIC=733.92
```

برای محاسبه arimafit ابتدا مرتبه MA را صفر و سپس یک و نهایتا دو در نظر می‌گیریم. همانطور که مشهود است کمترین مقدار معیار اطلاعاتی آکائیک مربوط به ARMA(۱،۲) است که بنابراین بهترین مرتبه برای MA برابر دو خواهد بود زیرا دارای کمترین مقدار برای معیار اطلاعاتی آکائیک است.

۲.۳. بررسی مانایی داده‌ها

یک مدل ARMA(۱،۲) را که در آن تفاضل‌گیری انجام نشده‌است،در نظر می‌گیریم. در یک مدل ARMA(p,q)، قسمت MA(q) مدل همیشه مانا است چرا که دارای مجموعه‌ای از فرایندهای نوفه سفید است که همه آنها ناهمبسته هستند و بنابراین کافی است که ما قسمت AR(p) مدل را از نظر مانایی بررسی کنیم. شرط مانایی قسمت AR(p) این است که همه ریشه‌های معادله مفسر آن خارج از دایره واحد قرار گیرند. در واقع همه ضرایب $\varphi_1, \varphi_2, \ldots, \varphi_p$ در معادله زیر، بیشتر از عدد ۱ باشند.

معادله ۱)

$$\varphi(B) = (1 - \phi_1 B - \varphi_2 B^2 - \ldots - \varphi_p B^p)$$

در این پژوهش چون مرتبه AR(۱) یک است، بنابراین تنها یک ریشه داریم و بنا به محاسبات این ریشه $\varphi_1 = 0.945062$ است و داخل دایره واحد قرار دارد که این به معنای عدم مانایی داده‌ها در مدل ARMA(p,q) است.

```
Series: x
ARIMA(1,0,2) with non-zero mean

Coefficients:
         ar1      ma1     ma2  intercept
      1.0581  -0.1724  0.2155   97192.71
s.e.  0.0364   0.2133  0.3105  151129.90

sigma^2 estimated as 447206710:  part log likelihood=-352.72

Training set error measures:
                    ME     RMSE     MAE        MPE     MAPE      MASE     ACF1
Training set -128.3556 20803.38 17681.7 -0.7105692 6.423064 0.8074769 0.027534
> abs(polyroot(c(1,- armafit1$coef[1:1])))
[1] 0.945062
```

بنابراین این احتمال وجود دارد که داده‌های این سری زمانی، غیرمانا (دارای ریشه واحد) باشند. لذا باید آزمون ریشه واحد را برای این داده‌ها انجام داده و چنانچه فرضیه عدم مانایی صحیح بود از مدل آریما استفاده کنیم. در این قسمت از آزمون دیکی فولر تعمیم‌یافته۴۷ برای تست عدم مانایی استفاده می‌کنیم. فرضیه صفر این آزمون غیر مانابودن داده‌هاست.

$H_0 = NonStationary$

$H_1 = Stationary$

```
> adf.test(x,alternative = "stationary")

  Augmented Dickey-Fuller Test

data:  x
Dickey-Fuller = -1.8349, Lag order = 3, p-value = 0.6363
alternative hypothesis: stationary
```

با توجه به نتایج بالا، فرضیه صفر آزمون دیکی فولر مبنی بر غیرمانایی داده‌ها با P-Value معادل ۰,٦٣٦٣ درصد پذیرفته می‌شود.

۳.۲.۱. آزمون ریشه واحد

اهمیت داشتن ریشه واحد برای یک فرایند تصادفی آن است که اگر شوکی به آن وارد شود، اثر دائمی دارد. یعنی اگر متغیر اقتصادی مانا نباشد، دارای ریشه واحد است و شوک‌هایی که به آن وارد می‌شود دائمی خواهند بود و اگر شوکی دائمی باشد نمی‌تواند ناشی از طرف تقاضا باشد. اغلب متغیرهای اقتصادی از جمله شاخص قیمت مصرف‌کننده عملا سری‌های غیرمانا هستند، بدین معنا که میانگین و واریانس آنها به زمان بستگی دارد و هر چه زمان پیش می‌رود این متغیرها تمایل دارند که از یک مقدار معین دور شوند. چنانچه این حرکت در یک جهت (بالا و پایین) باشد، این سری را سری زمانی دارای روند گویند.

در گام اول باید کلی‌ترین معادله ممکن را برای مدل برآورد کنیم. بنابراین مدل زیر را تخمین می‌زنیم:

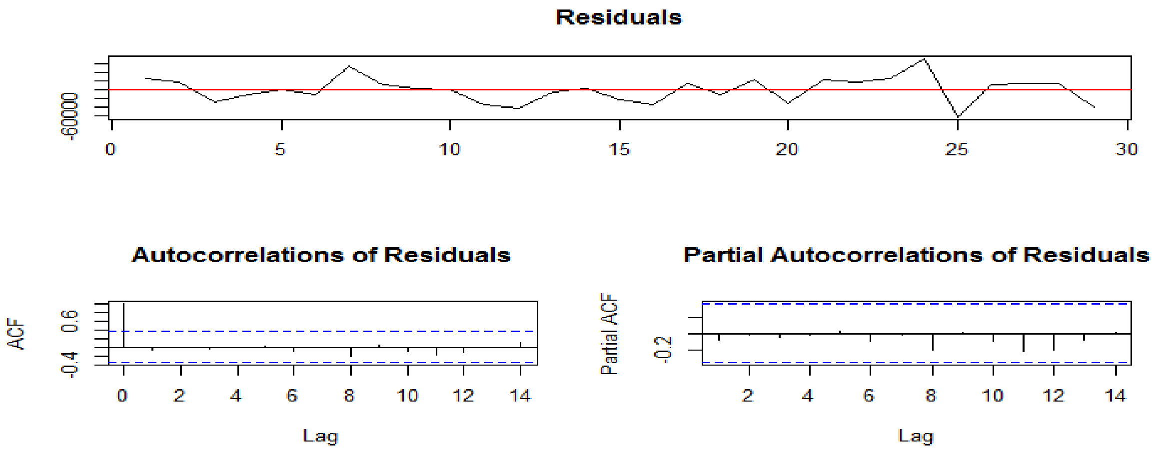
معادله ۲)

$$\Delta x_t = \alpha + \beta t + \delta x_{t-1} + \sum_{i=1}^{k} \delta_i \Delta x_{t-i} + \varepsilon_t$$

برای پیداکردن تعداد وقفه‌های بهینه یا به‌عبارتی تعداد بهینه Δx_{t-i} از معیارهای اطلاعاتی آکائیک و بیزین[^8] استفاده می‌کنیم. در این قسمت در وهله اول از دستور $ur.\text{df}$ استفاده می‌کنیم. در واقع این دستور آزمون دیکی فولر تکمیل شده را انجام می‌دهد. هنگامی که ما نوع آن را روی حالت Trend قرار می‌دهیم، هر دوی عرض از مبدا و روند را مد نظر قرار می‌دهد. همچنین عبارت Lags در آن به معنای تعداد وقفه برای متغیر درون‌زایی است که در آن شامل شده است. همانطور که در نمودار زیر مشهود است، خودهمبستگی و خودهمبستگی جزئی میان باقیمانده‌ها وجود ندارد.

```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  9.523e+02  1.332e+04   0.071    0.944
z.lag.1     -8.252e-02  8.481e-02  -0.973    0.340
tt           2.413e+03  1.158e+03   2.084    0.048
z.diff.lag  -9.870e-02  1.990e-01  -0.496    0.624
Value of test-statistic is: -0.973 4.1472 3.9764

Critical values for test statistics:
      1pct  5pct 10pct
tau3 -4.15 -3.50 -3.18
phi2  7.02  5.13  4.31
phi3  9.31  6.73  5.61
```



نمودار (۵): باقیمانده‌ها و خودهمبستگی بین آنها

در گام دوم آزمون $\delta = 0$ یا به عبارت دیگر $\rho = 1$ را که همان وجود ریشه واحد است، انجام می‌دهیم. عبارت $t-Value$ مربوط به $z.lag.1$ را با مقدار تابع نمونه‌ای τ_t مقایسه می‌کنیم. با توجه به اینکه آزمون τ یک دامنه است و سمت راست مقدار بحرانی آن فضای اطمینان است، بنابراین باید برای قبول فرضیه صفر مبنی بر $\delta = 0$، مقدار $t-Value$ از $tau3 = -3.50$ بیشتر باشد.در اینجا t-Value معادل ۰٫۹۷۳- است و بنابراین فرضیه صفر را نمی‌توانیم رد کنیم. پس به گام بعدی برای آزمون‌کردن روند و عرض از مبدا می‌رویم.

در گام سوم از آزمون ϕ_3 برای آزمودن فرضیه $\beta = 0$ توامان با $\delta = 0$، یعنی آزمون‌کردن معناداری جزء روند با فرض وجود ریشه واحد در سری زمانی x_t در مدل کلی بیان شده در بالا، استفاده می‌کنیم. بنابراین برای انجام این آزمون پس از برآورد آزمون ریشه واحد، تابع نمونه‌ای F را با تحمیل قید $\beta = \phi = 0$ به رگرسیون، محاسبه و با ϕ_3 مقایسه می‌کنیم. در ابتدا معادله کلی نتایج حاصل از مرحله یک و دو را اجرا می‌کنیم و سپس به تحمیل قید ذکر شده می‌پردازیم.

```
Coefficients:
               Estimate Std. Error t value Pr(>|t|)
(Intercept)  -1.818e+04  1.217e+04  -1.494    0.148
xlag[2:30]    1.168e-01  7.786e-02   1.500    0.146
trend[3:31]  -2.625e+02  1.127e+03  -0.233    0.818
dx1lag[1:29] -2.047e-01  1.910e-01  -1.072    0.294
F-statistic: 2.767 on 3 and 25 DF,  p-value: 0.06275
```

حال قید $\beta = \phi = 0$ را تحمیل می‌کنیم. همان‌طور که در گام یک مشاهده شد، آزمون دیکی‌فولر مقدار ϕ_3 مربوط به سری زمانی مورد نظر را محاسبه می‌کند و لذا کافی است که آن مقدار را با مقدار بحرانی ϕ_3 مقایسه کنیم که اگر کوچک‌تر باشد فرضیه صفر را نمی‌توانیم رد کنیم. مقدار $\phi_3 = 3.97$ است که از مقدار بحرانی کمتر است و بنابراین فرضیه صفر رد نمی‌شود.

در گام چهارم، به تست معناداری عرض از مبدا می‌پردازیم. تابع نمونه ای $\tau_{\alpha t}$ را برای آزمودن فرضیه $\alpha = 0$ به شرط $\delta = 0$، یعنی اعمال قید همزمان $\alpha = \delta = 0$ برای آزمون معناداری جز ثابت با فرض وجود ریشه واحد در سری زمانی مورد مطالعه، مورد استفاده قرار می‌دهیم. نرم افزار R همانطور که در گام سوم گفته شد، مقدار ϕ_2 مربوط به آزمون را محاسبه می کند که در اینجا $\phi_2 = 4.1472$ است، که کمتر از مقدار بحرانی بوده و فرضیه صفر را نمی‌توان رد کرد.

در گام پنجم، معادله رگرسیونی زیر را به روش حداقل مربعات معمولی برآورد می‌کنیم.

معادله ۳)

$$\Delta x_t = \alpha + \delta x_{t-1} + \sum_{i=1}^{k} \delta_i \Delta x_{t-i} + \varepsilon_t$$

بنابراین در این مرحله می‌توانیم با استفاده از دستور lm که مربوط به برازش یک مدل خطی است، این رگرسیون را انجام دهیم.

```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -9.186e+03  1.321e+04  -0.695   0.4932
z.lag.1      7.431e-02  4.165e-02   1.784   0.0865
z.diff.lag  -1.193e-01  2.117e-01  -0.564   0.5780
F-statistic: 1.608 on 2 and 25 DF,  p-value: 0.2202


Value of test-statistic is: 1.7841 3.8942

Critical values for test statistics:
      1pct  5pct 10pct
tau2 -3.58 -2.93 -2.60
phi1  7.06  4.86  3.94
```

که در این برازش مدل خطی dx تابعی از x_t دوره قبل، x_t با یک دوره وقفه و عرض از مبدا است، که در آن ۱ نشان‌دهنده وجود عرض از مبدا خواهد بود. با وجود این تابع رگرسیون خطی به گام بعد می‌رویم.

در گام ششم با استفاده از τ_μ فرضیه $\delta = 0$ را در معادله‌ای که در گام پنجم تخمین زده بودیم، بررسی می‌کنیم. حال چنانچه فرضیه صفر رد شود، آن‌گاه سری زمانی دارای ریشه واحد نیست و فرایند گامها متوقف می‌شود و در نتیجه سری زمانی مانا خواهد بود. اما اگر فرضیه صفر رد نشود، باید فرآیند آزمون را برای عرض از مبدا ادامه دهیم. مقدار آماره آزمون داده شده در مرحله پنجم را که متعلق به $tau2 = 1.784$ است، با مقدار بحرانی مقایسه نموده و مشخص می‌شود که فرضیه صفر را نمی‌توان رد کرد.

در گام هفتم هنگامی که در گام ششم، آزمون دارای ریشه واحد بودن مدل برازش شده در گام پنجم تایید شد، آنگاه با استفاده از قید $\alpha = 0$ و همزمان با ریشه واحد داشتن $\delta = 0$، معناداری جز ثابت را بررسی می‌کنیم. از آن‌جا که مقدار آماره ϕ_1 داده‌شده در گام پنجم که برابر ۳،۸۹ است و α اختلاف معناداری با صفر ندارد، فرضیه صفر را نمی‌توان رد کرد.

در گام هشتم باید معادله زیر را به روش حداقل مربعات معمولی برآورد کنیم:
معادله ٤)

$$\Delta x_t = \delta x_{t-1} + \sum_{i=1}^{k} \delta_i \Delta x_{t-i} + \varepsilon_t$$

که نتایج آن بدین صورت است:

```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
z.lag.1        0.04744    0.01540   3.081  0.00483
z.diff.lag    -0.08783    0.20470  -0.429  0.67140

Value of test-statistic is: 3.0811

Critical values for test statistics:
      1pct  5pct 10pct
tau1 -2.62 -1.95 -1.61
```

در گام نهم فرضیه وجود ریشه واحد $\delta = 0$ را باید در برآورد بالا آزمون کنیم و نهایتا اگر فرضیه صفر رد شود نتیجه خواهیم گرفت که سری زمانی فاقد ریشه واحد بوده و ماناست و در غیر این صورت، سری زمانی ریشه واحد داشته و مانا نیست. همان‌طور که در برآورد گام هشتم مشاهده شد، مقدار آماره آزمون $tau1 = 3.08$ و مقدار بحرانی آن در سطح ٥٪ معادل $tau1 = -1.95$ است و فرضیه صفر را نمی‌توان رد کرد و در نتیجه سری زمانی حداقل یک ریشه واحد دارد.

۲.۲.۳. آزمون ریشه‌های واحد چندگانه

تا این قسمت ما به وجود حداقل یک ریشه واحد پی برده‌ایم. برای مشخص‌کردن این که آیا ریشه واحد دیگری برای این سری زمانی وجود دارد یا خیر، باید با یکبار تفاضل‌گیری از سری زمانی اثر ریشه واحد کشف شده را در بالا از بین ببریم و آنگاه مجددا ۹ گام بالا را آزمون کنیم. به‌طور کلی یک سری زمانی از لحاظ نظری می‌تواند دارای $d \in \mathbb{Z}^+$ ریشه واحد باشد و این یعنی برای ماناسازی آن باید به تعداد d بار از سری زمانی مورد مطالعه تفاضل‌گیری کرد. اگر مانایی داده‌ها در این قسمت اثبات شود نتیجه می‌گیریم که سری زمانی شاخص قیمت مصرف‌کننده تنها یک ریشه واحد داشته است. بنابراین در این قسمت، از داده‌ها یک بار تفاضل‌گیری کرده و مجددا گامها را تکرار می‌کنیم.

برای داده‌های تفاضل‌گیری‌شده، معادله زیر را با روش حداقل مربعات معمولی برآورد کنیم:
معادله ٥)

$$\Delta x_t = \alpha + \beta t + \delta x_{t-1} + \sum_{i=1}^{k} \delta_i \Delta x_{t-i} + \varepsilon_t$$

که در آن منظور از x_t همان داده‌های تفاضل‌گیری‌شده است. در ادامه آزمون $\delta = 0$ یا به‌عبارت دیگر $\rho = 1$ را که همان وجود ریشه واحد است، انجام می‌دهیم. عبارت $t - Value$ مربوط به $z.lag.1$ را با مقدار تابع نمونه‌ای τ_t مقایسه می‌کنیم. با توجه به اینکه آزمون τ یک دامنه است و سمت راست مقدار بحرانی آن فضای اطمینان است، بنابراین باید برای قبول فرضیه صفر مبنی بر $\delta = 0$، مقدار $t - Value$ از $tau3 = -3.50$ بیشتر باشد. در اینجا $t - Value$ یا همان $tau3$ برابر ‎-٦،٠٢ است و بنابراین فرضیه صفر رد می‌شود که به‌معنای عدم وجود ریشه واحد در داده‌های سری زمانی تفاضل‌گیری شده است و در نتیجه مانایی آن اثبات می‌شود.

```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -3282.7843  8252.1809  -0.398   0.6940
z.lag.1        -1.1678     0.1939  -6.021 2.33e-06

Value of test-statistic is: -6.0213 12.1629 18.1289

Critical values for test statistics:
      1pct  5pct 10pct
tau3 -4.15 -3.50 -3.18
phi2  7.02  5.13  4.31
phi3  9.31  6.73  5.61
```

بنابراین فرایند آزمون‌کردن مجدد، در همین گام به پایان می‌رسد و در مجموع تنها یک ریشه واحد برای داده‌های سری زمانی هزینه مصرفی خانوار به قیمت ثابت سال ۱۳۸۳ وجود دارد.

۳.۳. برآورد مدل ARIMA (p,d,q)

به‌طور کلی یک سری زمانی ممکن است ابتدا نیاز به d بار تفاضل‌گیری برای رسیدن به مانایی داشته باشد و سری‌هایی که بدین صورت به‌دست می‌آیند، ممکن است خودهمبسته باشند. اگر این نوع خودهمبستگی با استفاده از فرایند ARMA(p,q) مدل‌سازی شود، آنگاه مدل‌سازی اصلی دارای فرمول‌بندی زیر است:

معادله ٦)

$$\phi(B)\Delta^d x_t = \theta(\mathrm{B})\varepsilon_t$$

که یک فرایند خودرگرسیو با میانگین متحرک انباشته از مرتبه‌های p، d و q است و x_t را انباشته از مرتبه d می‌نامند که به صورت I(d) نشان داده می‌شود. پس برای مدل ARIMA در این پژوهش، d=1 است .لذا می‌توان نتیجه گرفت که مرتبه AR برابر یک و مرتبه MA برابر صفر است، اما با توجه به اینکه ARMA (۱،۲) کمترین مقدار معیار اطلاعاتی آکائیک را می‌دهد، ما از عدد ۲ به عنوان مرتبه MA استفاده می‌کنیم.

```
Series: x
ARIMA(1,0,2) with non-zero mean

Coefficients:
         ar1      ma1     ma2  intercept
      0.3903  -0.6090  0.6236  12466.635
s.e.  0.2514   0.2175  0.1932   6018.864

sigma^2 estimated as 424267529:  log likelihood=-341.05
AIC=692.11   AICc=694.61   BIC=699.11

Training set error measures:
                    ME     RMSE      MAE     MPE    MAPE      MASE       ACF1
Training set -44.96471 20597.76 17893.66 140.474 354.196 0.7408678 0.07115756
> 
```

نهایتا در برآورد مدل ARIMA(۱،۱،۲) خواهیم داشت:

```
Series: x
ARIMA(1,1,2)

Coefficients:
         ar1      ma1     ma2
      0.5718  -0.6944  0.6726
s.e.  0.1917   0.1764  0.1901

sigma^2 estimated as 462636938:  log likelihood=-342.55
AIC=693.11   AICc=694.71   BIC=698.71

Training set error measures:
                   ME     RMSE      MAE     MPE     MAPE     MASE       ACF1
Training set 5610.081 21159.27 18739.07 1.43421 6.592072 0.855764 0.01419771
```

اگر ریشه این مدل را به‌دست آوریم، خارج از دایره واحد قرار می‌گیرد که بیانگر مانایی این سری زمانی است.

```
> abs(polyroot(c(1,- armafit1$coef[1:1])))
[1] 1.748786
```

در مورد معناداری ضرایب، آزمون زیر را انجام می‌دهیم:

```
> (1-pnorm(abs(armafit2$coef)/sqrt(diag(armafit2$var.coef))))*2
        ar1         ma1         ma2
0.025154413 0.005720843 0.022991982
```

همانطور که مشاهده می‌شود، تمام ضرایب معنادار هستند و این را می‌توان از طریق محاسبه مقدار احتمال و همچنین دستور آزمون ضرایب به‌دست آورد.

```
> coeftest(armafit2)

z test of coefficients:

    Estimate Std. Error z value Pr(>|z|)
ar1  0.51811    0.23140  2.2390 0.025154 *
ma1 -0.58127    0.21035 -2.7634 0.005721 **
ma2  0.57184    0.25152  2.2736 0.022992 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

حال اگر به جای مرتبه p=1 برای خودرگرسیو، از مرتبه های بالاتری استفاده کنیم. بنا به تخمین‌های زیر، مقدار احتمال و ضریب آماره آزمون، مرتبه های بالاتر بی معنا می‌شوند و تنها مرتبه مناسب برای این مدل همان p=1 خواهد بود. در محاسبات زیر آماره z برای ARIMA(۲،۱،۲) محاسبه شده است. مشهود است که ضریب ar2 معنادار نیست.

```
> coeftest(armafit2)

z test of coefficients:

    Estimate Std. Error z value Pr(>|z|)
ar1  0.77731    0.36233  2.1453 0.031928 *
ar2 -0.27403    0.40734 -0.6727 0.501117
ma1 -0.84894    0.29441 -2.8836 0.003932 **
ma2  0.84956    0.33412  2.5427 0.011001 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

همچنین این آزمون برای ARIMA(۳،۱،۲) نیز انجام شده‌است که همان نتایج بالا را به همراه دارد.

```
> coeftest(armafit2)

z test of coefficients:

    Estimate Std. Error z value Pr(>|z|)
ar1  0.35395    0.50630  0.6991 0.484490
ar2 -0.17833    0.26781 -0.6659 0.505479
ar3  0.33438    0.34173  0.9785 0.327830
ma1 -0.39322    0.52564 -0.7481 0.454410
ma2  0.63887    0.20309  3.1458 0.001657 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

۴.۳. آزمون پورتمن- تیو ۴٫بررسی کفایت مدل

در آزمون پورتمن-تیو برای تعیین خالص تصادفی (نوفه سفید) بودن یک سری زمانی، فرضیه صفر این‌گونه تعریف می‌شود که به‌طور مشترک ضرایب خودهمبستگی سری زمانی صفر باشد. باکس و پیرس تابع نمونه‌ای پورتمن-تیو زیر را پیشنهاد کردند:

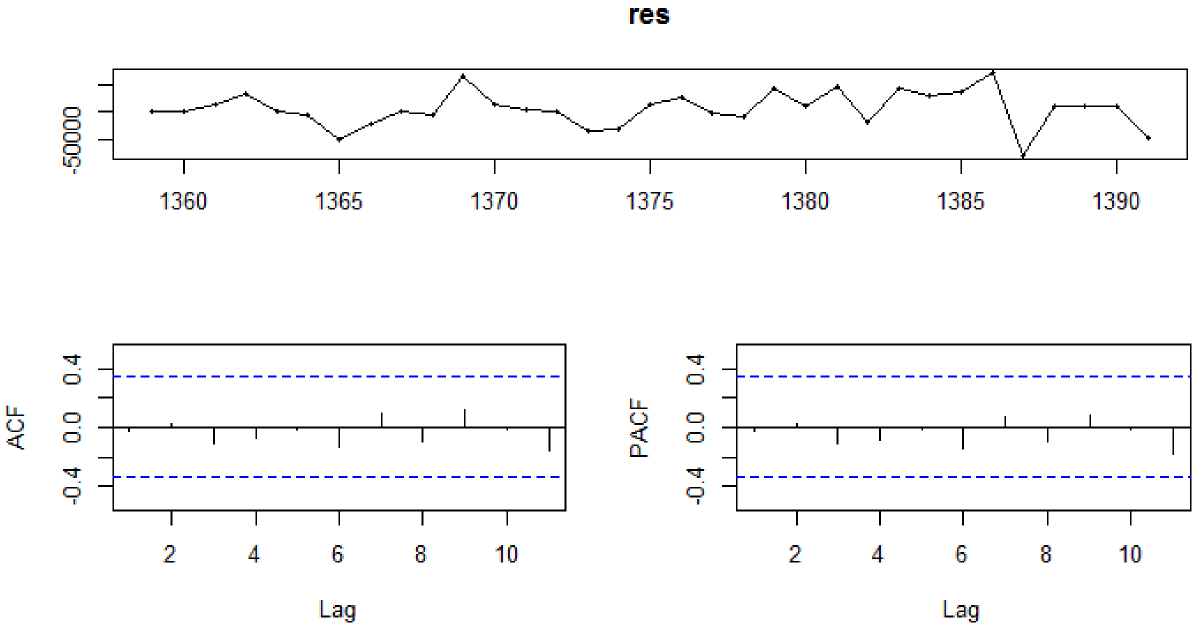
معادله ۷)

$$Q^*(m) = T\sum_{k=1}^{m} r_k^2$$

در روش باکس-پیرس فرضیه H_0 که در واقع بیانگر نوفه سفید بودن یک دنباله تصادفی است، به‌صورت زیر تعریف می‌شود:

$$H_0 : \rho_1 = \rho_2 = \ldots = \rho_m = 0$$

حال این آزمون را بر روی این مدل ARIMA(۱،۱،۲) اجرا می‌کنیم که نتایج آن در نمودار ۶) زیر مشاهده می‌شود.



نمودار (۶): آزمون مربوط به نمودار تابع خودهمبستگی و خودهمبستگی جزئی باقیمانده‌ها

تحلیل خود را از این آزمون برای باقیمانده‌های مدل ارائه می‌دهیم.

```
> Box.test(armafit2$residuals,lag=2)

 Box-Pierce test

data:  armafit2$residuals
X-squared = 0.1088, df = 2, p-value = 0.9471
```

همان‌طور که از نتایج مشاهده می‌شود، مقدار P-Value برای تمام ضرایب خودهمبستگی بسیار بالا است و فرضیه H_0 رد نخواهد شد. یعنی این فرض را که تمام ضرایب خودهمبستگی برابر صفر به‌طور همزمان برابر صفر می‌باشند، می‌پذیریم که این به معنی نوفه سفیدبودن سری زمانی باقیمانده‌های مدل است.

۵.۳. پیش‌بینی

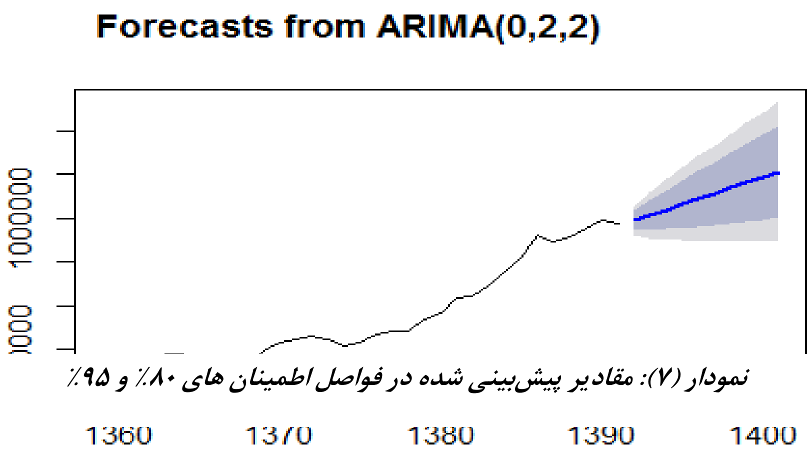
با توجه به اینکه مدل ارائه‌شده کفایت لازم را در آزمون‌های انجام شده تا به این‌جا داشته است، بنابراین می‌توانیم مقادیر آینده شاخص قیمت مصرف‌کننده به قیمت ثابت ۱۳۸۳ را برای سالهای بعد از ۱۳۹۱ پیش‌بینی کنیم.

در این قسمت از دستور forecast(arima2, h=10) برای پیش‌بینی استفاده می‌کنیم، که نتایج به‌صورت زیر خواهد بود.

جدول (۲): مقادیر حاصل از پیش‌بینی سری زمانی

سال	پیش‌بینی نقطه‌ای	حد پایین فاصله اطمینان ۸۰درصد	حد بالای فاصله اطمینان ۸۰درصد	حد پایین فاصله اطمینان ۹۵درصد	حد بالای فاصله اطمینان ۹۵درصد
۱۳۹۲	۹۸۸۹۰۸,۳	۹۴۶۳۵۶,۸	۱۰۳۱۴۶۰	۹۲۳۸۳۱,۵	۱۰۵۳۹۸۵
۱۳۹۳	۱۰۱۳۶۶۲,۹	۹۴۵۱۶۶,۹	۱۰۸۲۱۵۹	۹۰۸۹۰۷,۳	۱۱۱۸۴۱۹
۱۳۹۴	۱۰۳۸۴۱۷,۵	۹۴۹۱۳۵,۸	۱۱۲۷۶۹۹	۹۰۱۸۷۲,۹	۱۱۷۴۹۶۲
۱۳۹۵	۱۰۶۳۱۷۲,۲	۹۵۵۱۲۴,۶	۱۱۷۱۲۲۰	۸۹۷۹۲۷,۷	۱۲۲۸۴۱۷
۱۳۹۶	۱۰۸۷۹۲۶,۸	۹۶۲۱۲۵,۸	۱۲۱۳۷۲۸	۸۹۵۵۳۰,۸	۱۲۸۰۳۲۳
۱۳۹۷	۱۱۱۲۶۸۱,۴	۹۶۹۶۷۳,۹	۱۲۵۵۶۸۹	۸۹۳۹۷۰,۳	۱۳۳۱۳۹۲
۱۳۹۸	۱۱۳۷۴۳۶,۰	۹۷۷۵۱۳,۸	۱۲۹۷۳۵۸	۸۹۲۸۵۶,۱	۱۳۸۲۰۱۶
۱۳۹۹	۱۱۶۲۱۹۰,۶	۹۸۵۴۹۰,۷	۱۳۳۸۸۹۰	۸۹۱۹۵۱,۵	۱۴۳۲۴۳۰
۱۴۰۰	۱۱۸۶۹۴۵,۲	۹۹۳۵۰۴,۲	۱۳۸۰۳۸۶	۸۹۱۱۰۲,۷	۱۴۸۲۷۸۸
۱۴۰۱	۱۲۱۱۶۹۹,۸	۱۰۰۱۴۸۵,۸	۱۴۲۱۹۴۱	۸۹۰۲۰۵,۳	۱۵۳۳۱۹۴

نتایج پیش‌بینی در نمودار (۷ آورده شده است.



نمودار (۷): مقادیر پیش‌بینی شده در فواصل اطمینان های ۸۰٪ و ۹۵٪

نتیجه‌گیری

در این پژوهش، داده‌های سری زمانی شاخص قیمت مصرف‌کننده را از سال ۱۳۵۹ تا ۱۳۹۱ به قیمت‌های ثابت ۱۳۸۳ بررسی و با انجام آزمون ریشه واحد، عدم مانایی آن را اثبات می‌کنیم. در این راستا با استفاده از نرم‌افزار R-Project و طی ۹ گام، مشخص می‌شود که این سری زمانی حداقل یک ریشه واحد داشته و مانا نیست. در ادامه با یک‌بار تفاضل‌گیری از داده‌های سری زمانی و اجرای مجدد ۹ گام فوق، ثابت می‌کنیم که این سری زمانی تنها یک ریشه واحد دارد. سپس از آنجا که مدل‌های آریما، وابستگی زیادی به داده‌ها ندارند و همچنین از دقت مناسبی در پیش‌بینی برخوردارند، یک مدل ARIMA(۱،۱،۲) را برآورد می‌کنیم و با انجام آزمون معناداری ضرایب نشان می‌دهیم که اگر از ضرایب خودرگرسیو بالاتر از ۱ یعنی ۲ و ۳ در برآورد مدل آریما استفاده کنیم، مرتبه‌های بالاتر بی‌معنا شده و

مناسب‌ترین مدل، همان مدل ARIMA(۱،۱،۲) است. در ادامه برای اثبات کفایت مدل از تابع نمونه‌ای باکس-پیرس آزمون پورتمن-تیو برای بررسی تصادفی خالص بودن سری زمانی شاخص قیمت مصرف‌کننده استفاده می‌کنیم. در نهایت پس از حصول اطمینان از تصادفی خالص‌بودن سری زمانی، از مدل آریمای برآوردشده برای پیش‌بینی مقادیر آینده شاخص قیمت مصرف‌کننده از سال ۱۳۹۲ تا ۱٤۰۱ در دو سطح اطمینان ۸۰ و ۹۵ درصد استفاده می‌کنیم. همان‌طور که از نمودار پیش‌بینی مشخص است، برآورد نقطه‌ای شاخص قیمت مصرف-کننده روند صعودی داشته و از سال ۱۳۹۲ تا ۱٤۰۱ حدود ۲۲۳۰۰۰ واحد افزایش می‌یابد. از نتایج این مدل می‌توان برای کمک به تصمیم-گیران حوزه اقتصادی کشور برای اتخاذ تصمیمات و راهبردهای لازم در راستای کاهش نرخ تورم و شاخص قیمت مصرف‌کننده استفاده کرد.

۵. منابع

[1] G. C. Tiao, “Time Series: ARIMA methods,” Int. Encycl. Soc. Behav. Sci., pp. 316–321, 2015.

[2] K. Chapman, “The Consumer Price Index: A History and Source List,” Ref. Serv. Rev., vol. 13, no. 4, pp. 47–51, 1985.

[3] R. E. Abdel-Aal and A. Z. Al-Garni, “Forecasting monthly electric energy consumption in eastern Saudi Arabia using univariate time-series analysis,” Energy, vol. 22, no. 11, pp. 1059–1069, 1997.

[4] Y. Wang, J. Wang, G. Zhao, and Y. Dong, “Application of residual modification approach in seasonal ARIMA for electricity demand forecasting: A case study of China,” Energy Policy, vol. 48, pp. 284–294, ۲۰۱۲.

[5] V. Ş. Ediger, S. Akar, and B. Uğurlu, “Forecasting production of fossil fuel sources in Turkey using a comparative regression and ARIMA model,” Energy Policy, vol. 34, no. 18, pp. 3836–3846, 2006.

[6] V. Ş. Ediger and I. Berk, “Crude oil import policy of Turkey: Historical analysis of determinants and implications since 1968,” Energy Policy, vol. 39, no. 4, pp. 2132–2142, 2011.

[7] M. Tabesh, Z. Banafsheh, and A. J. Khoshkholgh, “Application of time series analyzing in forecasting daily demand of Tehran drinkable water,” in National Congress of Civil Engineering, 2004.

[8] S. Barak and S. S. Sadegh, “Forecasting energy consumption using ensemble ARIMA–ANFIS hybrid algorithm,” Int. J. Electr. Power Energy Syst., vol. 82, pp. 92–104, 2016.

[9] C. Yuan, S. Liu, and Z. Fang, “Comparison of China’s primary energy consumption forecasting by using ARIMA (the autoregressive integrated moving average) model and GM (1, 1) model,” Energy, vol. 100, pp. 384–390, 2016.

[10] P. Narayanan, A. Basistha, S. Sarkar, and S. Kamna, “Trend analysis and ARIMA modelling of pre-monsoon rainfall data for western India,” Comptes Rendus Geosci., vol. 345, no. 1, pp. 22–27, 2013.

[11] J. Jia, J. Zhao, H. Deng, and J. Duan, “Ecological footprint simulation and prediction by ARIMA model—A case study in Henan Province of China,” Ecol. Indic., vol. 10, no. 2, pp. 538–544, 2010.

[12] R. O. Yusuf, Z. Zainon Noor, A. Halilu Abba, M. Ariffin Abu Hassan, M. Rafee Majid, and N. Idris Medugu, “Predicting methane emissions from livestock in Malaysia using the ARIMA model,” Manag. Environ. Qual. An Int. J., vol. 25, no. 5, pp. 585–599, 2014.

[13] K. Taneja, S. Ahmad, K. Ahmad, and S. D. Attri, “Time series analysis of aerosol optical depth over New Delhi using Box–Jenkins ARIMA modeling approach,” Atmos. Pollut. Res., 2016.

[14] H. Karbasi Yazdi, Y. Nourifard, and H. Chenari Bouket, “Study of regression toward the mean phenomenon in Tehran stock exchange using unit root test,” Knowl. Invest., vol. 4, pp. 87–103, 2013.

[15] A. Jadevicius and S. Huston, “ARIMA modelling of Lithuanian house price index,” Int. J. Hous. Mark. Anal., vol. 8, no. 1, pp. 135–147, 2015.

[16] Y. Du, Y. Cai, M. Chen, W. Xu, H. Yuan, and T. Li, “A Novel Divide-and-Conquer Model for CPI Prediction Using ARIMA, Gray Model and BPNN,” Procedia Comput. Sci., vol. 31, pp. 842–851, 2014.

[17] C. N. Babu and B. E. Reddy, “A moving-average filter based hybrid ARIMA–ANN model for forecasting time series data,” Appl. Soft Comput., vol. 23, pp. 27–3٨, ٢٠١٤.

[18] A. Vosseler, “Bayesian model selection for unit root testing with multiple structural breaks,” Comput. Stat. Data Anal., 2014.

[19] A. Hepsen and M. Vatansever, “Forecasting future trends in Dubai housing market by using Box-Jenkins autoregressive integrated moving average,” Int. J. Hous. Mark. Anal., vol. 4, no. 3, pp. 210–223, 2011.

[20] S. Stevenson, “A comparison of the forecasting ability of ARIMA models,” J. Prop. Invest. Financ., vol. ٢٥, □□. ٣, □□. ٢٢٣□٢٤٠, ٢٠٠٧.

[21] I. Teymouri, M. Ahmadi, and K. Abolhassanzadeh, “Designing a mathematical model of railway passenger demand using time series-case study: Khorasan mainline,” in Rail Transportaion Conference, ٢٠٠٦.

[22] S. L. Ho and M. Xie, “The use of ARIMA models for reliability forecasting and analysis,” Comput. Ind. Eng., vol. 35, no. 1, pp. 213–216, 1998.

[23] C. Y. Lam, W. H. Ip, and C. W. Lau, “A business process activity model and performance measurement using a time series ARIMA intervention analysis,” Expert Syst. Appl., vol. 36, no. 3, pp. 6986–6994, ٢٠٠٩.

[24] M. Z. Babai, M. M. Ali, J. E. Boylan, and A. A. Syntetos, “Forecasting and inventory performance in a two-stage supply chain with ARIMA (0, 1, 1) demand: Theory and empirical analysis,” Int. J. Prod. Econ., vol. 143, no. 2, pp. 463–471, 2013.

[25] K. C. Gilbert and V. Chatpattananan, “An ARIMA supply chain model with a generalized ordering policy,” J. Model. Manag., vol. 1, no. 1, pp. 33–51, 2006.

[26] M. Blanchard and G. Desrochers, “Generation of autocorrelated wind speeds for wind energy conversion system studies,” Sol. Energy, vol. 33, no. 6, pp. 571–579, 1984.

[27] B. G. Brown, R. W. Katz, and A. H. Murphy, “Time series models to simulate and forecast wind speed and wind power,” J. Clim. Appl. Meteorol., vol. 23, no. 8, pp. 1184–1195, 1984.

[28] G. Box and G. Jenkins, Time series analysis: forecasting and control. 1976.

[29] Central Bank of Iran (CBI), Annual data of Consumer Price Index in Iran from 1359 to 1391: http://www.cbi.ir/datedlist/10807.aspx.

پی‌نوشت‌ها:

[1] *Autoregressive Moving Average (ARMA)*
[2] *Autoregressive Integrated Moving Average (ARIMA)*
[3] *Consumer Price Index (CPI)*
[4] *Nielsen, S. & Nielsen, H. 2008*
[5] *Ediger et al.*
[6] *Ediger and Akar*
[7] *Tabesh et al.*
[8] *Wang et al.*
[9] *Abdel-AAl and Al-Garni*
[10] *Barak and Sadegh*
[11] *Yuan et al.*
[12] *Residual Modification Models*
[13] *Seasonal ARIMA (SARIMA)*
[14] *Yusuf et al.*
[15] *Narayanan et al.*
[16] *Jia et al.*

[17] *Taneja et al.*
[18] *Mann Kendall*
[19] *Ecological Footprint*
[20] *Aerosol Optical Depth (AOD)*
[21] *Karbasi Yazdi et al.*
[22] *Jadevicius and Huston*
[23] *Du et al.*
[24] *Babu and Reddy*
[25] *Vosseler*
[26] *Hepsen and Vatansever*
[27] *Stevenson*
[28] *Divide and Conquer*
[29] *Gray Model (GM(1,1))*
[30] *Back Propagation*
[31] *Bayesian*
[32] *Unit Root Test*
[33] *OECD*
[34] *Dubai Residential Property Price Index (DRPPI)*
[35] *Teymouri et al.*
[36] *Ho and Xie*
[37] *Lam et al.*
[38] *Babai et al.*
[39] *Gilbert and Chatpattananan*
[40] *Ordinary Least Square*
[41] *Spurious Results*
[42] *Differencing*
[43] *Order of Differencing*
[44] *Diagnostic Checking*
[45] *White Noise*
[46] *Akaike Information Criterion*
[47] *Augmented Dicky Fuller*
[48] *Bayesian Information Criterion (Also Schwarz Information Criterion)*
[49] *Portman-Teau Test*
[50] *Box and Pierce*