

پیش‌بینی خسارت رشته بدنه اتومبیل با به‌کارگیری روش‌های سری زمانی (مورد مطالعه: شرکت بیمه پارسیان)

دکتر امیر تیمور پاینده نجف‌آبادی^۱

تاریخ دریافت مقاله: ۱۳۸۹/۱۱/۰۲

زهرا سلطانی رنانی^۲

تاریخ پذیرش مقاله: ۱۳۸۹/۱۲/۰۹

چکیده

بیمه اتومبیل در بازار بیمه ایران بالاترین حجم صدور و خسارت را دارد. بنابراین نیاز به تصمیم‌گیری‌های مالی و پیش‌بینی رفتار این بازار برای بیمه‌گران امری ضروری است. سری زمانی از روش‌های توانمند آماری است که با استفاده از روش‌های مبتنی بر فرآیندهای تصادفی به مدل‌سازی و پیش‌بینی رفتار متغیرهایی می‌پردازد که به‌گونه‌ای به زمان وابسته‌اند. مقاله حاضر میزان خسارت وارد شده در رشته بیمه بدنه اتومبیل (شرکت بیمه پارسیان از فروردین ماه سال ۱۳۸۵ الی تیرماه ۱۳۸۹) را به صورت یک متغیر زمانی در نظر گرفته است، سپس با استفاده از دو مدل سری زمانی باکس-جنکینز و ویتترز به مدل‌بندی رفتار این متغیر زمانی می‌پردازد. با توجه به اینکه مشاهدات، روند فصلی دارند از مدل‌های آریمای فصلی ضریبی باکس و جنکینز و ویتترز فصلی ضریبی استفاده شد. ابزار شاخص خطای MAPE نشان می‌دهد مدل ویتترز فصلی ضریبی برازش دقیق‌تر به داده‌ها ارائه می‌کند. در نهایت با استفاده از مدل منتخب، میزان خسارت برای یک دوره شش‌ماهه پیش‌بینی می‌شود.

واژگان کلیدی: بیمه بدنه اتومبیل، پیش‌بینی سری زمانی، مدل باکس-جنکینز، مدل

ویتترز، شاخص خطای MAPE

پژوهشگاه علوم انسانی و مطالعات فرهنگی

۱. عضو هیئت علمی دانشگاه شهید بهشتی، دانشکده علوم ریاضی (Email: Amirtpayandeh@Sbu.ac.ir)

۲. کارشناس ارشد آمار بیمه دانشگاه شهید بهشتی، دانشکده علوم ریاضی (نویسنده مسئول)

(Email: Z.soltani.renani@Gmail.com)

۱. مقدمه

اتخاذ تصمیم مناسب در بازارهای مالی از مهم‌ترین شاخص‌ها در حوزه‌های رقابتی شرکت‌ها به‌منظور تأمین بهینه منابع مالی برای بقا در محیط متلاطم تجاری است. پیش‌بینی رفتار در این‌گونه بازارها نقشی حیاتی در اتخاذ تصمیم‌های مناسب دارد. بیمه خودرو را می‌توان از جمله رایج‌ترین بیمه‌هایی دانست که در عصر حاضر افراد زیادی با آن سروکار دارند. تعداد زیاد وسایل نقلیه سبک و سنگین با کاربردهای متفاوت باعث وقوع حوادث و اتفاقاتی گوناگون در ابعادی وسیع و گسترده می‌شود که نتیجه آن زیان‌های جانی و مالی فراوان است و جبران اغلب این خسارت‌ها خارج از توانایی مالی دارندگان و یا رانندگان وسایل نقلیه موتوری است. به‌هرحال حادثه اتومبیل، گاهی موجب بروز زیان مالی می‌شود و زمانی فوت و صدمات جانی به‌بارمی‌آورد و این خسارت‌ها نیز گاهی به مالک یا راننده وسیله نقلیه مشخصی که تحت پوشش بیمه قرار گرفته وارد شده و زمانی فرد یا افرادی دیگر دچار زیان می‌شوند که اصطلاحاً به این افراد شخص ثالث می‌گویند.

۱-۱. انواع پوشش‌های بیمه‌ای برای وسایل نقلیه موتوری زمینی

- در بیمه اتومبیل گاهی تعهد بیمه‌گر، جبران خسارت وارده به خود وسیله نقلیه بیمه‌شده است؛ به‌عبارت‌دیگر بیمه‌گر با بیمه‌کردن وسیله نقلیه تعهد می‌کند چنانچه به‌علت هریک از حوادث مشمول بیمه، خسارتی به وسیله نقلیه وارد شده و یا کلاً از بین برود، خسارت را جبران کند. این پوشش بیمه‌ای به بیمه بدنه اتومبیل معروف شده است.

- زمانی که بیمه‌گر تعهد کند که اگر اتومبیل بیمه‌شده موجب بروز خسارت جانی یا مالی به اشخاص دیگری غیر از بیمه‌گذار گردد و بیمه‌گذار یا راننده، مسئول جبران خسارت باشد، در این‌صورت خسارت وارده را پرداخت نماید، این نوع بیمه را بیمه مسئولیت می‌نامند که اصطلاحاً به بیمه شخص ثالث اتومبیل معروف شده است.

وضعیت صنعت بیمه در ایران به گونه‌ای است که بیش از نیمی از حجم بازار صدور و خسارت به رشته اتومبیل اختصاص دارد و این رشته، گردش نقدینگی بالایی دارد. براین اساس بررسی نوسانات خسارت این رشته، در دوره‌های زمانی مختلف در راستای مدیریت بهینه منابع مالی شرکت‌های بیمه‌گر مفید به نظر می‌رسد و پیش‌بینی خسارت در دوره‌های آتی می‌تواند در کنترل و مدیریت بدهی‌های بیمه‌گران استفاده شود.

علاوه بر این بیمه مرکزی به عنوان یک نهاد ناظر در صنعت بیمه می‌تواند با به‌کارگیری روش‌های پیش‌بینی در صدور و خسارت رشته ثالث در کل صنعت بیمه کشور بر میزان کفایت منابع درآمدی در صندوق تأمین خسارت‌های بدنی بر جبران خسارت‌های مشمول موضوع ماده ۱۰ قانون بیمه اجباری شخص ثالث نظارت و مدیریت کند. پرداخت خسارت در رشته اتومبیل به چهار دسته بدنه، ثالث مالی، ثالث جانی و ثالث سرنشین تقسیم می‌شود. جدول ۱ اطلاعات پرداخت خسارت را در رشته اتومبیل به تفکیک سال پرداخت و بخش دولتی و خصوصی نشان می‌دهد.

جدول ۱. درصد مبلغ خسارت پرداختی کل بازار بیمه در رشته اتومبیل به تفکیک سال پرداخت

درصد مبلغ خسارت پرداختی				بازار بیمه	رشته
۱۳۸۶	۱۳۸۵	۱۳۸۴	۱۳۸۳		
%۷۸	%۷۴	%۶۹	%۷۱	بخش دولتی	اتومبیل
%۷۲	%۷۴	%۸۰	%۵۰	بخش غیردولتی	
%۷۷	%۷۴	%۷۰	%۷۰	کل بازار بیمه کشور	

بر اساس جدول ۱ رشته اتومبیل به‌طور تقریبی بیش از %۷۰ پرتفوی خسارت را در اختیار دارد و حجم پرداخت‌های مالی در این رشته بسیار هنگفت است. با این وجود بررسی نوسانات خسارت در این رشته به عنوان یکی از راهکارهای مدیریت ریسک در حوزه منابع مالی شرکت‌های بیمه‌گر ضروری به نظر می‌رسد.

در این مقاله بررسی سری زمانی تعدادی پرونده خسارت در رشته بدنه مطالعه شده و از دو روش معروف سری زمانی باکس-جنکینز و هموارسازی نمایی استفاده شده است. سپس با استفاده از بهترین مدل برازش‌شده، این‌گونه خسارات پیش‌بینی شده‌اند. مقاله حاضر به این‌صورت نگارش شده است: بخش دوم به مرور ادبیات موضوع پرداخته و کارهای ارائه‌شده در مورد مدل‌سازی و پیش‌بینی خسارات وارد شده به یک پرتفوی را بررسی می‌کند. مدل‌های سری زمانی مورد استفاده در این مقاله در بخش سوم معرفی و چگونگی استفاده از آنها در عمل مورد توجه قرار می‌گیرند. در بخش چهارم و پنجم مدل‌های سری زمانی معرفی شده استفاده شده‌اند و مناسب‌ترین مدل براساس معیار شاخص خطای MAPE انتخاب شده است. سرانجام براساس مدل ارائه‌شده، پیش‌بینی و نتیجه‌گیری صورت گرفته است.

۲. پیشینه تحقیق

باسونی و حبشی^۱ با استفاده از مدل سری زمانی باکس-جنکینز خسارات بیمه اتومبیل را در کشور کویت مدل‌سازی و براساس آن خسارات آینده در رشته را پیش‌بینی کرده‌اند. کامینس و گریپتراک^۲ در پژوهشی به بررسی پیش‌بینی خسارت رشته اتومبیل مربوط به کمپانی‌های بیمه در آمریکا توسط مدل‌های هموارسازی نمایی و مدل‌های اتورگرسیو پرداخته‌اند و نتایج این پژوهش نشان می‌دهد به‌منظور پیش‌بینی خسارت‌های مالی، مدل هموارسازی نمایی، برازش دقیق‌تری ارائه می‌کند ولی در مورد خسارت‌های جانی، دو روش چندان تفاوتی نداشته‌اند. بریج^۳ با استفاده از مدل‌های هموارسازی نمایی و رگرسیون خطی سری زمانی، خسارت‌های رشته بیمه

-
1. Bassiouni & Habashi, 2002
 2. Cummins & Griepentrog, 1985
 3. Berridge, 1998

اتومبیل را در یک کمپانی بیمه در نیوزیلند مدل سازی کرد. داری و لوی^۱ نشان دادند که می توان تعداد خسارت های واقع شده اما گزارش نشده^۲ را با استفاده از روش سری زمانی باکس- جنکینز مدل بندی و پیش بینی کرد. رنتالا و هیتکر^۳ در پژوهشی به بررسی کاربرد روش باکس- جنکینز در برآورد برخی شاخص های بیمه ای از جمله تعداد و مبلغ خسارت پرداخته اند و همچنین هاروی و فرناندس^۴ در پژوهشی، کاربرد روش های سری زمانی به ویژه روش باکس- جنکینز را در برآورد مجموع مبلغ و تعداد خسارت بیمه گر بررسی کردند.

۳. شرح روش ها

اساس روش های سری زمانی، برازش یک مدل آماری بر مبنای داده هایی است که در راستای زمان جمع آوری شده اند. یعنی با یافتن قانون حاکم بر داده های زمانی می توان مدلی را طراحی کرد که قابل تعمیم به آینده باشد. این گونه مدل های آماری را مدل های سری زمانی می نامیم. در ادامه به معرفی مختصر برخی از مدل های معروف سری زمانی می پردازیم.

۳-۱. مدل سری زمانی باکس - جنکینز یا ARIMA:^۵

مدل سری زمانی باکس-جنکینز یکی از مدل های سری زمانی پرتعداد است. الگوی عمومی باکس و جنکینز از مرتبه p, q, d, D, P, Q به این صورت نوشته می شود:

$$\phi_p(B)\phi_p(B^L)\nabla_L^D\nabla^d y_t^* = \delta + \theta_q(B)\theta_Q(B^L)a_t \quad (1)$$

که به مدل «آریمای فصلی ضربی باکس و جنکینز» مشهور است که در آن:

1. Doray & Louis, 1996
2. Incurred But Not Reported (IBNR)
3. Rantala & Hietikko, 1988
4. Harvey & Fernandes, 1989
5. Auto Regressive Integrated Moving Average (ARIMA)

$$\phi_p(B) = (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)_{-P}$$

مرتبه P؛ عملگر اتورگرسیو غیرفصلی از مرتبه P؛

$$\phi_p(B^L) = (1 - \phi_{1,L} B^L - \phi_{2,L} B^{2L} - \dots - \phi_{p,L} B^{pL})_{-P}$$

مرتبه P؛ عملگر اتورگرسیو فصلی از مرتبه P؛

$$\theta_q(B) = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q)_{-q}$$

مرتبه q؛ عملگر میانگین متحرک غیرفصلی از مرتبه q؛

$$\theta_Q(B^L) = (1 - \theta_{1,L} B^L - \theta_{2,L} B^{2L} - \dots - \theta_{Q,L} B^{QL})_{-Q}$$

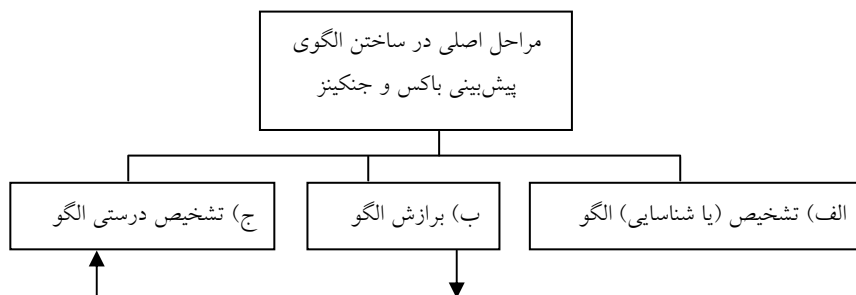
مرتبه Q؛ عملگر میانگین متحرک فصلی از مرتبه Q؛

$\delta = \mu \phi_p(B) \phi_p(B^L)$ - مقدار ثابت مدل نامیده می‌شود که در آن μ میانگین واقعی سری زمانی ایستایی است که مدل شده است؛

$\alpha_1, \alpha_{t-1}, \dots$ - جملات اغتشاش و تصادفی هستند که از توزیع نرمال پیروی می‌کنند؛
 $\delta, \phi_1, \phi_2, \dots, \phi_p, \phi_{1,L}, \dots, \phi_{p,L}, \theta_1, \theta_2, \dots, \theta_q, \theta_{1,L}, \theta_{2,L}, \dots, \theta_{q,L}$ - پارامترهای نامعلوم و مجهول‌اند که باید از داده‌های نمونه برآورد گردند؛

$B^k y_t = y_{t-k}$ - عملگر پس‌رو که به شکل $B^k y_t = y_{t-k}$ تعریف می‌شود؛
 $\nabla = 1 - B$ - عملگر غیرفصلی نامیده می‌شود و به شکل $\nabla = 1 - B$ تعریف می‌شود؛
 $\nabla_L = 1 - B^L$ - عملگر فصلی نامیده می‌شود و به شکل $\nabla_L = 1 - B^L$ تعریف می‌شود.
 به‌طور خلاصه در این مدل p و p به ترتیب مرتبه غیرفصلی و فصلی جز اتورگرسیو، q و Q به ترتیب مرتبه غیرفصلی و فصلی جزء میانگین متحرک‌اند. مراحل اصلی در ساختن الگوی پیش‌بینی باکس و جنکینز در نمودار ۱ خلاصه شده است.

نمودار ۲: مراحل اصلی در ساختن الگوی پیش‌بینی باکس-جنکینز



ایستایی سری زمانی، نرمال بودن و عدم برازش مدل سری زمانی به مانده‌ها فرضیات اساسی هستند که در مدل‌سازی سری زمانی همواره باید مورد توجه قرار گیرند. به زبان ساده، البته نه خیلی دقیق، ایستایی یک متغیر زمانی را می‌توان ثابت بودن آن متغیر در طول زمان تفسیر کرد. متغیرها یا سری زمانی به خودی خود ایستا نیستند. البته در هنگام مدل‌سازی باید آنها را ایستا نمود. با استفاده از دو روش «فرآیند تفاضلی کردن» و استفاده از «تبدیل‌های پایداری واریانس» می‌توان سری‌های نایستا را ایستا کرد.^۱

۲-۳. مدل سری زمانی SARIMA^۲:

مدل سری زمانی SARIMA حالت خاصی از رده کلی مدل‌های سری زمانی است که اگر چه لزوماً ایستا نیستند، اما همگن بوده و تفاضل مرتبه d آن یک فرآیند SARIMA فصلی ایستاست. ساخت مدل SARIMA یک فرآیند تکراری شامل شناسایی، برآورد پارامترها و بررسی درستی تشخیص است تا مناسب‌ترین مدل به دست آید. اولین قدم در این روش رسم نمودار سری زمانی متغیر زمانی است. با توجه به نمودار معمولاً ایده خوبی در مورد اینکه آیا سری زمانی دارای روند،

۱. برای اطلاع بیشتر رک: کرایر، ۱۳۸۴

2. Seasonal Autoregressive Integrated Moving Average (SARIMA)

ویژگی فصلی و واریانس غیرثابت است، به‌دست‌می‌آید. این درک معمولاً مبنایی را برای ایجاد یک تبدیل مناسب داده‌ها فراهم می‌سازد. جهت تأیید درجه تفاضلی کردن، نمودار خودهمبستگی و خودهمبستگی جزئی سری اولیه رسم و پس از شناسایی مدل سری زمانی، پارامترهای مدل برآورد گردید.

بعد از شناسایی و برآورد پارامترهای مدل، لازم است تا کفایت الگو را با بررسی اینکه آیا فرض‌های الگو برقرار هستند، ارزیابی کرد. فرض اساسی این است که سری باقی‌مانده‌ها، توزیع نرمال با واریانس ثابت دارند. جهت بررسی نرمال بودن باقیمانده‌ها می‌توان از آزمون آماری کلموگروف-اسمیرنوف و یا نمودار چندک‌های مانده‌ها در مقابل چندک‌های توزیع نرمال استفاده کرد.

۳-۳. مدل سری زمانی وینترز

در روش هموارسازی نمایی، تغییرات سری زمانی به‌صورت گام‌به‌گام و با استفاده از مشاهدات قبلی برآورد می‌شود. برآورد هر مشاهده به مشاهدات قبلی بستگی دارد و با ورود هر مشاهده جدید به‌نگام می‌شود. ضریب تأثیر مشاهده جدید و مشاهدات قبل از آن، پارامترهای مدل را تشکیل می‌دهند. برای استفاده از روش هموارسازی نمایی، نیازی به حصول ایستایی مشاهدات نداریم، کافی است نوع فصل (جمععی، ضربی) و تعداد مشاهدات تشکیل‌دهنده یک فصل را تعیین کنیم که با توجه به نمودار مشاهدات برحسب زمان قابل تشخیص است. روش‌های هموارسازی مانند مدل‌های دیگر سری زمانی به دو دسته عمده برای سری‌های فصلی و غیرفصلی تقسیم‌بندی می‌شوند. روش‌های هموارسازی مرتبه یک و حالت ویژه سری‌های زمانی غیرفصلی، روش وینترز و هموارسازی مرتبه دوم ویژه سری‌های زمانی فصلی به‌کار می‌رود. در این پژوهش به‌دلیل وجود اثر فصلی در مشاهدات از مدل وینترز استفاده شده است. برای استفاده از روش وینترز احتیاجی به حصول ایستایی مشاهدات نداریم، کافی است نوع فصل (جمععی، ضربی) و اینکه چه تعداد مشاهده تشکیل‌دهنده یک فصل

است را تعیین کنیم. تفاوت مدل‌های فصلی جمعی و فصلی ضربی در این است که در نوع ضربی، الگوی فصلی بودن با افزایش تعداد مشاهدات تغییر خواهد کرد، ولی در مدل جمعی الگوی فصلی به تعداد مشاهدات بستگی ندارد، که باتوجه به نمودار سری زمانی، نوع فصلی ضربی بودن مناسب تشخیص داده می‌شود. مسئله مهم دیگری که باقی می‌ماند انتخاب مقادیر مناسب برای سطح، روند و فصلی بودن سری زمانی است که مقادیری بین ۰ تا ۱ را می‌توانند، اختیار کنند. با یک جستجوی شبکه‌ای، مقادیر مختلفی را برای این مقادیر در نظر می‌گیریم و باتوجه به مقادیر شاخص‌های اندازگیری دقت مدل می‌توان بهترین ضرایب را انتخاب کرد.

در مدل فصلی با روند خطی، اگر \hat{S}_n و $\hat{\beta}_1(n)$ و $\hat{\beta}_0(n)$ مقادیر هموار شده اثرهای فصلی و شیب و عرض از مبدأ در زمان n باشند، پیش‌بینی در مبدأ در زمان n باشند، پیش‌بینی \hat{Y}_{n+1} در مبدأ n به کمک این رابطه محاسبه می‌شود:

$$[\overline{\text{TR}}_n + \hat{\beta}_1(n)] (1) = \hat{X}_n \times \hat{S}_{n+1} \quad (2)$$

که در آن، $\overline{\text{TR}}_n = [\hat{\beta}_0(n) + \hat{\beta}_1(n) \times n]$ در این روش، \hat{S}_t و $\hat{\beta}_1(n)$ و $\overline{\text{TR}}_n$ توسط روابط زیر به‌نگام می‌شوند:

$$\overline{\text{TR}}_t = \alpha \left(\frac{X_t}{\hat{S}_{t-r}} \right) + (1-\alpha) (\overline{\text{TR}}_{t-1} + \hat{\beta}_1(t-1)) \quad (3)$$

$$\hat{\beta}_1(t) = \beta (\overline{\text{TR}}_t - \overline{\text{TR}}_{t-1}) + (1-\beta) \hat{\beta}_1(t-1) \quad (4)$$

$$\hat{S}_t = \gamma \left(\frac{X_t}{\overline{\text{TR}}_t} \right) + (1-\gamma) \hat{S}_{t-s} \quad (5)$$

در این روابط $0 < \alpha < 1$ و $0 < \beta < 1$ و $0 < \gamma < 1$ ثابت‌های هموارسازی نامیده می‌شوند.

در رابطه (۳)، $\frac{X_t}{\hat{S}_{t-s}}$ برآوردی از TR_t است که به محض مشاهده x_t ، محاسبه می‌شود و $\hat{\beta}_1(t-1) + \overline{\text{TR}}_{t-1}$ برآورد دیگری از TR_t ، حاصل گسترش مدل

برازنده‌شده برای روند در زمان $t-1$ به زمان t است. پس TR_t در رابطه (۳) میانگین موزون این دو برآوردگر است. تعابیر مشابهی نیز برای \hat{S}_t و $\hat{\beta}_t(t)$ می‌توان ارائه کرد.

ویترز شیوه‌ای هم برای تعیین برآوردهای اولیه \overline{TR}_0 و $\hat{\beta}_1(0)$ و \hat{S}_{j-s} ، $(j=1,2,\dots,s)$ ، به این شرح ارائه کرده است. اگر داده‌ها برای k دوره تناوب کامل در دسترس باشند $(n=k \times s)$ و \overline{X}_i $(i=1,2,\dots,k)$ میانگین مقادیر مشاهده‌شده در دوره i ام باشد، \overline{TR}_0 و $\hat{\beta}_1(0)$ را به ترتیب، عرض از مبدأ و شیب خط واصل بین نقاط $(\frac{s}{2}, \overline{X}_1)$ و $(n - \frac{s}{2}, \overline{X}_k)$ اختیار می‌کنیم، یعنی:

$$\hat{\beta}_1(0) = \frac{\overline{X}_k - \overline{X}_1}{n - s} = \frac{\overline{X}_k - \overline{X}_1}{(k-1)s} \quad \overline{TR}_0 = \overline{X}_1 - \frac{s}{2} \hat{\beta}_1(0) \quad (6)$$

برای محاسبه برآوردهای اولیه اثرهای فصلی، این کمیت‌ها را محاسبه می‌کنیم:

$$S_t^\circ = \frac{X_t}{\overline{X}_i + \left(j - \frac{s+1}{2}\right) \hat{\beta}_1(0)} \quad \text{و} \quad t=1,2,\dots,n \quad (7)$$

در رابطه فوق، j نشان‌دهنده فصل است و i نشان‌دهنده دوره متناظر با زمان t ، که باتوجه به

$$t = j + (i-1)s, \quad i=1,2,\dots,k, \quad j=1,2,\dots,s$$

تعیین می‌شوند. مخرج کسر سمت راست تساوی (۷) برآوردی از \overline{TR}_t است که از خط ماربر نقطه میانی دوره i ام با شیب $\hat{\beta}_1(0)$ حاصل می‌شود. پس باتوجه به ساختار مدل کلاسیک ضربی، $S_j^\circ + (t-1)s$ ها به ازای $i=1,2,\dots,k$ برآوردهایی از S_j هستند. ویترز میانگین S_t° ها، یعنی $\overline{S}_j^\circ = \frac{1}{k} \sum_{j=1}^k S_{j+(i-1)s}^\circ$ را در نظر گرفت و سپس تصحیح شده آنها را به‌گونه‌ای که مجموعشان برابر با S شود، یعنی:

$$\hat{S}_{i-s} = \frac{\overline{S}_1^\circ}{\sum_{j=1}^s \overline{S}_j^\circ} \times S, \quad j=1,2,\dots,s \quad (8)$$

به عنوان برآوردهای اولیه اثرهای فصلی پیشنهاد کرد.^۱

شاخص انتخاب مناسب ترین مدل، شاخص خطایی MAPE است که به صورت

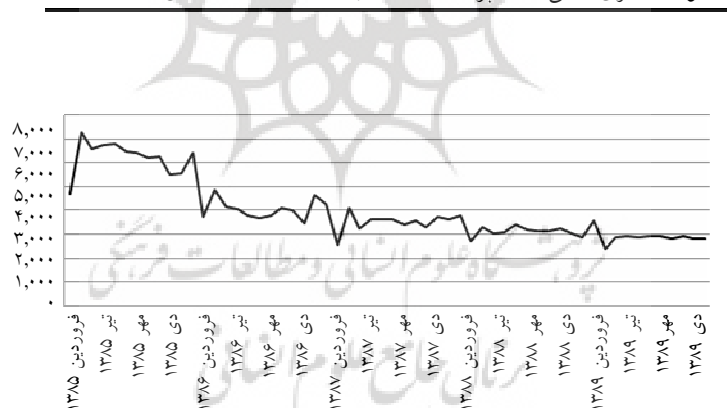
$$MAPE = \frac{1}{n} \sum_{t=1}^n \frac{|x_t - \hat{x}_t|}{x_t} \times 100\%$$

تعریف شده و کوچک بودن آن مناسب تر بودن مدل را نتیجه می دهد.

۴. مدل سازی خسارات وارد شده به بیمه اتومبیل شرکت بیمه پارسیان

بیمه پارسیان فعالیتش را همزمان با شروع فعالیت شرکت های بیمه خصوصی از اردیبهشت ماه ۱۳۸۲ آغاز کرده است. مشاهدات سال های ۱۳۸۲ و ۱۳۸۳ و ۱۳۸۴ به دلیل تغییرات در شرایط بازار و شرایط محیطی برای تحلیل مناسب نیست؛ لذا تعداد پرونده خسارت در رشته بدنه به صورت ماهیانه از ابتدای سال ۱۳۸۵ تا تیر ۱۳۸۹ به منظور برازش مدل به کاررفته است. از مشاهدات مرداد تا دی ماه سال ۱۳۸۹ به منظور ارزیابی سازگاری مدل با واقعیت استفاده شده است. نمودار ۲ رفتار زمانی تعداد پرونده خسارت بدنه از سال ۱۳۸۵ تا دی ۱۳۸۹ این شرکت را نشان می دهد.

نمودار ۲. سری زمانی تعداد پرونده خسارت بدنه از سال ۱۳۸۵ تا دی ماه ۱۳۸۹



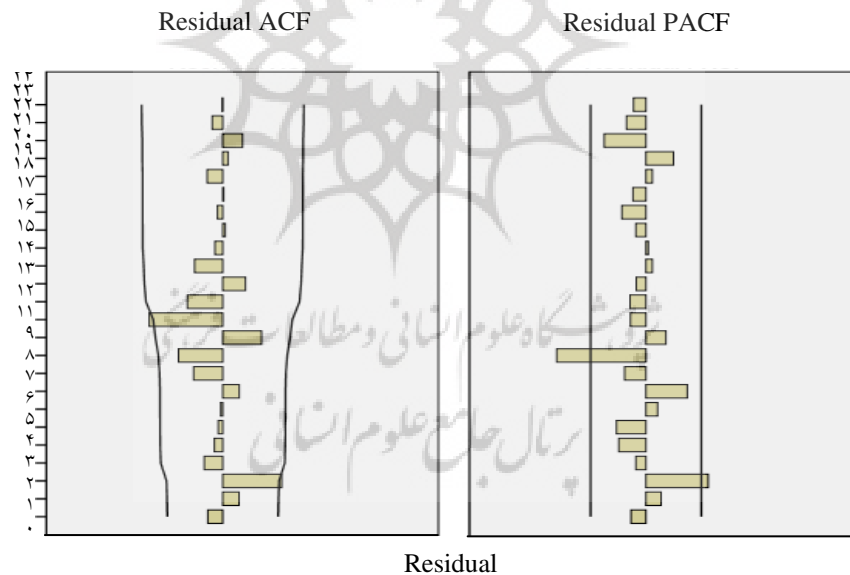
۱. برای اطلاع بیشتر رک: Hyndman & Koehler, 2008

همان‌طور که نمودار ۲ نشان می‌دهد وجود یک الگوی زمانی و فصلی در این سری زمانی مشهود است. بدین‌صورت که از سال ۱۳۸۵ تا ۱۳۸۹ در اسفندماه هر سال با افزایش تعداد پرداخت خسارت بدنه، در فروردین با کاهش آن و سپس در اردیبهشت با افزایش دوباره روبرو بوده‌ایم. به‌طورکلی تعداد پرونده خسارت بدنه در حال کاهش است. اکنون با استفاده از نرم‌افزارهای ۱۷ Spss و ۱۵ Minitab به برازش مدل‌های SARIMA و ویترز می‌پردازیم. شاخص انتخاب مناسب‌ترین مدل شاخص خطایی MAPE است.

۴-۱. مدل SARIMA

نخست به‌منظور ایستاکردن سری زمانی از تبدیل لگاریتم طبیعی و تفاضل‌گیری استفاده می‌شود. سپس با رسم نمودارهای خودهمبستگی و خودهمبستگی جزئی سری زمانی مرتبه اتورگرسیو و میانگین متحرک مدل تشخیص داده می‌شود. نمودار ۳ خودهمبستگی و خودهمبستگی جزئی سری زمانی را نشان می‌دهد. با استفاده از این نمودار می‌توان مدل $SARIMA(1,1,0)(0,1,0)$ را یک مدل مناسب برای داده‌ها انتخاب کرد.

نمودار ۳. نمودارهای همبستگی و همبستگی جزئی



اکنون به بررسی مناسب بودن این مدل می پردازیم. جدول ۲ نتیجه بررسی مدل برازش شده را براساس شاخص خطایی MAPE نشان می دهد. براساس مقدار معناداری این آماره می توان نتیجه گرفت که در سطح معناداری ۰/۰۵ مدل برازش شده مورد پذیرش قرار می گیرد.

جدول ۲. نتایج مدل SARIMA

نتایج برآورد مدل				
مقدار معناداری	درجه آزادی	مقدار آماره	MAPE	مدل SARIMA
۰/۰۲۶	۱۷	۳۰/۰۶۳	۹/۵۱۴	تعداد پرونده خسارت بدنه

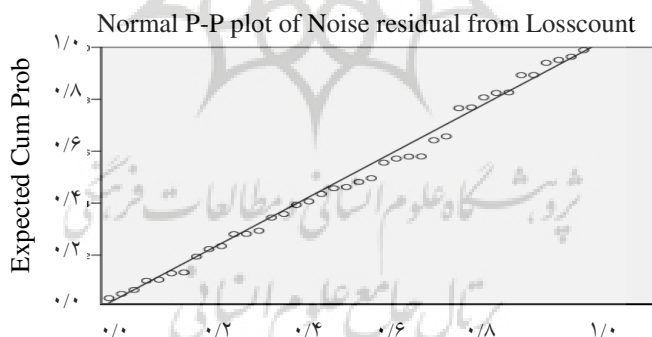
نرمال بودن مانده ها یکی دیگر از فرضیات اساسی مدل است که جدول ۳ براساس آزمون کلموگروف اسمیرنوف و نمودار چندک- چندک ۴ به بررسی آن پرداخته اند.

جدول ۳. آزمون کلموگروف اسمیرنوف مانده ها

میزان معناداری آزمون	مقدار آماره آزمون	مانده ها
۰/۹۳	۰/۵۴	ایستا و نرمال

باتوجه به میزان معناداری آزمون (۰/۹۳) می توان نرمال بودن مانده ها را در سطح معناداری ۰/۰۵ پذیرفت.

نمودار ۴. چندک های مانده ها در مقابل چندک های توزیع نرمال



باتوجه به نزدیک بودن نقاط به خط مبنا در نمودار چندک- چندک ۴ می توان درست بودن فرض نرمال بودن مانده ها را تأیید کرد.

۲-۴. مدل سری زمانی وینترز

برای به‌دست‌آوردن پارامترهای مدل وینترز از جستجوی شبکه استفاده می‌شود، که در نهایت بهترین مدل وینترز فصلی ضربی با کمترین درصد میانگین مطلق خطا انتخاب شد. جدول ۴ مدل وینترز برازش‌شده به داده‌ها را نشان می‌دهد. با توجه به مقادیر معناداری پارامترها می‌توان معناداری تمامی پارامترها را در سطح معناداری ۰/۰۵ نتیجه گرفت.

جدول ۴. نتایج برآورد پارامترها در مدل وینترز

متغیر	تبدیل	پارامترها	برآورد پارامترها	مقدار آماره آزمون	مقداری معناداری آزمون	MAPE
تعداد پرونده خسارت	لگاریتم طبیعی	α	۰/۴۲۵	۳/۹۷	۰	۶/۰۱
		β	۰/۰۰۱	۱/۲۷	۰/۰۶۲	
		γ	۰/۲۴	۰/۲۷	۰/۰۴۱	

با توجه به جدول ۴ می‌توان مدل وینترز فصلی ضربی با پارامترهای ۰/۴۵۲، ۰/۰۰۱ و ۰/۲۴ را به داده‌ها برازش کرد. جدول ۵ مناسب‌بودن مدل برازش‌شده را براساس شاخص خطایی MAPE بررسی می‌کند.

جدول ۵. نتایج مدل وینترز

مدل وینترز فصلی ضربی	نتایج برآورد مدل		
	MAPE	مقدار آماره	درجه آزادی
تعداد پرونده خسارت بدنه	۶/۰۱۰۴	۲۶/۶۷۳	۱۵

با توجه به مقدار کوچک معناداری آزمون (۰/۰۳۱۵) می‌توان مناسب‌بودن مدل را در سطح معناداری ۰/۰۵ نتیجه گرفت.

۵. نتیجه گیری و پیشنهادات

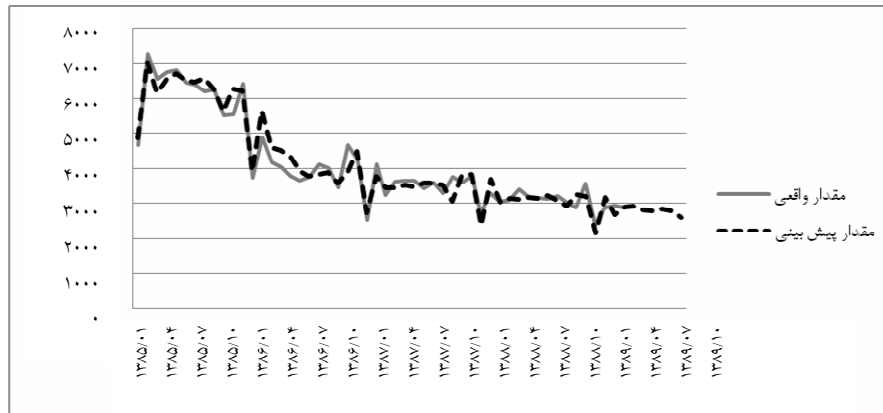
این تحقیق به منظور مقایسه عملکرد مدل سری زمانی باکس - جنکینز و ویتترز برای پیش بینی تعداد پرونده خسارت بدنه انجام شده است، برای این منظور مشاهدات سال های ۱۳۸۵ الی تیرماه ۱۳۸۹ در نظر گرفته شدند و از مشاهدات مرداد الی دی ماه ۱۳۸۹ برای ارزیابی سازگاری مدل استفاده می شود. با توجه به وجود روند فصلی در مشاهدات دو مدل آریمای فصلی و ویتترز فصلی ضریب برازش داده شد. بهترین مدل ها $(0,1,0)$ SARIMA $(1,1,0)$ و $(0,24,0)$ ، $(0,001,0)$ و $(0,45)$ Winter's به دست آمدند. اکنون با استفاده از شاخص خطایی MAPE به مقایسه دو مدل برازش شده می پردازیم. جدول ۶ مقدار MAPE را برای دو مدل برازش شده نشان می دهد.

جدول ۶. مقدار MAPE برای دو مدل برازش شده

	مدل ویتترز	مدل SARIMA
شاخص خطایی MAPE	۶٪	۱۰٪

باتوجه به مقدار شاخص خطایی MAPE مدل ویتترز، برازش بهتری به داده ها ارائه می کند. اکنون با استفاده از مدل ویتترز برازش شده به داده ها می توان به پیش بینی رفتار متغیر زمانی تعداد خسارات وارد شده به پرتفوی بیمه اتومبیل بیمه پارسیان پرداخت. نمودار ۵ و جدول ۷ شاخص ها و مقادیر پیش بینی این سری زمانی را برای یک دوره ۶ ماه نشان می دهد.

نمودار ۵. نمودار تعداد واقعی مشاهدات در مقابل تعداد پیش‌بینی شده توسط مدل وینترز



تاریخ	تعداد پرونده خسارت واقعی	تعداد پیش‌بینی شده	خطای پیش‌بینی
۱۳۸۹/۰۵	۲/۹۴۵	۲/۹۳۲	٪ ۰
۱۳۸۹/۰۶	۲/۹۱۹	۲/۸۲۴	٪ ۳
۱۳۸۹/۰۷	۲/۸۵۶	۲/۸۰۱	٪ ۲
۱۳۸۹/۰۸	۲/۹۱۷	۲/۸۵۹	٪ ۲
۱۳۸۹/۰۹	۲/۸۴۸	۲/۸۰۱	٪ ۲
۱۳۸۹/۱۰	۲/۸۴۰	۲/۶۰۳	٪ ۸
جمع شش ماه	۱۷/۳۲۵	۱۶/۸۲۰	٪ ۳

باتوجه به نتایج مدل‌بندی میزان خسارت توسط سری زمانی وینترز می‌توان گفت خطای پیش‌بینی در این روش تقریباً قابل اغماض است و بیمه‌گران می‌توانند در پیش‌بینی خسارت رشته اتومبیل و همچنین پیش‌بینی میزان صدور این رشته و رشته‌های مختلف به‌منظور تنظیم پرتفوی از این روش سری زمانی بهره بگیرند. در این تحقیق از روش‌های سری زمانی یک متغیره استفاده شده است و عوامل مؤثر دیگر نظیر نوع ماشین، جنسیت بیمه‌گذاران و ... که بر میزان خسارت به‌صورت

مستقیم (یا غیرمستقیم) در مدل وارد نشده‌اند، لذا می‌توان با شناسایی این عوامل و بررسی تأثیر آنها از روش‌های سری زمانی چند متغیره استفاده کرد.

منابع

۱. آمار و اطلاعات بیمه‌ای سال‌های ۱۳۸۴، ۱۳۸۵ و ۱۳۸۶، بیمه مرکزی ج.ا.ا. <http://www.centinsur.ir/frmHome_fa-IR.aspx>2010/3/25
۲. خزائی، مجتبی ۱۳۸۷، آشنایی با تحلیل سری‌های زمانی به کمک نرم افزار، پژوهشکده آمار، تهران، چ ۱.
۳. کرایر، جاناناتان ۱۳۸۴، تجزیه و تحلیل سری‌های زمانی، ترجمه نیرومند، حسین، دانشگاه فردوسی.
4. Bassiouni, MY & Habashi, MH, 2002, 'Forecasting compulsory motor insurance claims in Kuwait', *Insurance: Mathematics and Economics*, vol. 10, pp. 85-92.
5. Berridge, S 1998, *Forecasting claims in motor vehicle insurance*, Institute of Statistics and Operations Research, Victoria University of Wellington.
6. Cummins, D & Griepentrog, G 1985, 'Forecasting automobile insurance paid claim costs using econometric and ARIMA; models', *International Journal of Forecasting*, vol.1, pp.203-15.
7. Doray, D & Louis, H 1996, 'Constrained forecasting of the number of IBNR claims', *Journal of Actuarial Practice*, vol 4, pp.287-305.
8. Harvey, AC & Fernandes, C 1989, 'Time series models for insurance claims', *Journal of the Institute of Actuaries [JIA]*, vol. 116, pp. 513-28.
9. Hyndman, R & Koehler, A 2008, *Forecasting with exponential smoothing: the state space approach*, Springer Series in Statistics, Springer.
10. Rantala, J & Hietikko, H. 1988, 'An application of time series method to financial guarantee insurance', *European Journal of Operational Research*, vol. 87, no.3, pp.398-408.



پښتونستان د علوم او انساني مطالعات فریښتی
پرتال جامع علوم انسانی