پژوهشنامهٔ بیمه/ سال سی‌و سوم / شمارهٔ ۱/ بهار ۱۳۹۷/ شمارهٔ مسلسل ۱۲۹/ صفحات ۸۵ تا ۱۰۷
کد شناسایی دیجیتال (DOI):10.22056/jir.2018.82679.2056

مدل‌بندی توأم استوار داده‌های ذخیره‌سازی خسارات معوق رشته‌های بیمهٔ بدنه و بیمهٔ شخص ثالث اتومبیل یک شرکت بیمهٔ ایرانی: یک روش بیزی

تاریخ دریافت: ۱۳۹۶/۰۲/۰۱
تاریخ پذیرش: ۱۳۹۶/۱۰/۱۳

منیر گودرزی[۱]
محمد ذکایی[۲]

چکیده

ذخیره‌سازی خسارات معوق یکی از اساسی‌ترین مسائل در بیمهٔ عمومی است. در این مقاله، یک روش بیزی تعمیم‌یافته برای مدل‌بندی توأم داده‌های ذخیره‌سازی خسارات معوق رشته‌های بیمهٔ بدنه و بیمهٔ شخص ثالث اتومبیل یک شرکت بیمهٔ ایرانی با استفاده از توزیعهای tی استیودنت و پی‌یرسون نوع هفتم دومتغیره به کار گرفته می‌شود. هنگامی که داده‌ها از فرض نرمال‌بودن پیروی نمی‌کنند، توزیعهای دُم‌سنگینی چون tی استیودنت و پی‌یرسون نوع هفتم به استنباطهای استوارتری منجر می‌شوند. این توزیعها به ردهٔ توزیعهای آمیخته- مقیاس نرمال تعلق دارند. ساختار سلسله‌مراتبی این رده سبب می‌شود که در چارچوب بیزی، برآورد پارامترها به‌سادگی با استفاده از روشهای مونت کارلوی زنجیر مارکوفی انجام شود. برای میانگین توزیعهای نمونه‌گیری، سه مدل آنالیز واریانس، آنالیز کوواریانس، و قدم‌زدن تصادفی در نظر گرفته می‌شود. به‌علاوه، برای شناسایی نمونه‌های مؤثر، یک مطالعهٔ آنالیز حساسیت بر اساس واگرایی کولبک- لیب‌لر در مدلها انجام شده است. نتایج نشان می‌دهد که مدل قدم‌زدن تصادفی با توزیع tی استیودنت دومتغیره برای پرداختهای خسارت، عملکرد بهتری دارد.

واژگان کلیدی: *ذخیره‌سازی خسارات معوق، توزیعهای آمیخته- مقیاس، استنباط بیزی، حذف موردی، واگرایی کولبک- لیب‌لر.*

---

۱. دانشجوی دکتری آمار، دانشکدهٔ علوم ریاضی، دانشگاه شهید بهشتی (نویسندهٔ مسئول)،
m_goudarzi@sbu.ac.ir
۲. دانشیارگروه بیم‌سنجی، دانشکدهٔ علوم ریاضی، دانشگاه شهید بهشتی، zokaei@sbu.ac.ir

۱. مقدمه

یک بیمه‌نامه قراردادی است بین دوطرف، بیمه‌گر (شرکت بیمه) و بیمه‌گذار. در این قرارداد بیمه‌گر در قبال مبلغی ثابت (حق‌بیمه) متعهد می‌شود که در صورت بروز خسارت برای موضوع قرارداد، خسارت وارده را جبران و یا وجه معینی را بپردازد. در صورتی که یک رویداد (خسارت) اتفاق افتد، بیمه‌گذار یک ادعا به بیمه‌گر ارائه می‌کند. وجهی را که بیمه‌گر باید بپردازد، مقدار خسارت می‌نامند. در طول یک سال مالی، شرکت بیمه نمی‌تواند برخی از خسارات را بلافاصله تسویه کند، بنابراین باید برای پرداخت چنین خساراتی، ذخیره‌ای تحت عنوان ذخیرهٔ خسارت معوق[1] در نظر بگیرد، این امر به دلایل زیر است (Wüthrich and Merz, 2008):

۱. معمولاً در گزارش خسارتها تأخیر وجود دارد (فاصلهٔ بین وقوع خسارت و گزارش به شرکت بیمه طولانی است)؛ بنابراین گزارش یک خسارت ممکن است چند سال طول بکشد. برخی از دلایل تأخیر در گزارش عبارت‌اند از:

- خسارت در روزهای پایانی اتفاق افتاده و بیمه‌گذار فرصت نکرده است به بیمه‌گر اعلام کند؛
- بیمه‌گذار از تحت پوشش بیمه‌ای‌بودن خسارت اطمینان ندارد؛
- بیمه‌گذار از تحقق خطر منجر به خسارت بی‌خبر است؛
- بیمه‌گذار قادر نیست که به بیمه‌گر اطلاع بدهد.

۲. بعد از گزارش ممکن است چند سال طول بکشد تا یک خسارت به طور کامل تسویه شود.

۳. همچنین خسارتی که تسویه شده است (بسته شده است) ممکن است به هر دلیلی دوباره مطرح شود.

بدهی بابت خسارت معوق که توسط هر یک از مدیریتهای مختلف بیمه‌ای (باربری، آتش‌سوزی، اتومبیل، مهندسی، مسئولیت، و ...) در شرکتهای بیمه برآورد و

1. Outstanding Loss Reserves

برای انعکاس در حسابهای شرکت اعلام می‌شود، شامل موارد زیر است (بازنگری آیین‌نامهٔ ذخایر فنی مؤسسات بیمه ارائهٔ پیشنهادهای اصلاحی، ۱۳۹۵):

- خسارتهای واقع‌شده‌ای که گزارش نشده‌اند (IBNR [۱])؛
- خسارت گزارش‌شده‌ای که پرداخت نشده است (RBNS[۲])؛
- خسارتی که به طور کامل گزارش نشده است (IBNER[۳])؛
- هزینهٔ تسویهٔ خسارت.

محاسبهٔ ذخایر دارای اهمیت زیادی است زیرا کم‌برآوردی آن سبب می‌شود که شرکت بیمه نتواند تعهدات خود را انجام دهد و بیش‌برآوردی آن نیز سبب می‌شود که به صورت غیرلازم سرمایهٔ اضافی نگه داشته شود. لذا پیش‌بینی خسارات معوق و ذخیرهٔ مناسب برای پرداخت چنین ادعاهایی نقش مهمی در توانایی مالی و استمرار کسب‌وکار شرکتهای بیمه‌ای دارد؛ بنابراین پیش‌بینی مطالبات با روشهای مناسب و برآورد خطای پیش‌بینی، هدف ذخیره‌سازی است.

بردار تصادفی $\mathbf{X}_{ij} = (X_{ij}^{(1)}, X_{ij}^{(2)})' \in R^2$ را که در آن $X_{ij}^{(l)}$ نشان‌دهندهٔ میزان پرداختهای خسارات غیرتجمعی[۴] به‌وسیلهٔ یک شرکت بیمه برای سال مبدأ[۵] (سال یا دورهٔ صدور بیمه‌نامه) $i \in \{1,..I.,$ و سال تأخیر[۶] (تعداد سالها یا فصلهایی که تا پرداخت خسارت، تأخیر به وجود می‌آید) $j \in \{1,..J,$ در $l \in \{1,2\}$امین رشتهٔ بیمه است، در نظر بگیرید. داده‌های مربوط به پرداختهای خسارت در هر رشتهٔ بیمه همان‌گونه که در جدول ۱ دیده می‌شود، معمولاً به صورت یک مثلث تأخیر[۷] نشان داده می‌شوند.

---

1. Incurred But Not Reported
2. Reported But Not Settled
3. Incurred But Not Enough Reported
4.Incremental Paid Losses
5. Accident Year
6. Development Year
7. Run-Off Triangle

جدول ۱. مثلث تأخیر

سال مبدأ **i**	سال تأخیر **j**					
	۱	۲	...	j	...	J
۱						
۲						
⋮		مقادیر $X_{ij}^{(l)}$ (پرداختهای خسارت)				
I-j						
⋮				مقادیر $X_{ij}^{(l)}$ که پیش‌بینی می‌شوند.		
I						

مدلهای آماری متعددی برای پیش‌بینی میزان خسارات معوق در مثلث پایینی در جدول ۱ وجود دارد؛ برای مرور روشهای ذخیره‌سازی در یک رشتهٔ بیمه می‌توان به تیلور[1] (۲۰۰۰)، انگلند و ورال[2] (۲۰۰۲ و ۲۰۰٦)، و وتریچ و مرز[3] (۲۰۰۸) مراجعه کرد.

مدل‌بندی تؤام مثلثهای تأخیر مرتبط با رشته‌های مختلف بیمه‌ای به منظور درنظرگرفتن ساختارهای وابستگی بین آنها در محاسبهٔ ذخایر، یک مسئلهٔ جدید در نوشتگان ذخیره‌سازی تصادفی است، برای نمونه می‌توان براون[4] (۲۰۰٤)، هس[5] و همکاران (۲۰۰٦)، اشمیت[6] (۲۰۰٦)، مرز و وتریچ (۲۰۰۹b و ۲۰۰۹a)، ژانگ[7] (۲۰۱۰)، شی و فریز[8] (۲۰۱۱)، سالزمن[9] و وتریچ (۲۰۱۲)، ژانگ و همکاران (۲۰۱۲)، دی جونگ[10] (۲۰۱۲)، شی و همکاران (۲۰۱۲)، هپ[11] و وتریچ (۲۰۱۳)، مرز و

---

1. Taylor
2. England and Verrall
3. Wüthrich and Merz
4. Braun
5. Hess
6. Schmidt
7. Zhang
8. Shi and Frees
9. Salzmann
10. De Jong
11. Happ

همکاران (۲۰۱۳) و شی (۲۰۱٤) را نام برد. شی و همکاران (۲۰۱۲) یک مدل لگ نرمال بیزی را در پیش‌بینی خسارات معوق برای رشته‌های وابستهٔ بیمه‌ای پیشنهاد کردند.

در این مقاله، روش شی و همکاران (۲۰۱۲) را به حالتی که خسارات از توزیعهای t ی استیودنت و پی‌یرسون نوع هفتم[1] دومتغیره پیروی می‌کنند، تعمیم می‌دهیم و بر این اساس و در یک چارچوب بیزی به مدل‌بندی توأم داده‌های ذخیره‌سازی خسارات بیمهٔ بدنه و بیمهٔ شخص ثالث اتومبیل یک شرکت بیمهٔ ایرانی می‌پردازیم. توزیعهای t ی استیودنت و پی‌یرسون نوع هفتم، دُمهای سنگین‌تری از توزیع نرمال دارند و لذا اغلب به استنباطهای استواتری[2] منجر می‌شوند. همچنین برای تابع میانگین، مدلهای آنالیز واریانس (ANOVA [3])، آنالیز کوواریانس (ANCOVA [4])، و قدم‌زدن تصادفی[5] را در نظر می‌گیریم.

بعد از برازش مدل، مطالعات حساسیت برای کشف مشاهدات مؤثر، یعنی مشاهداتی که حذف آنها منجر به تغییرات اساسی در برآوردهای پارامترها یا تابعی از پارامترها در یک تحلیل آماری می‌شود، مورد علاقهٔ بسیاری از محققان است. یک روش شناخته‌شده برای شناسایی نمونه‌های مؤثر روش حذف موردی معرفی‌شده توسط کوک[6] (۱۹۸٦) است. این روش با موفقیت برای مدلهای آماری مختلفی به کار گرفته شده است. در چارچوب بیزی، چو[7] و همکاران (۲۰۰۹) تشخیصهای تأثیر حذف موردی برای مدلهای بقا را از طریق روش بیزی توسعه دادند. روش تشخیص بیزی برای مدل‌بندی توأم ذخیره‌سازی خسارات معوق تاکنون در تحقیقات قبلی مورد توجه قرار نگرفته است؛ بنابراین یکی دیگر از اهدف این مقاله، مطالعهٔ اندازه‌ای

---

1. Pearson Type VII
2. Robust
3. Analysis of Variance
4. Analysis of Covariance
5. Random Walk
6. Cook
7. Cho

تشخیصی برای شناسایی نمونه‌های مؤثر بر اساس واگرایی کولبک–لیب‌لر[۱] (K-L) پیشنهادشده توسط چو و همکاران (۲۰۰۹) در مدل‌بندی توأم خسارات در مسئلهٔ ذخیره‌سازی بیمه‌های غیرعمر است. معیارهای بیزی به‌آسانی با نرم‌افزاری استاندارد بیزی مانند OpenBUGS محاسبه می‌شوند.

ادامهٔ مقاله به صورت زیر سازمان‌دهی شده است: در بخش ۲، داده‌های خسارات رشته‌های بیمهٔ بدنه و بیمهٔ شخص ثالث اتومبیل یک شرکت بیمهٔ ایرانی توصیف شده‌اند. در بخش۳، مدل آماری و استنباط بیزی ارائه شده است. روش تشخیصی تأثیر نمونه‌ها به روش بیزی در بخش ٤ بررسی شده است. در بخش ۵ به تحلیل داده‌ها و بیان نتایج پرداخته و در پایان نتیجه‌گیری بیان شده است.

۲. توصیف داده‌ها

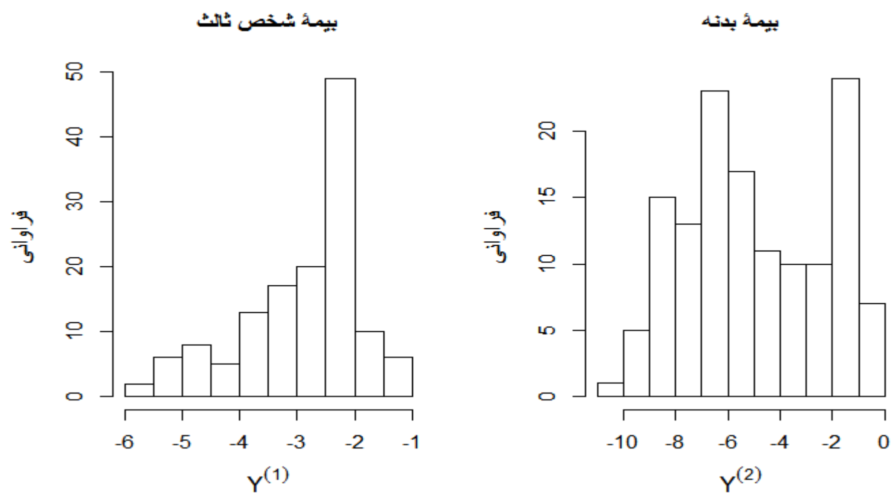
داده‌های مثلثهای تأخیر به‌کاررفته در این مقاله، پرداختهای خسارات و حق‌بیمه‌های عایدشده برای دو رشتهٔ بیمهٔ بدنه و بیمهٔ شخص ثالث اتومبیل در سالهای ۱۳۹۲–۱۳۹۵ یک شرکت بیمهٔ ایرانی است، که به صورت فصلی ثبت شده‌اند. ضریب همبستگی پی‌یرسون بین میزان پرداختهای خسارت سلولهای متناظر در مثلثهای تأخیر این دو رشتهٔ بیمه، ۰/٦۷ است که بیانگر وجود همبستگی بین آنهاست و با توجه به صفربودن پی- مقدار، همبستگی میان دو مثلث در سطح معنی‌داری ۰/۰۱ پذیرفته می‌شود. برای هر یک از دو رشتهٔ بیمه ($l=1,2$)، پرداختهای خسارت را با تقسیم بر حق‌بیمه‌های عایدشدهٔ سال وقوع متناظر آن نرمال می‌کنیم، یعنی $Y_{ij}^{(l)}$ را به صورت

$$Y_{ij}^{(l)} = \log(\frac{X_{ij}^{(l)}}{p_i^{(l)}}),\quad i, j=1,...,16,$$

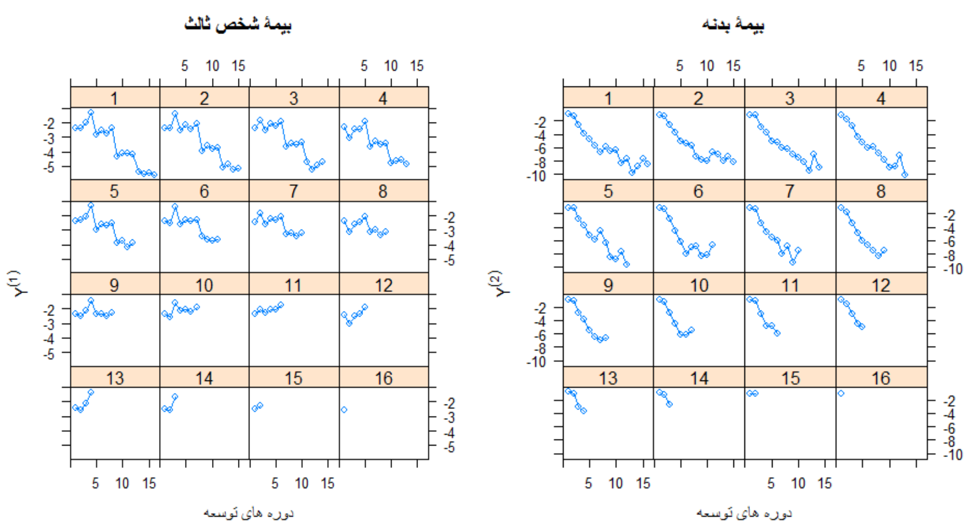
تعریف می‌کنیم، که در آن $p_i^{(l)}$ حق‌بیمهٔ عایدشدهٔ مربوط به فصل صدور iام و مثلث رشتهٔ lام است (Côté and et al., 2016). مقادیر لگاریتم پرداختهای خسارات نرمال‌شده در دو مثلث تأخیر مربوط به رشته‌های بیمهٔ بدنه و بیمهٔ شخص ثالث

1. Kullback-Leibler Divergence

اتومبیل در جدولهای ۳ و ٤ نشان داده شده‌اند. شکلهای ۲ و ۳ به‌ترتیب نمودارهای بافت‌نگار و سریهای زمانی لگاریتم پرداختهای خسارات نرمال‌شدهٔ دو رشتهٔ بیمه را به صورت جداگانه نشان می‌دهد.



شکل ۲. نمودارهای بافت‌نگار لگاریتم پرداختهای خسارات نرمال‌شدهٔ دو مثلث تأخیر



شکل ۳. نمودارهای سریهای زمانی لگاریتم پرداختهای خسارات نرمال‌شدهٔ دو مثلث تأخیر

۳. مدل آماری و استنباط بیزی

۳-۱. توزیع پی‌یرسون نوع هفتم چندمتغیره

توزیع پی‌یرسون نوع هفتم چندمتغیره (Pearson, 1916)، به صورت

$$f(y;\mu,\Lambda,m)=\frac{\Gamma(m)}{\pi^{\frac{d}{2}}\Gamma(m-\frac{d}{2})}|\Lambda|^{-\frac{1}{2}}[1+\Delta^2]^{-m},$$

تعریف می‌شود، که در آن μ بردار میانگین، $|\Lambda|$ نشان‌دهندهٔ دترمینان ماتریس کوواریانس Λ، $\Delta^2=(\mathbf{y}-\mu)^T\Lambda^{-1}(\mathbf{y}-\mu)$ یک فاصلهٔ ماهالانوبیس، m درجهٔ آزادی و d بعد $\mathbf{y}$ است. این توزیع به طور گسترده در مدل‌بندی بسیاری از پدیده‌ها مورد استفاده قرار می‌گیرد. فرناندز و استیل[1] (۲۰۰۰)، نشان دادند که توزیع پی‌یرسون نوع هفتم به ردهٔ توزیعهای آمیخته- مقیاس نرمال چندمتغیره تعلق دارد. بر این اساس، یک نمایش جایگزین سلسله مراتبی دو مرحله‌ای از توزیع $\mathbf{y}$ به صورت

$$\mathbf{Y}\,|\,\Lambda=\lambda\sim N_L(\boldsymbol{\mu},\lambda^{-1}\boldsymbol{\Sigma}),\ \Lambda\sim gamma(\frac{\nu_1}{2},\frac{\nu_2}{2}),$$

است، که در آن $N_L(.,.)$نشان‌دهندهٔ توزیع نرمال چندمتغیره، Λ یک متغیر تصادفی آمیزنده با تابع چگالی گاما و پارامترهای $\mathbf{v}=(\nu_1,\nu_2)$ است. هنگامی که $\nu_1=\nu_2$ باشد، $\mathbf{y}$ دارای توزیع tی استیودنت چندمتغیره با درجهٔ آزادی ν_1 خواهد بود.

۳-۲. مدل بیزی استوار توأم داده‌های ذخیره‌سازی خسارات معوق

فرض می‌کنیم که $Y_{ij}^{(l)}$ با توزیع پی‌یرسون نوع هفتم یا t ی استیودنت دومتغیره مدل‌بندی شوند؛ بنابراین بردار تصادفی $\mathbf{Y}_{ij}$را بر اساس نمایش سلسله مراتبی به صورت

$$\mathbf{Y}_{ij}\,\big|\,\Lambda_{ij}=\lambda_{ij}\sim N_L(\boldsymbol{\mu}_{ij},\lambda_{ij}^{-1}\boldsymbol{\Sigma}),\ \Lambda_{ij}\sim gamma(\frac{\nu_1}{2},\frac{\nu_2}{2}), \qquad (۱)$$

می‌توان در نظر گرفت، که در آن $\mathbf{y}_{ij}$، $\boldsymbol{\mu}_{ij}$ و Σبه صورت

1. Fernandez and Steel

$$\mathbf{y}_{ij}=\begin{bmatrix} y_{ij}^{(1)} \\ \vdots \\ y_{ij}^{(L)} \end{bmatrix},\ \boldsymbol{\mu}_{ij}=\begin{bmatrix} \mu_{ij}^{(1)} \\ \vdots \\ \mu_{ij}^{(L)} \end{bmatrix} \quad \text{و} \quad \Sigma=\begin{bmatrix} \sigma^{(1,1)} & \cdots & \sigma^{(1,L)} \\ \vdots & \ddots & \vdots \\ \sigma^{(L,1)} & \cdots & \sigma^{(L,L)} \end{bmatrix}$$

تعریف می‌شوند. ماتریس کوواریانس Σ بیانگر وابستگی دوبه‌دوی سلولها از یک مثلث و سلول متناظر از مثلث دیگر است.

برای lامین مثلث تأخیر، مدلهای زیر را برای میانگین در نظر می‌گیریم:

۱. مدل ANOVA: مدل ANOVA برای $i\in\{1,\dots,I\}$ و $i+j=1,\dots,J$، به صورت (Verral, 1991)

$$\mu_{ij}^{(l)}=\mu^{(l)}+\alpha_i^{(l)}+\beta_j^{(l)}, \qquad (۲)$$

است و در شرطهای $\sum_{i=1}^{I}\alpha_i^{(l)}=\sum_{j=1}^{J}\beta_j^{(l)}=0$ صدق می‌کند. در رابطهٔ (۲) اثرهای ثابت $\alpha_i^{(l)}$ و $\beta_j^{(l)}$ به‌ترتیب پارامترهای فصلهای مبدأ و تأخیر را نشان می‌دهند.

۲. مدل ANCOVA: در مدل ANCOVA، فرض می‌کنیم که اثرهای فصل مبدأ خطی هستند، بنابراین تابع میانگین (۲) به صورت

$$\mu_{ij}^{(l)}=\mu^{(l)}+i\,\alpha^{(l)}+\beta_j^{(l)},$$

تغییر می‌کند و در شرط $\sum_{j=1}^{J}\beta_j^{(l)}=0$ صدق می‌کند. در این مدل فرض می‌کنیم اثرهای فصل مبدأ، تصادفی و اثرهای فصل تأخیر ثابت هستند.

۳. مدل قدم‌زدن تصادفی: با استفاده از رابطهٔ (۲) و

$$\alpha_i^{(l)}=\alpha_{i-1}^{(l)}+h_i^{(l)}\ ,\ \ \beta_j^{(l)}=\beta_{j-1}^{(l)}+u_j^{(l)}$$

یک مدل قدم‌زدن تصادفی را برای تابع میانگین در نظر می‌گیریم که برای $i=2,\dots,I$ و $j=2,\dots,J-i+1$ در شرطهای $\alpha_1^{(l)}=\beta_1^{(l)}=0$ صدق می‌کنند. در این مدل فرض می‌کنیم اثرهای فصل مبدأ و تأخیر هر دو تصادفی هستند.

۳-۳. توزیعهای پیشینی و پسینی

برای کامل‌کردن توصیف بیزی مدلها، مشخص‌کردن توزیع پیشینی برای همهٔ پارامترهای نامعلوم ضروری است. به دلیل آنکه هیچ اطلاعاتی دربارهٔ پارامترهای مدلهای پیشنهادشده وجود ندارد، توزیعهای پیشینی مزدوج کم‌آگاهنده[1] نسبت داده می‌شود؛ بنابراین، بر اساس نوشتگان موضوع برای پارامترهای مدل ANOVA، پیشین‌های

$$\boldsymbol{\alpha}_i = \left(\alpha_i^{(1)},...,\alpha_i^{(L)}\right)' \sim N_L(\mathbf{0},\Sigma_{\boldsymbol{\alpha}_i}),\ \boldsymbol{\beta}_j = \left(\beta_j^{(1)},...,\beta_j^{(L)}\right)' \sim N_L(\mathbf{0},\Sigma_{\boldsymbol{\beta}_j}),$$
$$\tau_{\boldsymbol{\Sigma}} = \boldsymbol{\Sigma}^{-1} \sim Wishart(\mathbf{R},L),$$

را در نظر می‌گیریم، که در آن $\boldsymbol{\Sigma}_{\boldsymbol{\alpha}_i} = diag\,(\sigma^2_{\alpha_i^{(1)}},...,\sigma^2_{\alpha_i^{(L)}})$ و $\boldsymbol{\Sigma}_{\boldsymbol{\beta}_j} = diag\,(\sigma^2_{\beta_j^{(1)}},...,\sigma^2_{\beta_j^{(L)}})$ و R یک ماتریس نیمه‌مثبت با ابعاد $L\times L$ است. برای مدل ANCOVA پیشین‌های

$$\boldsymbol{\alpha} = \left(\alpha^{(1)},...,\alpha^{(L)}\right)' \sim N_L(\mathbf{0},\boldsymbol{\Sigma}_{\boldsymbol{\alpha}}),\ \ \tau_\alpha = \sigma_\alpha^{-2} \sim Gamma(\kappa_1,\upsilon_1),$$
$$\boldsymbol{\beta}_j = \left(\beta_j^{(1)},...,\beta_j^{(L)}\right)' \sim N_L(\mathbf{0},\Sigma_{\boldsymbol{\beta}_j}),\ \ \tau_{\boldsymbol{\Sigma}} = \boldsymbol{\Sigma}^{-1} \sim Wishart(\mathbf{R},L),$$

به کار گرفته می‌شود که در آن $\boldsymbol{\Sigma}_{\boldsymbol{\alpha}} = \sigma_\alpha^2 \mathbf{I}_{L\times L}$است، و برای مدل قدم‌زدن تصادفی، پیشین‌های

$$\tau_{\boldsymbol{\Sigma}} = \boldsymbol{\Sigma}^{-1} \sim Wishart(\mathbf{R},L),\ h_i^{(l)} \sim N\,(0,\sigma_h^2),\ \ \tau_h = \sigma_h^{-2} \sim Gamma(\kappa_1,\upsilon_1),$$
$$u_i^{(l)} \sim N\,(0,\sigma_u^2),\ \ \tau_u = \sigma_u^{-2} \sim Gamma(\kappa_2,\upsilon_2),$$

در نظر گرفته می‌شوند. توزیع پسینی توأم پارامترها با استفاده از قضیهٔ بیز، با ترکیب‌کردن توزیع درستنمایی در رابطهٔ (۱) و توزیعهای پیشینی به دست می‌آید. با توجه اینکه توزیعهای پسینی به‌دست‌آمده دارای شکل پیچیده‌ای هستند، برای به‌دست‌آوردن برآورد بیز پارامترهای مدلها می‌توان از روشهای MCMC مانند نمونه‌گیر گیبز[2] و الگوریتم متروپولیس-هستینگ[1] که با نمونه‌گیری مکرر از توزیع

1. Conjugate but Weakly Informative Prior
2. Gibbs Sampling

شرطی کامل هر پارامتر عمل می‌کنند، استفاده کرد. روشهای MCMC به‌آسانی با نرم‌افزارهای استاندارد بیزی مانند OpenBUGS محاسبه می‌شوند.

۳-٤. مقایسهٔ مدلها

برای مقایسهٔ مدلها، معیارهای اطلاع انحراف[۲] (DIC) (Spiegelhalter et al., 2002) و لگاریتم شبه درستنمایی حاشیه‌ای[۳] (LPML) را به کار برده‌ایم. فرض کنیم $\mathbf{y}_{ij}=(y_{ij}^{(1)},\dots,y_{ij}^{(L)})'$، و $\boldsymbol{\theta}$ مجموعهٔ همهٔ پارامترهای مدل باشد. معیار انحراف[٤] را که به صورت

$$D(\boldsymbol{\theta})=-2\log f(\mathbf{y}|\boldsymbol{\theta})=-2\sum_{i=1}^{I}\sum_{j=1}^{J+1-i}\log f(\mathbf{y}_{ij}|\boldsymbol{\theta}),$$

تعریف می‌شود، در نظر بگیرید. پس $\bar{D}(\boldsymbol{\theta})=E(D(\boldsymbol{\theta}))$ معیاری در برازش مدل است و می‌تواند با استفاده از نمونه‌های MCMC، به صورت

$$\bar{D}=-2\frac{1}{M}\sum_{i=1}^{I}\sum_{j=1}^{J+1-i}\sum_{m=1}^{M}\log f(\mathbf{y}_{ij}|\boldsymbol{\theta}^{(m)}),$$

تقریب زده شود، که در آن $\boldsymbol{\theta}^{(m)}$، mامین تکرار زنجیر MCMC مدل و M تعداد تکرارهاست. با درنظرگرفتن اندازهٔ انحراف، معیار بیزی DIC به صورت

$$DIC=\bar{D}+\hat{p}_D,$$

تعریف شده است، که در آن $\hat{p}_D=\bar{D}-\hat{D}$ با $\hat{D}=D(\frac{1}{M}\sum_{m=1}^{M}\theta^{(m)})$ است. هر اندازه مقدار DIC کوچکتر باشد، مدل بهتر برازش داده شده است. آمارهٔ عرض پیشگوی شرطی[٥] (CPO)، معیار مشهور دیگری است که معمولاً برای مقایسهٔ مدلها در متون

1. Metropolis Hastings Algorithm
2. Deviance Information Criterion
3. Log Pseudo Marginal Likelihood
4. Deviance
5. Conditional Predictive Ordinate

بیزی به کار گرفته می‌شود (برای توضیحات بیشتر دربارهٔ این آماره و کاربردهای آن به گلفند[1] و همکاران (۱۹۹۲) مراجعه کنید). آمارهٔ CPO به صورت

$$CPO_{ij} = p(\mathbf{y}_{ij} | \mathbf{y}_{(-ij)}) = \int_{\boldsymbol{\theta}\in\Theta} f(\mathbf{y}_{ij} | \boldsymbol{\theta}) \pi(\boldsymbol{\theta} | \mathbf{y}_{(-ij)}) = \left\{ \int_{\boldsymbol{\theta}\in\Theta} \frac{\pi(\boldsymbol{\theta}|\mathbf{y})}{f(\mathbf{y}_{ij}|\boldsymbol{\theta})} d\boldsymbol{\theta} \right\}^{-1},$$

تعریف می‌شود، که در آن $\mathbf{y}_{(-ij)}$ نشان‌دهندهٔ داده‌ها با حذف بردار $\mathbf{y}_{ij}$ است و $p(\mathbf{y}_{ij}|\mathbf{y}_{(-ij)})$ توزیع پیشگوی یک مشاهده جدید به شرط $\mathbf{y}_{(-ij)}$ است و $\pi(\boldsymbol{\theta}|\mathbf{y})$ توزیع پسینی $\boldsymbol{\theta}$ است. CPO_{ij}را می‌توان با استفاده از نمونه‌های MCMC به صورت

$$CPO_{ij} = \left\{ \frac{1}{M} \sum_{m=1}^{M} \frac{1}{f(\mathbf{y}_{ij}|\boldsymbol{\theta}^{(m)})} \right\}^{-1},$$

تقریب زد (Chen et al., 2000) و بر اساس آن معیار LPML به صورت

$$LPML = \sum_{i=1}^{I} \sum_{j=1}^{J+1-i} \log(CPO_{ij}),$$

تعریف می‌شود. یک LMPL بزرگتر نشان‌دهندهٔ عملکرد بهتر مدل است.

٤. مباحث تشخیصی تأثیر موردی بیزی

مباحث تشخیصی تأثیر، یک گام مهم در تحلیل مدلهای آماری است. فرض کنید $L(\boldsymbol{\theta}|\mathbf{y})$ و $L(\boldsymbol{\theta}|\mathbf{y}_{(-ij)})$ به‌ترتیب نشان‌دهندهٔ تابع درستنمایی بر اساس داده‌های کامل و تابع درستنمایی بر اساس داده‌ها با حذف ijامین نمونه باشند. توزیعهای پسینی برای داده‌های کامل و داده‌ها با حذف ijامین نمونه می‌تواند به‌ترتیب به صورت $\pi(\boldsymbol{\theta}|\mathbf{y}) \propto L(\boldsymbol{\theta}|\mathbf{y})\pi(\boldsymbol{\theta})$ و $\pi(\boldsymbol{\theta}|\mathbf{y}_{(-ij)}) \propto L(\boldsymbol{\theta}|\mathbf{y}_{(-ij)})\pi(\boldsymbol{\theta})$ تعریف شوند، که در آن $\pi(\boldsymbol{\theta})$ توزیع پیشینی $\boldsymbol{\theta}$ است. فرض کنید $K(P, P_{(ij)})$ نشان‌دهندهٔ فاصلهٔ کولبک- لیبلر بین P و $P_{(ij)}$ باشد، که در آن P نشان‌دهندهٔ $\pi(\boldsymbol{\theta}|\mathbf{y})$ و $P_{(ij)}$ نشان‌دهندهٔ $\pi(\boldsymbol{\theta}|\mathbf{y}_{(-ij)})$ است، لذا

1. Gelfand

$$K(P,P_{(ij)})=\int \pi(\boldsymbol{\theta}|\mathbf{y})\log\left\{\frac{\pi(\boldsymbol{\theta}|\mathbf{y})}{\pi(\boldsymbol{\theta}|\mathbf{y}_{(-ij)})}\right\}d\boldsymbol{\theta}$$

$$=-\log[E_{\boldsymbol{\theta}|\mathbf{y}}[\{f(\mathbf{y}_{ij}|\boldsymbol{\theta})\}^{-1}]]^{-1}+E_{\boldsymbol{\theta}|\mathbf{y}}[\log\{f(\mathbf{y}_{ij}|\boldsymbol{\theta})\}],$$

بنابراین $K(P,P_{(ij)})$ تأثیر حذف ijامین نمونه از داده‌های کامل روی توزیع پسینی توأم $\boldsymbol{\theta}$ را اندازه‌گیری می‌کند. برآورد مونت کارلوی $K(P,P_{(ij)})$ به صورت

$$\hat{K}(P,P_{(ij)})=-\log\left(\left\{\frac{1}{M}\sum_{m=1}^{M}\frac{1}{f(\mathbf{y}_{ij}|\boldsymbol{\theta}^{(m)})}\right\}^{-1}\right)+\frac{1}{M}\sum_{m=1}^{M}\log[f(\mathbf{y}_{ij}|\boldsymbol{\theta}^{(m)})],$$

به دست می‌آید. یک مشاهده با مقدار کولبک- لیب‌لر بزرگ به‌عنوان یک مشاهدۀ مؤثر در نظر گرفته می‌شود.

۵. تحلیل داده‌ها

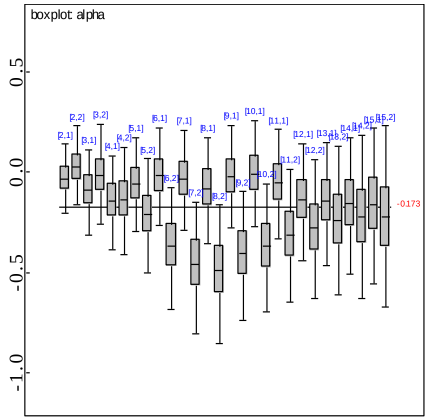
برای انجام مدلها با استفاده از روش شبیه‌سازی بیزی، با به کار بردن نرم‌افزار OpenBUGS، نمونه‌گیر گیبز را برای ۲۵۰۰۰۰ تکرار اجرا کردیم. پس از حذف اولین۵۰۰۰۰ تکرار به‌عنوان دورۀ داغیدن، برای کاهش همبستگی نمونه‌های شبیه‌سازی‌شده به صورت سیستماتیک هر ٤۰امین تکرار در زنجیر به کار گرفته شد. درنهایت،۵۰۰۰ نمونۀ شبیه‌سازی‌شده خواهیم داشت. برای ارزیابی همگرایی زنجیر مارکف، نمودارهای روند، نمودارهای خودهمبستگی و خطاهای مونت کارلوی پارامترها مورد بررسی قرار گرفت. جدول ۲، مقایسۀ بین مدلهای مختلف برازش‌داده‌شده را با استفاده از معیارهای انتخاب مدل بحث‌شده در بخش ۳-٤ نشان می‌دهد. به منظور ارزیابی توانایی پیش‌بینی یک سال جلوتر مدلها، ۱۵ فصل اول برای برازش داده‌ها به کار گرفته می‌شود.

جدول ۲. مقایسۀ بین مدلهای برازش‌داده‌شده

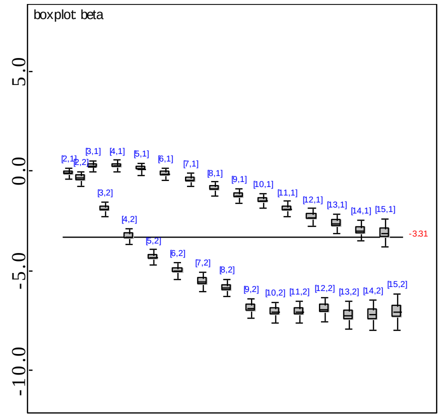
مدل میانگین	توزیع خسارت	DIC	LMPL
ANOVA	نرمال	٤۰۰/۹	-۲۰۵/۷۸
	پی‌یرسون نوع هفتم	۳۸۷/٤	-۱۹۸/۷۷
	t ی استیودنت	۳۸۱/٤	-۱۹۸/۵۷
ANCOVA	نرمال	۳۸۹/٤	-۲۰۱/۰۳

مدل میانگین	توزیع خسارت	DIC	LMPL
	پی‌یرسون نوع هفتم	۳۷٤	-۱۹٤/۷۲
	t ی استیودنت	۳٦۹/۸	-۱۹٥/۳٤
قدم‌زدن تصادفی	نرمال	۳٦۷/٦	-۱۸٦/۹٦
	پی‌یرسون نوع هفتم	۳٥۰	-۱۸۱/۳۹
	t ی استیودنت	**۳٤٥/٥**	**-۱۷۹/٦۳**

بر طبق معیارها، مدل قدم‌زدن تصادفی با توزیع t ی استیودنت دومتغیره برای لگاریتم پرداختهای خسارات نرمال‌شده، برازش بهتری دارد. در این مدل برای بیمهٔ شخص ثالث (مثلث ۱) و بیمهٔ بدنه (مثلث ۲)، میانگینهای پسینی (انحرافهای استاندارد) میانگینهای کل $\mu^{(1)}$، $\mu^{(2)}$ به‌ترتیب برابر با ۲/۳۲- (۰/۱۳) و ۰/۷٥- (۰/۱۷) است. شکل ٤، نمودار جعبه‌ای نمونه‌های پسینی اثر فصلهای مبدأ در مدل قدم‌زدن تصادفی با توزیع t ی استیودنت دومتغیره را برای دو مثلث نشان می‌دهد. یک روند غیرثابت در بین برآورد میانگین اثر فصلهای مبدأ در مثلث بیمهٔ شخص ثالث و یک روند کاهشی تا ۸امین فصل و پس از آن یک روند افزایشی در بین برآورد میانگین اثر فصلهای مبدأ در مثلث بیمهٔ بدنه مشاهده می‌شود. شکل ٥، نمودار جعبه‌ای نمونه‌های پسینی اثر فصلهای تأخیر در مدل قدم‌زدن تصادفی با توزیع t ی استیودنت دومتغیره برای دو مثلث را نشان می‌دهد. یک روند کاهشی در برآورد میانگین اثر فصلهای تأخیر از اولین تا آخرین فصل تأخیر در هر دو مثلث مشاهده می‌شود.



شکل ٤. نمودار جعبه‌ای نمونه‌های پسینی اثر فصلهای مبدأ برای مدل قدم‌زدن تصادفی با توزیع t ی استیودنت دومتغیره در دو مثلث



شکل ٥. نمودار جعبه‌ای نمونه‌های پسینی اثر فصلهای تأخیر برای مدل قدم‌زدن تصادفی با توزیع t ی استیودنت دومتغیره در دو مثلث

میزان ادعاهای پیش‌بینی‌شده مثلث پایینی که در واقع هدف نهایی در مسئلهٔ ذخیره‌سازی است در هر دو رشتهٔ بیمه بر اساس مدل قدم‌زدن تصادفی با توزیع t ی استیودنت دومتغیره برای لگاریتم پرداختهای خسارات نرمال‌شده در جدولهای ۳ و ٤ نشان داده شده است.

جدول ۳. مثلث تأخیر میزان لگاریتم پرداختهای خسارات نرمال‌شدهٔ بیمهٔ شخص ثالث (مثلث بالایی) و میزان لگاریتم پرداختهای خسارات نرمال‌شدهٔ پیش‌بینی‌شده بر اساس مدل قدم‌زدن تصادفی با توزیع tای استیودنت دومتغیره (مثلث پایینی)

سال خسارت	فصل خسارت	فصل توسعهٔ خسارت														
		۰	۱	۲	۳	٤	۵	٦	۷	۸	۹	۱۰	۱۱	۱۲	۱۳	۱٤
۱۳۹۲	بهار	-۲/۳۵	-۲/۳۹	-۱/۹۷	-۱/۳٤	-۲/۸۵	-۲/۵۲	-۲/۷۷	-۲/۳٦	-٤/۳۵	-٤/۰۸	-٤/۱۳	-٤/۱۹	-۵/٤	-۵/۵۸	-۵/٤۷
	تابستان	-۲/٤	-۲/۳۵	-۱/٤۳	-۲/۵٦	-۲/۱۸	-۲/٤٦	-۲/۰۵	-۳/۹۷	-۳/٦۲	-۳/۸	-۳/۷۱	-۵/۰۹	-٤/۸۵	-۵/۲٤	-۵/٤٦
	پاییز	-۲/۳۵	-۱/۸۲	-۲/۵۳	-۲/۱	-۲/۲۵	-۱/۹۳	-۳/٦۷	-۳/٤۷	-۳/٤۹	-۳/۳۸	-٤/٦۹	-۵/۲۱	-٤/۹۵	-۵/٤۳	-۵/۵۳
	زمستان	-۲/۳۲	-۳/۰۵	-۲/٤۵	-۲/٤۲	-۱/۹۳	-۳/٦۳	-۳/۲٦	-۳/۵۱	۳/٤۱	-٤/۸۲	-٤/٦٦	-٤/۵۹	-۵/۱	-۵/٤۸	-۵/۵۹
۱۳۹۳	بهار	-۲/٤	-۲/۳۲	-۲/۱	-۱/۲۸	-۲/۹٤	-۲/۵۷	-۲/٦۳	-۲/۵۳	-۳/۸٤	-۳/۷۵	-٤/۱۷	-٤/۷۱	-۵	-۵/۳۸	-۵/۵۷
	تابستان	-۲/٤	-۲/٤۹	-۱/۳۸	-۲/٦	-۲/۲٦	-۲/۳۷	-۲/۳۲	-۳/٤۱	-۳/٦۷	-۳/۷۲	-٤/۲٤	-٤/٦۷	-٤/۹۸	-۵/۳۵	-۵/٤٦
	پاییز	-۲/٤٤	-۱/۸۳	-۲/۵۹	-۲/۲	-۲/۲۷	-۲/۰۹	-۳/۲٤	-۳/۲۲	-۳/۳۹	-۳/۸۵	-٤/۲٦	-٤/٦۸	-٤/۹۹	-۵/۳٦	-۵/٤۸
	زمستان	-۲/۳٤	-۳/۱٤	-۲/۵٦	-۲/٤۱	-۲/۰٤	-۳/۱٤	-۳	-۳/۳۳	-۳/٦۳	-۳/۸۹	-٤/۳۱	-٤/۷۳	-۵/۰٤	-۵/٤۱	-۵/۵۱
۱۳۹٤	بهار	-۲/۳۲	-۲/۵۲	-۲/۱۳	-۱/٤٤	-۲/۳٦	-۲/۳٦	-۲/٤۹	-۳/۲۲	-۳/۵۸	-۳/۸۳	-٤/۲۵	-٤/٦٦	-٤/۹۸	-۵/۳۵	-۵/٤٦
	تابستان	-۲/۳٤	-۲/٦	-۱/۵۷	-۲/۱٦	-۲/۰۹	-۲/۲٤	-۲/۷۷	-۳/۲۱	-۳/۵۷	-۳/۸۲	-٤/۲۳	-٤/٦۵	-٤/۹۷	-۵/۳۵	-۵/٤٤
	پاییز	-۲/۳٦	-۲/۰۹	-۲/۳	-۲/۰۵	-۲/۰۵	-۲/۵۱	-۲/۸۱	-۳/۲۵	-۳/٦	-۳/۸٦	-٤/۲۸	-٤/۷	-۵/۰۱	-۵/۳۹	-۵/٤۹
	زمستان	-۲/٤٦	-۳/۰۲	-۲/۵۳	-۲/۳۲	-۲/۳۵	-۲/۵۹	-۲/۹	-۳/۳٤	-۳/۷	-۳/۹٦	-٤/۳٦	-٤/۷۸	-۵/۰۹	-۵/٤۸	-۵/۵۸
۱۳۹۵	بهار	-۲/٤	-۲/٦	-۲/۱۵	-۲/۲	-۲/۳۷	-۲/۵۹	-۲/۹۱	-۳/۳۵	-۳/۷۱	-۳/۹۷	-٤/۳۷	-٤/۷۹	-۵/۱	-۵/٤۹	-۵/۵۹
	تابستان	-۲/۵۱	-۲/۵۹	-۲/۲۵	-۲/۲۱	-۲/۳۷	-۲/٦	-۲/۹۲	-۳/۳۵	-۳/۷۱	-۳/۹۷	-٤/۳۸	-٤/۸۱	-۵/۱۱	-۵/٤۹	-۵/٦
	پاییز	-۲/۵	-۲/۵۸	-۲/۲۵	-۲/۲۳	-۲/۳٦	-۲/٦۱	-۲/۹۲	-۳/۳۵	-۳/۷۱	-۳/۹۷	-٤/۳۹	-٤/۸۱	-۵/۱۲	-۵/٤۹	-۵/۵۹

جدول ٤. مثلث تأخیر میزان لگاریتم پرداختهای خسارات نرمال‌شدهٔ بیمهٔ بدنه (مثلث بالایی) و میزان لگاریتم پرداختهای خسارات نرمال‌شدهٔ پیش‌بینی‌شده بر اساس مدل قدم‌زدن تصادفی با توزیع tای استیودنت دومتغیره (مثلث پایینی)

سال خسارت	فصل خسارت	فصل توسعهٔ خسارت														
		۰	۱	۲	۳	٤	۵	٦	۷	۸	۹	۱۰	۱۱	۱۲	۱۳	۱٤
۱۳۹۲	بهار	−۱/۰۱	−۱/۲۸	−۲/٦۳	−۳/۹۸	−٤/۸۱	−۵/٦۵	−٦/٦۸	−۵/۸٤	−٦/٤۸	−٦/٤	−۸/۲۵	−۷/٦٦	−۹/۸	−۸/۸۳	−۷/۷۵
	تابستان	−۱/۰۹	−۱/۳۳	−۲/۵٦	−۳/۷۸	−۵/۰۵	−۵/۳۹	−۵/٦۸	−۷/۲۸	−۷/۸۷	−۸/۰٤	−٦/٦۸	−٦/۹۸	−۸/۰٤	−۷/۳٦	−۷/۸۱
	پاییز	−۱/۰۹	−۱/۱٦	−۲/۹	−۳/٦۸	−۵/۰۱	−۵/۱۸	−٦/۰۸	−٦/۲۵	−٦/۹۸	−۷/۵۸	−۸/۱٦	−۹/٤	−٦/۹۸	−۸/۰٤	−۷/۸٦
	زمستان	−۱/۱٦	−۱/۸۱	−۲/۸۵	−٤/٤۲	−۵/۲۲	−٦/۰۲	−۵/۹۳	−٦/۸۷	−۷/۸۱	−۸/۹۳	−۸/۸۵	−۷/۱۲	−۸/۱۵	−۸/۱٤	−۷/۹۸
۱۳۹۳	بهار	−۱/۰٦	−۱/۱۸	−۲/۷	−۳/۷۱	−۵/۱۹	−۵/۸۳	−٤/٦	−٦/۳۳	−۸/٤	−۸/۷۱	−۷/٦	−۷/۹۳	−۸/۲۱	−۸/۲۲	−۸/۰۵
	تابستان	−۱/۰٤	−۱/۳	−۲/۷	−٤/۵۲	−٦/۱۳	−۸/۰٤	−٦/۹۷	−٦/۸	−۸/۳٤	−۸/۱۱	−۸/۲۳	۸/۰۹	−۸/۳۹	−۸/۳۹	−۸/۲۲
	پاییز	−۱/۱۱	−۱/۳	−۳/٤۳	−٤/٦۸	−۵/۵۸	−۵/۹٤	−۷/۹۲	−٦/۷۹	−۹/۲۸	−۸/۳۵	−۸/۳۳	−۸/۱۷	−۸/٤٦	−۸/٤۷	−۸/۳۱
	زمستان	−۱/۱۷	−۱/۸۳	−۳/۳٦	−٤/۸۹	−۵/۹۹	−٦/٦۳	−۷/٤۲	−۸/۲۳	−۸/۱٦	−۸/۳۸	−۸/۳۵	−۸/۲	−۸/۵	−۸/۵۱	−۸/۳۳
۱۳۹٤	بهار	−۰/۸۷	−۱/۱۵	−۲/۹٤	−۳/۸۲	−۵/٤۸	−٦/٤۵	−٦/۹۸	−۷/۰٦	−۸/۱	−۸/۲۹	−۸/۲۷	−۸/۱۱	−۸/٤۳	−۸/٤۳	−۸/۲۵
	تابستان	−۰/۸٦	−۱/۲۹	−۲/۹۲	−٤/۵	−٦/۱	−٦/۱۸	−٦/٦۹	−۷/۰۱	−۸/۰۵	−۸/۲۵	−۸/۲٤	−۸/۰٦	−۸/۳۷	−۸/۳۸	−۸/۲۱
	پاییز	−۰/۸٦	−۱/۰۵	−۳/۱۱	−٤/۷۹	−٤/۸۹	−٦/۰۷	−٦/٦٦	−٦/۹٤	−۷/۹۸	−۸/۱۹	−۸/۱۸	−۸/۰۳	−۸/۳۱	−۸/۳۲	−۸/۱۷
	زمستان	−۰/۹٦	−۱/٦۱	−۳	−٤/۵۸	−۵/٤	−٦/۰٦	−٦/٦۱	−٦/۹	−۷/۹۵	−۸/۱۷	−۸/۱۵	−۷/۹۸	−۸/۲۹	−۸/۲۹	−۸/۱۲
۱۳۹۵	بهار	−۰/۷۲	−۱/۱۲	−۲/۹۹	−٤/۲۸	−۵/۳۷	−٦/۰۱	−٦/۵۷	−٦/۸۸	−۷/۹۲	−۸/۱۳	−۸/۱۲	−۷/۹٤	−۸/۲٦	−۸/۲٦	−۸/۰۸
	تابستان	−۰/۸۷	−۱/۱٦	−۲/۹	−٤/۲۵	−۵/۳۳	−۵/۹۸	−٦/۵٦	−٦/۸۵	−۷/۸۹	−۸/۱	−۸/۰۹	−۷/۹۳	−۸/۲۱	−۸/۲۳	−۸/۰٦
	پاییز	−۰/۹۸	−−۱/۳٦	−۲/۹	−٤/۲۸	−۵/۳۵	−٦/۰۱	−٦/۵۵	−٦/۸۵	−۷/۹۱	−۸/۱	−۸/۰۸	−۷/۹۵	−۸/۲۳	−۸/۲۳	−۸/۰٦

میانگین پسینی و انحراف استاندارد ذخایر دو مثلث بر اساس مدلهای پیشنهادشده و بر اساس مدل نردبان زنجیره‌ای[1] در جدول ٥ گزارش شده است. برآورد ذخایر با استفاده از مدل نردبان زنجیره‌ای در دو مثلث به طور جداگانه و بدون درنظرگرفتن وابستگیها به‌دست‌آمده است. نتایج نشان می‌دهد که مدل نردبان زنجیره‌ای مقدار ذخایر را کمتر از روشهای پیشنهادشده برآورد می‌کند و انحراف استاندارد برآوردهای به‌دست‌آمده بر اساس این مدل نیز از انحراف استاندارد برآوردهای به‌دست‌آمده بر اساس مدلهای آنالیز کوواریانس و قدم‌زدن تصادفی بزرگتر است.

جدول ٥. برآوردهای ذخایر

مدل میانگین	توزیع خسارت	بیمهٔ شخص ثالث		بیمهٔ بدنه	
		میانگین	انحراف استاندارد	میانگین	انحراف استاندارد
ANOVA	نرمال	٣٥٣٢×١٠^{١٠}	٦٣٤٤×١٠^{٩}	١٠٣٢×١٠^{٩}	$٨٤٤/٩ \times ١٠^{٩}$
	پی‌یرسون نوع هفتم	٣٥٨٧×١٠^{١٠}	٦٤٢٨×١٠^{٩}	١٠١٠×١٠^{٩}	$٩٧٥/٩ \times ١٠^{٩}$
	t ی استیودنت	٣٦٠٤×١٠^{١٠}	٦٩٧٩×١٠^{٩}	$٩٩٢/٤ \times ١٠^{٩}$	$٨٥٣/٩ \times ١٠^{٩}$
ANCOVA	نرمال	٣٥٧٥×١٠^{١٠}	٤٤٤٧×١٠^{٩}	$٦٠٤/٥ \times ١٠^{٩}$	$٢٨٤/٨ \times ١٠^{٩}$
	پی‌یرسون نوع هفتم	٣٥٨٨×١٠^{١٠}	٤٤٦٦×١٠^{٩}	$٦٣٢/٣ \times ١٠^{٩}$	$٣٣٦/٧ \times ١٠^{٩}$
	t ی استیودنت	٣٥٩٩×١٠^{١٠}	٤٥٧١×١٠^{٩}	$٦٣٥/٣ \times ١٠^{٩}$	٣١٠×١٠^{٩}
قدم‌زدن تصادفی	نرمال	٣٤٩٣×١٠^{١٠}	٤٦٩٤×١٠^{٩}	$٨٢٠/٥ \times ١٠^{٩}$	$٣٧٢/١ \times ١٠^{٩}$
	پی‌یرسون نوع هفتم	٣٥١٩×١٠^{١٠}	٤٥٢٥×١٠^{٩}	$٨٢٠/١ \times ١٠^{٩}$	٣٦٣×١٠^{٩}
	t ی استیودنت	٣٥٥١×١٠^{١٠}	٤٧٦٨×١٠^{٩}	$٨٢٨/٨ \times ١٠^{٩}$	$٤٨٥/٩ \times ١٠^{٩}$
مدل نردبان زنجیره‌ای		٣١٣٤×١٠^{١٠}	٥٠٢٦×١٠^{٩}	٧٥٤×١٠^{٩}	٩٤٧×١٠^{٩}

همان‌گونه که قبلاً بیان شد، علی‌رغم موجودبودن اطلاعات برای ١٦ فصل، مدلها تنها به داده‌های مربوط به ١٥ فصل اول برازش داده شد و خسارات پرداخت‌شده در فصل ١٦ام پیش‌بینی و برای این مقادیر، میانگین اریبی نسبی[2]، با استفاده از فرمول

$$\overline{\text{Rel.Bias}} = {}^{1}\!\!\diagup_{NM} \sum_{m=1}^{M} \sum_{A} \left(\hat{y}_{ij}^{(l)}[m] - y_{ij}^{(l)}[obs.]\right) / \, y_{ij}^{(l)}[obs.],$$

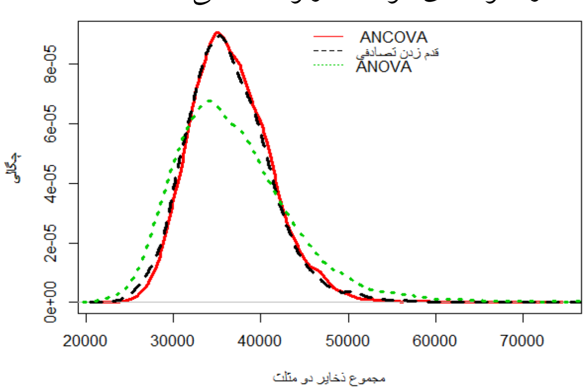
1. Chain-Ladder
2. Relative Bias

برای هر دو مثلث بر اساس مدل قدم‌زدن تصادفی با توزیع t ی استیودنت دومتغیره محاسبه شده است، که در رابطهٔ فوق $y_{ij}^{(l)}[obs.]$ و $\hat{y}_{ij}^{(l)}[m]$ به‌ترتیب مقدار مشاهده‌شده و پیش‌بینی‌شده در mامین تکرار زنجیر MCMC در مثلث lام و A مجموعهٔ اندیسهای مربوط به خسارات پرداخت‌شده در فصل ۱۶ام است که در داده‌های مورد مطالعه برای هر مثلث $N=۱۴$ است. همچنین این معیار برای پیش‌بینی‌های خسارات پرداخت‌شده در فصل ۱۶ام با استفاده از مدل نردبان زنجیره‌ای محاسبه شد.

جدول ٦. برآورد میانگین توان دوم خطای پیش‌بینی

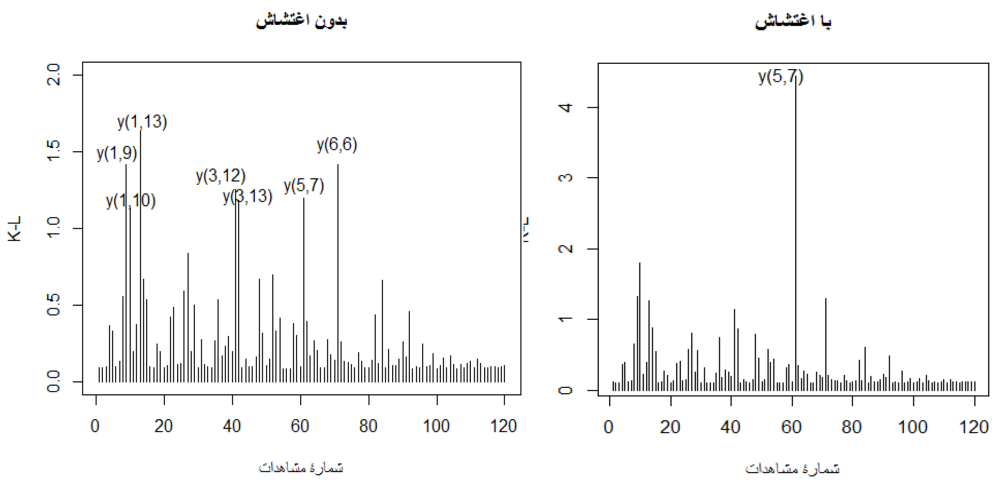
نوع بیمه	مدل	$\overline{\text{Rel.Bias}}$
شخص ثالث	نردبان زنجیره‌ای	۰/۴۵-
	قدم‌زدن تصادفی با توزیع t ی استیودنت	۰/۳۷-
بدنه	مدل نردبان زنجیره‌ای	۱/۲٤
	قدم‌زدن تصادفی با توزیع t ی استیودنت	۱/۰۱

جدول ٦ نشان می‌دهد که در هر دو مثلث عملکرد مدل قدم‌زدن تصادفی با توزیع tی استیودنت دومتغیره از مدل نردبان زنجیره‌ای بهتر بوده است. شکل ٦، نمودار توزیع پیش‌بینی پسینی مجموع ذخایر دو مثلث را بر اساس سه مدل ANOVA، ANCOVA و قدم‌زدن تصادفی با توزیع t ی استیودنت دومتغیره برای لگاریتم پرداختهای خسارات نرمال‌شده با استفاده از نمونه‌های مونت کارلو نشان می‌دهد.



شکل ٦. نمودارهای توزیع پیش‌بینی پسینی مجموع ذخایر دو مثلث بر اساس توزیع t ی استیودنت

برای کشف نقاط مؤثر در داده‌ها، اندازهٔ واگرایی کولبک–لیب‌لر که در بخش ٤ معرفی شد، بر اساس نمونه‌های پسینی پارامترهای مدل قدم‌زدن تصادفی با توزیع tی استیودنت دومتغیره محاسبه شد. شکل ۸، نمودار اندازهٔ واگرایی کولبک–لیب‌لر را برای همهٔ نمونه‌های دومتغیرهٔ دو رشتهٔ بیمه نشان می‌دهد. توجه به میزان پرداختهای خسارات در دوره‌های قبل و بعد در سطرها و ستونهای دو مثلث، اندازه‌های کولبک–لیب‌لر بالای نمونه‌های مؤثر شناسایی‌شده در شکل ۸ را در مقایسهٔ با سایر نمونه‌ها بیان می‌کند. همچنین یک طرح اغتشاش در مشاهدات را با $y_{57}^{(1)}+\delta$ که در آن $\delta=۲$ است، در نظر گرفتیم. همان‌گونه که در شکل ۸ مشاهده می‌شود، این مشاهدهٔ جدید به‌عنوان یک مشاهدهٔ پرت شناسایی شده است.



شکل ۷. اندازهٔ واگرایی کولبک– لیب‌لر برای مدل قدم‌زدن تصادفی با توزیع t ی استیودنت دومتغیره

٦. نتیجه‌گیری

در این مقاله با استفاده از استنباط بیزی، سه نوع مدل ANOVA، ANCOVA، و قدم‌زدن تصادفی با توزیعهای نرمال، توزیع پی‌یرسون نوع هفتم و tی استیودنت دومتغیره برای لگاریتم پرداختهای خسارات نرمال‌شده را برای داده‌های ذخیره‌سازی خسارات دومتغیرهٔ بیمهٔ بدنه و بیمهٔ شخص ثالث اتومبیل یک شرکت بیمهٔ ایرانی مورد بررسی قرار

دادیم. نتایج نشان می‌دهد که مدلهای با توزیعهای دُم‌سنگین عملکرد بهتری نسبت به توزیع نرمال دارند.

همچنین با استفاده از مدل نردبان زنجیره‌ای در دو مثلث به طور جداگانه و بدون درنظرگرفتن وابستگیها، برآورد ذخایر را به دست آوردیم؛ نتایج نشان می‌دهد که مدل نردبان زنجیره‌ای مقدار ذخایر را کمتر از روشهای پیشنهادشده برآورد می‌کند و انحراف استاندارد برآوردهای به‌دست‌آمده بر اساس این مدل نیز از انحراف استاندارد برآوردهای به‌دست‌آمده بر اساس مدلهای آنالیز کوواریانس و قدم‌زدن تصادفی بزرگتر است. به منظور ارزیابی دقت پیش‌بینی مدلها، میانگین اریبی نسبی مقادیر خسارات پرداخت‌شده در فصل ۱۶ام، با دو مدل قدم‌زدن تصادفی با توزیع t ی استیودنت دومتغیره و مدل نردبان زنجیره‌ای محاسبه شد؛ نتایج نشان داد که در هر دو مثلث عملکرد مدل قدم‌زدن تصادفی با توزیع t ی استیودنت دومتغیره از مدل نردبان زنجیره‌ای بهتر بوده است؛ بنابراین درنظرگرفتن وابستگیهای بین مثلثهای تأخیر در محاسبهٔ ذخایر می‌تواند به نتایج دقیق‌تری منجر شود. به‌علاوه، بر اساس واگرایی کولبک–لیب‌لر، نمونه‌های مؤثر بر توزیعهای پسینی توأم شناسایی شدند.

بنابراین با توجه به اینکه بر اساس آیین‌نامهٔ شمارهٔ ۵۸ بیمهٔ مرکزی، روش محاسبهٔ ذخایر خسارات معوق در شرکتهای بیمه در ایران محدود و فقط با درنظرگرفتن یک مثلث تأخیر است، بر اساس نتایج این تحقیق پیشنهاد می‌شود که به منظور واردکردن وابستگیهای بین مثلثهای تأخیر در محاسبهٔ ذخایر از روشهای بیزی استوار برای مدل‌بندی توأم خسارات معوق مربوط به رشته‌های مختلف بیمه‌ای استفاده شود. همچنین پیشنهاد می‌شود که پس از برازش مدلها، یک مطالعهٔ آنالیز حساسیت بر اساس واگرایی کولبک–لیب‌لربه منظور شناسایی نمونه‌های مؤثر انجام شود. روشها و معیارهای بیزی به‌آسانی با نرم‌افزاری استاندارد بیزی مانند OpenBUGS محاسبه می‌شوند.

منابع

۱. پژوهشکدهٔ بیمه، ۱۳۹۵. بازنگری آیین‌نامهٔ ذخایر فنی مؤسسات بیمه وارائهٔ پیشنهادهای اصلاحی.

2. Braun, C., 2004. The prediction error of the chain ladder method applied to correlated run-off triangles. *ASTIN Bulletin,* 34 (2), pp. 399-424.

3. Chen, M., Shao, Q. and Ibrahim, J., 2000. *Monte Carlo Methods in Bayesian Computation*. New York: Springer-Verlag.

4. Cho, H., Ibrahim J.G., Sinha, D. and Zhu, H., 2009. Bayesian case influence diagnostics for survival models. *Biometrics*, 65, pp.116-124.

5. Cook, R., 1986. Assessment of local influence. *Journal of the Royal Statistical Society: Series B (Methodological),*48, pp.133-169.

6. Côté, M.P., Genest, C. and Abdallah, A., 2016. Rank-based methods for modeling dependence between loss triangles. *European Actuarial Journal*, 6(2), pp. 377-408.

7. De Jong, P., 2012. Modeling dependence between loss triangles using copula. *North American Actuarial Journal*, 16(1), pp. 74-86.

8. England, P.D. and Verrall, R.J., 2002. Stochastic claims reserving in general insurance. *British Actuarial Journal,* 8(03), pp. 443-518.

9. England, P.D. and Verrall, R.J., 2006. Predictive distributions of outstanding liabilities in general insurance. *Annals of Actuarial Science*, 1, pp. 221-270.

10. Fernandez, C. and Steel, M.F. 2000. Bayesian regression analysis with scale mixtures of normals. *Econometric Theory*, *16*(01), pp. 80-101.

11. Gelfand, A., Dey, D. and Chang, H., 1992.Model determination using predictive distributions with implementation via sampling-based methods. *Bayesian Statistics,* 4, pp. 147-167.

12. Happ, S. and Wüthrich, M.V., 2013. Paid-incurred chain reserving method with dependence modeling. *ASTIN Bulletin*, 43(1), pp. 1-20.

13. Hess, K., Schmidt, K., and Zocher, M., 2006. Multivariate loss prediction in the multivariate additive model. *Insurance: Mathematics and Economics*, 39(2), pp. 185-191.

14. Merz, M., and Wüthrich, M.V., 2009a. Combining chain-ladder and additive loss reserving method for dependent lines of business. *Variance*, 3(2), pp. 270-291.

15. Merz, M., and Wüthrich, M.V., 2009b. Prediction error of the multivariate additive loss reserving method for dependent lines of business. *Variance*, 3(1), pp. 131-151.

16. Merz, M., and Wüthrich, M.V., and Hashorva, E., 2013. Dependence modelingin multivariate claims run-off triangles. *Annals of Actuarial Science*, 7(1), pp. 3-25.

17. Pearson, K. 1916. Mathematical contributions to the theory of evolution. XIX. Second supplement to a memoir on skew variation. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 216, pp.429-457.

18. Salzmann, R., and Wüthrich, M.V., 2012. Modeling accounting year dependence in run-off triangles. *European Actuarial Journal*, 2(2), pp. 227-242.

19. Schmidt, K., 2006. Optimal and additive loss reserving for dependent lines of business. *Casualty Actuarial Society Forum*, pp. 319-351.

20. Shi, P., 2014. A copula regression for modeling multivariate loss triangles and quantifying reserving variability. *ASTIN Bulletin*, 44(01), pp. 85-102.

21. Shi, P., Basu, S. and Meyers, G., 2012. A bayesian log-normal model for multivariate loss reserving. *North American Actuarial Journal*, 16(1), pp. 29–51.

22. Shi, P., and Frees, E., 2011.Dependent loss reserving using copulas. *ASTIN Bulletin*,, 41(2), pp. 449-486.

23. Spiegelhalter, D.J., Best, N.G., Carlin, B.P., and Lindevan der, A., 2002. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society Series B*, 64, pp. 583–616.

24. Taylor, G., 2000. *Loss Reserving: An Actuarial Perspective*. Dordrecht: Kluwer Academic Publishers.

25. Verrall, R., 1991.Chain ladder and maximum likelihood. *Journal of the* Institute *of* Actuaries, 118, pp. 489–499.

26. Wüthrich, M.V., and Merz, M. 2008. *Stochastic Claims Reserving in Insurance*. John Wiley Sons.

27. Zhang, Y., 2010. A general multivariate chain ladder model. *Insurance: Mathematics and Economics*,46, pp. 588-599.

28. Zhang, Y., Dukic, V. and Guszcza, J., 2012.A Bayesian non-linear model for forecasting insurance loss payments. *Journal of the Royal Statistical Society: Series A*, 175(2), pp. 637-656.