فصلنامه صنعت بیمه / سال بیست‌ویکم /شماره ۳/ پاییز ۱۳۸۵ /شماره مسلسل ۸۳/صفحات ۵۱ تا ۶۵

# توزیع پواسون تعمیم‌یافته و کاربرد آن در صنعت بیمه

عاطفه ضیغمی [۱]

## چکیده

اغلب در مسائل بیمه‌ای، از جمله محاسبهٔ حق بیمه، توزیع تعداد پیشامدها یا حوادث رخ داده در دوره‌ای از زمان یا مکان، مورد نیاز است. توزیع‌های متعددی از جمله پواسون، دوجمله‌ای، دوجمله‌ای منفی و غیره در این زمینه به‌کار می‌روند، اما یکی از توزیع‌های بسیار مناسب برای محاسبهٔ تعداد پیشامدها، توزیع پواسون تعمیم‌یافته است. در این مقاله توزیع پواسون تعمیم‌یافته معرفی و برخی از خصوصیات آماری آن بیان می‌شود و همچنین کاربردی ازآن در صنعت بیمه نشان داده خواهد شد.



## واژگان کلیدی

توزیع پواسون تعمیم‌یافته، بیش پراکندگی،کم پراکندگی، حق بیمه

## مقدمه

یکی از پرکاربردترین مدل‌های احتمال برای متغیرهای تصادفی صحیح مقدار، مدل پواسون است . این مدل، زمانی کاربرد دارد که میانگین و واریانس متغیر مورد نظر،

[۱]. مربی گروه اقتصاد بهداشت، دانشکده مدیریت واطلاع رسانی، دانشگاه علوم پزشکی ایران

یعنی تعداد پیشامدها درواحد زمان یا مکان ثابت و برابر باشند.اما اغلب الگوی داده‌های مشاهده شده چنین خصوصیتی ندارند و مدل پواسون برای آنها مناسب نیست. برای این‌گونه داده‌ها از توزیع دیگری استفاده می‌شود که توزیع پواسون تعمیم‌یافته نام دارد. توزیع پواسون تعمیم‌یافته علاوه بر سادگی ، به دلیل داشتن دو پارامتر $\lambda$ و $\theta$، انعطاف‌پذیر است. واریانس این مدل بسته به آن که پارامتر دوم، یعنی $\theta$، مثبت، صفر و یا منفی باشد، بزرگ‌تر، برابر و یا کمتر از میانگین است.درنتیجه دو ویژگی مهم از جمله بیش پراکندگی (بزرگ‌تر بودن واریانس از میانگین) و کم پراکندگی (کوچک‌تر بودن واریانس از میانگین) دارد. با توجه به داشتن چنین خصوصیاتی، توزیع پواسون تعمیم یافته می‌تواند برازش‌های خوبی برای انواع داده‌های مشاهده شده که عمدتاً از نوع دوجمله‌ای، پواسون و دوجمله‌ای منفی هستند، فراهم سازد.

به دلیل ماهیت انعطاف‌پذیری مدل پواسون تعمیم‌یافته کارهای فراوانی روی آن صورت گرفته وبرای مدل‌سازی پدیده‌های مختلف مانندمسائل بیمه‌ای به‌کاررفته است. برای مثال در بیمه اتوموبیل، معمولاً تصادفات خوشه ای که در بزرگراه‌ها اتفاق می‌افتد و دریک تصادف چندین خودرو آسیب می بینند و یا در بیمه آتش‌سوزی تعداد خانه‌ها یا محل‌هایی که به صورت خوشه‌ای آتش می‌گیرند و همچنین در بیمه از کارافتادگی بر اثر حادثه‌هایی که در کارگاه‌ها رخ می‌دهد ممکن است چندین نفر با هم آسیب ببینند، همه را می‌توان با کمک توزیع پواسون تعمیم یافته مدل‌بندی کرد. بنینگ و کورولف (۲۰۰۲)؛ کنسول (۱۹۸۳ و ۱۹۸۹).

## تعریف توزیع پواسون تعمیم‌یافته

متغیرتصادفی صحیح مقدار $X$ دارای توزیع پواسون تعمیم‌یافته $(GPD)$ [۲] است، اگر تابع جرم احتمال آن به صورت

$$P(X=x)=P_x(\lambda,\theta)=\lambda(\lambda+x\theta)^{x-1}\frac{e^{-(\lambda+x\theta)}}{x!},\quad x=0,1,2,... \qquad (1)$$

باشد که در آن $\lambda > 0$ و $\max(-1,-\lambda/m) < \theta < 1$، با توجه به این که $\lambda + x\theta$ باید همواره مثبت باشد، برای $\theta < 0$، $m$ را بزرگ‌ترین عدد صحیح مثبتی در نظر می‌گیریم که در شرط $\lambda + m\theta > 0$ صدق می‌کند. درنتیجه برای $x > m$، زمانی‌که $\theta < 0$، مقدار تابع جرم احتمال برابر صفر است.

اگر به رابطهٔ (۱) به دقت نگاه کنیم، متوجه می‌شویم که گویی در توزیع پواسون

$$P(X=x)=\frac{\eta^x e^{-\eta}}{x!}$$

به‌جای پارامتر $\eta$ از پارامتر $\lambda + x\theta > 0$ استفاده شده است. البته این حکم دقیقاً درست نیست ولی تقریباً چنین است.معنای این تعبیرآن است که پارامتر توزیع پواسون تعمیم‌یافته به ازای مقدار ثابتی از $\lambda$ به $x$، تعداد رخدادها، وابسته است. این وابستگی با پارامتر $\theta$ تعیین می‌شود.اگر $\theta > 0$، پارامتر $(\lambda + x\theta)$به‌ازای $x$های بزرگ مقدار بزرگ‌تری را اختیار می‌کند، یعنی احتمال وقوع مقادیر بزرگ $x$ مانند توزیع پواسون عادی به سرعت کاهش نمی‌یابد و همین امر باعث بزرگ‌ترشدن واریانس از میانگین

2. Generalized Poisson Distribution

می‌شود. اما اگر $\theta < 0$، ملاحظه کردیم که برای $x > m$، $P_x(\lambda,\theta) = 0$، یعنی احتمال وقوع $x$های بزرگ‌تر از $m$ را باید صفر درنظرگرفت، در غیر این‌صورت احتمال مقادیر نظیر، بزرگ‌تر از یک خواهد شد. بنابراین به علت محدودشدن دامنهٔ تغییرات متغیر تصادفی $X$، واریانس کوچک‌تر از میانگین خواهد شد.

بدین ترتیب تعمیمی از توزیع پواسون به‌دست آمده که مناسب مسائلی است که در آنها *بیش‌پراکندگی*(بزرگ‌تر بودن واریانس از میانگین) یا *کم‌پراکندگی*(کوچک‌تر بودن واریانس از میانگین) داریم.

میانگین و واریانس این مدل را به‌راحتی می‌توان با استفاده ازبسط لاگرانژ وشرط $\sum_{x=0}^{\infty} P_x(\lambda,\theta) = 1$ به‌دست آورد که به‌صورت زیر است.

$$\mu(\lambda,\theta) = \lambda(1-\theta)^{-1} \quad (٢)$$

$$\sigma^2(\lambda,\theta) = \lambda(1-\theta)^{-3} \quad (٣)$$

برای ملاحظه جزئیات محاسبه نگاه کنید به: ضیغمی (١٣٨٤)؛ بولمن (١٩٧٠)؛ کنسول و شنتون (١٩٧٢).

## برآورد مدل

اغلب در مسائل آماری هدف نهایی و مسئلهٔ مهم، استنباط از پارامترهای نامعلوم جامعه با استفاده از اطلاعات به‌دست آمده از نمونه است. این استنباط می‌تواند به‌صورت برآورد نقطه‌ای و بازه‌ای یا در چارچوب آزمون فرض‌های آماری باشد که در مورد پارامترهای جامعه یا توابعی از آنها بیان می‌شوند.

برآورد نقطه‌ای پارامترهای توزیع پواسون تعمیم‌یافته را با استفاده از روش ماکزیمم درست‌نمایی می‌توان به‌صورت زیر به‌دست آورد:

اگر $x_1$، $x_2$،...، $x_n$ نمونه‌ای تصادفی به اندازهٔ $n$ از توزیع پواسون تعمیم‌یافتهٔ تعریف شده (۱) و $n_x$ به‌ازای $x = 0,1,...,k$، فراوانی‌های مشاهده شده برای $k+1$ طبقهٔ مختلف در این نمونه تصادفی باشد، به‌طوری که $\sum_{x=0}^{k} n_x = n$ ، آن‌گاه تابع درست-نمایی به‌صورت زیر است.

$$
\begin{aligned}
L &= \prod_{x=0}^{k}[P_x(\lambda,\theta)]^{n_x} \\
&= \prod_{x=0}^{k}[\lambda^{n_x}(\lambda + x\theta)^{n_x(x-1)} e^{-(\lambda+x\theta)n_x} \Big/ (x!)^{n_x}] \qquad (٤) \\
&= \lambda^n e^{-\sum_{x=0}^{k} n_x(\lambda+x\theta)} [\prod_{x=0}^{k}(\lambda + x\theta)^{n_x(x-1)}] \Big/ \prod_{x=0}^{k}(x!)^{n_x}
\end{aligned}
$$

پس از ماکزیمم کردن این تابع، نقطهٔ اکسترمم تابع درست‌نمایی به‌صورت زیر در می‌آید:

$$\hat{\lambda} = \bar{x}(1-\hat{\theta}) \qquad (٥)$$

$$H(\hat{\theta}) = \sum_{x=0}^{k} n_x \frac{x(x-1)}{\bar{x} + (x-\bar{x})\hat{\theta}} - n\bar{x} = 0 \qquad (٦)$$

از حل دومعادله دومجهولی بالا مقادیر $\hat{\lambda}$ و $\hat{\theta}$ به دست می آیند که مستلزم محاسبات عددی به وسیله کامپیوتر است.

## مسئله آزمون فرض مدل

### به‌طوریکنواخت تواناترین آزمون برای یکی از پارامترها زمانی که دیگری معلوم و اندازهٔ نمونه بزرگ باشد.

کنسول و شنتون (۱۹۷۲) نشان دادندکه اگر متغیرتصادفی $X$ دارای توزیع پواسون تعمیم‌یافته با پارامترهای $(\lambda,\theta)$ باشد و اگر $\theta < 0.5$، متغیر استاندارد شدهٔ $Z = \dfrac{X-\mu}{\sigma}$ با افزایش $\lambda$ به سمت توزیع نرمال استاندارد میل می‌کند .توجه کنید که این تقریب به اندازهٔ نمونه بستگی ندارد.

در توزیع پواسون تعمیم یافته برای $\lambda$ های نسبتاً بزرگ $(\lambda = 8)$و $0 < \theta < 0.5$،شکل توزیع تقریباً متقارن و شبیه به توزیع نرمال است. در نتیجه آزمون برای فرض‌های زیر می‌تواند بر پایه تقریب نرمال باشد، حتی اگر اندازهٔ نمونه کوچک باشد:

$$H_0 : \theta = \theta_0 < 0.5$$
$$H_1 : \theta > \theta_0$$

در این صورت آمارهٔ آزمون از یک نمونهٔ تصادفی به اندازهٔ $n$ ، برابر $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} x_i$ است.ناحیهٔ بحرانی برای رد فرض صفر، باتوجه به فرض مقابل، در سطح معنی‌داری $\alpha$ برابر $\overline{X} > C$ است که $C$ از رابطهٔ زیر به‌دست می‌آید:

$$\alpha = P(\overline{X} > C | H_0)$$

برای آزمون فرض داریم؛

$$P(Z > Z_\alpha) \quad P[Z > \frac{\sqrt{n}(C - \lambda_0(1-\theta_0)^{-1})}{\sqrt{\lambda_0(1-\theta_0)^{-3}}}] = \tag{۷}$$

$$C = \lambda_0(1-\theta_0)^{-1} + Z_\alpha \sqrt{\frac{\lambda_0(1-\theta_0)^{-3}}{n}} \tag{۸}$$

اگر $\overline{X} > C$ فرض صفر پذیرفته نمی‌شود.

به طورمشابه، در صورتی که به آزمون $\begin{matrix} H_0 : \theta = \theta_0 < 0.5 \\ H_1 : \theta < \theta_0 \end{matrix}$ علاقه مند باشیم،

ناحیه‌بحرانی $\overline{X} < C$ خواهد بود که $c$ از رابطه $\alpha = P(\overline{X} < C | H_0)$ به‌دست می آید.

آزمون فرض برای پارامتر $\lambda$ زمانی‌که $\theta$ معلوم است نیز با روشی مشابه فوق به‌دست می‌آید.

## آزمون توزیع پواسون تعمیم‌یافته در برابر توزیع پواسون

می‌دانیم که همواره باید مدلی را برای برازش داده‌ها انتخاب کنیم که علاوه برداشتن پارامتر کمتر، بیشترین و دقیق‌ترین اطلاعات را در مورد جامعهٔ تحت بررسی در اختیار ما قرار دهد. از آنجا که توزیع پواسون تعمیم‌یافته دارای پارامتری بیشتر و پیچیده‌تر از توزیع پواسون است، اغلب انتخاب و استفاده از این توزیع جز در موارد ضروری عاقلانه نیست. در نتیجه ابتدا با استفاده از آزمون‌های آماری آزمون می‌کنیم که توزیع پواسون تعمیم‌یافته برای داده‌های مورد نظر مناسب است یا توزیع پواسون. برای این کار از دو آزمون زیر استفاده می شود.

**الف) آزمون واریانس:** آزمونی است که برای آزمون پواسون بودن توزیع، بسیار رواج دارد. آمارهٔ آزمون برای آزمون فرض $H_0: \theta = 0$ در برابر $H_1: \theta \neq 0$ به‌صورت

$$VT = \sum_{i=1}^{n} \frac{(X_i - \overline{X})^2}{\overline{X}} = (n-1)\frac{S^2}{\overline{X}} \qquad (۹)$$

است. این آماره آزمون مبتنی بر این فرض است که در صورت پواسون بودن توزیع $\overline{X}$ و $S^2$ هر دو برآوردگرهای $\lambda$ هستند و انتظار می رود که تقریباً مساوی باشند. در غیر این صورت بسته به بیش پراکندگی یا کم پراکندگی $S^2$ از $\overline{X}$ بزرگ‌تر یا کوچک‌تر خواهد بود.

زمانی که فرض صفر درست باشد، آزمون واریانس تقریباً دارای توزیع خی‌دو با درجهٔ آزادی $n-1$ است.در نتیجه ناحیهٔ بحرانی برای رد فرض صفر در آزمون واریانس برابر $VT > \chi^2_{\alpha,n-1}$ است که اگر فرض صفر رد شود، پذیرش توزیع پواسون تعمیم‌یافته برای داده‌های تحت مطالعه را نشان می‌دهد.

**ب)آزمون پاتهوف[3] و ویتینگهیل[4]:** همان‌طور که می دانیم ویژگی اصلی توزیع پواسون، برابر بودن میانگین و واریانس است. درحالی که در توزیع پواسون تعمیم یافته چنین خاصیتی وجود ندارد. در نتیجه می‌توانیم برای مقایسه این دو توزیع فرض زیر را بیان کنیم.

$$\begin{aligned} &H_0: \sigma^2 = \mu \\ &H_1: \sigma^2 \neq \mu \end{aligned} \qquad (۱۰)$$

برای آزمون این فرض پاتهوف و ویتینگهیل و بوهنینگ[5] آمارهٔ آزمون

3. Potthoff
4. Whittinghill
5. Bohning

$$O_2 = \sqrt{\frac{n-1}{2}}\left(\frac{S^2-\overline{X}}{\overline{X}}\right) = \frac{\sum(X_i-\overline{X})^2}{\overline{X}\sqrt{2(n-1)}} - \sqrt{\frac{n-1}{2}} \qquad (11)$$

را معرفی کردند. تحت فرض صفر این آماره دارای توزیع نرمال استاندارد است.ناحیهٔ بحرانی برای رد فرض صفر به‌وسیلهٔ آمارهٔ $O_2$ ،برابر $O_2 > z_\alpha$است که $z_\alpha$ حد بالای توزیع نرمال استاندارد در سطح $\alpha$ است.

## کاربرد مدل در صنعت بیمه

می‌دانیم که محاسبهٔ حق بیمه بر پایهٔ این فرض استوار است که بتوان مقادیر ادعاهای خسارت احتمالی را با پرداخت‌های ثابت (حق بیمه) جبران کرد. برای محاسبهٔ حق بیمه لازم است که توزیع مقدار ادعای انباشته در یک دورهٔ زمانی معین( معمولاً یک‌سال) را داشته باشیم. بدین منظور باید توزیع تعداد ادعاها در همان دوره و مبالغ ادعاهای خسارت را بدانیم. مثلاً اگر بخواهیم در یک نمایندگی بیمه، حق بیمه دریافتی ازافرادی را که دچار سانحه اتوموبیل شده‌اند محاسبه کنیم لازم است که ابتدا توزیع تعداد افراد را که در بازهٔ زمانی معین به این مرکزمراجعه کرده‌اند بیابیم و سپس براساس میزان خسارت و هزینه‌های مربوط، مدل مناسب را برازش دهیم. اگر توزیع تعداد ادعاها در یک دورهٔ معین، پواسون باشد، توزیع مبالغ ادعای انباشته در آن دورهٔ زمانی از توزیع پواسون مرکب تبعیت می‌کند. برای مثال در بیمه، اگر X میزان خسارت ( یک متغیر پیوسته و از توزیعی مثل گاما یا لوگ‌نرمال پیروی کند) وارد شده بر یک ادعا و N تعداد ادعاها ( یک متغیر گسسته و از توزیع پواسون تبعیت می کند ) باشد، در این حالت مجموع خسارت‌های وارد شده بر شرکت بیمه ( $Y = \sum_{i=1}^{N} X_i$ ) از توزیع پواسون مرکب پیروی می‌کنند. (آمباگاس پیتیا و بالاکریشنان (۱۹۹٤)).

حال در این قسمت با مثالی نشان می‌دهیم که در مواردی که تعداد ادعاها ($N$) از توزیع پواسون عادی تبعیت نمی‌کنند،می‌توانیم توزیع پواسون تعمیم‌یافته را به‌کار بریم و از مدل *پواسون مرکب تعمیم‌یافته* برای توزیع مبالغ ادعای انباشته در بازهٔ زمانی معین و همچنین محاسبهٔ‌حق بیمه استفاده کنیم. هدف آن است که یکی از کاربردهای مدل پواسون تعمیم‌یافته را در صنعت بیمه، در مواردی که مدل پواسون عادی به داده‌ها نمی‌برازد، نشان دهیم.

## معرفی داده‌ها

داده‌هایی که در این مقاله مورد استفاده واقع می‌شوند، مربوط به تعداد افرادی است که به علت سانحه اتوموبیل به یکی از نمایندگی‌های بیمه مراجعه کرده‌اند. این داده‌ها در مدت یک هفته ثبت شده‌اند. این داده‌ها، در جدول ۱ به‌صورت یک توزیع فراوانی ارائه شده‌اند که در آن متغیر تصادفی مورد نظر، تعداد مراجعه کنندگان در هرساعت است. ملاحظه می‌شودکه تعداد ساعاتی که هیچ مراجعه‌کننده نداشته ۷۸ مورد وتعداد ساعاتی که ۱،۲،۳و٤ مراجعه‌کننده داشته است ، به‌ترتیب ٤۳، ٢٦، ۱۷و۸ مورد است.

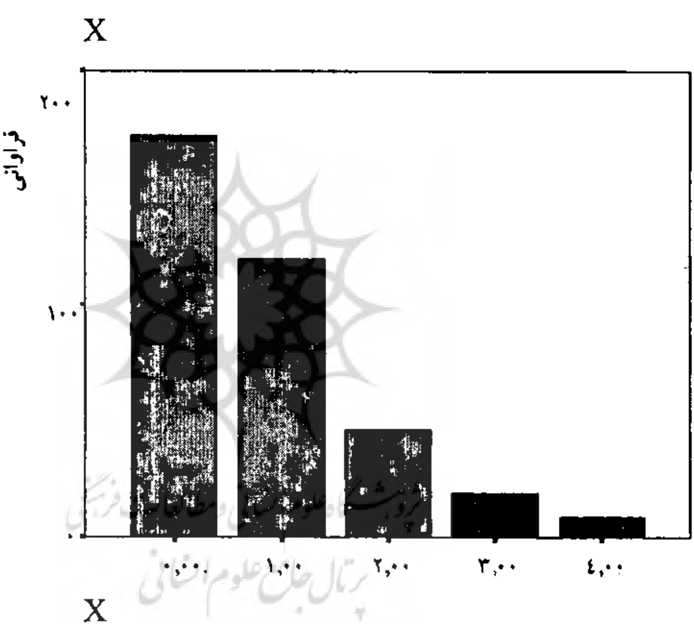
جدول ۱. توزیع فراوانی

| تعداد مراجعه‌کنندگان | فراوانی مشاهده شده |
|---|---|
| ۰ | ۷۸ |
| ۱ | ٤۳ |
| ۲ | ٢٦ |
| ۳ | ۱۷ |
| ٤ | ۸ |
| جمع | ۱۷۲ |

## برازش مدل

با توجه به نمودار فراوانی داده‌ها در شکل۱ و $L$ - شکل بودن نمودار این داده‌ها، می‌توان حدس زد که مدل پواسون برای برازش به این داده‌ها چندان مناسب نیست و توزیع پواسون تعمیم‌یافته برای مدل‌بندی آنها مناسب‌تر است.

**شکل ۱. نمودار توزیع فراوانی مراجعه‌کنندگان در نمونه**




حال می‌خواهیم که حدس خود را بیازماییم وتوزیع مناسبی برای برازش به این داده‌ها بیابیم. آزمونی را که بدین منظور به‌کار می‌بریم ، آزمون نیکویی برازش خی‌دو است که با استفاده از نرم‌افزار $SPSS$ برای دو فرض توزیع پواسون و توزیع پواسون تعمیم‌یافته انجام داده‌ایم.

**الف) آزمون نیکویی برازش مدل پواسون عادی به داده‌ها:** اگر آزمون نیکویی برازش را برای فرض

$H_0$ : داده‌ها، نمونه‌ای از توزیع پواسون عادی‌اند

$H_1$ : داده‌ها، نمونه‌ای از توزیع پواسون عادی نیستند

انجام دهیم، با توجه به خروجی برنامه کامپیوتری ملاحظه می‌کنیم که مقدار احتمال (۰/۰۲۶) از سطح معنی‌داری $\alpha =$ ۰/۰۵ کمتر است و همچنین مقدار آمارهٔ این آزمون با فرض توزیع پواسون ($\chi^2 =$ ۹/۸۶۲) از توزیع خی‌دو با ۳ (=۱-۱-۵) درجهٔ آزادی در سطح معنی‌داری $\alpha =$ ۰/۰۵ (۷/۸۱۵) بزرگ‌تر است، درنتیجه فرض پواسون پذیرفته نیست و درحقیقت تعداد مراجعه‌کنندگان به این مرکز از توزیع پواسون عادی تبعیت نمی‌کند.

جدول ۲. محاسبات مربوط به آزمون نیکویی برازش پواسون عادی

| | |
|---|---|
| آمارهٔ خی دو | ۹/۸۶۲ |
| درجهٔ آزادی | ۳ |
| مقدار احتمال | ۰/۰۲۶ |

**ب) آزمون نیکویی برازش مدل پواسون تعمیم‌یافته به داده‌ها:** حال اگر آزمون نیکویی برازش را برای فرض

$H_0$ : داده‌ها، نمونه‌ای از توزیع پواسون تعمیم‌یافته‌اند

$H_1$ : داده‌ها، نمونه‌ای از توزیع پواسون تعمیم‌یافته نیستند

انجام دهیم،ملاحظه می‌کنیم که آمارهٔ این آزمون ($\chi^2 =$ ۴/۲۶۳) از مقدار توزیع خی‌دو با ۲ (=۱-۲- ۵) درجه آزادی در سطح معنی‌داری $\alpha =$ ۰/۰۵ (۵/۹۹۱) کمتر است و همچنین ۰/۰۵ < ۰/۲۱۴= مقدار احتمال، درنتیجه فرض توزیع پواسون تعمیم‌یافته برای مدل‌بندی کردن تعداد مراجعه‌کنندگان به این مرکز در طی یک هفته پذیرفته می‌شود.

جدول ۳. محاسبات آزمون نیکویی برازش پواسون تعمیم یافته

| | |
|---|---|
| آمارهٔ خی دو | ٤/٢٦٣ |
| درجهٔ آزادی | ٢ |
| مقدار احتمال | ٠/٢١٤ |

## برآورد پارامترهای مدل

اکنون که از پیروی داده‌ها از توزیع پواسون تعمیم‌یافته اطمینان یافته‌ایم، می‌خواهیم که پارامترهای توزیع پواسون تعمیم‌یافته را با توجه به این داده‌ها برآورد کنیم و سپس بازه‌های اطمینان برای پارامترهای این توزیع را به‌دست آوریم و فرض‌های معمول را آزمون کنیم.

برآورد ماکزیمم درست‌نمایی پارامترهای اول و دوم توزیع پواسون تعمیم‌یافته با توجه به این داده‌ها عبارت‌اند از:

$$\hat{\theta} = 0/06234 \qquad \hat{\lambda} = 0/564$$

در نتیجه مدل مناسب برای برازش به این داده‌ها عبارت است از:

$$P(X=x) = (0/564)\{[(0/564)+x(0/06234)]\hat{}(x-1)\} * \exp\{-[(0/564)+x(0/06234)]\}/x!$$

اکنون که مدل پواسون تعمیم‌یافته مناسب برای تعداد مراجعه‌کنندگان به این مرکز را برازش دادیم، می‌توانیم توزیع مبلغ ادعای تجمعی را بیابیم و با کمک آن حق بیمهٔ پرداختی هر فرد را محاسبه کنیم.

امباگاس‌پیتیا[6] و بالاکریشنان (١٩٩٤) نشان دادند که اگر $S = Y_1 + Y_2 + \ldots + Y_N$ مبلغ ادعای خسارت تجمعی و $N$ (تعداد ادعاها در یک دورهٔ زمانی معین) دارای

6. Ambagaspitiya

توزیع پواسون تعمیم‌یافته باشد،آن‌گاه $S$ دارای توزیع پواسون مرکب تعمیم‌یافته به‌صورت

$$F_S(y)=\sum_{x=0}^{\infty}F^{*x}(y)\lambda(\lambda+x\theta)^{x-1}\frac{\exp\{-(\lambda+x\theta)\}}{x!}$$

است. همچنین داریم،

$$E(S)=(1-\theta)^{-1}E(Y)$$
$$Var(S)=(1-\theta)^{-1}V[Y]+\lambda(1-\theta)^{-3}E[Y^2]$$

بر طبق اصول محاسبه حق بیمه از جمله اصل امید ریاضی (بولمن (۱۹۷۰)) می‌توانیم حق بیمه پرداخت شده هرفرد را محاسبه کنیم. اگر مقدار حق بیمه را با $p$ نشان دهیم داریم،

$$\begin{aligned}p&=(1+\alpha)E(S)\\&=(1+\alpha)\lambda(1-\theta)^{-1}E(Y)\end{aligned}$$

اکنون با استفاده از مدل برازانده شده برای تعداد مراجعه‌کنندگان حق بیمه دریافتی از هر فرد بر طبق اصل امید ریاضی به‌وسیلهٔ رابطهٔ زیر به‌دست می‌آید:

$$p=(0/87)(1+\alpha)E(Y)$$

که در آن $E(Y)$ با توجه به توزیع مبالغ ادعاها به‌دست می‌آید.

## نتیجه گیری

در بسیاری از موارد کاربردی که با شمارش نوعی پدیده در واحد زمان یا مکان سرو کار داریم ، معمولاً از توزیع پواسون استفاده می شود. اما در موارد زیادی مشاهده شده است که توزیع پواسون برازش خوبی به داده‌ها ندارد، زیرا داده‌ها دارای بیش

پراکندگی یاکم پراکندگی‌اند. در این موارد توزیع پواسون تعمیم یافته برازش بهتری را فراهم می‌سازد. در این مقاله توزیع پواسون تعمیم یافته بررسی و با ذکر مثالی کاربرد آن نشان داده شد. در ادامه این پژوهش می‌توان به استنباط بیزی پارامترها و مسائل رگرسیونی درتوزیع پواسون تعمیم یافته اشاره‌کرد که می‌توانند موضوع پژوهش‌های آتی باشند.

## منابع


۱. ضیغمی، عاطفه (۱۳۸٤)، *توزیع پواسون تعمیم یافته و کاربرد آن در صنعت بیمه*، پایان نامه کارشناسی ارشد بیمه آمار، دانشگاه شهید بهشتی.

2. Ambagaspitiya,R.S. and Balakrishnan,N. (1994). "On the Compound Generalized Poisson Distribution ." *Astin Bulletin* , 24(2) , 255-263.

3. Bening,V.E. and Korolev,V.Yu. (2002). "Generalized Poisson Models and their Applications in Insurance and Finance." *Modern Probability and Statistics.*

4.Bohning,D. (1994). "A Note on a Test for Poisson Overdispersion." Biometrika, 81,418-419.

5. Buhlmann,H. (1970). "Mathematical Methods in Risk Theory."*Springer-Verlag*. New York.

6. Consul,P.C. (1983). "Lagrange Expansions." in *Encyclopedia of Statistical Science* ,vol.4, 454-456, Kotz,S. , Johnson,N. (editors).Wiley, New York.

7. Consul,P.C. (1989). "Generalized Poisson Distributions." *Properties and Application*, Marcel Decker,New York.

8. Consul,P.C. and Jain,G.C. (1973). "A Generalization of the Poisson Distribution." *Technometrics*, 15(4),791-799.

9. Consul,P.C. and Shenton,L.R. (1972). "Use of Lagrangian expansion for generating discrete generalized probability distribution." *SIAM, Journal of Appl.Math.* , 23(2), 239-248.