محاسبهٔ دقیق بیشینهٔ احتمالات پوشش بازه‌های اطمینان با کران‌های صعودی برای میانگین توزیع پوآسون

علی‌رضا شیروانی [۱]

تاریخ دریافت: ۱۳۹۳/۱۲/۱۷

تاریخ پذیرش: ۱۳۹۵/۷/۱۷

چکیده:

توزیع پوآسون یک مدل استاندارد برای تحلیل داده‌های شمارشی است و برآورد پارامتر میانگین این توزیع در عمل کاربرد زیادی دارد. تا به حال بازه‌های اطمینان متعددی برای میانگین توزیع پوآسون ارائه‌شده که همگی مجانبی هستند و مقایسهٔ دقیق آن‌ها اهمیت دارد. احتمال پوشش بازهٔ اطمینان $(L(X), U(X))$ برای میانگین متغیر تصادفی پوآسون X با پارامتر نامعلوم θ، تابعی نسبت به θ است. از آن‌جا که توزیع پوآسون گسسته است، تابع احتمال پوشش شکل بسته‌ای ندارد و با تغییر θ در فضای پارامتری تغییر می‌کند. بنا بر این محاسبهٔ بیشینه، کمینه و میانگین احتمالات پوشش به‌صورت دقیق برای بازه‌های اطمینان پارامتر θ کار بسیار مشکلی است. روش محاسبهٔ کمینه و میانگین دقیق احتمالات پوشش بازه‌های اطمینان با کران‌های صعودی برای پارامتر θ توسط وانگ [۱۱] ارائه‌شده است. در این مقاله روش محاسبهٔ دقیق بیشینهٔ احتمالات پوشش بازه‌های اطمینان با کران‌های صعودی برای پارامتر نامعلوم θ در توزیع پوآسون ارائه می‌شود. اگر تصمیم‌گیری با مقایسهٔ همزمان بازه‌های اطمینان بر اساس بیشینه، کمینه و میانگین احتمالات پوشش آن‌ها انجام شود، مطمئن‌تر خواهد بود.

واژه‌های کلیدی: توزیع پوآسون، بازهٔ اطمینان، احتمال پوشش، ضریب اطمینان، میانگین احتمال پوشش، بیشینهٔ احتمالات پوشش.

۱ مقدمه

فرض کنید X یک متغیر تصادفی پوآسون با میانگین نامعلوم θ باشد. احتمال پوشش بازهٔ اطمینان $(L(X), U(X))$ برای θ به‌کمک از احتمال این‌که بازهٔ تصادفی $(L(X), U(X))$ مقدار واقعی پارامتر θ را در بر داشته باشد. بنا بر این احتمال پوشش تابعی نسبت به θ است. ممکن است در توزیع‌های پیوسته، تابع احتمال پوشش، برای همهٔ نقاط فضای پارامتری یکسان باشد، اما در توزیع‌های گسسته، تابع احتمال پوشش با تغییر مقدار پارامتر مجهول در فضای پارامتری، تغییر می‌کند. به همین دلیل محاسبهٔ دقیق بیشینه، کمینه و میانگین احتمالات پوشش برای بازه‌های اطمینان پارامتر این توزیع، کار دشواری است. تا کنون چندین بازهٔ اطمینان مجانبی برای θ معرفی شده است. به‌عنوان مثال، خام‌کنگ [۶] بازهٔ اطمینان مجانبی AWC، گوآن [۳] بازهٔ اطمینان MSC و کابایلا و برن [۴] یک بازهٔ اطمینان دقیق و کوتاه برای پارامتر میانگین توزیع پوآسون ارائه کرده‌اند. شورتمن و مارتینز [۸] و ساهای و خورشید [۷] دقت چند بازهٔ اطمینان مجانبی ارائه‌شده برای پارامتر میانگین توزیع پوآسون را بررسی کرده‌اند. کابایلا و لوید [۵] نیز بازه‌های اطمینانی را برای پارامترهای مجهول توزیع‌های گسسته ارائه کرده‌اند. از آن‌جا که اکثر بازه‌های ارائه‌شده دارای کران‌های صعودی هستند، بازه‌های اطمینان با کران‌های صعودی اهمیت ویژه‌ای می‌یابند. بنا بر این در این‌جا بازه‌های اطمینان $(L(X), U(X))$ را به‌صورتی در نظر می‌گیریم که $L(X)$ و $U(X)$ توابعی صعودی نسبت به X باشند. وانگ [۱۱] روش محاسبهٔ دقیق ضریب اطمینان و میانگین احتمال پوشش بازه‌های اطمینان با کران‌های صعودی برای توزیع‌های گسستهٔ یک‌پارامتری را ارائه کرده است. ارزیابی بازه‌ها با معیار ضریب اطمینان خیلی

[۱] عضو هیئت علمی گروه آمار دانشگاه ولایت، ایرانشهر

محافظه‌کارانه است، زیرا اگر احتمال پوشش حتی در یک نقطه از فضای پارامتری کوچک باشد، ضریب اطمینان نیز کوچک خواهد بود. از طرفی میانگین احتمالات پوشش نیز به‌شدت تحت تأثیر کمینه و بیشینهٔ احتمالات پوشش قرار دارد. بنا بر این با در نظر گرفتن همزمان کمینه، میانگین و بیشینهٔ احتمالات پوشش، می‌توان مقایسهٔ دقیق‌تری بین بازه‌ها انجام داد. در این مقاله روش محاسبهٔ بیشینهٔ احتمالات پوشش بازه‌های اطمینان با کران‌های صعودی برای میانگین توزیع پوآسون ارائه می‌شود. با این روش می‌توانیم بیشینهٔ احتمالات پوشش و نقطه‌ای از فضای پارامتری را که بیشینهٔ احتمالات پوشش به‌ازای آن رخ می‌دهد به‌صورت دقیق به دست آوریم.

در بخش ۲ به شرایط مورد نیاز بازهٔ اطمینان و روش محاسبهٔ بیشینهٔ احتمالات پوشش دقیق می‌پردازیم. سپس در بخش ۳ چند بازهٔ اطمینان برای میانگین توزیع پوآسون معرفی کرده، آن‌ها را از نظر بیشینهٔ احتمالات پوشش دقیق مقایسه می‌کنیم. در ادامه، بخش ۴ به شبیه‌سازی اختصاص دارد. در نهایت در بخش ۵ بحث و نتیجه‌گیری آورده شده است.

۲ روش محاسبهٔ دقیق بیشینه احتمالات پوشش

فرض کنید X یک متغیر تصادفی پوآسون با پارامتر نامعلوم θ باشد. احتمال پوشش بازهٔ اطمینان $(L(X), U(X))$ برای θ به‌کمک از $P_\theta(\theta \in (L(X), U(X)))$؛ یعنی احتمال این‌که بازهٔ تصادفی $(L(X), U(X))$ مقدار واقعی θ را پوشش دهد.

در این مقاله بازه‌های اطمینانی را بررسی می‌کنیم که شرط زیر را داشته باشند:

گزارهٔ ۱.۲. برای بازهٔ اطمینان $(L(X), U(X))$ از $X_1 < X_2$، بتوان نتیجه گرفت که $L(X_1) \le L(X_2)$ و $U(X_1) \le U(X_2)$.

قبل از ارائهٔ نتیجهٔ اصلی، نمادگذاری‌ها را معرفی می‌کنیم:

فرض کنید l و u به‌ترتیب کران پایین و بالای فضای پارامتری باشند. برای یک بازهٔ اطمینان $(L(X), U(X))$، نقاط نخستیی و انتهایی متناظر با $x = 0, 1, 2, \ldots$ عبارت‌اند از:

$$L(0), L(1), L(2), \ldots$$

$$U(0), U(1), U(2), \ldots$$

این نقاط را به‌ترتیب صعودی مرتب می‌کنیم و آن‌ها را v_1 و v_2 و... نامگذاری می‌کنیم. برای $i = 1, 2, \ldots$ تعریف می‌کنیم:

$$w_i = \begin{cases} v_i, & l \le v_i \le u \\ l, & v_i < l \\ u, & v_i > u \end{cases} \qquad (1)$$

برای یک $\theta \in (v_i, v_{i+1})$ توابع $k_.$ و k_1 را به‌صورت زیر تعریف می‌کنیم:

$k_.(\theta)$: کوچک‌ترین مقدار x که برای آن، $U(x) > \theta$

$k_1(\theta)$: بزرگ‌ترین مقدار x که برای آن، $L(x) < \theta$

لم ۲.۲. اگر بازهٔ اطمینان $(L(X), U(X))$ گزارهٔ ۱.۲ را داشته باشد، خواهیم داشت:

$$P_\theta(\theta \in (L(X), U(X))) = \sum_{i=k_.(\theta)}^{k_1(\theta)} \frac{e^{-\theta}\theta^i}{i!}. \qquad (1)$$

اثبات. اگر قرار دهیم $X = k_.(\theta)$ و $Y = k_1(\theta)$، با توجه به تعریف $k_.$ و k_1 فقط چهار حالت زیر ممکن است:

حالت اول: $L(Y) < U(Y) < \theta < L(X) < U(X)$ که در این‌صورت θ پوشش داده نمی‌شود.

حالت دوم: $L(X) < L(Y) < \theta < U(X) < U(Y)$ که در این‌صورت طبق گزارهٔ ۱.۲ داریم $X \le Y$. حالت سوم: $L(X) < L(Y) < U(Y) < \theta < U(X)$.

حالت چهارم: $L(Y) < \theta < L(X) < U(X) < U(Y)$.

حالت‌های سوم و چهارم طبق گزارهٔ ۱.۲ امکان پذیر نیستند. بنا بر این یگانه حالت ممکن، حالت دوم است و این یعنی $k_.(\theta) \le k_1(\theta)$. از طرفی طبق تعریف $k_.$ و k_1 و گزارهٔ ۱.۲ داریم:

$$L(X) < \theta \Leftrightarrow X \le k_1(\theta)$$

$$U(X) > \theta \Leftrightarrow X \ge k_.(\theta)$$

بنا بر این:

$$\begin{aligned} P_\theta(\theta \in (L(X),U(X))) &= P_\theta(L(X) < \theta, U(X) > \theta) \\ &= P_\theta(X \leq k_1(\theta), X \geq k_0(\theta)) \\ &= P_\theta(k_0(\theta) \leq X \leq k_1(\theta)) \\ &= \sum_{i=k_0(\theta)}^{k_1(\theta)} \frac{e^{-\theta}\theta^i}{i!} \end{aligned}$$

و اثبات کامل است. □

تعریف ۳.۲. خانوادهٔ توابع چگالی یک‌پارامتری $\{f_\theta : \theta \in \Theta\}$ دارای خاصیت نسبت درست‌نمایی یکنوا[۲] در آمارهٔ T با تابع چگالی $g_\theta(t)$ است اگر $\frac{g_{\theta_2}(t)}{g_{\theta_1}(t)}$ به‌ازای هر $\theta_2 > \theta_1$، تابعی غیر نزولی نسبت به t باشد.

قضیهٔ ۴.۲. فرض کنید $f_\theta(x)$ یک خانواده از توابع چگالی روی $\mathbb{R}$ با خاصیت نسبت درست‌نمایی یکنوا (MLR) در x باشد. در این صورت برای هر $\theta_1 < \theta_2$، تابع توزیع تجمعی X در رابطهٔ زیر صدق می‌کند:

$$F_{\theta_2}(x) \leq F_{\theta_1}(x), \quad \forall x \in \mathbb{R}$$

اثبات. قضیهٔ ۸.۳.۱۷ در کسلا و برگر [۲] را ببینید. □

لم ۵.۲. ۱- برای یک m_1 ثابت که $m_1 > 0$، $\sum_{i=0}^{m_1} \frac{e^{-\theta}\theta^i}{i!}$ یک تابع نزولی نسبت به θ است.

۲- برای یک m_0 ثابت که $m_0 > 0$، $\sum_{i=m_0}^{+\infty} \frac{e^{-\theta}\theta^i}{i!}$ یک تابع صعودی نسبت به θ است.

۳- برای m_0 و m_1 ثابت که $0 < m_0 < m_1$، $\sum_{i=m_0}^{m_1} \frac{e^{-\theta}\theta^i}{i!}$ یک تابع تک‌مدی نسبت به θ است و بیشینهٔ این تابع در نقطهٔ $\theta = h(m_0, m_1)$ رخ می‌دهد، زمانی که:

$$h(m_0, m_1) = \sqrt[m_1 - m_0 + 1]{\frac{m_1!}{(m_0 - 1)!}}$$

اثبات. خانوادهٔ توابع چگالی پوآسون دارای خاصیت MLR در x است. بنا بر این طبق قضیهٔ ۴.۲ داریم:

$$F_{\theta_2}(x) \leq F_{\theta_1}(x), \quad \forall x \in \mathbb{R}, \quad \forall \theta_1 < \theta_2$$

و این، اثبات قسمت‌های اول و دوم لم را به دنبال دارد. حال به اثبات قسمت سوم لم می‌پردازیم. فرض کنید m_0 و m_1 دو عدد حسابی باشند و $m_0 \leq m_1$. تعریف می‌کنیم:

$$y(\theta) = \sum_{i=m_0}^{m_1} \frac{e^{-\theta}\theta^i}{i!}.$$

از تابع $y(\theta)$ نسبت به θ مشتق می‌گیریم:

$$\frac{\partial y(\theta)}{\partial \theta} = e^{-\theta}\theta^{m_0-1}\left(\frac{1}{(m_0-1)!} - \frac{\theta^{m_1-m_0+1}}{m_1!}\right).$$

برای یافتن نقاط اکسترمم تابع $y(\theta)$، مقدار مشتق به‌دست‌آمده را برابر با صفر قرار می‌دهیم:

$$\frac{\partial y(\theta)}{\partial \theta} = e^{-\theta}\theta^{m_0-1}\left(\frac{1}{(m_0-1)!} - \frac{\theta^{m_1-m_0+1}}{m_1!}\right) = 0$$

$$m_1! - (m_0-1)!\theta^{m_1-m_0+1} = 0$$

$$\theta = \sqrt[m_1-m_0+1]{\frac{m_1!}{(m_0-1)!}} \tag{2}$$

واضح است که مشتق $y(\theta)$ برای مقادیر θی کمتر از $\sqrt[m_1-m_0+1]{\frac{m_1!}{(m_0-1)!}}$ مثبت و برای مقادیر θی بیشتر از $\sqrt[m_1-m_0+1]{\frac{m_1!}{(m_0-1)!}}$ منفی است. بنا بر این تابع $y(\theta)$ تابعی تک‌مدی است که بیشینهٔ آن در نقطهٔ $\sqrt[m_1-m_0+1]{\frac{m_1!}{(m_0-1)!}}$ اتفاق می‌افتد و اثبات کامل است. □

فضای پارامتری برای میانگین توزیع پوآسون بازهٔ $(0, +\infty)$ است، اما در عمل برای قابل انجام شدن محاسبات باید فضای پارامتری را کراندار در نظر بگیریم که با توجه به اثبات لم ۴.۲ مشکلی در محاسبهٔ بیشینهٔ احتمالات پوشش در این حالت‌ها نخواهیم داشت.

قضیهٔ ۶.۲. برای یک بازهٔ اطمینان $(L(X), U(X))$ که گزارهٔ ۱.۲ را داشته باشد، بیشینهٔ احتمالات پوشش برابر است با بیشینهٔ احتمالات پوشش به‌ازای نقاط مجموعهٔ زیر:

$$S_1 = \{w_1, w_2, ..., a_1, a_2, ...\}. \tag{3}$$

به‌طوری که

$$a_i = h(k_0(\frac{w_i + w_{i+1}}{2}), k_1(\frac{w_i + w_{i+1}}{2}))$$

و تابع h در لم ۵.۲ تعریف شده است.

[۲] monotone likelihood ratio

اثبات. میتوان نوشت:

$$\max_{\theta\in\ (l,u)} P_\theta(\theta\in(L(X),U(X))) = \max_{i=۱,۲,\dots}\big(\max_{w_i<\theta<w_{i+۱}} P_\theta(\theta\in(L(X),U(X)))\big)$$

از طرفی طبق لم ۵.۲ داریم:

$$P_\theta(\theta\in(L(X),U(X))) = \sum_{i=k_۰(\theta)}^{k_۱(\theta)} \frac{e^{-\theta}\theta^i}{i!}$$

و در لم ۵.۲ ثابت کردیم تابع $\sum_{i=m_۰}^{m_۱} \frac{e^{-\theta}\theta^i}{i!}$ تابعی صعودی، نزولی، یا تک‌مدی است که بیشینهٔ آن در نقطهٔ $h(m_۰,m_۱)$ رخ می‌دهد. با توجه به این‌که برای θهای متعلق به فاصلهٔ $(w_i,w_{i+۱})$، $k_۰(\theta)$ و $k_۱(\theta)$ یکسان هستند، می‌توانیم $k_۰(\frac{w_i+w_{i+۱}}{۲})$ و $k_۱(\frac{w_i+w_{i+۱}}{۲})$ را در نظر بگیریم. بنا بر این $\max_{w_i<\theta<w_{i+۱}} P_\theta(\theta\in(L(X),U(X)))$ به‌ازای یکی از نقاط $\theta=w_i$، $\theta=w_{i+۱}$ و $\theta=a_i$ اتفاق می‌افتد. به عبارت دیگر بیشینهٔ احتمالات پوشش به ازای یکی از نقاط مجموعهٔ $S_۱$ اتفاق می‌افتد که درون فضای پارامتری نیز قرار دارد. بنا بر این اثبات کامل است. □

۱.۲ الگوریتم محاسبهٔ دقیق بیشینهٔ احتمالات پوشش

گام ۱: در ابتدا بررسی می‌شود که بازهٔ اطمینان مورد نظر، گزارهٔ ۱.۲ را دارا باشد. اگر گزارهٔ ۱.۲ برقرار نباشد، قادر به محاسبهٔ بیشینهٔ احتمالات پوشش نیستیم.

گام ۲: نقاط انتهایی بازه‌های اطمینان متناظر با $x=۰,۱,۲,\dots$ را که درون فضای پارامتری قرار دارند در نظر می‌گیریم.

گام ۳: نقاط $a_۱,a_۲,\dots$ ارائه‌شده در قضیهٔ ۶.۲ را به دست می‌آوریم.

گام ۴: احتمالات پوشش متناظر با نقاط به‌دست‌آمده در گام‌های دوم و سوم و کران‌های پایین و بالای فضای پارامتری را به دست می‌آوریم. بیشینهٔ احتمالات پوشش به‌دست‌آمده همان بیشینهٔ دقیق احتمالات پوشش بازهٔ مورد نظر است.

۳ مثال‌ها

در مثال‌های ۱.۳ و ۲.۳ با استفاده از الگوریتم به‌دست‌آمده در بخش ۲ بیشینهٔ احتمالات پوشش دو بازهٔ اطمینان مجانبی معرفی‌شده برای پارامتر میانگین توزیع پوآسون توسط گوآن [۳] را به‌صورت دقیق محاسبه و سپس آن‌ها را با معیار بیشینهٔ احتمالات پوشش مقایسه می‌کنیم.

مثال ۱.۳. فرض کنید $z_{\frac{\alpha}{۲}}$ چندک $\frac{\alpha}{۲}$ بالایی توزیع نرمال استاندارد باشد. بازهٔ اطمینان SC یا امتیاز با ضریب $۱-\alpha$ برای θ به‌صورت زیر معرفی شده است:

$$x+\frac{z^۲_{\frac{\alpha}{۲}}}{۲} \mp z_{\frac{\alpha}{۲}}\sqrt{x+\frac{z^۲_{\frac{\alpha}{۲}}}{۴}}$$

در این‌جا بیشینهٔ احتمالات پوشش را روی فضای پارامتری محدود شده به بازهٔ $(۰,۵)$ به دست می‌آوریم. جدول ۱ بازه‌های SC، ۹۵ درصدی به‌ازای $x=۰,۱,\dots,۱۰$ را نشان می‌دهد. با توجه به جدول ۱ برای $x\geq ۱۰$ بازه‌های اطمینان کاملاً خارج از بازهٔ $(۰,۵)$ قرار می‌گیرند. بنا بر این کافی است فقط بازه‌های اطمینان متناظر با $x=۰,۱,\dots,۹$ را در نظر بگیریم. با توجه به بازه‌ها واضح است که گزارهٔ ۱.۲ برقرار است. بنا بر این می‌توانیم بیشینهٔ احتمالات پوشش را به‌صورت دقیق محاسبه کنیم. برای این منظور در ابتدا کران‌های بازه‌های اطمینان به‌ازا $x=۰,۱,\dots,۹$ را به‌ترتیب صعودی مرتب می‌کنیم، یعنی ۰ و ۰/۱۷۶ و ۰/۵۴۸ و ۱/۰۲۰ و ۱/۵۵۵ و ۲/۱۳۶ و ۲/۷۵۰ و ۳/۳۹۱ و ۳/۸۴۲ و ۴/۰۵۴ و ۴/۷۳۵ و ۵/۶۶۵ و ... و ۱۵/۷۸۸ و ۱۷/۱۰۶. با توجه به این که فضای پارامتری، بازهٔ $(۰,۵)$ است، داریم:

$$w_۱=۰,\ w_۲=۰/۱۷۶,\ w_۳=۰/۵۴۸,\ w_۴=۱/۰۲۰,$$
$$w_۵=۱/۵۵۵,\ w_۶=۲/۱۳۶,\ w_۷=۲/۷۵۰,\ w_۸=۳/۳۹۱,$$
$$w_۹=۳/۸۴۲,\ w_{۱۰}=۴/۰۵۴,\ w_{۱۱}=۴/۷۳۵,\ w_{۱۲}=۵.$$

همچنین داریم، $a_۱=۰$ و $a_۲=۰/۱۷۶$ و $a_۳=۰/۵۴۸$ و $a_۴=۱/۰۲۰$ و $a_۵=۱/۵۵۵$ و $a_۶=۲/۱۳۶$ و $a_۷=۲/۷۵۰$ و $a_۸=۳/۳۹۱$ و $a_۹=۳/۳۸۰$ و $a_{۱۰}=۳/۷۶۴$ و $a_{۱۱}=۴/۱۴۷$.

احتمال پوشش در $w_۱$ و ... و $w_{۱۲}$ و $a_۱$ و ... و $a_{۱۱}$ به‌ترتیب برابر است با ۰/۹۹۹ و ۰/۸۳۸ و ۰/۸۹۵ و ۰/۹۱۶ و ۰/۹۲۷ و ۰/۹۳۴ و ۰/۹۳۹ و ۰/۹۴۳ و ۰/۹۵۸ و ۰/۹۲۸ و ۰/۹۳۹ و ۰/۹۶۱

و ۰/۹۹۹ و ۰/۸۳۸ و ۰/۸۹۵ و ۰/۹۱۶ و ۰/۹۲۷ و ۰/۹۳۴ و ۰/۹۳۹ و ۰/۹۴۳ و ۰/۹۴۴ و ۰/۹۶۲ و ۰/۹۵۸.

جدول ۱. فاصلهٔ اطمینان ۹۵ درصدی SC برای $x = ۰, ..., ۱۰$

x	$L(x)$	$U(x)$
۰	۰/۰۰۰	۳/۸۴۲
۱	۰/۱۷۶	۵/۶۶۵
۲	۰/۵۴۸	۷/۲۹۳
۳	۱/۰۲۰	۸/۸۲۱
۴	۱/۵۵۵	۱۰/۲۸۶
۵	۲/۱۳۶	۱۱/۷۰۶
۶	۲/۷۵۰	۱۳/۰۹۲
۷	۳/۳۹۱	۱۴/۴۵۱
۸	۴/۰۵۴	۱۵/۷۸۸
۹	۴/۷۳۵	۱۷/۱۰۶
۱۰	۵/۴۳۲	۱۸/۴۱۰

بنا بر این بیشینهٔ احتمالات پوشش بازهٔ اطمینان SC با ضریب ۰/۹۵ به‌کمک از ۰/۹۹۹. (عدد ارائه‌شده گرد شده است و مقدار دقیق آن تا ۶ رقم اعشار برابر است با ۰/۹۹۹۹۹۹) برای آشنا شدن با نحوهٔ محاسبهٔ احتمال پوشش، احتمال پوشش را به‌ازای $\theta = ۴/۰۵۴$ به‌دست می‌آوریم.

جدول ۲. فاصلهٔ اطمینان ۹۵ درصدی MSC برای $x = ۰, ..., ۱۰$

x	$L(x)$	$U(x)$
۰	−۰/۱۵۴	۳/۶۸۸
۱	۰/۰۲۳	۵/۵۱۱
۲	۰/۳۹۵	۷/۱۳۹
۳	۰/۸۶۶	۸/۶۶۸
۴	۱/۴۰۲	۱۰/۱۳۲
۵	۱/۹۸۲	۱۱/۵۵۲
۶	۲/۵۹۶	۱۲/۹۳۸
۷	۳/۲۳۷	۱۴/۲۹۷
۸	۳/۹۰۰	۱۵/۶۳۴
۹	۴/۵۸۱	۱۶/۹۵۳
۱۰	۵/۲۷۸	۱۸/۲۵۶

در ابتدا بررسی می‌کنیم که ۴/۰۵۴ به‌ازای چه xهایی در بازهٔ اطمینان قرار می‌گیرد. سپس احتمالات مشاهدهٔ آن xها را با فرض $\theta = ۴/۰۵۴$ با هم جمع می‌کنیم. با توجه به بازه‌های بالا، ۴/۰۵۴ در بازه‌های متناظر با $x = ۱, ۲, ۳, ۴, ۵, ۶, ۷$ قرار می‌گیرد. بنا بر این داریم:

$$P_{۴/۰۵۴}(۴/۰۵۴ \in (L(X), U(X))) = \sum_{x=۱}^{۷} \frac{e^{-x} x^{۴/۰۵۴}}{x!}$$
$$= ۰/۹۲۸۲$$

مثال ۲.۳. بازهٔ اطمینان MSC یا امتیاز متحرک[۳] با ضریب $۱ - \alpha$ برای θ به‌صورت زیر معرفی شده است:

$$x + ۰/۴۶ z^{۲}_{\frac{\alpha}{۲}} \mp z_{\frac{\alpha}{۲}} \sqrt{x + \frac{z^{۲}_{\frac{\alpha}{۲}}}{۴}}$$

در این‌جا نیز بیشینهٔ احتمالات پوشش را روی فضای پارامتری محدود شده به بازهٔ $(۰, ۵)$ به دست می‌آوریم. جدول ۲ بازه‌های نود و پنج درصدی MSC را به‌ازای $x = ۰, ۱, ..., ۱۰$ نشان می‌دهد. با توجه به جدول ۲ برای $x \geq ۱۰$ بازه‌های اطمینان کاملاً خارج از بازهٔ $(۰, ۵)$ قرار می‌گیرند. بنا بر این کافی است فقط بازه‌های اطمینان متناظر با $x = ۰, ۱, ..., ۹$ را در نظر بگیریم. با توجه به بازه‌ها واضح است که گزارهٔ ۱.۲ برقرار است. بنا بر این می‌توانیم بیشینهٔ احتمالات پوشش را به‌صورت دقیق محاسبه کنیم.

برای این منظور، در ابتدا کران‌های فواصل را به‌ترتیب صعودی مرتب می‌کنیم؛ یعنی ۰/۱۵۴− و ۰/۰۲۳ و ۰/۳۹۵ و ۰/۸۶۶ و ۱/۴۰۲ و ۱/۹۸۲ و ۲/۵۹۶ و ۳/۲۳۷ و ۳/۶۸۸ و ۳/۹۰۰ و ۴/۵۸۱ و ۵/۲۷۸ و ۱۶/۹۵۳ و ۱۸/۲۵۶ و $\cdots$. با توجه به این‌که فضای پارامتری بازهٔ $(۰, ۵)$ است، $w_۱ = ۰$ و $w_۲ = ۰/۰۲۳$ و $w_۳ = ۰/۳۹۵$ و $w_۴ = ۰/۸۶۶$ و $w_۵ = ۱/۴۰۲$ و $w_۶ = ۱/۹۸۲$ و $w_۷ = ۲/۵۹۶$ و $w_۸ = ۳/۲۳۷$ و $w_۹ = ۳/۶۸۸$ و $w_{۱۰} = ۳/۹۰۰$ و $w_{۱۱} = ۴/۵۸۱$ و $w_{۱۲} = ۵$. همچنین داریم $a_۱ = ۰$ و $a_۲ = ۰/۰۲۳$ و $a_۳ = ۰/۳۹۵$ و $a_۴ = ۰/۸۶۶$ و $a_۵ = ۱/۴۰۲$ و $a_۶ = ۱/۹۸۲$ و $a_۷ = ۲/۵۹۶$ و $a_۸ = ۳/۲۳۷$ و $a_۹ = ۳/۳۸۰$ و $a_{۱۰} = ۳/۷۶۴$ و $a_{۱۱} = ۴/۱۴۷$. احتمال پوشش در $w_۱$ و $\cdots$ و $w_{۱۲}$ و $a_۱$ و $\cdots$ و $a_{۱۱}$ به‌ترتیب برابر است با ۱/۰۰۰ و ۰/۹۷۷ و ۰/۹۴۰ و ۰/۹۴۲ و ۰/۹۴۶ و ۰/۹۴۹ و۰/۹۵۱ و ۰/۹۵۳ و ۰/۹۴۰ و۰/۹۳۴ و ۰/۹۴۶ و ۰/۹۶۱ و ۱/۰۰۰ و ۰/۹۷۷

[۳] moved score confidence interval

و ۰/۹۴۰ و ۰/۹۴۲ و ۰/۹۴۶ و ۰/۹۴۹ و ۰/۹۵۱ و ۰/۹۵۳ و ۰/۹۷۸ و ۰/۹۳۸ و ۰/۹۵۸. بنا بر این بیشینهٔ احتمالات پوشش بازهٔ MSC با ضریب ۰/۹۵ به‌کمک از ۱/۰۰۰.

۴ مطالعهٔ شبیه‌سازی

در این بخش در ابتدا دو بازهٔ اطمینان مجانبی AWC یا والد اصلاح‌شده با اضافه کردن احتمال در دم توزیع[۴] و WCC یا والد اصلاح‌شده با تصحیح پیوستگی[۵] برای پارامتر میانگین توزیع پوآسون را از مقالهٔ خام‌کنگ [۶] معرفی می‌کنیم و سپس برای اثبات نیرومندی الگوریتم ارائه‌شده در بخش ۲ از یک مطالعهٔ شبیه‌سازی استفاده می‌کنیم.

۱.۴ بازهٔ اطمینان AWC

بازهٔ اطمینان AWC با ضریب $\alpha - ۱$ برای θ به‌صورت زیر معرفی شده است:

$$x + \frac{z^{۲}_{\frac{\alpha}{۲}}}{۲} \mp z_{\frac{\alpha}{۲}}\sqrt{x}$$

۲.۴ بازهٔ اطمینان WCC

بازهٔ اطمینان WCC با ضریب $\alpha - ۱$ برای θ به‌صورت زیر معرفی شده است:

$$x \mp z_{\frac{\alpha}{۲}}\sqrt{x + ۰/۵}$$

جدول ۳ بیشینهٔ احتمالات پوشش بازه‌های اطمینان SC، MSC، AWC و WCC نود و پنج درصدی را، در صد هزار نقطه که به‌تصادف از توزیع یکنواخت روی فاصلهٔ $(۰,۵)$ انتخاب شده‌اند (در این‌جا فضای پارامتری را بازهٔ $(۰,۵)$ در نظر گرفته‌ایم). در کنار بیشینهٔ دقیق احتمالات پوشش محاسبه‌شده با الگوریتم ارائه‌شده در این مقاله نشان می‌دهد.

۵ بحث و نتیجه‌گیری

در این مقاله روش محاسبهٔ بیشینهٔ دقیق احتمالات پوشش بازه‌های اطمینان با کران‌های صعودی برای میانگین توزیع پوآسون را ارائه کردیم. با این روش می‌توانیم با محاسبهٔ احتمالات پوشش بازهٔ اطمینان در تعداد کمی از نقاط فضای پارامتری به بیشینهٔ دقیق احتمالات پوشش بازهٔ مورد نظر دست یابیم. نتایج شبیه‌سازی در بخش ۴ نشان می‌دهد که مقادیر بیشینهٔ تقریبی به‌دست آمده در مقایسه با بیشینهٔ دقیق احتمالات پوشش محاسبه‌شده برای بازه‌های اطمینان توسط الگوریتم بخش ۲ کوچک‌تر هستند، و این نشان‌دهندهٔ اهمیت زیاد روش پیشنهادی است. چرا که با محاسبهٔ احتمال پوشش در نقاطی بسیار کمتر به بیشینهٔ احتمالات پوشش بیشتری دست می‌یابیم. با توجه به جدول ۳ بر اساس بیشینهٔ احتمالات پوشش، بازه‌ها به‌ترتیب بهینگی عبارت‌اند از:

۱. MSC یا WCC

۲. SC

۳. AWC

جدول ۳. بیشینهٔ احتمالات پوشش بازه‌های اطمینان ۹۵ درصدی در صدهزار نقطهٔ فضای پارامتری که به‌تصادف انتخاب، در کنار بیشینهٔ دقیق احتمالات پوشش

بازهٔ اطمینان	بیشینهٔ تقریبی	بیشینهٔ دقیق
SC	۰/۹۹۹۹۴۲	۰/۹۹۹۹۹۹
MSC	۰/۹۹۹۹۱۴	۱
AWC	۰/۹۲۳۱۹۱	۰/۹۳۱۸۱۹
WCC	۱	۱

[۴] adding the tail probability of the wald confidence interval

[۵] wald interval with continuity correction

مراجع

[1] Byrne, J. and Kabaila, P. (2001). Short exact confidence intervals for the Poisson mean, *Communications in Statistics: Theory and Methods*, **30**, 257-261.

[2] Casella, G. and Berger, R. (2002). *Statistical inference*, 2nd ed. United States of America:Thomson Learning Academic Resource Center

[3] Guan, Y. (2011). Moved score confidence intervals for means of discrete distributions, *American Open Journal of Statistics*, **1**, 81-86.

[4] Kabaila, P. and Byrne, J. (2001). Exact short Poisson confidence intervals, *Canadian Journal of Statistics*, **29**, 99-106.

[5] Kabaila, P. and Lloyd, C.J. (1997). Tight upper confidence limits from discrete data, *Australian Journal of Statistics*, **39**, 193-204.

[6] Khamkong, M. (2012). Approximate confidence interval for the mean of Poisson distribution, *American Open Journal of Statistics*, **2**, 204-207.

[7] Sahai, H. and Khurshid, A. (1993). Confidence intervals for the mean of a Poisson distribution, *Biometrical Journal*, **35**, **7**, 857-867.

[8] Schwertman, N.C. and Martinez, (1994). Approximate Poisson confidence limits, *Communication in Statistics: Theory and Methods*, **23(5)**, 1507-1529.

[9] Shirvani, A. and Fathizadeh, M. (2016). Assessing the accuracy of approximate confidence intervals proposed for the mean of a Poisson distribution, *Journal of Modern Applied Statistical Methods*. to appear.

[10] Shirvani, A. and Toomaj, A. (2016). Exact confidence coefficient and maximum coverage probabilities of nested confidence intervals for geometric parameter, *European Journal of Pure and Applied Mathematics*, to appear.

[11] Wang, H. (2009). Exact average coverage probabilities and confidence coefficients of confidence intervals for discrete distributions, *Statistics and Computing*, **19**, 139-148.