ว.วิทย. มข. 37(1) 92-101 (2552)　　KKU Sci. J.37(1) 92-101 (2009)

ขอบเขตล่างและบนของความคลาดเคลื่อนสัมพัทธ์ระหว่างฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มทวินามนิเสธและตัวแปรสุ่มปัวซง

Lower and Upper Bounds of the Relative Error Between the Negative Binomial and Poisson Cumulative Distribution Functions

คณินทร์ ธีรภาพโอฬาร [1]

บทคัดย่อ

งานวิจัยนี้ได้ใช้วิธีสไตน์เชนเพื่อหาขอบเขตล่างและบนของความคลาดเคลื่อนสัมพัทธ์ระหว่างฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มทวินามนิเสธและตัวแปรสุ่มปัวซง โดยที่ค่าเฉลี่ยของตัวแปรสุ่มปัวซง $\lambda = n(1-p)$ เมื่อ n และ p คือ พารามิเตอร์ของการแจกแจงทวินามนิเสธ ขอบเขตดังกล่าวชี้ให้เห็นว่าฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มทวินามนิเสธสามารถประมาณด้วยฟังก์ชันการแจกแจงของตัวแปรสุ่มปัวซงได้ดีเมื่อ $q = 1-p$ หรือ λ มีค่าน้อย

Abstract

This paper uses the Stein-Chen method to determine upper and lower bounds on the relative error between the negative binomial and Poisson cumulative distribution functions, where the Poisson mean $\lambda = n(1-p)$ and n and p are parameters of the negative binomial distribution. For these bounds, it is pointed out that the negative binomial cumulative distribution function can be properly approximated by the Poisson cumulative distribution function if $q = 1-p$ or λ is small.

คำสำคัญ : การประมาณปัวซง, การแจกแจงปัวซง, การแจกแจงทวินามนิเสธ, ฟังก์ชันการแจกแจงสะสม, วิธีสไตน์เชน

Keywords: Poisson approximation, Poisson distribution, negative binomial distribution, cumulative distribution function, Stein-Chen method.

[1] ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ มหาวิทยาลัยบูรพา 20131
E-mail: kanint@buu.ac.th

1. บทนำ

ให้ X เป็นจำนวนความล้มเหลว (failure) ก่อนความสำเร็จครั้งที่ n ในลำดับของการทดลองย่อยแบร์นูลลี (Bernoulli trials) ที่เป็นอิสระต่อกัน โดยที่ความสำเร็จ (success) และความล้มเหลว (failure) ที่เกิดขึ้นในแต่ละการทดลองย่อยมีความน่าจะเป็นเท่ากับ $p \in (0,1)$ และ $q = 1-p$ ตามลำดับ แล้วการแจกแจงความน่าจะเป็นของ X สามารถเขียนได้เป็น $P(X=k) = \binom{n+k-1}{k} q^k p^n$ เมื่อ k = 0, 1, … *การแจกแจงนี้เรียกว่า การแจกแจงทวินามนิเสธที่มีพารามิเตอร์ n และ p* (negative binomial distribution with parameters n and p) ในกรณีที่ n = 1 จะเรียกการแจกแจงของ X ว่า *การแจกแจงเรขาคณิตที่มีพารามิเตอร์ p* (geometric distribution with parameter p) ซึ่งเป็นการแจกแจงของจำนวนความล้มเหลวก่อนความสำเร็จครั้งแรก (Johnson et al., 2005)

ให้ $Y_1, \ldots, Y_n$ เป็น n ตัวแปรสุ่มที่เป็นอิสระต่อกันและแต่ละ Y_i มีการแจกแจงเรขาคณิตแบบเดียวกัน นั่นคือ $P(Y_i = k) = pq^k$, k = 0, 1, … สำหรับทุก i = 1, 2, …, n ให้ $X = \sum_{i=1}^{n} Y_i$ แล้วเราจะได้ว่า X จะมีการแจกแจงทวินามนิเสธที่มีพารามิเตอร์ n และ p เช่นเดียวกับที่ได้กล่าวมาข้างต้น และเป็นที่ทราบกันทั่วไปว่าในกรณีที่ q มีค่าน้อยหรือมีค่าเข้าใกล้ 0 การแจกแจงทวินามนิเสธสามารถประมาณได้ด้วยการแจกแจงปัวซงที่มีค่าเฉลี่ย $\lambda = \frac{nq}{p}$ (ซึ่งมีค่าเท่ากับค่าเฉลี่ยของตัวแปรสุ่มทวินามนิเสธ) ในทำนองเดียวกันฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มทวินามนิเสธแทนด้วย $\mathcal{NB}(k;n,p) = \sum_{j=0}^{k} \binom{n+j-1}{j} q^j p^n$ สามารถประมาณด้วยฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มปัวซงแทนด้วย $\mathcal{P}(k;\lambda) = \sum_{j=0}^{k} \frac{e^{-\lambda}\lambda^j}{j!}$ เมื่อ k = 0, 1, … และ $\lambda = \frac{nq}{p}$ ได้เช่นเดียวกัน ในกรณีนี้ คณินทร์ ธีรภาพโอฬาร และคณะ (2551) ได้ใช้วิธีสไตน์เชน (Stein-Chen method) และฟังก์ชัน w (w-function) สร้างขอบเขตบนของความคลาดเคลื่อนสัมพัทธ์ระหว่างฟังก์ชันการแจกแจงสะสมของสองตัวแปรสุ่มเพื่อใช้วัดความคลาดเคลื่อนของการประมาณดังนี้

$$\left|1 - \frac{\mathcal{P}(x_0;\lambda)}{\mathcal{NB}(x_0;n,p)}\right| \le \frac{1-e^{-\lambda}}{p^{n+\min\{x_0, \lceil \lambda \rceil + 1\}}} \min\left\{\frac{(1-p^n)}{n}, \frac{q}{p(x_0+1)}\right\} \tag{1.1}$$

โดยที่ $x_0 \in \mathbb{N} \cup \{0\}$ และ $\lceil \lambda \rceil$ เป็นจำนวนเต็มที่น้อยที่สุดที่ไม่น้อยกว่า λ

นอกจากการประมาณการแจกแจงทวินามนิเสธด้วยการแจกแจงปัวซงที่มีค่าเฉลี่ย $\lambda = \frac{nq}{p}$ แล้ว ในทางทฤษฎีเรายังทราบว่าถ้า $n \to \infty$ และ $q \to 0$ โดยที่ nq เป็นค่าคงตัว แล้วจะได้ว่าการแจกแจงทวินามนิเสธจะลู่เข้าสู่การแจกแจงปัวซงด้วยค่าเฉลี่ย $\lambda = nq$ (Feller, 1968) ซึ่งแสดงว่าเราสามารถประมาณการแจกแจงทวินามนิเสธด้วยการแจกแจงปัวซงที่มีค่าเฉลี่ย $\lambda = nq$ ได้เช่นเดียวกัน สำหรับงานวิจัยนี้เราต้องการศึกษาการประมาณ

$\mathcal{NB}(k;n,p)$ ด้วย $\mathcal{P}(k;\lambda)$ ที่มีค่าเฉลี่ย $\lambda = nq$ โดยใช้วิธีสไตน์เชน และเกณฑ์ที่ใช้วัดความคลาดเคลื่อนของการประมาณในงานวิจัยนี้ คือ ขอบเขตล่างและบนของความคลาดเคลื่อนสัมพัทธ์ระหว่างฟังก์ชันการแจกแจงสะสมของสองตัวแปรสุ่มดังกล่าว ซึ่งคล้ายกับในงานวิจัยของ คณินทร์ ธีรภาพโอฬาร และคณะ (2551) โดยที่ความคลาดเคลื่อนสัมพัทธ์ระหว่างฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มทวินามนิเสธและตัวแปรสุ่มปัวซงที่ x_0 คือ $1-\dfrac{\mathcal{P}(x_0;\lambda)}{\mathcal{NB}(x_0;n,p)}$

2. วิธีดำเนินการวิจัย

ศึกษาการประมาณการแจกแจงด้วยการแจกแจงปัวซงในรูปของความคลาดเคลื่อนสัมพัทธ์จากงานวิจัยของ Teerapabolarn (2007) และงานวิจัยของ คณินทร์ ธีรภาพโอฬาร และคณะ (2551) ซึ่งจากการศึกษางานวิจัยดังกล่าว เราพบว่าวิธีสไตน์เชนเป็นวิธีที่สามารถหาผลลัพธ์ที่เราต้องการได้ ดังนั้นเราควรเริ่มต้นด้วยการสร้างบทตั้ง (lemma) เพื่อใช้ในการพิสูจน์ทฤษฎีบทหลัก (main theorem) ดังนี้

พิจารณาสมการของสไตน์ (Stein's equation) (Stein, 1986) สำหรับการแจกแจงปัวซงที่มีพารามิเตอร์ λ ซึ่ง (กำหนดฟังก์ชัน h) นิยามโดย

$$\lambda f(x+1) - xf(x) = h(x) - \wp_\lambda(h) \tag{2.1}$$

โดยที่

$$\wp_\lambda(h) = e^{-\lambda}\sum_{\ell=0}^{\infty} h(\ell)\frac{\lambda^\ell}{\ell!} \tag{2.2}$$

และ f และ h เป็นฟังก์ชันค่าจริงที่มีขอบเขตที่นิยามบนเซต $\mathbb{N}\cup\{0\}$

สำหรับ $x_0 \in \mathbb{N}\cup\{0\}$ ให้ $h_{x_0} : \mathbb{N}\cup\{0\} \to \mathbb{R}$ เป็นฟังก์ชันที่กำหนดโดย

$$h_{x_0}(x) = \begin{cases} 1\ , x \le x_0 \\ 0\ , x > x_0 \end{cases}$$

โดย Barbour et al. (1992) และสำหรับ $x \in \mathbb{N}\cup\{0\}$ จะได้ว่าผลเฉลย f เมื่อแทน h ด้วย h_{x_0} ในสมการ (2.1) สามารถเขียนได้เป็น

$$f(x) = \begin{cases} (x-1)!\lambda^{-x}e^{\lambda}[\wp_\lambda(h_{x_0})\wp_\lambda(1-h_{x-1})]\ , x > x_0 \\ (x-1)!\lambda^{-x}e^{\lambda}[\wp_\lambda(h_{x-1})\wp_\lambda(1-h_{x_0})]\ , x \le x_0 \\ 0 \qquad , x = 0 \end{cases} \tag{2.3}$$

บทตั้ง 2.1 *ให้* $x_0 \in \mathbb{N} \cup \{0\}$ *และ* $k \in \mathbb{N} \setminus \{1\}$ *แล้วจะได้ว่า*

$$0 \le \sup_{x \ge k} f(x) \le \lambda^{-1}(e^{\lambda}-1)\mathcal{P}(x_0;\lambda)\min\left\{\frac{1}{k},\frac{1}{x_0+1}\right\} \tag{2.4}$$

พิสูจน์ เราจะเห็นได้ชัดว่า $f(x) \ge 0$ สำหรับทุก $x \in \mathbb{N} \cup \{0\}$ ดังนั้นเราจะแสดงว่า

$$\sup_{x \ge k} f(x) \le \lambda^{-1}(e^{\lambda}-1)\mathcal{P}(x_0;\lambda)\min\left\{\frac{1}{k},\frac{1}{x_0+1}\right\}$$

โดยแบ่งการพิสูจน์ออกเป็น 2 กรณี ดังนี้

กรณีที่ 1: $k \le x \le x_0$

จากสมการ (2.3) เราจะได้ว่า

$$\begin{aligned}
f(x) &= \wp_{\lambda}(h_{x-1})(x-1)!\lambda^{-x}e^{\lambda}\wp_{\lambda}(1-h_{x_0}) \\
&\le \mathcal{P}(x_0;\lambda)(x-1)!\sum_{j=x_0+1}^{\infty}\frac{\lambda^{j-x}}{j!} \quad \text{(จากสมการ (2.2))} \\
&= \mathcal{P}(x_0;\lambda)(x-1)!\left\{\frac{\lambda^{(x_0+1)-x}}{(x_0+1)!}+\frac{\lambda^{(x_0+2)-x}}{(x_0+2)!}+\cdots\right\} \\
&= \mathcal{P}(x_0;\lambda)\left\{\frac{(x-1)!\lambda^{(x_0+1)-x}}{(x_0+1)!}+\frac{(x-1)!\lambda^{(x_0+2)-x}}{(x_0+2)!}+\cdots\right\} \\
&= \mathcal{P}(x_0;\lambda)\left\{\frac{\lambda^{(x_0+1)-x}}{(x_0+1)\binom{x_0}{x-1}[(x_0+1)-x]!}+\frac{\lambda^{(x_0+2)-x}}{(x_0+2)\binom{x_0+1}{x-1}[(x_0+2)-x]!}+\cdots\right\} \\
&\le \frac{\mathcal{P}(x_0;\lambda)}{x_0+1}\left\{\frac{\lambda}{2!}+\frac{\lambda^2}{3!}+\cdots\right\} \\
&= \frac{\mathcal{P}(x_0;\lambda)\lambda^{-1}}{x_0+1}\left\{\frac{\lambda^2}{2!}+\frac{\lambda^3}{3!}+\cdots\right\} \\
&\le \frac{\lambda^{-1}(e^{\lambda}-1)\mathcal{P}(x_0;\lambda)}{x_0+1}
\end{aligned}$$

และ $$f(x) \le \frac{\lambda^{-1}(e^{\lambda}-1)\mathcal{P}(x_0;\lambda)}{k}$$

กรณีที่ 2: $k \le x$ และ $x_0 + 1 \le x$

$$\begin{aligned}
f(x) &= \wp_\lambda(h_{x_0})(x-1)!\lambda^{-x}e^{\lambda}\wp_\lambda(1-h_{x-1}) \\
&= \mathcal{P}(x_0;\lambda)(x-1)!\sum_{j=x}^{\infty}\frac{\lambda^{j-x}}{j!} \qquad \text{(จากสมการ (2.2))} \\
&= \mathcal{P}(x_0;\lambda)(x-1)!\left\{\frac{1}{x!}+\frac{\lambda}{(x+1)!}+\cdots\right\} \\
&= \frac{\mathcal{P}(x_0;\lambda)}{x}\left\{1+\frac{\lambda}{x+1}+\cdots\right\} \\
&\le \frac{\mathcal{P}(x_0;\lambda)\lambda^{-1}}{x}\left\{\lambda+\frac{\lambda^2}{2!}+\cdots\right\} \\
&= \frac{\lambda^{-1}(e^{\lambda}-1)\mathcal{P}(x_0;\lambda)}{x}
\end{aligned}$$

ซึ่งทำให้เราได้

$$f(x) \le \frac{\lambda^{-1}(e^{\lambda}-1)\mathcal{P}(x_0;\lambda)}{k} \quad \text{และ} \quad f(x) \le \frac{\lambda^{-1}(e^{\lambda}-1)\mathcal{P}(x_0;\lambda)}{x_0+1}$$

ดังนั้นจากทั้ง 2 กรณี เราจะได้อสมการ (2.4) □

บทตั้ง 2.2 *ให้* $x_0 \in \mathbb{N} \cup \{0\}$ *และ* $\lambda = nq$ *แล้วอสมการต่อไปนี้เป็นจริง*

$$\frac{\mathcal{P}(x_0;\lambda)}{\mathcal{NB}(x_0;n,p)} \le \frac{e^{-\lambda}}{p^n} \tag{2.5}$$

พิสูจน์ เราเห็นได้ชัดในกรณี $x_0 = 0$ ดังนั้นจะแสดงว่าอสมการ (2.5) เป็นจริงในกรณี $x_0 \ge 1$ นั่นคือ

$$\begin{aligned}
\sum_{k=0}^{x_0}\binom{n+k-1}{k}q^k &= 1+\sum_{k=1}^{x_0}\frac{(n+k-1)!q^k}{k!(n-1)!} \\
&= 1+\sum_{k=1}^{x_0}\frac{(n+k-1)\cdots nq^k}{k!} \\
&= 1+\sum_{k=1}^{x_0}\frac{(nq)^k}{k!}\left\{\frac{(n+k-1)}{n}\cdots\frac{n}{n}\right\} \\
&\ge 1+\sum_{k=1}^{x_0}\frac{\lambda^k}{k!} \\
&= \sum_{k=0}^{x_0}\frac{\lambda^k}{k!}
\end{aligned}$$

และเราจะได้

$$\frac{\mathcal{P}(x_0;\lambda)}{\mathcal{NB}(x_0;n,p)} = \frac{e^{-\lambda}}{p^n}\frac{\sum_{k=0}^{x_0}\frac{\lambda^k}{k!}}{\sum_{k=0}^{x_0}\binom{n+k-1}{k}q^k}$$
$$\leq \frac{e^{-\lambda}}{p^n}$$

ซึ่งทำให้อสมการ (2.5) เป็นจริง □

3. ผลการวิจัย

ในงานวิจัยนี้ผลการวิจัยที่เราต้องการ คือ ขอบเขตล่างและบนของความคลาดเคลื่อนสัมพัทธ์ระหว่างฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มทวินามนิเสธและตัวแปรสุ่มปัวซง ซึ่งเราสามารถหาได้โดยใช้วิธีสไตน์เชนโดยเริ่มต้นด้วยสมการของสไตน์ (2.1) ทฤษฎีบทต่อไปนี้เป็นผลลัพธ์ของการประมาณการแจกแจงที่อยู่ในรูปแบบของขอบเขตล่างและบนของความคลาดเคลื่อนสัมพัทธ์ระหว่างฟังก์ชันการแจกแจงสะสมของสองตัวแปรสุ่มดังกล่าว

ทฤษฎีบท 3.1 *ให้ X เป็นตัวแปรสุ่มทวินามนิเสธดังที่ได้กล่าวมาแล้ว และ $\lambda = nq$ แล้วจะได้ว่า*

$$-\frac{1-e^{-\lambda}}{p^n}\min\left\{q, \frac{q}{p(x_0+1)}\right\} \leq 1 - \frac{\mathcal{P}(x_0;\lambda)}{\mathcal{NB}(x_0;n,p)} \leq 0 \tag{3.1}$$

โดยที่ $x_0 \in \mathbb{N}\cup\{0\}$

พิสูจน์ จากสมการ (2.1) เมื่อกำหนด $h = h_{x_0}$ และเราหาค่าคาดหมายตลอดสมการ (2.1) จะได้ว่า

$$P(X \leq x_0) - \sum_{j=0}^{x_0}\frac{e^{-\lambda}\lambda^j}{j!} = E[\lambda f(X+1) - Xf(X)]$$

หรือ

$$\mathcal{NB}(x_0;n,p) - \mathcal{P}(x_0;\lambda) = E[\lambda f(X+1) - Xf(X)]$$
$$= E\left[\sum_{i=1}^{n} qf(X+1) - \sum_{i=1}^{n} Y_i f(X)\right]$$
$$= \sum_{i=1}^{n} E\left[qf(X+1) - Y_i f(X)\right] \tag{3.2}$$

โดยที่ f นิยามเช่นเดียวกับสมการ (2.2)

ให้ $X_i = X - Y_i$ ดังนั้นสำหรับแต่ละ i เราจะได้ว่า

$$\begin{aligned}
E\big[qf(X+1) - Y_i f(X)\big] &= E\big[qf(X_i+Y_i+1) - Y_i f(X_i+Y_i)\big] \\
&= E\big\{E\big[(qf(X_i+Y_i+1) - Y_i f(X_i+Y_i)) \mid Y_i\big]\big\} \\
&= E\big[(qf(X_i+Y_i+1) - Y_i f(X_i+Y_i)) \mid Y_i = 0\big]P(Y_i=0) \\
&\quad + E\big[(qf(X_i+Y_i+1) - Y_i f(X_i+Y_i)) \mid Y_i = 1\big]P(Y_i=1) \\
&\quad + \sum_{k\geq 2} E\big[(qf(X_i+Y_i+1) - Y_i f(X_i+Y_i)) \mid Y_i = k\big]P(Y_i=k) \\
&= E\big[(qf(X_i+Y_i+1) - Y_i f(X_i+Y_i)) \mid Y_i = 0\big]p \\
&\quad + E\big[(qf(X_i+Y_i+1) - Y_i f(X_i+Y_i)) \mid Y_i = 1\big]pq \\
&\quad + \sum_{k\geq 2} E\big[(qf(X_i+Y_i+1) - Y_i f(X_i+Y_i)) \mid Y_i = k\big]pq^k \\
&= E\big[pqf(X_i+1)\big] + E\Big[pq^2 f(X_i+2) - pqf(X_i+1)\Big] \\
&\quad + \sum_{k\geq 2} E\Big[pq^{k+1} f(X_i+k+1) - kpq^k f(X_i+k)\Big] \\
&= \sum_{k\geq 2} E\Big[pq^k f(X_i+k) - kpq^k f(X_i+k)\Big] \\
&= \sum_{k\geq 2} (1-k)pq^k E\big[f(X_i+k)\big] \qquad (3.3)
\end{aligned}$$

จากสมการ (3.3) เราจะได้

$$E\big[qf(X+1) - Y_i f(X)\big] \leq 0 \qquad (3.4)$$

และจากสมการ (3.3) และโดยใช้บทตั้ง 2.1 เราจะได้

$$\begin{aligned}
E\big[qf(X+1) - Y_i f(X)\big] &= -\sum_{k\geq 2} (k-1)pq^k E\big[f(X_i+k)\big] \\
&\geq -\sum_{k\geq 2} (k-1)pq^k \sup_{x\geq k} f(x) \\
&\geq -\lambda^{-1}(e^{\lambda}-1)\mathcal{P}(x_0;\lambda)\sum_{k\geq 2}(k-1)pq^k \min\left\{\frac{1}{k}, \frac{1}{x_0+1}\right\} \\
&\geq -\lambda^{-1}(e^{\lambda}-1)\mathcal{P}(x_0;\lambda)\min\left\{q^2, \frac{q^2}{p(x_0+1)}\right\} \qquad (3.5)
\end{aligned}$$

โดยสมการ (3.2) และอสมการ (3.4) และ (3.5) เราจะได้

$$-(e^{\lambda}-1)\mathcal{P}(x_0;\lambda)\min\left\{q,\frac{q}{p(x_0+1)}\right\}\leq \mathcal{NB}(x_0;n,p)-\mathcal{P}(x_0;\lambda)\leq 0 \tag{3.6}$$

หารตลอดอสมการ (3.6) ด้วย $\mathcal{NB}(x_0;n,p)$ จะได้

$$-(e^{\lambda}-1)\frac{\mathcal{P}(x_0;\lambda)}{\mathcal{NB}(x_0;n,p)}\min\left\{q,\frac{q}{p(x_0+1)}\right\}\leq 1-\frac{\mathcal{P}(x_0;\lambda)}{\mathcal{NB}(x_0;n,p)}\leq 0$$

ดังนั้นโดยบทตั้ง 2.2 เราจึงได้อสมการ (3.1) ตามที่ต้องการ □

บทแทรก (corollary) ต่อไปนี้เป็นผลลัพธ์ที่หาได้จากการหารอสมการ (3.6) (ของการพิสูจน์ทฤษฎีบท 3.1) ด้วย $\mathcal{P}(x_0;\lambda)$

บทแทรก 3.1 *สำหรับ* $x_0\in\mathbb{N}\cup\{0\}$ *และ* $\lambda=nq$ *แล้วจะได้ว่า*

$$-(e^{\lambda}-1)\min\left\{q,\frac{q}{p(x_0+1)}\right\}\leq \frac{\mathcal{NB}(x_0;n,p)}{\mathcal{P}(x_0;\lambda)}-1\leq 0 \tag{3.7}$$

ข้อสังเกต

1. เราสังเกตได้ว่าขอบเขตล่างและบนของความคลาดเคลื่อนสัมพัทธ์ระหว่างฟังก์ชัน $\mathcal{NB}(x_0;n,p)$ และ $\mathcal{P}(x_0;\lambda)$ คือ $-\frac{1-e^{-\lambda}}{p^n}\min\left\{q,\frac{q}{p(x_0+1)}\right\}$ และ 0 ตามลำดับ

2. ถ้า q มีค่าเข้าใกล้ 0 (หรือ p มีค่าเข้าใกล้ 1) หรือ λ มีค่าน้อย แล้วขอบเขตล่างของอสมการ (3.1) และ (3.4) จะมีค่าเข้าใกล้ 0 ซึ่งแสดงให้เห็นว่าผลลัพธ์ของการประมาณฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มทวินามนิเสธด้วยฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มปัวซงจะมีความคลาดเคลื่อนน้อยลง หรือมีความถูกต้องมากขึ้น เมื่อ q มีค่าเข้าใกล้ 0 หรือ λ มีค่าน้อย

3. เมื่อเปรียบเทียบผลลัพธ์ในอสมการ (3.1) และผลลัพธ์ของ คณินทร์ ธีรภาพโอฬาร และคณะ (2551) ในอสมการ (1.1) เราจะเห็นว่าขอบเขตบนของอสมการ (3.1) ดีกว่าอสมการ (1.1) นอกจากนี้ขอบเขตในอสมการ (3.1) จะดีกว่าขอบเขตในอสมการ (1.1) ในทุกกรณีเมื่อ

$$\frac{q(1-e^{-nq})}{p^n}\leq\frac{1-e^{-nq/p}}{p^{n+\min\{x_0,\lceil\lambda\rceil+1\}}}\frac{1-p^n}{n}$$

4. ตัวอย่างเชิงตัวเลข

สำหรับหัวข้อนี้เราได้แสดงผลลัพธ์เชิงตัวเลขบางส่วนของการประมาณฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มทวินามนิเสธด้วยฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มปัวซงในทฤษฎีบท 3.1 ดังตัวอย่างต่อไปนี้

4.1. $n = 100$ และ $p = 0.999$ ดังนั้น $\lambda = 0.1$ และผลลัพธ์เชิงตัวเลขที่ได้ คือ

$$0 \ge 1 - \frac{\mathcal{P}(x_0;\lambda)}{\mathcal{NB}(x_0;n,p)} \ge \begin{cases} -0.00010518 & , x_0 = 0 \\ \dfrac{-0.00010528}{x_0+1} & , x_0 \ge 1 \end{cases}$$

4.2. $n = 500$ และ $p = 0.999$ ดังนั้น $\lambda = 0.5$ และผลลัพธ์เชิงตัวเลขที่ได้ คือ

$$0 \ge 1 - \frac{\mathcal{P}(x_0;\lambda)}{\mathcal{NB}(x_0;n,p)} \ge \begin{cases} -0.00064888 & , x_0 = 0 \\ \dfrac{-0.00064953}{x_0+1} & , x_0 \ge 1 \end{cases}$$

4.3. $n = 100$ และ $p = 0.99$ ดังนั้น $\lambda = 1.0$ และผลลัพธ์เชิงตัวเลขที่ได้ คือ

$$0 \ge 1 - \frac{\mathcal{P}(x_0;\lambda)}{\mathcal{NB}(x_0;n,p)} \ge \begin{cases} -0.01726953 & , x_0 = 0 \\ \dfrac{-0.01744397}{x_0+1} & , x_0 \ge 1 \end{cases}$$

4.4. $n = 250$ และ $p = 0.99$ ดังนั้น $\lambda = 2.5$ และผลลัพธ์เชิงตัวเลขที่ได้ คือ

$$0 \ge 1 - \frac{\mathcal{P}(x_0;\lambda)}{\mathcal{NB}(x_0;n,p)} \ge \begin{cases} -0.11324103 & , x_0 = 0 \\ \dfrac{-0.11438488}{x_0+1} & , x_0 \ge 1 \end{cases}$$

ตัวอย่าง 4.1 และ 4.2 เราได้กำหนดค่าพารามิเตอร์ n และ p เช่นเดียวกับตัวอย่างของ คณินทร์ ธีรภาพโอฬาร และคณะ (2551) และเมื่อเปรียบเทียบผลลัพธ์เชิงตัวเลขของทั้งสองตัวอย่างนี้และผลลัพธ์ของ คณินทร์ ธีรภาพโอฬาร และคณะ (2551) (เปรียบเทียบเฉพาะขอบเขตล่าง) เราพบว่าขอบเขตล่างที่ได้จากทั้งสองตัวอย่างนี้จะดีกว่าขอบเขตล่างของ คณินทร์ ธีรภาพโอฬาร และคณะ (2551) สำหรับทุก $x_0 \ge 1$ และในกรณีที่ p เป็นค่าคงตัว (พิจารณาจากตัวอย่างทั้งหมด) เราจะเห็นว่าขอบเขตของความคลาดเคลื่อนของการประมาณจะกว้างขึ้นตามค่า n หรือ λ ที่เพิ่มขึ้น แสดงว่า λ ที่มีค่าน้อยจะให้ผลของการประมาณดีกว่า λ ที่มีค่ามาก

5. สรุปผลการวิจัย

งานวิจัยนี้ได้ศึกษาการประมาณฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มทวินามนิเสธที่มีพารามิเตอร์ n และ p ด้วยฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มปัวซงที่มีค่าเฉลี่ย λ (ที่สอดคล้องกับ n และ p) ในรูปแบบขอบเขตล่างและบนของความคลาดเคลื่อนสัมพัทธ์ระหว่างฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มทั้งสองและในการศึกษาการประมาณดังกล่าวเรากำหนดให้ค่าเฉลี่ย $\lambda = nq$ ซึ่งแตกต่างจากค่าเฉลี่ย λ ในงานวิจัยของ คณินทร์ ธีรภาพโอฬาร และคณะ (2551) จากผลการวิจัยเราพบว่าผลลัพธ์ของการประมาณปัวซงนี้จะมีความถูกต้องและสามารถยอมรับได้ ถ้าขอบเขตล่างของความคลาดเคลื่อนสัมพัทธ์มีค่าเข้าใกล้ 0 หรืออาจกล่าวได้ว่า ผลลัพธ์ของการประมาณปัวซงที่มีค่าเฉลี่ย $\lambda = nq$ จะให้ผลการประมาณที่ดีและมีความถูกต้องเมื่อ q หรือ λ มีค่าน้อย ซึ่งสอดคล้องกับกฎของเลขจำนวนน้อย (the law of small number) และเมื่อเปรียบเทียบกับผลลัพธ์ของงานวิจัยนี้และผลลัพธ์ของ คณินทร์ ธีรภาพโอฬาร และคณะ (2551) เราพบว่าขอบเขตของความคลาดเคลื่อนของการประมาณในงานวิจัยนี้จะดีกว่าในทุกกรณีเมื่อ

$$\frac{q(1-e^{-nq})}{p^n} \le \frac{1-e^{-nq/p}}{p^{n+\min\{x_0,\lceil\lambda\rceil+1\}}}\frac{1-p^n}{n}$$

เอกสารอ้างอิง

คณินทร์ ธีรภาพโอฬาร, พัชรี วงษ์เกษม และเสาวรส ศรีสุข. (2551). ความคลาดเคลื่อนสัมพัทธ์ของการประมาณการแจกแจงทวินามนิเสธด้วยการแจกแจงปัวซง. วารสารวิทยาศาสตร์บูรพา. 13 : 26-32.

Barbour, A.D., Holst, L., and Janson, S. (1992). Poisson approximation. Oxford Studies in probability 2. Oxford : Clarendon Press.

Feller, W. (1968). An Introduction to Probability Theory and Its Applications. (vol. 1). New York : Wiley.

Johnson, N.L., Kotz, S., and Kemp, A.W. (2005). Univariate Discrete Distributions. (3rd ed.). New York : Wiley.

Stein, C.M. (1986). Approximate Computation of Expectations. Hayward California : IMS.

Teerapabolarn, K. (2007). A bound on the Poisson-binomial relative error. Statistical Methodology. 4 : 407-415.

❑❑❑❑❑