ความคลาดเคลื่อนสัมพัทธ์ของการประมาณการแจกแจงทวินามนิเสธด้วยการแจกแจงปัวซง

The Relative Error of Poisson Approximation to the Negative Binomial Distribution

คณินทร์ ธีรภาพโอฬาร*, พัชรี วงษ์เกษม และเสาวรส ศรีสุข

ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ มหาวิทยาลัยบูรพา บางแสน ชลบุรี 20131

Kanint Teerapabolarn, Patcharee Wongkasem and Saowaros Srisook

Department of Mathematics, Faculty of Science, Burapha University, Chonburi 20131, Thailand

บทคัดย่อ

งานวิจัยนี้ได้ใช้ฟังก์ชัน w (w-function) ร่วมกับเอกลักษณ์สไตน์เชน (Stein-Chen identity) ในการหาขอบเขตบน (upper bound) ของความคลาดเคลื่อนสัมพัทธ์สำหรับการประมาณการแจกแจงทวินามนิเสธ (ที่มีพารามิเตอร์ n และ p) ด้วยการแจกแจงปัวซง (ที่มีค่าเฉลี่ย $\lambda = \frac{n(1-p)}{p}$) โดยอยู่ในรูปแบบของฟังก์ชันการแจกแจงสะสมนอกจากนี้เรายังสามารถหาขอบเขต (bound) ของอัตราส่วนระหว่างฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มทวินามนิเสธ และตัวแปรสุ่มปัวซง ขอบเขตทั้งสองแบบข้างต้นได้ชี้ให้เห็นว่าแต่ละผลลัพธ์จะให้ผลการประมาณปัวซงที่ดี เมื่อ $q = 1-p$ หรือ λ มีค่าน้อย ในตอนท้ายของงานวิจัยเราได้ยกตัวอย่างเชิงตัวเลขเพื่อแสดงการประยุกต์ใช้ผลลัพธ์ที่ได้มาทั้งหมด

คำสำคัญ : การประมาณปัวซง; การแจกแจงทวินามนิเสธ; เอกลักษณ์สไตน์เชน; ฟังก์ชัน w

Abstract

This paper uses the w-function and the Stein-Chen identity to determine an upper bound of the relative error for approximating the negative binomial distribution with parameters n and p by the Poisson distribution with mean $\lambda = \frac{n(1-p)}{p}$ in the form of a cumulative distribution function. Furthermore, we also give a bound of the ratio between the negative binomial and Poisson cumulative distribution functions. For these upper bounds, it is pointed out that each result yields a good Poisson approximation if $q = 1-p$ or λ is small. Finally, we give some numerical examples to illustrate applications of all results obtained.

Key Words : Poisson approximation; negative binomial distribution; Stein-Chen identity; w-function

***Corresponding author.** E-mail: kanint@buu.ac.th

บทนำ

การแจกแจงทวินามนิเสธ (negative binomial distribution) ที่มีพารามิเตอร์ n และความน่าจะเป็น p สามารถพิจารณาในรูปของการแจกแจงของจำนวนความล้มเหลว (failure) ก่อนความสำเร็จ (success) ครั้งที่ n ในลำดับของการทดลองย่อยแบร์นูลลี (Bernoulli trials) ที่เป็นอิสระต่อกัน โดยที่ความสำเร็จและความล้มเหลวที่เกิดขึ้นในแต่ละการทดลองย่อยมีความน่าจะเป็นเท่ากับ p และ q = 1-p ตามลำดับ ในกรณีที่ n = 1 เราจะเรียกการแจกแจงทวินามนิเสธว่า การแจกแจงเรขาคณิต (geometric distribution) ซึ่งเป็นการแจกแจงของจำนวนความล้มเหลวก่อนความสำเร็จครั้งแรก

การแจกแจงทวินามนิเสธสามารถนำไปประยุกต์ใช้ในหลายๆ ด้าน ได้แก่ การวิเคราะห์ข่ายงาน การโทรคมนาคม การวิเคราะห์สินค้าคงคลัง และทางด้านพันธุศาสตร์ของประชากร (สำหรับการประยุกต์ในด้านอื่นๆ สามารถหาดูได้ใน Johnson *et al.* (2005) นอกจากนี้ยังมีการประยุกต์อีกมากที่เกี่ยวข้องกับการแจกแจงเรขาคณิต และก่อนที่เราจะประมาณการแจกแจงทวินามนิเสธด้วยการแจกแจงปัวซง เราจะให้นิยามของการแจกแจงของตัวแปรสุ่มทวินามนิเสธไว้ดังนี้

นิยาม 1.1 ให้ X เป็นจำนวนความล้มเหลวก่อนความสำเร็จครั้งที่ n ในลำดับของการทดลองย่อยแบร์นูลลีที่เป็นอิสระต่อกันด้วยความน่าจะเป็นของความสำเร็จ p (0 < p < 1) แล้วการแจกแจงความน่าจะเป็นของ X สามารถเขียนได้เป็น

$$p(k) = P(X = k) = \binom{n+k-1}{k} q^k p^n \quad \text{เมื่อ } k = 0, 1, \ldots$$

การแจกแจงนี้เรียกว่า "การแจกแจงทวินามนิเสธที่มีพารามิเตอร์ n และ p" (Johnson *et al.*, 2005)

เป็นที่ทราบกันโดยทั่วไปว่า ในกรณีที่ q มีค่าน้อย (มีค่าเข้าใกล้ 0) หรือ p มีค่ามาก (มีค่าเข้าใกล้ 1) การแจกแจงของตัวแปรสุ่มทวินามนิเสธสามารถประมาณได้ด้วยการแจกแจงปัวซงที่มีค่าเฉลี่ยเป็น $\lambda = \frac{nq}{p}$ (ค่าเฉลี่ยของตัวแปรสุ่มทวินามนิเสธ) ซึ่งทำให้ฟังก์ชันการแจกแจง สะสมของตัวแปรสุ่มทวินามนิเสธเขียนแทนด้วย $NB(k;n,p) = \sum_{j=0}^{k} \binom{n+j-1}{j} q^j p^n$ เมื่อ k = 0, 1,... สามารถประมาณได้ด้วยฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มปัวซง เขียนแทนด้วย $P(k;\lambda) = \sum_{j=0}^{k} \frac{e^{-\lambda}\lambda^j}{j!}$ เมื่อ k= 0, 1,... และเกณฑ์ที่ใช้วัดความถูกต้องของการประมาณการแจกแจงทวินามนิเสธด้วยการแจกแจงปัวซงในงานวิจัยนี้ คือ ขอบเขตบนของความคลาดเคลื่อนสัมพัทธ์ และขอบเขตของอัตราส่วนระหว่างฟังก์ชันการแจกแจงสะสมของทั้งสองตัวแปรสุ่มดังกล่าว ความคลาดเคลื่อนสัมพัทธ์และอัตราส่วนระหว่างฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มทวินาม นิเสธและตัวแปรสุ่มปัวซงที่ x_0 เราจะเขียนแทนด้วย $\left|1 - \frac{P(x_0;\lambda)}{NB(x_0;n,p)}\right|$ และ $\frac{NB(x_0;n,p)}{P(x_0;\lambda)}$ ตามลำดับ

วิธีดำเนินการวิจัย

ศึกษาแนวความคิดของความคลาดเคลื่อนสัมพัทธ์ และอัตราส่วนของการประมาณการแจกแจงทวินามด้วยการแจกแจงปัวซงจากงานวิจัยของ Teerapabolarn (2007) และศึกษาคุณสมบัติของฟังก์ชัน w ที่จะนำมาใช้ร่วมกับเอกลักษณ์สไตน์เชนจากงานวิจัยของ Majsnero-wska (1998) และ Cacoullos and Papathanasion (1989) ซึ่งจากการศึกษางานวิจัยดังกล่าวข้างต้นเราควรเริ่มต้นด้วยการสร้างบทตั้ง (lemma) เพื่อช่วยในการพิสูจน์ทฤษฎีบทหลัก (main theorem) ของงานวิจัยได้ดังนี้

พิจารณาเอกลักษณ์สไตน์เชนหรือเอกลักษณ์ของสไตน์ (Stein identity) (Stein, 1986) สำหรับการแจกแจงปัวซงที่มีพารามิเตอร์ λ ซึ่ง (กำหนดฟังก์ชัน h) นิยามโดย

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

โดยที่ $x \in \mathbb{N} \cup \{0\}$, $P_\lambda(h) = e^{-\lambda} \sum_{\ell=0}^{\infty} h(\ell) \frac{\lambda^\ell}{\ell!}$ และ f และ h เป็นฟังก์ชันค่าจริงที่มีขอบเขตที่นิยามบนเซต $\mathbb{N} \cup \{0\}$ ให้ $h_{x_0} : \mathbb{N} \cup \{0\} \to \mathbb{R}$ เป็นฟังก์ชันที่กำหนดโดย

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

โดยที่ $x \in \mathbb{N} \cup \{0\}$ ดังนั้นเราจะได้ผลเฉลย f เมื่อแทน h ด้วย h_{x_0} ในสมการ (2.1) (Barbour *et al.*, 1992) ดังสมการต่อไปนี้

$$
f(x)=\begin{cases}(x-1)!\lambda^{-x}e^{\lambda}[P_{\lambda}(h_{x_0})P_{\lambda}(1-h_{x-1})]\ , x_0<x\\ (x-1)!\lambda^{-x}e^{\lambda}[P_{\lambda}(h_{x-1})P_{\lambda}(1-h_{x_0})]\ , x_0\ge x\\ 0 \qquad\qquad\qquad\qquad\qquad\qquad , x=0\end{cases} \tag{2.2}
$$

บทตั้ง 2.1 ให้ $\Delta f(x) = f(x+1)-f(x)$, $x_0 \in \mathbb{N} \cup \{0\}$ และ $k \in \mathbb{N}$ แล้วจะได้ว่า

$$
\sup_{x\ge k}|\Delta f(x)| \le \lambda^{-1}(e^{\lambda}-1)P(x_0;\lambda)\min\left\{\frac{1}{k},\frac{1}{x_0+1}\right\} \tag{2.3}
$$

พิสูจน์ ดูพิสูจน์ได้จากบทตั้ง 2.1 ของ Teerapabolarn (2007) ร่วมกับบทตั้ง 2.2 ของ Teerapabolarn and Wongkasem (2007) #

บทตั้ง 2.2 ให้ $x_0 \in \mathbb{N} \cup \{0\}$ และ $\lambda = \frac{nq}{p}$ แล้วเราจะได้

$$
\frac{P(x_0;\lambda)}{NB(x_0;n,p)} \le \frac{e^{-\lambda}}{p^{n+\min\{x_0,\lceil\lambda\rceil+1\}}} \tag{2.4}
$$

โดยที่ $\lceil\lambda\rceil$ เป็นจำนวนเต็มที่น้อยที่สุดที่ไม่น้อยกว่า λ

พิสูจน์ เราเห็นได้ชัดในกรณี $x_0 = 0$ จึงเป็นการเพียงพอที่จะแสดงว่าสมการ (2.4) เป็นจริงในกรณี $x_0 \ge 1$ เราจะเห็นว่า

$$
\begin{aligned}
\sum_{k=0}^{x_0}\binom{n+k-1}{k}q^k &= 1+\sum_{k=1}^{x_0}\frac{(n+k-1)!q^k}{k!(n-1)!}\\
&= 1+\sum_{k=1}^{x_0}\frac{(n+k-1)\cdots nq^k}{k!}\\
&= 1+\sum_{k=1}^{x_0}\frac{\left[np^{-1}q\right]^k p^k}{k!}\\
&\quad \times\left\{\frac{(n+k-1)}{n}\cdots\frac{n}{n}\right\}\\
&\ge 1+\sum_{k=1}^{x_0}\frac{\lambda^k p^{x_0}}{k!}\\
&\ge p^{x_0}\sum_{k=0}^{x_0}\frac{\lambda^k}{k!}
\end{aligned} \tag{2.5}
$$

และในกรณีที่ $x_0 \ge \lceil\lambda\rceil+1$ เราจะได้ว่า $x_0-1 \ge \lceil\lambda\rceil \ge \lambda = np^{-1}(1-p) = np^{-1}-n$ ซึ่งทำให้ $\frac{n+x_0-1}{n} \ge p^{-1}$ ดังนั้น

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

และ

$$
\begin{aligned}
&\sum_{k=0}^{x_0}\binom{n+k-1}{k}q^k\\
&\ge p^{\lceil\lambda\rceil+1}\left\{1+\lambda+\cdots+\frac{\lambda^{\lceil\lambda\rceil+1}}{(\lceil\lambda\rceil+1)!}+\cdots+\frac{\lambda^{x_0}}{x_0!}\right\}\\
&\ge p^{\lceil\lambda\rceil+1}\sum_{k=0}^{x_0}\frac{\lambda^k}{k!}
\end{aligned} \tag{2.6}
$$

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

ดังนั้นโดยใช้สมการ (2.5) และ (2.6) เราจะได้ว่า

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

และ

$$\frac{P(x_0;\lambda)}{NB(x_0;n,p)} \le \frac{e^{-\lambda}}{p^n p^{\lceil \lambda \rceil + 1}}$$

เมื่อ $x_0 \ge \lceil \lambda \rceil + 1$ ซึ่งทำให้เราได้อสมการ (2.4) #

เมื่อเราพิจารณาแนวคิดของการประมาณปัวซงโดยใช้ฟังก์ชัน w จากงานวิจัยของ Majsnerowska (1998) เราพบว่า Majsnerowska (1998) ได้ปรับและประยุกต์รูปแบบความสัมพันธ์ของฟังก์ชัน w ที่สัมพันธ์กับตัวแปรสุ่มที่มีค่าจำนวนเต็มไม่เป็นลบ (non-negative integer-valued random variable) X ที่ปรากฏในงานวิจัยของ Cacoullos and Papathanasiou (1989) ให้อยู่ในรูปของความสัมพันธ์เวียนเกิด (recurrence relation) ดังนี้

$$w(k+1) = \frac{p(k)}{p(k+1)} w(k) - \frac{\mu - (k+1)}{\sigma^2}$$

$$\ge 0 \text{ เมื่อ } k = 0, 1, \ldots \qquad (2.7)$$

โดยที่ $w(0) = \frac{\mu}{\sigma^2}$ μ และ σ^2 แทนค่าเฉลี่ยและความแปรปรวนของตัวแปรสุ่ม X

บทตั้งต่อไปนี้เป็นผลที่ได้มาจากการประยุกต์ใช้สมการเวียนเกิด (2.7)

บทตั้ง 2.3 ให้ X เป็นตัวแปรสุ่มทวินามนิเสธที่มีการแจกแจงความน่าจะเป็นเหมือนในนิยาม 1.1 และ w(X) เป็นฟังก์ชัน w ที่สัมพันธ์กับตัวแปรสุ่ม X แล้ว

$$w(k) = \frac{(n+k)q}{\sigma^2 p} \quad \text{เมื่อ } k = 0, 1, \ldots \qquad (2.8)$$

โดยที่ $\sigma^2 = \frac{nq}{p^2}$

พิสูจน์ แทน $\mu = \frac{nq}{p}$ ใน (2.7) และจัดรูปแบบความสัมพันธ์เวียนเกิดใหม่เราจะได้

$$w(k) = \frac{nq}{\sigma^2 p} + w(k-1)\frac{p(k-1)}{p(k)} - \frac{k}{\sigma^2}$$

$$\text{เมื่อ } k = 1, 2, \ldots \qquad (2.9)$$

โดยที่ $w(0) = \frac{nq}{\sigma^2 p}$

เราจะแสดงว่า (2.8) เป็นจริงสำหรับทุก $k \in \mathbb{N} \cup \{0\}$ จากสมการ (2.9) เราจะได้

$$w(1) = \frac{nq}{\sigma^2 p} + w(0)\frac{p(0)}{p(1)} - \frac{1}{\sigma^2}$$

$$= \frac{nq}{\sigma^2 p} + \frac{nq}{\sigma^2 p}\frac{1}{nq} - \frac{1}{\sigma^2} = \frac{(n+1)q}{\sigma^2 p}$$

สมมติว่า $w(i) = \frac{(n+i)q}{\sigma^2 p}$ ต่อไปเราจะแสดงว่า

$w(i+1) = \frac{[n+(i+1)]q}{\sigma^2 p}$ จาก (2.9) เราจะได้

$$w(i+1) = \frac{nq}{\sigma^2 p} + w(i)\frac{p(i)}{p(i+1)} - \frac{i+1}{\sigma^2}$$

$$= \frac{nq}{\sigma^2 p} + \frac{(n+i)q}{\sigma^2 p}\frac{i+1}{(n+i)q} - \frac{i+1}{\sigma^2}$$

$$= \frac{[n+(i+1)]q}{\sigma^2 p}$$

ดังนั้นโดยวิธีอุปนัยเชิงคณิตศาสตร์ (mathematical induction) เราจะได้ (2.8) เป็นจริงสำหรับทุก $k \in \mathbb{N} \cup \{0\}$ #

ผลการวิจัย

ผลการวิจัยที่เราต้องการจากงานวิจัยนี้ คือ ขอบเขตบนของความคลาดเคลื่อนสัมพัทธ์ และขอบเขตของอัตราส่วนระหว่างฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มทวินามนิเสธและตัวแปรสุ่มปัวซง ผลการวิจัยดังกล่าวสามารถหาได้โดยการใช้คุณสมบัติเฉพาะของฟังก์ชัน w ร่วมกับเอกลักษณ์สไตน์เชน (2.1) คุณสมบัติเฉพาะของ ฟังก์ชัน w ที่สัมพันธ์กับตัวแปรสุ่ม X ได้กล่าวไว้ใน Cacoullos and Papathanasiou (1989) ดังนี้

"ถ้าฟังก์ชัน g สอดคล้องกับเงื่อนไข $E|w(X)\Delta g(X)| < \infty$ และ $E|(X-\mu)g(X)| < \infty$ แล้วจะได้ว่า $Cov(X,g(X)) = \sigma^2 E[w(X)\Delta g(X)]$" เมื่อ $\Delta g(x) = g(x+1) - g(x)$

ทฤษฎีบทต่อไปนี้เป็นผลลัพธ์ของการประมาณการแจกแจง (ผลการวิจัย) ที่อยู่ในรูปแบบของขอบเขตบนของความคลาดเคลื่อนสัมพัทธ์และขอบเขตของอัตราส่วนระหว่างฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มทวินามนิเสธและตัวแปรสุ่มปัวซง

ทฤษฎีบท 3.1 ให้ X เป็นตัวแปรสุ่มทวินามนิเสธที่มีการแจกแจงความน่าจะเป็นเหมือนในนิยาม 1.1 $x_0 \in \mathbb{N} \cup \{0\}$ และ $\mu = \lambda = \frac{nq}{p} < \infty$ แล้วอสมการต่อไปนี้เป็นจริง

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

$$\times \min\left\{\frac{(1-p^n)}{n}, \frac{q}{p(x_0+1)}\right\} \quad (3.1)$$

พิสูจน์ เราจะแสดงว่าอสมการ (3.1) เป็นจริง จากสมการ (2.1) เมื่อกำหนด h = h_{x0} เราจะได้

$$P(X \le x_0) - P(x_0;\lambda) = E[\lambda f(X+1) - Xf(X)]$$

หรือ

$$|NB(x_0;n,p) - P(x_0;\lambda)| = |E[\lambda f(X+1) - Xf(X)]| \quad (3.2)$$

โดยที่ f นิยามเช่นเดียวกับสมการ (2.2) จากสมการ (3.2) เราจะได้

$$\begin{aligned}
&|NB(x_0;n,p) - P(x_0;\lambda)| \\
&= |\lambda E[f(X+1)] - E[Xf(X)]| \\
&= |\lambda E[f(X+1)] - Cov(X,f(X)) - \mu E[f(X)]| \\
&= |E[\lambda f(X+1) - \mu f(X)] - Cov(X,f(X))| \\
&= |\lambda E[\Delta f(X)] - Cov(X,f(X))|
\end{aligned}$$

เนื่องจาก $E|w(X)\Delta f(X)| < \infty$ และ $\sup_{x\ge1} f(x) \le \min\{1,\lambda^{-1/2}\}$ (Barbour *et al.*, 1992) ซึ่งทำให้

$$\begin{aligned}
&|NB(x_0;n,p) - P(x_0;\lambda)| \\
&= |\lambda E[\Delta f(X)] - \sigma^2 E[w(X)\Delta f(X)]| \\
&= \left|E\left\{\left[\lambda - \sigma^2 w(X)\right]\Delta f(X)\right\}\right| \\
&\le E\left|\left\{\left[\lambda - \sigma^2 w(X)\right]\Delta f(X)\right\}\right| \\
&= \sum_{k=0}^{\infty}\left|\left[\lambda - \sigma^2 w(k)\right]\Delta f(k)\right|p(k) \quad \text{(โดยบทตั้ง 2.3)} \\
&= \sum_{k=0}^{\infty}\left|\left[\frac{nq}{p} - \frac{(n+k)q}{p}\right]\Delta f(k)\right|p(k) \\
&= \frac{q}{p}\sum_{k=1}^{\infty} k|\Delta f(k)|p(k) \\
&\le \frac{q}{p}\sum_{k=1}^{\infty} k \sup_{x\ge k}|\Delta f(x)|p(k) \\
&\le \lambda^{-1}(e^{\lambda}-1)P(x_0;\lambda)\frac{q}{p}\sum_{k=1}^{\infty} kp(k) \\
&\times \min\left\{\frac{1}{k}, \frac{1}{x_0+1}\right\} \quad \text{(โดยบทตั้ง 2.1)} \qquad (3.3)
\end{aligned}$$

เนื่องจาก

$$\begin{aligned}
&\frac{q}{p}\sum_{k=1}^{\infty} kp(k)\min\left\{\frac{1}{k}, \frac{1}{x_0+1}\right\} \\
&= \frac{q}{p}\sum_{k=1}^{\infty}\min\left\{p(k), \frac{kp(k)}{x_0+1}\right\} \\
&\le \frac{q}{p}\min\left\{\sum_{k=1}^{\infty}p(k), \frac{1}{x_0+1}\sum_{k=1}^{\infty}kp(k)\right\} \\
&= \frac{q}{p}\min\left\{1-p(0), \frac{\mu}{x_0+1}\right\} \\
&= \frac{q}{p}\min\left\{1-p^n, \frac{\lambda}{x_0+1}\right\} \qquad (3.4)
\end{aligned}$$

ดังนั้นโดยการแทนอสมการ (3.4) ลงในอสมการ (3.3) เราจะได้

$$\begin{aligned}
&|NB(x_0;n,p) - P(x_0;\lambda)| \\
&\le \lambda^{-1}(e^{\lambda}-1)P(x_0;\lambda)\frac{q}{p}\min\left\{1-p^n, \frac{\lambda}{x_0+1}\right\} \\
&= (e^{\lambda}-1)P(x_0;\lambda)\min\left\{\frac{1-p^n}{n}, \frac{q}{p(x_0+1)}\right\} \qquad (3.5)
\end{aligned}$$

หารอสมการ (3.5) ด้วย NB(x_0; n, p) ซึ่งจะได้

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

อสมการสุดท้ายได้มาโดยบทตั้ง 2.2 ดังนั้นอสมการ (3.1) เป็นจริง #

ทฤษฎีบท 3.2 ให้ X เป็นตัวแปรสุ่มทวินามนิเสธที่มีการแจกแจงความน่าจะเป็นเหมือนในนิยาม 1.1 และ $x_0 \in \mathbb{N} \cup \{0\}$ แล้วเราจะได้

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

$$\le 1+(e^{\lambda}-1)\min\left\{\frac{1-p^n}{n}, \frac{q}{p(x_0+1)}\right\} \qquad (3.6)$$

พิสูจน์ จากอสมการ (3.5) หารด้วย $P(x_0;\lambda)$ จะได้

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

หรือ

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

$$\le (e^{\lambda}-1)\min\left\{\frac{1-p^n}{n},\frac{q}{p(x_0+1)}\right\} \quad (3.7)$$

บวกตลอดอสมการ (3.7) ด้วย 1 เราจะได้อสมการ (3.6) ตามที่ต้องการ #

ข้อสังเกต

1. ถ้า q มีค่าเข้าใกล้ 0 (หรือ p มีค่าเข้าใกล้ 1) หรือ λ มีค่าน้อย แล้วขอบเขตบนของอสมการ (3.1) จะมีค่าเข้าใกล้ 0 ส่วนขอบเขตของอสมการ (3.4) จะมีค่าเข้าใกล้ 1 ซึ่งแสดงให้เห็นว่าผลลัพธ์ของการประมาณการแจกแจงทวินามนิเสธด้วยการแจกแจงปัวซงจะมีความถูกต้องมากขึ้น เมื่อ q มีค่าเข้าใกล้ 0 หรือ λ มีค่าน้อย

2. เมื่อ λ คงที่ เราจะเห็นว่าขอบเขตของการประมาณการแจกแจงของทั้งสองทฤษฎีจะแคบลงตามค่า n ที่เพิ่มขึ้น ซึ่งในกรณีนี้ เมื่อ n มีค่ามากจะทำให้การประมาณปัวซงได้ผลดี

ตัวอย่างเชิงตัวเลข

ในหัวข้อนี้เราได้ยกตัวอย่างเชิงตัวเลข เพื่อแสดงการประยุกต์ใช้ผลลัพธ์ของการประมาณการแจกแจงทวินามนิเสธด้วยการแจกแจงปัวซงที่อยู่ในทฤษฎีบท 3.1 และ 3.2

ตัวอย่าง 4.1 ให้ n = 100 และ p = 0.999 ดังนั้น λ = 0.1001001 และผลลัพธ์เชิงตัวเลขที่ได้มาจากทฤษฎีบท 3.1 และ 3.2 สามารถแสดงได้ดังนี้

$$\left|1-\frac{P(x_0;0.1001001)}{NB(x_0;100,0.999)}\right| \le \begin{cases} 0.00010023\ , & x_0=0 \\ 0.00005274\ , & x_0=1 \\ \dfrac{0.00010559}{x_0+1}, & x_0\ge 2 \end{cases}$$

และ

$$0.99989976 \le \frac{NB(x_0;100,0.999)}{P(x_0;0.1001001)} \le 1.00010024 \qquad , x_0=0$$

$$1-\frac{0.00010539}{x_0+1} \le \frac{NB(x_0;100,0.999)}{P(x_0;0.1001001)} \le 1-\frac{0.00010539}{x_0+1}, x_0\ge 1$$

ตัวอย่าง 4.2 ให้ n = 500 และ p = 0.999 เราจะได้ λ = 0.5005005 และผลลัพธ์เชิงตัวเลขที่หาได้ คือ

$$\left|1-\frac{P(x_0;0.5005005)}{NB(x_0;500,0.999)}\right| \le \begin{cases} 0.00051122\ , & x_0=0 \\ 0.00032534\ , & x_0=1 \\ \dfrac{0.00065134}{x_0+1}, & x_0\ge 2 \end{cases}$$

และ

$$0.99948865 \le \frac{NB(x_0;500,0.999)}{P(x_0;0.5005005)} \le 1.00051135 \qquad , x_0=0$$

$$1-\frac{0.00065020}{x_0+1} \le \frac{NB(x_0;500,0.999)}{P(x_0;0.5005005)} \le 1+\frac{0.00065020}{x_0+1}, x_0\ge 1$$

ตัวอย่าง 4.3 กำหนดให้ n = 5000 และ p = 0.999 แล้วจะได้ λ = 5.00500501 และผลลัพธ์เชิงตัวเลขของการประมาณการแจกแจงเป็นดังนี้

$$\left|1-\frac{P(x_0;5.00500501)}{NB(x_0;5000,0.999)}\right| \le \begin{cases} 0.02947655\ , & x_0=0,1,2,3,4 \\ 0.02477954\ , & x_0=5 \\ 0.02126087\ , & x_0=6 \\ \dfrac{0.14897504}{x_0+1}, & x_0\ge 7 \end{cases}$$

และ

$$0.97056759 \le \frac{NB(x_0;5000,0.999)}{P(x_0;5.00500501)} \le 1.02943241, x_0=0,1,2,3,4$$

$$1-\frac{0.14830614}{x_0+1} \le \frac{NB(x_0;5000,0.999)}{P(x_0;5.00500501)} \le 1+\frac{0.14830614}{x_0+1}, x_0\ge 5$$

ข้อสังเกต

1. จากตัวอย่างเชิงตัวเลขทั้งสามตัวอย่างที่ได้แสดงไปนั้น เราสามารถสังเกตได้ว่าขอบเขตของการประมาณปัวซงจะแคบลง เมื่อ x_0 มีค่าเพิ่มขึ้น หรือขอบเขต ของการประมาณเป็นฟังก์ชันลดสำหรับ x_0

2. เมื่อ p คงที่ เราจะเห็นว่าขอบเขตของการประมาณจะกว้างขึ้นตามค่า n หรือ λ ที่เพิ่มขึ้น แสดงว่า λ ที่มีค่าน้อยจะให้ผลของการประมาณดีกว่า λ ที่มีค่ามาก

สรุปผลการวิจัย

ขอบเขตบนของความคลาดเคลื่อนสัมพัทธ์ และขอบเขตของอัตราส่วนระหว่างฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มทวินามนิเสธที่มีพารามิเตอร์ n และ p และตัวแปรสุ่มปัวซงที่มีพารามิเตอร์ $\lambda = \frac{np}{p}$ (ค่าเฉลี่ยของตัวแปรสุ่มทวินามนิเสธ) เป็นเพียงเกณฑ์แบบหนึ่งที่ใช้วัดความถูกต้องของการประมาณปัวซง ซึ่งจากผลการวิจัยเราพบว่าผลลัพธ์ของการประมาณปัวซงจะมีความถูกต้องและสามารถยอมรับได้ ถ้าขอบเขตบนของความคลาดเคลื่อนสัมพัทธ์มีค่าเข้าใกล้ 0 หรือขอบเขตของอัตราส่วนมีค่าเข้าใกล้ 1 หรืออาจสรุปในลักษณะที่เกี่ยวข้องกับพารามิเตอร์ได้ว่า “ผลลัพธ์ของการประมาณปัวซงแต่ละแบบจะให้ผลการประมาณที่ดีและมีความถูกต้องเมื่อ q หรือ λ มีค่าน้อย” ซึ่งสอดคล้องกับกฎของเลขจำนวนน้อย (The Law of Small Number) ในทฤษฎีความน่าจะเป็น

กิตติกรรมประกาศ

คณะผู้วิจัยขอขอบคุณ คณะวิทยาศาสตร์ มหาวิทยาลัยบูรพา ชลบุรี ที่ได้จัดสรรเงินงบประมาณเงินรายได้ประจำปี 2550 เพื่อสนับสนุนการวิจัยครั้งนี้ และคณะผู้วิจัยขอขอบคุณกรรมการผู้อ่านผลงานวิจัยทุกท่านที่ได้กรุณาให้คำแนะนำและแก้ไขข้อบกพร่องต่าง ๆ จนทำให้ผลงานวิจัยนี้มีความถูกต้องและสมบูรณ์มากยิ่งขึ้น

เอกสารอ้างอิง

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

Cacoullos, T., & Papathanasiou, V. (1989). Characterization of distributions by variance bounds. *Statistics & Probability Letters, 7,* 351-356.

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

Majsnerowska, M. (1998). A note on Poisson approximation by w-functions. *Applicationes Mathematicae, 25,* 387-392.

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.

Teerapabolarn, K., & Wongkasem, P. (2007). Poisson approximation for independent geometric random variables. *International Mathematical Forum, 2,* 3211-3218.