การประมาณทวินามด้วยวิธีของสไตน์และฟังก์ชัน w

BINOMIAL APPROXIMATION WITH STEIN'S METHOD AND w-FUNCTIONS

คุณากร แซ่เจ็ง

มหาวิทยาลัยบูรพา

2560

การประมาณทวินามด้วยวิธีของสไตน์และฟังก์ชัน w

คุณากร แซ่เจ็ง

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตรมหาบัณฑิต
สาขาวิชาสถิติ
คณะวิทยาศาสตร์ มหาวิทยาลัยบูรพา
2560

BINOMIAL APPROXIMATION WITH STEIN'S METHOD AND w-FUNCTIONS

KHUNAKORN SAE-JENG

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE MASTER OF SCIENCE
IN STATISTICS
FACULTY OF SCIENCE
BURAPHA UNIVERSITY
2017

คณะกรรมการควบคุมวิทยานิพนธ์และคณะกรรมการสอบวิทยานิพนธ์ได้พิจารณาวิทยานิพนธ์ของ คุณากร แซ่เจ็ง ฉบับนี้แล้ว เห็นสมควรรับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิชาสถิติ ของมหาวิทยาลัยบูรพาได้

คณะกรรมการควบคุมวิทยานิพนธ์

.... อาจารย์ที่ปรึกษาหลัก
(ผู้ช่วยศาสตราจารย์ ดร. คณินทร์ ธีรภาพโอฬาร)

คณะกรรมการสอบวิทยานิพนธ์

.... ประธาน
(ศาสตราจารย์ ดร. กฤษณะ เนียมมณี)

.... กรรมการ
(ผู้ช่วยศาสตราจารย์ ดร. คณินทร์ ธีรภาพโอฬาร)

.... กรรมการ
(ดร. พัชรี วงษ์เกษม)

คณะวิทยาศาสตร์อนุมัติให้รับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิชาสถิติ ของมหาวิทยาลัยบูรพา

.... คณบดีคณะวิทยาศาสตร์
(ผู้ช่วยศาสตราจารย์ ดร. เอกรัฐ ศรีสุข)
วันที่ 31 เดือน พฤษภาคม พ.ศ. 2561

59910013: สาขาวิชา: สถิติ; วท.ม. (สถิติ)
คำสำคัญ: การประมาณทวินาม, ขอบเขตไม่เอกรูป, วิธีของสไตน์, ฟังก์ชัน w

คุณากร แซ่เจ็ง: การประมาณทวินามด้วยวิธีของสไตน์และฟังก์ชัน *w* (BINOMIAL APPROXIMATION WITH STEIN’S METHOD AND *w*-FUNCTIONS) คณะกรรมการควบคุมวิทยานิพนธ์: คณินทร์ ธีรภาพโอฬาร ปี พ.ศ. 2560

จุดมุ่งหมายของการศึกษาครั้งนี้ คือ หาขอบเขตไม่เอกรูปสำหรับประมาณฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มที่มีค่าจำนวนเต็มไม่เป็นลบด้วยฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มทวินามโดยใช้วิธีของสไตน์และฟังก์ชัน *w* โดยที่ผลการศึกษาอยู่ในรูปของระยะทางของคอลโมโกรอฟระหว่างฟังก์ชันการแจกแจงสะสมของทั้งสองฟังก์ชันพร้อมด้วยขอบเขตไม่เอกรูป ขอบเขตไม่เอกรูปที่ได้จากการศึกษาครั้งนี้สามารถใช้เป็นเกณฑ์วัดความแม่นยำของการประมาณ นั่นคือ ถ้าขอบเขตไม่เอกรูปดังกล่าวมีค่าน้อย แสดงว่าฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มทวินามมีความเหมาะสมที่จะใช้ประมาณฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มที่มีค่าจำนวนเต็มไม่เป็นลบ การประยุกต์ของการศึกษาครั้งนี้เป็นการใช้ผลลัพธ์ที่ได้ไปประมาณฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มบางชนิด ได้แก่ ฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มเรขาคณิตไฮเพอร์ ฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มเรขาคณิตไฮเพอร์เชิงลบ ฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มโพลยา และฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มทวินามบีต้า นอกจากนี้ได้มีการนำเสนอผลลัพธ์เชิงตัวเลขเพื่อแสดงความแม่นยำของการประมาณสำหรับทุกฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มดังที่ได้กล่าวมาข้างต้น

59910013: MAJOR: STATISTICS; M.Sc. (STATISTICS)
KEYWORDS: Binomial approximation, Non-uniform bound, Stein's method, w-functions

KHUNAKORN SAE-JENG: BINOMIAL APPROXIMATION WITH STEIN'S METHOD AND *w*-FUNCTIONS. ADVISORY COMMITTEE: KANINT TEERAPABOLARN, Ph.D. 2017

The aim of this study is to determine a non-uniform bound for approximating the cumulative distribution function of a non-negative integer-valued random variable by the binomial cumulative distribution function using Stein's method and *w*-functions, where the result of the study is in the form of the Kolmogorov distance between the two cumulative distribution functions together with its non-uniform bound. The non-uniform bound obtained in this study can be used as a criterion for measuring the accuracy of the approximation, that is, if the non-uniform is small, it indicates that the binomial cumulative distribution function is appropriate to approximate the cumulative distribution function of a non-negative integer-valued random variable. The applications of this study are a use of the obtained result to approximate some cumulative distribution functions such as the hypergeometric cumulative distribution function, the negative hypergeometric cumulative distribution function, the Pólya cumulative distribution function and the beta-binomial cumulative distribution function. Additionally, some numerical results are also provided to illustrate the accuracy of the approximation for all cumulative distribution functions that mentioned above.

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงได้ด้วยความกรุณาช่วยเหลือและให้คำปรึกษาอย่างดียิ่ง จากผู้ช่วยศาสตราจารย์ ดร.คณินทร์ ธีรภาพโอฬาร อาจารย์ที่ปรึกษาวิทยานิพนธ์ ที่ได้กรุณาถ่ายทอดความรู้ แนวคิด วิธีการ คำแนะนำ และตรวจสอบแก้ไขข้อบกพร่องต่าง ๆ ด้วยความเอาใจใส่ยิ่ง ผู้วิจัยกราบขอบพระคุณเป็นอย่างสูง

ขอขอบพระคุณคณะกรรมการสอบวิทยานิพนธ์ในครั้งนี้ ซึ่งประกอบด้วยศาสตราจารย์ ดร.กฤษณะ เนียมมณี จากจุฬาลงกรณ์มหาวิทยาลัยและอาจารย์ ดร.พัชรี วงษ์เกษม อาจารย์ประจำหลักสูตรที่ได้กรุณาให้คำแนะนำเพิ่มเติม แก้ไขข้อบกพร่อง และตรวจสอบความถูกต้องของวิทยานิพนธ์ฉบับนี้ให้มีความถูกต้องและเสร็จสมบูรณ์มากยิ่งขึ้น

ขอขอบคุณอาจารย์ทุกท่านที่คอยอบรมสั่งสอน ให้วิชาความรู้และคำแนะนำต่างๆ ตลอดระยะเวลาที่ข้าพเจ้าได้ศึกษาอยู่ที่มหาวิทยาลัยบูรพา

ขอขอบคุณน้อง ๆ นิสิตสาขาวิชาสถิติและคณิตศาสตร์ที่คอยเป็นกำลังใจที่ดีตลอดมา และนอกจากนี้ยังมีผู้ที่ให้ความร่วมมือช่วยเหลืออีกหลายท่าน ซึ่งผู้วิจัยไม่สามารถกล่าวนามในที่นี้ได้หมด จึงขอขอบคุณทุกท่านเหล่านั้นไว้ ณ โอกาสนี้ด้วย

คุณค่าทั้งหลายที่ผู้วิจัยได้รับจากวิทยานิพนธ์ฉบับนี้ ผู้วิจัยขอมอบแด่บิดา มารดา ผู้ที่อยู่เบื้องหลังในความสำเร็จครั้งนี้ ตลอดจนผู้มีพระคุณทุกท่าน

คุณากร แซ่เจ็ง

สารบัญ

หน้า

บทที่ 1
บทนำ

ความเป็นมาและความสำคัญของปัญหา

การประมาณทวินาม (Binomial approximation) เป็นการประมาณการแจกแจงของตัวแปรสุ่มไม่ต่อเนื่อง (Discrete random variable) ด้วยการแจกแจงทวินาม (Binomial distribution) โดยมีเกณฑ์ที่นิยมใช้วัดความแม่นยำของการประมาณ ได้แก่ เกณฑ์ที่วัดความแม่นยำโดยใช้ค่าคลาดเคลื่อนสัมบูรณ์ (Absolute error) ของการประมาณโดยตรง และเกณฑ์ที่วัดความแม่นยำโดยใช้ขอบเขตบน (Upper bound) ของการประมาณซึ่งเป็นเกณฑ์ที่นิยมใช้กันอย่างแพร่หลาย ในกรณีที่ใช้ขอบเขตบนเป็นเกณฑ์วัดความแม่นยำของการประมาณ Stein (1986) ได้นำเสนอวิธีของสไตน์ (Stein's method) มาประยุกต์ใช้กับการประมาณการแจกแจงของตัวแปรสุ่มไม่ต่อเนื่องด้วยการแจกแจงทวินามเป็นบุคคลแรก ต่อมา Ehm (1991) ได้ใช้วิธีของสไตน์หาขอบเขตเอกรูป (Uniform bound) สำหรับการประมาณการแจกแจงของผลรวมของตัวแปรสุ่มแบร์นูลลี (Bernoulli random variable) ที่แตกต่างกันและเป็นอิสระต่อกัน (Independent) Barbour et al. (1992) ได้ปรับสมการของสไตน์สำหรับการแจกแจงทวินามใน Stein (1986) ให้อยู่ในรูปแบบที่เหมาะสมเช่นเดียวกับการประมาณปัวซง (Poisson approximation) Soon (1996) ได้ใช้วิธีของสไตน์หาขอบเขตเอกรูปสำหรับการประมาณการแจกแจงของผลรวมของตัวแปรสุ่มแบร์นูลลีที่ไม่เป็นอิสระต่อกัน (Dependent) Wongkasem et al. (2008) ได้ใช้วิธีของสไตน์และฟังก์ชัน w หาขอบเขตเอกรูปสำหรับการประมาณการแจกแจงทวินามทั่วไป (Generalized binomial distribution) ด้วยการแจกแจงทวินาม Prukpousana and Teerapabolarn (2010) ได้ใช้วิธีของสไตน์สำหรับการแจกแจงทวินามและฟังก์ชัน w หาขอบเขตเอกรูปสำหรับการประมาณการแจกแจงเรขาคณิตไฮเพอร์เชิงลบ (Negative hypergeometric distribution) ด้วยการแจกแจงทวินาม และ Teerapabolarn (2011) ได้ใช้วิธีเช่นเดียวกับ Wongkasem et al. (2008) ในการหาขอบเขตไม่เอกรูป (Non-uniform bound) สำหรับการประมาณแบบจุดของการแจกแจงทวินามทั่วไปด้วยการแจกแจงทวินาม จากที่กล่าวมาข้างต้น จะเห็นได้ว่าการแจกแจงทวินามสามารถประมาณการแจกแจงของตัวแปรสุ่มไม่ต่อเนื่องที่เป็นตัวแปรสุ่มตัวเดียวหรือเป็นผลรวมของตัวแปรสุ่มหลายตัว แต่ในกรณีที่ต้องการศึกษาการประมาณการแจกแจงของตัวแปรสุ่มที่มีค่าจำนวนเต็มไม่เป็นลบในรูปทั่วไปเพียงหนึ่งตัวแปร (ไม่เจาะจงชนิดของการแจกแจง) ด้วยการแจกแจงทวินาม Teerapabolarn and Wongkasem (2011) ได้กำหนดตัวแปรสุ่มที่มีค่าจำนวนเต็มไม่เป็นลบมีลักษณะ ดังนี้

ให้ X แทนตัวแปรสุ่มที่มีค่าจำนวนเต็มไม่เป็นลบที่มีฟังก์ชันมวลความน่าจะเป็น (Probability mass function) $p_X(x) > 0$ สำหรับทุก $x \in \mathscr{X}$ เมื่อ $\mathscr{X}$ คือ เซตของค่าตัวแปรสุ่ม X ที่มีค่าคาดหมาย $\mu = E(X)$ และความแปรปรวน $\sigma^2 = Var(X) < \infty$ และให้ B แทนตัวแปรสุ่มทวินาม (Binomial random variable) ที่มีพารามิเตอร์ n $(n \in \mathbb{N})$ และ p $(0 < p = 1 - q < 1)$ ซึ่งมีฟังก์ชันมวลความน่าจะเป็นดังนี้

$$p_B(x) = \binom{n}{x} p^x q^{n-x} \quad ; x = 0, 1, 2, \ldots, n \tag{1.1}$$

โดยมีค่าเฉลี่ย $E(B) = np$ และความแปรปรวน $Var(B) = npq$ ในการประมาณการแจกแจงของตัวแปรสุ่ม X ด้วยการแจกแจงทวินาม Teerapabolarn and Wongkasem (2011) ได้ใช้วิธีของสไตน์และฟังก์ชัน w หาขอบเขตไม่เอกรูปสำหรับเมตริกแบบจุด (Point metric) ระหว่างการแจกแจงของตัวแปรสุ่ม X และการแจกแจงทวินามดังนี้

กรณีที่ $x_0 = 0$

$$d_0(X, B) \le \frac{1 - q^n}{np} \left\{ E \left| (n - X)p - \sigma^2 w(X) \right| + | np - \mu | \right\} \tag{1.2}$$

และถ้า $np = \mu$ แล้ว

$$d_0(X, B) \le \frac{1 - q^n}{np} E \left| (n - X)p - \sigma^2 w(X) \right| \tag{1.3}$$

กรณีที่ $x_0 \in \{1, \ldots, n\}$

$$d_{x_0}(X, B) \le \min \left\{ \frac{1 - p^n}{x_0 q}, \frac{1 - p^{n+1} - q^{n+1}}{(n+1)pq} \right\} \left\{ E \left| (n - X)p - \sigma^2 w(X) \right| + | np - \mu | \right\} \tag{1.4}$$

และถ้า $np = \mu$ แล้ว

$$d_{x_0}(X, B) \le \min \left\{ \frac{1 - p^n}{x_0 q}, \frac{1 - p^{n+1} - q^{n+1}}{(n+1)pq} \right\} E \left| (n - X)p - \sigma^2 w(X) \right| \tag{1.5}$$

โดยที่ $d_{x_0}(X, B) = | p_X(x_0) - p_B(x_0) |$ เมื่อ $x_0 \in \{0, \ldots, n\}$

พิจารณาผลลัพธ์ในอสมการ (1.2) ถึง (1.5) ผลลัพธ์ดังกล่าวมีความเหมาะสมสำหรับการประมาณการแจกแจงของตัวแปรสุ่ม X ด้วยการแจกแจงทวินามที่อยู่ในรูปของการประมาณแบบจุด (Pointwise approximation) เท่านั้น ซึ่งในกรณีที่ต้องการประมาณฟังก์ชันการแจกแจงสะสม

(Cumulative distribution function) ของตัวแปรสุ่ม X ด้วยฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มทวินาม ผลลัพธ์ในอสมการ (1.2) ถึง (1.5) จะทำให้ขอบเขตไม่เอกรูปมีค่ามากขึ้นเมื่อ x_0 มีค่าเพิ่มขึ้น ซึ่งอาจไม่เหมาะสมสำหรับการประมาณฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่ม X ดังนั้นการศึกษาครั้งนี้จึงต้องการหาผลลัพธ์ใหม่เพื่อใช้ประมาณฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่ม X ด้วยฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มทวินาม และโดยทั่วไปรูปแบบของการประมาณฟังก์ชันการแจกแจงสะสมจะอยู่ในรูปของผลต่างของฟังก์ชันการแจกแจงสะสมสองฟังก์ชัน ซึ่งจะเรียกว่าระยะทางของคอลโมโกรอฟ (Kolmogorov distance) และเกณฑ์ที่ใช้วัดความแม่นยำของการประมาณในกรณีนี้ คือ ขอบเขตไม่เอกรูปสำหรับระยะทางของคอลโมโกรอฟระหว่างฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มที่มีค่าจำนวนเต็มไม่เป็นลบและฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มทวินาม

วัตถุประสงค์ของการวิจัย

2.1 เพื่อหาขอบเขตไม่เอกรูปสำหรับระยะทางของคอลโมโกรอฟระหว่างฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มที่มีค่าจำนวนเต็มไม่เป็นลบและฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มทวินาม

2.2 เพื่อประยุกต์ใช้ขอบเขตไม่เอกรูปในการประมาณฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มเรขาคณิตไฮเพอร์ ฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มเรขาคณิตไฮเพอร์เชิงลบ ฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มโพลยา และฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มทวินามบีต้า

ประโยชน์ที่คาดว่าจะได้รับจากการวิจัย

ได้ขอบเขตไม่เอกรูปเพื่อใช้ตรวจสอบความแม่นยำในการประมาณฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มที่มีค่าจำนวนเต็มไม่เป็นลบด้วยฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มทวินาม

ขอบเขตของการวิจัย

ในการศึกษาครั้งนี้ ศึกษาเฉพาะการหาขอบเขตไม่เอกรูปสำหรับระยะทางของคอลโมโกรอฟระหว่างฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มที่มีค่าจำนวนเต็มไม่เป็นลบและฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มทวินาม และวิธีที่ใช้หาขอบเขตไม่เอกรูปดังกล่าว คือ วิธีของสไตน์สำหรับการแจกแจงทวินามและฟังก์ชัน w

บทที่ 2
เอกสารและงานวิจัยที่เกี่ยวข้อง

ทฤษฎีที่เกี่ยวข้อง

1.1 ขอบเขตบนของฟังก์ชันค่าจริง (Upper bound of real-valued function)

ให้ S เป็นเซตย่อยของ $\mathbb{R}$ และ f เป็นฟังก์ชันจาก S ไป $\mathbb{R}$ ถ้ามีจำนวนจริง u ซึ่ง $f(x) \le u$ สำหรับทุก $x \in S$ จะเรียก u ว่าเป็นขอบเขตบนของ f และถ้า u เป็นขอบเขตบนของ f ที่มีค่าน้อยกว่าหรือเท่ากับขอบเขตทุกตัวของ f แล้วจะเรียก u ว่า ขอบเขตบนที่น้อยที่สุด (Least upper bound หรือ supremum) ของ f แทนด้วย $\sup f$ เช่น $f(x) = \dfrac{x}{x+1}$ เมื่อ $x \ge 1$ ดังนั้น ขอบเขตบนของ f คือ จำนวนจริงที่มีค่ามากกว่าหรือเท่ากับ 1 และ $\sup_{x \ge 1} f(x) = 1$

1.2 ขอบเขตเอกรูปและไม่เอกรูป (Uniform and non-uniform bounds)

ให้ $\mathbb{F}$ และ $\mathbb{H}$ แทนฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มที่มีค่าจำนวนเต็มไม่เป็นลบ X และ Y ตามลำดับ และให้ $\delta(x_0)$ แทนขอบเขตบนของระยะทางระหว่าง $\mathbb{F}$ และ $\mathbb{H}$ ที่ $x_0 \in \mathbb{N} \cup \{0\}$ ซึ่งมีรูปแบบการประมาณการแจกแจงพร้อมด้วยขอบเขตบนเป็นดังนี้

$$\left|\mathbb{F}(x_0) - \mathbb{H}(x_0)\right| \le \delta\left(x_0\right)$$

ถ้า $\delta(x_0)$ มีค่าคงตัวสำหรับทุกค่าของ x_0 แล้วจะกล่าวว่า $\delta(x_0)$ เป็นขอบเขตเอกรูปของ $\left|\mathbb{F}(x_0) - \mathbb{H}(x_0)\right|$ บนเซต $\mathbb{N} \cup \{0\}$ และถ้า $\delta(x_0)$ มีค่าเปลี่ยนแปลงตามค่าของ x_0 แล้วจะกล่าวว่า $\delta(x_0)$ เป็นขอบเขตไม่เอกรูปของ $\left|\mathbb{F}(x_0) - \mathbb{H}(x_0)\right|$ บนเซต $\mathbb{N} \cup \{0\}$

1.3 ระยะทางของคอลโมโกรอฟ (Kolmogorov distance)

ให้ $\mathbb{F}$ และ $\mathbb{H}$ แทนฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มที่มีค่าจำนวนเต็มไม่เป็นลบ X และ Y ตามลำดับ ระยะทางของคอลโมโกรอฟระหว่าง $\mathbb{F}$ และ $\mathbb{H}$ สำหรับทุก $x_0 \in \mathbb{N} \cup \{0\}$ กำหนดได้ดังนี้

(1) สำหรับขอบเขตเอกรูป (ไม่เปลี่ยนแปลงตามค่าของ x_0) จะได้ว่า

$$d_K(X, H) = \sup_{x_0 \ge 0} \left|\mathbb{F}(x_0) - \mathbb{H}(x_0)\right|$$

(2) สำหรับขอบเขตไม่เอกรูป (เปลี่ยนแปลงตามค่าของ x_0) จะได้ว่า

$$d_{K_{x_0}}(X,H) = |\mathbb{F}(x_0) - \mathbb{H}(x_0)|$$

1.4 วิธีของสไตน์สำหรับการการแจกแจงทวินาม

วิธีของสไตน์ (Stein's method) ได้เริ่มต้นนำเสนอครั้งแรกโดย Stein (1972) ซึ่งเป็นวิธีที่ใช้ในการประมาณการแจกแจงของผลรวมของตัวแปรสุ่มที่ไม่เป็นอิสระต่อกัน (Dependent random variables) ด้วยการแจกแจงปรกติ (Normal distribution) ต่อมา Chen (1975) ได้ปรับปรุงและพัฒนาวิธีของสไตน์แบบเดิมมาสู่การประมาณการแจกแจงของผลรวมของตัวแปรสุ่มแบร์นูลลีที่ไม่เป็นอิสระต่อกัน (Dependent Bernoulli random variable) ด้วยการแจกแจงปัวซง (Poisson distribution) และ Stein (1986) ได้นำวิธีของสไตน์มาใช้กับการประมาณการแจกแจงของตัวแปรสุ่มไม่ต่อเนื่องด้วยการแจกแจงทวินาม และสมการของสไตน์สำหรับการแจกแจงทวินามที่มีพารามิเตอร์ n และ p เมื่อ $0 < p = 1 - q < 1$ (เมื่อกำหนดฟังก์ชัน h) ดังนี้

$$h(x) - B_{n,p}(h) = (n-x)pg(x+1) - qxg(x) \tag{2.1}$$

โดยที่ $B_{n,p}(h) = \sum_{k=0}^{n} h(k)\binom{n}{k} p^k q^{n-k}$ และ g และ h เป็นฟังก์ชันค่าจริงที่มีขอบเขตและนิยามบนเซต $\{0,1,\ldots,n\}$

1.5 ฟังก์ชัน w

ให้ X แทนตัวแปรสุ่มที่มีค่าเป็นจำนวนเต็มไม่เป็นลบดังที่นิยามในบทที่ 1 Cacoullos and Papathanasiou (1989) ได้นิยามฟังก์ชัน w ที่สัมพันธ์กับตัวแปรสุ่ม X หรือสัมพันธ์กับการแจกแจงของตัวแปรสุ่ม X ในรูปของความสัมพันธ์ดังนี้

$$w(x)p_X(x) = \frac{1}{\sigma^2}\sum_{i=0}^{x}(\mu - i)p_X(i), \quad x \in \mathscr{X} \tag{2.2}$$

ต่อมา Majsnerowska (1998) ได้ปรับรูปความสัมพันธ์ในสมการ (2.2) ให้อยู่ในรูปแบบของความสัมพันธ์เวียนเกิด (Recurrence relation) ดังนี้

$$w(x) = \frac{1}{\sigma^2}\left\{\mu + \frac{\sigma^2 w(x-1)p_X(x-1)}{p_X(x)} - x\right\}, \quad x \in \mathscr{X} \setminus \{0\} \tag{2.3}$$

และ $w(x) \ge 0$, $x \in \mathscr{X} \setminus \{0\}$ (2.4)

โดยที่ $w(0)=\dfrac{\mu}{\sigma^2}$ และ $p_X(x)>0$ สำหรับทุก $x\in\mathscr{X}$

รูปแบบความสัมพันธ์ต่อไปนี้เป็นสมบัติที่สำคัญของฟังก์ชัน w สำหรับการสร้างผลการวิจัยที่ต้องการซึ่ง Cacoullos and Papathanasiou (1989) ได้กำหนดไว้ดังนี้

ถ้าตัวแปรสุ่มที่มีค่าจำนวนเต็มไม่เป็นลบ X ที่มี $p_X(x)>0$ สำหรับทุก $x\in\mathscr{X}$ และมีความแปรปรวนจำกัด $0<\sigma^2<\infty$ แล้ว

$$E\left[(X-\mu)g(X)\right]=\sigma^2E[w(X)\Delta g(X)] \tag{2.5}$$

สำหรับฟังก์ชัน $g:\mathbb{N}\cup\{0\}\to\mathbb{R}$ ที่ทำให้ $E|w(X)\Delta g(X)|<\infty$ โดย $\Delta g(x)=g(x+1)-g(x)$ และจะได้ว่า $E[w(X)]=1$ (เมื่อกำหนด $g(x)=x$)

งานวิจัยที่เกี่ยวข้อง

ในการศึกษานี้จะศึกษาเกี่ยวกับการประมาณฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มที่มีค่าจำนวนเต็มไม่เป็นลบด้วยฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มทวินาม ซึ่งมีงานวิจัยที่เกี่ยวข้อง ดังนี้

Wongkasem et al. (2008) ได้ใช้วิธีของสไตน์สำหรับการแจกแจงทวินามและฟังก์ชัน w ที่สัมพันธ์กับตัวแปรสุ่มทวินามทั่วไปในการหาขอบเขตเอกรูปของการประมาณการแจกแจงทวินามทั่วไปด้วยการแจกแจงทวินามดังนี้

$$d_{TV}(GB,B)\le\left[1-\left(\frac{A}{A+B}\right)^{n+1}-\left(\frac{B}{A+B}\right)^{n+1}\right]\frac{n(n-1)\alpha\delta(\alpha)}{(n+1)(A+B-\alpha)} \tag{2.6}$$

โดยที่ $d_{TV}(GB,B)=\sup\limits_{E\subseteq\{0,\dots,n\}}|P(GB\in E)-P(B\in E)|$ คือ ระยะทางของความแตกต่างรวม (Total variation distance) ระหว่างการแจกแจงทวินามทั่วไป (การแจกแจงของตัวแปรสุ่ม GB) ที่มีพารามิเตอร์ A,B,n และ α และการแจกแจงทวินามที่มีพารามิเตอร์ n และ $\dfrac{A}{A+B}$

Prukpousana and Teerapabolarn (2010) ได้ใช้วิธีของสไตน์สำหรับการแจกแจงทวินามและฟังก์ชัน w ที่สัมพันธ์กับตัวแปรสุ่มเรขาคณิตไฮเพอร์เชิงลบ ในการหาขอบเขตเอกรูปสำหรับการประมาณการแจกแจงเรขาคณิตไฮเพอร์เชิงลบด้วยการแจกแจงทวินามดังนี้

$$d_{TV}(NH,B)\le\left(1-\left(\frac{r}{R+1}\right)^{S+1}-\left(\frac{R-r+1}{R+1}\right)^{S+1}\right)\frac{(S-1)S}{(S+1)(R+2)} \tag{2.7}$$

โดยที่ NH คือ ตัวแปรสุ่มเรขาคณิตไฮเพอร์เชิงลบที่มีพารามิเตอร์ R,S และ r

Teerapabolarn (2011) ได้ใช้วิธีของสไตน์สำหรับการแจกแจงทวินามและฟังก์ชัน w ที่สัมพันธ์กับตัวแปรทวินามทั่วไป ในการหาขอบเขตไม่เอกรูปสำหรับการประมาณแบบจุดของการแจกแจงทวินามทั่วไปด้วยการแจกแจงทวินามดังนี้

$$d_{x_0}(GB,B) \le \begin{cases} \dfrac{(np+q^n-1)B|\alpha|}{(A+B-\alpha)A} & , x_0 = 0 \\ \dfrac{(1-q^{n-1})qn|\alpha|}{A+B-\alpha} & , x_0 = 1 \\ \dfrac{n(n-1)|\alpha|}{A+B-\alpha}\min\left\{\dfrac{(1-p^n)p}{x_0}, \dfrac{1-p^{n+1}-q^{n+1}}{n+1}\right\} & , x_0 = \{2,\dots,n\} \end{cases} \tag{2.8}$$

โดยที่ $d_{x_0}(GB,B) = |p_{GB}(x_0) - p_B(x_0)|$ คือ เมตริกแบบจุดระหว่างการแจกแจงทวินามทั่วไปที่มีพารามิเตอร์ A,B,n และ α และการแจกแจงทวินามที่มีพารามิเตอร์ n และ $\dfrac{A}{A+B}$

Teerapabolarn and Wongkasem (2011) ได้ใช้วิธีของสไตน์และฟังก์ชัน w หาขอบเขตไม่เอกรูปบนเมตริกแบบจุดระหว่างการแจกแจงของตัวแปรสุ่ม X และการแจกแจงทวินาม ดังที่ได้แสดงในอสมการ (1.2) ถึง (1.5)

บทที่ 3
วิธีดำเนินการวิจัย

สำหรับวิธีดำเนินการวิจัยนั้น สามารถแบ่งการดำเนินการออกเป็น 3 ขั้นตอน ดังนี้

3.1 ศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับการประมาณการแจกแจงทวินาม โดยเฉพาะการประมาณการแจกแจงทวินามด้วยวิธีของสไตน์และฟังก์ชัน w จากงานวิจัยของ Teerapabolarn and Wongkasem (2011) เพื่อนำมาสร้างบทตั้งต่าง ๆ ที่จำเป็นสำหรับการสร้างและพิสูจน์ผลลัพธ์ของการวิจัยดังต่อไปนี้

สมการของสไตน์สำหรับการแจกแจงทวินามที่มีพารามิเตอร์ $n \in \mathbb{N}$ และ $0 < p < 1$ (เมื่อกำหนดฟังก์ชัน h) ซึ่ง Barbour et al. (1992) ได้ปรับเป็นดังนี้

$$h(x) - \mathbb{B}_{n,p}(h) = (n-x)pg(x+1) - qxg(x) \tag{3.1}$$

โดยที่ $\mathbb{B}_{n,p}(h) = \sum_{k=0}^{n} h(k)\binom{n}{k} p^k q^{n-k}$ และ g และ h เป็นฟังก์ชันค่าจริงที่มีขอบเขตนิยามบนเซต $\{0,\ldots,n\}$

สำหรับ $E \subseteq \{0,\ldots,n\}$ ให้ $h_E : \{0,\ldots,n\} \to \mathbb{R}$ เป็นฟังก์ชันที่กำหนดโดย

$$h_E(x) = \begin{cases} 1 & , x \in E \\ 0 & , x \notin E \end{cases} \tag{3.2}$$

และจาก Barbour et al. (1992) ที่กำหนดให้ $g_E : \mathbb{N} \cup \{0\} \to \mathbb{R}$ เป็นฟังก์ชันที่สอดคล้องกับสมการ (3.1) โดยที่ $g_E(0) = g_E(1)$ และ $g_E(x) = g_E(n)$ สำหรับทุก $x \ge n$ ดังนั้น ผลเฉลย g_E ของสมการ (3.1) คือ

$$g_E(x) = \frac{\mathbb{B}_{n,p}(h_{E \cap C_{x-1}}) - \mathbb{B}_{n,p}(h_E)\mathbb{B}_{n,p}(h_{C_{x-1}})}{x\binom{n}{x} p^x q^{n-x+1}} \tag{3.3}$$

โดยที่ $1 \le x \le n$ และ $C_x = \{0,\ldots,x\}$

เมื่อ $E = \{x_0\}$ สำหรับ $x_0 \in \{0,\ldots,n\}$ และให้ $h_{x_0} = h_{\{x_0\}}$ จะสามารถเขียนผลเฉลย $g_{x_0} = g_{\{x_0\}}$ ของสมการ (3.3) ได้ดังนี้

$$g_{x_0}(x)=\begin{cases}\dfrac{-\mathbb{B}_{n,p}(h_{x_0})\mathbb{B}_{n,p}(h_{C_{x-1}})}{x\binom{n}{x}p^x q^{n-x+1}} & , \quad 1\le x\le x_0\\[2ex] \dfrac{\mathbb{B}_{n,p}(h_{x_0})\mathbb{B}_{n,p}(1-h_{C_{x-1}})}{x\binom{n}{x}p^x q^{n-x+1}} & , \quad x_0\le x\le n\end{cases} \tag{3.4}$$

เมื่อ $E=C_{x_0}$ สำหรับ $x_0\in\{0,\ldots,n\}$ จะได้ผลเฉลย $g_{C_{x_0}}$ ของสมการ (3.3) ได้ดังนี้

$$g_{C_{x_0}}(x)=\begin{cases}\dfrac{\mathbb{B}_{n,p}(h_{C_{x-1}})\mathbb{B}_{n,p}(1-h_{C_{x_0}})}{x\binom{n}{x}p^x q^{n-x+1}} & , \quad 1\le x\le x_0\\[2ex] \dfrac{\mathbb{B}_{n,p}(h_{C_{x_0}})\mathbb{B}_{n,p}(1-h_{C_{x-1}})}{x\binom{n}{x}p^x q^{n-x+1}} & , \quad x_0\le x\le n\end{cases} \tag{3.5}$$

ให้ $\Delta g_{x_0}(x)=g_{x_0}(x+1)-g_{x_0}(x)$ และ $\Delta g_{C_{x_0}}(x)=g_{C_{x_0}}(x+1)-g_{C_{x_0}}(x)$ จะเห็นได้ว่า $\Delta g_{x_0}(x)=0$ และ $\Delta g_{C_{x_0}}(x)=0$ เมื่อ $x=0,n$ ดังนั้น สำหรับ $x\in\{1,\ldots,n-1\}$ โดยสมการ (3.4) จะได้ว่า

$$\Delta g_{x_0}(x)=\begin{cases}-\dfrac{\mathbb{B}_{n,p}(h_{x_0})\mathbb{B}_{n,p}(h_{C_x})}{(x+1)\binom{n}{x+1}p^{x+1}q^{n-x}}+\dfrac{\mathbb{B}_{n,p}(h_{x_0})\mathbb{B}_{n,p}(h_{C_{x-1}})}{x\binom{n}{x}p^x q^{n-x+1}} & , \ x_0>x\\[2ex] \dfrac{\mathbb{B}_{n,p}(h_{x_0})\mathbb{B}_{n,p}(1-h_{C_x})}{(x+1)\binom{n}{x+1}p^{x+1}q^{n-x}}+\dfrac{\mathbb{B}_{n,p}(h_{x_0})\mathbb{B}_{n,p}(h_{C_{x-1}})}{x\binom{n}{x}p^x q^{n-x+1}} & , \ x_0=x\\[2ex] \dfrac{\mathbb{B}_{n,p}(h_{x_0})\mathbb{B}_{n,p}(1-h_{C_x})}{(x+1)\binom{n}{x+1}p^{x+1}q^{n-x}}-\dfrac{\mathbb{B}_{n,p}(h_{x_0})\mathbb{B}_{n,p}(1-h_{C_{x-1}})}{x\binom{n}{x}p^x q^{n-x+1}} & , \ x_0<x\end{cases} \tag{3.6}$$

$$=\begin{cases}\dfrac{\mathbb{B}_{n,p}(h_{x_0})}{\binom{n}{x}p^x q^{n-x}}\left[-\dfrac{\mathbb{B}_{n,p}(h_{C_x})}{(n-x)p}+\dfrac{\mathbb{B}_{n,p}(h_{C_{x-1}})}{xq}\right] & , \ x_0>x\\[2ex] \dfrac{\mathbb{B}_{n,p}(h_{x_0})}{\binom{n}{x}p^x q^{n-x}}\left[\dfrac{\mathbb{B}_{n,p}(1-h_{C_x})}{(n-x)p}+\dfrac{\mathbb{B}_{n,p}(h_{C_{x-1}})}{xq}\right] & , \ x_0=x\\[2ex] \dfrac{\mathbb{B}_{n,p}(h_{x_0})}{\binom{n}{x}p^x q^{n-x}}\left[\dfrac{\mathbb{B}_{n,p}(1-h_{C_x})}{(n-x)p}-\dfrac{\mathbb{B}_{n,p}(1-h_{C_{x-1}})}{xq}\right] & , \ x_0<x\end{cases} \tag{3.7}$$

$$
=\begin{cases}
\dfrac{\binom{n}{x_0}p^{x_0}q^{n-x_0}}{\binom{n}{x}p^{x}q^{n-x}}\left[-\dfrac{\sum\limits_{k=0}^{x}\binom{n}{k}p^{k}q^{n-k}}{(n-x)p}+\dfrac{\sum\limits_{k=0}^{x-1}\binom{n}{k}p^{k}q^{n-k}}{xq}\right] & ,\ x_0>x\\[2ex]
\dfrac{\binom{n}{x_0}p^{x_0}q^{n-x_0}}{\binom{n}{x}p^{x}q^{n-x}}\left[\dfrac{\sum\limits_{k=x+1}^{n}\binom{n}{k}p^{k}q^{n-k}}{(n-x)p}+\dfrac{\sum\limits_{k=0}^{x-1}\binom{n}{k}p^{k}q^{n-k}}{xq}\right] & ,\ x_0=x\\[2ex]
\dfrac{\binom{n}{x_0}p^{x_0}q^{n-x_0}}{\binom{n}{x}p^{x}q^{n-x}}\left[\dfrac{\sum\limits_{k=x+1}^{n}\binom{n}{k}p^{k}q^{n-k}}{(n-x)p}-\dfrac{\sum\limits_{k=x}^{n}\binom{n}{k}p^{k}q^{n-k}}{xq}\right] & ,\ x_0<x
\end{cases}
\tag{3.8}
$$

และโดยสมการ (3.5) จะได้ว่า

$$
\Delta g_{C_{x_0}}(x)=\begin{cases}
\dfrac{\mathbb{B}_{n,p}(1-h_{C_{x_0}})\mathbb{B}_{n,p}(h_{C_x})}{(x+1)\binom{n}{x+1}p^{x+1}q^{n-x}}-\dfrac{\mathbb{B}_{n,p}(1-h_{C_{x_0}})\mathbb{B}_{n,p}(h_{C_{x-1}})}{x\binom{n}{x}p^{x}q^{n-x+1}} & ,\ x_0\ge x\\[2ex]
\dfrac{\mathbb{B}_{n,p}(h_{C_{x_0}})\mathbb{B}_{n,p}(1-h_{C_x})}{(x+1)\binom{n}{x+1}p^{x+1}q^{n-x}}-\dfrac{\mathbb{B}_{n,p}(h_{C_{x_0}})\mathbb{B}_{n,p}(1-h_{C_{x-1}})}{x\binom{n}{x}p^{x}q^{n-x+1}} & ,\ x_0<x
\end{cases}
\tag{3.9}
$$

$$
=\begin{cases}
\dfrac{\mathbb{B}_{n,p}(1-h_{C_{x_0}})}{\binom{n}{x}p^{x}q^{n-x}}\left[\dfrac{\mathbb{B}_{n,p}(h_{C_x})}{(n-x)p}-\dfrac{\mathbb{B}_{n,p}(h_{C_{x-1}})}{xq}\right] & ,\ x_0\ge x\\[2ex]
\dfrac{\mathbb{B}_{n,p}(h_{C_{x_0}})}{\binom{n}{x}p^{x}q^{n-x}}\left[\dfrac{\mathbb{B}_{n,p}(1-h_{C_x})}{(n-x)p}-\dfrac{\mathbb{B}_{n,p}(1-h_{C_{x-1}})}{xq}\right] & ,\ x_0<x
\end{cases}
\tag{3.10}
$$

$$
=\begin{cases}
\dfrac{\sum\limits_{j=x_0+1}^{n}\binom{n}{j}p^{j}q^{n-j}}{\binom{n}{x}p^{x}q^{n-x}}\left[\dfrac{\sum\limits_{k=0}^{x}\binom{n}{k}p^{k}q^{n-k}}{(n-x)p}-\dfrac{\sum\limits_{k=0}^{x-1}\binom{n}{k}p^{k}q^{n-k}}{xq}\right] & ,\ x_0\ge x\\[2ex]
\dfrac{\sum\limits_{j=0}^{x_0}\binom{n}{j}p^{j}q^{n-j}}{\binom{n}{x}p^{x}q^{n-x}}\left[\dfrac{\sum\limits_{k=x+1}^{n}\binom{n}{k}p^{k}q^{n-k}}{(n-x)p}-\dfrac{\sum\limits_{k=x}^{n}\binom{n}{k}p^{k}q^{n-k}}{xq}\right] & ,\ x_0<x
\end{cases}
\tag{3.11}
$$

ต่อไปจะสร้างบทตั้งต่าง ๆ ที่แสดงสมบัติที่สำคัญของฟังก์ชัน $g_{C_{x_0}}$ และ $\Delta g_{C_{x_0}}$ เพื่อที่จะนำไปใช้ในการพิสูจน์ทฤษฎีบทที่ต้องการ

บทตั้ง 3.1 ให้ $x, x_0 \in \{1,...,n\}$ แล้วจะได้ว่าสมการต่อไปนี้เป็นจริง

$$\Delta g_{x_0}(x)\begin{cases} >0 & , x_0 = x \\ <0 & , x_0 \neq x \end{cases} \tag{3.12}$$

และ

$$\Delta g_{C_{x_0}}(x)\begin{cases} >0 & , 1\le \mathrm{x} \le x_0 \\ <0 & , x_0 < \mathrm{x} \le n \end{cases} \tag{3.13}$$

พิสูจน์ เริ่มต้นจะแสดงว่าสมการ (3.12) เป็นจริงดังนี้ จากสมการ (3.8) เมื่อ $x_0 = x$ จะได้ว่า

$$\Delta g_{x_0}(x) > 0 \tag{3.14}$$

ต่อไปจะแสดงว่า $\Delta g_{x_0}(x) < 0$ เมื่อ $x_0 \neq x$

เมื่อ $1 \le x < x_0$ และจากสมการ (3.8) จะได้ว่า

$$\Delta g_{x_0}(x) = \frac{\binom{n}{x_0} p^{x_0} q^{n-x_0}}{\binom{n}{x} p^x q^{n-x}} \left[-\frac{\sum_{k=0}^{x} \binom{n}{k} p^k q^{n-k}}{(n-x)p} + \frac{\sum_{k=0}^{x-1} \binom{n}{k} p^k q^{n-k}}{xq} \right]$$

$$= \frac{\binom{n}{x_0} p^{x_0} q^{n-x_0}}{x\binom{n}{x} p^{x+1} q^{n-x+1}(n-x)} \left[-\sum_{k=0}^{x} \binom{n}{k} p^k q^{n-k+1} + \sum_{k=0}^{x-1} (n-x)\binom{n}{k} p^{k+1} q^{n-k} \right]$$

ให้ $\delta_1(x) = -\sum_{k=0}^{x} \binom{n}{k} p^k q^{n-k+1} + \sum_{k=0}^{x-1} (n-x)\binom{n}{k} p^{k+1} q^{n-k}$ จะได้ว่า

$$\delta_1(x) = -\sum_{k=0}^{x} \frac{x(n-k+1)}{n+1} \binom{n+1}{k} p^k q^{n-k+1} + \sum_{k=0}^{x-1} \frac{(k+1)(n-x)}{n+1} \binom{n+1}{k+1} p^{k+1} q^{n+1-(k+1)}$$

$$= -\sum_{k=0}^{x} \frac{x(n-k+1)}{n+1} \binom{n+1}{k} p^k q^{n-k+1} + \sum_{k=0}^{x} \frac{k(n-x)}{n+1} \binom{n+1}{k} p^k q^{n+1-k}$$

$$= -\sum_{k=0}^{x} \frac{\binom{n+1}{k} p^k q^{n-k+1}}{n+1} \left[-x(n+1) + kn \right]$$

$$< 0$$

ซึ่งจะได้ว่า

$$\Delta g_{x_0}(x) < 0 \qquad (3.15)$$

เมื่อ $x_0 < x < n$ และจากสมการ (3.8) จะได้ว่า

$$\Delta g_{x_0}(x) = \frac{\binom{n}{x_0} p^{x_0} q^{n-x_0}}{\binom{n}{x} p^x q^{n-x}} \left[\frac{\sum_{k=x+1}^{n} \binom{n}{k} p^k q^{n-k}}{(n-x)p} - \frac{\sum_{k=x}^{n} \binom{n}{k} p^k q^{n-k}}{xq} \right]$$

$$= \frac{\binom{n}{x_0} p^{x_0} q^{n-x_0}}{x \binom{n}{x} p^{x+1} q^{n-x+1} (n-x)} \left[\sum_{k=x+1}^{n} x \binom{n}{k} p^k q^{n-k+1} - \sum_{k=x}^{n} (n-x) \binom{n}{k} p^{k+1} q^{n-k} \right]$$

ให้ $\delta_2(x) = \sum_{k=x+1}^{n} x \binom{n}{k} p^k q^{n-k+1} - \sum_{k=x}^{n} (n-x) \binom{n}{k} p^{k+1} q^{n-k}$ จะได้ว่า

$$\delta_2(x) = \sum_{k=x+1}^{n} \frac{x(n-k+1)}{n+1} \binom{n+1}{k} p^k q^{n-k+1} - \sum_{k=x}^{n} \frac{(k+1)(n-x)}{n+1} \binom{n+1}{k+1} p^{k+1} q^{n+1-(k+1)}$$

$$= \sum_{k=x+1}^{n+1} \frac{x(n-k+1)}{n+1} \binom{n+1}{k} p^k q^{n-k+1} - \sum_{k=x+1}^{n+1} \frac{k(n-x)}{n+1} \binom{n+1}{k} p^k q^{n+1-k}$$

$$= \sum_{k=x+1}^{n+1} \frac{\binom{n+1}{k} p^k q^{n-k+1}}{n+1} \left[x(n+1) - kn \right]$$

$$< 0$$

ซึ่งจะได้ว่า

$$\Delta g_{x_0}(x) < 0 \qquad (3.16)$$

และจากอสมการ (3.15) และ (3.16) เมื่อ $x_0 \neq x$ จะได้ว่า

$$\Delta g_{x_0}(x) < 0 \tag{3.17}$$

ดังนั้นโดยการพิสูจน์ทั้งสองกรณี (อสมการ (3.14) และ (3.17)) จะได้ว่า อสมการ (3.12) เป็นจริง

ต่อไปจะแสดงว่า (3.13) เป็นจริงดังนี้ เมื่อ $1 \leq x \leq x_0$ โดยใช้สมการ (3.11) จะได้ว่า

$$\Delta g_{C_{x_0}}(x) = \frac{\sum_{j=x_0+1}^{n}\binom{n}{j}p^j q^{n-j}}{x\binom{n}{x}p^{x+1}q^{n-x+1}(n-x)}\left[\sum_{k=0}^{x} x\binom{n}{k}p^k q^{n-k+1} - \sum_{k=0}^{x-1}(n-x)\binom{n}{k}p^{k+1}q^{n-k}\right]$$

$$= \frac{\sum_{j=x_0+1}^{n}\binom{n}{j}p^j q^{n-j}}{x\binom{n}{x}p^{x+1}q^{n-x+1}(n-x)}\sum_{k=0}^{x}\frac{\binom{n+1}{k}p^k q^{n-k+1}}{n+1}\left[x(n+1)-kn\right]$$

> 0 (โดยอสมการ (3.15))

ซึ่งจะได้ว่า

$$\Delta g_{C_{x_0}}(x) > 0 \tag{3.18}$$

เมื่อ $x_0 < x < n$ และจากสมการ (3.11) จะได้ว่า

$$\Delta g_{C_{x_0}}(x) = \frac{\sum_{k=0}^{x_0}\binom{n}{k}p^k q^{n-k}}{\binom{n}{x}p^x q^{n-x}}\left[\frac{\sum_{k=x+1}^{n}\binom{n}{k}p^k q^{n-k}}{(n-x)p} - \frac{\sum_{k=x}^{n}\binom{n}{k}p^k q^{n-k}}{xq}\right]$$

$$= \frac{\sum_{j=0}^{x_0}\binom{n}{j}p^j q^{n-j}}{x\binom{n}{x}p^{x+1}q^{n-x+1}(n-x)}\left[\sum_{k=x+1}^{n} x\binom{n}{k}p^k q^{n-k+1} - \sum_{k=x}^{n}(n-x)\binom{n}{k}p^{k+1}q^{n-k}\right]$$

$$= \frac{\sum_{j=0}^{x_0}\binom{n}{j}p^j q^{n-j}}{x\binom{n}{x}p^{x+1}q^{n-x+1}(n-x)}\sum_{k=x+1}^{n+1}\frac{\binom{n+1}{k}p^k q^{n-k+1}}{n+1}\left[x(n+1)-kn\right]$$

< 0 (โดยอสมการ (3.16))

ซึ่งจะได้ว่า

$$\Delta g_{C_{x_0}}(x) < 0 \tag{3.19}$$

ดังนั้นโดยการพิสูจน์ทั้งสองกรณี (อสมการ (3.18) และ (3.19)) จะได้ว่าอสมการ (3.13) เป็นจริง □

บทตั้ง 3.2 $\Delta g_{C_{x_0}}$ เป็นฟังก์ชันเพิ่มสำหรับ $x \in \{1,\ldots,x_0\}$

พิสูจน์ ให้ $\Delta^2 g_{C_{x_0}}(x) = \Delta g_{C_{x_0}}(x+1) - \Delta g_{C_{x_0}}(x)$ จะแสดงว่า $\Delta^2 g_{C_{x_0}}(x) > 0$ สำหรับ $x \in \{1,\ldots,x_0-1\}$ จากสมการ (3.11) จะได้ว่า

$$\begin{aligned}
\Delta^2 g_{C_{x_0}}(x) = {} & \frac{\sum_{j=x_0+1}^{n}\binom{n}{j}p^j q^{n-j}}{\binom{n}{x+1}p^{x+1}q^{n-x-1}}\left[\frac{\sum_{k=0}^{x+1}\binom{n}{k}p^k q^{n-k}}{(n-x-1)p} - \frac{\sum_{k=0}^{x}\binom{n}{k}p^k q^{n-k}}{(x+1)q}\right] \\
& - \frac{\sum_{j=x_0+1}^{n}\binom{n}{j}p^j q^{n-j}}{\binom{n}{x}p^{x}q^{n-x}}\left[\frac{\sum_{k=0}^{x}\binom{n}{k}p^k q^{n-k}}{(n-x)p} - \frac{\sum_{k=0}^{x-1}\binom{n}{k}p^k q^{n-k}}{xq}\right] \\
= {} & \frac{\sum_{j=x_0+1}^{n}\binom{n}{j}p^j q^{n-j}}{(x+1)\binom{n}{x+1}p^{x+2}q^{n-x}(n-x-1)}\sum_{k=0}^{x+1}\frac{\binom{n+1}{k}p^k q^{n-k+1}}{n+1}\big[(x+1)(n+1)-kn\big] \\
& - \frac{\sum_{j=x_0+1}^{n}\binom{n}{j}p^j q^{n-j}}{x\binom{n}{x}p^{x+1}q^{n-x+1}(n-x)}\sum_{k=0}^{x}\frac{\binom{n+1}{k}p^k q^{n-k+1}}{n+1}\big[x(n+1)-kn\big] \\
= {} & \frac{\sum_{j=x_0+1}^{n}\binom{n}{j}p^j q^{n-j}}{\binom{n}{x}p^{x+2}q^{n-x}(n-x)(n+1)}\sum_{k=0}^{x+1}\binom{n+1}{k}p^k q^{n-k+1}\left[\frac{(x+1)(n+1)-kn}{n-x-1}\right] \\
& - \frac{\sum_{j=x_0+1}^{n}\binom{n}{j}p^j q^{n-j}}{\binom{n}{x}p^{x+2}q^{n-x}(n-x)(n+1)}\sum_{k=0}^{x}\binom{n+1}{k}p^{k+1} q^{n+1-(k+1)}\left[\frac{x(n+1)-kn}{x}\right]
\end{aligned}$$

$$= \frac{\sum_{j=x_0+1}^{n}\binom{n}{j}p^j q^{n-j}}{\binom{n}{x}p^{x+2}q^{n-x}(n-x)(n+1)}\left\{\sum_{k=0}^{x+1}\binom{n+1}{k}p^k q^{n-k+1}\left[\frac{(x+1)(n+1)-kn}{n-x-1}\right]\right.$$

$$\left. -\sum_{k=0}^{x}\binom{n+1}{k+1}\frac{k+1}{n+1-k}p^{k+1}q^{n+1-(k+1)}\left[\frac{x(n+1)-kn}{x}\right]\right\}$$

$$= \frac{\sum_{j=x_0+1}^{n}\binom{n}{j}p^j q^{n-j}}{\binom{n}{x}p^{x+2}q^{n-x}(n-x)(n+1)}\left\{\sum_{k=0}^{x+1}\binom{n+1}{k}p^k q^{n-k+1}\left[\frac{(x+1)(n+1)-kn}{n-x-1}\right]\right.$$

$$\left. -\sum_{k=0}^{x+1}\binom{n+1}{k}\frac{k}{n+2-k}p^{k}q^{n+1-k}\left[\frac{x(n+1)-(k-1)n}{x}\right]\right\}$$

$$= \frac{\sum_{j=x_0+1}^{n}\binom{n}{j}p^j q^{n-j}}{\binom{n}{x}p^{x+2}q^{n-x}(n-x)(n+1)}\sum_{k=0}^{x+1}\binom{n+1}{k}p^k q^{n-k+1}\left\{\left[\frac{x(n+1)-kn+n+1}{n-x-1}\right]\right.$$

$$\left. -\left[\frac{x(n+1)-kn+n}{n+2-k}\right]\frac{k}{x}\right\}$$

$$> 0$$

ดังนั้น $\Delta g_{C_{x_0}}(x)$ เป็นฟังก์ชันเพิ่มสำหรับ $x \in \{1,\ldots,x_0\}$ □

บทตั้ง 3.3 ให้ $x, x_0 \in \{1,\ldots,n\}$ แล้วจะได้ว่าอสมการต่อไปนี้เป็นจริง

$$g_{C_{x_0}}(x) \le g_{C_{x_0}}(x_0+1) \tag{3.20}$$

พิสูจน์ จากบทตั้ง 3.1 (อสมการ 3.13) จะได้ว่า $g_{C_{x_0}}$ เป็นฟังก์ชันเพิ่มสำหรับ $x \in \{1,\ldots,x_0\}$ และเป็นฟังก์ชันลดสำหรับ $x \in \{x_0+1,\ldots,n\}$ ดังนั้นจะได้ $g_{C_{x_0}}(x) \le g_{C_{x_0}}(x_0)$ สำหรับ $x \in \{1,\ldots,x_0\}$ และ $g_{C_{x_0}}(x) \le g_{C_{x_0}}(x_0+1)$ สำหรับ $x \in \{x_0+1,\ldots,n\}$ เนื่องจาก

$$g_{C_{x_0}}(x_0+1) - g_{C_{x_0}}(x_0) = \frac{\mathbb{B}_{n,p}(h_{C_{x_0}})\mathbb{B}_{n,p}(1-h_{C_{x0}})}{(x_0+1)\binom{n}{x_0+1}p^{x_0+1}q^{n-x_0}} - \frac{\mathbb{B}_{n,p}(1-h_{C_{x_0-1}})\mathbb{B}_{n,p}(1-h_{C_{x_0}})}{x_0\binom{n}{x_0}p^{x_0+1}q^{n-x_0+1}}$$

$$= \frac{\mathbb{B}_{n,p}(1-h_{C_{x0}})}{\binom{n}{x_0}p^{x_0}q^{n-x_0}}\left(\frac{\mathbb{B}_{n,p}(h_{C_{x_0}})}{(n-x_0)p} - \frac{\mathbb{B}_{n,p}(h_{C_{x_0-1}})}{x_0 q}\right)$$

โดยใช้การพิสูจน์ของอสมการ (3.18) จะได้ว่า $\dfrac{\mathbb{B}_{n,p}(1-h_{C_{x0}})}{\binom{n}{x_0}p^{x_0}q^{n-x_0}}\left(\dfrac{\mathbb{B}_{n,p}(h_{C_{x_0}})}{(n-x_0)p}-\dfrac{\mathbb{B}_{n,p}(1-h_{C_{x_0-1}})}{x_0 q}\right)\geq 0$

นั่นคือ $g_{C_{x_0}}(x_0+1)-g_{C_{x_0}}(x_0)\geq 0$ หรือ $g_{C_{x_0}}(x)\leq g_{C_{x_0}}(x_0+1)$ สำหรับทุกค่าของ x เมื่อ $x\in\{1,\ldots,n\}$ ดังนั้นจะได้อสมการ (3.20) เป็นจริง □

บทตั้ง 3.4 ให้ $x,x_0\in\{1,\ldots,n\}$ แล้วจะได้ว่า

1. $\sup\limits_{x\geq 1}\left\{g_{C_0}(x)\right\}\leq \dfrac{1-q^n}{np}$ (3.21)

2. $\sup\limits_{x\geq 1}\left|\Delta g_{C_0}(x)\right|\leq \dfrac{np+q^n-1}{(n-1)np^2}$ (3.22)

3. $\sup\limits_{x\geq 1}\left\{g_{C_{x_0}}(x)\right\}\leq \xi(x_0)$ (3.23)

โดยที่ $\xi(x_0)=\begin{cases}\dfrac{(n-x_0-1)\left[1-P(B\leq x_0)\right]}{(n-x_0)\left[(n-1)p-x_0\right]} & ,\ x_0<(n-1)p\\[2ex] \dfrac{(x_0+2)P(B\leq x_0)}{(x_0+1)\left[x_0+2-(n-1)p\right]} & ,\ x_0\geq (n-1)p\end{cases}$ และ $P(B\leq x_0)=\sum\limits_{k=0}^{x_0}p_B(k)$

4. $\sup\limits_{x\geq 1}\left|\Delta g_{C_{x_0}}(x)\right|\leq \min\left\{\dfrac{1-p^n}{x_0 q},\dfrac{1-p^{n+1}-q^{n+1}}{(n+1)pq}\right\}$ (3.24)

พิสูจน์ 1. จากสมการ (3.4) จะได้ว่า $g_{x_0}(x)=\dfrac{\mathbb{B}_{n,p}(h_{x_0})\mathbb{B}_{n,p}(1-h_{c_{x-1}})}{x\binom{n}{x}p^x q^{n-x+1}}$ เมื่อ $x_0<x$ ดังนั้นจะได้

ว่า $g_0(x)=\dfrac{\mathbb{B}_{n,p}(h_0)\mathbb{B}_{n,p}(1-h_{c_{x-1}})}{x\binom{n}{x}p^x q^{n-x+1}}=\dfrac{q^n\mathbb{B}_{n,p}(1-h_{c_{x-1}})}{x\binom{n}{x}p^x q^{n-x+1}}$ ต่อไปจะแสดงว่า $g_0(x+1)-g_0(x)\leq 0$

$$\begin{aligned}g_0(x+1)-g_0(x)&=\frac{q^n\mathbb{B}_{n,p}(1-h_{c_x})}{(x+1)\binom{n}{x+1}p^{x+1}q^{n-x}}-\frac{q^n\mathbb{B}_{n,p}(1-h_{c_{x-1}})}{x\binom{n}{x}p^x q^{n-x+1}}\\&=\frac{q^n}{\binom{n}{x}p^x q^{n-x}}\left[\frac{\mathbb{B}_{n,p}(1-h_{c_x})}{(n-x)p}-\frac{\mathbb{B}_{n,p}(1-h_{c_{x-1}})}{xq}\right]\\&=\frac{q^n}{\binom{n}{x}p^x q^{n-x}}\delta_2(x)\\&\leq 0\ (\text{โดย } \delta_2(x)<0)\end{aligned}\tag{3.25}$$

และทำให้ $g_0(x+1)-g_0(x)\le 0$ หรือ $g_0(x+1)\le g_0(x)$ นั่นคือ $g_0(x)\le\cdots\le g_0(2)\le g_0(1)$ ดังนั้น

$$\begin{aligned} g_0(x) &\le g_0(1) \\ &= \frac{\mathbb{B}_{n,p}(h_0)\mathbb{B}_{n,p}(1-h_{c_0})}{\binom{n}{1}pq^n} \\ &= \frac{q^n(1-q^n)}{npq^n} \\ &= \frac{(1-q^n)}{np} \end{aligned}$$

ซึ่งจะได้อสมการ (3.21)

2. ดูรายละเอียดของการพิสูจน์ได้ใน Teerapabolarn (2011)

3. จะแสดงว่า $\sup_{x\ge 1}\left\{g_{C_{x_0}}(x)\right\}\le \xi(x_0)$ นั่นคือ จะต้องแสดงว่า $g_{C_{x_0}}(x)\le \xi(x_0)$ สำหรับทุกค่า $x\in\{1,\ldots,n\}$ โดยบทตั้ง 3.3 สำหรับทุก $x_0, x\in\{1,\ldots,n\}$ จะได้ว่า

$$\begin{aligned} g_{C_{x_0}}(x) &\le g_{C_{x_0}}(x_0+1) \\ &= \frac{\mathbb{B}_{n,p}(h_{C_{x_0}})\mathbb{B}_{n,p}(1-h_{C_{x_0}})}{(x_0+1)\binom{n}{x_0+1}p^{x_0+1}q^{n-x_0}} \end{aligned}$$

$$= \frac{P(B\le x_0)\left[1-P(B\le x_0)\right]}{(x_0+1)\binom{n}{x_0+1}p^{x_0+1}q^{n-x_0}} \tag{3.26}$$

Barbour et al. (1992) ได้แสดงว่า

$$P(B\le x_0)\le \frac{p(n-x_0-1)\binom{n}{x_0}p^{x_0}q^{n-x_0}}{(n-1)p-x_0} \tag{3.27}$$

เมื่อ $x_0<(n-1)p$ และเมื่อ $x_0>np-1-q$ จะได้ว่า

$$1-P(B\le x_0)\le \frac{(x_0+2)q\binom{n}{x_0+1}p^{x_0+1}q^{n-(x_0+1)}}{x_0+2-(n+1)p} \tag{3.28}$$

โดยอสมการ (3.26) ถึง (3.28) จะได้

$$g_{C_{x_0}}(x) \le \frac{\dfrac{p(n-x_0-1)\binom{n}{x_0}p^{x_0}q^{n-x_0}}{(n-1)p-x_0}\left[1-P(B\le x_0)\right]}{(x_0+1)\binom{n}{x_0+1}p^{x_0+1}q^{n-x_0}}$$

$$=\frac{(n-x_0-1)\left[1-P(B\le x_0)\right]}{(n-x_0)\left[(n-1)p-x_0\right]} \tag{3.29}$$

เมื่อ $x_0<(n-1)p$ และจะได้

$$g_{C_{x_0}}(x) \le \frac{P(B\le x_0)\dfrac{(x_0+2)q\binom{n}{x_0+1}p^{x_0+1}q^{n-(x_0+1)}}{x_0+2-(n+1)p}}{(x_0+1)\binom{n}{x_0+1}p^{x_0+1}q^{n-x_0}}$$

$$=\frac{(x_0+2)P(B\le x_0)}{(x_0+1)\left[x_0+2-(n+1)p\right]} \tag{3.30}$$

เมื่อ $x_0\ge(n-1)p$ (เพราะว่า $(n-1)p>np-1-q$) ดังนั้น จากอสมการ (3.29) และ (3.30) ทำให้ได้ว่า

$$\sup_{x\ge1}\left\{g_{C_{x_0}}(x)\right\}\le\xi(x_0)$$

โดยที่

$$\xi(x_0)=\begin{cases}\dfrac{(n-x_0-1)\left[1-P(B\le x_0)\right]}{(n-x_0)\left[(n-1)p-x_0\right]} & ,\ x_0<(n-1)p\\[2ex] \dfrac{(x_0+2)P(B\le x_0)}{(x_0+1)\left[x_0+2-(n-1)p\right]} & ,\ x_0\ge(n-1)p\end{cases}$$

ซึ่งทำให้ได้อสมการ (3.23) ตามต้องการ

4. Teerapabolarn and Wongkasem (2011) ได้แสดงว่า

$$\Delta g_{x_0}(x_0)\le\min\left\{\frac{1-p^n}{x_0q},\frac{1-p^{n+1}-q^{n+1}}{(n+1)pq}\right\} \tag{3.31}$$

ในทำนองเดียวกันจากอสมการ (3.12) จะได้ว่า

$$\Delta g_x(x) \le \min\left\{\frac{1-p^n}{xq}, \frac{1-p^{n+1}-q^{n+1}}{(n+1)pq}\right\} \tag{3.32}$$

ต่อไปจะแสดงว่า $\sup_{x\ge 1}\left|\Delta g_{C_{x_0}}(x)\right| \le \min\left\{\frac{1-p^n}{x_0 q}, \frac{1-p^{n+1}-q^{n+1}}{(n+1)pq}\right\}$ นั่นคือ จะต้องแสดงว่า $\left|\Delta g_{C_{x_0}}(x)\right| \le \min\left\{\frac{1-p^n}{x_0 q}, \frac{1-p^{n+1}-q^{n+1}}{(n+1)pq}\right\}$ สำหรับทุก $x \in \{1,\ldots,n\}$ ซึ่งจะแบ่งการพิสูจน์ออกเป็น 2 กรณีดังนี้

1. กรณีที่ $1 \le x \le x_0$

$$\begin{aligned}
0 &\le \Delta g_{C_{x_0}}(x) \quad \text{(โดยอสมการ (3.13))}\\
&\le \Delta g_{C_{x_0}}(x_0) \quad \text{(โดยบทตั้ง 2.2)}\\
&= \sum_{k=0}^{x_0} \Delta g_k(x_0)\\
&= \Delta g_0(x_0) + \Delta g_1(x_0) + \cdots + \Delta g_{x_0}(x_0)\\
&\le \Delta g_{x_0}(x_0) \quad \text{(โดยอสมการ (3.25) และ (3.12))}\\
&\le \min\left\{\frac{1-p^n}{x_0 q}, \frac{1-p^{n+1}-q^{n+1}}{(n+1)pq}\right\} \quad \text{(โดยอสมการ (3.31))}
\end{aligned}$$

และจะได้ว่า

$$\left|\Delta g_{C_{x_0}}(x)\right| \le \min\left\{\frac{1-p^n}{x_0 q}, \frac{1-p^{n+1}-q^{n+1}}{(n+1)pq}\right\} \tag{3.33}$$

2. กรณีที่ $x_0 < x \le n$

$$\begin{aligned}
0 &\le -\Delta g_{C_{x_0}}(x) \quad \text{(โดยอสมการ (3.13))}\\
&= -\sum_{k=0}^{x_0} \Delta g_k(x)\\
&= -\Delta g_0(x) - \Delta g_1(x) - \cdots - \Delta g_{x_0}(x)\\
&= -\Delta g_0(x) - \Delta g_1(x) - \cdots - \Delta g_{x_0}(x) - \Delta g_{x_0+1}(x) - \cdots - \Delta g_x(x)\\
&\quad + \Delta g_{x_0+1}(x) + \cdots + \Delta g_x(x)\\
&= -\sum_{k=0}^{x} \Delta g_k(x) + \Delta g_{x_0+1}(x) + \cdots + \Delta g_x(x)\\
&= -\Delta g_{C_x}(x) + \Delta g_{x_0+1}(x) + \cdots + \Delta g_x(x)\\
&\le \Delta g_x(x) \quad \text{(โดยอสมการ (3.13) และ (3.12))}
\end{aligned}$$

$$\leq \min\left\{\frac{1-p^n}{xq}, \frac{1-p^{n+1}-q^{n+1}}{(n+1)pq}\right\} \quad \text{(โดยอสมการ (3.32))}$$

$$\leq \min\left\{\frac{1-p^n}{x_0 q}, \frac{1-p^{n+1}-q^{n+1}}{(n+1)pq}\right\}$$

และจะได้ว่า

$$\left|\Delta g_{C_{x_0}}(x)\right| \leq \min\left\{\frac{1-p^n}{x_0 q}, \frac{1-p^{n+1}-q^{n+1}}{(n+1)pq}\right\} \tag{3.34}$$

ดังนั้น จากการอสมการ (3.33) และ (3.34) จะได้ว่า $\left|\Delta g_{C_{x_0}}(x)\right| \leq \min\left\{\frac{1-p^n}{x_0 q}, \frac{1-p^{n+1}-q^{n+1}}{(n+1)pq}\right\}$

สำหรับทุก $x \in \{1,\ldots,n\}$ ซึ่งจะได้อสมการ (3.24) ตามที่ต้องการ □

3.2 ดำเนินการหาขอบเขตไม่เอกรูปสำหรับระยะทางของคอลโมโกรอฟระหว่างฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มที่มีค่าจำนวนเต็มไม่เป็นลบและฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มทวินามโดยใช้วิธีของสไตน์และฟังก์ชัน w

3.3 นำผลลัพธ์ที่ได้ไปใช้ในการประมาณฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มไม่ต่อเนื่องสี่การแจกแจง คือ การแจกแจงเรขาคณิตไฮเพอร์ (Hypergeometric distribution) การแจกแจงเรขาคณิตไฮเพอร์เชิงลบ การแจกแจงโพลยา (Pólya distribution) และการแจกแจงทวินามบีต้า (Beta binomial distribution)

บทที่ 4
ผลการวิจัย

ผลการวิจัยที่ต้องการคือ ผลลัพธ์ของการประมาณการแจกแจงของตัวแปรสุ่มที่มีค่าจำนวนเต็มไม่เป็นลบด้วยการแจกแจงทวินามที่อยู่ในรูปของขอบเขตไม่เอกรูปสำหรับระยะทางของคอลโมโกรอฟระหว่างฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มที่มีค่าจำนวนเต็มไม่เป็นลบและฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มทวินาม

ทฤษฎีบท 4.1 ให้ X แทนตัวแปรสุ่มที่มีค่าจำนวนเต็มไม่เป็นลบดังที่นิยามแล้วข้างต้น และ $w(X)$ เป็นฟังก์ชัน w ที่สัมพันธ์กับตัวแปรสุ่ม X แล้วจะได้ผลลัพธ์ดังต่อไปนี้

(1). กรณีที่ $x_0 = 0$ จะได้ว่า

$$d_{K_0}(X,B) \le \frac{np+q^n-1}{(n-1)np^2} E\left|(n-X)p-\sigma^2 w(X)\right| + \frac{1-q^n}{np}|np-\mu| \tag{4.1}$$

และถ้า $np = \mu$ แล้ว

$$d_{K_0}(X,B) \le \frac{np+q^n-1}{(n-1)np^2} E\left|(n-X)p-\sigma^2 w(X)\right| \tag{4.2}$$

(2). กรณีที่ $x_0 \in \{1,\ldots,n\}$

$$d_{K_{x_0}}(X,B) \le \min\left\{\frac{1-p^n}{x_0 q}, \frac{1-p^{n+1}-q^{n+1}}{(n+1)pq}\right\} E\left|(n-X)p-\sigma^2 w(X)\right| + \xi(x_0)|np-\mu| \tag{4.3}$$

และถ้า $np = \mu$ แล้ว

$$d_{K_{x_0}}(X,B) \le \min\left\{\frac{1-p^n}{x_0 q}, \frac{1-p^{n+1}-q^{n+1}}{(n+1)pq}\right\} E\left|(n-X)p-\sigma^2 w(X)\right| \tag{4.4}$$

โดยที่ $\xi(x_0)$ ได้นิยามไว้ในอสมการ (3.23)

พิสูจน์ จากสมการ (3.1) เมื่อแทน h ด้วย $h_{C_{x_0}}$ และแทน x ด้วยตัวแปรสุ่ม X แล้วหาค่าคาดหมายตลอดสมการจะได้

$$E\left[h_{C_{x_0}}(X)-\mathbb{B}_{n,p}(h_{C_{x_0}})\right] = E\left[(n-X)pg(X+1)-qXg(X)\right]$$

หรือ $$\left|P(X\le x_0)-P(B\le x_0)\right| = \left|E\left[(n-X)pg(X+1)-qXg(X)\right]\right|$$

หรือ $$d_{K_{x_0}}(X,B) = \left|E\left[(n-X)pg(X+1)-qXg(X)\right]\right|$$

โดยที่ $g=g_{C_{x_0}}$ ได้นิยามไว้ในสมการ (3.5) ต่อไปแสดงการหาขอบเขตไม่เอกรูปสำหรับ $d_{K_{x_0}}(X,B)$ ดังนี้

$$\begin{aligned}
d_{K_{x_0}}(X,B) &= \left|E[npg(X+1)]-E[pXg(X+1)]-E[qXg(X)]\right| \\
&= \left|npE[g(X+1)]-pE[Xg(X+1)]-qE[Xg(X)]\right| \\
&= \left|npE[g(X+1)]-pE[Xg(X+1)]-E[Xg(X)]+pE[Xg(X)]\right| \quad (\text{โดย } q=1-p) \\
&= \left|npE[g(X+1)]-\left[pE[Xg(X+1)]-pE[Xg(X)]\right]-E[Xg(X)]\right| \\
&= \left|npE[g(X+1)]-\left[pE[X(g(X+1)-g(X))]\right]-E[Xg(X)]\right| \\
&= \left|npE[g(X+1)]-pE[X\Delta g(X)]-E[Xg(X)]\right| \\
&= \left|npE[g(X+1)]-pE[X\Delta g(X)]-E[(X-\mu+\mu)g(X)]\right| \\
&= \left|npE[g(X+1)]-pE[X\Delta g(X)]-E[(X-\mu)g(X)+\mu g(X)]\right| \\
&= \left|npE[g(X+1)]-pE[X\Delta g(X)]-E[(X-\mu)g(X)]-\mu E[g(X)]\right| \\
&= |npE[g(X+1)]-pE[X\Delta g(X)]-E[(X-\mu)g(X)]-\mu E[g(X)]+npE[g(X)] \\
&\quad -npE[g(X)]| \\
&= |npE[g(X+1)-g(X)]-pE[X\Delta g(X)]-E[(X-\mu)g(X)]-\mu E[g(X)] \\
&\quad +npE[g(X)]| \\
&= \left|npE[\Delta g(X)]-pE[X\Delta g(X)]-E[(X-\mu)g(X)]+(np-\mu)E[g(X)]\right|
\end{aligned}$$

เนื่องจาก $E\left|w(X)\Delta g(X)\right| \le \sup_{x\ge 1}|\Delta g(x)|\,E\,|w(X)| = \sup_{x\ge 1}|\Delta g(x)| < \infty$ ดังนั้นโดยอสมการ (2.5) จะได้ว่า

$$\begin{aligned}
d_{K_{x_0}}(X,B) &= \left|npE[\Delta g(X)]-pE[X\Delta g(X)]-\sigma^2E[w(X)\Delta g(X)]+(np-\mu)E[g(X)]\right| \\
&= \left|E\left[np\Delta g(X)-pX\Delta g(X)-\sigma^2 w(X)\Delta g(X)\right]+(np-\mu)E[g(X)]\right| \\
&\le \left|E\left[np\Delta g(X)-pX\Delta g(X)-\sigma^2 w(X)\Delta g(X)\right]\right|+\left|(np-\mu)E[g(X)]\right| \\
&\le E\left|\left[np\Delta g(X)-pX\Delta g(X)-\sigma^2 w(X)\Delta g(X)\right]\right|+\left|np-\mu\right|E\left|g(X)\right|
\end{aligned}$$

$$\leq E\left\{\left|(n-X)p-\sigma^2 w(X)\right|\left|\Delta g(X)\right|\right\}+\left|np-\mu\right|E\left|g(X)\right|$$

เมื่อ $x_0 = 0$ โดยใช้บทตั้ง 3.4 (อสมการ (3.21) และ (3.22)) จะได้ว่า

$$d_{K_0}(X,B) \quad \leq \sup_{x\geq 1}|\Delta g(x)|\,E\left|(n-X)p-\sigma^2 w(X)\right|+\sup_{x\geq 1}\{g(x)\}\left|np-\mu\right|$$

$$\leq \frac{np+q^n-1}{(n-1)np^2}E\left|(n-X)p-\sigma^2 w(X)\right|+\frac{1-q^n}{np}\left|np-\mu\right| \tag{4.5}$$

เมื่อ $x_0 \in \{1,...,n\}$ โดยใช้บทตั้ง 3.4 (อสมการ (3.23) และ (3.24)) จะได้ว่า

$$d_{K_{x_0}}(X,B) \leq \sup_{x\geq 1}|\Delta g(x)|\,E\left|(n-X)p-\sigma^2 w(X)\right|+\sup_{x\geq 1}\{g(x)\}\left|np-\mu\right|$$

$$\leq \min\left\{\frac{1-p^n}{x_0 q},\frac{1-p^{n+1}-q^{n+1}}{(n+1)pq}\right\}E\left|(n-X)p-\sigma^2 w(X)\right|+\xi(x_0)\left|np-\mu\right| \tag{4.6}$$

ดังนั้นจากอสมการ (4.5) และ (4.6) จะได้ทฤษฎีบท 4.1 ตามต้องการ □

บทแทรก 4.1 ถ้า $(n-x)p-\sigma^2 w(x)\geq 0$ สำหรับทุก $x\in\mathscr{X}$ หรือ $(n-x)p-\sigma^2 w(x)<0$ สำหรับทุก $x\in\mathscr{X}$ แล้วจะได้ว่า

(1). กรณีที่ $x_0 = 0$ จะได้ว่า

$$d_{K_0}(X,B)\leq \frac{np+q^n-1}{(n-1)np^2}\left|(n-\mu)p-\sigma^2\right|+\frac{1-q^n}{np}\left|np-\mu\right| \tag{4.7}$$

และถ้า $np=\mu$ แล้ว

$$d_{K_0}(X,B)\leq \frac{np+q^n-1}{(n-1)np^2}\left|\mu q-\sigma^2\right| \tag{4.8}$$

(2). กรณีที่ $x_0 \in \{1,...,n\}$

$$d_{K_{x_0}}(X,B)\leq \min\left\{\frac{1-p^n}{x_0 q},\frac{1-p^{n+1}-q^{n+1}}{(n+1)pq}\right\}\left|(n-\mu)p-\sigma^2\right|+\xi(x_0)\left|np-\mu\right| \tag{4.9}$$

และถ้า $np=\mu$ แล้ว

$$d_{K_{x_0}}(X,B)\leq \min\left\{\frac{1-p^n}{x_0 q},\frac{1-p^{n+1}-q^{n+1}}{(n+1)pq}\right\}\left|\mu q-\sigma^2\right| \tag{4.10}$$

พิสูจน์ ให้ $(n-x)p-\sigma^2 w(x) \geq 0$ สำหรับทุก $x \in \mathscr{X}$ แล้วจะได้ว่า

$$
\begin{aligned}
E\left|(n-X)p-\sigma^2 w(X)\right| &= E\left[(n-X)p-\sigma^2 w(X)\right] \\
&= E\left[(n-X)p\right]-E\left[\sigma^2 w(X)\right] \\
&= E(np-Xp)-\sigma^2 E\left[w(X)\right] \\
&= E(np)-pE(X)-\sigma^2 \text{ (โดย } E\left[w(X)\right]=1\text{)} \\
&= np-\mu p-\sigma^2 \text{ (โดย } E(X)=\mu\text{)} \\
&= (n-\mu)p-\sigma^2 \qquad (4.11)
\end{aligned}
$$

ให้ $(n-x)p-\sigma^2 w(x) < 0$ สำหรับทุก $x \in \mathscr{X}$ แล้วจะได้ว่า

$$
\begin{aligned}
E\left|(n-X)p-\sigma^2 w(X)\right| &= E\left[\sigma^2 w(X)-(n-X)p\right] \\
&= E\left[\sigma^2 w(X)\right]-E\left[(n-X)p\right] \\
&= \sigma^2 E\left[w(X)\right]-E\left[(np-Xp)\right] \\
&= \sigma^2-E(np)+pE(X) \\
&= \sigma^2-np+\mu p \\
&= \sigma^2-(n-\mu)p \qquad (4.12)
\end{aligned}
$$

จากสมการ (4.11) และ (4.12) จะได้ว่า $E\left|(n-X)p-\sigma^2 w(X)\right| = \left|(n-\mu)p-\sigma^2\right|$ และถ้า $np=\mu$ แล้วจะได้ว่า $E\left|(n-X)p-\sigma^2 w(X)\right| = \left|\mu q-\sigma^2\right|$ ซึ่งเมื่อแทนลงในทฤษฎีบท 4.1 จะได้ผลลัพธ์ดังที่ได้แสดงไว้ในบทแทรก 4.1 □

หมายเหตุ พิจารณาทฤษฎีบท 4.1 กรณีที่ $np=\mu$ ผลลัพธ์ที่อยู่ในรูปของขอบเขตเอกรูปสามารถเขียนได้ดังนี้

$$d_{K_{x_0}}(X,B) \leq \frac{1-p^{n+1}-q^{n+1}}{(n+1)pq} E\left|(n-X)p-\sigma^2 w(X)\right| \qquad (4.13)$$

สำหรับทุก $x_0 \in \{0,1,\dots,n\}$ ซึ่งผลลัพธ์ในทฤษฎีบท 4.1 จะดีกว่าผลลัพธ์ในอสมการ (4.13) ก็ต่อเมื่อ

$$\frac{np+q^n-1}{(n-1)np^2} < \frac{1-p^{n+1}-q^{n+1}}{(n+1)pq} \text{ และ } \frac{1-p^n}{x_0} < \frac{1-p^{n+1}-q^{n+1}}{(n+1)p}$$

และจากบทแทรก 1 (2) ในงานวิจัยของ Teerapabolarn (2011) จะได้ว่า

$\frac{np+q^n-1}{(n-1)np^2} < \frac{1-p^n}{nq} < \frac{1-p^{n+1}-q^{n+1}}{(n+1)pq}$ ซึ่งแสดงว่าอสมการแรกเป็นจริงเสมอ และอสมการ

$\frac{1-p^n}{x_0} < \frac{1-p^{n+1}-q^{n+1}}{(n+1)p}$ ก็ต่อเมื่อ $x_0 > \frac{(n+1)p(1-p^n)}{1-p^{n+1}-q^{n+1}}$ ดังนั้นผลลัพธ์ในทฤษฎีบท 4.1 จะดีกว่าผลลัพธ์ในอสมการ (4.13) ก็ต่อเมื่อ $x_0 > \frac{(n+1)p(1-p^n)}{1-p^{n+1}-q^{n+1}}$

การประยุกต์ใช้ผลการวิจัย

การประยุกต์ใช้ผลการวิจัยที่จะกล่าวถึงต่อไปนี้ เป็นการนำเสนอตัวอย่างของการประยุกต์ใช้ผลลัพธ์ของการประมาณทวินามในทฤษฎีบท 4.1 และบทแทรก 4.1 เพื่อประมาณฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มที่มีการแจกแจงสี่การแจกแจง คือ การแจกแจงเรขาคณิตไฮเพอร์ การแจกแจงเรขาคณิตไฮเพอร์เชิงลบ การแจกแจงโพลยา และการแจกแจงทวินามบีต้า

ตัวอย่าง 4.1 การแจกแจงเรขาคณิตไฮเพอร์

สมมุติว่าถุงใบหนึ่งบรรจุลูกบอลสีขาว m ลูก และลูกบอลสีดำ $N-m$ ลูก หยิบลูกบอลอย่างสุ่มขึ้นมาหนึ่งลูกแบบไม่ใส่กลับคืนลงไปในถุง ทำซ้ำในลักษณะนี้จำนวน n ครั้ง ให้ X แทนจำนวนลูกบอลสีขาวที่หยิบได้จากการสุ่มหยิบลูกบอล n ครั้ง แล้วจะได้ว่า X เป็นตัวแปรสุ่มเรขาคณิตไฮเพอร์ที่มีพารามิเตอร์ N, m และ n และมีฟังก์ชันมวลความน่าจะเป็น ดังนี้

$$p_X(x) = \frac{\binom{m}{x}\binom{N-m}{n-x}}{\binom{N}{n}} \quad , x = 0,1,\ldots,\min\{n,m\} \tag{4.14}$$

โดยมีค่าเฉลี่ย $\mu = E(X) = \frac{nm}{N}$ และความแปรปรวน $\sigma^2 = Var(X) = \frac{nm(N-n)(N-m)}{N^2(N-1)}$

จากสมการ (2.3) และ (2.4) จะได้ว่าฟังก์ชัน w ที่สัมพันธ์กับตัวแปรสุ่มเรขาคณิตไฮเพอร์ คือ $w(x) = \frac{(n-x)(m-x)}{N\sigma^2} \geq 0$

ในกรณีที่ $\min\{n,m\} = n$ และให้ $p = \frac{m}{N}$ เมื่อแทนค่า $p = \frac{m}{N}$ ลงในทฤษฎีบท 4.1 ซึ่งจะได้ว่า $(n-x)p - \sigma^2 w(x) = \frac{(n-x)m}{N} - \frac{(n-x)(m-x)}{N} = \frac{(n-x)x}{N} \geq 0$ สำหรับทุก $0 \leq x \leq n$ ดังนั้น โดยประยุกต์ใช้บทแทรก 4.1 จะได้ผลลัพธ์ดังต่อไปนี้

บทแทรก 4.2 ถ้า $\min\{n,m\}=n$ ให้ $p=\frac{m}{N}$ แล้วจะได้ว่า

$$d_{K_{x_0}}(X,B) \le \begin{cases} \dfrac{(np+q^n-1)(N-m)}{m(N-1)} & , \quad x_0=0 \\ \min\left\{\dfrac{1-p^n}{x_0}, \dfrac{1-p^{n+1}-q^{n+1}}{(n+1)p}\right\}\dfrac{nm(n-1)}{N(N-1)} & , \quad 1\le x_0 \le n \end{cases} \tag{4.15}$$

จากอสมการ (4.15) จะสังเกตได้ว่าการประมาณฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มเรขาคณิตไฮเพอร์ด้วยฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มทวินามจะมีผลการประมาณที่ดีเมื่อ N มีค่ามากและ m และ n มีค่าน้อย

ตัวอย่างเชิงตัวเลข

● ให้ $m=15,\ n=10$ และ $N=60$ จะได้ $p=0.25$ แล้วผลลัพธ์เชิงตัวเลขในอสมการ (4.15) สามารถแสดงได้ดังนี้

$$d_{K_{x_0}}(X,B) \le \begin{cases} 0.079135 & , \quad x_0=0 \\ 0.132818 & , \quad x_0=1,2 \\ \dfrac{0.381356}{x_0} & , \quad x_0=3,\ldots,10 \end{cases}$$

● ให้ $m=15,\ n=10$ และ $N=120$ จะได้ $p=0.125$ แล้วผลลัพธ์เชิงตัวเลขในอสมการ (4.15) สามารถแสดงได้ดังนี้

$$d_{K_{x_0}}(X,B) \le \begin{cases} 0.030181 & , \quad x_0=0 \\ 0.052928 & , \quad x_0=1 \\ \dfrac{0.094538}{x_0} & , \quad x_0=2,\ldots,10 \end{cases}$$

ผลลัพธ์เชิงตัวเลขได้แสดงให้เห็นว่าการประมาณฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มเรขาคณิตไฮเพอร์ด้วยฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มทวินามจะมีผลการประมาณที่เหมาะสมเมื่อ N มีค่ามากและ m และ n มีค่าน้อย ซึ่งสอดคล้องกับผลการประมาณในอสมการ (4.15)

ในกรณีที่ $\min\{n,m\}=m$ และให้ $p=\frac{n}{N}$ เมื่อแทนค่า n ด้วย m และ $p=\frac{n}{N}$ ลงในทฤษฎีบท 4.1 ซึ่งจะได้ว่า $(n-x)p-\sigma^2 w(x)=\frac{(m-x)n}{N}-\frac{(n-x)(m-x)}{N}=\frac{(m-x)x}{N}\ge 0$ สำหรับทุก $0\le x\le m$ ดังนั้นโดยประยุกต์ใช้บทแทรก 4.1 จะได้ผลลัพธ์ดังต่อไปนี้

บทแทรก 4.3 ถ้า $\min\{n,m\}=n$ ให้ $p=\dfrac{n}{N}$ แล้วจะได้ว่า

$$d_{K_{x_0}}(X,B)\le\begin{cases}\dfrac{(mp+q^m-1)(N-n)}{n(N-1)} & , \quad x_0=0\\[2ex] \min\left\{\dfrac{1-p^m}{x_0},\dfrac{1-p^{m+1}-q^{m+1}}{(m+1)p}\right\}\dfrac{nm(m-1)}{N(N-1)} & , \quad 1\le x_0\le m\end{cases} \tag{4.16}$$

จากอสมการ (4.16) จะสังเกตได้ว่าการประมาณฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มเรขาคณิตไฮเพอร์ด้วยฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มทวินามจะมีผลการประมาณที่ดีเมื่อ N มีค่ามากและ m และ n มีค่าน้อยเช่นเดียวกัน

ตัวอย่างเชิงตัวเลข

● ให้ $m=15,\ n=10$ และ $N=100$ จะได้ $p=0.15$ แล้วผลลัพธ์เชิงตัวเลขในอสมการ (4.16) สามารถแสดงได้ดังนี้

$$d_{K_{x_0}}(X,B)\le\begin{cases}0.121578 & , \quad x_0=0\\ 0.081821 & , \quad x_0=1,2\\ \dfrac{0.212121}{x_0} & , \quad x_0=3,\ldots,15\end{cases}$$

● ให้ $m=15,\ n=10$ และ $N=150$ จะได้ $p=0.1$ แล้วผลลัพธ์เชิงตัวเลขในอสมการ (4.16) สามารถแสดงได้ดังนี้

$$d_{K_{x_0}}(X,B)\le\begin{cases}0.066325 & , \quad x_0=0\\ 0.047843 & , \quad x_0=1\\ \dfrac{0.093960}{x_0} & , \quad x_0=2,\ldots,15\end{cases}$$

ผลลัพธ์เชิงตัวเลขได้แสดงให้เห็นว่าการประมาณฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มเรขาคณิตไฮเพอร์ด้วยฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มทวินามจะมีผลการประมาณที่เหมาะสมเมื่อ N มีค่ามากและ m และ n มีค่าน้อย ซึ่งสอดคล้องกับผลการประมาณในอสมการ (4.16)

ตัวอย่าง 4.2 การแจกแจงเรขาคณิตไฮเพอร์เชิงลบ

สมมุติว่าถุงใบหนึ่งบรรจุลูกบอลสีขาว n ลูก และลูกบอลสีดำ $N-n$ ลูก หยิบลูกบอลอย่างสุ่มขึ้นมาหนึ่งลูกแบบไม่ใส่กลับคืนในลงไปในถุง ทำซ้ำในลักษณะนี้จนกระทั่งได้ลูกบอลสีดำจำนวน m ลูก ให้ X แทนจำนวนลูกบอลสีขาวที่หยิบได้ก่อนที่จะได้ลูกบอลสีดำลูกที่ m แล้วจะได้ว่า X เป็นตัวแปรสุ่มเรขาคณิตไฮเพอร์เชิงลบที่มีพารามิเตอร์ N, n และ m และมีฟังก์ชันมวลความน่าจะเป็น ดังนี้

$$p_X(x)=\frac{\binom{m+x-1}{x}\binom{N-m-x}{n-x}}{\binom{N}{n}} \quad , \; x=0,1,\ldots,n \tag{4.17}$$

โดยที่ $\mu=E(X)=\dfrac{nm}{N-n+1}$ และ $\sigma^2=Var(X)=\dfrac{nm(N-n-m+1)(N+1)}{(N-n+1)^2(N-n+2)}$ คือ ค่าเฉลี่ยและค่าความแปรปรวนของ X

จากสมการ (2.3) และ (2.4) จะได้ว่าฟังก์ชัน w ที่สัมพันธ์กับตัวแปรสุ่มเรขาคณิตไฮเพอร์เชิงลบ คือ $w(x)=\dfrac{(m+x)(n-x)}{(N-n+1)\sigma^2}\geq 0$ ให้ $p=\dfrac{m}{N-n+1}$ และแทนค่า p ลงในทฤษฎีบท 4.1 ซึ่งจะได้ว่า $(n-x)p-\sigma^2w(x)=\dfrac{(n-x)m}{N-n+1}-\dfrac{(m+x)(n-x)}{N-n+1}=-\dfrac{(n-x)x}{N-n+1}\leq 0$ สำหรับทุก $0\leq x\leq n$ ดังนั้นโดยประยุกต์ใช้บทแทรก 4.1 จะได้ผลลัพธ์ดังต่อไปนี้

บทแทรก 4.4 ให้ $x_0\in\{0,1,\ldots,n\}$ ถ้า $p=\dfrac{m}{N-n+1}$ แล้วจะได้ว่า

$$d_{K_{x_0}}(X,B)\leq\begin{cases}\dfrac{(np+q^n-1)(N-m-n+1)}{m(N-n+2)} & , \quad x_0=0\\ \min\left\{\dfrac{1-p^n}{x_0},\dfrac{1-p^{n+1}-q^{n+1}}{(n+1)p}\right\}\dfrac{mn(n-1)}{(N-n+1)(N-n+2)} & , \quad 1\leq x_0\leq n\end{cases} \tag{4.18}$$

จากอสมการ (4.18) จะสังเกตได้ว่าการประมาณฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มเรขาคณิตไฮเพอร์เชิงลบด้วยฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มทวินามจะมีผลการประมาณที่ดีเมื่อ N มีค่ามากและ m และ n มีค่าน้อย

ตัวอย่างเชิงตัวเลข

● ให้ $m=10,\; n=15$ และ $N=100$ จะได้ $p=0.116279$ แล้วผลลัพธ์เชิงตัวเลขในอสมการ (4.18) สามารถแสดงได้ดังนี้

$$d_{K_{x_0}}(X,B) \le \begin{cases} 0.078687 & , \quad x_0 = 0 \\ 0.129987 & , \quad x_0 = 1,2 \\ \dfrac{0.280674}{x_0} & , \quad x_0 = 3,\ldots,15 \end{cases}$$

- ให้ $m = 10,\ n = 15$ และ $N = 200$ จะได้ $p = 0.053763$ แล้วผลลัพธ์เชิงตัวเลขในอสมการ (4.18) สามารถแสดงได้ดังนี้

$$d_{K_{x_0}}(X,B) \le \begin{cases} 0.022867 & , \quad x_0 = 0 \\ 0.041197 & , \quad x_0 = 1 \\ \dfrac{0.060376}{x_0} & , \quad x_0 = 2,\ldots,15 \end{cases}$$

ผลลัพธ์เชิงตัวเลขได้แสดงให้เห็นว่าการประมาณฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มเรขาคณิตไฮเพอร์เชิงลบด้วยฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มทวินามจะมีผลการประมาณที่เหมาะสมเมื่อ N มีค่ามากและ m และ n มีค่าน้อย ซึ่งสอดคล้องกับผลการประมาณในอสมการ (4.18)

ตัวอย่าง 4.3 การแจกแจงโพลยา

สมมุติว่าถุงใบหนึ่งบรรจุลูกบอลสีขาว m ลูก และลูกบอลสีดำ $N-m$ ลูก หยิบลูกบอลอย่างสุ่มขึ้นมาหนึ่งลูก แล้วใส่กลับคืนลงในถุงพร้อมด้วยลูกบอลสีเดียวกับลูกบอลที่สุ่มได้อีกหนึ่งลูก ทำซ้ำในลักษณะนี้เป็นจำนวน n ครั้ง ให้ X แทนจำนวนลูกบอลสีขาวที่สุ่มได้จากการสุ่มหยิบลูกบอล n ครั้ง แล้วจะได้ว่า X เป็นตัวแปรสุ่มโพลยาที่มีพารามิเตอร์ N, n และ m โดยมีฟังก์ชันมวลความน่าจะเป็น ดังนี้

$$p_X(x) = \frac{\binom{m+x-1}{x}\binom{N-m+n-x-1}{n-x}}{\binom{N+n-1}{n}},\ x = 0,\ldots,n \tag{4.19}$$

โดยมีค่าเฉลี่ย $\mu = E(X) = \dfrac{nm}{N}$ และความแปรปรวน $\sigma^2 = Var(X) = \dfrac{nm(N+n)(N-m)}{N^2(N+1)}$

จากสมการ (2.3) และ (2.4) จะได้ว่าฟังก์ชัน w ที่สัมพันธ์กับตัวแปรสุ่มโพลยา คือ $w(x) = \dfrac{(m+x)(n-x)}{N\sigma^2} \ge 0$ ให้ $p = \dfrac{m}{N}$ และแทนค่า p ลงใน ทฤษฎีบท 4.1 ซึ่งจะได้ว่า $(n-x)p - \sigma^2 w(x) = \dfrac{(n-x)m}{N} - \dfrac{(m+x)(n-x)}{N} = -\dfrac{(n-x)x}{N} \le 0$ สำหรับทุก $0 \le x \le n$ ดังนั้นโดยประยุกต์ใช้บทแทรก 4.1 จะได้ผลลัพธ์ดังต่อไปนี้

บทแทรก 4.5 ให้ $x_0 \in \{0,1,...,n\}$ ถ้า $p = \frac{m}{N}$ แล้วจะได้ว่า

$$d_{K_{x_0}}(X,B) \le \begin{cases} \dfrac{(np+q^n-1)(N-m)}{m(N+1)} & , \quad x_0 = 0 \\ \min\left\{\dfrac{1-p^n}{x_0}, \dfrac{1-p^{n+1}-q^{n+1}}{(n+1)p}\right\}\dfrac{mn(n-1)}{N(N+1)} & , \quad 1 \le x_0 \le n \end{cases} \tag{4.20}$$

จากอสมการ (4.20) จะสังเกตได้ว่าการประมาณฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มโพลยาด้วยฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มทวินามจะมีผลการประมาณที่ดี เมื่อ N มีค่ามากและ m และ n มีค่าน้อย

ตัวอย่างเชิงตัวเลข

● ให้ $m = 10,\ n = 10$ และ $N = 100$ จะได้ $p = 0.10$ แล้วผลลัพธ์เชิงตัวเลขในอสมการ (4.20) สามารถแสดงได้ดังนี้

$$d_{K_{x_0}}(X,B) \le \begin{cases} 0.031070 & , \quad x_0 = 0 \\ 0.055587 & , \quad x_0 = 1 \\ \dfrac{0.089109}{x_0} & , \quad x_0 = 2,...,10 \end{cases}$$

● ให้ $m = 10,\ n = 10$ และ $N = 150$ จะได้ $p = 0.066667$ แล้วผลลัพธ์เชิงตัวเลขในอสมการ (4.20) สามารถแสดงได้ดังนี้

$$d_{K_{x_0}}(X,B) \le \begin{cases} 0.015602 & , \quad x_0 = 0 \\ 0.028817 & , \quad x_0 = 1 \\ \dfrac{0.039735}{x_0} & , \quad x_0 = 2,...,10 \end{cases}$$

ผลลัพธ์เชิงตัวเลขได้แสดงให้เห็นว่าการประมาณฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มโพลยาด้วยฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มทวินามจะมีผลการประมาณที่เหมาะสม เมื่อ N มีค่ามากและ m และ n มีค่าน้อย ซึ่งสอดคล้องกับผลการประมาณในอสมการ (4.20)

ตัวอย่าง 4.4 การแจกแจงทวินามบีต้า

ถ้า X เป็นตัวแปรสุ่มที่ได้มาจากการแจกแจงทวินามที่มีพารามิเตอร์ $n \in \mathrm{N}$ แต่ความน่าจะเป็นของความสำเร็จ p เป็นตัวแปรสุ่มที่มีการแจกแจงบีต้าที่มีพารามิเตอร์ $\alpha > 0$ และ $\beta > 0$ แล้วจะได้ว่า X เป็นตัวแปรสุ่มทวินามบีต้าที่มีพารามิเตอร์ n, α และ β โดยมีฟังก์ชันมวลความน่าจะเป็น ดังนี้

$$p_X(x)=\binom{n}{x}\frac{B(\alpha+x,\beta+n-x)}{B(\alpha,\beta)},\ x=0,...,n \tag{4.21}$$

และมีค่าเฉลี่ย $\mu=E(X)=\dfrac{n\alpha}{\alpha+\beta}$ และความแปรปรวน $\sigma^2=Var(X)=\dfrac{n\alpha\beta(n+\alpha+\beta)}{(\alpha+\beta+1)(\alpha+\beta)^2}$

จากสมการ (2.3) และ (2.4) จะได้ว่าฟังก์ชัน w ที่สัมพันธ์กับตัวแปรสุ่มทวินามบีต้า คือ $w(x)=\dfrac{(\alpha+x)(n-x)}{(\alpha+\beta)\sigma^2}\geq 0$ ให้ $p=\dfrac{\alpha}{\alpha+\beta}$ และแทนค่า p ลงในทฤษฎีบท 4.1 ซึ่งจะได้ว่า $(n-x)p-\sigma^2w(x)=\dfrac{(n-x)\alpha}{\alpha+\beta}-\dfrac{(\alpha+x)(n-x)}{\alpha+\beta}=-\dfrac{(n-x)x}{\alpha+\beta}\leq 0$ สำหรับ $0\leq x\leq n$ ดังนั้นโดยประยุกต์ใช้บทแทรก 4.1 จะได้ผลลัพธ์ดังต่อไปนี้

บทแทรก 4.6 ให้ $x_0\in\{0,1,...,n\}$ ถ้า $p=\dfrac{\alpha}{\alpha+\beta}$ แล้วจะได้ว่า

$$d_{K_{x_0}}(X,B)\leq\begin{cases}(np+q^n-1)\dfrac{\beta}{\alpha(\alpha+\beta+1)} & ,\quad x_0=0\\ \min\left\{\dfrac{1-p^n}{x_0},\dfrac{1-p^{n+1}-q^{n+1}}{(n+1)p}\right\}\dfrac{\alpha n(n-1)}{(\alpha+\beta)(\alpha+\beta+1)} & ,\quad 1\leq x_0\leq n\end{cases} \tag{4.22}$$

จากอสมการ (4.22) จะสังเกตได้ว่าการประมาณฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มทวินามบีต้าด้วยฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มทวินามจะมีผลการประมาณที่ดีเมื่อ $\alpha+\beta$ มีค่ามากและ α และ n มีค่าน้อย

ตัวอย่างเชิงตัวเลข

- ให้ $n=30,\ \alpha=100,\ \beta=1000$ และ $\alpha+\beta=1100$ จะได้ $p=0.090909$ แล้วผลลัพธ์เชิงตัวเลขในอสมการ (4.22) สามารถแสดงได้ดังนี้

$$d_{K_{x_0}}(X,B)\leq\begin{cases}0.016209 & ,\quad x_0=0\\ 0.024162 & ,\quad x_0=1,2\\ \dfrac{0.071836}{x_0} & ,\quad x_0=3,...,30\end{cases}$$

- ให้ $n=10,\ \alpha=100,\ \beta=1000$ และ $\alpha+\beta=1100$ จะได้ $p=0.090909$ แล้วผลลัพธ์เชิงตัวเลขในอสมการ (4.22) สามารถแสดงได้ดังนี้

$$d_{K_{x_0}}(X,B)\leq\begin{cases}0.004566 & ,\quad x_0=0\\ 0.004827 & ,\quad x_0=1\\ \dfrac{0.007431}{x_0} & ,\quad x_0=2,...,10\end{cases}$$

ผลลัพธ์เชิงตัวเลขได้แสดงให้เห็นว่าการประมาณฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มทวินามบีต้าด้วยฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มทวินามจะมีผลการประมาณที่เหมาะสม เมื่อ β มีค่ามากและ α และ n มีค่าน้อย ซึ่งสอดคล้องกับผลการประมาณในอสมการ (4.22)

บทที่ 5
สรุปและอภิปรายผล

สรุปและอภิปรายผล

ขอบเขตไม่เอกรูปสำหรับระยะทางของคอลโมโกรอฟระหว่างฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มที่มีค่าจำนวนเต็มไม่เป็นลบและฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มทวินาม เป็นผลลัพธ์ในอีกรูปแบบหนึ่งที่หาได้โดยใช้วิธีของสไตน์และฟังก์ชัน w และขอบเขตไม่เอกรูปดังกล่าวสามารถใช้เป็นเกณฑ์ในการวัดความแม่นยำของการประมาณฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มที่มีค่าจำนวนเต็มไม่เป็นลบด้วยฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มทวินาม ซึ่งผลลัพธ์ที่ได้จากการศึกษาครั้งนี้ สามารถแสดงได้ดังนี้

กรณีที่ $x_0 = 0$ จะได้ว่า

$$d_{K_0}(X,B) \le \frac{np+q^n-1}{(n-1)np^2} E\left|(n-X)p-\sigma^2 w(X)\right| + \frac{1-q^n}{np}|np-\mu| \tag{5.1}$$

และถ้า $np=\mu$ แล้ว

$$d_{K_0}(X,B) \le \frac{np+q^n-1}{(n-1)np^2} E\left|(n-X)p-\sigma^2 w(X)\right| \tag{5.2}$$

กรณีที่ $x_0 \in \{1,\dots,n\}$

$$d_{K_{x_0}}(X,B) \le \min\left\{\frac{1-p^n}{x_0 q}, \frac{1-p^{n+1}-q^{n+1}}{(n+1)pq}\right\} E\left|(n-X)p-\sigma^2 w(X)\right| + \xi(x_0)|np-\mu| \tag{5.3}$$

และถ้า $np=\mu$ แล้ว

$$d_{K_{x_0}}(X,B) \le \min\left\{\frac{1-p^n}{x_0 q}, \frac{1-p^{n+1}-q^{n+1}}{(n+1)pq}\right\} E\left|(n-X)p-\sigma^2 w(X)\right| \tag{5.4}$$

โดยที่

$$\xi(x_0) = \begin{cases} \dfrac{(n-x_0-1)\left[1-P(B \le x_0)\right]}{(n-x_0)\left[(n-1)p-x_0\right]} & , x_0 < (n-1)p \\ \dfrac{(x_0+2)P(B \le x_0)}{(x_0+1)\left[x_0+2-(n-1)p\right]} & , x_0 \ge (n-1)p \end{cases} \quad \text{และ} \quad P(B \le x_0) = \sum_{k=0}^{x_0} p_B(k)$$

และสังเกตได้ว่าขอบเขตไม่เอกรูปสำหรับการประมาณฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มที่มีค่าจำนวนเต็มไม่เป็นลบด้วยฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มทวินามดังที่ได้แสดงข้างต้นจะมีค่าน้อยลงเมื่อกำหนดให้ $np = \mu$

สำหรับการประยุกต์ใช้ผลลัพธ์ในเชิงทฤษฎี เป็นการนำผลลัพธ์ที่ได้ไปประมาณฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มเรขาคณิตไฮเพอร์ ฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มเรขาคณิตไฮเพอร์เชิงลบ ฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มโพลยา และฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มทวินามบีต้า นอกจากนี้ได้มีการนำเสนอผลลัพธ์เชิงตัวเลขเพื่อแสดงความแม่นยำของการประมาณข้างต้น ถ้าขอบเขตไม่เอกรูปสำหรับระยะทางคอลโมโกรอฟในแต่ละผลลัพธ์มีค่าน้อย แสดงว่าในการประมาณข้างต้นด้วยฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มทวินามมีความแม่นยำและเหมาะสม

ข้อเสนอแนะ

เนื่องจากขอบเขตไม่เอกรูปสำหรับระยะทางของคอลโมโกรอฟในการศึกษาครั้งนี้ เป็นขอบเขตไม่เอกรูปที่มีฟังก์ชัน w ดังนั้นในการนำผลลัพธ์เชิงทฤษฎีไปประยุกต์ใช้จะต้องหาฟังก์ชัน w ที่สัมพันธ์กับการแจกแจงที่ต้องการประมาณ

บรรณานุกรม

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*(5), 351-356.

Chen, L. H. Y. (1975). Poisson approximation for dependent trials. *Annals of Probability, 3*(3), 535-545.

Ehm, W. (1991). Binomial approximation to the Poisson binomial distribution. *Statistics & Probability Letters, 11*(1), 7-16.

Majsnerowska, M. (1998). A note on Poisson approximation by w-functions. *Applications Mathematical, 25*(3), 387-392.

Prukpousana, K., & Teerapabolarn, K. (2010). Binomial Approximation to the Negative Hypergeometric Distribution. *Khon Kaen University Science Journal, 38*(4), 606-616. (In thai).

Soon, S. Y. T. (1996). Binomial approximation for dependent indicators. *Statistica Sinica, 6*(3), 703-714

Stein, C. M. (1972). A bound for the error in normal approximation to the distribution of a sum of dependent random variables. *In Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, 3*, 583-602.

Stein, C. M. (1986). *Approximate computation of expectations*. Hayward California : IMS.

Teerapabolarn, K. (2011). A non-uniform bound on pointwise approximation of generalized binomial distribution by binomial distribution. *Srinakharinwirot Science Journal, 27*(1), 37-52. (In thai).

Teerapabolarn, K., & Wongkasem, P. (2011). On pointwise binomial approximation by w-function. *International Journal of Pure and Applied Mathematics, 71*(1), 57-66.

Wongkasem, P., Teerapabolarn, K., Gulasirima, R., & (2008). On approximating a Generalized Binomial by Binomial and Poisson Distributions. *International Journal of statistics and Systems, 3*(2), 113-124.