บทที่ 1

บทนำ

1.1. ความเป็นมาและความสำคัญของปัญหา

ในการอนุมานเกี่ยวกับประชากรนั้นจะอาศัยข้อมูลที่รวบรวมมาจากกลุ่มตัวอย่างมาทำการศึกษา หรือที่เรียกว่า ข้อมูลตัวอย่าง (sample data) และนำผลของการศึกษาข้อมูลตัวอย่างนั้นไปใช้ในการอธิบายลักษณะที่สนใจศึกษา หรือที่เรียกว่า ประชากร (population) การอาศัยข้อมูลเพียงบางส่วนหรือข้อมูลตัวอย่างเพื่อนำไปอธิบายลักษณะของประชากรที่สนใจศึกษานี้เปรียบเสมือนการสรุปผลจากกรณีเฉพาะไปสู่กรณีทั่วไป เราจึงเรียกว่า การอนุมาน (inference) ดังนั้นการศึกษาข้อมูลตัวอย่างและใช้วิธีการทางสถิติมาทำการหาข้อสรุปเกี่ยวกับประชากรจากตัวอย่างนั้น จึงเรียกกันว่า การอนุมานเชิงสถิติ (statistical inference) ในการอนุมานอาจเป็นการทำนาย หรือการตัดสินใจเกี่ยวกับลักษณะบางอย่างของประชากรตามที่ผู้วิจัยสนใจ การอนุมานทางสถิติต้องอาศัยทฤษฎีการแจกแจงความน่าจะเป็น การที่ได้ทราบการแจกแจงความน่าจะเป็นของตัวอย่างสุ่มจะเป็นประโยชน์อย่างยิ่งต่อการอนุมานทางสถิติ

การหาการแจกแจงความน่าจะเป็นของตัวแปรสุ่มที่ต้องการในหลายปัญหาเป็นไปได้ยากหรือมีรูปแบบที่ซับซ้อนไม่เหมาะสมหรือไม่ง่ายต่อการนำไปใช้ประโยชน์ ในตัวแปรสุ่มดังกล่าว มีจำนวนไม่น้อยที่พบว่าการแจกแจงนั้นขึ้นอยู่กับจำนวนเต็มซึ่งอาจเป็นพารามิเตอร์ของการแจกแจง หรือเป็นขนาดตัวอย่าง n กล่าวคือ การแจกแจงที่ขึ้นอยู่กับขนาดตัวอย่าง n นั้น เมื่อ n มีค่าใหญ่ การแจกแจงอาจจะลู่เข้า (converge) หรือเบนเข้าหาการแจกแจงที่ง่ายกว่า หรือเป็นที่รู้จักกันดีที่สามารถใช้เป็นการแจกแจงโดยประมาณได้ เพราะฉะนั้น การประมาณการแจกแจงความน่าจะเป็นของตัวแปรสุ่มที่ขึ้นอยู่กับจำนวนเต็มจึงเป็นสิ่งที่น่าจะทำการศึกษา โดยมีจุดสนใจเพื่อให้ทราบถึงค่าของจำนวนเต็มว่าควรมีค่าเป็นเท่าใดจึงจะทำให้การประมาณการแจกแจงความน่าจะเป็นของตัวแปรสุ่มดังกล่าวมีประสิทธิภาพในสถานการณ์ต่าง ๆ ดังนั้นการหาขนาดตัวอย่าง n ที่เหมาะสมในแต่ละสถานการณ์จึงมีความเหมาะสมกว่าการหาขนาดตัวอย่าง n จากหลักเกณฑ์โดยทั่วไป ซึ่งเป็นการกำหนดคร่าว ๆ จึงทำให้สามารถ

เกิดความคลาดเคลื่อนสูงได้ในบางกรณี เพราะขนาดตัวอย่าง n ขึ้นกับปัจจัยอื่น ๆ เช่น รูปแบบการแจกแจงดั้งเดิม เป็นต้น สำหรับงานวิจัยฉบับนี้จะทำการศึกษาการประมาณการแจกแจงที่สำคัญและถูกกล่าวถึงอย่างมาก 3 กรณีด้วยกัน กรณีแรก ได้แก่ การประมาณการแจกแจงไฮเปอร์จีออเมตริกด้วยการแจกแจงทวินาม การประมาณการแจกแจงความน่าจะเป็นกรณีดังกล่าวนี้ขึ้นกับค่าพารามิเตอร์ของการแจกแจงไฮเปอร์จีออเมตริก คือ ขนาดประชากรทั้งหมด N และขนาดตัวอย่าง n กล่าวคือ เมื่อขนาดประชากร N มีค่าใหญ่มากเมื่อเทียบกับขนาดตัวอย่าง n การแจกแจงไฮเปอร์จีออเมตริกจะลู่เข้าสู่การแจกแจงทวินาม, กรณีที่ 2 ได้แก่ การประมาณการแจกแจงทวินามด้วยการแจกแจงปัวส์ซอง, และกรณีที่ 3 ได้แก่ การประมาณการแจกแจงไฮเปอร์จีออเมตริก, การแจกแจงทวินาม, และการแจกแจงปัวส์ซอง ด้วยการแจกแจงปกติ การประมาณการแจกแจงความน่าจะเป็นดังกล่าวขึ้นกับ ขนาดตัวอย่าง n กล่าวคือ เมื่อขนาดตัวอย่าง n มีค่าใหญ่ การแจกแจงที่แท้จริงจะลู่เข้าสู่การแจกแจงที่ใช้ประมาณ

การประมาณการแจกแจงความน่าจะเป็นไฮเปอร์จีออเมตริกด้วยการแจกแจงทวินาม และการประมาณการแจกแจงความน่าจะเป็นทวินามด้วยการแจกแจงปัวส์ซอง มีจุดมุ่งหมายสำคัญ คือ เพื่อความง่ายและความรวดเร็วในการหาค่าความน่าจะเป็น เป็นการแก้ปัญหารูปแบบการแจกแจงความน่าจะเป็นของตัวแปรสุ่มที่ซับซ้อน จึงทำให้มีความเหมาะสม หรือง่ายต่อการนำไปใช้ประโยชน์ และมีความถูกต้องสูง ทั้งนี้ขึ้นอยู่กับขนาดตัวอย่าง n สำหรับกรณีการประมาณการแจกแจงที่จะกล่าวถึงในส่วนถัดไป มีจุดมุ่งหมายสำคัญเพื่อใช้ในการอนุมานโดยเน้นการแจกแจงทางด้านหางของการแจกแจง

ในการอนุมานเชิงสถิตินั้น จะทำการศึกษาและสรุปผลจากตัวอย่างเพื่ออนุมานว่าประชากรมีลักษณะเช่นไร ดังนั้นตัวสถิติที่จะนำไปใช้ในการอนุมานไม่ว่าจะเป็นการประมาณค่าหรือการทดสอบสมมุติฐาน จะเป็นฟังก์ชันของตัวแปรสุ่ม สมมติว่าขนาดตัวอย่างเท่ากับ n และพิจารณาตัวอย่างสุ่มซึ่งประกอบขึ้นด้วยตัวแปรสุ่ม n ตัว $(X_1,\ldots,X_n)$ การหาการแจกแจงความน่าจะเป็นของฟังก์ชันของตัวแปรสุ่ม n ตัวนี้ บ่อยครั้งจะกระทำได้ยากจึงมีความจำเป็นต้องหาทางประมาณการแจกแจงความน่าจะเป็นที่ถูกต้อง ทฤษฎีที่เป็นประโยชน์ต่อการหาการแจกแจงความน่าจะเป็นโดยประมาณ ก็คือ ทฤษฎีบทลิมิตต่าง ๆ ซึ่งเป็นทฤษฎีที่ว่าด้วยคุณสมบัติของตัวแปรสุ่มที่เป็นฟังก์ชันของตัวแปรสุ่ม n ตัว เมื่อ n มีขนาดใหญ่ $(n \to \infty)$ ทฤษฎีลิมิตที่มีความสำคัญมากในเรื่องของการอนุมานเชิงสถิติ คือ ทฤษฎีบทลิมิต-สู่ส่วนกลาง (Central Limit Theorem)

ทฤษฎีบทลิมิตสู่ส่วนกลาง (Central Limit Theorem) เป็นทฤษฎีที่ว่าด้วยการแจกแจงขีดจำกัด (limiting distribution) ของตัวแปรสุ่ม $\overline{X}$ หรือผลบวก $\sum X$ ซึ่งตัวอย่างสุ่มที่ใช้ไม่จำเป็นต้องมาจากการแจกแจงแบบปกติ สำหรับงานวิจัยฉบับนี้จะทำการศึกษา 3 กรณี คือ การประมาณการแจกแจงทวินาม, การแจกแจงไฮเปอร์จีออเมตริก, และการแจกแจงปัวส์ซอง ด้วยการแจกแจงปกติ และเหตุผลที่ใช้การแจกแจงปกติก็เนื่องจากจะทำให้การอนุมานเป็นไปได้ง่าย และมีความถูกต้องที่ยอมรับได้ โดยขึ้นอยู่กับขนาดตัวอย่าง n

1.2. วัตถุประสงค์ของการวิจัย

1.2.1. เพื่อหาขนาดตัวอย่างมากสุดที่ควรใช้ในการประมาณการแจกแจงไฮเปอร์จีออเมตริกด้วยการแจกแจงทวินาม

1.2.2. เพื่อหาขนาดตัวอย่างน้อยสุดที่ควรใช้ในการประมาณการแจกแจงทวินามด้วยการแจกแจงปัวส์ซอง

1.2.3. เพื่อหาขนาดตัวอย่างน้อยสุดที่ควรใช้ในการประมาณการแจกแจงทวินาม, การแจกแจงปัวส์ซอง, และการแจกแจงไฮเปอร์จีออเมตริก ด้วยการแจกแจงปกติ

1.3. ข้อตกลงเบื้องต้น

การวิจัยครั้งนี้ ได้กล่าวถึงการแจกแจง 4 การแจกแจง ดังนี้

1.3.1 การแจกแจงทวินาม (Binomial Distribution : $B(n,p)$)

ในการทดลองแบบทวินาม n ครั้ง แต่ละครั้งความน่าจะเป็นของการเกิดความสำเร็จ (Sucess) เป็น p ความน่าจะเป็นของการเกิดความไม่สำเร็จเป็น $1-p$ ถ้า X แทนจำนวนครั้งของความสำเร็จ การแจกแจงความน่าจะเป็นของ X เรียกว่าการแจกแจงทวินาม ด้วยพารามิเตอร์ n และ p มีฟังก์ชันความน่าจะเป็น คือ

$$P(X=x) = \binom{n}{x} p^x (1-p)^{n-x} \qquad ; \ x = 0,1,2,\ldots,n \text{ และ } 0<p<1$$

1.3.2 การแจกแจงปัวส์ซอง (Poisson Distribution : $Poi(\lambda)$)

กำหนดให้ X เป็นจำนวนครั้งของความสำเร็จที่เกิดขึ้นในช่วงเวลาหรือขอบเขตที่กำหนดให้ จากการทดลองแบบปัวส์ซอง X จะมีการแจกแจงปัวส์ซอง ด้วยพารามิเตอร์ λ และฟังก์ชันความน่าจะเป็น คือ

$$P(X=x) = \frac{e^{-\lambda}\lambda^{x}}{x!} \qquad ; \; x = 0,1,2,\ldots$$

1.3.3 การแจกแจงไฮเปอร์จีออเมตริก

(Hypergeomatric Distribution : $H(N,M,n)$)

ถ้า X แทนจำนวนสิ่งของพวกแรกในการสุ่มตัวอย่างขนาด n สิ่ง จากของทั้งหมด N สิ่ง ซึ่งประกอบด้วยของ 2 พวก พวกแรกมี M สิ่ง พวกที่สองมี $N-M$ สิ่ง X จะมีการแจกแจงไฮเปอร์จีออเมตริก ด้วยพารามิเตอร์ N, M และ n และฟังก์ชันความน่าจะเป็น คือ

$$P(X=x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}} \qquad ; \; x = 0,1,2,\ldots,\mathrm{Min}\,(n,M),$$

$$x \le M \quad \text{และ} \quad n-x \le N-M$$

1.3.4 การแจกแจงปกติ (Normal Distribution : $N(\mu,\sigma^2)$)

ตัวแปรสุ่ม X มีการแจกแจงความน่าจะเป็นแบบปกติ ด้วยพารามิเตอร์ μ และ σ^2 และฟังก์ชันหนาแน่น คือ

$$f(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma}\, e^{-\frac{1}{2\sigma^2}(x-\mu)^2} \qquad ; \; -\infty < x < \infty$$

1.4. <u>ขอบเขตการวิจัย</u>

1.4.1 การประมาณการแจกแจงไฮเปอร์จีออเมตริกด้วยการแจกแจงทวินาม

การแจกแจงทวินามที่ใช้ประมาณการแจกแจงไฮเปอร์จีออเมตริกจะมีพารามิเตอร์ดังนี้ คือ ขนาดตัวอย่างสุ่ม $= n$ และ ความน่าจะเป็นของการเกิดความสำเร็จ $p = \dfrac{M}{N}$

1.4.1.1 กำหนดขนาดความคลาดเคลื่อนที่มากที่สุดระหว่างการแจกแจงที่แท้จริงกับการแจกแจงที่ใช้ประมาณ (ε) = 0.05, 0.04, 0.03, 0.02, 0.01, 0.005, และ 0.001

1.4.1.2 กำหนดขนาดประชากรทั้งหมด (N) = 30, 50, 100, ... , 500

1.4.1.3 กำหนดขนาดประชากรย่อยที่สนใจ (M) = 1, 2, 3, ..., $N-1$

1.4.2 การประมาณการแจกแจงทวินามด้วยการแจกแจงปัวส์ซอง

การแจกแจงปัวส์ซองที่ใช้ประมาณการแจกแจงทวินามจะมีพารามิเตอร์ คือ $\lambda = np$

1.4.2.1 กำหนดขนาดความคลาดเคลื่อนที่มากที่สุดระหว่างการแจกแจงที่แท้จริงกับการแจกแจงที่ใช้ประมาณ (ε) = 0.05, 0.04, 0.03, 0.02, 0.01, 0.005, และ 0.001

1.4.2.2 กำหนดค่า λ เป็นค่าคงที่ แล้วทำการหาขนาดตัวอย่างน้อยสุดที่ควรใช้ในการประมาณการแจกแจงความน่าจะเป็นในแต่ละกรณี

1.4.3 การประมาณการแจกแจงด้วยการแจกแจงปกติ

1.4.3.1 การประมาณการแจกแจงทวินามด้วยการแจกแจงปกติ

การแจกแจงปกติที่ใช้ประมาณการแจกแจงทวินามจะมีพารามิเตอร์ดังนี้ คือ มีค่าเฉลี่ย $= np$ และ ค่าความแปรปรวน $= npq$

1.4.3.1.1 เปลี่ยนค่าความน่าจะเป็นของการเกิดความสำเร็จ p เป็นค่าต่าง ๆ ตั้งแต่ 0.01-0.50 เพื่อหาขนาดตัวอย่าง n น้อยสุดที่ควรใช้ในการประมาณการแจกแจงในแต่ละค่า p

1.4.3.1.2 ระดับนัยสำคัญของการทดสอบ (α) ที่ศึกษา คือ 0.10, 0.05, และ 0.01

1.4.3.1.3 ระดับนัยสำคัญของการทดสอบทวินาม (Binomial Test) คือ 0.05

1.4.3.2 การประมาณการแจกแจงปัวส์ซองด้วยการแจกแจงปกติ

การแจกแจงปกติที่ใช้ประมาณการแจกแจงปัวส์ซองจะมีพารามิเตอร์ดังนี้ คือ มีค่าเฉลี่ยและความแปรปรวน $=\lambda$

1.4.3.2.1 หาค่า λ ที่ทำให้การแจกแจงปกติประมาณการแจกแจงปัวส์ซองได้ดี

1.4.3.2.2 ระดับนัยสำคัญของการทดสอบ (α) ที่ศึกษา คือ 0.10, 0.05, และ 0.01

1.4.3.2.3 ระดับนัยสำคัญของการทดสอบทวินาม (Binomial Test) คือ 0.05

1.4.3.3 การประมาณการแจกแจงไฮเปอร์จีออเมตริกด้วยการแจกแจงปกติ

การแจกแจงปกติที่ใช้ประมาณการแจกแจงไฮเปอร์จีออเมตริกจะมีพารามิเตอร์ดังนี้ คือ มีค่าเฉลี่ย $= nM/N$ และความแปรปรวน $= \dfrac{n.(N-n)}{(N-1)}\left(\dfrac{M}{N}\right)\left(1-\dfrac{M}{N}\right)$

1.4.3.3.1 กำหนดขนาดประชากร (N) = 30, 50, 100, ... , 500

1.4.3.3.2 ระดับนัยสำคัญของการทดสอบ (α) ที่ศึกษา คือ 0.10, 0.05, และ 0.01

1.4.3.3.3 ระดับนัยสำคัญของการทดสอบทวินาม (Binomial Test) คือ 0.05

1.4.3.4 ขั้นตอนการประมาณค่าระดับนัยสำคัญ หรือค่าความน่าจะเป็นของความผิดพลาดประเภทที่ 1 จะจำลองให้มีสถานการณ์ตามที่กำหนดข้างต้นโดยใช้เทคนิคการจำลองแบบมอนติคาร์โล (Monte Carlo Simulation Technique) ทำซ้ำ 10,000 รอบ (n^{*} = 10,000)

1.4.4 ทดสอบยืนยันผลการวิจัยด้วย การทดสอบเทียบความกลมกลืนกัน โดยใช้ไคสแควร์ (Chi-Square Goodness-of-Fit Test) ณ ระดับนัยสำคัญของการทดสอบ คือ 0.05

1.5 <u>เกณฑ์การตัดสินใจ</u>

1.5.1. การประมาณการแจกแจงใดการแจกแจงหนึ่ง ด้วยการแจกแจงอื่น ๆ นั้น ในการวิจัยนี้มีวัตถุประสงค์เพื่อหาขนาดตัวอย่าง n ที่จะทำให้การประมาณการแจกแจงดังกล่าวมีประสิทธิภาพเป็นไปตามความต้องการ จึงต้องมีการเปรียบเทียบจากพื้นฐานเดียวกัน ดังนั้นในการวิจัยครั้งนี้จึงใช้เกณฑ์ในการพิจารณาถึง 2 ลักษณะตามประเภทของการแจกแจงเพื่อความเหมาะสมตามแต่กรณี ดังนี้

<u>กรณีที่ 1</u>. การประมาณการแจกแจงแบบไม่ต่อเนื่องด้วยการแจกแจงแบบไม่ต่อเนื่อง ได้แก่

- การประมาณการแจกแจงไฮเปอร์จีออเมตริกด้วยการแจกแจงทวินาม
- การประมาณการแจกแจงทวินามด้วยการแจกแจงปัวส์ซอง

ใช้หลักเกณฑ์ในการพิจารณาค่าความคลาดเคลื่อนระหว่างความน่าจะเป็นของการแจกแจงจริงกับความน่าจะเป็นของการแจกแจงที่ใช้ประมาณการ ดังแสดงได้ดังสูตร

$$\underset{i}{\text{Max}} \left| P_i - \hat{P}_i \right| \quad \leq \quad \varepsilon \qquad ; i = 0,1,2,..,n$$

โดย P_i คือ ความน่าจะเป็นของการแจกแจงจริง ที่จะเกิดลักษณะที่ i

$\hat{P}_i$ คือ ความน่าจะเป็นของการแจกแจงที่ใช้ประมาณ ที่จะเกิดลักษณะที่ i

ε คือ ค่าความคลาดเคลื่อนที่กำหนดขึ้น

ในกรณีนี้เราจะคัดเลือกขนาดตัวอย่าง n ที่ทำให้ ความคลาดเคลื่อนระหว่างความน่าจะเป็นของการแจกแจงจริงกับความน่าจะเป็นของการแจกแจงที่ใช้ประมาณทุกค่า มีค่าไม่มากกว่าค่าความคลาดเคลื่อนที่กำหนดก็จะถือว่าค่า n นั้นเหมาะสม

<u>กรณีที่ 2</u>. การประมาณการแจกแจงแบบไม่ต่อเนื่องด้วยการแจกแจงแบบต่อเนื่อง ได้แก่

- การประมาณการแจกแจงทวินามด้วยการแจกแจงปกติ
- การประมาณการแจกแจงปัวส์ซองด้วยการแจกแจงปกติ
- การประมาณการแจกแจงไฮเปอร์จีออเมตริกด้วยการแจกแจงปกติ

ใช้หลักเกณฑ์การพิจารณา แบ่งเป็น 2 ขั้นตอน ดังนี้

1. ประมาณค่าระดับนัยสำคัญของการทดสอบ หรือค่าความน่าจะเป็นของความผิดพลาดประเภทที่ 1 จากการทดลอง โดยทำการจำลองข้อมูลแล้วใช้หลักการตามสมการ

$$P\left(\left|\frac{S_n - E[S_n]}{\sqrt{V[S_n]}}\right| \geq Z_{\alpha/2}\right) \approx \alpha$$

เมื่อได้ค่าประมาณของระดับนัยสำคัญ หรือค่าความน่าจะเป็นของความผิดพลาดประเภทที่ 1 จากการทดลองแล้ว จะทำการพิจารณาความแตกต่างระหว่างค่าระดับนัยสำคัญที่แท้จริงกับค่าระดับนัยสำคัญที่ประมาณการได้ ซึ่งถ้าขนาดตัวอย่างมีความเหมาะสมค่าที่ประมาณการได้ควรไม่แตกต่างกันอย่างมีนัยสำคัญ เกณฑ์ที่ใช้ในการทดสอบความแตกต่างนี้มาจากการทดสอบทวินาม (Binomial Test)

2. การทดสอบทวินาม (Binomial Test)

มีรูปแบบเป็นดังนี้

$$\alpha \pm Z_{\alpha^*/2}\sqrt{\frac{\alpha(1-\alpha)}{n^*}}$$

โดยที่ค่า α มาจากระดับนัยสำคัญที่แท้จริง เมื่อทำการหาค่าประมาณการของระดับนัยสำคัญ หรือค่าความน่าจะเป็นของความผิดพลาดประเภทที่ 1 จากการทดลองได้แล้ว จะนำมาเปรียบเทียบกับช่วงดังกล่าวถ้าอยู่ภายในช่วง กล่าวได้ว่าระดับนัยสำคัญจริงและประมาณการไม่มีความแตกต่างกัน แสดงว่าขนาดตัวอย่างมีความเหมาะสม

1.5.2 ใช้การทดสอบเทียบความกลมกลืนกัน (Goodness-of-Fit Test)

เป็นการทดสอบเกี่ยวกับกลุ่มประชากรที่สนใจ ว่าจะมีลักษณะการแจกแจงตามทฤษฎีใด ๆ เป็นดังที่คาดไว้หรือตั้งสมมติฐานไว้หรือไม่ การทดสอบขึ้นกับการเปรียบเทียบจำนวนความถี่ที่เกิดจากการทดลองหรือที่สังเกตได้ กับจำนวนความถี่ที่คาดหมายซึ่งหาได้จากการตั้งสมมติฐานไว้ โดยพิจารณาว่าค่าความถี่นั้นเท่ากันหรือแตกต่างกันหรือไม่ ถ้าผลต่างของความถี่จากการทดลองแตกต่างกับความถี่จากสมมติฐานอย่างไม่มีนัยสำคัญ ก็จะถือว่าค่าจากการทดลองนั้นมีลักษณะการแจกแจงตามสมมติฐาน แต่ถ้าความแตกต่างที่ได้มีค่ามากจนถือว่าแตกต่างอย่างมีนัยสำคัญ จะสรุปว่าการแจกแจงค่าที่ได้จากการทดลองจะไม่เป็นไปตามสมมติฐานที่คาดไว้

ค่าสถิติ χ_c^2 ที่ใช้ทดสอบจากสูตร

$$\chi_c^2 = \sum_{i=1}^{k} \frac{\left(O_i - E_i\right)^2}{E_i}$$

โดยที่ O_i คือ ความถี่ที่ได้จากการทดลอง

E_i คือ ความถี่ที่คาดหมาย คำนวณตามสมมติฐาน

นำค่า χ_c^2 ที่คำนวณได้เปรียบเทียบกับ ค่าวิกฤต $\chi^2_{\alpha,k-m-1}$ ซึ่งเป็นเปอร์เซ็นไทล์ที่ $100(1-\alpha)$ ของตัวแปรสุ่มไคสแควร์ องศาอิสระ $k-m-1$

1.5.3. ค่าพารามิเตอร์ของการแจกแจงที่มีผลต่อการประมาณการแจกแจงความน่าจะเป็นกรณีต่าง ๆ สำหรับงานวิจัยครั้งนี้ ได้แก่

- การประมาณการแจกแจงไฮเปอร์จีออเมตริกด้วยการแจกแจงทวินาม การประมาณดังกล่าวขึ้นกับพารามิเตอร์ของการแจกแจงไฮเปอร์จีออเมตริก คือ ขนาดประชากรทั้งหมด N และขนาดตัวอย่าง n ในการศึกษาเราจะกำหนดค่าพารามิเตอร์ N เป็นค่าคงที่หนึ่ง ๆ และทำการคัดเลือกค่าพารามิเตอร์ n ที่จะทำให้การประมาณมีประสิทธิภาพ เป็นไปตามเงื่อนไข

- การประมาณการแจกแจงทวินามด้วยการแจกแจงปัวส์ซอง, การประมาณการแจกแจงทวินาม, การแจกแจงปัวส์ซอง, และการแจกแจงไฮเปอร์จีออเมตริก ด้วยการแจกแจงปกติ การประมาณดังกล่าวขึ้นกับพารามิเตอร์ของการแจกแจง คือ ขนาดตัวอย่าง n

1.5.4 การประมาณการแจกแจงไฮเปอร์จีออเมตริกด้วยการแจกแจงทวินามค่าพารามิเตอร์ n ที่ได้จะเป็นค่ามากสุดที่ควรใช้ในการประมาณการ ทั้งนี้จากคำกล่าวที่ว่า ถ้าขนาดตัวอย่าง n มีค่าน้อยเมื่อเทียบกับขนาดประชากร N การแจกแจงไฮเปอร์จีออเมตริกจะเข้าใกล้การแจกแจงทวินาม ในงานวิจัยฉบับนี้จะทำการกำหนดขนาดพารามิเตอร์ N เป็นค่าคงที่หนึ่ง ๆ แล้วคัดเลือก ขนาดพารามิเตอร์ n โดยเริ่มศึกษาจากพารามิเตอร์ n ที่มาก และทำการลดค่าพารามิเตอร์ดังกล่าวลงจนกระทั่งการประมาณการแจกแจงมีประสิทธิภาพเป็นไปตามเงื่อนไขที่ต้องการ

สำหรับการประมาณการแจกแจงอื่น ๆ นั้น ได้แก่ การประมาณการแจกแจงทวินามด้วยการแจกแจงปัวส์ซอง และ การประมาณการแจกแจงไฮเปอร์จีออเมตริก, การแจกแจงทวินาม, และการแจกแจงปัวส์ซอง ด้วยการแจกแจงปกติ ค่าพารามิเตอร์ n ที่ได้จะเป็นค่าน้อยสุดที่ควรใช้ในการประมาณการแจกแจง ทั้งนี้เพราะการประมาณการแจกแจงดังกล่าวจะประมาณได้ดีขึ้นเมื่อขนาดพารามิเตอร์ n มีค่ามากขึ้น สำหรับงานวิจัยฉบับนี้จะเริ่มศึกษาจากพารามิเตอร์ n ที่น้อย และทำการเพิ่มขนาด n ขึ้นจนกระทั่งการประมาณการแจกแจงมีประสิทธิภาพเป็นไปตามเงื่อนไขที่ต้องการ

1.5.5 การประมาณการแจกแจงไฮเปอร์จีออเมตริก, การแจกแจงทวินาม,และการแจกแจงปัวส์ซองด้วยการแจกแจงปกติ

การประมาณการแจกแจงดังกล่าวทั้ง 3 กรณี สามารถอธิบายได้โดยทฤษฎีบทลิมิตสู่ส่วนกลาง (Central Limit Theorem) ดังนี้

กำหนดให้ $X_1, X_2, \ldots, X_n$ เป็นตัวแปรสุ่มที่มีการแจกแจงเดียวกัน และเป็นอิสระต่อกัน โดยมีค่าเฉลี่ย μ และความแปรปรวน $\sigma^2 < \infty$ ในงานวิจัยนี้จะสนใจ $S_n = \sum_{i=1}^{n} X_i$ จะได้ว่า $Z_n = \dfrac{S_n - E[S_n]}{\sqrt{V[S_n]}}$ ลู่เข้าในเชิงการแจกแจงสู่ตัวแปรสุ่ม Z ซึ่งเป็นตัวแปรสุ่มที่มีค่า $\mu = 0$ และ $\sigma^2 = 1$ กล่าวคือ Z_n จะลู่เข้าสู่การแจกแจงปกติมาตรฐาน

1.5. ประโยชน์ของการวิจัย

1.5.1 ช่วยแก้ปัญหารูปแบบการแจกแจงความน่าจะเป็นของตัวแปรสุ่มที่มีความซับซ้อนให้สามารถหาความน่าจะเป็นได้ง่ายและมีความรวดเร็วมากขึ้น

1.5.2 การประมาณการแจกแจงต่าง ๆ ด้วยการแจกแจงปกตินั้น จะทำให้การอนุมานเป็นไปได้ง่ายขึ้น

1.5.3 สามารถนำผลของการวิจัยมาสรุปเป็นตารางที่แสดงค่าขนาดตัวอย่างที่ควรใช้ในการประมาณการแจกแจง เพื่ออำนวยประโยชน์ให้แก่ผู้ที่ต้องการใช้งาน