การเปรียบเทียบวิธีการประมาณแบบช่วงสำหรับพารามิเตอร์ของการแจกแจงพหุนาม

นายบัณฑิต ฉัตรตะวัน

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาสถิติศาสตรมหาบัณฑิต
สาขาวิชาสถิติ ภาควิชาสถิติ
คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2551

A COMPARISON ON METHODS OF INTERVAL ESTIMATION FOR PARAMETER OF MULTINOMIAL DISTRIBUTION

Mr. Bandit Chattawan

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Statistics Program in Statistics
Department of Statistics
Faculty of Commerce and Accountancy
Chulalongkorn University
Academic Year 2008

หัวข้อวิทยานิพนธ์	การเปรียบเทียบวิธีการประมาณแบบช่วงสำหรับพารามิเตอร์ของการแจกแจงพหุนาม
โดย	นายบัณฑิต ฉัตรตะวัน
สาขาวิชา	สถิติ
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก	รองศาสตราจารย์ ดร.ธีระพร วีระถาวร

คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญามหาบัณฑิต

....คณบดีคณะพาณิชยศาสตร์และการบัญชี
(รองศาสตราจารย์ ดร.อรรณพ ต้นละมัย)

คณะกรรมการสอบวิทยานิพนธ์

....ประธานกรรมการ
(รองศาสตราจารย์ ดร.สุพล ดุรงค์วัฒนา)

....อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก
(รองศาสตราจารย์ ดร.ธีระพร วีระถาวร)

....กรรมการ
(อาจารย์ ดร. อรุณี กำลัง)

....กรรมการภายนอกมหาวิทยาลัย
(รองศาสตราจารย์ ดร. มนต์ทิพย์ เทียนสุวรรณ)

บัณฑิต ฉัตรตะวัน : การเปรียบเทียบวิธีการประมาณแบบช่วงสำหรับพารามิเตอร์ของการแจกแจงพหุนาม (A COMPARISON ON METHODS OF INTERVAL ESTIMATION FOR PARAMETER OF MULTINOMIAL DISTRIBUTION) อ. ที่ปรึกษาวิทยานิพนธ์หลัก : รศ.ดร. ธีระพร วีระถาวร, 90 หน้า

งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาวิธีการประมาณค่าแบบช่วงสำหรับพารามิเตอร์ของการแจกแจงพหุนาม 3 วิธี คือ วิธีการประมาณแบบปกติ (Normal method) วิธีการประมาณแบบคิวเซนเบอร์รี่และเฮิร์ส (Quesenbery and Hurst's method) และวิธีการประมาณแบบเอฟ (F-method) ณ ระดับความเชื่อมั่น 90% , 95% และ 99% และกำหนดขนาดตัวอย่าง 50, 100, 200, 500, 1000 ผลการทดลองแบ่งออกเป็น 4, 6, 8 และ 10 กลุ่ม ซึ่งในแต่ละกลุ่มแบ่งลักษณะค่าสัดส่วนประชากร p_i ออกเป็น 3 ประเภท คือ เท่ากัน เบ้ และสมมาตร ข้อมูลที่ใช้วิจัยมาจากการจำลองด้วยโปรแกรม MATLAB และใช้เทคนิคมอนติคาร์โล โดยทำซ้ำ 2,000 ครั้ง ผลการวิจัยสรุปได้ดังนี้

1. วิธีการประมาณแบบช่วงของวิธีปกติ จะให้ค่าสัมประสิทธิ์ช่วงความเชื่อมั่นและระยะห่างของบริเวณความเชื่อมั่นเฉลี่ยต่ำสุด และสามารถคลุมค่าสัดส่วนประชากรดังนี้

กรณีที่ขนาดตัวอย่างมีค่าตั้งแต่ 500 ขึ้นไป วิธีการประมาณนี้จะให้ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นไม่ต่ำกว่าระดับที่กำหนด และสามารถคลุมค่าสัดส่วนประชากรจำนวน 4 กลุ่ม จากค่าสัดส่วนประชากรจำนวน 4, 6, 8 และ 10 กลุ่ม

กรณีที่ขนาดตัวอย่างมีค่าตั้งแต่ 1000 ขึ้นไป วิธีการประมาณนี้จะให้ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นไม่ต่ำกว่าระดับที่กำหนด และสามารถคลุมค่าสัดส่วนประชากรจำนวน 6 กลุ่ม จากค่าสัดส่วนประชากรจำนวน 4, 6, 8 และ 10 กลุ่ม

2. วิธีการประมาณแบบคิวเซนเบอร์รี่และเฮิร์ส จะให้ค่าระดับความเชื่อมั่นไม่ต่ำกว่าระดับความเชื่อมั่นที่กำหนดทุกกรณี ซึ่งวิธีนี้จะให้ค่าสัมประสิทธิ์ช่วงความเชื่อมั่นและระยะห่างบริเวณความเชื่อมั่นเฉลี่ยสูงสุด

3. วิธีการประมาณแบบเอฟ จะให้ค่าระดับความเชื่อมั่นไม่ต่ำกว่าระดับความเชื่อมั่นที่กำหนดทุกกรณี ซึ่งวิธีนี้จะให้ค่าสัมประสิทธิ์ช่วงความเชื่อมั่นสูงกว่าวิธีการประมาณแบบปกติและให้ค่าระยะห่างบริเวณช่วงความเชื่อมั่นเฉลี่ยใกล้เคียงกับค่าประมาณแบบช่วงของวิธีปกติ

ภาควิชา.......สถิติ.......	ลายมือชื่อนิสิต....บัณฑิต ฉัตรตะวัน....
สาขาวิชา.......สถิติ.......	ลายมือชื่อ อ. ที่ปรึกษาวิทยานิพนธ์หลัก....................
ปีการศึกษา.......2551.......	

4982191626: MAJOR STATISTICS

KEYWORDS: PROPORTION/ MULTINOMIAL/ INTERVAL ESTIMATION

BANDIT CHATTAWAN: A COMPARISON ON METHODS OF INTERVAL ESTIMATION FOR PARAMETER OF MULTINOMIAL DISTRIBUTION , Ph.D., 90 pp.

The objective of this research was to compare the interval estimation for the parameters of multinomial distribution. The estimation uses method of Normal, Quesenbery and Hurst's and F with sample of size 50, 100, 200, 500 and 1000 from multinomial distributions of groups 4, 6, 8 and 10. For each group we divided into three cases; equality, bias, and symmetry. All of which were considered at 90%, 95% and 99% confidence levels.The data in this research were obtained from MATLAB program by using Monte Carlo technique and this process was repeated 2000 times. The conclusion of this study are as follow:

1. The Normal method performs the lowest average confidence levels and the length in term of average confidence region length. The average confidence levels are not lower than the given confidence levels when the sample size is bigger than 500 for 4 groups and bigger than 1000 for 6 groups from 4, 6, 8 and 10 groups of proportions.

2. The Quesenberry an Hurst method performs the highest average confidence levels and the length in term of average confidence region length. The average confidence levels are not lower than the given confidence levels.

3. The F method performs the average confidence levels higher than the normal method but its length in term of average confidence region length is near the normal method's. The average confidence levels are not lower than the given confidence levels.

Department: ....Statistics.........

Field of Study: ....Statistics.........

Academic Year: ......2008.............

Student's Signature. Bandit Chattawan.......

Advisor's Signature. Theeraporn Veerahan

กิตติกรรมประกาศ

ผู้วิจัยขอกราบขอบพระคุณรองศาสตราจารย์ ดร.ธีระพร วีระถาวร อาจารย์ที่ปรึกษาวิทยานิพนธ์เป็นอย่างสูงที่กรุณาสละเวลาในการให้ความรู้ คำแนะนำและคำปรึกษา ตลอดจนแก้ไขข้อบกพร่องต่างๆ ด้วยความเมตตาและเอาใจใส่อย่างดียิ่งจนทำให้วิทยานิพนธ์ ฉบับนี้สำเร็จลุล่วงได้ด้วยดี

ผู้วิจัยขอกราบขอบพระคุณรองศาสตราจารย์ ดร.สุพล ดุรงค์วัฒนา ประธานกรรมการสอบวิทยานิพนธ์ อาจารย์ ดร. อรุณี กำลัง กรรมการสอบวิทยานิพนธ์ และรองศาสตราจารย์ ดร. มนต์ทิพย์ เทียนสุวรรณ กรรมการสอบวิทยานิพนธ์จากมหาวิทยาลัยมหิดล ที่กรุณาให้คำแนะนำ ตรวจสอบ และแก้ไขวิทยานิพนธ์ฉบับนี้ให้สมบูรณ์ยิ่งขึ้น ตลอดจนครูอาจารย์ทุกท่านที่ได้ประสิทธิ์ประสาทวิชาความรู้ให้แก่ผู้วิจัยตั้งแต่การศึกษาขั้นต้นจนถึงปัจจุบัน

สุดท้ายนี้ ผู้วิจัยขอกราบขอบพระคุณ คุณพ่อ คุณแม่ที่ส่งเสริมและสนับสนุนด้านการเรียนของผู้วิจัยและเป็นกำลังใจให้เสมอมาจนสำเร็จการศึกษา และขอขอบคุณเพื่อนๆ พี่ๆ น้องๆ ทุกคนที่เป็นกำลังใจและให้ความช่วยเหลือเสมอมา

สารบัญ

หน้า

สารบัญตาราง

ตารางที่ หน้า

ตารางที่ หน้า

สารบัญภาพ

บทที่ 1

บทนำ

1.1 ที่มาและความสำคัญของปัญหา

การแจกแจงทวินาม (binomial distribution) เป็นการแจกแจงที่อธิบายถึงเหตุการณ์ของการทดลองต่างๆ โดยมีผลการทดลอง 2 กลุ่ม นั่นคือ สิ่งที่เราสนใจ (success) โดยมีความน่าจะเป็นที่เกิดเหตุการณ์ที่สนใจเท่ากับ p และสิ่งที่เราไม่สนใจ (failure) โดยมีความน่าจะเป็นที่จะเกิดเป็น $1-p$ ในการทดลองทั้งหมด n ครั้ง ที่เป็นอิสระซึ่งกันและกันซึ่งความน่าจะเป็นที่จะเกิดเหตุการณ์ที่สนใจเท่ากับ p คงที่ตลอดการทดลอง

การแจกแจงพหุนาม (multinomial distribution) เป็นการแจกแจงที่พัฒนามาจากการแจกแจงทวินามโดยที่ผลการทดลองของเหตุการณ์ที่จะเกิดขึ้น k กลุ่ม

ตัวแปรสุ่ม $\underset{\sim}{X} \sim (X_1, X_2, ..., X_k)'$ มีการแจกแจงแบบพหุนาม ที่มีฟังก์ชันความน่าจะเป็นร่วมอยู่ในรูปของ

$$p(x_1, x_2, ..., x_k) = \frac{n!}{x_1! x_2! ... x_k!} p_1^{x_1} p_2^{x_2} ... p_k^{x_k} \;;\; x_i = 0,1,...,n \text{ และ }$$

$$\sum_{i=1}^{k} x_i = n$$

$$\therefore \quad E(X_i) = np_i$$

$$Var(X_i) = np_i(1-p_i)$$

และ $\quad Cov(X_{ij}) = -p_i p_j$

ซึ่งการแจกแจงนี้จะใช้ในการอธิบายการทดลองที่มีผลลัพธ์มากกว่า 2 กลุ่มในการทดลอง โดยที่การทดลองแต่ละครั้งจะมีความน่าจะเป็นคงที่ p_i ซึ่งจะเกิดผลลัพธ์ $i, i = 1,...,k$

การประมาณค่าพารามิเตอร์ p_i สามารถประมาณได้ทั้งแบบจุด (point estimation) และการประมาณแบบช่วง (interval estimation) ซึ่งงานวิจัยนี้จะทำการประมาณพารามิเตอร์ p_i โดยใช้การประมาณแบบช่วง

การประมาณแบบช่วงสามารถประมาณได้ด้วยวิธีปกติ (normal method) ซึ่งจากทฤษฎีบทค่าจำกัดสู่ส่วนกลาง (Central Limit Theorem) สรุปได้ว่าถ้าจำนวนครั้งของการทดลอง n มีมากพอค่าสัดส่วนตัวอย่าง $(\hat{p}_i)$ จะมีแนวโน้มลู่เข้าส่วนกลางด้วยการแจกแจงแบบปกติที่มีค่าเฉลี่ยเป็น p_i และความแปรปรวนเป็น $p_i(1-p_i)/n$ และช่วงความเชื่อมั่นขนาด $1-a$ อยู่ในรูปของ

$$\left(\hat{p}_i - z_{1-a/(2k)} \sqrt{\frac{\hat{p}_i(1-\hat{p}_i)}{n}}, \hat{p}_i + z_{1-a/(2k)} \sqrt{\frac{\hat{p}_i(1-\hat{p}_i)}{n}} \right)$$

เมื่อ $Z_{1-a/(2k)}$ แทนตัวแปรสุ่มปกติมาตรฐาน ที่ให้ค่าฟังก์ชันความน่าจะเป็นสะสมเท่ากับ $1-a/(2k)$ ณ ระดับความเชื่อมั่น $100(1-a)\%$ โดยประมาณ

ในกรณีที่ค่าสัดส่วนตัวอย่างมีแนวโน้มลู่เข้าสู่การแจกแจงปกติ ในปี ค.ศ. 1927 วิลสัน (Edwin B. Wilson) ได้เสนอวิธีการประมาณแบบช่วงในรูปของ

$$\left(\hat{p}+\frac{z_{1-a/2}^{2}}{2n}\pm z_{1-a/2}\sqrt{\left\{\hat{p}(1-\hat{p})+z_{1-a/2}^{2}/4n\right\}/n}\right)$$

ซึ่งต่อมาการประมาณดังกล่าวได้ถูกพัฒนาโดย คิวเซนเบอร์รี่และเฮิร์ส (Quesenberry and Hurst)ในปี ค.ศ. 1964 และ จอห์นสันและโคลท์ซ (Johnson and Koltz)ในปี ค.ศ. 1969 โดยอยู่ในรูปแบบของ

$$\left(\frac{c_{k-1,1-a}^{2}+2n_i\pm\sqrt{c_{k-1,1-a}^{2}\left(c_{k-1,1-a}^{2}+4\frac{n_i}{n}(n-n_i)\right)}}{2\left(n+c_{k-1,1-a}^{2}\right)}\right)$$

กำหนดให้ n_i แทน จำนวนเหตุการณ์ที่สนใจ

n แทน จำนวนครั้งของการทดลอง

เมื่อ n มีขนาดเล็ก การสร้างช่วงความเชื่อมั่น ณ ระดับ $100(1-a)\%$ ของ p จะต้องพิจารณาถึงค่าขอบเขตบน (P_L) และขอบเขตล่าง (P_U) เมื่อ $P(Y\le y\mid p=P_L)=\dfrac{a}{2}$ และ $P(Y\ge y\mid p=P_U)=\dfrac{a}{2}$ ตามลำดับ ในปี ค.ศ. 1969 จอห์นสันและโคลท์ซ (Johnson and Koltz) ได้เสนอการประมาณค่าโดยใช้วิธีของ คอปเปอร์เพียร์สัน (Clopper-Pearson) ซึ่งต่อมา วิธีดังกล่าวได้รับการพัฒนาโดย ลีมิสและไทรวิดี้ (Leemis and Trivedi) ในปี ค.ศ. 1996 ซึ่งจะคำนวณค่า P_L และ P_U ในเทอมของการแจกแจงของ F ดังนี้

$$\left(\frac{1}{1+\dfrac{n-n_i+1}{n_iF_{(2n,2(n-n_i+1)),1-a/2k}}}\ ,\ \frac{1}{1+\dfrac{n-n_i}{(n_i+1)F_{(2(n_i+1),2(n-n_i)),a/2k}}}\right)$$

1.2 วัตถุประสงค์ของการวิจัย

การวิจัยมีจุดประสงค์ดังนี้

1.2.1 เพื่อศึกษาวิธีการประมาณค่าแบบช่วงสำหรับพารามิเตอร์ของการแจกแจงแบบพหุนามซึ่งใช้วิธีการประมาณต่างๆ ดังนี้

1.2.1.1 วิธีการประมาณแบบปกติ (Normal method)

1.2.1.2 วิธีการประมาณแบบคิวเซนเบอร์รี่ และ เฮิร์ส (Quesenberry and Hurst's method)

1.2.1.3 วิธีการประมาณแบบเอฟ (F method)

1.2.2 ศึกษาและเปรียบเทียบวิธีการประมาณทั้งสาม ด้วยค่าระดับความเชื่อมั่น และความน่าจะเป็นของช่วงที่คำนวณได้จากแต่ละวิธี ณ ระดับความเชื่อมั่น 3 ระดับ คือ 90% , 95% และ 99%

1.2.3 เพื่อหาข้อสรุปในการเลือกใช้วิธีประมาณแบบช่วงที่เหมาะสมในแต่ละสถานการณ์ของขนาดตัวอย่าง n และจำนวนกลุ่ม k ณ ระดับความเชื่อมั่น $100(1-a)\%$

1.3 สมมติฐานการวิจัย

การประมาณค่าสัดส่วนประชากรโดยใช้วิธีการประมาณแบบเอฟ น่าจะให้ค่าสัมประสิทธิ์ของบริเวณความเชื่อมั่นมากที่สุด และจะให้ค่าระยะห่างของค่าบริเวณช่วงความเชื่อมั่นต่ำสุด เนื่องจากวิธีการดังกล่าวอาศัยเพียงความสัมพันธ์ของการแจกแจงบีตาและเอฟ โดยแตกต่างจากวิธีอื่นซึ่งจะต้องทำการประมาณค่าพารามิเตอร์กรณีที่ขนาดตัวอย่างมีขนาดมากและต้องใช้ทฤษฎีค่าจำกัดสู่ส่วนกลาง

1.4 ข้อตกลงเบื้องต้น

ขีดจำกัดความเชื่อมั่นของค่าประมาณแบบช่วงทั้ง 3 วิธีมีดังนี้

<u>1.4.1 วิธีการประมาณแบบปกติ</u>

ขีดจำกัดความเชื่อมั่นล่าง (p_L) คือ $\hat{p}_i - z_{1-a/(2k)}\sqrt{\frac{\hat{p}_i(1-\hat{p}_i)}{n}}$

และขีดจำกัดความเชื่อมั่นบน (p_U) คือ $\hat{p}_i + z_{1-a/(2k)}\sqrt{\frac{\hat{p}_i(1-\hat{p}_i)}{n}}$

1.4.2 วิธีการประมาณแบบคิวเซนเบอร์รี่ และ เฮิร์ส

ขีดจำกัดความเชื่อมั่นล่าง (p_L) คือ

$$\left(\frac{c^{2}_{k-1,1-a} + 2n_i - \sqrt{c^{2}_{k-1,1-a}\left(c^{2}_{k-1,1-a} + 4\frac{n_i}{n}(n-n_i)\right)}}{2\left(n + c^{2}_{k-1,1-a}\right)}\right)$$

และขีดจำกัดความเชื่อมั่นบน (p_U) คือ

$$\left(\frac{c^{2}_{k-1,1-a} + 2n_i + \sqrt{c^{2}_{k-1,1-a}\left(c^{2}_{k-1,1-a} + 4\frac{n_i}{n}(n-n_i)\right)}}{2\left(n + c^{2}_{k-1,1-a}\right)}\right)$$

1.4.3 วิธีการประมาณแบบเอฟ

ขีดจำกัดความเชื่อมั่นล่าง (p_L) คือ $\dfrac{n_i}{n_i + (n-n_i+1)F_{(2(n-n_i+1),2n_i),1-\frac{a}{2k}}}$

และขีดจำกัดความเชื่อมั่นบน (p_U) คือ $\dfrac{(n_i+1)F_{(2(n_i+1),2(n-n_i)),1-\frac{a}{2k}}}{n-n_i+(n_i+1)F_{(2(n_i+1),2(n-n_i)),1-\frac{a}{2k}}}$

สัญลักษณ์ที่ใช้กำหนดข้างต้นใช้แทนความหมายดังนี้

$\hat{p}_i$ เป็นค่าสัดส่วนตัวอย่างกลุ่มที่ i มีค่าเท่ากับ $\dfrac{n_i}{n}$

n_i เป็นจำนวนครั้งของผลสำเร็จ (ในตัวอย่าง)

n เป็นจำนวนครั้งของการทดลอง

k เป็นจำนวนกลุ่มของการแจกแจงพหุนาม

$Z_{1-a/(2k)}$ เป็นค่าของตัวแปรสุ่มปกติมาตรฐาน ที่ให้ค่าฟังก์ชันความน่าจะเป็นสะสมเท่ากับ $1-a/(2k)$

$c^{2}_{k-1,1-a}$ เป็นค่าของตัวแปรสุ่มไค-สแควร์ ที่ให้ค่าฟังก์ชันความน่าจะเป็นสะสมเท่ากับ $1-a$ และมีองศาความเป็นอิสระเท่ากับ $k-1$

และ $F_{(x,y),1-\frac{a}{2}}$ เป็นค่าของตัวแปรสุ่มเอฟ ที่ให้ค่าฟังก์ชันความน่าจะเป็นสะสมเท่ากับ $1-a/2$ และมีองศาความเป็นอิสระ x และ y

1.5 **ขอบเขตของการวิจัย**

1.5.1 กำหนดให้ขนาดตัวอย่างที่ใช้ n มีค่าตั้งแต่ 50, 100, 200, 500 และ 1000

1.5.2. กำหนดค่า p_i ในแต่ละกลุ่มโดยแบ่งเป็น 3 กรณีย่อย

1.5.2.1 กรณีที่ค่า p_i มีลักษณะเท่ากัน

1.5.2.2 กรณีที่ค่า p_i มีลักษณะสมมาตรแบ่งเป็น 2 กรณี

ก) กรณี $i \le \frac{k}{2}$ จะได้ว่า

$$p_i = \frac{i}{\left(k/2\right)\left(k/2+1\right)} \; ; \; i = 1,\ldots,k \text{ และ } k = 4,6,8,10$$

ข) กรณี $i > \frac{k}{2}$ จะได้ว่า

$$p_i = \frac{k-i+1}{\left(k/2\right)\left(k/2+1\right)} \; ; \; i = 1,\ldots,k \text{ และ } k = 4,6,8,10$$

1.5.2.3 กรณีที่ค่า p_i มีลักษณะเบ้จะได้ว่า

$$p_i = \frac{2i}{k(k+1)} \; ; \; i = 1,\ldots,k \text{ และ } k = 4,6,8,10$$

และจะจัด p_i โดยแบ่งตามกรณีย่อยทั้งสาม ดังนี้

$k=4$ $\quad mul\left(\frac{1}{4},\frac{1}{4},\frac{1}{4},\frac{1}{4}\right)$

$mul\left(\frac{1}{6},\frac{2}{6},\frac{2}{6},\frac{1}{6}\right)$

$mul\left(\frac{1}{10},\frac{2}{10},\frac{3}{10},\frac{4}{10}\right)$

$k=6$ $\quad mul\left(\frac{1}{6},\frac{1}{6},\frac{1}{6},\frac{1}{6},\frac{1}{6},\frac{1}{6}\right)$

$mul\left(\frac{1}{12},\frac{2}{12},\frac{3}{12},\frac{3}{12},\frac{2}{12},\frac{1}{12}\right)$

$mul\left(\frac{1}{21},\frac{2}{21},\frac{3}{21},\frac{4}{21},\frac{5}{21},\frac{6}{21}\right)$

$k=8$ $\quad mul\left(\frac{1}{8},\frac{1}{8},\frac{1}{8},\frac{1}{8},\frac{1}{8},\frac{1}{8},\frac{1}{8},\frac{1}{8}\right)$

$mul\left(\frac{1}{20},\frac{2}{20},\frac{3}{20},\frac{4}{20},\frac{4}{20},\frac{3}{20},\frac{2}{20},\frac{1}{20}\right)$

$mul\left(\frac{1}{36},\frac{2}{36},\frac{3}{36},\frac{4}{36},\frac{5}{36},\frac{6}{36},\frac{7}{36},\frac{8}{36}\right)$

$$k = 10 \quad mul\left(\frac{1}{10},\frac{1}{10},\frac{1}{10},\frac{1}{10},\frac{1}{10},\frac{1}{10},\frac{1}{10},\frac{1}{10},\frac{1}{10},\frac{1}{10}\right)$$

$$mul\left(\frac{1}{30},\frac{2}{30},\frac{3}{30},\frac{4}{30},\frac{5}{30},\frac{5}{30},\frac{4}{30},\frac{3}{30},\frac{2}{30},\frac{1}{30}\right)$$

$$mul\left(\frac{1}{55},\frac{2}{55},\frac{3}{55},\frac{4}{55},\frac{5}{55},\frac{6}{55},\frac{7}{55},\frac{8}{55},\frac{9}{55},\frac{10}{55}\right)$$

1.5.3. กำหนดระดับความเชื่อมั่น 3 ระดับคือ 90% , 95% และ 99%

1.6 **ประโยชน์ที่คาดว่าจะได้รับ**

1.6.1 เป็นแนวทางในการตัดสินใจเลือกใช้วิธีประมาณค่าสัดส่วนประชากรให้เหมาะสมกับแต่ละสถานการณ์

1.6.2 เป็นแนวทางในการศึกษาวิจัยเพื่อเปรียบเทียบการประมาณแบบช่วง สำหรับวิธีประมาณแบบอื่นหรือการแจกแจงแบบอื่นต่อไป

บทที่ 2

ทฤษฎีที่เกี่ยวข้องกับงานวิจัย

เนื่องจากงานวิจัยนี้ได้นำเสนอวิธีการประมาณค่าสัดส่วนประชากรแบบช่วง 3 วิธีด้วยกัน ซึ่งแต่ละวิธีมีทฤษฎีที่เกี่ยวข้องแตกต่างกันไป ดังนั้นในบทนี้จะนำเสนอเกี่ยวกับทฤษฎีบทต่างๆ ที่เกี่ยวข้องกับงานวิจัย ดังต่อไปนี้

2.1 ทฤษฎีบทค่าจำกัดสู่ส่วนกลาง

เมื่อสุ่มตัวอย่างขนาดใหญ่จากประชากรที่มีการแจกแจงใดๆ ซึ่งมีความแปรปรวนหาค่าได้ ค่าเฉลี่ยตัวอย่างจะมีการแจกแจงลู่เข้าสู่การแจกแจงปกติ กล่าวคือ

$$\overline{X} \sim N\left(m, s^{2}_{\overline{X}}\right)$$

จะได้ว่า

$$Z = \frac{\overline{X} - m}{s_{x} / \sqrt{n}} \sim N(0,1)$$

2.2 การแจกแจงแบร์นูลลี

การทดลองใดๆที่ให้ผลการทดลองได้ 2 กลุ่มคือ สิ่งที่สนใจและสิ่งที่ไม่สนใจ การทดลองนั้นจะเรียกว่า การทดลองแบร์นูลลี (Bernoulli experiment) กล่าวคือ X เป็นตัวแปรสุ่มของการแจกแจงแบร์นูลลีซึ่งเขียนแทนด้วย $(X \sim Ber(p))$ จะได้ว่า

$$X = \begin{cases} 1 & \text{เมื่อเกิดสิ่งที่สนใจ ด้วยความน่าจะเป็น } p \\ 0 & \text{เมื่อเกิดสิ่งที่ไม่สนใจ ด้วยความน่าจะเป็น } 1-p \end{cases}$$

และฟังก์ชันความน่าจะเป็นของ X อยู่ในรูปของ

$$p_X(x) = p^x (1-p)^{1-x} \quad , \quad x = 0,1$$

เมื่อ $E(X) = p$

และ $Var(X) = p(1-p)$

2.3 การแจกแจงทวินาม

เมื่อทำการทดลองแบร์นูลลี n ครั้ง ที่เป็นอิสระซึ่งกันและกัน การทดลองเกี่ยวกับความน่าจะเป็นของจำนวนครั้งซึ่งเกิดเหตุการณ์ที่สนใจ (x) จะเรียกว่า การทดลองทวินาม (binomial experiment)

ซึ่งฟังก์ชันความน่าจะเป็นของ X อยู่ในรูปของ

$$p_X(x) = \binom{n}{x} p^x (1-p)^{n-x} \qquad , x = 0,1,\ldots,n$$

เมื่อ $E(X) = np$

และ $Var(X) = np(1-p)$

2.4 การแจกแจงพหุนาม

การแจกแจงพหุนามเป็นกรณีที่เราสนใจศึกษาการทดลองที่จำแนกได้มากกว่า 2 กลุ่ม กล่าวคือ ทำการทดลองสุ่มซ้ำๆ กัน n ครั้ง และแต่ละครั้งมีผลการทดลองที่เป็นไปได้จำนวน k กลุ่ม การทดลองนั้นจะเรียกว่า การทดลองพหุนาม (multinomial experiment)

และฟังก์ชันความน่าจะเป็นของ X อยู่ในรูปของ

$$P(X_1 = x_1, X_2 = x_2, \ldots, X_k = x_k) = \frac{n!}{x_1! x_2! \ldots x_k!} p_1^{x_1} p_2^{x_2} \ldots p_k^{x_k}, x_i = 0,1,\ldots,n$$

เมื่อ $E(X_i) = np_i$

และ $Var(X_i) = np_i(1-p_i)$

[1]<u>ตัวอย่างการแจกแจงพหุนาม</u> เจ้าของร้านขายของแห่งหนึ่งได้รวบรวมข้อมูลไว้ว่าลูกค้าที่เข้ามาในร้านของเขาจะมีพฤติกรรมดังนี้ 45% จะซื้อโทรทัศน์ธรรมดา 15% จะซื้อโทรทัศน์สี อีก 40% จะแวะชมเท่านั้น ถ้ามีลูกค้า 5 คนเข้าร้านของเขาในวันใดวันหนึ่ง จงหาความน่าจะเป็นที่เข้าจะขายโทรทัศน์ชนิดธรรมดาได้ 2 เครื่องและชนิดสีได้ 1 เครื่องในวันนั้น โดยสมมติว่าลูกค้าแต่ละคนจะซื้อโทรทัศน์

ให้ $n_1 =$ จำนวนคนที่ซื้อโทรทัศน์ธรรมดา

ให้ $n_2 =$ จำนวนคนที่ซื้อโทรทัศน์สี

ให้ $n_3 =$ จำนวนคนที่จะแวะชมเท่านั้น

ให้ $\underset{\sim}{p}(3\times 1) = \left(\frac{45}{100}, \frac{15}{100}, \frac{40}{100}\right)$

$$m(n_1,\ldots,n_k\,;\underset{\sim}{p}) = \binom{n}{n_1, n_2, \ldots, n_k} p_1^{n_1}\, p_2^{n_2} \ldots p_k^{n_k}$$

$$\therefore m(2,1,2\,;\underset{\sim}{p}) = \left(\frac{5!}{2!\,1!\,2!}\right)(0.45)^2(0.15)\,(0.4)^2$$

$$= 0.1458$$

2.5 [2]**ความสัมพันธ์ระหว่างสถิติอันดับกับการแจกแจงบีตา**

กำหนดให้ $X_{(1)} < X_{(2)} < \ldots < X_{(n)}$ คือสถิติอันดับของตัวแปรเชิงสุ่ม $X_1, X_2, \ldots, X_n$ ซึ่งมาจากประชากรที่มีการแจกแจงแบบต่อเนื่อง โดยมีฟังก์ชันการแจกแจง(สะสม) $F_X(x)$ และฟังก์ชันความหนาแน่น $f_X(x)$ แล้ว จะได้ว่าฟังก์ชันความหนาแน่นของ $X_{(j)}$ คือ

$$P_{X_{(j)}}(x) = \frac{n!}{(j-1)!(n-j)!} f_X(x)[F_X(x)]^{j-1}[1-F_X(x)]^{n-j}$$

ดังนั้นถ้าตัวแปรสุ่ม $X_1, X_2, \ldots, X_n$ มีการแจกแจงแบบเอกรูปในช่วง (0,1) จะได้ว่า $f_X(x) = 1$ เมื่อ $x \in (0,1)$ และ $F_X(x) = x$ เมื่อ $x \in (0,1)$ และเมื่อต้องการหาฟังก์ชันความหนาแน่นของสถิติอันดับที่ j จะได้ว่า

$$P_{X_{(j)}}(x) = \frac{n!}{(j-1)!(n-j)!} x^{j-1}(1-x)^{n-j}$$

$$= \frac{\Gamma(n+1)}{\Gamma(j)\Gamma(n-j+1)} x^{j-1}(1-x)^{(n-j+1)-1}$$

---

[1] ธีระพร วีระถาวร, <u>ความน่าจะเป็นกับการประยุกต์</u>. พิมพ์ครั้งที่ 2. (กรุงเทพฯ : วิทยพัฒน์, 2539), หน้า 153.

[2] สงขลา สำเภาเงิน, “การเปรียบเทียบวิธีการประมาณค่าแบบช่วงสำหรับสัดส่วนประชากรของการแจกแจงแบบทวินาม,” (วิทยานิพนธ์ปริญญามหาบัณฑิต สาขาวิชาสถิติ คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย, 2542), หน้า 16.

ซึ่งก็คือฟังก์ชันความหนาแน่นของการแจกแจงบีตาที่มีพารามิเตอร์คือ j และ $n-j+1$ นั่นก็คือตัวสถิติอันดับที่ j จากตัวแปรสุ่มที่มีการแจกแจงเอกรูปในช่วง (0,1) จะมีการแจกแจงแบบบีตาที่มีพารามิเตอร์ คือ j และ $n-j+1$

2.6 [3]ความสัมพันธ์ระหว่างการแจกแจงบีตากับการแจกแจงเอฟ

ในปี ค.ศ. 1952 ฮาลด์ (Hald) ได้อธิบายความสัมพันธ์ระหว่างฟังก์ชันบีตาที่ไม่สมบูรณ์กับการแจกแจงเอฟดังนี้

ถ้า B คือตัวแปรสุ่มบีตาที่มีพารามิเตอร์ (u_1,u_2) เราสามารถสร้างตัวแปรสุ่มเอฟที่มีพารามิเตอร์ $2u_1$ และ $2u_2$ ได้จากความสัมพันธ์ดังต่อไปนี้

$$f=\frac{u_2 x}{u_1(1-x)} \quad ; 0\le x\le 1$$

เมื่อ x คือค่าที่สอดคล้องกับ $I_X(u_1,u_2)$ ซึ่งก็คือฟังก์ชันความน่าจะเป็นสะสมของตัวแปรสุ่มบีตา หรือที่เรียกว่าอัตราส่วนของฟังก์ชันบีตาที่ไม่สมบูรณ์ (Incomplete Beta function ratio) ซึ่งมีรูปแบบดังนี้

$$\begin{aligned}F(x)&=\frac{1}{B(u_1,u_2)}\int_0^X w^{u_1-1}(1-w)^{u_2-1}\,dw\\&=\frac{B_X(u_1,u_2)}{B(u_1,u_2)}\\&=I_X(u_1,u_2)\end{aligned}$$

โดยที่

$B(y,n-y+1)=\int_0^p w^{y-1}(1-w)^{n-y}\,dw$ คือ ฟังก์ชันบีตา (Beta function)

$B(y,n-y+1)=\int_0^1 w^{y-1}(1-w)^{n-y}\,dw$ คือ ฟังก์ชันบีตาที่ไม่สมบูรณ์ (Incomplete Beta function)

---

[3] สงขลา สำเภาเงิน, "การเปรียบเทียบวิธีการประมาณค่าแบบช่วงสำหรับสัดส่วนประชากรของการแจกแจงแบบทวินาม," (วิทยานิพนธ์ปริญญามหาบัณฑิต สาขาวิชาสถิติ คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย, 2542), หน้า 17.

2.7 [4]การประมาณค่าแบบช่วง

การประมาณค่าแบบช่วงเป็นการประมาณค่าพารามิเตอร์ของประชากรในรูปแบบช่วงโดยใช้ข้อมูลตัวอย่าง การประมาณแบบช่วงนั้นจะบอกถึงค่าต่ำสุดและค่าสูงสุดของพารามิเตอร์ที่เป็นไปได้ ในกรณีที่มีพารามิเตอร์หลายตัวจะเรียกว่า บริเวณความเชื่อมั่น (confidence regions)

กำหนดให้ $X_1, X_2, \ldots, X_n$ เป็นตัวแปรสุ่มจากการแจกแจงซึ่งมี q เป็นพารามิเตอร์ที่ไม่ทราบค่า

สมมติให้ $L(X_1, X_2, \ldots, X_n)$ และ $U(X_1, X_2, \ldots, X_n)$ เป็นตัวสถิติซึ่ง $L < U$ และ

$$P(L(X_1, X_2, \ldots, X_n) < q < U(X_1, X_2, \ldots, X_n)) = 1 - a$$

ถ้า $x_1, x_2, \ldots, x_n$ เป็นค่าสังเกตของ $X_1, X_2, \ldots, X_n$ ดังนั้น $l(x_1, x_2, \ldots, x_n)$ และ $u(x_1, x_2, \ldots, x_n)$ เป็นค่าสังเกตของ $L(X_1, X_2, \ldots, X_n)$ และ $U(X_1, X_2, \ldots, X_n)$ ตามลำดับ

เมื่อทราบค่าของตัวแปรสุ่ม และค่าตัวสถิติ L และ U จะสามารถสร้างช่วง (l, u) และเรียกช่วง (l, u) ที่ได้ว่า $(1-a)100\%$ ช่วงความเชื่อมั่น เมื่อ l เป็นขีดจำกัดความเชื่อมั่นล่าง (lower confidence limit) และ u ว่าขีดจำกัดความเชื่อมั่นบน (upper confidence limit)

2.8 [5]การประมาณค่าสัดส่วนประชากรแบบช่วง

ในการหาค่าประมาณแบบช่วงสำหรับค่าสัดส่วนของประชากร ที่ระดับความเชื่อมั่น $(1-a)100\%$ เมื่อกำหนดค่า p ใดๆ โดยที่ $0 < p < 1$ ในปี ค.ศ. 1982 ลาร์สัน (Larson) ได้อธิบายถึงวิธีการประมาณค่าสัดส่วนของประชากรแบบช่วงไว้ดังนี้

ถ้ากำหนดให้ $h_1(p)$ และ $h_2(p)$ คือขีดจำกัดความเชื่อมั่นล่างและขีดจำกัดความเชื่อมั่นบนของ p ที่ระดับความเชื่อมั่น $(1-a)100\%$ แล้วจะได้ว่า

$$P\{h_1(p) < p < h_2(p) \mid Y = y\} = 1 - a$$

นั่นก็คือในการประมาณค่า p ที่ระดับความเชื่อมั่น $(1-a)100\%$ สามารถหาได้จากค่าจำนวนเต็มที่มีค่ามากที่สุดของ $h_1(p)$ ที่ทำให้ $P\{Y \le h_1(p)\} \le a/2$ และค่าจำนวนเต็มที่เล็กที่สุดของ

---

[4] สงขลา สำเภาเงิน, "การเปรียบเทียบวิธีการประมาณค่าแบบช่วงสำหรับสัดส่วนประชากรของการแจกแจงแบบทวินาม," (วิทยานิพนธ์ปริญญามหาบัณฑิต สาขาวิชาสถิติ คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย, 2542), หน้า 18.

[5] สงขลา สำเภาเงิน, "การเปรียบเทียบวิธีการประมาณค่าแบบช่วงสำหรับสัดส่วนประชากรของการแจกแจงแบบทวินาม," (วิทยานิพนธ์ปริญญามหาบัณฑิต สาขาวิชาสถิติ คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย, 2542), หน้า 19.

$h_2(p)$ ที่ทำให้ $P\{Y \geq h_2(p)\} \leq {}^{a}\!/_{2}$ เนื่องจากการแจกแจงทวินามเป็นการแจกแจงที่ไม่ต่อเนื่อง ดังนั้นจึงไม่สามารถหาค่าให้เท่ากับ ${}^{a}\!/_{2}$ ได้

และเนื่องจากเหตุการณ์ $P\{Y \leq h_1(p)\} \leq {}^{a}\!/_{2}$ และ $P\{Y \geq h_2(p)\} \leq {}^{a}\!/_{2}$ เป็นเหตุการณ์ที่ไม่เกิดร่วมกัน (mutually disjoint) จะได้ว่า

$$P\{Y \leq h_1(p) \text{ หรือ } Y \geq h_2(p)\} \leq \frac{a}{2} + \frac{a}{2} = a$$

ดังนั้นเมื่อพิจารณาส่วนเติมเต็มของเหตุการณ์ จะพบว่ามีค่าความน่าจะเป็นอย่างน้อยที่สุด $1-a$ นั่นก็คือ

$$\begin{aligned} P\{h_1(p) \leq Y \leq h_2(p)\} &= 1 - P\{Y \leq h_1(p) \text{หรือ} Y \geq h_2(p)\} \\ &\geq 1-a \end{aligned}$$

เมื่อพิจารณาฟังก์ชัน $h_1(p)$ และ $h_2(p)$ ซึ่งเป็นฟังก์ชันหนึ่งต่อหนึ่งจึงสามารถหาฟังก์ชันผกผัน (inverse function) ของเหตุการณ์ได้ นั่นคือเมื่อ $Y = h_1(p)$ จะได้ว่า $p = h_1^{-1}(Y)$ ในทำนองเดียวกันกับฟังก์ชัน $h_2(p)$ ดังนั้น

$$P\left\{h_1^{-1}(Y) \leq p \leq h_2^{-1}(Y)\right\} \geq 1-a$$

แสดงว่า $\left\{h_1^{-1}(Y), h_2^{-1}(Y)\right\}$ คือค่าประมาณแบบช่วงของ p ที่ระดับความเชื่อมั่น $(1-a)100\%$

เมื่อพิจารณา $h_1(p)$ เนื่องจากคือจำนวนเต็มที่มีค่ามากที่สุดที่ทำให้ $P\{\ Y \leq h_1(p)\ \} \leq {}^{a}\!/_{2}$ ดังนั้นจึงได้ว่า $\sum_{j=0}^{h_1(p)} \binom{n}{j} p^j (1-p)^{n-j} \leq {}^{a}\!/_{2}$

สมมติได้ค่าสังเกตคือ $Y = y$ จะได้ว่า $h_1^{-1}(y)$ คือขีดจำกัดความเชื่อมั่นบนของ p และเนื่องจาก $h_1\left(h_1^{-1}(y)\right) = y$ ดังนั้นการหาขีดจำกัดความเชื่อมั่นบนของ p คือการหาค่า p ที่ทำให้

$$\sum_{j=0}^{y} \binom{n}{j} p^j (1-p)^{n-j} = {}^{a}\!/_{2}$$

ทำนองเดียวกัน $h_2\left(h_2^{-1}(y)\right) = y$ ดังนั้นขีดจำกัดความเชื่อมั่นล่างของ p ก็คือค่าของ p ที่ทำให้

$$\sum_{j=y}^{n} \binom{n}{j} p^j (1-p)^{n-j} = {}^{a}\!/_{2}$$

ดังนั้นสามารถสรุปได้ว่า ถ้า Y เป็นตัวแปรสุ่มที่มีการแจกแจงแบบทวินาม ที่มีพารามิเตอร์ n และ p ให้ p_L และ p_U คือค่าที่ทำให้

$$\sum_{j=y}^{n} \binom{n}{j} {p_L}^j (1-p_L)^{n-j} = {}^{a}\!/_{2}$$

$$\sum_{j=0}^{y} \binom{n}{j} {p_U}^j (1-p_U)^{n-j} = {}^{a}\!/_{2}$$

ดังนั้น จะได้ว่า (p_L, p_U) คือค่าประมาณสัดส่วนของประชากรแบบช่วง ณ ระดับความเชื่อมั่น $(1-a)100\%$

จากสูตรการหาขีดจำกัดความเชื่อมั่นสำหรับค่าสัดส่วนประชากร ซึ่งต้องอาศัยการแก้สมการและคำนวณค่าแบบแฟคทอเรียล พบว่าถ้า n มีค่ามากจะประสบปัญหาในการประมาณค่า ดังนั้นจึงต้องอาศัยการประมาณโดยวิธีการอื่นๆ ช่วยในการหาขีดจำกัดความเชื่อมั่นสำหรับค่าสัดส่วนประชากร

2.9.1 **วิธีการประมาณปกติ**

จากทฤษฎีบทค่าจำกัดสู่ส่วนกลางกล่าวว่า ถ้าผลคูณระหว่างขนาดตัวอย่างกับค่าสัดส่วนประชากรตัวอย่างมีมากพอ $\hat{p}_i$ จะมีการแจกแจงแบบปกติโดยประมาณ ซึ่งมีค่าเฉลี่ย p_i และความแปรปรวน $p_i(1-p_i)/n$ กล่าวคือ

$$\hat{p}_i \sim N\left(p_i, \frac{p_i(1-p_i)}{n}\right)$$

ดังนั้น

$$Z = \frac{\hat{p}_i - p_i}{\sqrt{p_i(1-p_i)/n}} \sim N(0,1)$$

จะได้ว่า

$$P\left(-Z_{1-a/2k} < Z < Z_{1-a/2k}\right) = 1-a$$

กล่าวคือ

$$P\left(-Z_{1-a/2k} < \frac{\hat{p}_i - p_i}{\sqrt{p_i(1-p_i)/n}} < Z_{1-a/2k}\right) = 1-a$$

$$\therefore \quad P\left(\hat{p}_i - Z_{1-a/2k}\sqrt{\frac{p_i(1-p_i)}{n}} < p_i < \hat{p}_i + Z_{1-a/2k}\sqrt{\frac{p_i(1-p_i)}{n}}\right) = 1-a$$

และได้ค่าประมาณแบบช่วงคือ

$$\left(\hat{p}_i - Z_{1-a/2k}\sqrt{\hat{p}_i(1-\hat{p}_i)/n} \quad , \quad \hat{p}_i + Z_{1-a/2k}\sqrt{\hat{p}_i(1-\hat{p}_i)/n}\right)$$

2.9.2 [6]วิธีการประมาณแบบคิวเซนเบอร์รี่ และ เฮิร์ส (Quesenberry and Hurst's Method)

ในปี ค.ศ. 1900 เพียร์สัน (Pearson) ได้แสดงว่า ถ้า $E(n_i) = np_i$ ขนาดที่ใหญ่พอสำหรับทุก $i = 1, 2, \ldots, k$ จะได้ว่า

$$c^2 = \sum_{i=1}^{k} (n_i - np_i)^2 \Big/ np_i \tag{1}$$

มีการแจกแจงโดยประมาณเป็นไคสแควร์ ด้วย $(k-1)$ ระดับขั้นความเสรี ในปี ค.ศ. 1954 ค็อกเครน (Cochran) กล่าวว่า ถ้า n มีขนาดมากพอ np_i อย่างน้อยเท่ากับ 5 การประมาณจะทำได้ดี ซึ่งจะได้ว่า

$$P\left\{\sum_{i=1}^{k} \left(np_i^2 / p_i\right) - n \le c^2_{a,k-1}\right\} = 1 - a\,, \quad p_i = n_i / n \tag{2}$$

เมื่อ $c^2_{a,k-1}$ เป็น ขอบเขตบนของค่าเปอร์เซ็นต์ a ของการแจกแจงไคสแควร์

กำหนด a, k และ N

$$\sum_{i=1}^{k} p_i^2 / p_i \le \left(c^2_{a,k-1} / n\right) + 1 \equiv C \tag{3}$$

และกำหนดให้

$$\sum_{i=1}^{k} p_i = 1 \tag{4}$$

สมการ (3) จะเป็นขอบเขตของบริเวณเพราะฉะนั้นจะหาค่ามากสุดและน้อยสุด

จากสมการที่ (3) จะได้ว่า

$$p_i = p_i^2 \Bigg/ \left(C - \sum_{\substack{j=1 \\ j \ne i}}^{k} p_j^2 / p_j\right), \quad (i = 1, 2, \ldots, k) \tag{5}$$

ดังนั้นฟังก์ชันที่มีค่ามากสุดและน้อยสุด คือ

$$Q = (1 - l)p_i^2 \Bigg/ \left(C - \sum_{\substack{j=1 \\ j \ne i}}^{k} p_i^2 / p_i\right) - l\left(\sum_{\substack{j=1 \\ j \ne i}}^{k} p_j - 1\right) \tag{6}$$

เมื่อ l เป็นตัวคูณของลาแกรน(*Lagrange* Multiplier) โดยที่เมื่อหาอนุพันธ์เทียบกับ p_i และ p_m และจัดเท่ากับศูนย์ จะได้ว่า

---

[6] Quesenberry, C.P. and Hurst, D. C.. Large-Sample Simultaneous Confidence Interval for Multinomial Proportions. Technometrics 6 : 191-195

$$l p_i^2/(l-1)p_i^2 = p_i^2/p_i^2 \quad ; \qquad i, j, m = 1, 2, \ldots, k$$
$$l p_i^2/(l-1)p_i^2 = p_m^2/p_m^2 \quad ; \qquad j \neq i \,,\, m \neq i \,,\, m \neq j \qquad (7)$$

ดังนั้น

$$p_m = p_m p_j / p_j \qquad (8)$$

จาก $(k-2)$ สมการ และจากสมการที่ (3) และ (4) จะได้ว่า

$$p_i = \frac{C + 2p_i - 1 \pm \left[(C + 2p_i - 1)^2 - 4C2p_i^2\right]^{\frac{1}{2}}}{2C} \qquad (9)$$

หรือ อาจจะเขียนได้ใหม่ในรูปที่คำนวณได้ง่ายขึ้นได้

$$p_i = \frac{c^2 + 2n_i \pm \left\{c^2\left[c^2 + 4n_i(n - n_i)/n\right]\right\}^{\frac{1}{2}}}{2(n + c^2)} \qquad (10)$$

2.9.3 [7]**วิธีการประมาณแบบเอฟ** (F Method)

การประมาณค่าสัดส่วนประชากร โดยใช้การแจกแจงเอฟนั้น ในปี ค.ศ. 1986 บริทซ์ (Blyth) ได้แสดงวิธีประมาณค่าโดยอาศัยความสัมพันธ์ระหว่างสถิติอันดับกับการแจกแจงบีตา ดังนี้

ถ้า $X_{(1)} < X_{(2)} < \ldots < X_{(n)}$ เป็นตัวสถิติอันดับของตัวแปรสุ่ม $X_1, X_2, \ldots, X_n$ ที่มีการแจกแจงแบบเอกรูปในช่วง $(0,1)$ โดยอาศัยคุณสมบัติของสถิติอันดับ จะได้ว่า $X_{(y)}$ มีการแจกแจงแบบบีตาที่มีพารามิเตอร์ y และ $n-y+1$ ถ้ากำหนดให้ Y คือจำนวน X_i ซึ่งมีค่าน้อยกว่า p ดังนั้น Y จึงมีการแจกแจงทวินามที่มีพารามิเตอร์ n และ p เนื่องจากเหตุการณ์ $Y \geq y$ ก็คือเหตุการณ์ $X_{(y)} \leq p$ ดังนั้นจึงสามารถหาค่า $P(Y \geq y)$ โดยพิจารณาจากการหาค่า $P(X_{(y)} \leq p)$ ซึ่งสามารถนำมาใช้ประโยชน์ในการหาขีดจำกัดความเชื่อมั่นล่างของค่าประมาณแบบช่วงได้ดังนี้

$$\begin{aligned} P(Y \geq y) &= P(X_{(y)} \leq p) \\ &= \frac{\Gamma(n+1)}{\Gamma(y)\Gamma(n-y+1)} \int_0^p x^{y+1}(1-x)^{n-y}\,dx \end{aligned}$$

กำหนดให้ $t = (n-y+1)x / y(1-x)$

จะได้ว่า

---

[7] Leemis, L.M. and Trivedi, K.S.. A Comparison of Approximate Interval Estimators for the Binomial Parameter. The American Statistician 50 : 63-68.

$$P(Y \geq y) = \frac{\Gamma(n+1)}{\Gamma(y)\Gamma(n-y+1)}\left(\frac{n-y+1}{y}\right)^{n-y+1} \times \int_0^{(n-y+1)p/y(1-p)} \frac{t^{y-1}}{\left(\frac{n-y+1}{y}+t\right)} dt$$

$$= P\left(F_{2y,2(n-y+1)} < \frac{(n-y+1)p}{y(1-p)}\right)$$

ในปี ค.ศ. 1970 จอห์นสันและโคลท์ซ (Johnson and Kotz) กล่าวว่าจะสามารถหาค่าขีดจำกัดความเชื่อมั่นล่างที่ $(a/2)100\%$ ของตัวแปรสุ่มเอฟที่มีระดับขั้นความเสรี $2y$ และ $2(n-y+1)$ ได้ดังสมการข้างล่าง

$$I_{2yF/2(n-y+1)+2yF}\{y, n-y+1\} = a/2$$

ดังนั้น การหาขีดจำกัดความเชื่อมั่นล่างสำหรับค่าสัดส่วนประชากรจะสามารถเขียนได้ในรูปของฟังก์ชันบีตาดังนี้

$$I_{P_L}\{y, n-y+1\} = a/2$$

เพราะฉะนั้น

$$P_L = \frac{2yF_{(2y,2(n-y+1)),a/2}}{2(n-y+1)+2yF_{(2y,2(n-y+1)),a/2}}$$

หรือสามารถเขียนได้ในรูปของ

$$P_L = \frac{y}{y+(n-y+1)F_{(2y,2(n-y+1)),1-a/2}}$$

ในทำนองเดียวกันจะได้ขีดจำกัดความเชื่อมั่นบนคือ

$$P_U = \frac{(y+1)F_{(2(y+1),2(n-y)),1-a/2}}{n-y+(y+1)F_{(2(y+1),2(n-y)),1-a/2}}$$

และได้ค่าประมาณแบบช่วงที่พิจารณาจากการแจกแจงเอฟดังนี้

$$\left(\frac{y}{y+(n-y+1)F_{(2(n-y+1),2y),1-a/2}} \;\; , \;\; \frac{(y+1)F_{(2(y+1),2(n-y)),1-a/2}}{n-y+(y+1)F_{(2(y+1),2(n-y)),1-a/2}} \right)$$

ในทำนองเดียวเราสามารถแสดงได้ว่า กรณีที่มีจำนวนพารามิเตอร์ k ตัว จะได้ค่าประมาณแบบช่วงดังนี้

$$\left(\frac{y_i}{y_i+(n-y_i+1)F_{(2(n-y_i+1),2y_i),1-a/2k}} \;\; , \;\; \frac{(y_i+1)F_{(2(y_i+1),2(n-y_i)),1-a/2k}}{n-y_i+(y_i+1)F_{(2(y_i+1),2(n-y_i)),1-a/2k}} \right)$$

y_i แทน จำนวนครั้งของเหตุการณ์ที่สนใจ i , $i = 1,2,...,k$

2.10 [8]ช่วงความเชื่อมั่นเชิงซ้อน (Multiple Confidence Intervals)

ในการหาบริเวณความเชื่อมั่นเราอาจจะเริ่มต้นโดยการหาช่วงความเชื่อมั่นสำหรับแต่ละสมาชิก q_j , $i = 1,2,...,k$ ของ $\underset{\sim}{q}(k \times 1)$ และรวมช่วงความเชื่อมั่นเหล่านี้เข้าด้วยกัน ก็จะได้บริเวณความเชื่อมั่นสำหรับ $\underset{\sim}{q}$

ให้ w_i เป็นช่วงความเชื่อมั่นขนาด $= 1 - a_i$ สำหรับ q_i , $i = 1,2,...,k$ และสมมติว่า w_i ไม่ได้ครอบคลุม q_i , $j \neq i$

ให้ E_i แทนเหตุการณ์ซึ่ง w_i บรรจุค่าจริง q_i

จะได้ว่า $P(E_i) \geq 1 - a_i$ และ $P(E_i^C) \leq a_i$

ให้ w เป็นบริเวณซึ่งได้จากการรวม $w_1, w_2, ..., w_k$

<u>กรณีที่ 1 ถ้า $E_1, E_2, ..., E_k$ เป็นอิสระซึ่งกันและกัน จะได้ว่า</u>

$$\begin{aligned} P(\underset{\sim}{q} \in w) &= P\left(\bigcap_{i=1}^{k} E_i \right) \\ &= \prod_{i=1}^{k} P(E_i) \\ &\geq \prod_{i=1}^{k} (1 - a_i) \end{aligned}$$

---

[8] ธีระพร วีระถาวร, <u>การอนุมานเชิงสถิติขั้นกลาง : โครงสร้างและความหมาย</u>. (กรุงเทพฯ : จุฬาลงกรณ์มหาวิทยาลัย, 2536), หน้า 241.

$\therefore$ สัมประสิทธิ์ความเชื่อมั่นสำหรับ $w = \prod_{i=1}^{k} (1 - a_i)$

<u>กรณีที่ 2 ถ้า $E_1, E_2, ..., E_k$ ไม่เป็นอิสระซึ่งกันและกัน</u>

$$P(\underset{\sim}{q} \in w) = P\left(\bigcap_{i=1}^{k} E_i\right)$$

$$= 1 - P\left(\bigcup_{i=1}^{k} E_i^C\right)$$

$$> 1 - \sum_{i=1}^{k} P(E_i^C)$$

$$\geq 1 - \sum_{i=1}^{k} a_i$$

$\therefore$ สัมประสิทธิ์ความเชื่อมั่นสำหรับ w จะมีค่าอย่างน้อยที่สุด $1 - \sum_{i=1}^{k} a_i$

2.11 เกณฑ์การเปรียบเทียบ

การตรวจสอบวิธีการประมาณแบบช่วงสำหรับพารามิเตอร์ของการแจกแจงพหุนามทั้ง 3 วิธีจะทำการตรวจสอบค่าระดับความเชื่อมั่นและเปรียบเทียบค่าสัมประสิทธิ์ของบริเวณความเชื่อมั่นที่กำหนดได้จากแต่ละสถานการณ์ในการทดลองซ้ำ 2000 ครั้ง โดยมีรายละเอียดการตรวจสอบดังนี้ ถ้าทำการทดสอบสมมติฐานว่าง

$H_0 : p \geq p_0$

เทียบกับ $H_1 : p < p_0$

จะได้ว่า

$$-Z_{1-a_0} < \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} < 1$$

เพราะฉะนั้น

$$-Z_{1-a_0} \sqrt{\frac{p_0(1-p_0)}{n}} < \hat{p} - p_0$$

$$p_0 - Z_{1-a_0} \sqrt{\frac{p_0(1-p_0)}{n}} < \hat{p}$$

เนื่องจาก $0 < \hat{p} < 1$ จะได้ว่า

$$p_0 - Z_{1-a_0}\sqrt{\frac{p_0(1-p_0)}{n}} < \hat{p} < 1$$

ดังนั้นความเชื่อมั่นของการยอมรับสมมติฐานว่าง H_0 คือ

$$\left(p_0 - Z_{1-a_0}\sqrt{\frac{p_0(1-p_0)}{n}} \ , \ 1 \right)$$

กล่าวคือวิธีการประมาณให้ค่าระดับความเชื่อมั่นไม่ต่ำกว่า p_0 ที่กำหนดไว้ ถ้า $\hat{p}$ มีค่าอยู่ในช่วง

$$\left(p_0 - Z_{1-a_0}\sqrt{\frac{p_0(1-p_0)}{n}} \ , \ 1 \right)$$

เมื่อ a_0 คือระดับนัยสำคัญที่กำหนดในการทดสอบ ซึ่งการวิจัยครั้งนี้กำหนดค่า $a_0 = 0.05$

p^* คือระดับความเชื่อมั่น

$\hat{p}^{**}$ คือระดับความเชื่อมั่นที่ได้จากการทดลอง

p_0^{***} คือระดับความเชื่อมั่นที่กำหนด

(0.90 , 0.95 และ 0.99)

และ n จำนวนครั้งของการทดลองซึ่งเท่ากับ 2,000 รอบ

1. ณ ระดับความเชื่อมั่น 90% ต้องการทดสอบสมมติฐานว่าง

$H_0 : p \geq 0.90$

เทียบกับ $H_1 : p < 0.90$

จะได้ว่าวิธีการให้ค่าระดับความเชื่อมั่นไม่ต่ำกว่าระดับความเชื่อมั่นที่กำหนด ถ้า $\hat{p}$ มีค่าอยู่ในช่วง

---

* p แทน ค่าระดับความเชื่อมั่นที่แท้จริง $(1-a)$

** $\hat{p}$ แทน ค่าระดับความเชื่อมั่นที่ได้จากการจำลองข้อมูลซึ่งคำนวณได้จาก $1-\sum_{i=1}^{k} a_j$ กรณีช่วงความเชื่อมั่นไม่อิสระซึ่งกันและกัน และคำนวณได้จาก $\prod_{i=1}^{k}(1-a_i)$ กรณีช่วงความเชื่อมั่นอิสระซึ่งกันและกัน

*** p_0 แทน ค่าระดับความเชื่อมั่นที่เป็นเกณฑ์กำหนด แบ่งเป็น ณ ระดับความเชื่อมั่นที่ 90% , 95% และ 99%

$$\left(0.90-1.645\sqrt{\frac{0.90(0.10)}{2000}}\ ,1\right)$$

ซึ่งเท่ากับ

$$(0.8890,1)$$

2. ณ ระดับความเชื่อมั่น 95% ต้องการทดสอบสมมติฐานว่าง

$H_0 : p \geq 0.95$

เทียบกับ $H_1 : p < 0.95$

จะได้ว่าวิธีการให้ค่าระดับความเชื่อมั่นไม่ต่ำกว่าระดับความเชื่อมั่นที่กำหนด ถ้า $\hat{p}$ มีค่าอยู่ในช่วง

$$\left(0.95-1.645\sqrt{\frac{0.95(0.05)}{2000}}\ ,1\right)$$

ซึ่งเท่ากับ

$$(0.9420,1)$$

3. ณ ระดับความเชื่อมั่น 99% ต้องการทดสอบสมมติฐานว่าง

$H_0 : p \geq 0.99$

เทียบกับ $H_1 : p < 0.99$

จะได้ว่าวิธีการให้ค่าระดับความเชื่อมั่นไม่ต่ำกว่าระดับความเชื่อมั่นที่กำหนดถ้า $\hat{p}$ มีค่าอยู่ในช่วง

$$\left(0.99-1.645\sqrt{\frac{0.99(0.01)}{2000}}\ ,1\right)$$

ซึ่งเท่ากับ

$$(0.9863,1)$$

เพราะฉะนั้นค่าระดับความเชื่อมั่นที่ได้จากการทดลองจะต้องมีค่าไม่ต่ำกว่า 0.8890, 0.9420, 0.9863 ณ ระดับความเชื่อมั่น 90%, 95%, 99% ตามลำดับ จึงจะถือว่าวิธีการประมาณนั้นให้ค่าระดับความเชื่อมั่นไม่ต่ำกว่าเกณฑ์ที่กำหนดและจะนำเฉพาะวิธีการประมาณที่ให้ค่าระดับความเชื่อมั่นไม่ต่ำกว่าเกณฑ์ที่กำหนดไปทำการเปรียบเทียบค่าของช่วงความเชื่อมั่นต่อไป โดยจะทำการพิจารณาเปรียบเทียบค่าสัมประสิทธิ์บริเวณความเชื่อมั่นของแต่ละวิธีซึ่งจะหาได้ว่า

สัมประสิทธิ์บริเวณความเชื่อมั่น $= \begin{cases} 1-\sum_{i=1}^{k} a_j & \text{กรณีช่วงความเชื่อมั่นไม่อิสระซึ่งกันและกัน} \\ \prod_{i=1}^{k}(1-a_i) & \text{กรณีช่วงความเชื่อมั่นอิสระซึ่งกันและกัน} \end{cases}$

จากการประมาณทั้ง 3 วิธี ช่วงความเชื่อมั่นของแต่ละพารามิเตอร์ที่ได้จากการประมาณแบบคิวเซนเบอร์รี่และเฮิร์สจะไม่เป็นอิสระซึ่งกันและกัน แต่วิธีการประมาณแบบปกติ และวิธีการประมาณแบบเอฟจะให้ช่วงความเชื่อมั่นของพารามิเตอร์ที่เป็นอิสระซึ่งกันและกัน

ดังนั้นการประมาณแบบคิวเซนเบอร์นี่และเฮิร์ส จะมีสัมประสิทธิ์บริเวณความเชื่อมั่นเท่ากับ $1-\sum_{i=1}^{k} a_j$ ส่วนวิธีการประมาณแบบปกติ และการประมาณแบบเอฟจะมีสัมประสิทธิ์บริเวณความเชื่อมั่นเท่ากับ $\prod_{i=1}^{k}(1-a_i)$ วิธีการประมาณใดที่ให้ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นเฉลี่ยมากที่สุด จะสรุปได้ว่าวิธีการประมาณนั้นจะให้ช่วงประมาณที่เหมาะสมที่สุดสำหรับสถานการณ์นั้นๆ เพราะค่าสัมประสิทธิ์บริเวณความเชื่อมั่นจะแสดงถึงความแม่นยำในการประมาณค่าแบบช่วงของวิธีนั้นๆ

นอกจากเกณฑ์ข้างต้นแล้วจะใช้ค่าระยะห่างของค่าบริเวณช่วงความเชื่อมั่นของแต่ละวิธีมาพิจารณาประกอบการเปรียบเทียบโดยจะคำนวณระยะห่างดังกล่าวด้วยวิธี Mahalanobis distance

$$D_{ijk} = (L_{ik} - U_{ik})'(nI \times S^{-1})(L_{ik} - U_{ik})$$

เมื่อ L_{ik} คือค่าขอบเขตล่างของช่วงความเชื่อมั่นของพารามิเตอร์ตัวที่ i วิธีที่ k

U_{ik} คือค่าขอบเขตบนของช่วงความเชื่อมั่นของพารามิเตอร์ตัวที่ i วิธีที่ k

S คือเมทริกซ์ความแปรปรวนร่วมของการแจกแจงพหุนาม

n คือขนาดตัวอย่าง

และ I คือเมทริกซ์เอกลักษณ์ (Identity Matrix)

ซึ่งวิธีการใดให้ค่าระยะห่างของค่าบริเวณช่วงความเชื่อมั่นเฉลี่ยต่ำจะถือว่าวิธีนั้นเหมาะสมในสถานการณ์นั้นๆ

บทที่ 3

วิธีดำเนินการวิจัย

ในงานวิจัยนี้ต้องการเปรียบเทียบวิธีการประมาณค่าสัดส่วนประชากรแบบช่วง 3 วิธี คือ วิธีการประมาณแบบปกติ วิธีการประมาณแบบคิวเซนเบอร์รี่และเฮิร์ส และวิธีการประมาณแบบเอฟ นั่นคือให้ค่าระดับความเชื่อมั่นไม่ต่ำกว่าเกณฑ์ที่กำหนดและมีค่าสูงที่สุดจะเป็นวิธีการประมาณที่เหมาะสม โดยจะทำการเปรียบเทียบที่ระดับความเชื่อมั่น 3 ระดับ คือ 90%, 95% และ 99% ขนาดตัวอย่างมีค่า 50 100 200 500 และ 1000 ค่าสัดส่วนประชากรแบ่ง ออกเป็น 3 ประเภท* คือ เท่ากัน เบ้ และสมมาตร และแบ่งจำนวนกลุ่มประชากรเป็น 4, 6, 8 และ 10 กลุ่ม ข้อมูลที่ใช้ในการทำวิจัยได้มาจากการจำลองด้วยเทคนิคมอนติคาร์โลจำนวน 2000 รอบ และเขียนด้วยโปรแกรม MATLAB

แผนการดำเนินการวิจัย และขั้นตอนในการดำเนินงานวิจัย ตลอดจนโปรแกรมที่ใช้เสนอเป็นรายละเอียดดังนี้

3.1 **แผนการดำเนินการวิจัย**

กำหนดสถานการณ์ต่างๆ เพื่อการเปรียบเทียบดังนี้

3.1.1 ขนาดตัวอย่าง (n) มีค่า 50 100 200 500 และ1000

3.1.2 กำหนดค่าสัดส่วนประชากร (p_i) เป็น 3 ลักษณะ ค่าสัดส่วนตัวอย่างมีลักษณะเท่ากัน ค่าสัดส่วนตัวอย่างมีลักษณะเบ้ ค่าสัดส่วนตัวอย่างมีลักษณะสมมาตร

3.1.3 กำหนดจำนวนกลุ่มของสัดส่วนประชากร (k) 4, 6, 8 และ 10 กลุ่ม

3.1.4 กำหนดค่าระดับความเชื่อมั่น 3 ระดับ คือ 90%, 95% และ 99%

ดังนั้น จำนวนสถานการณ์ที่ใช้ในการวิจัยมีจำนวน $= 5\times 3\times 4\times 3$

$= 180$ สถานการณ์

โดยที่ในแต่ละสถานการณ์ทดลองจะทำการเปรียบเทียบค่าระดับความเชื่อมั่นและค่าความยาวเฉลี่ยของค่าประมาณแบบช่วงที่คำนวณจากวิธีการประมาณทั้ง 3 วิธี เพื่อหาวิธีประมาณค่าที่เหมาะสมที่สุดของแต่ละสถานการณ์

---

* เท่ากัน หมายถึง กรณีที่ค่าสัดส่วนประชากรมีการกระจายเท่ากัน

เบ้ หมายถึง กรณีที่ค่าสัดส่วนประชากรมีการกระจายแบบเบ้

สมมาตร หมายถึง กรณีที่ค่าสัดส่วนประชากรมีการกระจายแบบสมมาตร

3.2 ขั้นตอนในการดำเนินการวิจัย

การดำเนินการวิจัยมีขั้นตอนต่อไปนี้

3.2.1 สร้างข้อมูลที่ใช้ในการวิจัย

3.2.2 คำนวณค่าประมาณแบบช่วง

3.2.3 คำนวณค่าสัมประสิทธิ์ช่วงความเชื่อมั่นที่ได้จากการทดลอง

3.2.4 คำนวณค่าระยะห่างของค่าบริเวณความเชื่อมั่น

3.2.5 เปรียบเทียบค่าสัมประสิทธิ์ช่วงความเชื่อมั่นที่ได้จากการทดลองและค่าระยะห่างของค่าบริเวณความเชื่อมั่น

3.2.6 สรุปผลการวิจัยในแต่ละสถานการณ์

ซึ่งรายละเอียดของแต่ละขั้นตอนมีดังนี้

<u>3.2.1 สร้างข้อมูลที่ใช้ในการวิจัย</u>

เนื่องจากข้อมูลที่ใช้ในการวิจัยต้องใช้เทคนิคมอนติคาร์โลในการจำลองข้อมูล จึงต้องเริ่มตั้งแต่สร้างตัวเลขสุ่มที่มีการแจกแจงเอกรูปในช่วง (0,1) เพื่อนำไปใช้ในการสร้างตัวแปรสุ่มที่มีการแจกแจงแบบพหุนามต่อไป โดยมีรายละเอียดดังนี้

<u>การสร้างตัวเลขสุ่ม</u>

การสร้างตัวเลขสุ่มที่มีการแจกแจงเอกรูปในช่วง (0,1) สร้างโดยใช้ฟังก์ชันย่อยคือ

```
FUNCTION  rand( )
```

รายละเอียดแสดงในภาคผนวก

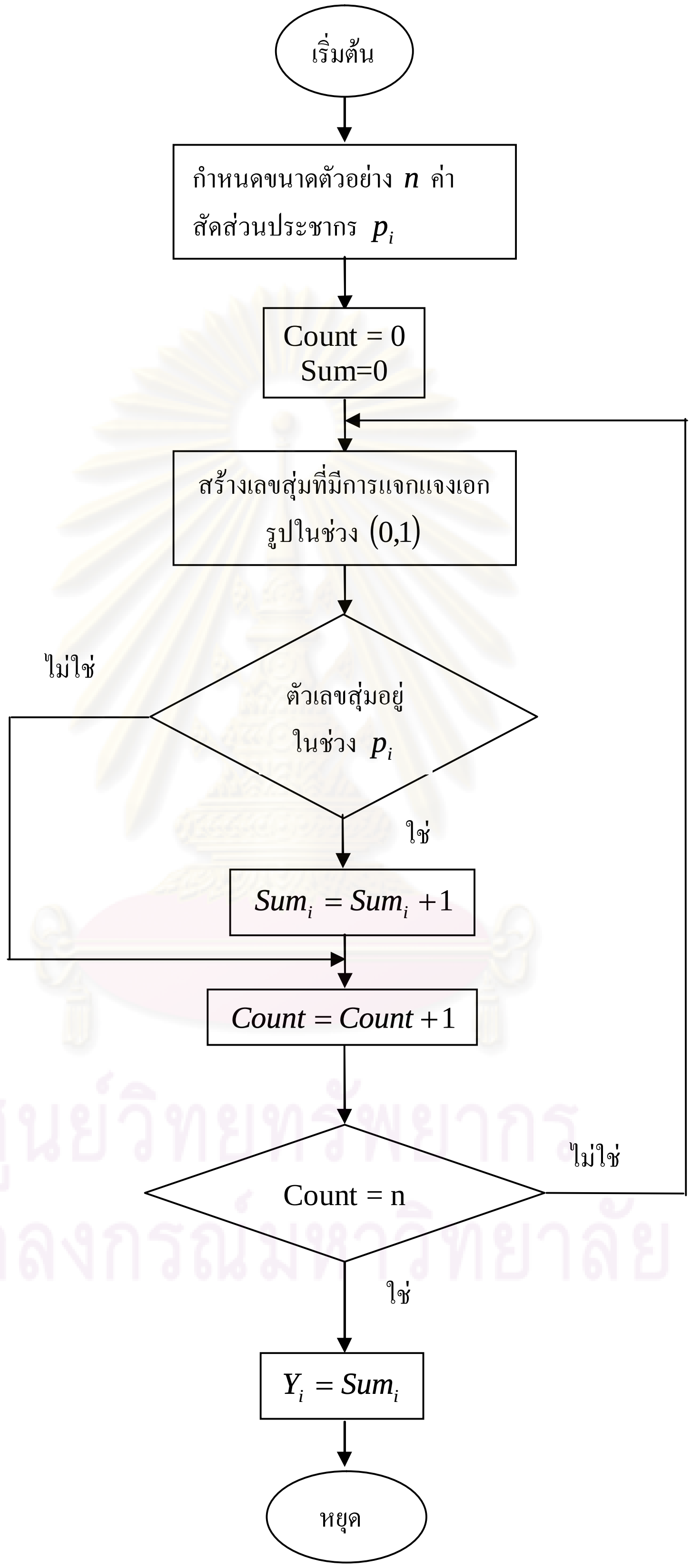
rand( ) คือ ฟังก์ชันในการสร้างตัวเลขสุ่มที่มีการแจกแจงเอกรูปในช่วง (0,1) 1 ค่า

<u>การสร้างตัวแปรสุ่มที่มีการแจกแจงพหุนาม</u>

การสร้างตัวแปรสุ่มที่มีการแจกแจงพหุนาม 1 ค่า มีหลักในการสร้างดังนี้

1. ในการสร้างตัวแปรสุ่ม Y ให้มีการแจกแจงพหุนามที่มีพารามิเตอร์ n และ p_i ; $i = 1,2,...,k$ สามารถทำได้โดยการสร้างตัวแปรสุ่มที่มีเหตุการณ์ที่เป็นไปได้ k เหตุการณ์ จำนวน n ตัว ซึ่งต่างก็เป็นอิสระซึ่งกันต่อกัน กำหนดให้ $X_1, X_2,..., X_n$ คือตัวแปรสุ่มแต่ละตัวที่ได้จากการนำค่าของตัวแปรสุ่มมาสร้างตามเงื่อนไข คือจะนับเป็นผลของเหตุการณ์ที่สนใจถ้าค่านั้นมีค่าอยู่ในช่วงพารามิเตอร์ p_i แต่ละตัว

2. เมื่อได้ตัวแปรสุ่มที่มีการแจกแจงดังกล่าว แล้วนำค่าของ $X_1, X_2,..., X_n$ มาบวกเข้าด้วยกัน ผลบวกที่ได้จะกำหนดให้เป็นค่าของตัวแปรสุ่ม X หรือก็คือ จำนวนผลสำเร็จจากการทดลอง n ครั้ง และจากขั้นตอนนี้จะได้ตัวแปรสุ่ม X ที่มีการแจกแจงพหุนามโดยมีพารามิเตอร์คือ n และ p_i



รูปที่ 3.1 แสดงแผนผังงานสำหรับสร้างตัวแปรสุ่มทวินาม

3.2.2 **คำนวณค่าประมาณแบบช่วง**

เมื่อสร้างตัวแปรสุ่มที่มีการแจกแจงพหุนาม โดยมีพารามิเตอร์คือ n และ p_i ได้ 1 ค่า จะได้ค่าสัดส่วนตัวอย่าง คือ $\hat{p}_i = \dfrac{n_i}{n}$ ซึ่งค่าเหล่านี้จะนำไปใช้ในหาค่าประมาณแบบช่วงจากวิธีการประมาณทั้ง 3 วิธี ดังนี้

3.2.2.1 วิธีการประมาณแบบปกติ มีสูตรการประมาณค่าช่วงความเชื่อมั่น คือ

$$p_L = \hat{p}_i - z_{1-a/(2k)}\sqrt{\frac{\hat{p}_i(1-\hat{p}_i)}{n}}$$

และ $$p_U = \hat{p}_i + z_{1-a/(2k)}\sqrt{\frac{\hat{p}_i(1-\hat{p}_i)}{n}}$$

ผังงานสำหรับการคำนวณค่าประมาณแบบช่วงจากการประมาณแบบปกติ แสดงในรูปที่ 3.2

3.2.2.2 วิธีการประมาณแบบคิวเซนเบอร์รี่และเฮิร์ส มีสูตรการประมาณค่าช่วงความเชื่อมั่น คือ

$$p_L = \left(\frac{c^2_{k-1,1-a} + 2n_i - \sqrt{c^2_{k-1,1-a}\left(c^2_{k-1,1-a} + 4\frac{n_i}{n}(n-n_i) \right)}}{2\left(n + c^2_{k-1,1-a}\right)} \right)$$

และ $$p_U = \left(\frac{c^2_{k-1,1-a} + 2n_i + \sqrt{c^2_{k-1,1-a}\left(c^2_{k-1,1-a} + 4\frac{n_i}{n}(n-n_i) \right)}}{2\left(n + c^2_{k-1,1-a}\right)} \right)$$

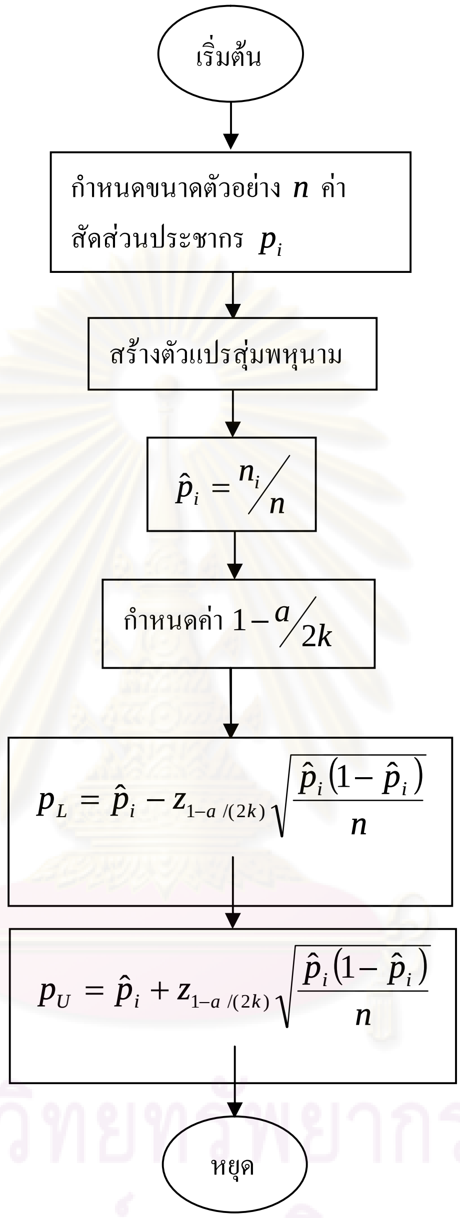
ผังงานสำหรับการคำนวณค่าประมาณแบบช่วงจากการประมาณแบบปกติ แสดงในรูปที่ 3.3

3.2.2.3 วิธีการประมาณแบบเอฟ มีสูตรการประมาณค่าช่วงความเชื่อมั่น คือ

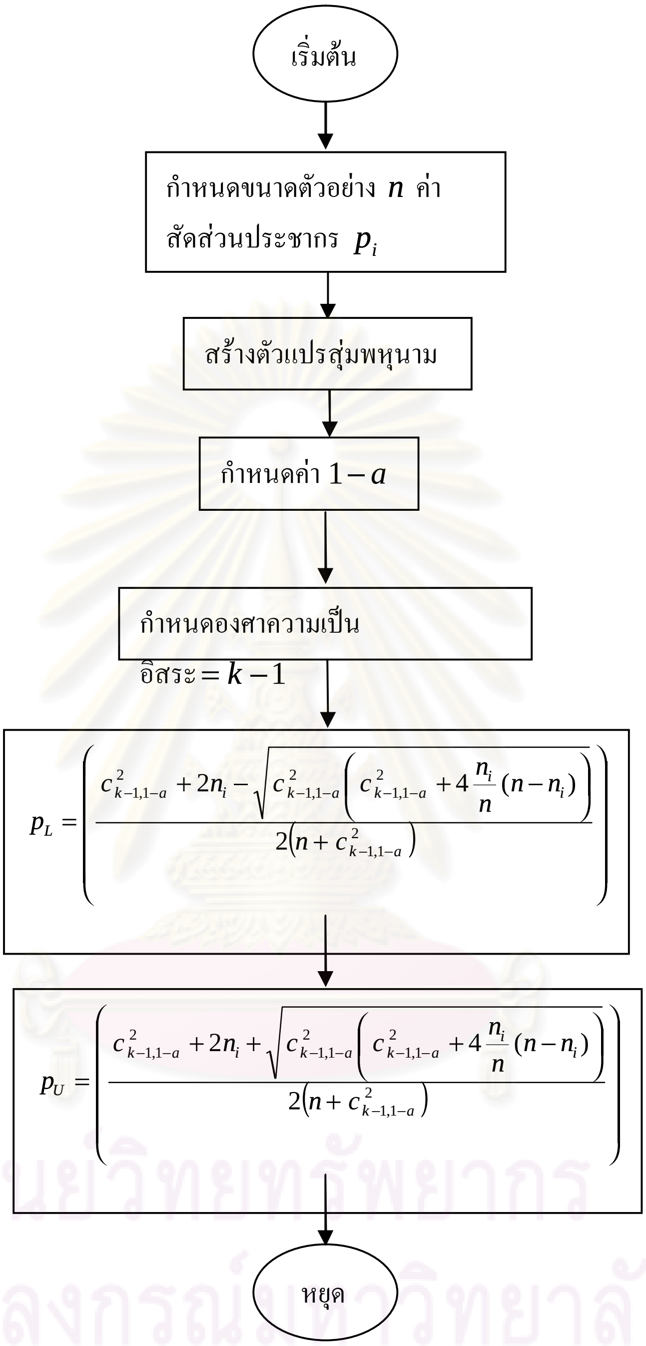
$$p_L = \frac{n_i}{n_i + (n - n_i + 1)F_{(2(n-n_i+1),2n_i),1-\frac{a}{2k}}}$$

และ $$p_U = \frac{(n_i + 1)F_{(2(n_i+1),2(n-n_i)),1-\frac{a}{2k}}}{n - n_i + (n_i + 1)F_{(2(n_i+1),2(n-n_i)),1-\frac{a}{2k}}}$$

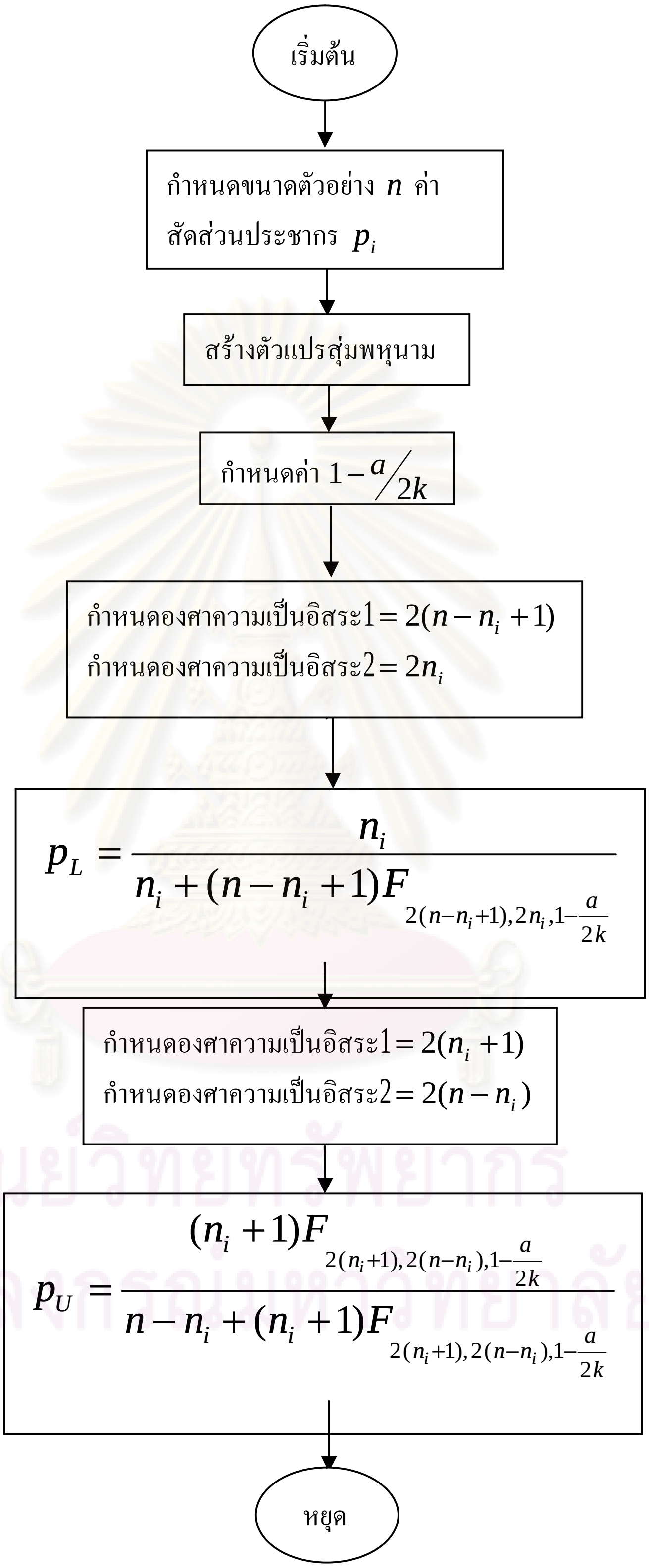
ผังงานสำหรับการคำนวณค่าประมาณแบบช่วงจากการประมาณแบบปกติ แสดงในรูปที่ 3.4



<u>รูปที่ 3.2 แสดงแผนผังการคำนวณค่าประมาณแบบช่วงจากวิธีการประมาณแบบปกติ</u>



รูปที่ 3.3 แสดงแผนผังการคำนวณค่าประมาณแบบช่วงจากวิธีการประมาณแบบคิวเซนเบอร์รี่และเฮิร์ส



<u>รูปที่ 3.4 แสดงแผนผังการคำนวณค่าประมาณแบบช่วงจากวิธีการประมาณแบบเอฟ</u>

3.2.3 การคำนวณค่าสัมประสิทธิ์ช่วงความเชื่อมั่นที่ได้จากการทดลอง

เมื่อทำการคำนวณค่าประมาณแบบช่วงของแต่ละกลุ่มและแต่ละวิธีการแล้ว ขั้นต่อไปจะคำนวณสัมประสิทธิ์ช่วงความเชื่อมั่นที่ได้จากการทดลอง เพื่อทำการตรวจสอบว่าค่าประมาณแบบช่วงแต่ละตัวที่คำนวณได้ครอบคลุมค่าพารามิเตอร์ p_i หรือไม่ โดยจะทำการนับสะสมจำนวนครั้งที่ค่าประมาณแบบช่วงครอบคลุมค่าพารามิเตอร์จากการคำนวณซ้ำ 2000 ครั้งในแต่ละค่าสัดส่วนประชากร นำผลบวกในแต่ละค่าสัดส่วนประชากรที่ครอบคลุมหารด้วย 2000 แล้วจึงนำค่าสัดส่วนที่ได้มาคำนวณหาสัมประสิทธิ์ช่วงความเชื่อมั่นที่ได้จากการทดลองตามวิธีการประมาณนั้นๆ

3.2.4 การคำนวณค่าระยะห่างของค่าบริเวณความเชื่อมั่น

เมื่อคำนวณหาค่าสัมประสิทธิ์ช่วงความเชื่อมั่นจากการทดลองแล้ว ขั้นต่อไปจะทำการคำนวณหาค่าระยะห่างของค่าบริเวณความเชื่อมั่น โดยจะหาผลต่างระหว่างขีดจำกัดความเชื่อมั่นบนและขีดจำกัดความเชื่อมั่นล่างของแต่ละสัดส่วนประชากร คำนวณหาค่าเมทริกซ์ความแปรปรวนร่วมของการแจกแจงพหุนาม และหาระยะห่างดังกล่าวด้วยวิธี Mahalanobis distance ของแต่ละรอบนำระยะห่างของแต่ละรอบมาบวกกัน แล้วจึงหารด้วย 2000 ค่าที่ได้จะเป็นค่าระยะห่างของบริเวณช่วงความเชื่อมั่น

3.2.5 การเปรียบเทียบค่าสัมประสิทธิ์ช่วงความเชื่อมั่นที่ได้จากการทดลองและค่าระยะห่างของค่าบริเวณความเชื่อมั่น

เมื่อได้ค่าสัมประสิทธิ์ช่วงความเชื่อมั่นจากการทดลองแล้ว จะนำค่าดังกล่าวมาเปรียบเทียบกับค่าระดับความเชื่อมั่นที่กำหนด เพื่อตรวจสอบว่าวิธีการประมาณแบบใดให้ค่าระดับความเชื่อมั่นจากการทดลองต่ำกว่าระดับความเชื่อมั่นที่กำหนด ซึ่งจะตรวจสอบอาศัยการทดสอบสมมติฐานโดยใช้ตัวสถิติ Z นั่นคือวิธีการใดให้ค่าระดับความเชื่อมั่นไม่ต่ำกว่า 0.8890, 0.9420 และ 0.9864 ที่ระดับความเชื่อมั่น 90%, 95% และ 99% ตามลำดับ จะถือว่าวิธีการนั้นให้ค่าระดับความเชื่อมั่นจากการทดลองไม่ต่ำกว่าระดับความเชื่อมั่นที่กำหนด

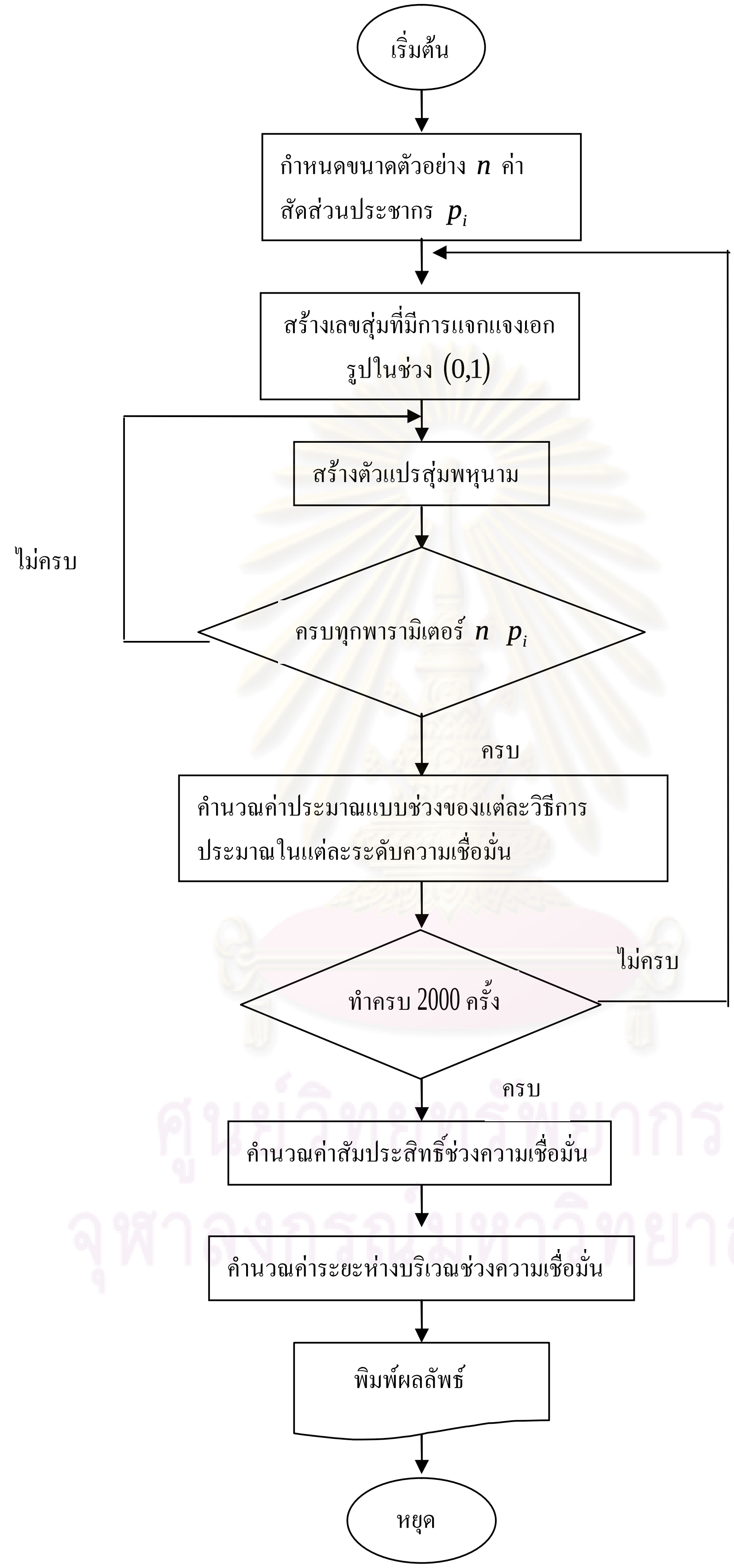
การเปรียบเทียบระยะห่างของค่าบริเวณความเชื่อมั่น จะทำการเปรียบเทียบว่าวิธีการใดให้ค่าระยะห่างบริเวณความเชื่อมั่นต่ำสุดสำหรับแต่ละสถานการณ์ทดลอง

3.2.6 การสรุปผลการวิจัยในแต่ละสถานการณ์

เมื่อทำการเปรียบเทียบค่าสัมประสิทธิ์ความเชื่อมั่นและระยะห่างของบริเวณช่วงความเชื่อมั่นของตัวประมาณแบบช่วงสำหรับแต่ละสถานการณ์แล้ว จะทำการสรุปผลการทดลองว่าวิธีการประมาณใดที่เหมาะสมกับการประมาณในสถานการณ์นั้นๆ

3.3 ขั้นตอนการทำงานของโปรแกรม

ขั้นตอนการทำงานของโปรแกรมการคำนวณสัมประสิทธิ์ช่วงความเชื่อมั่น และระยะห่างบริเวณช่วงความเชื่อมั่น ที่ประมาณโดยวิธีการประมาณทั้ง 3 วิธีสามารถสรุปเป็นผังงานได้ดังนี้



<u>รูปที่ 3.5 แสดงผังงานการคำนวณค่าสัมประสิทธิ์ช่วงความเชื่อมั่น และค่าระยะห่างบริเวณช่วงความเชื่อมั่น</u>

บทที่ 4

ผลการวิจัย

การวิจัยนี้ต้องการเปรียบเทียบวิธีการประมาณแบบช่วงสำหรับค่าสัดส่วนประชากรของการแจกแจงแบบพหุนามทั้ง 3 วิธี ซึ่งระกอบด้วย วิธีการประมาณแบบปกติ วิธีการประมาณแบบคิวเซนเบอร์รี่และเฮิร์ส และวิธีการประมาณแบบเอฟ เพื่อหาว่าวิธีการใดเหมาะสมสำหรับการประมาณค่าในแต่ละสถานการณ์ โดยจะใช้หลักระดับความเชื่อมั่นในการพิจารณา กล่าวคือ ถ้าวิธีการใดให้ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นเฉลี่ยไม่ต่ำกว่าระดับความเชื่อมั่นที่กำหนดและมีระยะห่างของค่าบริเวณช่วงความเชื่อมั่นเฉลี่ยต่ำสุด จะถือว่าวิธีการประมาณดังกล่าวเป็นวิธีการประมาณที่เหมาะสมกับสถานการณ์นั้น

การนำเสนอผลงานจะแบ่งออกเป็น 2 ตอน เพื่อทำการเปรียบเทียบ จะประกอบด้วย

ตอนที่ 1เปรียบเทียบค่าสัมประสิทธิ์บริเวณความเชื่อมั่น

ตอนที่ 2เปรียบเทียบระยะห่างของค่าบริเวณความเชื่อมั่น

ในการอธิบายจะกำหนดลักษณะต่างๆ ดังนี้

วิธีที่ 1 หมายถึงวิธีการประมาณแบบปกติ

วิธีที่ 2 หมายถึงวิธีการประมาณแบบคิวเซนเบอร์รี่ และเฮิร์ส

วิธีที่ 3 หมายถึงวิธีการประมาณแบบเอฟ

p เท่า หมายถึงค่าสัดส่วนประชากรที่มีการกระจายแบบเท่ากัน

p สม หมายถึงค่าสัดส่วนประชากรที่มีการกระจายแบบสมมาตร

p เบ้ หมายถึงค่าสัดส่วนประชากรที่มีการกระจายแบบเบ้

n หมายถึงขนาดตัวอย่าง

k หมายถึงจำนวนค่าสัดส่วนของประชากร

4.1 การเปรียบเทียบค่าสัมประสิทธิ์บริเวณความเชื่อมั่น

จะทำการเปรียบเทียบค่าสัมประสิทธิ์บริเวณความเชื่อมั่นจากการทดลองที่ได้จากวิธีการประมาณทั้ง 3 วิธี ณ ระดับความเชื่อมั่นที่กำหนด ซึ่งได้จากการทดสอบสมมติฐานที่ระดับความเชื่อมั่น 95% โดยมีรายละเอียด ดังนี้

ณ ระดับความเชื่อมั่น 90%

ถ้าวิธีการประมาณแบบใดให้ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นจากการทดลองไม่ต่ำกว่า 0.8890 จะถือว่าวิธีการนั้นให้ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นไม่ต่ำกว่าเกณฑ์ที่กำหนด

ณ ระดับความเชื่อมั่น 95%

ถ้าวิธีการประมาณแบบใดให้ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นจากการทดลองไม่ต่ำกว่า 0.9420 จะถือว่าวิธีการนั้นให้ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นไม่ต่ำกว่าเกณฑ์ที่กำหนด

ณ ระดับความเชื่อมั่น 99%

ถ้าวิธีการประมาณแบบใดให้ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นจากการทดลองไม่ต่ำกว่า 0.9863 จะถือว่าวิธีการนั้นให้ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นไม่ต่ำกว่าเกณฑ์ที่กำหนด

สำหรับการเสนอการเปรียบเทียบมีรายละเอียดการเสนอ ดังนี้

ณ ระดับความเชื่อมั่น 90% นำเสนอในตารางที่ 4.1-4.5
ณ ระดับความเชื่อมั่น 95% นำเสนอในตารางที่ 4.7-4.11
ณ ระดับความเชื่อมั่น 99% นำเสนอในตารางที่ 4.13-4.17

<u>ตอนที่ 2 การเปรียบเทียบค่าระยะห่างบริเวณความเชื่อมั่น</u>

การเปรียบเทียบระยะห่างบริเวณความเชื่อมั่นเฉลี่ย จะทำการเปรียบเทียบเฉพาะสถานการณ์ซึ่งวิธีประมาณให้ค่าไม่ต่ำกว่าระดับความเชื่อมั่นที่กำหนด เพื่อหาวิธีประมาณที่ให้ระยะห่างบริเวณความเชื่อมั่นเฉลี่ยต่ำที่สุด

การเสนอการเปรียบเทียบมีรายละเอียดการเสนอ ดังนี้

ณ ระดับความเชื่อมั่น 90% นำเสนอในตารางที่ 4.19-4.23
ณ ระดับความเชื่อมั่น 95% นำเสนอในตารางที่ 4.24-4.28
ณ ระดับความเชื่อมั่น 99% นำเสนอในตารางที่ 4.29-4.33

ตารางที่ 4.1 ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นเฉลี่ย ณ ระดับความเชื่อมั่น 90% และขนาดตัวอย่าง 50

P	วิธีที่ 1	วิธีที่ 2	วิธีที่ 3
P เท่า** , K=4	0.8098	0.9640*	0.9461*
P สม** , K=4	0.8251	0.9495*	0.9476*
P เบ้** , K=4	0.8345	0.9620*	0.9476*
P เท่า** , K=6	0.8529	0.9865*	0.9325*
P สม** , K=6	0.8046	0.9865*	0.9481*
P เบ้** , K=6	0.8089	0.9780*	0.9611*
P เท่า** , K=8	0.7102	0.9890*	0.9714*
P สม** , K=8	0.8121	0.9855*	0.9520*
P เบ้** , K=8	0.7804	0.9915*	0.9655*
P เท่า**,K=10	0.7287	0.9980*	0.9709*
P สม**, K=10	0.8261	0.9950*	0.9762*
P เบ้**, K=10	0.7207	0.9880*	0.9602*

ตารางที่ 4.2 ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นเฉลี่ย ณ ระดับความเชื่อมั่น 90% และขนาดตัวอย่าง 100

P	วิธีที่ 1	วิธีที่ 2	วิธีที่ 3
P เท่า** , K=4	0.8848	0.9520*	0.9256*
P สม** , K=4	0.8995	0.9505*	0.9219*
P เบ้** , K=4	0.8738	0.9650*	0.9328*
P เท่า** , K=6	0.8773	0.9830*	0.9496*
P สม** , K=6	0.8468	0.9810*	0.9491*
P เบ้** , K=6	0.8507	0.9845*	0.9357*
P เท่า** , K=8	0.7992	0.9910*	0.9312*
P สม** , K=8	0.7924	0.9885*	0.9292*
P เบ้** , K=8	0.7902	0.9920*	0.9439*
P เท่า**, K=10	0.7727	0.9950*	0.9331*
P สม**, K=10	0.8232	0.9965*	0.9607*
P เบ้**, K=10	0.7206	0.9905*	0.9368*

ตารางที่ 4.3 ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นเฉลี่ย ณ ระดับความเชื่อมั่น 90% และขนาดตัวอย่าง 200

P	วิธีที่ 1	วิธีที่ 2	วิธีที่ 3
P เท่า** , K=4	0.8889	0.9600*	0.9158*
P สม** , K=4	0.8949*	0.9575*	0.9361*
P เบ้** , K=4	0.8879	0.9575*	0.9214*
P เท่า** , K=6	0.8818	0.9880*	0.9264*
P สม** , K=6	0.8863	0.9925*	0.9367*
P เบ้** , K=6	0.8735	0.9820*	0.9334*
P เท่า** , K=8	0.8499	0.9950*	0.9246*
P สม** , K=8	0.8502	0.9945*	0.9396*
P เบ้** , K=8	0.7982	0.9915*	0.9363*
P เท่า**, K=10	0.8191	0.9980*	0.9322*
P สม**, K=10	0.8252	0.9975*	0.9411*
P เบ้**, K=10	0.7465	0.9955*	0.9335*

* ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นที่ไม่ต่ำกว่าเกณฑ์

** P เท่า หมายถึง ค่าสัดส่วนประชากรที่มีการกระจายเท่ากัน

P สม หมายถึง ค่าสัดส่วนประชากรที่มีการกระจายสมมาตร

P เบ้ หมายถึง ค่าสัดส่วนประชากรที่มีการกระจายเบ้

37

ตารางที่ 4.4 ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นเฉลี่ย ณ ระดับความเชื่อมั่น 90% และขนาดตัวอย่าง 500

P	วิธีที่ 1	วิธีที่ 2	วิธีที่ 3
P เท่า** , K=4	0.9046*	0.9630*	0.9158*
P สม** , K=4	0.9000*	0.9605*	0.9158*
P เบ้** , K=4	0.9023*	0.9585*	0.9224*
P เท่า** , K=6	0.8854	0.9845*	0.9151*
P สม** , K=6	0.8768	0.9820*	0.9217*
P เบ้** , K=6	0.8958*	0.9915*	0.9226*
P เท่า** , K=8	0.8745	0.9955*	0.9190*
P สม** , K=8	0.8793	0.9980*	0.9209*
P เบ้** , K=8	0.8524	0.9925*	0.9111*
P เท่า** , K=10	0.8632	0.9965*	0.9108*
P สม**, K=10	0.8431	0.9965*	0.9237*
P เบ้** , K=10	0.8488	0.9985*	0.9256*

ตารางที่ 4.5 ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นเฉลี่ย ณ ระดับความเชื่อมั่น 90% และขนาดตัวอย่าง 1000

P	วิธีที่ 1	วิธีที่ 2	วิธีที่ 3
P เท่า** , K=4	0.8972*	0.9540*	0.9055*
P สม** , K=4	0.9037*	0.9620*	0.9153*
P เบ้** , K=4	0.8972*	0.9535*	0.9102*
P เท่า** , K=6	0.8863	0.9855*	0.9064*
P สม** , K=6	0.8972*	0.9890*	0.9147*
P เบ้** , K=6	0.8963*	0.9925*	0.9221*
P เท่า** , K=8	0.8870	0.9955*	0.9172*
P สม** , K=8	0.8820	0.9960*	0.9038*
P เบ้** , K=8	0.8915*	0.9980*	0.9279*
P เท่า** , K=10	0.8908*	0.9985*	0.9182*
P สม** , K=10	0.8601	0.9985*	0.9209*
P เบ้** , K=10	0.8825	0.9990*	0.9312*

* ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นที่ไม่ต่ำกว่าเกณฑ์

** P เท่า หมายถึง ค่าสัดส่วนประชากรที่มีการกระจายเท่ากัน

P สม หมายถึง ค่าสัดส่วนประชากรที่มีการกระจายสมมาตร

P เบ้ หมายถึง ค่าสัดส่วนประชากรที่มีการกระจายเบ้

38

จากตารางที่ 4.1-4.5 ได้แสดงการเปรียบเทียบค่าสัมประสิทธิ์บริเวณความเชื่อมั่นเฉลี่ยจากการทดลองซึ่งได้จากวิธีการประมาณทั้ง 3 วิธี ณ ระดับความเชื่อมั่น 90% เมื่อกำหนดขนาดตัวอย่างเท่ากับ 50 100 200 500 และ1000 ค่าสัดส่วนประชากรเท่ากับ 4 6 8 และ 10 กลุ่ม โดยแบ่งเป็นกลุ่มละ 3 ประเภทตามการกระจายของค่าสัดส่วน คือ เท่ากัน เบ้ และสมมาตร ผลการทดลองสรุปได้ดังนี้

<u>1. วิธีการประมาณแบบปกติ (วิธีที่ 1)</u> ให้ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นเฉลี่ยที่ **<u>ไม่ต่ำกว่า</u>** ระดับความเชื่อมั่นที่กำหนด

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 200 และค่าสัดส่วนประชากร 4 กลุ่มที่มีการกระจายแบบสมมาตร

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 500 และค่าสัดส่วนประชากร 4 กลุ่มที่มีการกระจายแบบเท่า

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 500 และค่าสัดส่วนประชากร 4 กลุ่มที่มีการกระจายแบบเบ้

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 500 และค่าสัดส่วนประชากร 4 กลุ่มที่มีการกระจายแบบสมมาตร

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 500 และค่าสัดส่วนประชากร 6 กลุ่มที่มีการกระจายแบบเบ้

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 1000 และค่าสัดส่วนประชากร 4 กลุ่มที่มีการกระจายแบบเท่า

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 1000 และค่าสัดส่วนประชากร 4 กลุ่มที่มีการกระจายแบบเบ้

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 1000 และค่าสัดส่วนประชากร 4 กลุ่มที่มีการกระจายแบสมมาตร

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 1000 และค่าสัดส่วนประชากร 6 กลุ่มที่มีการกระจายแบบเบ้

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 1000 และค่าสัดส่วนประชากร 6 กลุ่มที่มีการกระจายแบบสมมาตร

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 1000 และค่าสัดส่วนประชากร 8 กลุ่มที่มีการกระจายแบบเบ้

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 1000 และค่าสัดส่วนประชากร 10 กลุ่มที่มีการกระจายแบบเท่า

จากการทดลองได้พบว่าเมื่อจำนวนขนาดตัวอย่าง (n) เพิ่มขึ้น ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นเฉลี่ยจะมีแนวโน้มเพิ่มขึ้น เพราะเมื่อขนาดประชากรตัวอย่างมาก และความแปรปรวนจำกัดประชากรจะลู่เข้าสู่การแจกแจงแบบปกติจึงทำให้การประมาณจะถูกต้องมากยิ่งขึ้น แต่เมื่อจำนวนกลุ่มสัดส่วนของประชากร (k) มากขึ้น ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นเฉลี่ยจะมีแนวโน้มลดลงเพราะเมื่อมีกลุ่มสัดส่วนประชากรเพิ่มมากขึ้นจะทำให้ประชากรตัวอย่างในแต่ละกลุ่มลดลงจึงไม่ลู่

เข้าสู่การแจกแจงแบบปกติ ดังนั้นค่าสัมประสิทธิ์บริเวณความเชื่อมั่นเฉลี่ยจึงลดลง วิธีนี้จึงไม่เหมาะสมเมื่อจำนวนกลุ่มสัดส่วนของประชากรมีค่ามาก

2. วิธีการประมาณแบบคิวเซนเบอร์รี่ และ เฮิร์ส (วิธีที่ 2) ให้ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นเฉลี่ยที่ **ไม่ต่ำกว่า** ระดับความเชื่อมั่นที่กำหนดทุกสถานการณ์การทดลอง

จากการทดลองที่ได้พบว่าเมื่อจำนวนกลุ่มสัดส่วนของประชากร (k) มากขึ้น ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นเฉลี่ยจะมีแนวโน้มเพิ่มขึ้น เนื่องจากวิธีการประมาณแบบคิวเซนเบอร์รี่และเฮิร์สเป็นวิธีการประมาณแบบสถิติที่ไม่ใช้พารามิเตอร์ (non-parametric) จะมีการแจกแจงแบบไคสแควร์โดยประมาณ เมื่อกลุ่มสัดส่วนประชากรมากขึ้นจะทำให้ค่าองศาความเป็นอิสระมีค่ามากขึ้นส่งผลให้ช่วงความเชื่อมั่นของวิธีดังกล่าวมีช่วงยาวมากขึ้นซึ่งจะครอบคลุมค่าพารามิเตอร์ที่แท้จริง จึงทำให้เมื่อจำนวนกลุ่มตัวอย่างมาก จะมีค่าสัมประสิทธิ์บริเวณความเชื่อมั่นเฉลี่ยเพิ่มขึ้น

3. วิธีการประมาณแบบเอฟ (วิธีที่ 3) ให้ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นเฉลี่ยที่ **ไม่ต่ำกว่า** ระดับความเชื่อมั่นที่กำหนดทุกสถานการณ์การทดลอง

จากการทดลองที่ได้พบว่าเมื่อจำนวนขนาดตัวอย่าง (n) เพิ่มขึ้น ค่าสัมประสิทธิ์บริเวณบริเวณความเชื่อมั่นเฉลี่ยจะมีแนวโน้มลดลง เพราะประชากรที่มีขนาดเล็กประชากรจะมีลักษณะเบ้ซึ่งเหมาะสมกับวิธีการประมาณแบบเอฟ แต่เมื่อขนาดประชากรตัวอย่างมีขนาดมากขึ้น จะทำให้ข้อมูลลู่เข้าการแจกแจงแบบปกติ วิธีนี้จึงไม่เหมาะสมเมื่อจำนวนประชากรมีค่ามาก

ตารางที่ 4.6 สรุปค่าสัมประสิทธิ์บริเวณความเชื่อมั่นเฉลี่ยที่ไม่ต่ำกว่าเกณฑ์ที่กำหนดของวิธีการประมาณแบบปกติ ณ ระดับความเชื่อมั่น 90%

ขนาดตัวอย่าง	ประเภทสัดส่วนประชากร	จำนวนกลุ่มของสัดส่วนประชากร (k)			
		4	6	8	10
50	เท่ากัน	-	-	-	-
	เบ้	-	-	-	-
	สมมาตร	-	-	-	-
100	เท่ากัน	-	-	-	-
	เบ้	-	-	-	-
	สมมาตร	-	-	-	-
200	เท่ากัน	-	-	-	-
	เบ้	-	-	-	-
	สมมาตร	√	-	-	-
500	เท่ากัน	√	-	-	-
	เบ้	√	√	-	-
	สมมาตร	√	-	-	-
1000	เท่ากัน	√	-	-	√
	เบ้	√	√	√	-
	สมมาตร	√	√	-	-

ตารางที่ 4.7 ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นเฉลี่ย ณ ระดับความเชื่อมั่น 95% และขนาดตัวอย่าง 50

P	วิธีที่ 1	วิธีที่ 2	วิธีที่ 3
P เท่า** , K=4	0.9134	0.9885*	0.9811*
P สม** , K=4	0.9252	0.9820*	0.9728*
P เบ้** , K=4	0.9083	0.9760*	0.9708*
P เท่า** , K=6	0.8573	0.9940*	0.9752*
P สม** , K=6	0.8140	0.9910*	0.9792*
P เบ้** , K=6	0.8785	0.9885*	0.9772*
P เท่า** , K=8	0.7157	0.9975*	0.9625*
P สม** , K=8	0.8808	0.9950*	0.9649*
P เบ้** , K=8	0.8170	0.9910*	0.9616*
P เท่า**,K=10	0.7378	0.9975*	0.9920*
P สม**, K=10	0.8248	0.9955*	0.9826*
P เบ้**, K=10	0.8335	0.9945*	0.9876*

ตารางที่ 4.2 ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นเฉลี่ย ณ ระดับความเชื่อมั่น 95% และขนาดตัวอย่าง 100

P	วิธีที่ 1	วิธีที่ 2	วิธีที่ 3
P เท่า** , K=4	0.9256	0.9875*	0.9674*
P สม** , K=4	0.9181	0.9795*	0.9587*
P เบ้** , K=4	0.9209	0.9805*	0.9645*
P เท่า** , K=6	0.8701	0.9980*	0.9674*
P สม** , K=6	0.8857	0.9935*	0.9723*
P เบ้** , K=6	0.8693	0.9895*	0.9709*
P เท่า** , K=8	0.7996	0.9970*	0.9601*
P สม** , K=8	0.8358	0.9965*	0.9538*
P เบ้** , K=8	0.8904	0.9960*	0.9533*
P เท่า**,K=10	0.7727	0.9965*	0.9544*
P สม**, K=10	0.8405	0.9975*	0.9758*
P เบ้**, K=10	0.8598	0.9970*	0.9704*

ตารางที่ 4.9 ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นเฉลี่ย ณ ระดับความเชื่อมั่น 95% และขนาดตัวอย่าง 200

P	วิธีที่ 1	วิธีที่ 2	วิธีที่ 3
P เท่า** , K=4	0.9394	0.9840*	0.9664*
P สม** , K=4	0.9332	0.9830*	0.9645*
P เบ้** , K=4	0.9418	0.9810*	0.9645*
P เท่า** , K=6	0.9410	0.9945*	0.9674*
P สม** , K=6	0.9268	0.9970*	0.9709*
P เบ้** , K=6	0.9059	0.9940*	0.9621*
P เท่า** , K=8	0.9265	0.9970*	0.9509*
P สม** , K=8	0.8855	0.9960*	0.9467*
P เบ้** , K=8	0.9115	0.9990*	0.9563*
P เท่า**,K=10	0.9195	0.9995*	0.9723*
P สม**, K=10	0.8557	0.9995*	0.9709*
P เบ้**, K=10	0.9011	0.9995*	0.9709*

* ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นที่ไม่ต่ำกว่าเกณฑ์

** P เท่า หมายถึง ค่าสัดส่วนประชากรที่มีการกระจายเท่ากัน

P สม หมายถึง ค่าสัดส่วนประชากรที่มีการกระจายสมมาตร

P เบ้ หมายถึง ค่าสัดส่วนประชากรที่มีการกระจายเบ้

42

ตารางที่ 4.10 ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นเฉลี่ย ณ ระดับความเชื่อมั่น 95% และขนาดตัวอย่าง 500

P	วิธีที่ 1	วิธีที่ 2	วิธีที่ 3
P เท่า** , K=4	0.9399	0.9795*	0.9557*
P สม** , K=4	0.9485*	0.9765*	0.9567*
P เบ้** , K=4	0.9466*	0.9840*	0.9625*
P เท่า** , K=6	0.9486*	0.9955*	0.9660*
P สม** , K=6	0.9419	0.9965*	0.9679*
P เบ้** , K=6	0.9277	0.9930*	0.9577*
P เท่า** , K=8	0.9232	0.9985*	0.9361*
P สม** , K=8	0.9255	0.9985*	0.9372*
P เบ้** , K=8	0.9311	0.9975*	0.9418
P เท่า** , K=10	0.9223	0.9995*	0.9704*
P สม** , K=10	0.9209	0.9995*	0.9685*
P เบ้** , K=10	0.9056	0.9995*	0.9655*

ตารางที่ 4.11 ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นเฉลี่ย ณ ระดับความเชื่อมั่น 95% และขนาดตัวอย่าง 1000

P	วิธีที่ 1	วิธีที่ 2	วิธีที่ 3
P เท่า** , K=4	0.9471*	0.9800*	0.9582*
P สม** , K=4	0.9471*	0.9850*	0.9611*
P เบ้** , K=4	0.9480*	0.9795*	0.9577*
P เท่า** , K=6	0.9462*	0.9930*	0.9505*
P สม** , K=6	0.9515*	0.9955*	0.9563*
P เบ้** , K=6	0.9453*	0.9950*	0.9553*
P เท่า** , K=8	0.9363	0.9980*	0.9352
P สม** , K=8	0.9415	0.9980*	0.9457*
P เบ้** , K=8	0.9359	0.9970*	0.9328
P เท่า** , K=10	0.9459*	1*	0.9660*
P สม** , K=10	0.9402	1*	0.9651*
P เบ้** , K=10	0.9311	0.9995*	0.9651*

* ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นที่ไม่ต่ำกว่าเกณฑ์

** P เท่า หมายถึง ค่าสัดส่วนประชากรที่มีการกระจายเท่ากัน

P สม หมายถึง ค่าสัดส่วนประชากรที่มีการกระจายสมมาตร

P เบ้ หมายถึง ค่าสัดส่วนประชากรที่มีการกระจายเบ้

43

จากตารางที่ 4.7-4.11 ได้แสดงการเปรียบเทียบค่าสัมประสิทธิ์บริเวณความเชื่อมั่นเฉลี่ยจากการทดลองซึ่งได้จากวิธีการประมาณทั้ง 3 วิธี ณ ระดับความเชื่อมั่น 95% เมื่อกำหนดขนาดตัวอย่างเท่ากับ 50 100 200 500 และ1000 ค่าสัดส่วนประชากรเท่ากับ 4 6 8 และ 10 กลุ่ม โดยแบ่งเป็นกลุ่มละ 3 ประเภทตามการกระจายของค่าสัดส่วน คือ เท่ากัน เบ้ และสมมาตร ผลการทดลองสรุปได้ดังนี้

<u>1. วิธีการประมาณแบบปกติ (วิธีที่ 1)</u> ให้ค่าสัมประสิทธิ์ช่วงความเชื่อมั่นเฉลี่ยที่ **<u>ไม่ต่ำกว่า</u>** ระดับความเชื่อมั่นที่กำหนด

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 500 ค่าสัดส่วนประชากร 4 กลุ่มที่มีการกระจายแบบเบ้

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 500 ค่าสัดส่วนประชากร 4 กลุ่มที่มีการกระจายแบบสมมาตร

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 500 ค่าสัดส่วนประชากร 6 กลุ่มที่มีการกระจายแบบเท่ากัน

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 1000 ค่าสัดส่วนประชากร 4 กลุ่มที่มีการกระจายแบบเท่ากัน

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 1000 ค่าสัดส่วนประชากร 4 กลุ่มที่มีการกระจายแบบเบ้

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 1000 ค่าสัดส่วนประชากร 4 กลุ่มที่มีการกระจายแบบสมมาตร

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 1000 ค่าสัดส่วนประชากร 6 กลุ่มที่มีการกระจายแบบเท่ากัน

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 1000 ค่าสัดส่วนประชากร 6 กลุ่มที่มีการกระจายแบบเบ้

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 1000 ค่าสัดส่วนประชากร 6 กลุ่มที่มีการกระจายแบบสมมาตร

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 1000 ค่าสัดส่วนประชากร 10 กลุ่มที่มีการกระจายแบบเท่ากัน

จากการทดลองได้พบว่าเมื่อจำนวนขนาดตัวอย่าง (n) เพิ่มขึ้น ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นเฉลี่ยจะมีแนวโน้มเพิ่มขึ้น เพราะเมื่อขนาดประชากรตัวอย่างมาก และความแปรปรวนจำกัดประชากรจะลู่เข้าสู่การแจกแจงแบบปกติจึงทำให้การประมาณจะถูกต้องมากยิ่งขึ้น แต่เมื่อจำนวนกลุ่มสัดส่วนของประชากร (k) มากขึ้น ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นเฉลี่ยจะมีแนวโน้มลดลง เพราะเมื่อมีกลุ่มสัดส่วนประชากรเพิ่มมากขึ้นจะทำให้ประชากรตัวอย่างในแต่ละกลุ่มลดลงจึงไม่ลู่เข้าสู่การแจกแจงแบบปกติ ดังนั้นค่าสัมประสิทธิ์บริเวณความเชื่อมั่นเฉลี่ยจึงลดลง วิธีนี้จึงไม่เหมาะสมเมื่อจำนวนกลุ่มสัดส่วนของประชากรมีค่ามาก

<u>2. วิธีการประมาณแบบคิวเซนเบอร์รี่ และ เฮิร์ส (วิธีที่ 2)</u> ให้ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นเฉลี่ยที่ **<u>ไม่ต่ำกว่า</u>** ระดับความเชื่อมั่นที่กำหนดทุกสถานการณ์การทดลอง

จากการทดลองที่ได้พบว่าเมื่อจำนวนกลุ่มสัดส่วนของประชากร (k) มากขึ้น ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นเฉลี่ยจะมีแนวโน้มเพิ่มขึ้น เนื่องจากวิธีการประมาณแบบคิวเซนเบอร์รี่และเฮิร์สเป็นวิธีการประมาณแบบสถิติที่ไม่ใช้พารามิเตอร์ (non-parametric) จะมีการแจกแจงแบบไคสแควร์โดยประมาณ เมื่อกลุ่มสัดส่วนประชากรมากขึ้นจะทำให้ค่าองศาความเป็นอิสระมีค่ามากขึ้นส่งผลให้ช่วงความเชื่อมั่นของวิธีดังกล่าวมีช่วงยาวมากขึ้นซึ่งจะครอบคลุมค่าพารามิเตอร์ที่แท้จริง จึงทำให้เมื่อจำนวนกลุ่มตัวอย่างมาก จะมีค่าสัมประสิทธิ์บริเวณความเชื่อมั่นเฉลี่ยเพิ่มขึ้น

<u>3. วิธีการประมาณแบบเอฟ (วิธีที่ 3)</u> ให้ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นเฉลี่ยที่ **<u>ไม่ต่ำกว่า</u>** ระดับความเชื่อมั่นที่กำหนดทุกสถานการณ์การทดลอง

จากการทดลองที่ได้พบว่าเมื่อจำนวนขนาดตัวอย่าง (n) เพิ่มขึ้น ค่าสัมประสิทธิ์บริเวณบริเวณความเชื่อมั่นเฉลี่ยจะมีแนวโน้มลดลง เพราะประชากรที่มีขนาดเล็กประชากรจะมีลักษณะเบ้ซึ่งเหมาะสมกับวิธีการประมาณแบบเอฟ แต่เมื่อขนาดประชากรตัวอย่างมีขนาดมากขึ้น จะทำให้ข้อมูลลู่เข้าการแจกแจงแบบปกติ วิธีนี้จึงไม่เหมาะสมเมื่อจำนวนประชากรมีค่ามาก

ตารางที่ 4.12 สรุปค่าสัมประสิทธิ์บริเวณความเชื่อมั่นเฉลี่ยที่ไม่ต่ำกว่าเกณฑ์ที่กำหนดของวิธีการประมาณแบบปกติ ณ ระดับความเชื่อมั่น 95%

ขนาดตัวอย่าง	ประเภทสัดส่วนประชากร	จำนวนกลุ่มของสัดส่วนประชากร (k)			
		4	6	8	10
50	เท่ากัน	-	-	-	-
	เบ้	-	-	-	-
	สมมาตร	-	-	-	-
100	เท่ากัน	-	-	-	-
	เบ้	-	-	-	-
	สมมาตร	-	-	-	-
200	เท่ากัน	-	-	-	-
	เบ้	-	-	-	-
	สมมาตร	-	-	-	-
500	เท่ากัน	-	√	-	-
	เบ้	√	-	-	-
	สมมาตร	√	-	-	-
1000	เท่ากัน	√	√	-	√
	เบ้	√	√	-	-
	สมมาตร	√	√	-	-

ตารางที่ 4.13 ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นเฉลี่ย ณ ระดับความเชื่อมั่น 99% และขนาดตัวอย่าง 50

P	วิธีที่ 1	วิธีที่ 2	วิธีที่ 3
P เท่า** , K=4	0.9757	0.9955*	0.9955*
P สม** , K=4	0.9392	0.9950*	0.9925*
P เบ้** , K=4	0.9402	0.9960*	0.9955*
P เท่า** , K=6	0.8719	0.9985*	0.9940*
P สม** , K=6	0.8252	0.9965*	0.9960*
P เบ้** , K=6	0.9153	0.9985*	0.9965*
P เท่า** , K=8	0.9288	1*	0.9980*
P สม** , K=8	0.8955	0.9995*	0.9965*
P เบ้** , K=8	0.9251	0.9985*	0.9965*
P เท่า**,K=10	0.7339	1*	0.9980*
P สม**, K=10	0.8729	0.9985*	0.9965*
P เบ้**, K=10	0.8733	0.9990*	0.9970*

ตารางที่ 4.14 ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นเฉลี่ย ณ ระดับความเชื่อมั่น 99% และขนาดตัวอย่าง 100

P	วิธีที่ 1	วิธีที่ 2	วิธีที่ 3
P เท่า** , K=4	0.9777	0.9955*	0.9940*
P สม** , K=4	0.9723	0.9975*	0.9950*
P เบ้** , K=4	0.9673	0.9985*	0.9960*
P เท่า** , K=6	0.9486	1*	0.9955*
P สม** , K=6	0.9529	0.9995*	0.9940*
P เบ้** , K=6	0.9273	0.9995*	0.9935*
P เท่า** , K=8	0.9167	1*	0.9935*
P สม** , K=8	0.8991	0.9985*	0.9950*
P เบ้** , K=8	0.9343	0.9990*	0.9960*
P เท่า**,K=10	0.9233	0.9995*	0.9950*
P สม**, K=10	0.8936	0.9995*	0.9960*
P เบ้**, K=10	0.9142	1*	0.9955*

ตารางที่ 4.15 ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นเฉลี่ย ณ ระดับความเชื่อมั่น 99% และขนาดตัวอย่าง 200

P	วิธีที่ 1	วิธีที่ 2	วิธีที่ 3
P เท่า** , K=4	0.9851	0.9955*	0.9900*
P สม** , K=4	0.9811	0.9970*	0.9900*
P เบ้** , K=4	0.9821	0.9965*	0.9920*
P เท่า** , K=6	0.9811	0.9990*	0.9915*
P สม** , K=6	0.9772	0.9990*	0.9945*
P เบ้** , K=6	0.9708	0.9990*	0.9940*
P เท่า** , K=8	0.9694	0.9995*	0.9965*
P สม** , K=8	0.9583	1*	0.9886*
P เบ้** , K=8	0.9462	1*	0.9950*
P เท่า**,K=10	0.9588	1*	0.9910*
P สม**, K=10	0.9078	1*	0.9940*
P เบ้**, K=10	0.9329	1*	0.9945*

* ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นที่ไม่ต่ำกว่าเกณฑ์

** P เท่า หมายถึง ค่าสัดส่วนประชากรที่มีการกระจายเท่ากัน

P สม หมายถึง ค่าสัดส่วนประชากรที่มีการกระจายสมมาตร

P เบ้ หมายถึง ค่าสัดส่วนประชากรที่มีการกระจายเบ้

ตารางที่ 4.16 ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นเฉลี่ย ณ ระดับความเชื่อมั่น 99% และขนาดตัวอย่าง 500

P	วิธีที่ 1	วิธีที่ 2	วิธีที่ 3
P เท่า** , K=4	0.9836	0.9950*	0.9885*
P สม** , K=4	0.9866*	0.9975*	0.9905*
P เบ้** , K=4	0.9895*	0.9965*	0.9910*
P เท่า** , K=6	0.9876*	1*	0.9915*
P สม** , K=6	0.9856	0.9995*	0.9945*
P เบ้** , K=6	0.9802	0.9990*	0.9900*
P เท่า** , K=8	0.9821	1*	0.9930*
P สม** , K=8	0.9816	0.9995*	0.9935*
P เบ้** , K=8	0.9772	1*	0.9925*
P เท่า**,K=10	0.9772	1*	0.9905*
P สม**, K=10	0.9684	1*	0.9910*
P เบ้**, K=10	0.9587	1*	0.9935*

ตารางที่ 4.17 ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นเฉลี่ย ณ ระดับความเชื่อมั่น 99% และขนาดตัวอย่าง 1000

P	วิธีที่ 1	วิธีที่ 2	วิธีที่ 3
P เท่า** , K=4	0.9920*	0.9980*	0.9930*
P สม** , K=4	0.9866*	0.9965*	0.9890*
P เบ้** , K=4	0.9890*	0.9960*	0.9905*
P เท่า** , K=6	0.9871*	0.9985*	0.9910*
P สม** , K=6	0.9876*	1*	0.9930*
P เบ้** , K=6	0.9915*	0.9995*	0.9945*
P เท่า** , K=8	0.9876*	0.9995*	0.9925*
P สม** , K=8	0.9846	1*	0.9950*
P เบ้** , K=8	0.9856	1*	0.9945*
P เท่า**,K=10	0.9895*	1*	0.9920*
P สม**, K=10	0.9762	1*	0.9945*
P เบ้**, K=10	0.9738	0.9995*	0.9935*

* ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นที่ไม่ต่ำกว่าเกณฑ์

** P เท่า หมายถึง ค่าสัดส่วนประชากรที่มีการกระจายเท่ากัน

P สม หมายถึง ค่าสัดส่วนประชากรที่มีการกระจายสมมาตร

P เบ้ หมายถึง ค่าสัดส่วนประชากรที่มีการกระจายเบ้

จากตารางที่ 4.13-4.17 ได้แสดงการเปรียบเทียบค่าสัมประสิทธิ์บริเวณความเชื่อมั่นเฉลี่ยจากการทดลองซึ่งได้จากวิธีการประมาณทั้ง 3 วิธี ณ ระดับความเชื่อมั่น 99% เมื่อกำหนดขนาดตัวอย่างเท่ากับ 50 100 200 500 และ1000 ค่าสัดส่วนประชากรเท่ากับ 4 6 8 และ 10 กลุ่ม โดยแบ่งเป็นกลุ่มละ 3 ประเภทตามการกระจายของค่าสัดส่วน คือ เท่ากัน เบ้ และสมมาตร ผลการทดลองสรุปได้ดังนี้

<u>1. วิธีการประมาณแบบปกติ (วิธีที่ 1)</u> ให้ค่าสัมประสิทธิ์ช่วงความเชื่อมั่นเฉลี่ยที่ **<u>ไม่ต่ำกว่า</u>** ระดับความเชื่อมั่นที่กำหนด

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 500 ค่าสัดส่วนประชากร 4 กลุ่มที่มีการกระจายแบบเบ้

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 500 ค่าสัดส่วนประชากร 4 กลุ่มที่มีการกระจายแบบสมมาตร

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 500 ค่าสัดส่วนประชากร 6 กลุ่มที่มีการกระจายแบบเท่ากัน

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 1000 ค่าสัดส่วนประชากร 4 กลุ่มที่มีการกระจายแบบเท่ากัน

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 1000 ค่าสัดส่วนประชากร 4 กลุ่มที่มีการกระจายแบบเบ้

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 1000 ค่าสัดส่วนประชากร 4 กลุ่มที่มีการกระจายแบบสมมาตร

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 1000 ค่าสัดส่วนประชากร 6 กลุ่มที่มีการกระจายแบบเท่ากัน

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 1000 ค่าสัดส่วนประชากร 6 กลุ่มที่มีการกระจายแบบเบ้

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 1000 ค่าสัดส่วนประชากร 6 กลุ่มที่มีการกระจายแบบสมมาตร

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 1000 ค่าสัดส่วนประชากร 8 กลุ่มที่มีการกระจายแบบเท่ากัน

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 1000 ค่าสัดส่วนประชากร 10 กลุ่มที่มีการกระจายแบบเท่ากัน

จากการทดลองได้พบว่าเมื่อจำนวนขนาดตัวอย่าง (n) เพิ่มขึ้น ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นเฉลี่ยจะมีแนวโน้มเพิ่มขึ้น เพราะเมื่อขนาดประชากรตัวอย่างมาก และความแปรปรวนจำกัดประชากรจะลู่เข้าสู่การแจกแจงแบบปกติจึงทำให้การประมาณจะถูกต้องมากยิ่งขึ้น แต่เมื่อจำนวนกลุ่มสัดส่วนของประชากร (k) มากขึ้น ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นเฉลี่ยจะมีแนวโน้มลดลง เพราะเมื่อมีกลุ่มสัดส่วนประชากรเพิ่มมากขึ้นจะทำให้ประชากรตัวอย่างในแต่ละกลุ่มลดลงจึงไม่ลู่เข้าสู่การแจกแจงแบบปกติ ดังนั้นค่าสัมประสิทธิ์บริเวณความเชื่อมั่นเฉลี่ยจึงลดลง วิธีนี้จึงไม่เหมาะสมเมื่อจำนวนกลุ่มสัดส่วนของประชากรมีค่ามาก

2. วิธีการประมาณแบบคิวเซนเบอร์รี่ และ เฮิร์ส (วิธีที่ 2) ให้ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นเฉลี่ยที่ **ไม่ต่ำกว่า** ระดับความเชื่อมั่นที่กำหนดทุกสถานการณ์การทดลอง

จากการทดลองที่ได้พบว่าเมื่อจำนวนกลุ่มสัดส่วนของประชากร (k) มากขึ้น ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นเฉลี่ยจะมีแนวโน้มเพิ่มขึ้น เนื่องจากวิธีการประมาณแบบคิวเซนเบอร์รี่และเฮิร์สเป็นวิธีการประมาณแบบสถิติที่ไม่ใช้พารามิเตอร์ (non-parametric) จะมีการแจกแจงแบบไคสแควร์โดยประมาณ เมื่อกลุ่มสัดส่วนประชากรมากขึ้นจะทำให้ค่าองศาความเป็นอิสระมีค่ามากขึ้นส่งผลให้ช่วงความเชื่อมั่นของวิธีดังกล่าวมีช่วงยาวมากขึ้นซึ่งจะครอบคลุมค่าพารามิเตอร์ที่แท้จริง จึงทำให้เมื่อจำนวนกลุ่มตัวอย่างมาก จะมีค่าสัมประสิทธิ์บริเวณความเชื่อมั่นเฉลี่ยเพิ่มขึ้น

3. วิธีการประมาณแบบเอฟ (วิธีที่ 3) ให้ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นเฉลี่ยที่ **ไม่ต่ำกว่า** ระดับความเชื่อมั่นที่กำหนดทุกสถานการณ์การทดลอง

จากการทดลองที่ได้พบว่าเมื่อจำนวนขนาดตัวอย่าง (n) เพิ่มขึ้น ค่าสัมประสิทธิ์บริเวณบริเวณความเชื่อมั่นเฉลี่ยจะมีแนวโน้มลดลง เพราะประชากรที่มีขนาดเล็กประชากรจะมีลักษณะเบ้ซึ่งเหมาะสมกับวิธีการประมาณแบบเอฟ แต่เมื่อขนาดประชากรตัวอย่างมีขนาดมากขึ้น จะทำให้ข้อมูลลู่เข้าการแจกแจงแบบปกติ วิธีนี้จึงไม่เหมาะสมเมื่อจำนวนประชากรมีค่ามาก

ตารางที่ 4.18 สรุปค่าสัมประสิทธิ์บริเวณความเชื่อมั่นเฉลี่ยที่ไม่ต่ำกว่าเกณฑ์ที่กำหนดของวิธีการประมาณแบบปกติ ณ ระดับความเชื่อมั่น 99%

ขนาดตัวอย่าง	ลักษณะสัดส่วนประชากร	จำนวนกลุ่มของสัดส่วนประชากร (k)			
		4	6	8	10
50	เท่ากัน	-	-	-	-
	เบ้	-	-	-	-
	สมมาตร	-	-	-	-
100	เท่ากัน	-	-	-	-
	เบ้	-	-	-	-
	สมมาตร	-	-	-	-
200	เท่ากัน	-	-	-	-
	เบ้	-	-	-	-
	สมมาตร	-	-	-	-
500	เท่ากัน	-	√	-	-
	เบ้	√	-	-	-
	สมมาตร	√	-	-	-
1000	เท่ากัน	√	√	√	√
	เบ้	√	√	-	-
	สมมาตร	√	√	-	-

<u>ข้อสังเกตที่ได้จากการทดลอง</u>

จากการทดลองที่ได้พบว่า เมื่อจำนวนกลุ่มสัดส่วนประชากรมีค่าน้อย และขนาดตัวอย่างมีค่ามาก วิธีการประมาณแบบปกติ (วิธีที่1) และวิธีการประมาณแบบคิวเซนเบอร์รี่ และเฮิร์สจะให้ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นจากการทดลองที่มีค่าใกล้เคียงกัน ทั้งนี้สามารถอธิบายได้ดังนี้

จากรูปแบบการประมาณแบบช่วงของวิธีประมาณคิวเซนเบอร์รี่และเฮิร์ส

$$\frac{c^2_{k-1,1-a} + 2n_i \pm \sqrt{c^2_{k-1,1-a}\left(c^2_{k-1,1-a} + 4\frac{n_i}{n}(n-n_i)\right)}}{2\left(n + c^2_{k-1,1-a}\right)}$$

เมื่อจำนวนกลุ่มสัดส่วนประชากรมากจะมีผลให้ตัวสถิติ $c^2_{k-1,1-a}$ มีค่าใกล้เคียงกับตัวสถิติ $Z^2_{1-a/2}$ จะได้ว่า

$$\frac{Z^2_{1-a/2} + 2n_i \pm \sqrt{Z^2_{1-a/2}\left(Z^2_{1-a/2} + 4\frac{n_i}{n}(n-n_i)\right)}}{2\left(n + Z^2_{1-a/2}\right)}$$

หรือเท่ากับ

$$\frac{\left(\frac{Z^2_{1-a/2}}{n} + 2\frac{n_i}{n} \pm Z_{1-a/2}\sqrt{\left(\frac{Z^2_{1-a/2}}{n^2} + 4\frac{n_i}{n^3}(n-n_i)\right)}\right)n}{2n\left(1 + \frac{Z^2_{1-a/2}}{n}\right)}$$

หรือเท่ากับ

$$\frac{\left(\frac{Z^2_{1-a/2}}{n} + 2p_i \pm Z_{1-a/2}\sqrt{\left(\frac{Z^2_{1-a/2}}{n^2} + 4\frac{p_i(1-p_i)}{n}\right)}\right)n}{2n\left(1 + \frac{Z^2_{1-a/2}}{n}\right)}$$

เมื่อขนาดตัวอย่างมีค่ามาก จะมีผลให้ $\frac{Z^2_{1-a/2}}{n}, \frac{Z^2_{1-a/2}}{n^2}$ มีค่าน้อยมาก ทำให้ค่าประมาณแบบช่วงจากวิธีการประมาณนี้สามารถเขียนได้เป็น

$$\frac{\left(2p_i \pm Z_{1-a/2}\sqrt{\left(4\frac{p_i(1-p_i)}{n}\right)}\right)n}{2n}$$

ตารางที่ 4.19 ระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ย ณ ระดับความเชื่อมั่น 90% และขนาดตัวอย่าง 50

P	วิธีที่ 1	วิธีที่ 2	วิธีที่ 3
P เท่า** , K=4	-	0.1156	0.1095
P สม** , K=4	-	0.1385	0.1300
P เบ้** , K=4	-	0.1742	0.1604
P เท่า** , K=6	-	0.4602	0.3324
P สม** , K=6	-	0.6414	0.4398
P เบ้** , K=6	-	0.8775	0.5705
P เท่า** , K=8	-	1.2359	0.7187
P สม** , K=8	-	2.0462	1.0692
P เบ้** , K=8	-	3.0177	1.4635
P เท่า**,K=10	-	2.7485	1.3200
P สม**, K=10	-	5.3832	2.2155
P เบ้**, K=10	-	8.4082	3.1953

ตารางที่ 4.20 ระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ย ณ ระดับความเชื่อมั่น 90% และขนาดตัวอย่าง 100

P	วิธีที่ 1	วิธีที่ 2	วิธีที่ 3
P เท่า** , K=4	-	0.0294	0.0265
P สม** , K=4	-	0.0347	0.0312
P เบ้** , K=4	-	0.0430	0.0383
P เท่า** , K=6	-	0.1130	0.0781
P สม** , K=6	-	0.1507	0.1011
P เบ้** , K=6	-	0.1943	0.1255
P เท่า** , K=8	-	0.2908	0.1640
P สม** , K=8	-	0.4414	0.2301
P เบ้** , K=8	-	0.6083	0.2956
P เท่า**,K=10	-	0.6154	0.2907
P สม**, K=10	-	1.0806	0.4455
P เบ้**, K=10	-	1.5791	0.5907

ตารางที่ 4.21 ระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ย ณ ระดับความเชื่อมั่น 90% และขนาดตัวอย่าง 200

P	วิธีที่ 1	วิธีที่ 2	วิธีที่ 3
P เท่า** , K=4	-	0.0074	0.0065
P สม** , K=4	0.0071	0.0087	0.0076
P เบ้** , K=4	-	0.0106	0.0092
P เท่า** , K=6	-	0.0280	0.0188
P สม** , K=6	-	0.0366	0.0242
P เบ้** , K=6	-	0.0455	0.0295
P เท่า** , K=8	-	0.0702	0.0388
P สม** , K=8	-	0.1011	0.0536
P เบ้** , K=8	-	0.1310	0.0662
P เท่า**,K=10	-	0.1438	0.0677
P สม**, K=10	-	0.2315	0.0999
P เบ้**, K=10	-	0.3134	0.1248

** P เท่า หมายถึง ค่าสัดส่วนประชากรที่มีการกระจายเท่ากัน

P สม หมายถึง ค่าสัดส่วนประชากรที่มีการกระจายสมมาตร

P เบ้ หมายถึง ค่าสัดส่วนประชากรที่มีการกระจายเบ้

53

ตารางที่ 4.22 ระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ย ณ ระดับความเชื่อมั่น 90% และขนาดตัวอย่าง 500

P	วิธีที่ 1	วิธีที่ 2	วิธีที่ 3
P เท่า** , K=4	9.6843e-004	0.0012	0.0010
P สม** , K=4	0.0011	0.0014	0.0012
P เบ้** , K=4	0.0014	0.0017	0.0014
P เท่า** , K=6	-	0.0045	0.0029
P สม** , K=6	-	0.0057	0.0037
P เบ้ , K=6	0.0042	0.0070	0.0045
P เท่า** , K=8	-	0.0110	0.0059
P สม** , K=8	-	0.0152	0.0081
P เบ้** , K=8	-	0.0188	0.0099
P เท่า** , K=10	-	0.0219	0.0103
P สม**, K=10	-	0.0329	0.0148
P เบ้** , K=10	-	0.0415	0.0180

ตารางที่ 4.23 ระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ย ณ ระดับความเชื่อมั่น 90% และขนาดตัวอย่าง 1000

P	วิธีที่ 1	วิธีที่ 2	วิธีที่ 3
P เท่า** , K=4	2.4163e-004	2.9942e-004	2.4893e-004
P สม** , K=4	2.8192e-004	3.4994e-004	2.9142e-004
P เบ้** , K=4	3.3912e-004	4.2226e-004	3.5245e-004
P เท่า** , K=6	-	0.0011	7.1370e-004
P สม** , K=6	8.7369e-004	0.0014	9.0988e-004
P เบ้** , K=6	0.0010	0.0017	0.0011
P เท่า** , K=8	-	0.0027	0.0015
P สม** , K=8	-	0.0037	0.0020
P เบ้ , K=8	0.0022	0.0045	0.0024
P เท่า** , K=10	0.0024	0.0054	0.0025
P สม** , K=10	-	0.0078	0.0036
P เบ้** , K=10	-	0.0097	0.0043

** P เท่า หมายถึง ค่าสัดส่วนประชากรที่มีการกระจายเท่ากัน
P สม หมายถึง ค่าสัดส่วนประชากรที่มีการกระจายสมมาตร
P เบ้ หมายถึง ค่าสัดส่วนประชากรที่มีการกระจายเบ้

จากตารางที่ 4.19-4.23 ได้แสดงการเปรียบเทียบค่าระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ยจากวิธีการประมาณทั้ง 3 วิธี ณ ระดับความเชื่อมั่น 90% เมื่อกำหนดขนาดตัวอย่างเท่ากับ 50 100 200 500 และ1000 ค่าสัดส่วนประชากรเท่ากับ 4 6 8 และ 10 กลุ่ม โดยแบ่งเป็นกลุ่มละ 3 ประเภทตามการกระจายของค่าสัดส่วน คือ เท่ากัน เบ้ และสมมาตร ผลการทดลองสรุปได้ดังนี้

1. วิธีการประมาณแบบปกติ (วิธีที่ 1)

วิธีนี้ให้ค่าระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ย **ต่ำสุด** ในทุกระดับการทดลอง และ**ครอบคลุม**ค่าสัดส่วนประชากร

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 200 ค่าสัดส่วนประชากร 4 กลุ่มที่มีการกระจายแบบสมมาตร

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 500 ค่าสัดส่วนประชากร 4 กลุ่มที่มีการกระจายแบบเท่ากัน

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 500 ค่าสัดส่วนประชากร 4 กลุ่มที่มีการกระจายแบบเบ้

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 500 ค่าสัดส่วนประชากร 4 กลุ่มที่มีการกระจายแบบสมมาตร

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 500 ค่าสัดส่วนประชากร 6 กลุ่มที่มีการกระจายแบบเบ้

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 1000 ค่าสัดส่วนประชากร 4 กลุ่มที่มีการกระจายแบบเท่ากัน

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 1000 ค่าสัดส่วนประชากร 4 กลุ่มที่มีการกระจายแบบเบ้

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 1000 ค่าสัดส่วนประชากร 4 กลุ่มที่มีการกระจายแบบสมมาตร

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 1000 ค่าสัดส่วนประชากร 6 กลุ่มที่มีการกระจายแบบเบ้

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 1000 ค่าสัดส่วนประชากร 6 กลุ่มที่มีการกระจายแบบสมมาตร

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 1000 ค่าสัดส่วนประชากร 8 กลุ่มที่มีการกระจายแบบเบ้

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 1000 ค่าสัดส่วนประชากร 10 กลุ่มที่มีการกระจายแบบเท่ากัน

จากการศึกษาที่ได้พบว่าเมื่อขนาดตัวอย่าง (n) เพิ่มขึ้น ค่าระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ยมีแนวโน้มลดลง เพราะเมื่อขนาดตัวอย่างเพิ่มขึ้นจะส่งผลให้ค่าความผิดพลาดจากการประมาณลดลงจึงทำให้มีการกระจายลดลงตามไปด้วย และเมื่อจำนวนกลุ่มของค่าสัดส่วนประชากร (k) เพิ่มมากขึ้นจะมีผลให้ค่าระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ยมีแนวโน้มเพิ่มขึ้น เพราะเมื่อเพิ่มจำนวนพารามิเตอร์ที่ประมาณจะทำให้มีค่าความผิดพลาดจากการประมาณเพิ่มขึ้นจึงมีผลทำให้ค่าระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ยมีแนวโน้มเพิ่มขึ้น

2. วิธีการประมาณแบบคิวเซนเบอร์รี่ และ เฮิร์ส (วิธีที่ 2)

วิธีนี้ให้ค่าระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ย **สูงสุด** ในทุกระดับการทดลอง และครอบคลุมค่าสัดส่วนประชากรในทุกระดับการทดลอง

จากการศึกษาที่ได้พบว่าเมื่อขนาดตัวอย่าง (n) เพิ่มขึ้น ค่าระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ยมีแนวโน้มลดลง เพราะเมื่อขนาดตัวอย่างเพิ่มขึ้นจะส่งผลให้ค่าความผิดพลาดจากการประมาณลดลงจึงทำให้มีการกระจายลดลงตามไปด้วย และเมื่อจำนวนกลุ่มของค่าสัดส่วนประชากร (k) เพิ่มมากขึ้นจะมีผลให้ค่าระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ยมีแนวโน้มเพิ่มขึ้น เพราะเมื่อเพิ่มจำนวนพารามิเตอร์ที่ประมาณจะทำให้มีค่าความผิดพลาดจากการประมาณเพิ่มขึ้นจึงมีผลทำให้ค่าระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ยมีแนวโน้มเพิ่มขึ้น

3. วิธีการประมาณแบบเอฟ (วิธีที่ 3)

วิธีการนี้ไม่สามารถให้ค่าระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ย ต่ำสุด หรือสูงสุดได้ในทุกระดับการทดลอง และครอบคลุมค่าสัดส่วนประชากรในทุกระดับการทดลอง

จากการศึกษาที่ได้พบว่าเมื่อขนาดตัวอย่าง (n) เพิ่มขึ้น ค่าระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ยมีแนวโน้มลดลง เพราะเมื่อขนาดตัวอย่างเพิ่มขึ้นจะส่งผลให้ค่าความผิดพลาดจากการประมาณลดลงจึงทำให้มีการกระจายลดลงตามไปด้วย และเมื่อจำนวนกลุ่มของค่าสัดส่วนประชากร (k) เพิ่มมากขึ้นจะมีผลให้ค่าระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ยมีแนวโน้มเพิ่มขึ้น เพราะเมื่อเพิ่มจำนวนพารามิเตอร์ที่ประมาณจะทำให้มีค่าความผิดพลาดจากการประมาณเพิ่มขึ้นจึงมีผลทำให้ค่าระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ยมีแนวโน้มเพิ่มขึ้น

ตารางที่ 4.24 ระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ย ณ ระดับความเชื่อมั่น 95% และขนาดตัวอย่าง 50

P	วิธีที่ 1	วิธีที่ 2	วิธีที่ 3
P เท่า** , K=4	-	0.1419	0.1328
P สม** , K=4	-	0.1714	0.1581
P เบ้** , K=4	-	0.2174	0.1946
P เท่า** , K=6	-	0.5449	0.3949
P สม** , K=6	-	0.7691	0.5232
P เบ้** , K=6	-	1.0653	0.6809
P เท่า** , K=8	-	1.4383	0.8316
P สม** , K=8	-	2.4239	1.2492
P เบ้** , K=8	-	3.6075	1.7120
P เท่า**,K=10	-	3.1642	1.5425
P สม**, K=10	-	6.3104	2.6109
P เบ้**, K=10	-	9.8838	3.7426

ตารางที่ 4.25 ระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ย ณ ระดับความเชื่อมั่น 95% และขนาดตัวอย่าง 100

P	วิธีที่ 1	วิธีที่ 2	วิธีที่ 3
P เท่า** , K=4	-	0.0364	0.0324
P สม** , K=4	-	0.0432	0.0382
P เบ้** , K=4	-	0.0535	0.0467
P เท่า** , K=6	-	0.1348	0.0935
P สม** , K=6	-	0.1823	0.1218
P เบ้** , K=6	-	0.2360	0.1502
P เท่า** , K=8	-	0.3404	0.1900
P สม** , K=8	-	0.5269	0.2694
P เบ้** , K=8	-	0.7303	0.3430
P เท่า**,K=10	-	0.7126	0.3413
P สม**, K=10	-	1.2762	0.5252
P เบ้**, K=10	-	1.8865	0.6988

ตารางที่ 4.26 ระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ย ณ ระดับความเชื่อมั่น 95% และขนาดตัวอย่าง 200

P	วิธีที่ 1	วิธีที่ 2	วิธีที่ 3
P เท่า** , K=4	-	0.0092	0.0079
P สม** , K=4	-	0.0109	0.0093
P เบ้** , K=4	-	0.0133	0.0114
P เท่า** , K=6	-	0.0335	0.0226
P สม** , K=6	-	0.0440	0.0291
P เบ้** , K=6	-	0.0550	0.0354
P เท่า** , K=8	-	0.0822	0.0450
P สม** , K=8	-	0.1203	0.0626
P เบ้** , K=8	-	0.1576	0.0771
P เท่า**,K=10	-	0.1665	0.0797
P สม**, K=10	-	0.2721	0.1177
P เบ้**, K=10	-	0.3719	0.1469

** P เท่า หมายถึง ค่าสัดส่วนประชากรที่มีการกระจายเท่ากัน

P สม หมายถึง ค่าสัดส่วนประชากรที่มีการกระจายสมมาตร

P เบ้ หมายถึง ค่าสัดส่วนประชากรที่มีการกระจายเบ้

ตารางที่ 4.27 ระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ย ณ ระดับความเชื่อมั่น 95% และขนาดตัวอย่าง 500

P	วิธีที่ 1	วิธีที่ 2	วิธีที่ 3
P เท่า** , K=4	-	0.0015	0.0012
P สม** , K=4	0.0014	0.0017	0.0015
P เบ้** , K=4	0.0017	0.0021	0.0018
P เท่า** , K=6	0.0034	0.0053	0.0035
P สม** , K=6	-	0.0069	0.0045
P เบ้** , K=6	-	0.0084	0.0054
P เท่า** , K=8	-	0.0128	0.0069
P สม** , K=8	-	0.0179	0.0095
P เบ้** , K=8	-	0.0223	0.0114
P เท่า**,K=10	-	0.0253	0.0121
P สม**, K=10	-	0.0383	0.0175
P เบ้**, K=10	-	0.0488	0.0213

ตารางที่ 4.28 ระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ย ณ ระดับความเชื่อมั่น 95% และขนาดตัวอย่าง 1000

P	วิธีที่ 1	วิธีที่ 2	วิธีที่ 3
P เท่า** , K=4	3.0004e-004	3.7390e-004	3.0793e-004
P สม** , K=4	3.5005e-004	4.3714e-004	3.6038e-004
P เบ้** , K=4	4.2085e-004	5.2762e-004	4.3549e-004
P เท่า** , K=6	8.3859e-004	0.0013	8.6362e-004
P สม** , K=6	0.0011	0.0017	0.0011
P เบ้** , K=6	0.0013	0.0021	0.0013
P เท่า** , K=8	-	0.0032	0.0017
P สม** , K=8	-	0.0044	0.0023
P เบ้** , K=8	-	0.0053	0.0027
P เท่า**,K=10	0.0029	0.0062	0.0030
P สม**, K=10	-	0.0091	0.0042
P เบ้**, K=10	-	0.0112	0.0051

** P เท่า หมายถึง ค่าสัดส่วนประชากรที่มีการกระจายเท่ากัน

P สม หมายถึง ค่าสัดส่วนประชากรที่มีการกระจายสมมาตร

P เบ้ หมายถึง ค่าสัดส่วนประชากรที่มีการกระจายเบ้

58

จากตารางที่ 4.24-4.28 ได้แสดงการเปรียบเทียบค่าระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ยจากวิธีการประมาณทั้ง 3 วิธี ณ ระดับความเชื่อมั่น 95% เมื่อกำหนดขนาดตัวอย่างเท่ากับ 50 100 200 500 และ1000 ค่าสัดส่วนประชากรเท่ากับ 4 6 8 และ 10 กลุ่ม โดยแบ่งเป็นกลุ่มละ 3 ประเภทตามการกระจายของค่าสัดส่วน คือ เท่ากัน เบ้ และสมมาตร ผลการทดลองสรุปได้ดังนี้

1. วิธีการประมาณแบบปกติ (วิธีที่ 1)

วิธีนี้ให้ค่าระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ย **ต่ำสุด** ในทุกระดับการทดลอง และ **ครอบคลุม**ค่าสัดส่วนประชากร

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 500 ค่าสัดส่วนประชากร 4 กลุ่มที่มีการกระจายแบบเบ้

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 500 ค่าสัดส่วนประชากร 4 กลุ่มที่มีการกระจายแบบสมมาตร

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 500 ค่าสัดส่วนประชากร 6 กลุ่มที่มีการกระจายแบบเท่ากัน

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 1000 ค่าสัดส่วนประชากร 4 กลุ่มที่มีการกระจายแบบเท่ากัน

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 1000 ค่าสัดส่วนประชากร 4 กลุ่มที่มีการกระจายแบบเบ้

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 1000 ค่าสัดส่วนประชากร 4 กลุ่มที่มีการกระจายแบบสมมาตร

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 1000 ค่าสัดส่วนประชากร 6 กลุ่มที่มีการกระจายแบบเท่ากัน

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 1000 ค่าสัดส่วนประชากร 6 กลุ่มที่มีการกระจายแบบเบ้

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 1000 ค่าสัดส่วนประชากร 6 กลุ่มที่มีการกระจายแบบสมมาตร

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 1000 ค่าสัดส่วนประชากร 10 กลุ่มที่มีการกระจายแบบเท่ากัน

วิธีนี้จากการศึกษาที่ได้พบว่าเมื่อขนาดตัวอย่าง(n)เพิ่มขึ้น ค่าระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ยมีแนวโน้มลดลง เพราะเมื่อขนาดตัวอย่างเพิ่มขึ้นจะส่งผลให้ค่าความผิดพลาดจากการประมาณลดลงจึงทำให้มีการกระจายลดลงตามไปด้วย และเมื่อจำนวนกลุ่มของค่าสัดส่วนประชากร (k)เพิ่มมากขึ้นจะมีผลให้ค่าระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ยมีแนวโน้มเพิ่มขึ้น เพราะเมื่อเพิ่มจำนวนพารามิเตอร์ที่ประมาณจะทำให้มีค่าความผิดพลาดจากการประมาณเพิ่มขึ้นจึงมีผลทำให้ค่าระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ยมีแนวโน้มเพิ่มขึ้น

2. วิธีการประมาณแบบคิวเซนเบอร์รี่ และ เฮิร์ส (วิธีที่ 2)

วิธีนี้ให้ค่าระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ย **สูงสุด** ในทุกระดับการทดลอง และครอบคลุมค่าสัดส่วนประชากรในทุกระดับการทดลอง

จากการศึกษาที่ได้พบว่าเมื่อขนาดตัวอย่าง (n) เพิ่มขึ้น ค่าระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ยมีแนวโน้มลดลง เพราะเมื่อขนาดตัวอย่างเพิ่มขึ้นจะส่งผลให้ค่าความผิดพลาดจากการประมาณลดลงจึงทำให้มีการกระจายลดลงตามไปด้วย และเมื่อจำนวนกลุ่มของค่าสัดส่วนประชากร (k) เพิ่มมากขึ้นจะมีผลให้ค่าระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ยมีแนวโน้มเพิ่มขึ้น เพราะเมื่อเพิ่มจำนวนพารามิเตอร์ที่ประมาณจะทำให้มีค่าความผิดพลาดจากการประมาณเพิ่มขึ้นจึงมีผลทำให้ค่าระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ยมีแนวโน้มเพิ่มขึ้น

3. วิธีการประมาณแบบเอฟ (วิธีที่ 3)

วิธีการนี้ไม่สามารถให้ค่าระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ย ต่ำสุด หรือสูงสุดได้ ในทุกระดับการทดลอง และครอบคลุมค่าสัดส่วนประชากรในทุกระดับการทดลอง

จากการศึกษาที่ได้พบว่าเมื่อขนาดตัวอย่าง (n) เพิ่มขึ้น ค่าระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ยมีแนวโน้มลดลง เพราะเมื่อขนาดตัวอย่างเพิ่มขึ้นจะส่งผลให้ค่าความผิดพลาดจากการประมาณลดลงจึงทำให้มีการกระจายลดลงตามไปด้วย และเมื่อจำนวนกลุ่มของค่าสัดส่วนประชากร (k) เพิ่มมากขึ้นจะมีผลให้ค่าระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ยมีแนวโน้มเพิ่มขึ้น เพราะเมื่อเพิ่มจำนวนพารามิเตอร์ที่ประมาณจะทำให้มีค่าความผิดพลาดจากการประมาณเพิ่มขึ้นจึงมีผลทำให้ค่าระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ยมีแนวโน้มเพิ่มขึ้น

ตารางที่ 4.29 ระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ย ณ ระดับความเชื่อมั่น 99% และขนาดตัวอย่าง 50

P	วิธีที่ 1	วิธีที่ 2	วิธีที่ 3
P เท่า** , K=4	-	0.1979	0.1858
P สม** , K=4	-	0.2405	0.2208
P เบ้** , K=4	-	0.3137	0.2736
P เท่า** , K=6	-	0.7225	0.5388
P สม** , K=6	-	1.0463	0.7202
P เบ้** , K=6	-	1.4771	0.9412
P เท่า** , K=8	-	1.8594	1.1372
P สม** , K=8	-	3.2204	1.7161
P เบ้** , K=8	-	4.8926	2.3880
P เท่า**,K=10	-	4.0199	2.0618
P สม**, K=10	-	8.2441	3.5410
P เบ้**, K=10	-	13.1499	5.1842

ตารางที่ 4.30 ระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ย ณ ระดับความเชื่อมั่น 99% และขนาดตัวอย่าง 100

P	วิธีที่ 1	วิธีที่ 2	วิธีที่ 3
P เท่า** , K=4	-	0.0517	0.0462
P สม** , K=4	-	0.0619	0.0545
P เบ้** , K=4	-	0.0776	0.0667
P เท่า** , K=6	-	0.1814	0.1295
P สม** , K=6	-	0.2502	0.1692
P เบ้** , K=6	-	0.3304	0.2093
P เท่า** , K=8	-	0.4463	0.2643
P สม** , K=8	-	0.7105	0.3752
P เบ้** , K=8	-	1.0068	0.4819
P เท่า**,K=10	-	0.9191	0.4600
P สม**, K=10	-	1.7028	0.7140
P เบ้**, K=10	-	2.5708	0.9600

ตารางที่ 4.31 ระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ย ณ ระดับความเชื่อมั่น 99% และขนาดตัวอย่าง 200

P	วิธีที่ 1	วิธีที่ 2	วิธีที่ 3
P เท่า** , K=4	-	0.0132	0.0114
P สม** , K=4	-	0.0156	0.0134
P เบ้** , K=4	-	0.0193	0.0164
P เท่า** , K=6	-	0.0454	0.0315
P สม** , K=6	-	0.0602	0.0406
P เบ้** , K=6	-	0.0764	0.0495
P เท่า** , K=8	-	0.1082	0.0633
P สม** , K=8	-	0.1617	0.0881
P เบ้** , K=8	-	0.2147	0.1085
P เท่า**,K=10	-	0.2151	0.1081
P สม**, K=10	-	0.3624	0.1605
P เบ้**, K=10	-	0.5090	0.2029

** P เท่า หมายถึง ค่าสัดส่วนประชากรที่มีการกระจายเท่ากัน
P สม หมายถึง ค่าสัดส่วนประชากรที่มีการกระจายสมมาตร
P เบ้ หมายถึง ค่าสัดส่วนประชากรที่มีการกระจายเบ้

61

ตารางที่ 4.32 ระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ย ณ ระดับความเชื่อมั่น 99% และขนาดตัวอย่าง 500

P	วิธีที่ 1	วิธีที่ 2	วิธีที่ 3
P เท่า** , K=4	-	0.0022	0.0018
P สม** , K=4	0.0021	0.0025	0.0021
P เบ้** , K=4	0.0025	0.0031	0.0026
P เท่า** , K=6	0.0048	0.0073	0.0049
P สม** , K=6	-	0.0094	0.0063
P เบ้** , K=6	-	0.0115	0.0076
P เท่า** , K=8	-	0.0169	0.0098
P สม** , K=8	-	0.0238	0.0134
P เบ้** , K=8	-	0.0299	0.0162
P เท่า**,K=10	-	0.0326	0.0165
P สม**, K=10	-	0.0502	0.0239
P เบ้**, K=10	-	0.0651	0.0292

ตารางที่ 4.33 ระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ย ณ ระดับความเชื่อมั่น 99% และขนาดตัวอย่าง 1000

P	วิธีที่ 1	วิธีที่ 2	วิธีที่ 3
P เท่า** , K=4	4.3963e-004	5.4155e-004	4.4836e-004
P สม** , K=4	5.1275e-004	6.3356e-004	5.2444e-004
P เบ้** , K=4	6.1762e-004	7.6742e-004	6.3466e-004
P เท่า** , K=6	0.0012	0.0018	0.0012
P สม** , K=6	0.0015	0.0023	0.0016
P เบ้** , K=6	0.0018	0.0028	0.0019
P เท่า** , K=8	0.0023	0.0042	0.0024
P สม** , K=8	-	0.0058	0.0033
P เบ้** , K=8	-	0.0071	0.0039
P เท่า**,K=10	0.0039	0.0080	0.0041
P สม**, K=10	-	0.0118	0.0058
P เบ้**, K=10	-	0.0147	0.0070

** P เท่า หมายถึง ค่าสัดส่วนประชากรที่มีการกระจายเท่ากัน
P สม หมายถึง ค่าสัดส่วนประชากรที่มีการกระจายสมมาตร
P เบ้ หมายถึง ค่าสัดส่วนประชากรที่มีการกระจายเบ้

62

จากตารางที่ 4.29-4.33 ได้แสดงการเปรียบเทียบค่าระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ยจากวิธีการประมาณทั้ง 3 วิธี ณ ระดับความเชื่อมั่น 99% เมื่อกำหนดขนาดตัวอย่างเท่ากับ 50 100 200 500 และ1000 ค่าสัดส่วนประชากรเท่ากับ 4 6 8 และ 10 กลุ่ม โดยแบ่งเป็นกลุ่มละ 3 ประเภทตามการกระจายของค่าสัดส่วน คือ เท่ากัน เบ้ และสมมาตร ผลการทดลองสรุปได้ดังนี้

1. วิธีการประมาณแบบปกติ (วิธีที่ 1)

วิธีนี้ให้ค่าระยะห่างของค่าบริเวณชความเชื่อมั่นเฉลี่ย **ต่ำสุด** ในทุกระดับการทดลอง และ **ครอบคลุม**ค่าสัดส่วนประชากร

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 500 ค่าสัดส่วนประชากร 4 กลุ่มที่มีการกระจายแบบเบ้

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 500 ค่าสัดส่วนประชากร 4 กลุ่มที่มีการกระจายแบบสมมาตร

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 500 ค่าสัดส่วนประชากร 6 กลุ่มที่มีการกระจายแบบเท่ากัน

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 1000 ค่าสัดส่วนประชากร 4 กลุ่มที่มีการกระจายแบบเท่ากัน

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 1000 ค่าสัดส่วนประชากร 4 กลุ่มที่มีการกระจายแบบเบ้

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 1000 ค่าสัดส่วนประชากร 4 กลุ่มที่มีการกระจายแบบสมมาตร

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 1000 ค่าสัดส่วนประชากร 6 กลุ่มที่มีการกระจายแบบเท่ากัน

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 1000 ค่าสัดส่วนประชากร 6 กลุ่มที่มีการกระจายแบบเบ้

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 1000 ค่าสัดส่วนประชากร 6 กลุ่มที่มีการกระจายแบบสมมาตร

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 1000 ค่าสัดส่วนประชากร 8 กลุ่มที่มีการกระจายแบบเท่ากัน

เมื่อขนาดตัวอย่างมีค่าเท่ากับ 1000 ค่าสัดส่วนประชากร 10 กลุ่มที่มีการกระจายแบบเท่ากัน

จากการศึกษาที่ได้พบว่าเมื่อขนาดตัวอย่าง (n) เพิ่มขึ้น ค่าระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ยมีแนวโน้มลดลง เพราะเมื่อขนาดตัวอย่างเพิ่มขึ้นจะส่งผลให้ค่าความผิดพลาดจากการประมาณลดลงจึงทำให้มีการกระจายลดลงตามไปด้วย และเมื่อจำนวนกลุ่มของค่าสัดส่วนประชากร (k) เพิ่มมากขึ้นจะมีผลให้ค่าระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ยมีแนวโน้มเพิ่มขึ้น เพราะเมื่อเพิ่มจำนวนพารามิเตอร์ที่ประมาณจะทำให้มีค่าความผิดพลาดจากการประมาณเพิ่มขึ้นจึงมีผลทำให้ค่าระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ยมีแนวโน้มเพิ่มขึ้น

2. วิธีการประมาณแบบคิวเซนเบอร์รี่ และ เฮิร์ส (วิธีที่ 2)

วิธีนี้ให้ค่าระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ย **สูงสุด** ในทุกระดับการทดลอง และครอบคลุมค่าสัดส่วนประชากรในทุกระดับการทดลอง

จากการศึกษาที่ได้พบว่าเมื่อขนาดตัวอย่าง (n) เพิ่มขึ้น ค่าระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ยมีแนวโน้มลดลง เพราะเมื่อขนาดตัวอย่างเพิ่มขึ้นจะส่งผลให้ค่าความผิดพลาดจากการประมาณลดลงจึงทำให้มีการกระจายลดลงตามไปด้วย และเมื่อจำนวนกลุ่มของค่าสัดส่วนประชากร (k) เพิ่มมากขึ้นจะมีผลให้ค่าระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ยมีแนวโน้มเพิ่มขึ้น เพราะเมื่อเพิ่มจำนวนพารามิเตอร์ที่ประมาณจะทำให้มีค่าความผิดพลาดจากการประมาณเพิ่มขึ้นจึงมีผลทำให้ค่าระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ยมีแนวโน้มเพิ่มขึ้น

3. วิธีการประมาณแบบเอฟ (วิธีที่ 3)

วิธีการนี้ไม่สามารถให้ค่าระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ย ต่ำสุด หรือสูงสุดได้ในทุกระดับการทดลอง และครอบคลุมค่าสัดส่วนประชากรในทุกระดับการทดลอง

จากการศึกษาที่ได้พบว่าเมื่อขนาดตัวอย่าง (n) เพิ่มขึ้น ค่าระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ยมีแนวโน้มลดลง เพราะเมื่อขนาดตัวอย่างเพิ่มขึ้นจะส่งผลให้ค่าความผิดพลาดจากการประมาณลดลงจึงทำให้มีการกระจายลดลงตามไปด้วย และเมื่อจำนวนกลุ่มของค่าสัดส่วนประชากร (k) เพิ่มมากขึ้นจะมีผลให้ค่าระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ยมีแนวโน้มเพิ่มขึ้น เพราะเมื่อเพิ่มจำนวนพารามิเตอร์ที่ประมาณจะทำให้มีค่าความผิดพลาดจากการประมาณเพิ่มขึ้นจึงมีผลทำให้ค่าระยะห่างของค่าบริเวณความเชื่อมั่นเฉลี่ยมีแนวโน้มเพิ่มขึ้น

ข้อสังเกตที่ได้จากการทดลอง

1. วิธีการประมาณแบบคิวเซนเบอร์รี่ และเฮิร์สจากการทดลองจะพบว่าจะให้ระยะห่างของบริเวณความเชื่อมั่นสูงที่สุดในสามวิธี เนื่องมาจากวิธีการดังกล่าวมีฐานมาจากการประมาณแบบสถิติที่ไม่ใช้พารามิเตอร์(non-parametric) ซึ่งครอบคลุมพารามิเตอร์ที่กำหนดได้มากกว่า แต่จะให้อำนาจการทดสอบที่ต่ำกว่า สามารถนำไปใช้งานได้ง่าย

2. วิธีการประมาณแบบเอฟ เมื่อพิจารณาระยะห่างของบริเวณความเชื่อมั่นเฉลี่ยพบว่าไม่สามารถให้ค่าระยะห่างเฉลี่ยที่ต่ำสุด แต่เมื่อขนาดตัวอย่างเพิ่มขึ้น ระยะห่างจากวิธีดังกล่าวจะให้ค่าระยะห่างที่ใกล้เคียงกับวิธีการประมาณแบบปกติ

บทที่ 5
สรุปผลการวิจัยและข้อเสนอแนะ

การวิจัยครั้งนี้มีวัตถุประสงค์ เพื่อหาข้อสรุปในการเลือกวิธีการประมาณค่าแบบช่วงที่เหมาะสมของพารามิเตอร์ของการแจกแจงแบบพหุนาม 3 วิธี ดังนี้

<u>1. วิธีประมาณค่าแบบปกติ (Normal Method)</u>

รูปแบบของค่าประมาณแบบช่วง (p_L, p_U) จะมีค่า p_L และ p_U อยู่ในรูปของ

ขีดจำกัดความเชื่อมั่นล่าง (p_L) : $\hat{p}_i - z_{1-a/(2k)}\sqrt{\dfrac{\hat{p}_i(1-\hat{p}_i)}{n}}$

และขีดจำกัดความเชื่อมั่นบน (p_U) : $\hat{p}_i + z_{1-a/(2k)}\sqrt{\dfrac{\hat{p}_i(1-\hat{p}_i)}{n}}$

<u>2. วิธีประมาณค่าแบบคิวเซนเบอร์รี่และเฮิร์ส (Quesenbery and Hurst's Method)</u>

รูปแบบของค่าประมาณแบบช่วง (p_L, p_U) จะมีค่า p_L และ p_U อยู่ในรูปของ

ขีดจำกัดความเชื่อมั่นล่าง (p_L) :

$$\left(\frac{c^2_{k-1,1-a} + 2n_i - \sqrt{c^2_{k-1,1-a}\left(c^2_{k-1,1-a} + 4\dfrac{n_i}{n}(n-n_i)\right)}}{2\left(n + c^2_{k-1,1-a}\right)}\right)$$

และขีดจำกัดความเชื่อมั่นบน (p_U) :

$$\left(\frac{c^2_{k-1,1-a} + 2n_i + \sqrt{c^2_{k-1,1-a}\left(c^2_{k-1,1-a} + 4\dfrac{n_i}{n}(n-n_i)\right)}}{2\left(n + c^2_{k-1,1-a}\right)}\right)$$

3. วิธีประมาณค่าแบบเอฟ (F Method)

รูปแบบของค่าประมาณแบบช่วง (p_L, p_U) จะมีค่า p_L และ p_U อยู่ในรูปของ

ขีดจำกัดความเชื่อมั่นล่าง (p_L) :
$$\frac{n_i}{n_i + (n - n_i + 1)F_{(2(n-n_i+1), 2n_i), 1-\frac{a}{2k}}}$$

ขีดจำกัดความเชื่อมั่นบน (p_U) :
$$\frac{(n_i + 1)F_{(2(n_i+1), 2(n-n_i)), 1-\frac{a}{2k}}}{n - n_i + (n_i + 1)F_{(2(n_i+1), 2(n-n_i)), 1-\frac{a}{2k}}}$$

การเปรียบเทียบวิธีการประมาณกระทำในสถานการณ์ต่างๆ ดังนี้

1. ขนาดตัวอย่าง จะใช้เท่ากับ 50, 100, 200, 500, 1000
2. กำหนดค่า p_i ในแต่ละกลุ่มจะกำหนดโดยแบ่งเป็น 3 กรณีย่อย
    - 2.1 กรณีที่ค่า p_i มีการกระจายแบบเท่ากัน
    - 2.2 กรณีที่ค่า p_i มีการกระจายแบบสมมาตร
    - 2.3 กรณีที่ค่า p_i มีการกระจายแบบเบ้
3. ณ ระดับความเชื่อมั่น 90%, 95% และ 99%

ในการเปรียบเทียบวิธีการประมาณค่าสัดส่วน จะทำการเปรียบเทียบค่าสัมประสิทธิ์บริเวณความเชื่อมั่น และระยะห่างบริเวณความเชื่อมั่นที่ได้ประกอบการตัดสินใจในการเลือกตัวแบบ โดยมีเกณฑ์ว่าวิธีการใดให้ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นสูงที่สุด และให้ค่าระยะห่างบริเวณความเชื่อมั่นต่ำ จะถือว่าวิธีการดังกล่าวเป็นวิธีการประมาณที่เหมาะสมที่สุด

การศึกษาวิจัยใช้ข้อมูลที่ได้จากการจำลองด้วยเทคนิคมอนติคาร์โล สำหรับสถานการณ์ต่างๆ ที่กำหนดโดยทำการทดลองซ้ำ 2000 ครั้งในแต่ละสถานการณ์วิจัยซึ่งสามารถสรุปได้ดังหัวข้อต่อไปนี้

5.1 สรุปผลการทดลอง

ในการเปรียบเทียบสัมประสิทธิ์บริเวณความเชื่อมั่น และค่าระยะห่างบริเวณความเชื่อมั่นเฉลี่ยของวิธีการประมาณทั้งสามวิธีได้ผลสรุปดังนี้

จากการเปรียบเทียบสัมประสิทธิ์บริเวณความเชื่อมั่นจากการทดลอง กับระดับความเชื่อมั่นที่กำหนดพบว่า วิธีการประมาณแบบเอฟและวิธีการประมาณแบบคิวเซนเบอร์รี่และเฮิร์ส จะให้ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นไม่ต่ำกว่าระดับความเชื่อที่กำหนดในทุกระดับค่าของขนาดตัวอย่างและค่าสัดส่วนประชากรที่ทำการศึกษาเปรียบเทียบ ส่วนวิธีการประมาณแบบปกติ ให้ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นไม่ต่ำกว่าระดับความเชื่อมั่นที่กำหนด เฉพาะบางสถานการณ์ที่ทำการศึกษาเปรียบเทียบเท่านั้น ดังผลลัพธ์ต่อไปนี้

1. การเปรียบเทียบค่าสัมประสิทธิ์บริเวณความเชื่อมั่นของช่วงความเชื่อมั่นที่คำนวณได้จากวิธีการประมาณทั้ง 3 วิธี กับค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด

1.1 วิธีการประมาณแบบปกติ จะให้ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด ซึ่งแสดงรายละเอียดในตารางที่ 5.1

ตารางที่ 5.1 แสดงระดับขนาดตัวอย่างและลักษณะค่าสัดส่วน ที่วิธีประมาณแบบปกติ ให้ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นไม่ต่ำกว่าระดับความเชื่อมั่นที่กำหนด ณ ระดับความเชื่อมั่น 90%, 95% และ 99%

ขนาดตัวอย่าง	ประเภทสัดส่วนประชากร	ณ ระดับความเชื่อมั่น 90%				ณ ระดับความเชื่อมั่น 95%				ณ ระดับความเชื่อมั่น 99%			
		จำนวนกลุ่มของสัดส่วนประชากร (k)				จำนวนกลุ่มของสัดส่วนประชากร (k)				จำนวนกลุ่มของสัดส่วนประชากร (k)			
		4	6	8	10	4	6	8	10	4	6	8	10
50	เท่ากัน	-	-	-	-	-	-	-	-	-	-	-	-
	เบ้	-	-	-	-	-	-	-	-	-	-	-	-
	สมมาตร	-	-	-	-	-	-	-	-	-	-	-	-
100	เท่ากัน	-	-	-	-	-	-	-	-	-	-	-	-
	เบ้	-	-	-	-	-	-	-	-	-	-	-	-
	สมมาตร	-	-	-	-	-	-	-	-	-	-	-	-
200	เท่ากัน	-	-	-	-	-	-	-	-	-	-	-	-
	เบ้	-	-	-	-	-	-	-	-	-	-	-	-
	สมมาตร	√	-	-	-	-	-	-	-	-	-	-	-
500	เท่ากัน	√	-	-	-	-	√	-	-	-	√	-	-
	เบ้	√	√	-	-	√	-	-	-	√	-	-	-
	สมมาตร	√	-	-	-	√	-	-	-	√	-	-	-
1000	เท่ากัน	√	-	-	√	√	√	-	√	√	√	√	√
	เบ้	√	√	√	-	√	√	-	-	√	√	-	-
	สมมาตร	√	√	-	-	√	√	-	-	√	√	-	-

68

จะเห็นได้ว่าเมื่อกลุ่มของสัดส่วนประชากรมีจำนวนน้อยและข้อมูลมีขนาดตัวอย่างมาก วิธีการประมาณแบบปกติจะให้ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นไม่ต่ำกว่าระดับความเชื่อมั่นที่กำหนด เมื่อกลุ่มของสัดส่วนประชากรมีจำนวนมากขึ้นพบว่าวิธีนี้จะให้ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นที่ผ่านเกณฑ์กำหนดน้อยลง

1.2 <u>วิธีการประมาณแบบคิวเซนเบอร์รี่และเฮิร์ส</u> จะให้ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นจากการทดลอง ไม่ต่ำกว่าระดับความเชื่อมั่นที่กำหนดในทุกสถานการณ์

1.3 <u>วิธีการประมาณแบบเอฟ</u> จะให้ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นจากการทดลอง ไม่ต่ำกว่าระดับความเชื่อมั่นที่กำหนดในทุกสถานการณ์

จากวิธีการประมาณทั้ง 3 สรุปได้ว่า วิธีการประมาณแบบคิวเซนเบอร์รี่และเฮิร์สจะให้ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นที่สูงที่สุด รองลงมาคือวิธีการประมาณแบบเอฟ และวิธีการประมาณแบบปกติ ตามลำดับ

<u>2. การเปรียบเทียบระยะห่างบริเวณความเชื่อมั่นเฉลี่ย</u>

ในการเปรียบเทียบระยะห่างจะพิจารณาเฉพาะสถานการณ์ที่วิธีการประมาณให้ค่าสัมประสิทธิ์ช่วงความเชื่อมั่นไม่ต่ำกว่าค่าสัมประสิทธิ์ช่วงความเชื่อมั่นที่กำหนดเท่านั้น ซึ่งสามารถสรุปได้ว่า วิธีการประมาณแบบปกติ จะให้ค่าระยะห่างบริเวณความเชื่อมั่นเฉลี่ยต่ำที่สุด รองลงมาคือวิธีการประมาณแบบเอฟ และวิธีการประมาณแบบคิวเซนเบอร์รี่และเฮิร์ส ตามลำดับ

5.2 **ข้อเสนอแนะ**

ผลการวิจัยแบ่งข้อเสนอแนะเป็น 2 ด้าน ดังนี้

5.2.1 **การนำไปใช้ประโยชน์**

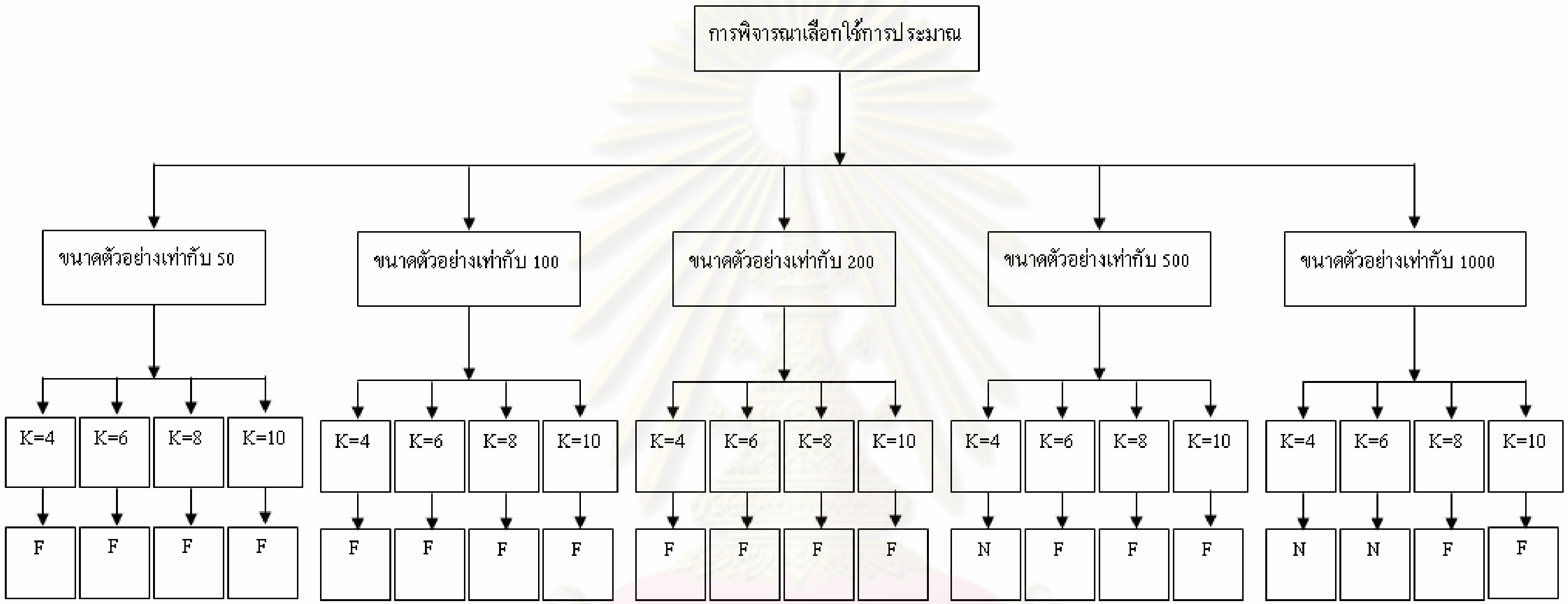
1. กรณีที่ข้อมูลมีขนาดตัวอย่างมากและมีจำนวนกลุ่มของสัดส่วนประชากรน้อย ควรเลือกใช้วิธีการประมาณแบบปกติ เพราะจะให้ค่าสัมประสิทธิ์ที่ใกล้เคียงกันกับอีกสองวิธีที่เหลือแต่จะให้ระยะห่างของบริเวณความเชื่อมั่นต่ำกว่า กล่าวคือ มีความแม่นยำ และมีการกระจายต่ำ

2. วิธีการประมาณแบบคิวเซนเบอร์รี่ และเฮิร์ส เป็นวิธีการประมาณที่ใช้ง่ายและใช้ได้ทุกกรณี มีค่าสัมประสิทธิ์บริเวณความเชื่อมั่นสูง แต่ต้องระวังเรื่องการกระจายเนื่องจากวิธีดังกล่าวให้ระยะห่างบริเวณความเชื่อมั่นสูง

3. กรณีที่ข้อมูลมีขนาดตัวอย่างต่ำ วิธีการประมาณแบบเอฟ จะเป็นทางเลือกที่ดี เพราะวิธีนี้จะมีค่าสัมประสิทธิ์บริเวณความเชื่อมั่นเฉลี่ยสูง (ซึ่งแตกต่างจากวิธีการประมาณแบบปกติ) และมีระยะห่างบริเวณความเชื่อมั่นต่ำ แต่วิธีนี้จะไม่ดีในกรณีที่ขนาดตัวอย่างใหญ่ เพราะเมื่อตัวอย่างมีขนาดใหญ่จะทำให้ข้อมูลลู่เข้าสู่การแจกแจงแบบปกติซึ่งไม่เหมาะกับการประมาณแบบเอฟ

ข้อสรุปในการเลือกใช้วิธีการประมาณในเชิงทฤษฎี

จากการพิจารณาค่าสัมประสิทธิ์บริเวณความเชื่อมั่น ประกอบกับระยะห่างบริเวณความเชื่อมั่นเฉลี่ย ถ้าพิจารณา ณ ระดับความเชื่อมั่น 90%, 95% และ 99% วิธีการใดที่ให้ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นที่ไม่ต่ำกว่าระดับความเชื่อมั่นที่กำหนด และให้ค่าระยะห่างบริเวณความเชื่อมั่นเฉลี่ยต่ำสุดจะถือว่าวิธีดังกล่าวเป็นวิธีที่เหมาะสมที่สุดในแต่ละสถานการณ์ ดังจะสรุปได้ดังรูปที่ 5.1



F แทนด้วย การประมาณแบบเอฟ

N แทนด้วย การประมาณแบบปกติ

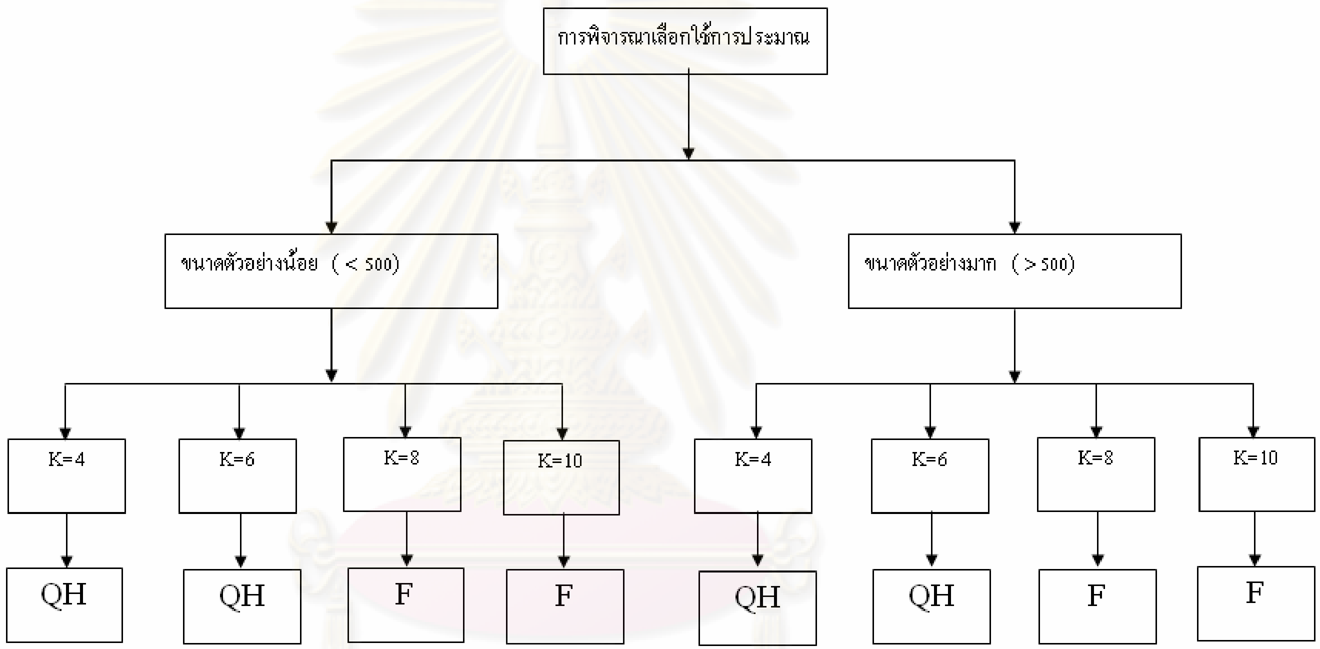
QH แทนด้วย การประมาณแบบคิวเซนเบอร์รี่และเฮิร์ส

รูปที่ 5.1 สรุปการพิจารณาเลือกใช้วิธีการประมาณในเชิงทฤษฎี

<u>ข้อสรุปในการเลือกใช้วิธีการประมาณในเชิงปฏิบัติ</u>

เมื่อพิจารณาเปรียบเทียบค่าสัมประสิทธิ์บริเวณความเชื่อมั่น และระยะห่างบริเวณความเชื่อมั่นเฉลี่ยของวิธีประมาณแบบเอฟ และวิธีประมาณแบบคิวเซนเบอร์รี่และเฮิร์ส พบว่าในกรณีที่ประชากรมีขนาดน้อย (น้อยกว่า 500) และมีจำนวนกลุ่มของค่าสัดส่วนประชากร<u>ตั้งแต่ 6 กลุ่มลงมา</u> จะ ค่าสัมประสิทธิ์บริเวณความเชื่อมั่นของวิธีคิวเซนเบอร์รี่และเฮิร์สจะมีค่าสูงกว่าวิธีการประมาณแบบเอฟ แต่จะให้ค่าระยะห่างบริเวณความเชื่อมั่นเฉลี่ยใกล้เคียงกัน เพราะฉะนั้นในเชิงปฏิบัติจึงควรเลือกใช้การประมาณ จึงควรเลือกใช้วิธีการประมาณคิวเซนเบอร์รี่และเฮิร์ส ในกรณีดังกล่าว

ปัจจัยสำคัญอีกประการหนึ่งที่เลือกวิธีคิวเซนเบอร์รี่และเฮิร์ส คือ วิธีการดังนี้จะทำการพิจารณา**ค่าความผิดพลาดของพารามิเตอร์ทุกตัวในการประมาณค่าพารามิเตอร์แต่ละตัว** ซึ่งแตกต่างจากวิธีการประมาณค่าแบบปกติ และเอฟ ที่ไม่มีการพิจารณาค่าความผิดพลาดของพารามิเตอร์ตัวอื่นประกอบ ดังนั้นในสถานการณ์ค่าระยะห่างบริเวณความเชื่อมั่นมีค่าไม่แตกต่างกันมากจึงควรเลือกใช้วิธีการประมาณแบบคิวเซนเบอร์รี่และเฮิร์ส



F แทนด้วย การประมาณแบบเอฟ

N แทนด้วย การประมาณแบบปกติ

QH แทนด้วย การประมาณแบบคิวเซนเบอร์รี่และเฮิร์ส

รูปที่ 5.2 สรุปการพิจารณาเลือกใช้วิธีการประมาณในเชิงปฏิบัติ

73

5.2.2 ด้านการศึกษาวิจัย

1. ขนาดตัวอย่างดังกล่าวอาจจะไม่เหมาะสมกับวิธีการประมาณแบบปกติ เพราะวิธีดังกล่าวต้องใช้ขนาดตัวอย่างที่มาก ซึ่งแตกต่างจากวิธีการประมาณแบบคิวเซนเบอร์รี่และเฮิร์สและวิธีการประมาณแบบเอฟ จึงควรศึกษาเพิ่มเติมว่าวิธีการแต่ละวิธีควรมีขนาดตัวอย่างเท่าใดเมื่อเทียบกับจำนวนกลุ่มสัดส่วนประชากรที่กำหนด จึงจะสามารถครอบคลุมค่าสัดส่วนประชากรได้ทุกค่าที่ทำการทดลอง

2. การกำหนดค่าสัดส่วนประชากร ซึ่งในงานวิจัยครั้งนี้แบ่งค่าสัดส่วนประชากรออกเป็น 3 ประเภทกว้างๆซึ่งอาจจะยังไม่เพียงพอ ดังนั้นเพื่อความถูกต้องและครบถ้วนจึงควรทำการศึกษาเพิ่มเติมในกรณีที่ประเภทค่าสัดส่วนประชากรมีความหลากหลายมากกว่านี้

3. นอกเหนือจากวิธีการประมาณทั้ง 3 ที่กล่าวมา พบว่ายังมีวิธีการประมาณค่าแบบอื่นๆ ที่น่าสนใจอีก ดังเช่น วิธีการประมาณค่าแบบสคอร์ ซึ่งวิธีนี้ได้ใช้วิธีการของเบส์ (Bayes method) มาช่วยในการประมาณ วิธีการประมาณค่าแบบปัวร์ซอง ซึ่งวิธีนี้เหมาะสำหรับกรณีที่ขนาดตัวอย่างมีค่ามาก และค่าสัดส่วนประชากรมีค่าน้อย วิธีแปลงแบบอาร์คไซน์ เป็นวิธีการประมาณแบบช่วงที่ใช้หลักทฤษฎีค่าจำกัดสู่ส่วนกลางเช่นเดียวกับวิธีการประมาณแบบปกติ วิธีนี้จะให้ค่าระยะห่างบริเวณความเชื่อมั่นที่มีค่าน้อย ซึ่งเป็นวิธีที่น่าสนใจและควรได้ทำการศึกษาวิจัยเปรียบเทียบต่อไป

รายการอ้างอิง

ภาษาไทย

ธีระพร วีระภาวร. การอนุมานเชิงสถิติขั้นกลาง : โครงสร้างและความหมาย. กรุงเทพมหานคร : สำนักพิมพ์จุฬาลงกรณ์มหาวิทยาลัย, 2536.

ธีระพร วีระถาวร. ความน่าจะเป็นกับ การประยุกต์. พิมพ์ครั้งที่ 2. กรุงเทพมหานคร : วิทยพัฒน์, 2539.

สงขลา สำเภาเงิน. การเปรียบเทียบวิธีประมาณค่าแบบช่วงสำหรับค่าสัดส่วนประชากรของการแจกแจงแบบทวินาม. วิทยานิพนธ์ปริญญามหาบัณฑิต, สาขาวิชาสถิติ คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย, 2542.

ภาษาอังกฤษ

Leemis, L.M. and Trivedi, K.S.. A Comparison of Approximate Interval Estimators for the Binomial Parameter. The American Statistician 50 : 63-68.

Quesenberry, C.P. and Hurst, D. C.. Large-Sample Simultaneous Confidence Interval for Multinomial Proportions. Technometrics 6 : 191-195

ภาคผนวก

ภาคผนวก ก

โปรมแกรมที่ใช้ในงานวิจัย

โปรแกรมสำหรับคำนวณค่าระดับความเชื่อมั่น และค่าความยาวเฉลี่ยของตัวประมาณแบบช่วงของค่าสัดส่วนประชากร ที่ประมาณโดยใช้วิธีต่างๆ ประกอบด้วย

1. วิธีการประมาณแบบปกติ มีสูตรการประมาณค่าช่วงความเชื่อมั่น คือ

$$p_L = \hat{p}_i - z_{1-a/(2k)}\sqrt{\frac{\hat{p}_i(1-\hat{p}_i)}{N}}$$

และ $$p_U = \hat{p}_i + z_{1-a/(2k)}\sqrt{\frac{\hat{p}_i(1-\hat{p}_i)}{N}}$$

2. วิธีการประมาณแบบคิวเซนเบอร์รี่และเฮิร์ส มีสูตรการประมาณค่าช่วงความเชื่อมั่น คือ

$$p_L = \left(\frac{c^2_{k-1,1-a} + 2n_i - \sqrt{c^2_{k-1,1-a}\left(c^2_{k-1,1-a} + 4\frac{n_i}{N}(N-n_i) \right)}}{2\left(N + c^2_{k-1,1-a}\right)} \right)$$

และ $$p_U = \left(\frac{c^2_{k-1,1-a} + 2n_i + \sqrt{c^2_{k-1,1-a}\left(c^2_{k-1,1-a} + 4\frac{n_i}{N}(N-n_i) \right)}}{2\left(N + c^2_{k-1,1-a}\right)} \right)$$

3. วิธีการประมาณแบบเอฟ มีสูตรการประมาณค่าช่วงความเชื่อมั่น คือ

$$p_L = \frac{n_i}{n_i + (N - n_i + 1)F_{2(N-n_i+1),2n_i,1-\frac{a}{2k}}}$$

และ $$p_U = \frac{(n_i + 1)F_{2(n_i+1),2(N-n_i),1-\frac{a}{2k}}}{N - n_i + (n_i + 1)F_{2(n_i+1),2(N-n_i),1-\frac{a}{2k}}}$$

รายละเอียดของโปรแกรมการคำนวณค่าสัมประสิทธิ์บริเวณความเชื่อมั่นและค่าระยะห่างบริเวณความเชื่อมั่น แสดงในหน้าถัดไป

กรณี ขนาดตัวอย่างเท่ากับ 50 ค่าสัดส่วนตัวอย่าง 4 กลุ่ม

```
i=2000;
y1=0;
y2=0;
y3=0;
y4=0;
j1=0;
j1_1=0;
j1_2=0;
j1_3=0;
j1_4=0;
j2_1=0;
j2_2=0;
j2_3=0;
j2_4=0;
j3_1=0;
j3_2=0;
j3_3=0;
j3_4=0;
p_1=1/4;
p_2=1/4;
p_3=1/4;
p_4=1/4;
n_c=50;
c=2000;
b_M1=0;
b_M2=0;
b_M3=0;
k=4;
```

```
while i>0
n=50;
while n>0
myrand=rand();
if myrand>0 & myrand<=0.25
y1=y1+1;
elseif myrand>0.25 & myrand<=0.5
y2=y2+1;
elseif myrand>0.5 & myrand<=0.75
y3=y3+1;
else
y4=y4+1;
end
n=n-1;
end
if  y1>0 & y2>0 & y3>0 & y4>0
Round=2000-i+1
y=[y1 y2 y3 y4];
p(1,1)=y1/n_c;
p(1,2)=y2/n_c;
p(1,3)=y3/n_c;
p(1,4)=y4/n_c;
Total_p(Round,:)=p;
int1_p1(Round,1)=(chi2inv(0.9,k-1)+2*y1-sqrt(chi2inv(0.9,k-1)*(chi2inv(0.9,k-1)+4*y1*((n_c-
y1)/n_c))))/(2*(n_c+chi2inv(0.9,k-1)));
int1_p1(Round,2)=(chi2inv(0.9,k-1)+2*y1+sqrt(chi2inv(0.9,k-1)*(chi2inv(0.9,k-1)+4*y1*((n_c-
y1)/n_c))))/(2*(n_c+chi2inv(0.9,k-1)));
int1_p2(Round,1)=(chi2inv(0.9,k-1)+2*y2-sqrt(chi2inv(0.9,k-1)*(chi2inv(0.9,k-1)+4*y2*((n_c-
y2)/n_c))))/(2*(n_c+chi2inv(0.9,k-1)));
int1_p2(Round,2)=(chi2inv(0.9,k-1)+2*y2+sqrt(chi2inv(0.9,k-1)*(chi2inv(0.9,k-1)+4*y2*((n_c-
y2)/n_c))))/(2*(n_c+chi2inv(0.9,k-1)));
```

```
int1_p3(Round,1)=(chi2inv(0.9,k-1)+2*y3-sqrt(chi2inv(0.9,k-1)*(chi2inv(0.9,k-1)+4*y3*((n_c-
y3)/n_c))))/(2*(n_c+chi2inv(0.9,k-1)));
int1_p3(Round,2)=(chi2inv(0.9,k-1)+2*y3+sqrt(chi2inv(0.9,k-1)*(chi2inv(0.9,k-1)+4*y3*((n_c-
y3)/n_c))))/(2*(n_c+chi2inv(0.9,k-1)));
int1_p4(Round,1)=(chi2inv(0.9,k-1)+2*y4-sqrt(chi2inv(0.9,k-1)*(chi2inv(0.9,k-1)+4*y4*((n_c-
y4)/n_c))))/(2*(n_c+chi2inv(0.9,k-1)));
int1_p4(Round,2)=(chi2inv(0.9,k-1)+2*y4+sqrt(chi2inv(0.9,k-1)*(chi2inv(0.9,k-1)+4*y4*((n_c-
y4)/n_c))))/(2*(n_c+chi2inv(0.9,k-1)));
int2_p1(Round,1)=p(1,1)-norminv(1-(0.1/(2*k)),0,1)*sqrt(p(1,1)*(1-p(1,1))/n_c);
int2_p1(Round,2)=p(1,1)+norminv(1-(0.1/(2*k)),0,1)*sqrt(p(1,1)*(1-p(1,1))/n_c);
int2_p2(Round,1)=p(1,2)-norminv(1-(0.1/(2*k)),0,1)*sqrt(p(1,2)*(1-p(1,2))/n_c);
int2_p2(Round,2)=p(1,2)+norminv(1-(0.1/(2*k)),0,1)*sqrt(p(1,2)*(1-p(1,2))/n_c);
int2_p3(Round,1)=p(1,3)-norminv(1-(0.1/(2*k)),0,1)*sqrt(p(1,3)*(1-p(1,3))/n_c);
int2_p3(Round,2)=p(1,3)+norminv(1-(0.1/(2*k)),0,1)*sqrt(p(1,3)*(1-p(1,3))/n_c);
int2_p4(Round,1)=p(1,4)-norminv(1-(0.1/(2*k)),0,1)*sqrt(p(1,4)*(1-p(1,4))/n_c);
int2_p4(Round,2)=p(1,4)+norminv(1-(0.1/(2*k)),0,1)*sqrt(p(1,4)*(1-p(1,4))/n_c);
int3_p1(Round,1)=y1/(y1+(n_c-y1+1)*finv(1-(0.1/(2*k)),2*(n_c-y1+1),2*y1));
int3_p1(Round,2)=((y1+1)*finv(1-(0.1/(2*k)),2*(y1+1),2*(n_c-y1)))/(n_c-y1+((y1+1)*finv(1-
(0.1/(2*k)),2*(y1+1),2*(n_c-y1))));
int3_p2(Round,1)=y2/(y2+(n_c-y2+1)*finv(1-(0.1/(2*k)),2*(n_c-y2+1),2*y2));
int3_p2(Round,2)=((y2+1)*finv(1-(0.1/(2*k)),2*(y2+1),2*(n_c-y2)))/(n_c-y2+((y2+1)*finv(1-
(0.1/(2*k)),2*(y2+1),2*(n_c-y2))));
int3_p3(Round,1)=y3/(y3+(n_c-y3+1)*finv(1-(0.1/(2*k)),2*(n_c-y3+1),2*y3));
int3_p3(Round,2)=((y3+1)*finv(1-(0.1/(2*k)),2*(y3+1),2*(n_c-y3)))/(n_c-y3+((y3+1)*finv(1-
(0.1/(2*k)),2*(y3+1),2*(n_c-y3))));
int3_p4(Round,1)=y4/(y4+(n_c-y4+1)*finv(1-(0.1/(2*k)),2*(n_c-y4+1),2*y4));
int3_p4(Round,2)=((y4+1)*finv(1-(0.1/(2*k)),2*(y4+1),2*(n_c-y4)))/(n_c-y4+((y4+1)*finv(1-
(0.1/(2*k)),2*(y4+1),2*(n_c-y4))));
if int1_p1(Round,1)<0
int1_p1(Round,1)=0;
end
```

```
if int1_p2(Round,1)<0
int1_p2(Round,1)=0;
end
if int1_p3(Round,1)<0
int1_p3(Round,1)=0;
end
if int1_p4(Round,1)<0
int1_p4(Round,1)=0;
end
if int1_p1(Round,2)>1
int1_p1(Round,2)=1;
end
if int1_p2(Round,2)>1
int1_p2(Round,2)=1;
end
if int1_p3(Round,2)>1
int1_p3(Round,2)=1;
end
if int1_p4(Round,2)>1
int1_p4(Round,2)=1;
end
if int2_p1(Round,1)<0
int2_p1(Round,1)=0;
end
if int2_p2(Round,1)<0
int1_p2(Round,1)=0;
end
if int2_p3(Round,1)<0
int2_p3(Round,1)=0;
end
if int2_p4(Round,1)<0
```

```
int2_p4(Round,1)=0;
end
if int2_p1(Round,2)>1
int2_p1(Round,2)=1;
end
if int2_p2(Round,2)>1
int2_p2(Round,2)=1;
end
if int2_p3(Round,2)>1
int2_p3(Round,2)=1;
end
if int2_p4(Round,2)>1
int2_p4(Round,2)=1;
end
if int3_p1(Round,1)<0
int1_p1(Round,1)=0;
end
if int3_p2(Round,1)<0
int1_p2(Round,1)=0;
end
if int3_p3(Round,1)<0
int1_p3(Round,1)=0;
end
if int3_p4(Round,1)<0
int3_p4(Round,1)=0;
end
if int3_p1(Round,2)>1
int3_p1(Round,2)=1;
end
if int3_p2(Round,2)>1
int3_p2(Round,2)=1;
```

```
end
if int3_p3(Round,2)>1
int3_p3(Round,2)=1;
end
if int3_p4(Round,2)>1
int3_p4(Round,2)=1;
end
if p_1>int1_p1(Round,1) & p_1<int1_p1(Round,2) & p_2>int1_p2(Round,1) &
p_2<int1_p2(Round,2) & p_3>int1_p3(Round,1) & p_3<int1_p3(Round,2) &
p_4>int1_p4(Round,1) & p_4<int1_p4(Round,2)
j1=j1+1 ;
end
if p_1>int1_p1(Round,1) & p_1<int1_p1(Round,2)
j1_1=j1_1+1 ;
else
j1_1=j1_1+0;
end
if p_2>int1_p2(Round,1) & p_2<int1_p2(Round,2)
j1_2=j1_2+1 ;
else
j1_2=j1_2+0;
end
if p_3>int1_p3(Round,1) & p_3<int1_p3(Round,2)
j1_3=j1_3+1 ;
else
j1_3=j1_3+0;
end
if p_4>int1_p4(Round,1) & p_4<int1_p4(Round,2)
j1_4=j1_4+1 ;
else
j1_4=j1_4+0;
```

```
end
if p_1>int2_p1(Round,1) & p_1<int2_p1(Round,2)
j2_1=j2_1+1 ;
else
j2_1=j2_1+0;
end
if p_2>int2_p2(Round,1) & p_2<int2_p2(Round,2)
j2_2=j2_2+1 ;
else
j2_2=j2_2+0;
end

if p_3>int2_p3(Round,1) & p_3<int2_p3(Round,2)
j2_3=j2_3+1 ;
else
j2_3=j2_3+0;
end
if p_4>int2_p4(Round,1) & p_4<int2_p4(Round,2)
j2_4=j2_4+1 ;
else
j2_4=j2_4+0;
end

if p_1>int3_p1(Round,1) & p_1<int3_p1(Round,2)
j3_1=j3_1+1 ;
else
j3_1=j3_1+0;
end
if p_2>int3_p2(Round,1) & p_2<int3_p2(Round,2)
j3_2=j3_2+1 ;
else
```

```
j3_2=j3_2+0;
end
if p_3>int3_p3(Round,1) & p_3<int3_p3(Round,2)
j3_3=j3_3+1 ;
else
j3_3=j3_3+0;
end
if p_4>int3_p4(Round,1) & p_4<int3_p4(Round,2)
j3_4=j3_4+1 ;
else
j3_4=j3_4+0;
end
Round_vs_int1(Round,1)=j1_1/Round;
Round_vs_int1(Round,2)=j1_2/Round;
Round_vs_int1(Round,3)=j1_3/Round;
Round_vs_int1(Round,4)=j1_4/Round;

Round_vs_int2(Round,1)=j2_1/Round;
Round_vs_int2(Round,2)=j2_2/Round;
Round_vs_int2(Round,3)=j2_3/Round;
Round_vs_int2(Round,4)=j2_4/Round;

Round_vs_int3(Round,1)=j3_1/Round;
Round_vs_int3(Round,2)=j3_2/Round;
Round_vs_int3(Round,3)=j3_3/Round;
Round_vs_int3(Round,4)=j3_4/Round;

mean_R_int1=Round_vs_int1;
mean_R_int2=Round_vs_int2;
mean_R_int3=Round_vs_int3;
```

```
if Round>1
mean_R_int1=mean(Round_vs_int1);
mean_R_int2=mean(Round_vs_int2);
mean_R_int3=mean(Round_vs_int3);
end

varianceint1(Round,1)=sum(((Round_vs_int1(:,1)-mean_R_int1(1,1)).^2)/Round);
varianceint1(Round,2)=sum(((Round_vs_int1(:,2)-mean_R_int1(1,2)).^2)/Round);
varianceint1(Round,3)=sum(((Round_vs_int1(:,3)-mean_R_int1(1,3)).^2)/Round);
varianceint1(Round,4)=sum(((Round_vs_int1(:,4)-mean_R_int1(1,4)).^2)/Round);
varianceint2(Round,1)=sum(((Round_vs_int1(:,1)-mean_R_int2(1,1)).^2)/Round);
varianceint2(Round,2)=sum(((Round_vs_int2(:,2)-mean_R_int2(1,2)).^2)/Round);
varianceint2(Round,3)=sum(((Round_vs_int2(:,3)-mean_R_int2(1,3)).^2)/Round);
varianceint2(Round,4)=sum(((Round_vs_int2(:,4)-mean_R_int2(1,4)).^2)/Round);
varianceint3(Round,1)=sum(((Round_vs_int2(:,1)-mean_R_int3(1,1)).^2)/Round);
varianceint3(Round,2)=sum(((Round_vs_int3(:,2)-mean_R_int3(1,2)).^2)/Round);
varianceint3(Round,3)=sum(((Round_vs_int3(:,3)-mean_R_int3(1,3)).^2)/Round);
varianceint3(Round,4)=sum(((Round_vs_int3(:,4)-mean_R_int3(1,4)).^2)/Round);

variancetot1(Round,1)=sum(varianceint1(Round,:))
variancetot2(Round,1)=sum(varianceint2(Round,:))
variancetot3(Round,1)=sum(varianceint3(Round,:))

if mean_R_int1>0 & mean_R_int2>0 & mean_R_int3
CV1(Round,1)=sqrt(varianceint1(Round,1))/mean_R_int1(1,1);
CV1(Round,2)=sqrt(varianceint1(Round,2))/mean_R_int1(1,2);
CV1(Round,3)=sqrt(varianceint1(Round,3))/mean_R_int1(1,3);
CV1(Round,4)=sqrt(varianceint1(Round,4))/mean_R_int1(1,4);
CV2(Round,1)=sqrt(varianceint2(Round,1))/mean_R_int2(1,1);
CV2(Round,2)=sqrt(varianceint2(Round,2))/mean_R_int2(1,2);
CV2(Round,3)=sqrt(varianceint2(Round,3))/mean_R_int2(1,3);
```

```
CV2(Round,4)=sqrt(varianceint2(Round,4))/mean_R_int2(1,4);
CV3(Round,1)=sqrt(varianceint3(Round,1))/mean_R_int3(1,1);
CV3(Round,2)=sqrt(varianceint3(Round,2))/mean_R_int3(1,2);
CV3(Round,3)=sqrt(varianceint3(Round,3))/mean_R_int3(1,3);
CV3(Round,4)=sqrt(varianceint3(Round,4))/mean_R_int3(1,4);
end

k1_1=j1/2000;

k2_1=j2_1/2000;
k2_2=j2_2/2000;
k2_3=j2_3/2000;
k2_4=j2_4/2000;

k3_1=j3_1/2000;
k3_2=j3_2/2000;
k3_3=j3_3/2000;
k3_4=j3_4/2000;

S=[n_c*p_1^2 -p_1*p_2 -p_1*p_3 -p_1*p_4;-p_1*p_2 n_c*p_2^2 -p_2*p_3 -p_2*p_4;-p_1*p_3
-p_2*p_3 n_c*p_3^2 -p_3*p_4;-p_1*p_4 -p_2*p_4 -p_3*p_4 n_c*p_4^2];
D1(1,1)=int1_p1(Round,2)-int1_p1(Round,1);
D1(1,2)=int1_p2(Round,2)-int1_p2(Round,1);
D1(1,3)=int1_p3(Round,2)-int1_p3(Round,1);
D1(1,4)=int1_p4(Round,2)-int1_p4(Round,1)
D2(1,1)=int2_p1(Round,2)-int2_p1(Round,1);
D2(1,2)=int2_p2(Round,2)-int2_p2(Round,1);
D2(1,3)=int2_p3(Round,2)-int2_p3(Round,1);
D2(1,4)=int2_p4(Round,2)-int2_p4(Round,1);
D3(1,1)=int3_p1(Round,2)-int3_p1(Round,1);
D3(1,2)=int3_p2(Round,2)-int3_p2(Round,1);
```

```
D3(1,3)=int3_p3(Round,2)-int3_p3(Round,1);
D3(1,4)=int3_p4(Round,2)-int3_p4(Round,1);
D_maha1(Round,1)=D1*inv(S)*D1';
D_maha2(Round,1)=D2*inv(S)*D2';
D_maha3(Round,1)=D3*inv(S)*D3';
Maha1_Total(Round,1)=D_maha1(Round,1)+b_M1;
Maha2_Total(Round,1)=D_maha2(Round,1)+b_M2;
Maha3_Total(Round,1)=D_maha3(Round,1)+b_M3;
b_M1=Maha1_Total(Round,1);
b_M2=Maha2_Total(Round,1);
b_M3=Maha3_Total(Round,1);
y1=0;
y2=0;
y3=0;
y4=0;
i=i-1;
end
y1=0;
y2=0;
y3=0;
y4=0;
end

prob1=k1_1
prob2=k2_1*k2_2*k2_3*k2_4
prob3=k3_1*k3_2*k3_3*k3_4
Maha_Mean1=Maha1_Total(2000,1)/2000
Maha_Mean2=Maha2_Total(2000,1)/2000
Maha_Mean3=Maha3_Total(2000,1)/2000
```

ประวัติผู้เขียนวิทยานิพนธ์

นายบัณฑิต ฉัตรตะวัน เกิดวันศุกร์ที่ 1 เมษายน พ.ศ. 2526 สำเร็จการศึกษาปริญญาวิทยาศาสตรบัณฑิต (วท.บ.) ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ มหาวิทยาลัยมหิดล ในปีการศึกษา 2547 และเข้าศึกษาต่อในหลักสูตรสถิติศาสตรมหาบัณฑิต (สต.ม.) สาขาสถิติ คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย ในปี พ.ศ. 2549