บทที่ 1

บทนำ

สำนักงานวิทยทรัพยากร จุฬาลงกรณ์มหาวิทยาลัย

1.1 ความเป็นมาและความสำคัญของปัญหา

ในการศึกษาการประมาณค่าพารามิเตอร์โดยทั่วไป สามารถทำการประมาณได้ใน 2 รูปแบบ คือการประมาณค่าแบบจุด (Point Estimation) และการประมาณค่าแบบช่วง (Interval Estimation) การประมาณค่าแบบจุดเป็นการประมาณค่าพารามิเตอร์ด้วยค่าๆหนึ่ง ซึ่งการประมาณค่าแบบจุด ค่าประมาณที่ได้จะคลาดเคลื่อนไปจากค่าพารามิเตอร์เพียงใด ขึ้นอยู่กับการเลือกใช้ตัวประมาณที่เหมาะสม ส่วนการประมาณค่าแบบช่วงเป็นการประมาณที่จะให้ช่วงๆหนึ่ง ซึ่งมีคุณสมบัติว่าช่วงที่ประมาณได้จะคลอบคลุมค่าที่แท้จริงด้วยความเชื่อมั่นระดับหนึ่ง โดยที่การประมาณค่าแบบช่วงเป็นการประมาณโดยอาศัยตัวประมาณแบบจุด และการแจกแจงความน่าจะเป็นของตัวประมาณนั้น ซึ่งผลจากการประมาณจะทำให้ผู้วิจัยเชื่อมั่นได้ในระดับหนึ่งว่าช่วงที่ประมาณได้คลุมค่าพารามิเตอร์ที่สนใจศึกษา จะเห็นได้ว่าการประมาณค่าแบบช่วงสามารถบอกขอบเขตของค่าประมาณได้ดีกว่าการประมาณค่าแบบจุด ที่ให้ค่าประมาณที่ได้เป็นเพียงค่าๆหนึ่ง

การทดลองที่มีผลการทดลองเป็น 2 ลักษณะ คือ ลักษณะที่สนใจ หรือเกิดผลสำเร็จ (Success) และลักษณะที่ไม่สนใจ หรือไม่เกิดผลสำเร็จ (Failure) โดยมีค่าความน่าจะเป็นของการเกิดผลสำเร็จเท่ากับ p และค่าความน่าจะเป็นของการไม่เกิดผลสำเร็จเท่ากับ $1-p$ การแจกแจงที่อธิบายสถานการณ์เช่นนี้เรียกว่า การแจกแจงแบบแบร์นูลลี (Bernoulli Distribution) และเมื่อทำการทดลองแบร์นูลลีจำนวน n ครั้ง โดยที่การทดลองในแต่ละครั้งเป็นอิสระต่อกันและมีค่าความน่าจะเป็นของการเกิดผลสำเร็จเท่ากับ p คงที่ตลอดการทดลอง จะได้ว่าจำนวนครั้งของผลสำเร็จที่เกิดขึ้นจากการทดลอง n ครั้ง จะมีการแจกแจงแบบทวินาม (Binomial Distribution) โดยมีพารามิเตอร์ n และ p

สำหรับการวิจัยครั้งนี้ผู้วิจัยสนใจศึกษาการประมาณค่าแบบช่วงสำหรับผลต่างระหว่างค่าสัดส่วนของประชากร 2 กลุ่มที่มีการแจกแจงแบบแบร์นูลลีและเป็นอิสระต่อกัน สำหรับในประชากรที่ 1 กำหนดให้ $X_{11}, X_{12}, \ldots\ldots, X_{1n_1}$ เป็นตัวแปรสุ่มที่มีการแจกแจงแบบแบร์นูลลี ที่เป็นอิสระกัน โดยมีพารามิเตอร์ p_1 และในประชากรที่ 2 กำหนดให้ $X_{21}, X_{22}, \ldots\ldots, X_{2n_2}$ เป็นตัวแปรสุ่มที่มีการแจกแจงแบบแบร์นูลลีที่เป็นอิสระกัน โดยมีพารามิเตอร์ p_2 โดยที่ $0 \le p_1, p_2 \le 1$ จะได้ว่า $Y_1 = \sum_{j=1}^{n_1} X_{1j}$ และ $Y_2 = \sum_{j=1}^{n_2} X_{2j}$ เป็นจำนวนครั้งของการเกิดลักษณะที่สนใจจากการทดลอง

แบบแบร์นูลลีที่เป็นอิสระกัน n_1 และ n_2 ครั้งตามลำดับ ดังนั้น Y_1 และ Y_2 จึงเป็นตัวแปรสุ่มที่มีการแจกแจงแบบทวินาม โดยมีพารามิเตอร์คือ n_1, p_1 และ n_2, p_2 และเมื่อ n_1 และ n_2 มีขนาดใหญ่ จะได้ว่า Y_1 และ Y_2 จะมีการแจกแจงแบบปกติโดยประมาณ $N(n_1 p_1, n_1 p_1(1-p_1))$ และ $N(n_2 p_2, n_2 p_2(1-p_2))$ ตามลำดับ

ในการอนุมานเกี่ยวกับพารามิเตอร์ p_1, p_2 หรือค่าสัดส่วนประชากร (Population Proportion) ของประชากรที่ 1 และ 2 ในการประมาณค่าแบบจุดจะได้ว่า ตัวประมาณภาวะน่าจะเป็นสูงสุด (Maximum Likelihood Estimator) ของ p_1 และ p_2 คือ $\hat{p}_1 = \frac{Y_1}{n_1}$ และ $\hat{p}_2 = \frac{Y_2}{n_2}$ หรือค่าสัดส่วนตัวอย่าง (Sample Proportion) ซึ่งมีคุณสมบัติไม่เอนเอียง (Unbiased Estimator) จากทฤษฎีบทขีดจำกัดส่วนกลาง (The Central Limit Theorem) ถ้าขนาดตัวอย่าง n_1 และ n_2 มีขนาดใหญ่ จะได้ว่า $\hat{p}_1$ และ $\hat{p}_2$ จะมีแนวโน้มลู่เข้าสู่การแจกแจงแบบปกติที่มีค่าเฉลี่ย p_1 และ p_2 ตามลำดับ และค่าความแปรปรวน $p_1(1-p_1)/n_1$ และ $p_2(1-p_2)/n_2$ ตามลำดับ

การที่ตัวอย่างสุ่มทั้งสองชุดมาจากประชากรที่มีการแจกแจงเป็นอิสระกัน ดังนั้นจากทฤษฎีบทขีดจำกัดส่วนกลาง เมื่อ n_1 และ n_2 มีขนาดใหญ่ จะได้ว่าการแจกแจงของผลต่างระหว่างค่าสัดส่วนของสองตัวอย่าง $(\hat{p}_1 - \hat{p}_2)$ จะสามารถประมาณได้ด้วยการแจกแจงแบบปกติที่มีค่าเฉลี่ย $p_1 - p_2$ และค่าความแปรปรวน $\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}$ และเมื่อกำหนดระดับนัยสำคัญ α สามารถหาช่วงความเชื่อมั่น $(1-\alpha)100\%$ สำหรับผลต่างระหว่างค่าสัดส่วนของสองประชากร $(p_1 - p_2)$ ได้คือ

$$(\hat{p}_1 - \hat{p}_2) \pm Z_{1-\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}$$

เมื่อ $Z_{1-\alpha/2}$ คือเปอร์เซ็นต์ไทล์ (Percentile) ที่ $(1-\alpha/2)100$ ของการแจกแจงแบบปกติมาตรฐานที่มีค่าเฉลี่ยเป็น 0 และความแปรปรวนเป็น 1 ช่วงที่คำนวณได้นี้เรียกว่า ช่วงความเชื่อมั่น $(1-\alpha)100\%$ สำหรับผลต่างระหว่างค่าสัดส่วนของสองประชากร รูปแบบของวิธีการประมาณข้างต้นจะเรียกว่า วิธีการประมาณแบบฉบับ (The Classical Method) สามารถให้ช่วงประมาณที่ดีหรือให้ค่าสัมประสิทธิ์ความเชื่อมั่นไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด เมื่อขนาดตัวอย่างมีขนาดใหญ่

การประมาณค่าสำหรับผลต่างระหว่างค่าสัดส่วนประชากร 2 กลุ่มที่มีการแจกแจงแบบแบร์นูลลีและเป็นอิสระต่อกัน โดยอาศัยการแจกแจงแบบปกติดังกล่าว จำเป็นที่ตัวอย่างต้องมีขนาดใหญ่มาก อย่างไรก็ตาม ในทางปฏิบัติการรวบรวมข้อมูลจากตัวอย่าง บางครั้งไม่อาจหา

ข้อมูลได้จำนวนมาก เนื่องจากมีข้อจำกัดเกี่ยวกับเวลาและงบประมาณ การประมาณค่าแบบช่วงสำหรับผลต่างระหว่างค่าสัดส่วนของสองประชากร โดยใช้วิธีการประมาณแบบฉบับ จึงอาจให้ช่วงความเชื่อมั่นที่ไม่ได้คุณสมบัติตามที่คาดหวัง

ด้วยสาเหตุดังกล่าวจึงเป็นที่น่าสนใจในการศึกษาเกี่ยวกับวิธีการประมาณค่าแบบช่วงสำหรับผลต่างระหว่างค่าสัดส่วนของสองประชากรในกรณีที่ขนาดตัวอย่างไม่ใหญ่มากนัก การจะให้การแจกแจงแบบปกติประมาณการแจกแจงแบบทวินามได้ใกล้เคียง ในกรณีขนาดตัวอย่างไม่ใหญ่ จำเป็นต้องมีการใช้ค่าปรับแก้เพื่อความต่อเนื่อง (Continuity Correction) เนื่องจากการแจงแจงแบบทวินามเป็นการแจกแจงแบบไม่ต่อเนื่อง ในขณะที่การแจงแจงแบบปกติเป็นการแจกแจงแบบต่อเนื่อง จากศึกษาพบว่ามีงานวิจัยที่ทำการศึกษาเกี่ยวกับเรื่องนี้โดยความแตกต่างของแต่ละวิธี ขึ้นอยู่กับรูปแบบค่าปรับแก้เพื่อความต่อเนื่องและค่าความคลาดเคลื่อนมาตรฐาน อยู่หลายงานที่น่าสนใจคือ

Hauck and Anderson (1986) ได้ศึกษาวิธีการสร้างช่วงความเชื่อมั่น สำหรับผลต่างระหว่างค่าสัดส่วนในประชากร 2 กลุ่ม ที่มีการใช้ค่าปรับแก้ไขเพื่อความต่อเนื่อง และมีการปรับค่าความคลาดเคลื่อนมาตรฐานโดยใช้ $n_i - 1$ แทน $n_i; i = 1,2$ เนื่องจากค่า $Var(\hat{p}_1 - \hat{p}_2)$ ที่มีการปรับค่าแบบนี้เป็นค่าประมาณที่ไม่เอนเอียงของความแปรปรวนของผลต่างค่าสัดส่วนประชากร และนำมาเปรียบเทียบกับวิธีการที่ใช้ความคลาดเคลื่อนมาตรฐานร่วมกับค่าปรับแก้เพื่อความต่อเนื่องรูปแบบอื่นๆ พบว่า วิธีที่เขาเสนอเป็นวิธีที่ให้ค่าสัมประสิทธิ์ความเชื่อมั่นจากการทดลองเข้าใกล้และไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนดมากที่สุดในบรรดาวิธีการที่นำมาเปรียบเทียบ

Peskan (1990) ได้อธิบายถึงวิธีการทั่วไปในการหาขีดจำกัดความเชื่อมั่นล่างและบน ที่ถูกต้องสำหรับพารามิเตอร์ที่ไม่ทราบค่า เมื่อประชากรมีการแจกแจงแบบไม่ต่อเนื่องใดๆ วิธีการนี้สามารถนำมาประยุกต์เพื่อหาช่วงความเชื่อมั่นสำหรับ $p_1 - p_2$ ได้ และได้เสนอค่าปรับแก้เพื่อความต่อเนื่องเช่นกัน โดย Peskan กล่าวว่าวิธีการที่เขาเสนอนี้มีความเหมาะสมมากกว่าวิธีการของ Hauck and Anderson เมื่อพิจารณาค่าสัมประสิทธิ์ความเชื่อมั่นและค่าความยาวเฉลี่ยของช่วงเชื่อมั่น

อย่างไรก็ตามงานวิจัยของ Hauck and Anderson กับงานวิจัยของ Peskan มีขอบเขตไม่เหมือนกัน ทั้งในเรื่องขนาดของตัวอย่าง ค่าสัดส่วนของลักษณะที่สนใจในประชากร ทำให้ไม่สามารถสรุปได้ถึงความเหมาะสมในกรณีทั่วๆไป ดังนั้นจึงมีงานวิจัยในประเทศไทยที่ทำการศึกษาเปรียบเทียบวิธีการเหล่านี้

ธาริณี คงคาธเนศ (2539) ได้ทำวิทยานิพนธ์เรื่องการเปรียบเทียบวิธีการประมาณแบบช่วงสำหรับผลต่างระหว่างค่าสัดส่วนของสองประชากร โดยทำการเปรียบเทียบวิธีการประมาณ 4 วิธีได้แก่วิธีการประมาณแบบฉบับ (The Classical Method) วิธีการประมาณโดยใช้ค่าปรับแก้เพื่อความต่อเนื่องของ Hauck and Anderson (The Estimation Method Using Continuity Correction by Hauck And Anderson) วิธีการประมาณโดยใช้ค่าปรับแก้เพื่อความต่อเนื่องของ Peskan (The Estimation Method Using Continuity Correction by Peskun) รวมทั้งวิธีการประมาณโดยใช้ค่าปรับแก้เพื่อความต่อเนื่องของ Yate (The Estimation Method Using Continuity Correction by Yate) ผลการวิจัยพบว่า ในกรณีส่วนใหญ่วิธีที่ใช้ค่าปรับแก้เพื่อความต่อเนื่องจะให้ค่าสัมประสิทธิ์ความเชื่อมั่นจากการทดลองไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด และเมื่อพิจารณาค่าความยาวเฉลี่ยของช่วงเชื่อมั่น สำหรับในกรณีขนาดตัวอย่างทั้งสองมีขนาดเล็ก วิธีการประมาณของ Peskun จะให้ค่าความยาวเฉลี่ยต่ำสุด และวิธีการประมาณของ Hauck and Anderson จะให้ค่าความยาวเฉลี่ยต่ำสุดในกรณีที่ขนาดตัวอย่างทั้งสองมีขนาดปานกลาง ส่วนวิธีการประมาณแบบฉบับจะให้ค่าความยาวเฉลี่ยต่ำสุดเมื่อขนาดตัวอย่างทั้งสองมีขนาดใหญ่

จารุภา ตรีระพงศ์ (2541) ได้ทำวิทยานิพนธ์เรื่อง การศึกษาเปรียบเทียบวิธีการประมาณแบบช่วงสำหรับผลต่างระหว่างค่าสัดส่วนของสองประชากร ได้แก่ วิธีการประมาณโดยใช้ค่าปรับแก้เพื่อความต่อเนื่องของ Hauck and Anderson (The Estimation Method Using Continuity Correction by Hauck And Anderson) วิธีการประมาณโดยใช้ค่าปรับแก้เพื่อความต่อเนื่องของ Peskan (The Estimation Method Using Continuity Correction by Peskun) และวิธีการประมาณแบบฉบับ (The Classical Method) ผลของการวิจัย พบว่า วิธีที่ใช้ค่าปรับแก้เพื่อความต่อเนื่องจะให้ค่าสัมประสิทธิ์ความเชื่อมั่นจากการทดลองไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด และเมื่อพิจารณาค่าความยาวเฉลี่ยของช่วงเชื่อมั่น วิธีการประมาณแบบฉบับให้ค่าความยาวเฉลี่ยต่ำสุด รองลงมาคือ วิธีการประมาณโดยใช้ค่าปรับแก้เพื่อความต่อเนื่องของ Hauck and Anderson และวิธีการประมาณโดยใช้ค่าปรับแก้เพื่อความต่อเนื่องของ Peskan ให้ค่าความยาวเฉลี่ยสูงสุด เมื่อพิจารณาทั้งค่าสัมประสิทธิ์ความเชื่อมั่นและค่าความยาวเฉลี่ยของช่วงเชื่อมั่น สรุปได้ว่า วิธีการประมาณโดยใช้ค่าปรับแก้เพื่อความต่อเนื่องของ Hauck and Anderson มีความเหมาะสมที่สุดในการใช้ประมาณผลต่างระหว่างค่าสัดส่วนของสองประชากร

นอกจากนี้ยังมีงานวิจัยต่างประเทศที่ทำการศึกษาพร้อมทั้งนำเสนอวิธีการใหม่ และเปรียบเทียบวิธีการประมาณหลายๆวิธี นอกเหนือจากการใช้ค่าปรับแก้เพื่อความต่อเนื่อง เช่น

Newcombe (1998) ได้ทำการศึกษาวิธีการสร้างช่วงความเชื่อมั่นสำหรับผลต่างระหว่างค่าสัดส่วนของลักษณะที่สนใจในประชากร 2 กลุ่ม โดยได้เสนอวิธีการประมาณของ Newcombe (The Newcombe's Confidence Interval) ซึ่งเป็นวิธีการประมาณที่ขั้นตอนแรก คือ หาช่วงความเชื่อมั่นของค่าสัดส่วนของแต่ละประชากรโดยใช้วิธีรากของสมการกำลังสอง ขั้นตอนต่อไป คือ หาช่วงความเชื่อมั่นของผลต่างค่าสัดส่วนของสองประชากร (L,U) โดยกำหนดให้ $L=\hat{\theta}-\delta$ และ $U=\hat{\theta}-\varepsilon$ โดยที่ δ เป็นระยะทางระหว่างจุด (l_1,u_2) และจุด $(\hat{p}_1,\hat{p}_2)$ และ ε เป็นระยะทางระหว่างจุด (l_2,u_1) และจุด $(\hat{p}_1,\hat{p}_2)$ จากผลการวิจัยที่ได้พบว่า วิธีการใหม่ที่เขาเสนอนั้นเป็นวิธีการที่เหมาะสมเพราะในกรณีส่วนใหญ่แล้วเป็นวิธีที่ให้ค่าสัมประสิทธิ์ความเชื่อมั่นจากการทดลองไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด และให้ค่าความยาวเฉลี่ยต่ำสุดเมื่อค่า p มีค่าเข้าใกล้ 0.5

Jeffrey (2001) ได้เสนอวิธีการประมาณค่าแบบช่วงสำหรับค่าสัดส่วนประชากร คือ วิธีการประมาณของ Jeffrey (The Jeffrey 's Confidence Interval) ที่นำแนวคิดของวิธีการเบส์มาประยุกต์ใช้ในการประมาณค่าแบบช่วง ซึ่งเขาเสนอว่าควรมีการแจกแจงแรกเริ่มเป็น $Beta(1/2,1/2)$ ซึ่งทำให้ได้ค่าประมาณแบบช่วงของค่าสัดส่วนที่เหมาะสม และจากแนวคิดนี้สามารถนำมาประยุกต์ใช้ในกรณีการประมาณค่าแบบช่วงสำหรับผลต่างของค่าสัดส่วน 2 ประชากรได้ จากสูตรช่วงความเชื่อมั่นที่ได้เห็นได้ว่าเป็นวิธีที่สามารถทำการคำนวณได้ง่าย และมีประสิทธิภาพใกล้เคียงกับวิธีการอื่นที่มีขั้นตอนการคำนวณยุ่งยากกว่า

Lawrence D. Brown และ Xuefeng Li (2003) ได้เสนอวิธีการประมาณช่วงความเชื่อมั่นอีกวิธีหนึ่งที่น่าสนใจ คือ วิธีการประมาณแบบจัดค่ากลางใหม่ (The Recentered Confidence Interval) ซึ่งเป็นวิธีที่ทำการสร้างค่าพารามิเตอร์ p ใหม่จากค่าสัดส่วนประชากรของประชากรที่ 1 และ 2 (p_1 และ p_2) แล้วนำมาหาช่วงความเชื่อมั่น โดยวิธีการนี้เป็นวิธีที่เหมาะสมเนื่องจากส่วนใหญ่จะให้ค่าสัมประสิทธิ์ความเชื่อมั่นจากการทดลองไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนดและให้ค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นที่ต่ำ ยกเว้นกรณีเมื่อขนาดตัวอย่างทั้งสองมีขนาดไม่เท่ากันและผลต่างระหว่างค่าสัดส่วนตัวอย่างมีค่าเข้าใกล้ 0

จากการวิจัยที่กล่าวในข้างต้น และด้วยเหตุที่ยังไม่พบงานวิจัยที่ทำการศึกษาเปรียบเทียบประสิทธิภาพของวิธีการประมาณช่วงความเชื่อมั่นสำหรับผลต่างระหว่างค่าสัดส่วนในประชากร 2 กลุ่ม ด้วยวิธีการทั้ง 3 วิธีข้างต้นคือ วิธีการประมาณของ Newcombe วิธีการประมาณของ Jeffrey วิธีการประมาณแบบจัดค่ากลางใหม่ และวิธีการประมาณแบบฉบับพร้อมกัน ผู้วิจัยจึงมี

ความสนใจที่จะทำการศึกษาการประมาณช่วงความเชื่อมั่นสำหรับผลต่างระหว่างค่าสัดส่วนในประชากร 2 กลุ่ม และเปรียบเทียบวิธีการประมาณช่วงความเชื่อมั่น 4 วิธีดังนี้

1. วิธีการประมาณแบบฉบับ (The Classical Method)
2. วิธีการประมาณของนิวคอมบ์ (The Newcombe's Confidence Interval)
3. วิธีการประมาณของเจฟฟรีย์ (The Jeffrey 's Confidence Interval)
4. วิธีการประมาณแบบจัดค่ากลางใหม่ (The Recentered Confidence Interval)

ผู้วิจัยสนใจที่จะทำการเปรียบเทียบประสิทธิภาพของวิธีการประมาณช่วงความเชื่อมั่นดังกล่าว เพื่อศึกษาหาข้อสรุปในการนำวิธีการประมาณแต่ละวิธีที่มีความเหมาะสมไปใช้ในแต่ละสถานการณ์ต่อไป

1.2 **วัตถุประสงค์ของการวิจัย**

ในการวิจัยนี้มีวัตถุประสงค์ดังนี้

1. เพื่อศึกษาวิธีการประมาณแบบช่วงสำหรับผลต่างระหว่างค่าสัดส่วนของลักษณะที่สนใจในประชากร 2 กลุ่มที่มีการแจกแจงแบบแบร์นูลลีและเป็นอิสระต่อกัน จำนวน 4 วิธีดังนี้
   - วิธีการประมาณแบบฉบับ (The Classical Method)
   - วิธีการประมาณของนิวคอมบ์ (The Newcombe's Confidence Interval)
   - วิธีการประมาณของเจฟฟรีย์ (The Jeffrey 's Confidence Interval)
   - วิธีการประมาณแบบจัดค่ากลางใหม่ (The Recentered Confidence Interval)
2. เปรียบเทียบสัมประสิทธิ์ความเชื่อมั่น และความยาวเฉลี่ยของช่วงเชื่อมั่นที่คำนวณได้จากแต่ละวิธี ที่ระดับความเชื่อมั่น 90%, 95% และ 99%

1.3 **สมมติฐานของการวิจัย**

ในการวิจัยนี้มีวัตถุประสงค์ดังนี้

การประมาณค่าแบบช่วงสำหรับผลต่างระหว่างค่าสัดส่วนของสองประชากรที่เป็นอิสระกันโดยใช้วิธีการประมาณแบบจัดค่ากลางใหม่จะให้ค่าสัมประสิทธิ์ความเชื่อมันไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด และให้ค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นต่ำกว่าการประมาณวิธีอื่นๆ

1.4 ข้อตกลงเบื้องต้น

การวิจัยครั้งนี้มีข้อตกลงเบื้องต้นสำหรับการดำเนินงานวิจัย ดังนี้

1) ในแต่ละประชากรมีการแจกแจงแบบแบร์นูลลีที่เป็นอิสระกัน

2) ขนาดตัวอย่าง n_1, n_2 ในประชากรที่ 1 และ 2 เป็นข้อมูลที่ทราบค่า

3) ค่าสัดส่วนความสำเร็จของแต่ละประชากร (p_1, p_2) เป็นพารามิเตอร์ที่ไม่ทราบค่า

1.5 ขอบเขตการวิจัย

ในการศึกษาวิธีการประมาณค่าแบบช่วงสำหรับผลต่างระหว่างค่าสัดส่วนของลักษณะที่สนใจในประชากร 2 กลุ่มที่เป็นอิสระกัน ให้ขนาดตัวอย่างที่สุ่มมาจากประชากรกลุ่มที่ 1 และ 2 คือ n_1 และ n_2 ตามลำดับและค่าความน่าจะเป็นที่เกิดลักษณะที่สนใจในประชากรกลุ่มที่ 1 และ 2 คือ p_1 และ p_2 ตามลำดับ และกำหนดขอบเขตของการวิจัยดังนี้

1. กำหนดให้ขนาดตัวอย่างของประชากรที่ 2 (n_2) มีค่าเท่ากับ 5, 10, 20, 30, 40, 50, 60, 70 และพิจารณาจำแนกเป็น 2 กรณี ตามความแตกต่างของขนาดตัวอย่างทั้งสอง

1) ขนาดตัวอย่างของสองประชากรมีค่าเท่ากัน $(n_1 = n_2)$

2) ขนาดตัวอย่างของสองประชากรมีค่าไม่เท่ากัน โดยที่ขนาดตัวอย่างของประชากรที่ 1 มีค่ามากกว่าขนาดตัวอย่างของประชากรที่ 2 $(n_1 > n_2)$จำแนกตามเปอร์เซ็นต์ผลต่างระหว่างขนาดตัวอย่างของสองประชากรซึ่งมีค่าเท่ากับ 20%, 40%, 60%, 80%, 100%, 140%, 180% และ 200% (รายละเอียดอยู่ในหัวข้อที่ 3.1)

2. กำหนดค่า p_1, p_2 ให้มีค่าตั้งแต่ 0.1 ถึง 0.9 โดยเพิ่มค่าทีละ 0.1 ซึ่งจะให้ค่าผลต่างระหว่างค่าสัดส่วนสองประชากร $(p_1 - p_2)$ มีความแตกต่างกันตั้งแต่ 0 ถึง 0.8 โดยเพิ่มค่าทีละ 0.1 และกำหนดให้ $p_1 \geq p_2$ (รายละเอียดอยู่ในหัวข้อที่ 3.1)

3. กำหนดค่าสัมประสิทธิ์ความเชื่อมั่น $(1-\alpha)100\%$ คือ 90 % , 95% , 99%

1.6 เกณฑ์ที่ใช้ในการประเมินประสิทธิภาพของช่วงความเชื่อมั่น

หลักเกณฑ์ในการพิจารณาเปรียบเทียบประสิทธิภาพของวิธีการประมาณค่าแบบช่วงสำหรับผลต่างของค่าสัดส่วนในประชากร 2 กลุ่มมีดังนี้

1) เปรียบเทียบค่าสัมประสิทธ์ความเชื่อมั่น (Confidence Coefficient) ที่คำนวณได้จากแต่ละวิธีการประมาณกับค่าสัมประสิทธิ์ที่กำหนด ในการตรวจสอบว่าวิธีการใดให้ค่าสัมประสิทธิ์ความเชื่อมั่นจากการทดลองไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนดได้หรือไม่นั้น ผู้วิจัยอาศัยการทดสอบสมมติฐานโดยใช้ตัวสถิติ Z ที่ระดับนัยสำคัญ 0.05 (รายละเอียดอยู่ในหัวข้อที่ 2.5)

2) ค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นที่ต่ำกว่าเป็นวิธีการที่มีประสิทธิภาพมากกว่า ทั้งนี้ในการเปรียบเทียบค่าความยาวเฉลี่ยของช่วงความเชื่อมั่น จะเปรียบเทียบเฉพาะในกรณีวิธีการประมาณที่ให้ค่าสัมประสิทธิ์ความเชื่อมั่นจากการทดลองไม่ต่ำกว่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด

1.7 ประโยชน์ที่คาดว่าจะได้รับ

ประโยชน์ที่คาดว่าจะได้รับในการวิจัยครั้งนี้

1) ทำให้ทราบประสิทธิภาพของวิธีการประมาณแต่ละวิธีในการประมาณค่าแบบช่วงสำหรับผลต่างระหว่างค่าสัดส่วนของสองประชากรที่เป็นอิสระกัน ภายใต้สถานการณ์ต่างๆ

2) ผลที่ได้จากการวิจัยครั้งนี้จะเป็นแนวทางในการเลือกใช้วิธีการประมาณค่าแบบช่วงสำหรับผลต่างระหว่างค่าสัดส่วนของสองประชากรที่เป็นอิสระกันให้เหมาะสมสำหรับแต่ละสถานการณ์

3) เป็นแนวทางในการศึกษาเปรียบเทียบหาวิธีการประมาณค่าแบบช่วงด้วยวิธีอื่นๆ สำหรับผลต่างระหว่างค่าสัดส่วนของสองประชากรต่อไป