บทที่ 1
บทนำ

1.1. ที่มาและความสำคัญของปัญหา

ปัจจุบันนี้วิชาสถิติมีความสำคัญอย่างมากในงานวิจัยต่างๆโดยเฉพาะทางด้านวิทยาศาสตร์ วิศวกรรมศาสตร์ และสังคมศาสตร์ ดังนั้นผู้วิจัยจะต้องมีความรู้ทางด้านสถิติพอสมควรโดยเฉพาะขั้นตอนในการวิเคราะห์และการสรุปผลการทดลองซึ่งเป็นขั้นตอนที่ต้องใช้ความแม่นยำจากการคำนวณ ในชีวิตประจำวันคอมพิวเตอร์จึงมีบทบาทต่อการวิเคราะห์ผลทางสถิติสูงหรืออาจกล่าวได้ว่าเป็นปัจจัยที่จำเป็นในการวิเคราะห์ข้อมูล เนื่องจากในปัจจุบันข้อมูลที่นำมาวิเคราะห์มีจำนวนมากและต้องการความเร็วที่ใช้ในการคำนวณสูงเนื่องจากมีการแข่งขันด้านความเร็วมากขึ้นเรื่อย ๆ ทำให้การวิเคราะห์ข้อมูลในปัจจุบันไม่สามารถคำนวณด้วยมือได้ การวิจัยครั้งนี้ก็เช่นเดียวกันเป็นการคำนวณที่มีความซับซ้อนและต้องนำคอมพิวเตอร์มาใช้โดยการวิจัยนี้จะทำการคำนวณฟังก์ชันการแจกแจงสะสมไคกำลังสองไร้ศูนย์กลาง การคำนวณนี้มีความสำคัญในด้านการทดสอบสมมติฐานทางสถิติซึ่งในการทดสอบสมมติฐานทางสถิติของกรณีการวิเคราะห์ข้อมูลพหุนาม (multinomial data analysis) เช่น การทดสอบความเป็นอิสระ(test of independence) การทดสอบภาวะเอกพันธ์ (test of homogeneity) การทดสอบภาวะสารูปสนิทดี (goodness of fit test) เพื่อหาอำนาจการทดสอบ (power of the test) ภายใต้สมมติฐานแย้ง $H_a: \mu_i \neq \mu_j$ สำหรับทุกค่า $i \neq j$ เป็นต้น การทดสอบสมมติฐานดังกล่าวข้างต้นจะกระทำการได้โดยใช้การแจกแจงไคกำลังสองไร้ศูนย์กลาง

การวิจัยครั้งนี้จะนำฟังก์ชันการแจกแจงสะสมไคกำลังสองไร้ศูนย์กลางมาใช้ในการคำนวณเนื่องจากฟังก์ชันนี้มีประโยชน์ในการหาบริเวณวิกฤต(critical region) ในปัจจุบันโปรแกรมที่ใช้ในการคำนวณการแจกแจงสะสมไคกำลังสองไร้ศูนย์กลางนั้นมีหลายโปรแกรมซึ่งใช้ฟังก์ชันการแจกแจงสะสมในการคำนวณต่างกัน ดังนั้นการพิจารณาประสิทธิภาพในการคำนวณของโปรแกรมที่ใช้จึงมีความจำเป็นเป็นอย่างยิ่ง เนื่องจากอนุกรมใดที่มีความเร็วในการคำนวณสูงและให้ค่าความคลาดเคลื่อนต่ำก็ถือว่าเป็นโปรแกรมที่มีประสิทธิภาพดีและเป็นการใช้ทรัพยากรอย่างประหยัดและคุ้มค่า

ในปีค.ศ.1990 แอชอาวร์และแอ็บเดล-ซาแมด (Ashour & Abdel-Samad) ได้ทำการเปรียบเทียบวิธีการคำนวณฟังก์ชันการแจกแจงสะสมไคกำลังสองไร้ศูนย์กลางโดยพิจารณากรณีที่องศาความเป็นอิสระเป็นค่าใด ๆ กับวิธีการคำนวณของแพ็ตเนค (Patnaik) ซึ่งคิดในปี ค.ศ. 1949 และในปีค.ศ.1995 มีการวิจัยเกี่ยวกับประสิทธิภาพในการคำนวณการแจกแจงสะสมไคกำลังสองไร้ศูนย์กลางโดยพิจารณากรณีที่องศาความเป็นอิสระเป็นเลขคู่ การวิจัยครั้งนั้นได้ทำการเปรียบเทียบวิธีการคำนวณของแอชอาวร์และแอ็บเดล-ซาแมด รูเบ็น(Ruben) ซึ่งคิดในปีค.ศ.1974 และซีเกิล (Siegel) ซึ่งคิดในปีค.ศ.1979 แต่การวิจัยครั้งนี้จะทำการเปรียบเทียบประสิทธิภาพในการคำนวณของฟังก์ชันการแจกแจงสะสมไคกำลังสองไร้ศูนย์กลางกรณีที่องศาความเป็นอิสระเป็นเลขคี่โดยเปรียบเทียบวิธีการคำนวณค่าแต่ละวิธีดังนี้

วิธีที่ 1 วิธีนี้เป็นวิธีการคำนวณค่าฟังก์ชันการแจกแจงสะสมไคกำลังสองไร้ศูนย์กลางโดยวิธีการของแพ็ตเนค(ค่าจริง) ซึ่งใช้ในการคำนวณค่าฟังก์ชันการแจกแจงสะสมไคกำลังสองไร้ศูนย์กลางในกรณีที่องศาความเป็นอิสระ เป็นค่าใด ๆ ในการวิจัยครั้งนี้ใช้วิธีนี้เป็นตัวควบคุม

วิธีที่ 2 วิธีนี้เป็นวิธีการคำนวณค่าฟังก์ชันการแจกแจงสะสมไคกำลังสองไร้ศูนย์กลางโดยวิธีการของแอชอาวร์และแอ็บเดล-ซาแมด ซึ่งวิธีการคำนวณวิธีนี้ปรับมาจากวิธีที่ 1 และใช้ในการคำนวณค่าฟังก์ชันการแจกแจงสะสมไคกำลังสองไร้ศูนย์กลางในกรณีที่องศาความเป็นอิสระเป็นค่าใด ๆ

วิธีที่ 3 วิธีนี้เป็นวิธีการคำนวณค่าฟังก์ชันการแจกแจงสะสมไคกำลังสองไร้ศูนย์กลางโดยวิธีการของแอชอาวร์และแอ็บเดล-ซาแมด ซึ่งใช้ในการคำนวณค่าฟังก์ชันการแจกแจงสะสมไคกำ-ลังสองไร้ศูนย์กลางในกรณีที่องศาความเป็นอิสระเป็นเป็นเลขคี่

วิธีที่ 4 วิธีนี้เป็นวิธีการคำนวณค่าฟังก์ชันการแจกแจงสะสมไคกำลังสองไร้ศูนย์กลางโดยวิธีการของรูเบ็น เบิก และโกวินดาราจูลู ซึ่งใช้ในการคำนวณค่าฟังก์ชันการแจกแจงสะสมไคกำลังสองไร้ศูนย์กลางในกรณีที่องศาความเป็นอิสระเป็นเลขคี่

สำหรับการวัดประสิทธิภาพของฟังก์ชันที่ใช้คำนวณทั้ง 4 วิธีนั้นจะวัดจากค่าความคลาด-เคลื่อนและเวลาที่ใช้ในการคำนวณ

1.2. วัตถุประสงค์ของการวิจัย

เพื่อเปรียบเทียบวิธีการประมาณค่าฟังก์ชันการแจกแจงสะสมไคกำลังสองไร้ศูนย์กลาง ระหว่างวิธีการคำนวณของแพ็ตเนคซึ่งใช้ในกรณีที่องศาความเป็นอิสระเป็นค่าใด ๆ (วิธีที่ 1) ซึ่งใช้เป็นตัวควบคุมกับวิธีการของแอชอาวร์และแอ็บเดล-ซาเมคซึ่งใช้ในกรณีที่องศาความเป็นอิสระเป็นค่าใด ๆ (วิธีที่ 2) วิธีการของแอชอาวร์และแอ็บเดล-ซาเมคซึ่งใช้ในกรณีที่องศาความเป็นอิสระเป็นเลขคี่ (วิธีที่ 3) และ วิธีการของรูเบ็น เบิก และโกวินดาราจูลูซึ่งใช้ในกรณีที่องศาความเป็นอิสระเป็นเลขคี่ (วิธีที่ 4) ว่าวิธีใดมีประสิทธิภาพในการคำนวณดีกล่าวคือวิธีใดที่ให้ค่าความคลาดเคลื่อนต่ำและทำการคำนวณได้เร็วที่สุด

1.3. สมมติฐานการวิจัย

เมื่อพิจารณาเปรียบเทียบกับค่าจริงซึ่งได้จากการคำนวณวิธีการของแพ็ตเนค (วิธีที่ 1) เป็นตัวควบคุม วิธีการประมาณค่าฟังก์ชันการแจกแจงสะสมของแอชอาวร์และแอ็บเดล-ซาเมคซึ่งใช้ในกรณีองศาความเป็นอิสระเป็นค่าใด ๆ (วิธีที่ 2) จะมีความคลาดเคลื่อนน้อยที่สุด รองลงมาคือวิธีการประมาณค่าฟังก์ชันการแจกแจงสะสมของแอชอาวร์และแอ็บเดล-ซาเมคซึ่งใช้ในกรณีองศาความเป็นอิสระเป็นเลขคี่ (วิธีที่ 3) และวิธีการประมาณค่าฟังก์ชันการแจกแจงสะสมของรูเบ็น เบิก และโกวินดาราจูลู (วิธีที่ 4) จะมีความคลาดเคลื่อนมากที่สุด แต่อย่างไรก็ตามในการพิจารณาวัดประสิทธิภาพของวิธีการคำนวณแต่ละวิธีนั้นวิธีที่มีความคลาดเคลื่อนน้อยอาจใช้เวลาในการคำนวณสูงจึงพิจารณาร่วมกับเวลาที่ใช้ในการคำนวณโดยที่วิธีที่ 4 จะใช้เวลาในการคำนวณน้อยที่สุด รองลงมาคือวิธีที่ 2 ส่วนวิธีที่ 3 จะใช้เวลาในการคำนวณมากที่สุด

1.4. ข้อตกลงเบื้องต้น

กำหนดให้ $\underset{\sim}{x}$ มีการแจกแจงปกติ นั่นคือ $\underset{\sim}{x} \sim N_n(\underset{\sim}{\mu}, I_n)$, $\underset{\sim}{\mu} \neq \underset{\sim}{0}$ ดังนั้น $u = \underset{\sim}{x}'\underset{\sim}{x}$ มีการแจกแจงไคกำลังสองไร้ศูนย์กลาง $\chi^2(n;\delta)$ โดยมีค่าเฉลี่ย $n+2\delta$ และความแปรปรวน $2n+8\delta$

1.5. ขอบเขตการวิจัย

1.5.1. การวิจัยนี้จะพิจารณาการประมาณค่าฟังก์ชันการแจกแจงสะสมไคกำลังสองไร้ศูนย์กลางกรณีที่องศาความเป็นอิสระเป็นเลขคี่เนื่องจากวิธีที่ 3 และวิธีที่ 4 สามารถคำนวณได้เฉพาะกรณีองศาความเป็นอิสระเป็นเลขคี่ โดยกำหนดให้องศาความเป็นอิสระเท่ากับ 1, 3, 5, 9, 15, 21, 25, 35, 49 และ 99

1.5.2. การวิจัยนี้จะพิจารณาการประมาณค่าฟังก์ชันการแจกแจงสะสมไคกำลังสองไร้ศูนย์กลางและค่าฟังก์ชันการแจกแจงสะสมไคกำลังสองไร้ศูนย์กลางโดยที่จะพิจารณาวิธีการประมาณค่าฟังก์ชันการแจกแจงสะสมไคกำลังสองไร้ศูนย์กลางของแอชอาวร์และแอ็บเดล-ชาแมด (วิธีที่ 2 และวิธีที่ 3) รูเบิน เบิก และโกวินดาราจูลู(วิธีที่ 4) กับค่าฟังก์ชันการแจกแจงสะสมไคกำลังสองไร้ศูนย์กลางของแพ็ตเนค(วิธีที่ 1)

1.5.2.1.ฟังก์ชันการแจกแจงสะสมไคกำลังสองไร้ศูนย์กลางของแพ็ตเนคเป็นฟังก์ชันที่ใช้ในการคำนวณค่าฟังก์ชันการแจกแจงสะสมไคกำลังสองไร้ศูนย์กลางกรณีที่องศาความเป็นอิสระเป็นค่าใด ๆ ซึ่งแสดงได้ดังสมการที่ (1.1)

$$F_n^{\delta}(x)=\frac{e^{-\frac{1}{2}(x+\delta)}x^{\frac{1}{2}n-1}}{2^{\frac{1}{2}n}\Gamma(\frac{1}{2})n}\{1+\frac{1}{n}\left(\frac{x\delta}{2}\right)+\frac{1}{n(n+2).2!}\left(\frac{x\delta}{2}\right)^2+\ldots\} \tag{1.1}$$

1.5.2.2.ฟังก์ชันการแจกแจงสะสมไคกำลังสองไร้ศูนย์กลางของแอชอาวร์และแอ็บเดล-ชาแมดเป็นฟังก์ชันที่ใช้ในการคำนวณค่าฟังก์ชันการแจกแจงสะสมไคกำลังสองไร้ศูนย์กลางกรณีที่องศาความเป็นอิสระเป็นค่าใด ๆ ซึ่งแสดงได้ดังสมการที่ (1.2)

$$F_n^{\delta}(x)=e^{-\delta/2}f_n(x)\sum_{i=0}^{\infty}\frac{C_i(z,m)}{i!}\sum_{j=0}^{\infty}C_j(y,m+i) \tag{1.2}$$

เมื่อ $f_n(x)=\dfrac{e^{-x/2}(\frac{x}{2})^{n/2}}{\Gamma(\frac{n}{2}+1)}$, $C_j(z,k)=\dfrac{z}{(k+j)C_{j-1}(z,k)}$, $j=1,2,\ldots$

$C_0(z,k)=$, $z=\dfrac{\delta y}{2}$, $y=\dfrac{x}{2}$ และ $m=\dfrac{n}{2}$

1.5.2.3.ฟังก์ชันการแจกแจงสะสมไคกำลังสองไร้ศูนย์กลางของแอชอาวร์และแอ็บเดล-ซาแมดเป็นฟังก์ชันที่ใช้ในการคำนวณค่าฟังก์ชันการแจกแจงสะสมไคกำลังสองไร้ศูนย์กลางกรณีที่องศาความเป็นอิสระเป็นเลขคี่ ซึ่งแสดงได้ดังสมการที่ (1.3)

$$F_{2m+1}^{\delta}(x) = F_{2m+1}(x) + \left(\frac{2}{\pi}\right)^{\frac{1}{2}} e^{-\frac{x}{2}} . \{ \sum_{r=m+1}^{\infty} \frac{x^{r-\frac{1}{2}}}{1.3.5\ldots(2r-1)} F_{2r-2m}(\delta) \} \tag{1.3}$$

เมื่อ $n = 2m+1$

1.5.2.4.ฟังก์ชันการแจกแจงสะสมไคกำลังสองไร้ศูนย์กลางของ รูเบ็น เบิก และโกวินดาราจูลูเป็นฟังก์ชันที่ใช้ในการคำนวณค่าฟังก์ชันการแจกแจงสะสมไคกำลังสองไร้ศูนย์กลางกรณีที่องศาความเป็นอิสระเป็นเลขคี่ ซึ่งแสดงได้ดังสมการที่ (1.4)

$$F_{n}^{\delta}(x) = F_{1}^{\delta}(x) - e^{-(\delta+x)/2} \sum_{j=1}^{k} (x/\delta)^{(2j-1)/4} I_{j-(1/2)}(\sqrt{\delta x}) \tag{1.4}$$

เมื่อ $F_1^{\delta}(x) = \Phi(a) - \Phi(b)$ และ $\Phi(.)$ แทนฟังก์ชันการแจกแจงสะสมของการแจกแจงปกติมาตรฐาน

$$a = \sqrt{\lambda} + \sqrt{x} \text{ , } b = \sqrt{\lambda} - \sqrt{x} \text{ , } I_{v+1}(y) = I_{v-1}(y) - (\frac{2v}{y}) I_v(y) \text{ , } n = 2k+1 \text{ ,}$$

$$I_{\frac{1}{2}}(y) = \sqrt{\frac{2}{\pi y}} \sinh y \quad \text{และ} \quad I_{\frac{3}{2}}(y) = \sqrt{\frac{2}{\pi y}} [\cosh y - (\sinh y)/y]$$

1.6. ประโยชน์ที่คาดว่าจะได้รับ

1.ผู้อ่านสามารถเลือกใช้วิธีการประมาณค่าฟังก์ชันการแจกแจงสะสมของการแจกแจงไค-กำลังสองไร้ศูนย์กลางในแต่ละสถานการณ์ได้อย่างเหมาะสม

2.การวิจัยครั้งนี้สามารถใช้เป็นแนวทางในการค้นหาและปรับปรุงประสิทธิภาพของวิธีการประมาณค่าฟังก์ชันการแจกแจงสะสมไคกำลังสองไร้ศูนย์กลางซึ่งเหมาะสมในอนาคตได้

1.7. เกณฑ์ในการตัดสินใจ

เกณฑ์ที่ใช้ในการตัดสินใจว่าค่าประมาณจากฟังก์ชันการแจกแจงสะสมไคกำลังสองไร้ศูนย์กลางจะพิจารณาเป็น 2 เกณฑ์ด้วยกันคือ เกณฑ์ความคลาดเคลื่อนและเกณฑ์เวลา โดยในขั้นแรกจะพิจารณาจากค่าความคลาดเคลื่อนจากวิธีการคำนวณค่าประมาณฟังก์ชันการแจกแจงสะสมไคกำลังสองไร้ศูนย์กลางกับค่าจริง แต่ถ้าผลลัพธ์ที่ได้ไม่ชัดเจนจะพิจารณาเวลาที่ใช้ในการคำนวณโดยวิธีอื่น ๆเทียบกับวิธีที่ใช้เวลาน้อยที่สุด

ในการพิจารณาเกณฑ์ความคลาดเคลื่อนจะพิจารณาค่าความคลาดเคลื่อนสัมพัทธ์ของแต่ละวิธี โดยพิจารณาว่าวิธีใดที่มีค่าความคลาดเคลื่อนสัมพัทธ์ต่ำที่สุดจะเป็นวิธีที่ดีที่สุด เกณฑ์นี้จะใช้ในการตัดสินใจว่าค่าประมาณจากฟังก์ชันการแจกแจงสะสมใดที่มีความถูกต้องมากที่สุดซึ่งจะพิจารณาเปรียบเทียบค่าความคลาดเคลื่อนของค่าประมาณโดยวิธีของแอชอาวร์และแอ็บเดล-ซาแมด รูเบ็น เบิก และโกวินดาราจูลู กับค่าประมาณโดยวิธีของแพ็ตเนค สูตรที่ใช้ในการคำนวณค่า RE อาจแสดงได้ดังนี้

สูตรที่ใช้ในการคำนวณค่า RE คือ

$$RE = \frac{\left|\hat{\theta}_i - \hat{\theta}_0\right|}{\hat{\theta}_0} \times 100$$

เมื่อ $\hat{\theta}_i$ หมายถึง ค่าที่ได้จากการประมาณโดยวิธีอื่น ๆ

และ $\hat{\theta}_0$ หมายถึง ค่าที่ได้จากการประมาณโดยวิธีของแพ็ตเนค

ส่วนการพิจารณาว่าวิธีใดมีประสิทธิภาพสูงที่สุดจะพิจารณาจากการนำค่าความคลาดเคลื่อนสัมพัทธ์แต่ละค่าที่คำนวณได้มาเฉลี่ยเรียกว่าค่าความคลาดเคลื่อนสัมพัทธ์เฉลี่ย(average relative error (ARE)) และถ้าค่าเฉลี่ยความคลาดเคลื่อนสัมพัทธ์ของวิธีใดที่มีค่าน้อยที่สุดจะเป็นวิธีที่มีประสิทธิภาพดี สูตรที่ใช้ในการคำนวณค่า ARE อาจแสดงได้ดังนี้

สูตรที่ใช้ในการคำนวณค่าความมีประสิทธิภาพคือ

$$ARE = \frac{\sum_{i=1}^{p} RE}{p}$$

เมื่อ ARE หมายถึง ค่าความคลาดเคลื่อนสัมพัทธ์เฉลี่ย (Average Relative Error)
และ p หมายถึง จำนวนตัวอย่างที่ใช้ในการเปรียบเทียบ

เกณฑ์ที่สองที่ใช้ในการพิจารณาค่าประมาณจากฟังก์ชันการแจกแจงสะสมไคกำลังสองไร้ศูนย์กลางคือ เกณฑ์เวลาจะพิจารณาค่าเวลาสัมพัทธ์(relative time (RT))ที่ใช้ในการคำนวณของแต่ละวิธี โดยพิจารณาว่าวิธีใดมีค่าเวลาสัมพัทธ์ต่ำที่สุดจะเป็นวิธีที่ดีที่สุด เกณฑ์นี้จะใช้ในการตัดสิน-ใจว่าค่าประมาณจากฟังก์ชันการแจกแจงสะสมใดที่มีความเร็วในการคำนวณสูงที่สุดซึ่งจะพิจารณาเปรียบเทียบเวลาที่ใช้ในการคำนวณของแต่ละวิธีกับวิธีที่ใช้เวลาในการคำนวณต่ำที่สุด

สูตรที่ใช้ในการคำนวณค่า RT อาจแสดงได้ดังนี้

สูตรที่ใช้ในการคำนวณค่า RT คือ

$$RT = \frac{\left|\hat{T}_i - \hat{T}_{\min.}\right|}{\hat{T}_{\min.}} \times 100$$

เมื่อ $\hat{T}_i$ หมายถึง เวลาที่ใช้ในการคำนวณ โดยวิธีอื่น ๆ
และ $\hat{T}_{\min.}$ หมายถึง เวลาที่ใช้ในการคำนวณ โดยวิธีที่ใช้เวลาน้อยที่สุด

ส่วนการพิจารณาว่าวิธีใดมีประสิทธิภาพสูงที่สุดจะพิจารณาจากการนำค่าเวลาสัมพัทธ์แต่ละค่ามาเฉลี่ยเรียกว่า เวลาสัมพัทธ์เฉลี่ย (average relative time (ART)) และถ้าค่าเฉลี่ยเวลาสัมพัทธ์ของวิธีใดที่มีค่าน้อยที่สุดจะเป็นวิธีที่มีประสิทธิภาพดี สูตรที่ใช้ในการคำนวณค่า ART อาจแสดงได้ดังนี้

สูตรที่ใช้ในการคำนวณค่าความมีประสิทธิภาพคือ

$$ART = \frac{\sum_{i=1}^{p} RT}{p}$$

เมื่อ ART หมายถึง ค่าเวลาสัมพัทธ์เฉลี่ย (Average Relative Time)

และ p หมายถึง จำนวนตัวอย่างที่ใช้ในการเปรียบเทียบ