เรื่อง

การศึกษาประสิทธิภาพของตัวประมาณค่าส่วนเบี่ยงเบนมาตรฐานในการแจกแจงปรกติ

A study of performance of standard deviation estimator in the normal distribution

โดย

นายภูมิพัฒน์	**มีเจริญวรานนท์**	**เลขทะเบียน**	**6109680667**
นางสาวกุลรัศมิ์	**ทรัพย์อุดมมาก**	**เลขทะเบียน**	**6109680725**
นาวสาวธมลวรรณ	**กิจวรวุฒิ**	**เลขทะเบียน**	**6109680758**

รายงานนี้เป็นส่วนหนึ่งของวิชา ส.495 โครงงานพิเศษ 2

ตามหลักสูตรวิทยาศาสตรบัณฑิต สาขาวิชาสถิติ

คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์

ปีการศึกษา 2564

พรรวรรณ แสงทวี
11/พค/2565

หัวข้อโครงงานพิเศษ	การศึกษาประสิทธิภาพของตัวประมาณค่าส่วนเบี่ยงเบนมาตรฐานในการแจกแจงปรกติ
คณะผู้จัดทำ	นายภูมิพัฒน์ มีเจริญวรานนท์ นางสาวกุลรัศมิ์ ทรัพย์อุดมมาก นาวสาวธมลวรรณ กิจวรวุฒิ
ชื่อปริญญา	วิทยาศาสตรบัณฑิต
หลักสูตร/สาขา	หลักสูตรวิทยาศาสตรบัณฑิต สาขาวิชาสถิติ
คณะ/มหาวิทยาลัย	คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์
อาจารย์ที่ปรึกษางานวิจัย	ผศ.ดร.ภทรวรรณ แสงนวกิจ
ปีการศึกษา	2564

บทคัดย่อ

การแจกแจงปรกติเป็นการแจกแจงหนึ่งที่พบได้บ่อยครั้งในการทำงานวิจัย การประมาณค่าส่วนเบี่ยงเบนมาตรฐานในการแจกแจงปรกติที่มีผู้นำเสนอไปแล้วนั้นอาศัยวิธีแบบดั้งเดิมและวิธีภาวะน่าจะเป็นสูงสุด แต่อย่างไรก็ตามผู้วิจัยเห็นว่าการนำเสนอการประมาณค่าส่วนเบี่ยงเบนมาตรฐานจากตัวประมาณที่มีสมบัติเป็นตัวประมาณไม่เอนเอียงจะทำให้ได้ตัวประมาณที่เหมาะสมในการวิเคราะห์ข้อมูลมากขึ้น ดังนั้นในงานวิจัยครั้งนี้จึงได้สนใจสร้างช่วงความเชื่อมั่น ศึกษาประสิทธิภาพ และเปรียบเทียบประสิทธิภาพของตัวประมาณดังกล่าวกับตัวประมาณด้วยวิธีดั้งเดิมและวิธีภาวะน่าจะเป็นสูงสุด ผลที่ได้พบว่าช่วงความเชื่อมั่นในการศึกษาครั้งนี้มีประสิทธิภาพที่ดี คือมีค่าความน่าจะเป็นคุ้มรวมใกล้เคียงกับค่าความน่าจะเป็นที่กำหนดหลายกรณีในการจำลอง นอกจากนั้นตัวประมาณที่ได้นำเสนอนั้นได้ถูกนำไปใช้กับข้อมูลจริงทางด้านราคาน้ำมันซึ่งเกี่ยวข้องกับราคาของน้ำมันแก๊สโซฮอล 91 ผลที่ได้จากการจำลองข้อมูลและสถานการณ์จริงมีความสอดคล้องกัน

คำสำคัญ : วิธีภาวะน่าจะเป็นสูงสุด, ส่วนเบี่ยงเบนมาตรฐาน, ช่วงความเชื่อมั่น, ความน่าจะเป็นคุ้มรวม

กิตติกรรมประกาศ

โครงงานเรื่อง "การศึกษาประสิทธิภาพของตัวประมาณค่าส่วนเบี่ยงเบนมาตรฐานในการแจกแจงปรกติ" สามารถประสบความสำเร็จลุล่วงด้วยดี เนื่องจากได้รับความอนุเคราะห์ ความกรุณาและการสนับสนุนจาก ผศ.ดร.ภัทรวรรณ แสงนวกิจ อาจารย์ที่ปรึกษาโครงงาน โดยให้คำปรึกษา คำแนะนำ ข้อคิดเห็นและตรวจสอบความถูกต้องจนปรับปรุงแก้ไขข้อบกพร่องต่าง ๆ ทั้งให้ความเมตตาและเสียสละเวลาแก่คณะผู้วิจัยมาโดยตลอด

ขอขอบพระคุณ ผศ.ดร.มณฑิรา ดวงสาพล และ ผศ.เบญจมาศ ตุลยนิติกุล คณะกรรมการสอบและอาจารย์ประจำสาขาคณิตศาสตร์และสถิติทุกท่าน สำหรับคำแนะนำในการปรับปรุงแก้ไขข้อบกพร่องต่าง ๆ รวมถึงเจ้าหน้าที่สาขาคณิตศาสตร์และสถิติที่เอื้อเฟื้อสถานที่ในการทำงานของคณะผู้วิจัย

ขอขอบพระคุณคณะผู้จัดทำที่ให้ความช่วยเหลือ ตลอดจนคำแนะนำที่เป็นประโยชน์ในการจัดทำโครงงานครั้งนี้ให้ออกมาได้อย่างดี

คณะผู้วิจัยหวังว่าโครงงานฉบับนี้คงมีประโยชน์เป็นอย่างมากสำหรับผู้ที่สนใจในเรื่องการศึกษาประสิทธิภาพของตัวประมาณค่าส่วนเบี่ยงเบนมาตรฐานในการแจกแจงปรกติ จึงขอมอบส่วนดีทั้งหมดนี้ให้คณะอาจารย์ที่ประสงค์สิทธิประสาทวิชา หากมีข้อผิดพลาดประการใดคณะผู้วิจัยต้องขออภัยและน้อมรับไว้ ณ ที่นี้ด้วย

นายภูมิพัฒน์ มีเจริญวรานนท์
นางสาวกุลรัศมิ์ ทรัพย์อุดมมาก
นาวสาวธมลวรรณ กิจวรวุฒิ

สารบัญ

สารบัญรูปภาพ ตารางและกราฟ

รูปภาพ **หน้า**

ตาราง **หน้า**

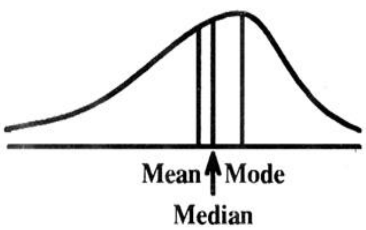
กราฟ **หน้า**

บทที่ 1

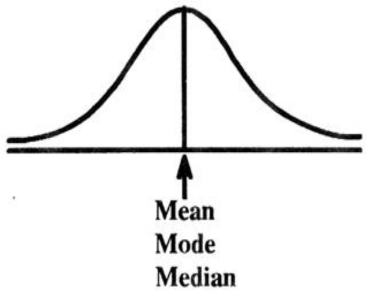
บทนำ

1.1 ที่มาและความสำคัญของปัญหา

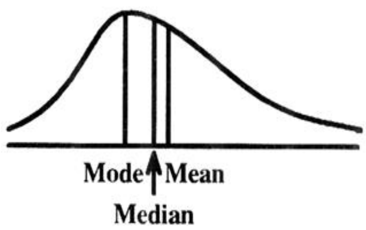
การประมาณค่าพารามิเตอร์เป็นเครื่องมือทางคณิตศาสตร์และสถิติที่สำคัญที่จะช่วยให้สามารถวิเคราะห์ข้อมูลและวางแผนการทำงานในหลากหลายด้าน เช่น งานทางด้านเศรษฐศาสตร์ การบริหารการตลาด การบริหารประเทศ และการทดลองวิทยาศาสตร์ แต่ปัญหาส่วนใหญ่ที่พบมักจะเป็นการนำวิธีการประมาณค่าพารามิเตอร์มาใช้ไม่สอดคล้องกับลักษณะของข้อมูลที่มีอยู่ ทั้งนี้เนื่องมาจากข้อมูลมีลักษณะหรือมีการแจกแจงได้หลายรูปแบบ เช่น ข้อมูลที่มีลักษณะสมมาตร และข้อมูลที่มีลักษณะไม่สมมาตร (เบ้ซ้ายหรือเบ้ขวา) ถ้าหากนักวิเคราะห์ข้อมูลเลือกใช้วิธีการทางสถิติที่ไม่เหมาะสมอาจส่งผลทำให้เกิดความคลาดเคลื่อนในการอนุมานทางสถิติได้



รูปที่ 1.1 การแจกแจงเบ้ซ้าย



รูปที่ 1.2 การแจกแจงปรกติ



รูปที่ 1.3 การแจกแจงเบ้ขวา

สำหรับข้อมูลตามธรรมชาติที่มักจะพบบ่อยครั้ง คือ การแจกแจงปรกติ (Normal distribution) ตัวแปรสุ่มที่เกี่ยวข้องกับการแจกแจงนี้จะเป็นตัวแปรที่มีลักษณะแบบต่อเนื่อง โดยจุดกึ่งกลาง คือ จุดที่มีความถี่สูงสุด ซึ่งจะแสดงค่าเฉลี่ยเลขคณิต ค่ามัธยฐาน และค่าฐานนิยม ค่าดังกล่าวจะมีค่าเท่ากัน อีกทั้ง ณ จุดกึ่งกลางจะมีแกนสมมาตร คือ เส้นตรงที่ลากผ่านจุดโด่งสุดของเส้นโค้งนั้นตั้งฉากกับแกนนอน และแกนสมมาตรจะแบ่งพื้นที่ใต้เส้นโค้งปรกติออกเป็น 2 ส่วนเท่ากัน ถ้าหากนำพื้นที่ดังกล่าวมาพิจารณาผลรวมของพื้นที่ใต้เส้นโค้งปรกติจะมีค่าเท่ากับ 1 เสมอ นอกจากนี้เส้นโค้งจะเข้าใกล้แกนนอน แต่จะไม่ตัดแกนนอนดังแสดงในรูปที่ 1.2

ถ้าตัวแปรสุ่ม X มีการแจกแจงปรกติด้วยค่าเฉลี่ยเท่ากับ μ ความแปรปรวนเท่ากับ σ^2 หรือใช้สัญลักษณ์แทนด้วย $X \sim N\left(\mu,\sigma^2\right)$ แล้ว X จะมีฟังก์ชันความหนาแน่นความน่าจะเป็น (Probability density function) คือ

$$f\left(x\middle|\mu,\sigma^2\right)=\frac{1}{\sqrt{2\pi\sigma^2}}e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}\ ;\ -\infty<x<\infty,\ -\infty<\mu<\infty,\ 0<\sigma^2<\infty$$

โดยที่ μ คือ พารามิเตอร์แทนค่าเฉลี่ย และ σ^2 คือ พารามิเตอร์แทนความแปรปรวน เมื่อ $\pi=3.141$ และ $e=2.718$

การแจกแจงปรกติจะมีพารามิเตอร์สองตัว คือ ค่าเฉลี่ยประชากร (Population mean) แทนด้วย μ และความแปรปรวนประชากร (Population variance) แทนด้วย σ^2 ดังนั้นหากหารากที่สองของความแปรปรวนประชากรแล้วจะเรียกว่าค่าดังกล่าวว่า ส่วนเบี่ยงเบนมาตรฐาน (Population standard deviation) แทนด้วย σ ในการอนุมานทางสถิติที่อิงพารามิเตอร์ (Parametric method) ตัวสถิติที่เราสร้างขึ้นมักจะอาศัยตัวแปรสุ่มที่มีพื้นมาจากการแจกแจงปรกติ เช่น การประมาณค่าแบบช่วงของค่าเฉลี่ยประชากรจะอาศัยปริมาณหมุน (Pivotal quantity) ที่มาจากการแจกแจงปรกติมาตรฐาน และการสร้างตัวสถิติทดสอบอัตราส่วนภาวะน่าจะเป็นจะอาศัยทฤษฎีขีดจำกัดกลาง (Central limit theorem) และการแจกแจงปรกติมาตรฐาน เนื่องจากในการแจกแจงปรกติมีพารามิเตอร์ที่ไม่ทราบค่า ได้แก่ μ และ σ^2 ตามที่ได้กล่าวไป ดังนั้นจึงต้องมีการสุ่มตัวอย่างเพื่อนำค่าสถิติที่ได้มาประมาณค่าพารามิเตอร์ สำหรับการประมาณค่าเฉลี่ยนั้นอาจอาศัยค่าเฉลี่ยตัวอย่าง มัธยฐาน หรือฐานนิยม ซึ่งเรียกว่า วิธีการวัดแนวโน้มเข้าสู่ส่วนกลาง ส่วนการประมาณค่าความแปรปรวนประชากร โดยทั่วไปมักใช้ความแปรปรวนตัวอย่างซึ่งหาได้จาก $S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$ และมีหน่วยในการวัด คือ หน่วยของข้อมูลยกกำลังสอง ตัวสถิตินี้เป็นตัวประมาณไม่เอนเอียงของ σ^2 อย่างไรก็ตามในการรายงานผลเกี่ยวกับการกระจายของข้อมูลโดยใช้ค่าความแปรปรวนที่มีหน่วยยกกำลังสองอาจทำให้เกิดความเข้าใจผิดไปได้ เนื่องจากเป็นตัวเลขเชิงปริมาณที่จะทำให้เกิดความรู้สึกว่ามีค่ามาก ดังนั้นจึงถอดรากที่สองจากความแปรปรวนตัวอย่างเพื่อให้ได้ค่าส่วนเบี่ยงเบนมาตรฐาน หรือ $S=\sqrt{S^2}$ ซึ่งตัวประมาณนี้จะมีหน่วยของข้อมูลเป็นไปตามตัวแปรที่ศึกษาและเข้าใจง่ายขึ้น ในกรณีนี้ตัวสถิติ S จึงถือว่าเป็นตัวแทนของพารามิเตอร์ σ

จากการประมาณค่าส่วนเบี่ยงเบนมาตรฐานข้างต้น ผู้วิจัยเล็งเห็นว่าการประมาณค่า σ^2 ด้วย S^2 เพราะ S^2 เป็นตัวประมาณที่มีสมบัติที่ดีสำหรับ σ^2 นั่นคือเป็นตัวประมาณไม่เอนเอียง แต่อย่างไรก็ตาม ผู้วิจัยยังมีคำถามต่อไปว่าการประมาณค่า σ ด้วย S จะเหมาะสมหรือไม่ และตัวประมาณ S จะยังคงมีสมบัติที่ดีทางสถิติเช่นเดียวกับ S^2 กล่าวคือเป็นตัวประมาณไม่เอนเอียงของ σ หรือไม่ ซึ่งหาก S ไม่ได้มีสมบัติที่ดีทางสถิติในเรื่องความไม่เอนเอียงอาจทำให้การประมาณค่าส่วนเบี่ยงเบนมาตรฐานประชากรเบี่ยงเบนไปจากค่าจริง และมีความคลาดเคลื่อนสูงในสถานการณ์ใดสถานการณ์หนึ่งได้ ดังนั้นในงานวิจัยฉบับนี้จึงต้องการที่จะศึกษาวิธีการประมาณค่าส่วนเบี่ยงเบนมาตรฐานในการแจกแจงปรกติ เพื่อให้ได้ตัวประมาณที่มีสมบัติไม่เอนเอียงในการประมาณค่าต่อไป

1.2 วัตถุประสงค์

1. เพื่อศึกษาวิธีการประมาณค่าส่วนเบี่ยงเบนมาตรฐานที่มีสมบัติเป็นตัวประมาณไม่เอนเอียงสำหรับส่วนเบี่ยงเบนมาตรฐานประชากรในการแจกแจงปรกติ

2. เพื่อศึกษาประสิทธิภาพของตัวประมาณที่นำเสนอในข้อที่ 1. เทียบตัวประมาณค่าส่วนเบี่ยงเบนมาตรฐานที่ใช้วิธีดั้งเดิมและวิธีภาวะน่าจะเป็นสูงสุด

3. เพื่อเปรียบเทียบช่วงความเชื่อมั่นสำหรับส่วนเบี่ยงเบนมาตรฐานประชากรในการแจกแจงปรกติ

1.3 ขอบเขตการศึกษา

1. กำหนดให้ $X = \left(X_1, X_2, ...X_n\right)$ เป็นตัวอย่างสุ่มขนาด n จากประชากรที่มีการแจกแจงปรกติ ซึ่งมีค่าเฉลี่ยเท่ากับ μ และความแปรปรวนเท่ากับ σ^2 และฟังก์ชันความหนาแน่นความน่าจะเป็นของ X คือ

$$f\left(x \middle| \mu, \sigma^2\right) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} ;\ -\infty < x < \infty,\ -\infty < \mu < \infty,\ 0 < \sigma^2 < \infty$$

2. เปรียบเทียบประสิทธิภาพของตัวประมาณแบบจุดสำหรับส่วนเบี่ยงเบนมาตรฐานที่สนใจในการศึกษาครั้งนี้และเป็นตัวประมาณไม่เอนเอียง (แทนด้วย $\hat{S}$) กับตัวประมาณค่าส่วนเบี่ยงเบนมาตรฐานที่ใช้วิธีดั้งเดิม (แทนด้วย S) และวิธีภาวะน่าจะเป็นสูงสุด (แทนด้วย S_{MLE})

3. เปรียบเทียบประสิทธิภาพของตัวประมาณจากการจำลอง (Simulation) ภายใต้สถานการณ์ต่าง ๆ โดยใช้โปรแกรม R

4. กำหนดค่าพารามิเตอร์ในสถานการณ์จำลองเพื่อศึกษาประสิทธิภาพของตัวประมาณ คือ μ เท่ากับ -5 0 และ 5 และกำหนด σ เท่ากับ 0.5 1 2 5 และ 10 สำหรับการกำหนดค่า σ หลายระดับนั้นเพื่อแทนการกระจายข้อมูลจากน้อยไปหามาก

5. กำหนดขนาดตัวอย่างที่ใช้ในการจำลอง คือ n เท่ากับ 10 20 30 50 100 และ 300 เพื่อแทนขนาดตัวอย่างจากน้อยไปหามาก

6. กำหนดจำนวนการทำซ้ำ M เท่ากับ 10,000 รอบในแต่ละสถานการณ์จำลอง

7. ศึกษาประสิทธิภาพของตัวประมาณแบบจุดจากค่าความเอนเอียง (Bias) และค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Mean squared error)

8. ศึกษาประสิทธิภาพของช่วงความเชื่อมั่น $(1-\alpha)100\%$ สำหรับส่วนเบี่ยงเบนมาตรฐานประชากร โดยพิจารณาจากค่าความน่าจะเป็นคุ้มรวม (Coverage probability) และค่าเฉลี่ยความกว้างของช่วงความเชื่อมั่น (Expected length) ของช่วงความเชื่อมั่นที่อยู่ในรูปแบบดังนี้

$$\hat{\theta} \pm t_{\alpha/2}\sqrt{Var(\hat{\theta})}$$

เมื่อ $\hat{\theta}$ คือ ตัวประมาณแบบจุดใด ๆ ของส่วนเบี่ยงเบนมาตรฐานประชากร $Var(\hat{\theta})$ คือ ตัวประมาณค่าความแปรปรวนของ $\hat{\theta}$ และ $t_{\alpha/2}$ คือ ค่าเปอร์เซ็นไทล์ที่ $(\alpha/2)100\%$ ของการแจกแจงที (t-distribution) ช่วงความเชื่อมั่นที่แสดงไว้ข้างต้นจะถูกนำมาเปรียบเทียบประสิทธิภาพกับช่วงความเชื่อมั่นของส่วนเบี่ยงเบนมาตรฐานจากวิธีดั้งเดิม นั่นคือ

$$\frac{\sqrt{(n-1)}}{\sqrt{\chi^2_{\frac{\alpha}{2},n-1}}}S \le \sigma \le \frac{\sqrt{(n-1)}}{\sqrt{\chi^2_{1-\frac{\alpha}{2},n-1}}}S$$

เมื่อ $\chi^2_{\frac{\alpha}{2},n-1}$ คือ ค่าเปอร์เซ็นไทล์ที่ $(\alpha/2)100\%$ ของการแจกแจงไคกำลังสอง (Chi-square distribution) ที่มีองศาเสรีเป็น $n-1$

1.4 ประโยชน์ที่คาดว่าจะได้รับ

1. ทราบวิธีการสร้างตัวประมาณค่าพารามิเตอร์ที่มีสมบัติที่ดีในเรื่องความไม่เอนเอียง และเป็นแนวทางในการประมาณค่าพารามิเตอร์ในการแจกแจงอื่น ๆ ต่อไป

2. ทราบสถานการณ์ที่เหมาะสมในการเลือกใช้ตัวประมาณค่าส่วนเบี่ยงเบนมาตรฐานในการแจกแจงปรกติ

บทที่ 2
ทฤษฎีทางสถิติและงานวิจัยที่เกี่ยวข้อง

ในการศึกษาประสิทธิภาพของตัวประมาณค่าส่วนเบี่ยงเบนมาตรฐานในการแจกแจงปรกติมีทฤษฎีและงานวิจัยที่เกี่ยวข้องดังนี้

2.1 ส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation)

ส่วนเบี่ยงเบนมาตรฐาน คือ ระยะห่างเฉลี่ยของค่าสังเกตจากค่าเฉลี่ย ซึ่งเป็นค่าที่แสดงถึงการกระจายของข้อมูลว่าเบี่ยงเบนไปจากค่าเฉลี่ยมากน้อยเพียงใด ถ้าหากค่าส่วนเบี่ยงเบนมาตรฐานมีค่าน้อยจะบ่งชี้ว่าข้อมูลมีการกระจายใกล้กับค่าเฉลี่ย แต่หากค่าส่วนเบี่ยงเบนมาตรฐานมีขนาดใหญ่จะบ่งชี้ว่าข้อมูลมีกระจายอยู่ไกลจากค่าเฉลี่ยนั่นเอง สูตรในการคำนวณค่าส่วนเบี่ยงเบนมาตรฐานโดยทั่วไปมีดังนี้

1. ส่วนเบี่ยงเบนมาตรฐานของประชากร

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(X_i - \mu)^2}{N}}$$

2. ส่วนเบี่ยงเบนมาตรฐานตัวอย่าง

$$S = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}}$$

2.2 ค่าคาดหมาย (Expected Value)

<u>นิยาม 1</u> ให้ X เป็นตัวแปรสุ่มไม่ต่อเนื่องที่มีฟังก์ชันมวลความน่าจะเป็นคือ $f(x)$.
ค่าคาดหมายของ X คือ

$$E(X) = \sum_{all\ x} xf(x)$$

<u>นิยาม 2</u> ให้ X เป็นตัวแปรสุ่มต่อเนื่องที่มีฟังก์ชันความหนาแน่นความน่าจะเป็น $f(x)$.
ค่าคาดหมายของ X คือ

$$E(X) = \int_{-\infty}^{\infty} xf(x)\,dx$$

<u>ทฤษฎีบท 3</u> ให้ X และ Y เป็นตัวแปรสุ่มที่ค่าคาดหมายหาค่าได้ และ a, b เป็นค่าคงที่
คุณสมบัติของค่าคาดหมายมีดังนี้

1. $E(b) = b$
2. $E(aX + b) = aE(X) + b$
3. $E(aX \pm bY) = aE(X) + bE(Y)$
4. สำหรับฟังก์ชัน g และ h ใด ๆ จะได้ว่า $E[g(X) \pm h(Y)] = E[g(X)] \pm E[h(Y)]$
5. ถ้า X และ Y เป็นตัวแปรสุ่มที่เป็นอิสระต่อกัน จะได้ว่า $E(XY) = E(X)E(Y)$
6. ถ้า X และ Y เป็นตัวแปรสุ่มที่เป็นอิสระต่อกัน และ สำหรับฟังก์ชัน g และ h ใด ๆ จะได้ว่า $E[g(X)h(Y)] = E[g(X)]E[h(Y)]$

<u>ทฤษฎีบท 4</u> ความแปรปรวนของตัวแปรสุ่ม X คือ

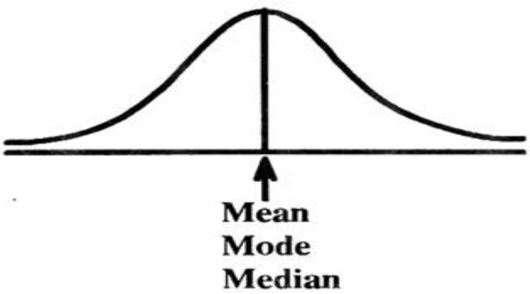
$$Var(X) = E(X^2) - E^2(X)$$

<u>ทฤษฎีบท 5</u> ให้ X และ Y เป็นตัวแปรสุ่มที่ค่าคาดหมายหาค่าได้ และ a, b เป็นค่าคงที่ จะได้ว่า

$$Var(aX + b) = a^2 Var(X)$$

2.3 การแจกแจงปรกติ (Normal Distribution)

การแจกแจงปรกติเป็นตัวแปรสุ่มชนิดต่อเนื่อง ทฤษฎีต่าง ๆ ในทางสถิติมักจะใช้พื้นฐานของการแจกแจงนี้ในการสร้างสูตรในการคำนวณ ลักษณะกราฟของข้อมูลที่มีการแจกแจงปรกติจะเป็นเส้นโค้งที่มีลักษณะเป็นรูประฆังที่สมมาตร และเมื่อแบ่งครึ่งโค้งการแจกแจงปรกติจะมีพื้นที่เท่ากันดังแสดงใน รูปที่ 2.1



รูปที่ 2.1 เส้นโค้งปรกติ

ถ้าตัวแปรสุ่ม X มีการแจกแจงปรกติด้วยค่าเฉลี่ยเท่ากับ μ ความแปรปรวนเท่ากับ σ^2 หรือใช้สัญลักษณ์แทนด้วย $X \sim N(\mu, \sigma^2)$ มีฟังก์ชันความหนาแน่นความน่าจะเป็น (Probability density function) คือ

$$f\left(x\middle|\mu,\sigma^2\right)=\frac{1}{\sqrt{2\pi\sigma^2}}e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad ; \; -\infty < x < \infty,\; -\infty < \mu < \infty,\; 0 < \sigma^2 < \infty$$

โดยที่ μ คือ พารามิเตอร์แทนค่าเฉลี่ย และ σ^2 คือ พารามิเตอร์แทนความแปรปรวน เมื่อ $\pi = 3.141$ และ $e = 2.718$ ฟังก์ชันก่อกำเนิดโมเมนต์ของการแจกแจงปรกติที่มีค่าเฉลี่ยเท่ากับ μ และความแปรปรวน σ^2 คือ $M_X\left(t\right)=e^{\mu t+\frac{1}{2}\sigma^2 t^2}$

2.4 ตัวประมาณที่ไม่เอนเอียง (Unbiasedness)

ความไม่เอนเอียง คือ สมบัติตัวประมาณที่มีค่าเฉลี่ยเท่ากันกับค่าพารามิเตอร์ของตัวประมาณนั้น ๆ เช่น สมมติให้ θ เป็นพารามิเตอร์ ดังนั้นตัวสถิติ $\hat{\theta}$ เป็นตัวประมาณที่ไม่เอนเอียงของ θ ก็ต่อเมื่อ $E\left(\hat{\theta}\right)=\theta$ ซึ่งค่าความเอนเอียงของ $\hat{\theta}$ จะหาได้จาก $Bias\left(\hat{\theta}\right)=E\left(\hat{\theta}\right)-\theta$

ในกรณีการประมาณค่าเฉลี่ย ถ้ากำหนดให้ $X_1, X_2, ..., X_n$ เป็นตัวอย่างสุ่มขนาด ที่มีการแจกแจงปรกติด้วยค่าเฉลี่ยเท่ากับ μ และความแปรปรวนเท่ากับ σ^2 จะได้ว่า $E\left(\bar{X}\right)=\mu$ ดังนั้น $\bar{X}$ เป็นตัวประมาณที่ไม่เอนเอียงของ μ

ในกรณีการประมาณค่าความแปรปรวน ถ้าให้ตัวประมาณของ σ^2 คือ $S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$ จะได้ว่า $E\left(S^2\right)=\sigma^2$ แต่หากพิจารณาตัวประมาณค่าความแปรปรวนด้วยวิธีภาวะน่าจะเป็นสูงสุด $S_{MLE}^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2$ แล้ว จะพบว่า $E\left(S_{MLE}^2\right)=\frac{(n-1)\sigma^2}{n}\neq\sigma^2$ ดังนั้นจึงสรุปได้ว่า S^2 เป็นตัวประมาณที่ไม่เอนเอียงของ σ^2 แต่ S_{MLE}^2 เป็นตัวประมาณที่เอนเอียงของ σ^2 (การพิสูจน์แสดงไว้ในบทที่ 3)

2.5 ตัวประมาณคงเส้นคงวา (Consistency)

ความคงเส้นคงวา คือ สมบัติของตัวประมาณแบบจุดที่จะต้องมีค่าประมาณเข้าใกล้ค่าจริงของประชากร เมื่อขนาดตัวอย่างเพิ่มขึ้น นิยามของความคงเส้นคงวามีดังนี้ ตัวสถิติ $\hat{\theta}$ จะเป็นตัวประมาณคงเส้นคงวาของพารามิเตอร์ θ ก็ต่อเมื่อ $P\left(\left|\hat{\theta}-\theta\right|<\varepsilon\right)\to 1$ เมื่อ $n\to\infty$ และ $\varepsilon>0$ ความคงเส้นคงวาของตัวประมาณพิจารณาได้จากทฤษฎีดังนี้ ตัวสถิติ $\hat{\theta}$ จะเป็นตัวประมาณคงเส้นคงวาของพารามิเตอร์ θ ก็ต่อเมื่อ

1. $\hat{\theta}$ เป็นตัวประมาณที่ไม่เอนเอียงของ θ เมื่อ $n\to\infty$ กล่าวคือ $\lim_{n\to\infty}E\left(\hat{\theta}\right)=\theta$ และ

2. ความแปรปรวนของ $\hat{\theta}$ มีค่าเข้าสู่ศูนย์ เมื่อ $n \to \infty$ กล่าวคือ $\lim_{n\to\infty} Var\left(\hat{\theta}\right) = 0$

2.6 ตัวประมาณภาวะน่าจะเป็นสูงสุด (Maximum Likelihood Estimator)

วิธีภาวะน่าจะเป็นสูงสุดเป็นวิธีการหนึ่งที่พบบ่อยและเป็นที่นิยมใช้ในทางสถิติ เนื่องจากวิธีการสร้างตัวประมาณค่าด้วยวิธีนี้จะได้ตัวประมาณค่าพารามิเตอร์ที่มีคุณสมบัติที่ดีหลายอย่าง

กำหนดให้ $X_1, X_2, ..., X_n$ เป็นตัวอย่างสุ่มขนาด n จากประชากรหนึ่งที่มีพารามิเตอร์แทนด้วย θ ซึ่งเป็นค่าคงที่ที่ไม่ทราบค่า และให้ $\hat{\theta} = \hat{\theta}\left(X_1, X_2, ..., X_n\right)$ เป็นตัวประมาณของ θ แล้ว $\hat{\theta}\left(x_1, x_2, ..., x_n\right)$ จะเป็นค่าประมาณของ θ โดยที่ $X_i = x_i$ ทุกค่า $i = 1, 2, 3, ..., n$

กำหนดให้ $X_1, X_2, ..., X_n$ เป็นตัวอย่างสุ่มขนาด n จากประชากรหนึ่งที่มี pdf. หรือ pmf. เป็น $f\left(x\right)$ และมีพารามิเตอร์แทนด้วย θ ซึ่งเป็นค่าคงที่ที่ไม่ทราบค่า และกำหนดให้

$$L\left(\theta; X_1, X_2, ..., X_n\right) = L\left(\theta\right) = f\left(x_1, x_2, ..., x_n; \theta\right)$$

แทนฟังก์ชันภาวะน่าจะเป็นของ θ ส่วน $f\left(x_1, x_2, ..., x_n; \theta\right)$ คือฟังก์ชันความหนาแน่นความน่าจะเป็นร่วม (Joint pdf.) หรือ ฟังก์ชันมวลความน่าจะเป็นร่วม (Joint pmf.) ของ $X_1, X_2, ..., X_n$ ที่มีพารามิเตอร์ θ ในกรณีที่ $X_1, X_2, ..., X_n$ เป็นอิสระต่อกันและมีการแจกแจงเหมือนกัน (Independence and identically distributed)

การหาฟังก์ชันภาวะน่าจะเป็นของ θ จะได้จาก

$$L\left(\theta\right) = \prod_{i=1}^{n} f\left(x_i; \theta\right) = \prod_{i=1}^{n} f\left(x_i\right)$$

ตัวประมาณภาวะน่าจะเป็นสูงสุดคือ θ ที่ทำให้ $L\left(\theta\right)$ มีค่าสูงสุด หรือกล่าวได้ว่า ถ้าให้ $\hat{\theta}$ เป็นตัวประมาณภาวะน่าจะเป็นสูงสุดของ θ แล้ว $L\left(\hat{\theta}\right) = \max_{\theta} L\left(\theta\right)$ เมื่อ $\hat{\theta} = \hat{\theta}\left(X_1, X_2, ..., X_n\right)$

2.7 ความคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Squared Error)

ความคลาดเคลื่อนกำลังสองเฉลี่ย หมายถึง ค่าคาดหมายของกำลังสองของความแตกต่างระหว่างตัวประมาณ $\hat{\theta}$ และพารามิเตอร์ θ ที่กำลังจะประมาณค่าคือ $MSE\left(\hat{\theta}\right) = E\left[\left(\hat{\theta} - \theta\right)^2\right]$

โดยมีความสัมพันธ์คือ $E\left[\left(\hat{\theta}-\theta\right)^2\right]=Var\left(\hat{\theta}\right)+\left\{E\left[\hat{\theta}\right]-\theta\right\}^2=Var\left(\hat{\theta}\right)+Bias\left(\hat{\theta}\right)^2$ (ผลบวกของความแปรปรวนของตัวประมาณกับกำลังสองของความเอนเอียงของตัวประมาณ)

ถ้า $\hat{\theta}$ เป็นตัวประมาณที่ไม่เอนเอียงของพารามิเตอร์ $\hat{\theta}$

แล้ว $MSE\left(\hat{\theta}\right)=E\left[\left(\hat{\theta}-\theta\right)^2\right]=Var\left(\hat{\theta}\right)$

2.8 การแจกแจงค่าเฉลี่ยตัวอย่าง (Mean of Random Sample Distribution)

<u>ทฤษฎีบท</u> ให้ $X_1,\ldots,X_n$ เป็นตัวอย่างสุ่มขนาด n จากการแจกแจงปรกติที่มีค่าเฉลี่ย μ และความแปรปรวน σ^2 จะได้ว่า $\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$ มีการแจกแจงปรกติที่มีค่าเฉลี่ยคือ μ และความแปรปรวนเท่ากับ σ^2/n

พิสูจน์ เนื่องจาก $X_1,\ldots,X_n$ เป็นตัวอย่างสุ่มจากการแจกแจงปรกติที่มีค่าเฉลี่ยคือ μ และความแปรปรวนเท่ากับ σ^2 นั่นคือ $X_i, i=1,2,3,\ldots,n$ แต่ละตัวเป็นอิสระกันและมีการแจกแจงปรกติเหมือนกันที่มี $E\left(X_i\right)=\mu$ และ $V\left(X_i\right)=\sigma^2$ นอกจากนั้นยังจะเห็นได้ว่า

$$\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i=\frac{1}{n}X_1+\frac{1}{n}X_2+\ldots+\frac{1}{n}X_n \quad =a_1X_1+a_2X_2+\ldots+a_nX_n$$

โดยที่ $a_i=\frac{1}{n}, i=1,2,\ldots,n$ นั่นคือ $\bar{X}$ สามารถเขียนในรูปผลรวมเชิงเส้นของ $X_1,\ldots,X_n$ ดังนั้น ตามทฤษฎีบท สามารถสรุปได้ว่า $\bar{X}$ มีการแจกแจงปรกติที่มีค่าเฉลี่ย

$$E\left[\bar{X}\right]=E\left[\frac{1}{n}X_1+\frac{1}{n}X_2+\ldots+\frac{1}{n}X_n\right] \quad =\frac{1}{n}E\left(X_1\right)+\frac{1}{n}E\left(X_2\right)+\ldots+\frac{1}{n}E\left(X_n\right)$$

$$=\frac{1}{n}\mu+\frac{1}{n}\mu+\ldots+\frac{1}{n}\mu \quad =\mu$$

และมีความแปรปรวน คือ

$$V\left[\bar{X}\right]=V\left[\frac{1}{n}X_1+\frac{1}{n}X_2+\ldots+\frac{1}{n}X_n\right] \quad =\frac{1}{n^2}V\left(X_1\right)+\frac{1}{n^2}V\left(X_2\right)+\ldots+\frac{1}{n^2}V\left(X_n\right)$$

$$=\frac{1}{n^2}\sigma^2+\frac{1}{n^2}\sigma^2+\ldots+\frac{1}{n^2}\sigma^2 \quad =\frac{\sigma^2}{n}$$

นั่นหมายความว่า การแจกแจงค่าตัวอย่างของ $\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$ จากประชากรที่มีการแจกแจงปรกติจะมีการแจกแจงปรกติที่มีค่าเฉลี่ยคือ μ และความแปรปรวนเท่ากับ $\frac{\sigma^2}{n}$

2.9 หลักการจำลองข้อมูล (Simulation)

ในการวิจัยหรือการศึกษาบางเรื่องนั้นผู้วิจัยไม่สามารถหาข้อมูลที่มีลักษณะตรงตามที่ต้องการมาศึกษาได้โดยง่าย ดังนั้นจึงจำเป็นต้องอาศัยข้อมูลจากการจำลอง (Simulated data) มาช่วยให้งานวิจัยสามารถดำเนินต่อไปได้อย่างสมบูรณ์ โดยการจำลองข้อมูลนั้นจะใช้ในการสร้างตัวเลขสุ่มให้มีลักษณะเป็นไปตามที่ต้องการและสามารถใช้ในการตรวจสอบประสิทธิภาพของตัวประมาณทางสถิติได้อีกด้วย

2.9.1 เกณฑ์การวัดประสิทธิภาพของตัวประมาณค่าพารามิเตอร์แบบจุด

ตัวประมาณค่าแบบจุดของพารามิเตอร์ที่สนใจมีเกณฑ์ที่นิยมใช้ในการวัดประสิทธิภาพ 2 เกณฑ์ คือ ความเอนเอียงและความคลาดเคลื่อนกำลังสองเฉลี่ย หากตัวประมาณค่าแบบจุดใดมีค่าความเอนเอียงเท่ากับหรือเข้าใกล้ศูนย์ และมีค่าความคลาดเคลื่อนกำลังสองเฉลี่ยต่ำ จะถือว่าตัวประมาณแบบจุดนั้นมีประสิทธิภาพที่ดีกว่าตัวประมาณที่นำมาเปรียบเทียบ สำหรับวิธีการหาค่าวัดประสิทธิภาพในการจำลองมีดังนี้

1. ความเอนเอียง (Bias)

จากนิยาม $\hat{\theta}$ เป็นตัวประมาณที่ไม่เอนเอียง (Unbiased estimator) ของพารามิเตอร์ θ ก็ต่อเมื่อ $E\left(\hat{\theta}\right)=\theta$ และความเอนเอียงของตัวประมาณ $\hat{\theta}$ คือ $Bias\left(\hat{\theta}\right)=E\left(\hat{\theta}\right)-\theta$ เมื่อ $E\left(\hat{\theta}\right)$ คือ ค่าความหวังของตัวประมาณ θ ในการจำลองข้อมูล $E\left(\hat{\theta}\right)$ ประมาณได้จาก $\hat{E}\left(\hat{\theta}\right)=\frac{1}{M}\sum_{i=1}^{M}\hat{\theta}_i$ ดังนั้น $Bias\left(\hat{\theta}\right)=\frac{1}{M}\sum_{i=1}^{M}\hat{\theta}_i-\theta$ เมื่อ M คือ จำนวนรอบของการทำซ้ำในการจำลอง และ $\hat{\theta}_i$ คือ ค่าประมาณในรอบการทำซ้ำในรอบที่ i

2. ความคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Squared Error)

จากนิยาม ความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณ $\hat{\theta}$ คือ $MSE\left(\hat{\theta}\right)=E\left(\hat{\theta}-\theta\right)^2$ $=\mathrm{Var}\left(\hat{\theta}\right)+\left[Bias(\theta)\right]^2$ เมื่อ $Var\left(\hat{\theta}\right)$ คือ ค่าความแปรปรวนของตัวประมาณ θ ในการจำลองข้อมูล $Var\left(\hat{\theta}\right)$ ประมาณได้จาก $V\hat{a}r\left(\hat{\theta}\right)=\frac{1}{M}\sum_{i=1}^{M}\left[\hat{\theta}_i-\hat{E}\left(\hat{\theta}\right)\right]^2$ ดังนั้น $MSE\left(\hat{\theta}\right)=Var\left(\hat{\theta}\right)+\left[Bias(\theta)\right]^2$ $=\frac{1}{M}\left[\hat{\theta}_i-\hat{E}\left(\hat{\theta}\right)\right]^2+\left(\frac{1}{M}\sum_{i=1}^{M}\hat{\theta}_i-\theta\right)^2$

2.9.2 เกณฑ์การวัดประสิทธิภาพของช่วงความเชื่อมั่นสำหรับพารามิเตอร์

ช่วงความเชื่อมั่นมีเกณฑ์ที่นิยมใช้ในการวัดประสิทธิภาพ 2 เกณฑ์ คือ ความน่าจะเป็นคุ้มรวมและความยาวช่วงเฉลี่ย หากช่วงความเชื่อมั่นใดมีค่าความน่าจะเป็นคุ้มรวมมากกว่าหรือเท่ากับระดับความเชื่อมั่นที่กำหนด $(1-\alpha)$ และมีค่าเฉลี่ยความกว้างของช่วงความเชื่อมั่นต่ำกว่า จะถือว่าช่วงความเชื่อมั่นมีประสิทธิภาพที่ดีกว่าช่วงความเชื่อมั่นที่นำมาเปรียบเทียบ สำหรับวิธีการหาค่าวัดประสิทธิภาพในการจำลองมีดังนี้

1. ความน่าจะเป็นคุ้มรวม (Coverage Probability)

การหาค่าความน่าจะเป็นคุ้มรวมของช่วงความเชื่อมั่น จะใช้ในการตรวจสอบว่าช่วงความเชื่อมั่นที่คำนวณได้นั้นครอบคลุมค่าพารามิเตอร์ที่สนใจด้วยความน่าจะเป็นมาน้อยเพียงใด ในการจำลองข้อมูลค่าความน่าจะเป็นคุ้มรวมหาได้จากวิธีการดังนี้ $CP = \dfrac{c\left(L \le \theta \le U\right)}{M}$ เมื่อ $c\left(L \le \theta \le U\right)$ คือ จำนวนครั้งที่ขอบเขตล่าง (Lower Limit: L) และขอบเขตบน (Upper Limit: U) ของช่วงความเชื่อมั่น $(1-\alpha)100\%$ ครอบคลุมพารามิเตอร์ที่สนใจ ทั้งนี้ความยาวของช่วงความเชื่อมั่นสามารถคำนวณได้จาก $Length = U - L$

2. ช่วงความเชื่อมั่นของค่าความน่าจะเป็นคุ้มรวม

ช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่าความน่าจะเป็นคุ้มรวม (CP) อยู่ในรูปแบบ

$$c_0 \pm z_{\frac{\alpha}{2}}\sqrt{\frac{c_0\left(1-c_0\right)}{M}}$$

เมื่อ c_0 คือค่าความน่าจะเป็นคุ้มรวม

z คือเปอร์เซ็นไทล์ที่ $(\alpha/2)100\%$ ของการแจกแจงปรกติมาตรฐาน

M คือจำนวนรอบของการทำซ้ำในการจำลองข้อมูล

ในงานวิจัยนี้พิจารณา $c_0 = 0.95$ $z_{0.025} = 1.96$ และ $M = 10,000$ คำนวณช่วงความเชื่อมั่น 95% ของค่าความน่าจะเป็นคุ้มรวม คือ $[0.9457, 0.9543]$ ดังนั้นผู้วิจัยจะตัดสินใจว่าตัวประมาณแบบช่วงจะมีประสิทธิภาพที่ดี เมื่อค่าความน่าจะเป็นคุ้มรวมที่ได้จากการจำลองข้อมูลอยู่ในช่วง $[0.9457, 0.9543]$

3. ค่าเฉลี่ยความกว้างของช่วงความเชื่อมั่นที่จะใช้ในการจำลองข้อมูล ซึ่งเป็นการหาค่าเฉลี่ยของความยาวของช่วงความเชื่อมั่นจากการทำซ้ำจำนวน M ครั้ง ซึ่งหาได้จากวิธีการดังนี้

$$EL = \frac{1}{M}\sum_{i=1}^{M}\left(U_i - L_i\right)$$

เมื่อ U_i และ L_i คือ ขอบเขตบนและขอบเขตล่างในรอบการทำซ้ำที่ i ตามลำดับ

2.10 งานวิจัยที่เกี่ยวข้อง

Mahamoudvand and Hassani (2009) นำเสนอวิธีการประมาณค่าแบบช่วงสำหรับสัมประสิทธิ์การแปรผันในการแจกแจงปรกติ โดยอาศัยตัวประมาณแบบจุดที่มีสมบัติไม่เอนเอียง เมื่อตัวอย่างมีขนาดใหญ่สำหรับค่าสัมประสิทธิ์การแปรผันประชากร จากนั้นจึงหาค่าความแปรปรวนของตัวประมาณแบบจุดดังกล่าวเพื่อนำไปสร้างช่วงความเชื่อมั่นสำหรับสัมประสิทธิ์การแปรผัน จากการจำลองในสถานการณ์ที่ค่าสัมประสิทธิ์การแปรผันประชากรน้อยกว่า 0.5 พบว่าช่วงความเชื่อมั่นที่เขานำเสนอให้ค่าความน่าจะเป็นคุ้มรวมใกล้เคียงกับความน่าจะเป็นที่กำหนด แต่เมื่อสัมประสิทธิ์การแปรผันประชากรมีค่าตั้งแต่ 0.5 ขึ้นไปช่วงความเชื่อมั่นกลับให้ค่าความน่าจะเป็นคุ้มรวมต่ำกว่าความน่าจะเป็นที่กำหนดเป็นอย่างมากทั้งที่ตัวอย่างมีขนาดใหญ่

Donner ans Tou (2010) นำเสนอวิธีการหาช่วงความเชื่อมั่นสำหรับฟังก์ชันของพารามิเตอร์ในการแจกแจงใด ๆ โดยอาศัยวิธีการสร้างช่วงความเชื่อมั่นสำหรับตัวอย่างขนาดใหญ่ วิธีดังกล่าวเรียกว่า Method of Variance Estimates Recovery ในงานวิจัยของ Donner ans Tou (2010) ได้นำเสนอสูตรปิดที่ใช้ในการหาช่วงความเชื่อมั่นสำหรับฟังก์ชันของพารามิเตอร์ที่อยู่ในรูปผลบวก ผลต่าง และอัตราส่วนไว้ด้วย

วรพจน์ แซ่หลี (2554) ได้นำเสนอช่วงความเชื่อมั่นสำหรับสัมประสิทธิ์การแปรผันในการแจกแจงปรกติสำหรับตัวอย่างขนาดใหญ่ ในงานวิจัยนี้เขาได้ทำการพิสูจน์หาความแปรปรวนของ $\hat{\tau}$ โดยประมาณ เมื่อได้ปริมาณหมุน (Pivot) ที่มีการแจกแจงปรกติมาตรฐานแล้วเขาจึงสร้างช่วงความเชื่อมั่นโดยอาศัยวิธีของ Wald (Wilson, 1927) อย่างไรก็ตามในการศึกษาดังกล่าวไม่ได้มีการศึกษาประสิทธิภาพของช่วงความเชื่อมั่นแต่อย่างใด

บทที่ 3
วิธีดำเนินการศึกษา

ในการศึกษาประสิทธิภาพของตัวประมาณค่าส่วนเบี่ยงเบนมาตรฐานในการแจกแจงปรกติมีวิธีการดำเนินการดังต่อไปนี้

3.1 สมบัติของตัวประมาณค่าพารามิเตอร์ในการแจกแจงปรกติ

กำหนดให้ $X_1, X_2, \ldots, X_n$ เป็นตัวอย่างสุ่มจากประชากรที่มีการแจกแจงปรกติซึ่งมีค่าเฉลี่ย คือ μ และความแปรปรวน คือ σ^2 โดยที่ μ และ σ^2 ต่างเป็นค่าคงตัวที่ไม่ทราบค่า ในการประมาณค่าพารามิเตอร์ μ และ σ^2 โดยทั่วไปนั้นมักจะอาศัยวิธีภาวะน่าจะเป็นสูงสุดหรือการประมาณค่า σ^2 ด้วยตัวประมาณไม่เอนเอียง ซึ่งในหัวข้อต่อไปผู้วิจัยจะกล่าวถึงขั้นตอนการประมาณค่าพารามิเตอร์จากวิธีดังกล่าวและสมบัติบางประการของตัวประมาณ

3.1.1 ตัวประมาณค่าพารามิเตอร์ μ

ในการประมาณค่า μ ที่อาศัยวิธีภาวะน่าจะเป็นสูงสุดสามารถหาได้จากขั้นตอนดังต่อไปนี้

1. พิจารณาตัวประมาณภาวะน่าจะเป็นสูงสุดของพารามิเตอร์ μ

พิจารณาฟังก์ชันความหนาแน่นความน่าจะเป็นของตัวแปรสุ่ม X จาก

$$f(x\,;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}\left(x-\mu\right)^2\right)$$

จากนั้นหาฟังก์ชันภาวะน่าจะเป็นสูงสุด (Likelihood function) ของ μ และ σ^2 โดยพิจารณาจากฟังก์ชันความหนาแน่นความน่าจะเป็นร่วม (Joint probability density function) ของ $X_1, X_2, \ldots, X_n$ ซึ่งจะได้ว่า

$$\begin{aligned} L(\mu,\sigma^2) &= \prod_{i=1}^{n} f(x\,;\mu,\sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}\left(x_i-\mu\right)^2\right) \\ &= \left(2\pi\sigma^2\right)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(x_i-\mu\right)^2\right) \end{aligned}$$

ดังนั้นจะได้ ล็อก-ฟังก์ชันภาวะน่าจะเป็นสูงสุด (Log-likelihood function) คือ

$$\begin{aligned} \ln L(\mu,\sigma^2) &= -\frac{n}{2}\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(x_i-\mu\right)^2 \\ &= -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(x_i-\mu\right)^2 \end{aligned}$$

เมื่อหาอนุพันธ์ย่อยอันดับที่หนึ่งเทียบกับ μ แล้วจะได้

$$\frac{\partial}{\partial\mu}\ln L(\mu,\sigma^2) = -\frac{1}{2\sigma^2}\sum_{i=1}^{n}2\left(x_i-\mu\right)\left(-1\right) = \frac{1}{\sigma^2}\sum_{i=1}^{n}\left(x_i-\mu\right)$$

และหาค่าสุดขีดจากการนำค่าของอนุพันธ์ย่อยเทียบกับ μ มีค่าเป็น 0 เราจะพบว่า

$$\frac{\partial}{\partial\mu}\ln L(\mu,\sigma^2) = 0 \quad \Leftrightarrow \quad \frac{1}{\sigma^2}\sum_{i=1}^{n}\left(x_i-\hat{\mu}_{MLE}\right) = 0$$

$$\sum_{i=1}^{n}\left(x_i-\hat{\mu}_{MLE}\right) = 0 \quad \Leftrightarrow \quad \sum_{i=1}^{n}x_i - n\hat{\mu}_{MLE} = 0$$

ดังนั้น ตัวประมาณค่าของ μ คือ $\hat{\mu}_{MLE} = \dfrac{1}{n}\sum_{i=1}^{n}X_i = \bar{X}$

และพบว่า $\dfrac{\partial^2}{\partial\mu^2}\ln L(\mu,\sigma^2) = \dfrac{1}{\sigma^2}\sum_{i=1}^{n}(-1) = -\dfrac{n}{\sigma^2} < 0$ เมื่อ $-\infty < \mu < \infty$, $\sigma^2 > 0$, $n > 1$ ซึ่งทำให้ค่าของ $\ln L(\mu,\sigma^2)$ มีค่าสูงสุด จากการพิสูจน์ข้างต้นจึงสรุปได้ว่า $\bar{X}$ เป็นตัวประมาณภาวะน่าจะเป็นสูงสุดของ μ

2. พิสูจน์สมบัติความไม่เอนเอียงของตัวประมาณภาวะน่าจะเป็นสูงสุดสำหรับ μ

จากข้อ 1. เราทราบว่าตัวประมาณ $\hat{\mu}_{MLE} = \dfrac{1}{n}\sum_{i=1}^{n}X_i = \bar{X}$ และเนื่องจาก $X_i \sim N\left(\mu,\sigma^2\right)$ จะได้ว่า $\bar{X} \sim N\left(\mu,\dfrac{\sigma^2}{n}\right)$ เมื่อพิจารณาค่าคาดหมายของ $\bar{X}$ จะได้ว่า $E\left(\bar{X}\right) = \mu$ ดังนั้นจึงสามารถสรุปได้ว่า $\hat{\mu}_{MLE} = \bar{X}$ เป็นตัวประมาณที่ไม่เอนเอียงของ μ

3. พิสูจน์สมบัติความคงเส้นคงวาของตัวประมาณภาวะน่าจะเป็นสูงสุดสำหรับ μ

ในทำนองเดียวกัน จากข้อ 1. เราทราบว่า $\hat{\mu}_{MLE} = \bar{X}$ และ $\bar{X} \sim N\left(\mu,\dfrac{\sigma^2}{n}\right)$ เมื่อพิจารณาค่าความแปรปรวนของ $\bar{X}$ ได้ว่า $Var(\bar{X}) = \dfrac{\sigma^2}{n}$ และ $E\left(\bar{X}\right) = \mu$ นอกจากนั้นเมื่อทำการพิจารณาค่าดังกล่าวเมื่อตัวอย่างมีขนาดใหญ่ ($n \to \infty$) พบว่า

$$\lim_{n\to\infty}E(\hat{\mu}_{MLE}) = \lim_{n\to\infty}E(\bar{X}) = \lim_{n\to\infty}\mu = \mu \text{ และ}$$

$$\lim_{n\to\infty}Var(\hat{\mu}_{MLE}) = \lim_{n\to\infty}Var(\bar{X}) = \lim_{n\to\infty}\frac{\sigma^2}{n} = 0 \;;\; 0 < \sigma^2 < \infty$$

ดังนั้นจึงสรุปได้ว่า $\hat{\mu}_{MLE} = \bar{X}$ เป็นตัวประมาณคงเส้นคงวาของ μ

3.1.2 สมบัติของตัวประมาณพารามิเตอร์ σ^2 ในการแจกแจงปรกติ

การประมาณค่า σ^2 โดยทั่วไปนั้น เรามักพบว่านักวิจัยใช้ตัวประมาณ $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$ ซึ่งตัวประมาณดังกล่าวมีสมบัติดังนี้

1. พิสูจน์สมบัติความไม่เอนเอียงของตัวประมาณ S^2

กำหนดให้ $X_i \sim N\left(\mu,\sigma^2\right)$ และ $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$ จากความสัมพันธ์ระหว่างการแจกแจงปรกติและการแจกแจงไคกำลังสองจะได้ว่า

$$\frac{1}{\sigma^2}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2 = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)} \tag{1.1}$$

นั่นคือตัวแปรดังกล่าวจะมีค่าคาดหมายเท่ากับ

$$E\left[\frac{(n-1)S^2}{\sigma^2}\right] = n-1 \quad \Leftrightarrow \quad \frac{(n-1)}{\sigma^2}E(S^2) = n-1$$

และ $$E(S^2) = \sigma^2$$

ดังนั้นจึงกล่าวได้ว่า $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$ เป็นตัวประมาณที่ไม่เอนเอียงของ σ^2

2. พิสูจน์สมบัติความคงเส้นคงวาของ S^2

จากสมการ (1.1) เราพบว่า $Var\left[\frac{(n-1)S^2}{\sigma^2}\right] = 2(n-1)$ อาศัยทฤษฎีที่เกี่ยวข้องกับการพิสูจน์ความคงเส้นคงวาของตัวประมาณจะพบว่า

$$\lim_{n\to\infty} E(S^2) = \lim_{n\to\infty}\sigma^2 = \sigma^2 \text{ และ}$$

$$\lim_{n\to\infty} Var(S^2) = \lim_{n\to\infty}\frac{2\sigma^4}{(n-1)} = 0 \; ; \; 0 < \sigma^2 < \infty$$

ดังนั้นจึงสรุปได้ว่า $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$ เป็นตัวประมาณคงเส้นคงวาของ σ^2

3.1.3 สมบัติของตัวประมาณภาวะน่าจะเป็นสูงสุดของพารามิเตอร์ σ^2 ในการแจกแจงปรกติ

1. พิจารณาตัวประมาณภาวะน่าจะเป็นสูงสุดของพารามิเตอร์ σ^2

จากฟังก์ชันความหนาแน่นความน่าจะเป็นของตัวแปรสุ่ม X หรือ $f(x\,;\mu,\sigma^2)$ และ ล็อกฟังก์ชันภาวะน่าจะเป็นสูงสุด $\ln L(\mu,\sigma^2)$ เมื่อหาอนุพันธ์ย่อยอันดับที่หนึ่งเทียบกับ σ^2 จะได้ว่า

$$\frac{\partial}{\partial\sigma^2}\ln L(\mu,\sigma^2) = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\mu)^2$$

หาค่าสุดขีดจากการนำค่าของอนุพันธ์ย่อยข้างต้นเทียบกับ σ^2 มีค่าเป็น 0 จะพบว่า

$$\frac{\partial}{\partial\sigma^2}\ln L(\mu,\sigma^2) = 0$$

$$-\frac{n}{2\hat{\sigma}^2_{MLE}} + \frac{1}{2\hat{\sigma}^4_{MLE}}\sum_{i=1}^{n}(x_i-\hat{\mu}_{MLE})^2 = 0$$

$$\frac{1}{2\hat{\sigma}^4_{MLE}}\sum_{i=1}^{n}(x_i-\hat{\mu}_{MLE})^2 = \frac{n}{2\hat{\sigma}^2_{MLE}}$$

$$\hat{\sigma}^2_{MLE} = \frac{1}{n}\sum_{i=1}^{n}(x_i-\hat{\mu}_{MLE})^2$$

และพบว่า $$\frac{\partial^2}{\partial\sigma^4}\ln L(\mu,\sigma^2) = \frac{n}{2\left[\frac{1}{n}\sum_{i=1}^{n}(x_i-\hat{\mu}_{MLE})^2\right]^2} - \frac{\left[\sum_{i=1}^{n}(x_i-\hat{\mu}_{MLE})^2\right]}{\left[\frac{1}{n}\sum_{i=1}^{n}(x_i-\hat{\mu}_{MLE})^2\right]^3}$$

$$= \frac{1}{\left[\sum_{i=1}^{n}(x_i-\hat{\mu}_{MLE})^2\right]^2}\left(\frac{n^3}{2}-n^3\right)$$

$$= \frac{1}{\left[\sum_{i=1}^{n}(x_i-\hat{\mu}_{MLE})^2\right]^2}\left(-\frac{1}{2}n^3\right) < 0 \ ; \ \sigma^2 > 0 \ , n > 1$$

ทำให้ค่าของ $\ln L(\mu,\sigma^2)$ มีค่าสูงสุด และถ้ากำหนดให้ $S^2_{MLE} = \hat{\sigma}^2_{MLE}$ แล้ว เราจะสามารถสรุปได้ว่า $S^2_{MLE} = \frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2$ เป็นตัวประมาณภาวะน่าจะเป็นสูงสุดของ σ^2

2. พิสูจน์สมบัติความไม่เอนเอียงของตัวประมาณภาวะน่าจะเป็นสูงสุดสำหรับพารามิเตอร์ σ^2

เนื่องจาก $X_i \sim N\left(\mu,\sigma^2\right)$ และตัวประมาณภาวะน่าจะเป็นสูงสุดสำหรับพารามิเตอร์ σ^2 คือ $S^2_{MLE} = \frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2$

อาศัยทฤษฎี $$\frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\bar{X})^2 = \frac{nS^2_{MLE}}{\sigma^2} \sim \chi^2_{(n-1)} \qquad (1.2)$$

ในการหาค่าคาดหมายของตัวแปรสุ่มจะได้ดังนี้

$$E\left[\frac{nS_{MLE}^2}{\sigma^2}\right] = n-1 \quad \Leftrightarrow \quad \frac{n}{\sigma^2}E(S_{MLE}^2) = n-1$$

$$E\left(S_{MLE}^2\right) = \left(\frac{n-1}{n}\right)\sigma^2$$

ดังนั้น $S_{MLE}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$ เป็นตัวประมาณเอนเอียงสำหรับ σ^2 โดยมีค่าความเอนเอียงของตัวประมาณเท่ากับ $Bias(S_{MLE}^2) = \left(\frac{n-1}{n}\right)\sigma^2 - \sigma^2 = \frac{-\sigma^2}{n}$

3. พิสูจน์สมบัติความคงเส้นคงวาของตัวประมาณภาวะน่าจะเป็นสูงสุดสำหรับพารามิเตอร์ σ^2

จากสมการ (1.2) เราทราบแล้วว่า $Var\left[\frac{nS_{MLE}^2}{\sigma^2}\right] = 2\left(n-1\right)$ ดังนั้น

$$\frac{n^2}{\sigma^4}Var(S_{MLE}^2) = 2\left(n-1\right)$$

$$Var(S_{MLE}^2) = \left(\frac{2\left(n-1\right)}{n^2}\right)\sigma^4$$

เมื่อพิจารณาค่าคาดหมายและค่าความแปรปรวนของ S_{MLE}^2 เมื่อตัวอย่างมีขนาดใหญ่ จะพบว่า

$$\lim_{n\to\infty} E(S_{MLE}^2) = \lim_{n\to\infty}\left(\frac{n-1}{n}\right)\sigma^2 = \lim_{n\to\infty}\left(1-\frac{1}{n}\right)\sigma^2 = \sigma^2$$

$$\lim_{n\to\infty} Var(S_{MLE}^2) = \lim_{n\to\infty}\left(\frac{2\left(n-1\right)}{n^2}\right)\sigma^4 = \lim_{n\to\infty}\left(\frac{2}{n}-\frac{2}{n^2}\right)\sigma^4 = 0 \; ; \; 0<\sigma^2<\infty$$

ดังนั้น $S_{MLE}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$ จึงเป็นตัวประมาณที่คงเส้นคงวาของ σ^2

3.2 สมบัติของตัวประมาณค่าส่วนเบี่ยงเบนมาตรฐานในการแจกแจงปรกติ

จากหัวข้อ 3.1 ผู้วิจัยทราบแล้วว่า S^2 และ S_{MLE}^2 มีสมบัติบางประการในเรื่องความไม่เอนเอียงและความคงเส้นคงวาอย่างไร โดยเฉพาะอย่างยิ่ง S^2 มีสมบัติที่ดีถึงสองประการ คือ เป็นตัวประมาณไม่เอนเอียงและคงเส้นคงวาสำหรับ μ และ σ^2 ตามลำดับ ส่วน S_{MLE}^2 เป็นตัวประมาณคงเส้นคงวาของ σ^2 อย่างไรก็ตามผู้วิจัยยังคงต้องการทราบว่าตัวประมาณ S และ S_{MLE} ซึ่งได้จากวิธีการที่ง่าย คือ ถอดรากที่สองของตัวประมาณ S^2 และ S_{MLE}^2 ตามลำดับเท่านั้น จะยังคงเป็นตัวประมาณแบบจุดที่มีสมบัติที่ดีหรือไม่เพื่อ

ใช้ในการประมาณค่าส่วนเบี่ยงเบนมาตรฐานของประชากร σ การพิสูจน์สมบัติที่กล่าวไปของ S และ S_{MLE} ตามทฤษฎีมีดังต่อไปนี้

3.2.1 พิสูจน์ค่าคาดหมายที่มีความสัมพันธ์กับตัวประมาณ S

การหาค่าคาดหมายของตัวประมาณ S จะเริ่มต้นการแจกแจงของ $\frac{1}{\sigma^2}\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2=\frac{(n-1)S^2}{\sigma^2}$ ซึ่งมีการแจกแจงไคกำลังสองด้วยค่าองศาเสรีเท่ากับ $n-1$

ถ้ากำหนดให้ $Y=\frac{1}{\sigma^2}\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2=\frac{(n-1)S^2}{\sigma^2}\sim\chi^2_{(n-1)}$ และพิจารณาการแจกแจงของ $W=\sqrt{Y}$ จะได้ฟังก์ชันความหนาแน่นความน่าจะเป็นของ Y คือ

$$f_Y(y)=\frac{y^{\frac{(n-1)}{2}-1}e^{\frac{-y}{2}}}{\Gamma\left(\frac{n-1}{2}\right)2^{\frac{n-1}{2}}},\ y>0,\ n-1>0$$

เนื่องจาก $w\ =\ \sqrt{y}, y>0, n-1>0$

จะได้ $y\ =\ w^2$ และ $J\ =\ \frac{dy}{dw}\ =\ 2w$

ดังนั้น $f_W(w)=f_Y(w^2)\cdot|2w|$

$$=\frac{(w^2)^{\frac{(n-1)}{2}-1}e^{\frac{-w^2}{2}}}{\Gamma\left(\frac{n-1}{2}\right)2^{\frac{(n-1)}{2}}}(2w)=\frac{w^{(n-1)-1}e^{\frac{-w^2}{2}}}{\Gamma\left(\frac{n-1}{2}\right)2^{\frac{(n-1)}{2}-1}}\ ,w>0$$

$$E(W)=\int_0^\infty w\frac{w^{(n-1)-1}e^{\frac{-w^2}{2}}}{\Gamma\left(\frac{n-1}{2}\right)2^{\frac{(n-1)}{2}-1}}dw=\int_0^\infty\frac{w^{(n-1)}e^{\frac{-w^2}{2}}}{\Gamma\left(\frac{n-1}{2}\right)2^{\frac{(n-1)}{2}-1}}dw$$

เมื่อพิจารณา $\int_0^\infty w^{(n-1)}e^{\frac{-w^2}{2}}dw$ จากการอินทิเกรต โดยการแทนค่า นั่นคือ กำหนดให้ $u=\frac{w^2}{2}$ และ $du\ =\ wdw$ ได้ว่า $w=\sqrt{2u}$ และ $dw=\frac{du}{\sqrt{2u}}$ นอกจากนั้นจะได้ว่า

$$\int_0^\infty w^{(n-1)}e^{\frac{-w^2}{2}}dw=\int_0^\infty\sqrt{2u}^{(n-1)}e^{-u}\frac{du}{\sqrt{2u}}=\int_0^\infty\sqrt{2u}^{n-2}e^{-u}du$$

$$=\sqrt{2}^{n-2}\int_0^\infty u^{\frac{n}{2}-1}e^{-u}du=2^{\frac{n}{2}-1}\Gamma\left(\frac{n}{2}\right)$$

ดังนั้น $E(W)=\dfrac{2^{\frac{n}{2}-1}\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)2^{\frac{(n-1)}{2}-1}}$ และ $E(\sqrt{Y})=\sqrt{2}\dfrac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}$

เราจึงสรุปได้ว่า $E\left(\sqrt{Y}\right)=E\left[\sqrt{\frac{1}{\sigma^2}\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2}\right]=E\left(\sqrt{\frac{(n-1)S^2}{\sigma^2}}\right)=\sqrt{2}\dfrac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}$

3.2.2 ความเอนเอียงของตัวประมาณ S

ใช้ข้อมูลในหัวข้อที่ผ่านมาในการพิจารณาค่าคาดหมายของ S ซึ่งจะได้ว่า

$$E\left(\sqrt{\frac{(n-1)S^2}{\sigma^2}}\right)=\sqrt{2}\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}$$

$$\frac{\sqrt{n-1}}{\sigma}E(S)=\sqrt{2}\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}$$

และ $$E(S)=\sqrt{\frac{2}{n-1}}\cdot\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}\sigma\neq\sigma$$

ดังนั้นจึงสรุปได้ว่า $S=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2}$ เป็นตัวประมาณที่เอนเอียงของ σ โดยค่าความเอนเอียงของตัวประมาณจะหาได้จาก $Bias(S)=E(S)-\sigma=\left[\sqrt{\frac{2}{n-1}}\cdot\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}-1\right]\sigma$

3.2.3 ความแปรปรวนของตัวประมาณ S

เนื่องจาก $E\left(W\right)=E\left(\sqrt{Y}\right)=E\left[\sqrt{\frac{1}{\sigma^2}\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2}\right]=\sqrt{2}\,\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}$ และ

$$E\left(W^2\right)=E\left(Y\right)=E\left[\frac{1}{\sigma^2}\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2\right]=n-1$$

เมื่อพิจารณา $Var(W)=E(W^2)-\left[E(W)\right]^2=E(Y)-\left[E(W)\right]^2$

$$=(n-1)-\left[\sqrt{2}\,\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}\right]^2=(n-1)-2\left[\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}\right]^2$$

แล้ว พบว่า $Var\left[\sqrt{\frac{(n-1)S^2}{\sigma^2}}\right]=(n-1)-2\left[\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}\right]^2$

$$\frac{(n-1)}{\sigma^2}Var(S)=(n-1)-2\left[\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}\right]^2$$

ดังนั้น $Var(S)=\left(1-\frac{2}{(n-1)}\left[\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}\right]^2\right)\sigma^2$

3.2.4 ความเอนเอียงของตัวประมาณ S_{MLE}

การพิสูจน์จะทำในทำนองเดียวกันกับหัวข้อที่ผ่านมา นั่นคือ จากที่เราทราบแล้วว่า $\frac{nS_{MLE}^2}{\sigma^2}\sim\chi_{(n-1)}^2$

จึงทำให้ $E\left(\sqrt{\frac{nS_{MLE}^2}{\sigma^2}}\right)=\sqrt{2}\,\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}$ และ $E\left(S_{MLE}\right)=\sqrt{\frac{2}{n}}\,\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}\sigma$

ดังนั้น $S_{MLE}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2}$ จึงเป็นตัวประมาณเอนเอียงของ σ โดยมีค่าความเอนเอียงเท่ากับ

$$Bias(S_{MLE})=E(S_{MLE})-\sigma=\left[\sqrt{\frac{2}{n}}\cdot\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}-1\right]\sigma$$

3.2.5 ความแปรปรวนของตัวประมาณ S_{MLE}

จากทฤษฎี $\frac{1}{\sigma^2}\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2=\frac{nS_{MLE}^2}{\sigma^2}\sim\chi^2_{(n-1)}$ และข้อมูลในหัวข้อ 3.2.3 เราพิสูจน์ได้ว่า

$$E\left(W^2\right)=E\left(Y\right)=E\left[\frac{1}{\sigma^2}\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2\right]=n-1$$

และ

$$E\left(W\right)=E\left(\sqrt{Y}\right)=E\left[\sqrt{\frac{1}{\sigma^2}\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2}\right]=\sqrt{2}\,\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}$$

นอกจากนั้นเมื่อหาค่าความแปรปรวนของ W จะได้

$$Var\left(W\right)=E\left(W^2\right)-\left[E\left(W\right)\right]^2$$

$$Var\left(\sqrt{\frac{1}{\sigma^2}\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2}\right)=Var\left(\sqrt{\frac{nS_{MLE}^2}{\sigma^2}}\right)=\left(n-1\right)-\left[\sqrt{2}\,\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}\right]^2$$

$$\frac{n}{\sigma^2}Var\left(S_{MLE}\right)=\left(n-1\right)-\left[\sqrt{2}\,\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}\right]^2$$

ดังนั้น ความแปรปรวนของตัวประมาณ $S_{MLE}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2}$ คือ

$$Var\left(S_{MLE}\right)=\left(\left(1-\frac{1}{n}\right)-\frac{2}{n}\left[\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}\right]^2\right)\sigma^2$$

3.3 ตัวประมาณไม่เอนเอียงสำหรับส่วนเบี่ยงเบนมาตรฐาน

จากข้อมูลที่ได้จากการพิสูจน์เกี่ยวกับสมบัติของตัวประมาณในหัวข้อที่ผ่านมา ทำให้ผู้วิจัยสามารถพิจารณาตัวประมาณไม่เอนเอียงสำหรับส่วนเบี่ยงเบนมาตรฐานได้ดังนี้

3.3.1 ตัวประมาณไม่เอนเอียงสำหรับตัวประมาณส่วนเบี่ยงเบนมาตรฐาน

จากหัวข้อ 3.2.3 ได้ว่า $E(S)=\sqrt{\frac{2}{n-1}}\cdot\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}\sigma$ และ $E\left[\sqrt{\frac{n-1}{2}}\frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}S\right]=\sigma$

ดังนั้น $\hat{S}=\sqrt{\frac{n-1}{2}}\frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}S$ เป็นตัวประมาณไม่เอนเอียงของ σ

และพบว่า

$$\hat{S}=\sqrt{\frac{n-1}{2}}\frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}S=\sqrt{\frac{n-1}{2}}\frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2}=\frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}\sqrt{\frac{1}{2}\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2}$$

ต่อมาผู้วิจัยสามารถพิสูจน์ความแปรปรวนของ $\hat{S}$ ได้ดังต่อไปนี้

จากหัวข้อ 3.3.1 เราพบว่า $\hat{S}=\sqrt{\frac{n-1}{2}}\frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}S$ เมื่อ $S=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2}$ และ

$$Var(S)=\left(1-\frac{2}{(n-1)}\left[\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}\right]^2\right)\sigma^2$$

เมื่อพิจารณา $$Var\left(\hat{S}\right)=Var\left[\sqrt{\frac{n-1}{2}}\frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}S\right]$$

จะได้ว่าความแปรปรวนของ $\hat{S}$ คือ

$$Var\left(\hat{S}\right)=\left(\sqrt{\frac{n-1}{2}}\frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}\right)^{2}Var\left[S\right]$$

$$=\left(\sqrt{\frac{n-1}{2}}\frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}\right)^{2}\left(1-\frac{2}{\left(n-1\right)}\left[\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}\right]^{2}\right)\sigma^{2}$$

$$=\left(\frac{\left(n-1\right)}{2}\left(\frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}\right)^{2}-1\right)\sigma^{2}$$

3.3.2 ตัวประมาณไม่เอนเอียงสำหรับตัวประมาณส่วนเบี่ยงเบนมาตรฐานภาวะน่าจะเป็นสูงสุด

จากหัวข้อ 3.2.4 ได้พิสูจน์ว่า $E\left(S_{MLE}\right)=\sqrt{\frac{2}{n}}\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}\sigma$ และ $E\left[\sqrt{\frac{n}{2}}\frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}S_{MLE}\right]=\sigma$

ดังนั้น $\tilde{S}=\sqrt{\frac{n}{2}}\frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}S_{MLE}$ เป็นตัวประมาณไม่เอนเอียงของ σ

และพบว่า

$$\tilde{S}=\sqrt{\frac{n}{2}}\frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}S_{MLE}=\sqrt{\frac{n}{2}}\frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2}=\frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}\sqrt{\frac{1}{2}\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2}$$

ต่อมาผู้วิจัยสามารถพิสูจน์ความแปรปรวนของ $\tilde{S}$ ได้ดังต่อไปนี้

จากหัวข้อ 3.3.2 เราได้ว่า $\tilde{S}=\sqrt{\frac{n}{2}}\frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}S_{MLE}$ เมื่อ $S_{MLE}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2}$ และ

$$Var\left(S_{MLE}\right)=\left(\left(1-\frac{1}{n}\right)-\frac{2}{n}\left[\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}\right]^2\right)\sigma^2$$

เมื่อพิจารณา $Var\left(\tilde{S}\right)=Var\left[\sqrt{\frac{n}{2}}\frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}S_{MLE}\right]=\left(\sqrt{\frac{n}{2}}\frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}\right)^2 Var\left(S_{MLE}\right)$

จึงพบว่าความแปรปรวนของ $\tilde{S}$ คือ

$$Var\left(\tilde{S}\right)=\left(\sqrt{\frac{n}{2}}\frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}\right)^2\left(\left(1-\frac{1}{n}\right)-\frac{2}{n}\left[\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}\right]^2\right)\sigma^2$$

$$=\left(\frac{(n-1)}{2}\left(\frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}\right)^2-1\right)\sigma^2$$

3.3.3 สมบัติของตัวประมาณที่มีสมบัติไม่เอนเอียง ($\hat{S}$ และ $\tilde{S}$)

จาก 3.3.1

พบว่า $\hat{S}=\frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}\sqrt{\frac{1}{2}\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2}$ และ $Var\left(\hat{S}\right)=\left(\frac{(n-1)}{2}\left(\frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}\right)^2-1\right)\sigma^2$

และ 3.3.2

พบว่า $\tilde{S} = \dfrac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} \sqrt{\frac{1}{2}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}$ และ $Var\left(\tilde{S}\right) = \left(\dfrac{(n-1)}{2} \left(\dfrac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} \right)^2 - 1 \right) \sigma^2$

ผู้วิจัยได้พบว่า $\hat{S} = \tilde{S}$ และยิ่งไปกว่านั้น $Var\left(\hat{S}\right) = Var\left(\tilde{S}\right)$

จึงสรุปว่า ตัวประมาณที่จะศึกษาต่อคือ $\hat{S} = \dfrac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} \sqrt{\frac{1}{2}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}$ เท่านั้น

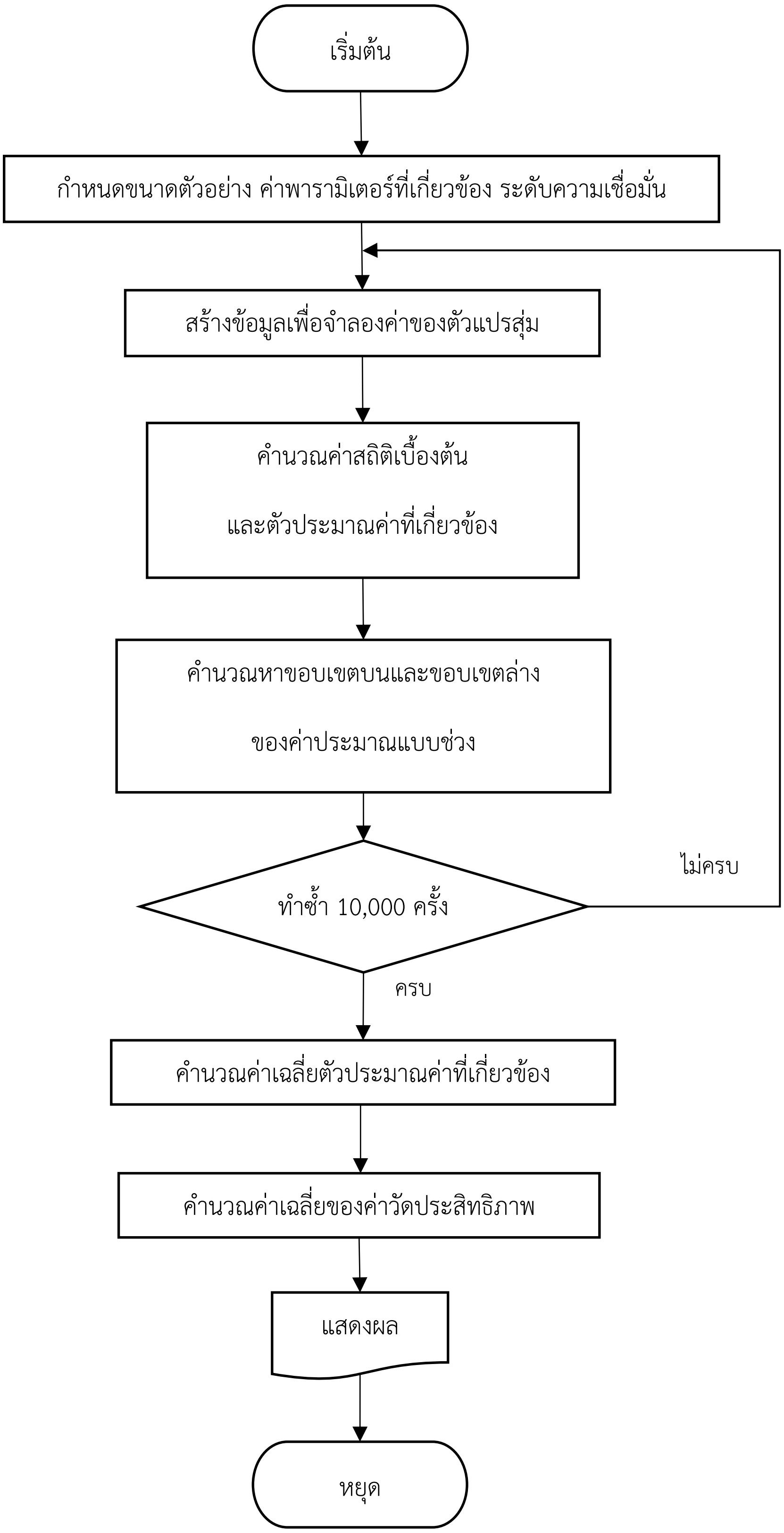
3.4 สรุปตัวประมาณค่า σ ความแปรปรวนของตัวประมาณ และค่าความเอนเอียงของตัวประมาณ

จากการศึกษาตัวประมาณแบบดั้งเดิม (S และ S_{MLE}) และตัวประมาณที่มีสมบัติไม่เอนเอียง ($\hat{S}$) ในครั้งนี้ สามารถสรุปสูตรได้ดังตารางที่ 3.1

ตารางที่ 3.1 เปรียบเทียบตัวประมาณส่วนเบี่ยงเบนมาตรฐาน

ตัวประมาณ	ค่าความเอนเอียงของตัวประมาณ	ความแปรปรวนของตัวประมาณ
$S=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2}$	$Bias(S)=\left[\sqrt{\frac{2}{n-1}}\cdot\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}-1\right]\sigma$	$Var(S)=\left[1-\frac{2}{(n-1)}\left(\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}\right)^2\right]\sigma^2$
$S_{MLE}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2}$	$Bias(S_{MLE})=\left[\sqrt{\frac{2}{n}}\cdot\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}-1\right]\sigma$	$Var(S_{MLE})=\left[\left(1-\frac{1}{n}\right)-\frac{2}{n}\left(\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}\right)\right]\sigma^2$
$\hat{S}=\frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}\sqrt{\frac{1}{2}\sum_{i=1}^{n}(X_i-\bar{X})^2}$	$Bias(\hat{S})=0$ ($\hat{S}$ เป็นตัวประมาณที่ไม่เอนเอียง)	$Var(\hat{S})=\left[\frac{(n-1)}{2}\left(\frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}\right)^2-1\right]\sigma^2$

3.5 ผังการจำลองข้อมูล



รูปที่ 3.1 ผังการจำลองข้อมูล

บทที่ 4

ผลจากการดำเนินงานวิจัย

ผลการวิจัยที่นำเสนอในหัวข้อต่อไปนี้ได้จากการจำลองข้อมูลที่มีการแจกแจงปรกติโดยมีการทำซ้ำ 10,000 รอบในแต่ละสถานการณ์ ผลการจำลองแบ่งเป็น 3 ตอนดังนี้

4.1 ประสิทธิภาพของตัวประมาณแบบจุดจากค่าความเอนเอียง (Bias) และค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Mean squared error)

4.2 ประสิทธิภาพของช่วงความเชื่อมั่นสำหรับส่วนเบี่ยงเบนมาตรฐานประชากรโดยพิจารณาจากค่าความน่าจะเป็นคุ้มรวม (Coverage probability) และค่าเฉลี่ยความกว้างของช่วงความเชื่อมั่น (Expected length)

4.3 ตัวอย่างการคำนวณจากข้อมูลจริง

สำหรับสัญลักษณ์ที่ใช้ในตารางมีความหมายดังนี้

S หมายถึงตัวประมาณ $S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2}$

S.mle หมายถึงตัวประมาณ $S_{MLE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2}$

S.hat หมายถึงตัวประมาณ $\hat{S} = \frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}\sqrt{\frac{1}{2}\sum_{i=1}^{n}(X_i - \bar{X})^2}$

4.1 ประสิทธิภาพของตัวประมาณแบบจุดจากค่าความเอนเอียง (Bias) และค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Mean squared error)

ตารางที่ 4.1 4.2 และ 4.3 แสดงค่าความเอนเอียง ความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ย กรณี $\mu = 0$ $\mu = 5$ และ $\mu = -5$ ตามลำดับ ผลการวิจัยในกรณี กรณี $\mu = 0$ พบว่าค่าความเอนเอียง ของตัวประมาณ $\hat{S}$ มีค่าน้อยที่สุดในทุกกรณีศึกษา และ 2 ตัวประมาณที่เหลือมีค่าความเอนเอียงมากกว่า (ดูกราฟที่ 4.1-4.5) ซึ่งสอดคล้องกับผลจากการพิสูจน์ก่อนหน้านี้ เมื่อพิจารณาขนาดตัวอย่างน้อยกว่า 50 พบว่าค่าความแปรปรวนของตัวประมาณ (ดูกราฟที่ 4.6-4.10) และค่าคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณ $\hat{S}$ (ดูกราฟที่ 4.11-4.15) มีค่ามากกว่า 2 ตัวประมาณที่เหลือเล็กน้อย แต่เมื่อขนาดตัวอย่างที่ใหญ่ขึ้นพบว่าค่าความแปรปรวนของตัวประมาณและค่าคลาดเคลื่อนกำลังสองเฉลี่ยของทั้ง 3 ตัวประมาณมีค่าใกล้เคียงกันมาก ผู้วิจัยพบว่าผลลัพธ์ กรณีที่ $\mu = 5$ และ $\mu = -5$ ให้ข้อสรุปของผลลัพธ์สอดคล้องกันกับกรณี $\mu = 0$

ตารางที่ 4.1 ค่าความเอนเอียง ความแปรปรวน และความคลาดเคลื่อนกำลังสองเฉลี่ย กรณี $\mu = 0$

			Bias			Variance			MSE		
μ	σ	n	S	S.mle	S.hat	S	S.mle	S.hat	S	S.mle	S.hat
0	0.5	10	-0.0139	-0.0388	-0.0002	0.0132	0.0119	0.0139	0.0134	0.0134	0.0139
		20	-0.0053	-0.0178	0.0013	0.0065	0.0062	0.0067	0.0066	0.0065	0.0067
		30	-0.0045	-0.0128	-0.0002	0.0042	0.0041	0.0043	0.0042	0.0042	0.0043
		50	-0.0023	-0.0073	0.0002	0.0025	0.0025	0.0025	0.0025	0.0025	0.0025
		100	-0.0014	-0.0039	-0.0001	0.0013	0.0013	0.0013	0.0013	0.0013	0.0013
		300	-0.0003	-0.0012	0.0001	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004
	1	10	-0.0282	-0.0780	-0.0008	0.0533	0.0480	0.0564	0.0541	0.0541	0.0564
		20	-0.0096	-0.0347	0.0035	0.0263	0.0250	0.0270	0.0264	0.0262	0.0270
		30	-0.0104	-0.0270	-0.0018	0.0175	0.0169	0.0178	0.0176	0.0176	0.0178
		50	-0.0048	-0.0148	0.0003	0.0100	0.0098	0.0101	0.0100	0.0100	0.0101
		100	-0.0034	-0.0084	-0.0009	0.0050	0.0049	0.0050	0.0050	0.0050	0.0050
		300	-0.0009	-0.0026	-0.0001	0.0016	0.0016	0.0017	0.0016	0.0016	0.0017
	2	10	-0.0577	-0.1574	-0.0031	0.2167	0.1950	0.2291	0.2200	0.2198	0.2291
		20	-0.0277	-0.0776	-0.0016	0.1019	0.0968	0.1047	0.1027	0.1029	0.1047
		30	-0.0150	-0.0484	0.0022	0.0679	0.0656	0.0691	0.0681	0.0680	0.0691
		50	-0.0111	-0.0311	-0.0009	0.0409	0.0401	0.0413	0.0410	0.0410	0.0413
		100	-0.0066	-0.0166	-0.0016	0.0200	0.0198	0.0201	0.0200	0.0200	0.0201
		300	-0.0015	-0.0048	0.0002	0.0066	0.0066	0.0066	0.0066	0.0066	0.0066
	5	10	-0.1408	-0.3902	-0.0042	1.3112	1.1801	1.3859	1.3310	1.3323	1.3859
		20	-0.0619	-0.1869	0.0035	0.6421	0.6100	0.6592	0.6459	0.6449	0.6592
		30	-0.0469	-0.1302	-0.0040	0.4159	0.4020	0.4231	0.4181	0.4190	0.4231
		50	-0.0238	-0.0738	0.0016	0.2520	0.2470	0.2546	0.2526	0.2524	0.2546
		100	-0.0109	-0.0359	0.0017	0.1269	0.1256	0.1275	0.1270	0.1269	0.1275
		300	-0.0028	-0.0112	0.0014	0.0417	0.0415	0.0418	0.0417	0.0417	0.0418
	10	10	-0.2587	-0.7586	0.0152	5.3965	4.8569	5.7042	5.4634	5.4323	5.7044
		20	-0.1222	-0.3724	0.0085	2.5932	2.4635	2.6623	2.6081	2.6022	2.6624
		30	-0.0768	-0.2436	0.0091	1.7220	1.6646	1.7520	1.7279	1.7240	1.7521
		50	-0.0567	-0.1567	-0.0059	1.0157	0.9954	1.0261	1.0189	1.0199	1.0261
		100	-0.0086	-0.0587	0.0166	0.4973	0.4923	0.4998	0.4973	0.4957	0.5001
		300	-0.0075	-0.0242	0.0008	0.1674	0.1668	0.1677	0.1674	0.1674	0.1677

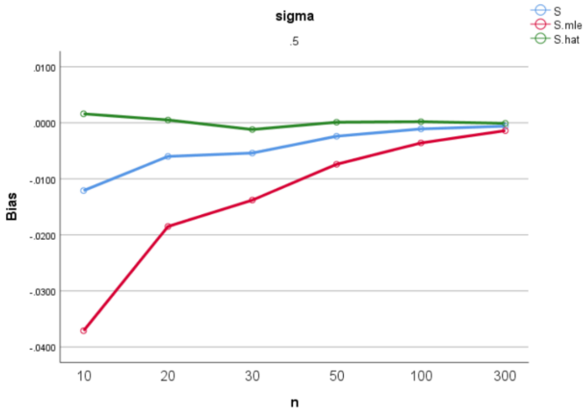
ตารางที่ 4.2 ค่าความเอนเอียง ความแปรปรวน และความคลาดเคลื่อนกำลังสองเฉลี่ย กรณี $\mu = 5$

			Bias			Variance			MSE		
μ	σ	n	S	S.mle	S.hat	S	S.mle	S.hat	S	S.mle	S.hat
.	0.5	10	-0.0121	-0.0371	0.0016	0.0136	0.0122	0.0143	0.0137	0.0123	0.0152
		20	-0.0060	-0.0185	0.0005	0.0063	0.0059	0.0064	0.0065	0.0062	0.0068
		30	-0.0054	-0.0138	-0.0012	0.0043	0.0041	0.0043	0.0043	0.0041	0.0044
		50	-0.0024	-0.0074	0.0001	0.0025	0.0024	0.0025	0.0025	0.0025	0.0026
		100	-0.0011	-0.0036	0.0002	0.0013	0.0013	0.0013	0.0013	0.0012	0.0013
		300	-0.0006	-0.0014	-0.0001	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004
	1	10	-0.0267	-0.0767	0.0007	0.0546	0.0491	0.0577	0.0547	0.0489	0.0604
		20	-0.0122	-0.0372	0.0009	0.0261	0.0248	0.0268	0.0262	0.0248	0.0274
		30	-0.0080	-0.0247	0.0006	0.0170	0.0164	0.0173	0.0172	0.0166	0.0177
		50	-0.0033	-0.0134	0.0018	0.0102	0.0100	0.0103	0.0102	0.0100	0.0104
		100	-0.0019	-0.0069	0.0006	0.0050	0.0050	0.0051	0.0050	0.0050	0.0051
		300	-0.0004	-0.0021	0.0004	0.0017	0.0017	0.0017	0.0017	0.0017	0.0017
	2	10	-0.0495	-0.1496	0.0054	0.2133	0.1920	0.2255	0.2195	0.1960	0.2421
		20	-0.0263	-0.0762	-0.0001	0.1044	0.0992	0.1072	0.1045	0.0991	0.1095
		30	-0.0181	-0.0514	-0.0009	0.0675	0.0652	0.0687	0.0686	0.0662	0.0707
		50	-0.0122	-0.0322	-0.0021	0.0411	0.0403	0.0416	0.0406	0.0398	0.0414
		100	-0.0056	-0.0156	-0.0006	0.0202	0.0200	0.0203	0.0202	0.0200	0.0203
		300	-0.0020	-0.0054	-0.0004	0.0067	0.0067	0.0068	0.0067	0.0067	0.0067
	5	10	-0.1265	-0.3766	0.0105	1.3441	1.2097	1.4207	1.3712	1.2239	1.5122
		20	-0.0804	-0.2050	-0.0153	0.6340	0.6023	0.6509	0.6489	0.6152	0.6796
		30	-0.0454	-0.1287	-0.0025	0.4253	0.4111	0.4327	0.4286	0.4139	0.4418
		50	-0.0215	-0.0715	0.0040	0.2541	0.2490	0.2567	0.2548	0.2497	0.2594
		100	-0.0097	-0.0347	0.0029	0.1264	0.1252	0.1271	0.1263	0.1250	0.1274
		300	-0.0044	-0.0128	-0.0002	0.0422	0.0421	0.0423	0.0418	0.0416	0.0419
	10	10	-0.2720	-0.7712	0.0015	5.3902	4.8512	5.6975	5.4655	4.8780	6.0273
		20	-0.1446	-0.3941	-0.0141	2.6124	2.4818	2.6820	2.6060	2.4705	2.7292
		30	-0.0968	-0.2632	-0.0111	1.7350	1.6772	1.7652	1.7130	1.6544	1.7656
		50	-0.0654	-0.1653	-0.0146	1.0332	1.0125	1.0438	1.0150	0.9943	1.0333
		100	-0.0335	-0.0834	-0.0083	0.5040	0.4990	0.5066	0.5036	0.4985	0.5080
		300	-0.0023	-0.0190	0.0061	0.1709	0.1704	0.1712	0.1674	0.1668	0.1679

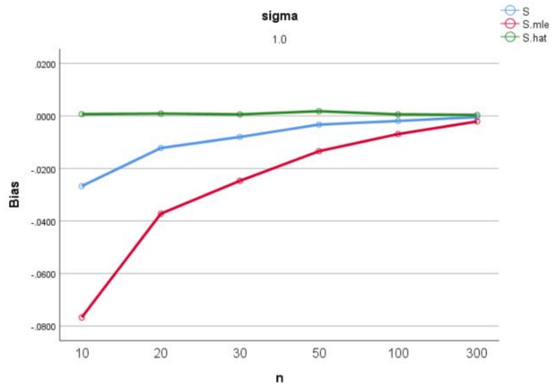
ตารางที่ 4.3 ค่าความเอนเอียง ความแปรปรวน และความคลาดเคลื่อนกำลังสองเฉลี่ย กรณี $\mu = -5$

			Bias			Variance			MSE		
μ	σ	n	S	S.mle	S.hat	S	S.mle	S.hat	S	S.mle	S.hat
-5	0.5	10	-0.0144	-0.0393	-0.0008	0.0135	0.0122	0.0143	0.0136	0.0122	0.0150
		20	-0.0055	-0.0180	0.0010	0.0065	0.0062	0.0067	0.0066	0.0062	0.0069
		30	-0.0030	-0.0113	0.0013	0.0043	0.0041	0.0044	0.0043	0.0042	0.0044
		50	-0.0022	-0.0072	0.0004	0.0025	0.0025	0.0025	0.0025	0.0025	0.0026
		100	-0.0011	-0.0037	0.0001	0.0013	0.0013	0.0013	0.0013	0.0012	0.0013
		300	-0.0004	-0.0012	0.0001	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004
	1	10	-0.0223	-0.0725	0.0052	0.0532	0.0479	0.0563	0.0551	0.0492	0.0608
		20	-0.0113	-0.0363	0.0018	0.0266	0.0253	0.0273	0.0262	0.0249	0.0275
		30	-0.0081	-0.0248	0.0004	0.0173	0.0167	0.0176	0.0172	0.0166	0.0177
		50	-0.0045	-0.0146	0.0005	0.0102	0.0100	0.0103	0.0102	0.0100	0.0104
		100	-0.0039	-0.0089	-0.0014	0.0051	0.0051	0.0052	0.0050	0.0050	0.0051
		300	-0.0001	-0.0017	0.0008	0.0017	0.0017	0.0017	0.0017	0.0017	0.0017
	2	10	-0.0618	-0.1612	-0.0073	0.2137	0.1923	0.2259	0.2169	0.1936	0.2393
		20	-0.0248	-0.0748	0.0014	0.1057	0.1004	0.1085	0.1047	0.0993	0.1096
		30	-0.0222	-0.0554	-0.0050	0.0678	0.0656	0.0690	0.0683	0.0660	0.0704
		50	-0.0086	-0.0286	0.0016	0.0411	0.0403	0.0416	0.0408	0.0400	0.0415
		100	-0.0059	-0.0159	-0.0008	0.0205	0.0203	0.0206	0.0202	0.0200	0.0203
		300	-0.0022	-0.0055	-0.0005	0.0068	0.0068	0.0068	0.0067	0.0067	0.0067
	5	10	-0.1496	-0.3985	-0.0133	1.3492	1.2143	1.4261	1.3592	1.2131	1.4989
		20	-0.0778	-0.2025	-0.0127	0.6547	0.6220	0.6722	0.6501	0.6163	0.6809
		30	-0.0458	-0.1291	-0.0029	0.4288	0.4145	0.4363	0.4286	0.4139	0.4418
		50	-0.0281	-0.0780	-0.0026	0.2562	0.2510	0.2588	0.2542	0.2490	0.2588
		100	-0.0082	-0.0332	0.0044	0.1258	0.1246	0.1265	0.1263	0.1250	0.1274
		300	-0.0061	-0.0144	-0.0019	0.0422	0.0421	0.0423	0.0418	0.0416	0.0419
	10	10	-0.2642	-0.7638	0.0095	5.4918	4.9426	5.8049	5.4792	4.8899	6.0426
		20	-0.1431	-0.3927	-0.0127	2.5909	2.4614	2.6600	2.6061	2.4707	2.7294
		30	-0.0991	-0.2656	-0.0134	1.7098	1.6528	1.7395	1.7117	1.6532	1.7643
		50	-0.0545	-0.1544	-0.0036	1.0354	1.0147	1.0461	1.0172	0.9965	1.0355
		100	-0.0257	-0.0757	-0.0005	0.4993	0.4943	0.5019	0.5043	0.4992	0.5088
		300	-0.0144	-0.0311	-0.0060	0.1709	0.1703	0.1712	0.1670	0.1664	0.1674

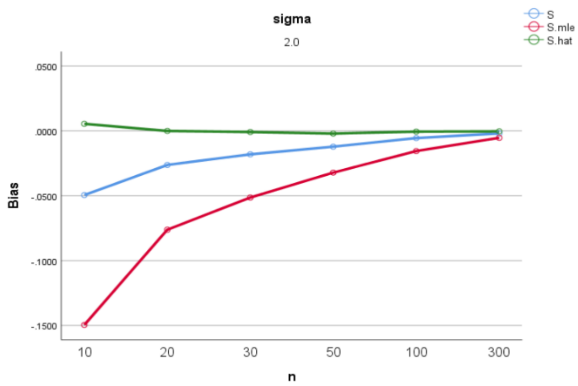
กราฟแสดงค่าความเอนเอียง กรณี $\mu = 0$



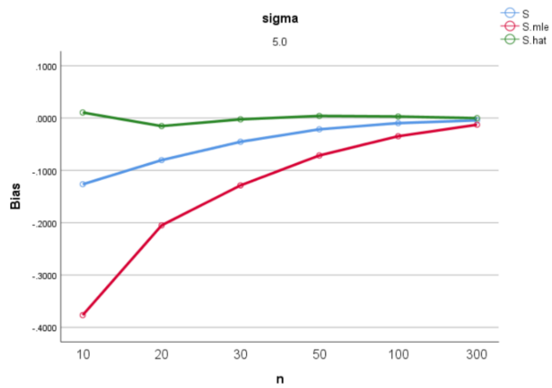
กราฟที่ 4.1 Bias กรณีที่ $\sigma = 0.5$



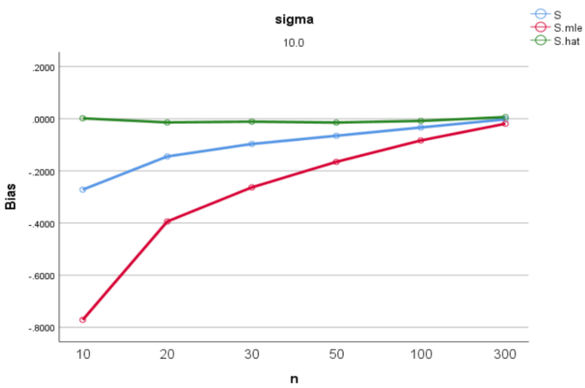
กราฟที่ 4.2 Bias กรณีที่ $\sigma = 1$



กราฟที่ 4.3 Bias กรณีที่ $\sigma = 2$

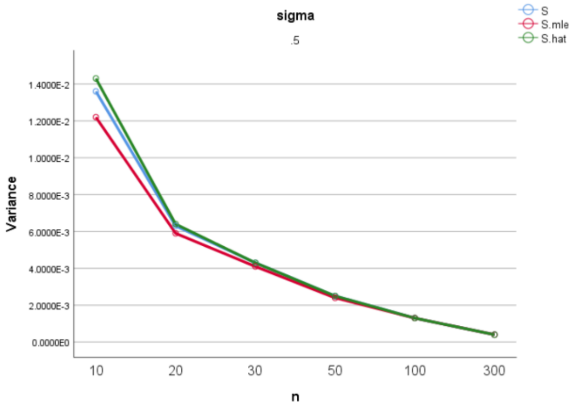


กราฟที่ 4.4 Bias กรณีที่ $\sigma = 5$

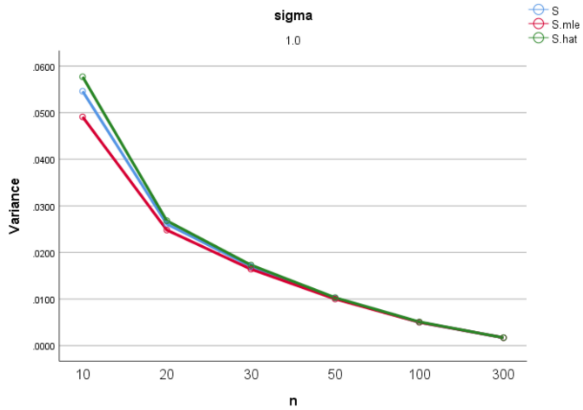


กราฟที่ 4.5 Bias กรณีที่ $\sigma = 10$

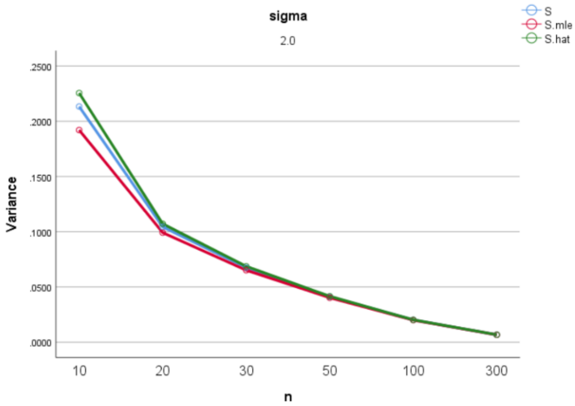
กราฟแสดงค่าความแปรปรวน กรณี $\mu = 0$



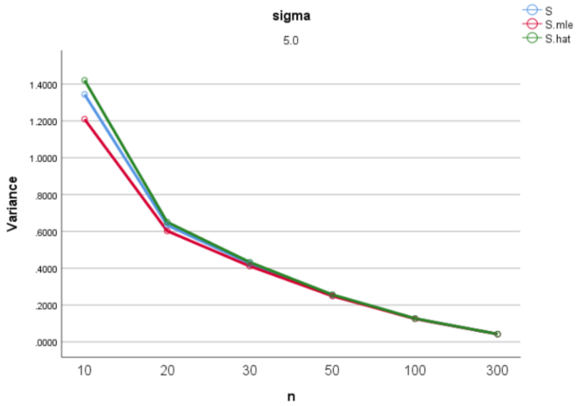
กราฟที่ 4.6 Variance กรณีที่ $\sigma = 0.5$



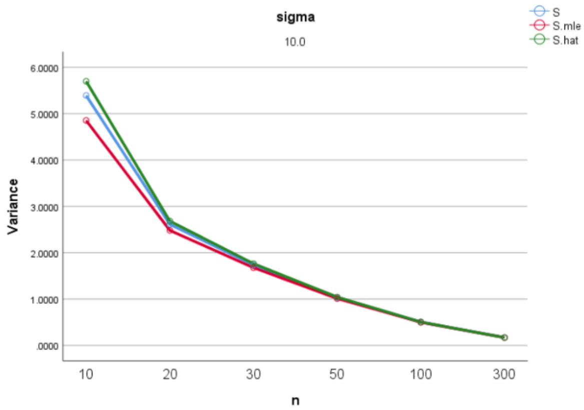
กราฟที่ 4.7 Variance กรณีที่ $\sigma = 1$



กราฟที่ 4.8 Variance กรณีที่ $\sigma = 2$

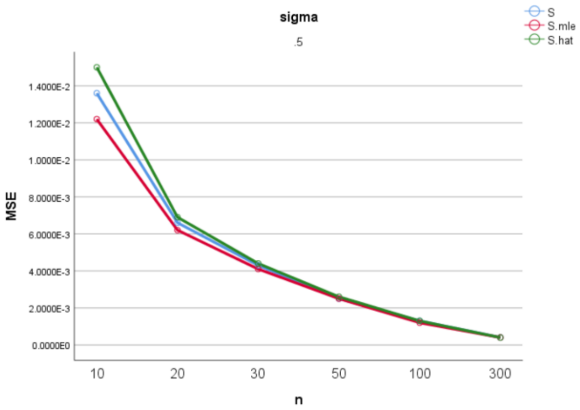


กราฟที่ 4.9 Variance กรณีที่ $\sigma = 5$

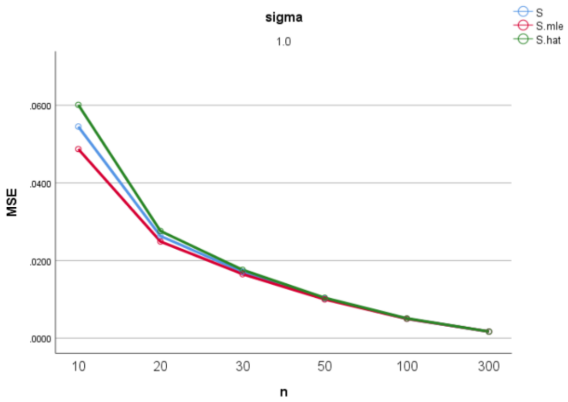


กราฟที่ 4.10 Variance กรณีที่ $\sigma = 10$

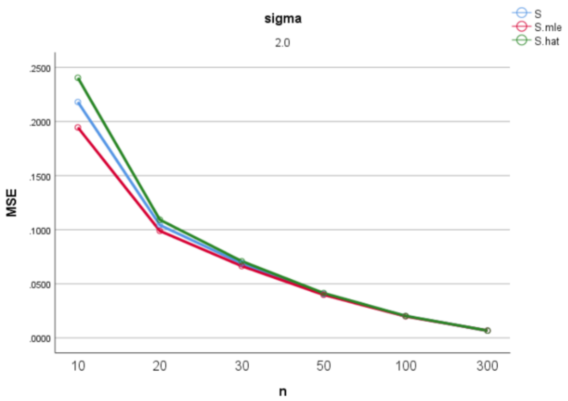
กราฟแสดงค่าความคลาดเคลื่อนกำลังสองเฉลี่ย กรณี $\mu = 0$



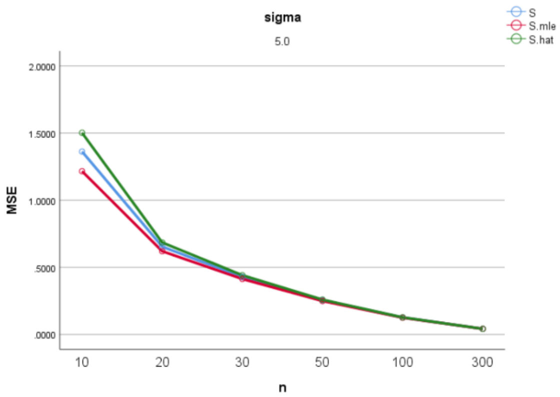
กราฟที่ 4.11 MSE กรณีที่ $\sigma = 0.5$



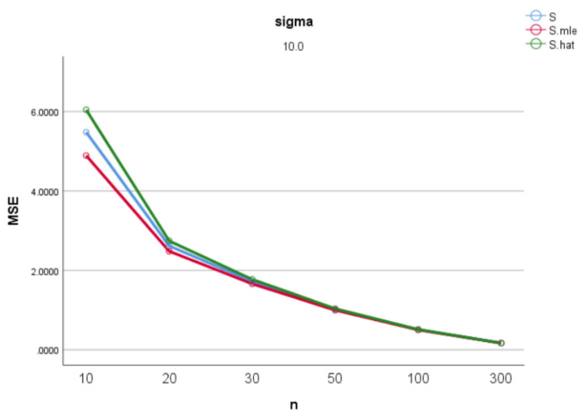
กราฟที่ 4.12 MSE กรณีที่ $\sigma = 1$



กราฟที่ 4.13 MSE กรณีที่ $\sigma = 2$



กราฟที่ 4.14 MSE กรณีที่ $\sigma = 5$



กราฟที่ 4.15 MSE กรณีที่ $\sigma = 10$

4.2 ประสิทธิภาพของช่วงความเชื่อมั่นสำหรับส่วนเบี่ยงเบนมาตรฐานประชากรโดยพิจารณาจากค่าความน่าจะเป็นคุ้มรวม (Coverage probability) และค่าเฉลี่ยความกว้างของช่วงความเชื่อมั่น (Expected length)

ตารางที่ 4.4 4.5 และ 4.6 แสดงค่าความน่าจะเป็นคุ้มรวม (CP) และค่าเฉลี่ยความกว้างของช่วงความเชื่อมั่น (EL) กรณี $\mu = 0$, กรณี $\mu = 5$ และ $\mu = -5$ ตามลำดับ กราฟที่ 4.16-4.20 แสดงค่าความน่าจะเป็นคุ้มรวม และ กราฟที่ 4.21-4.25 แสดงค่าเฉลี่ยความกว้างของช่วงความเชื่อมั่นสำหรับ

กำหนดสัญลักษณ์ที่ใช้ในตารางดังนี้

S หมายถึง ช่วงความเชื่อมั่น $(1-\alpha)100\%$ ที่อยู่ในรูปแบบ

$$S \pm t_{\alpha/2}\sqrt{Var(S)}$$

เมื่อ $S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2}$

S.mle หมายถึง ช่วงความเชื่อมั่น $(1-\alpha)100\%$ ที่อยู่ในรูปแบบ

$$S_{MLE} \pm t_{\alpha/2}\sqrt{Var(S_{MLE})}$$

เมื่อ $S_{MLE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2}$

S.hat หมายถึง ช่วงความเชื่อมั่น $(1-\alpha)100\%$ ที่อยู่ในรูปแบบ

$$\hat{S} \pm t_{\alpha/2}\sqrt{Var(\hat{S})}$$

เมื่อ $\hat{S} = \frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}\sqrt{\frac{1}{2}\sum_{i=1}^{n}(X_i - \bar{X})^2}$

S_chisq หมายถึง ช่วงความเชื่อมั่น $(1-\alpha)100\%$ ที่อยู่ในรูปแบบ

$$\frac{\sqrt{(n-1)}}{\sqrt{\chi^2_{\frac{\alpha}{2},n-1}}}S \le \sigma \le \frac{\sqrt{(n-1)}}{\sqrt{\chi^2_{1-\frac{\alpha}{2},n-1}}}S$$

ตารางที่ 4.4 ค่าความน่าจะเป็นคุ้มรวมและค่าเฉลี่ยความกว้างของช่วงความเชื่อมั่น กรณี $\mu = 0$

			CP				EL			
μ	σ	n	S	S.mle	S.hat	S	S.mle	S.hat	S	S.mle
0	0.5	10	0.9251	0.8850	0.9424	**0.9470**	0.5108	0.4597	0.5399	0.5528
		20	0.9367	0.9126	**0.9457**	**0.9495**	0.3337	0.3170	0.3426	0.3458
		30	0.9403	0.9241	**0.9484**	**0.9498**	0.2650	0.2561	0.2696	0.2710
		50	**0.9462**	0.9372	**0.9500**	**0.9503**	0.2015	0.1975	0.2036	0.2041
		100	0.9453	0.9402	**0.9474**	**0.9492**	0.1404	0.1390	0.1412	0.1415
		300	**0.9481**	**0.9479**	**0.9491**	0.9452	0.0804	0.0801	0.0805	0.0805
	1	10	0.9211	0.8794	0.9360	**0.9518**	1.0211	0.9190	1.0793	1.1069
		20	0.9375	0.9146	**0.9465**	**0.9511**	0.6680	0.6346	0.6858	0.6900
		30	0.9353	0.9177	0.9434	**0.9527**	0.5292	0.5115	0.5384	0.5426
		50	0.9439	0.9338	**0.9480**	**0.9495**	0.4030	0.3950	0.4072	0.4082
		100	**0.9478**	0.9432	**0.9496**	**0.9490**	0.2807	0.2779	0.2821	0.2831
		300	**0.9514**	**0.9500**	**0.9523**	**0.9495**	0.1607	0.1602	0.1610	0.1611
	2	10	0.9169	0.8753	0.9350	**0.9497**	2.0408	1.8367	2.1571	2.2207
		20	0.9324	0.9103	0.9441	**0.9514**	1.3303	1.2638	1.3657	1.3800
		30	0.9428	0.9248	**0.9503**	**0.9510**	1.0615	1.0261	1.0799	1.0858
		50	0.9419	0.9315	**0.9464**	**0.9495**	0.8054	0.7893	0.8137	0.8166
		100	**0.9483**	0.9441	**0.9500**	**0.9504**	0.5615	0.5559	0.5643	0.5656
		300	**0.9488**	**0.9465**	**0.9501**	**0.9492**	0.3215	0.3204	0.3221	0.3222
	5	10	0.9227	0.8813	0.9392	**0.9512**	5.1056	4.5951	5.3967	5.5347
		20	0.9374	0.9121	**0.9477**	**0.9470**	3.3306	3.1641	3.4194	3.4502
		30	0.9417	0.9270	**0.9480**	0.9547	2.6486	2.5604	2.6947	2.7086
		50	0.9451	0.9356	**0.9486**	**0.9504**	2.0151	1.9748	2.0358	2.0413
		100	**0.9489**	0.9419	**0.9513**	**0.9467**	1.4053	1.3912	1.4124	1.4151
		300	**0.9511**	**0.9495**	**0.9515**	**0.9481**	0.8040	0.8013	0.8053	0.8058
	10	10	0.9237	0.8789	0.9396	**0.9511**	10.2354	9.2118	10.8189	11.0877
		20	0.9354	0.9140	0.9442	**0.9494**	6.6623	6.3291	6.8398	6.9197
		30	0.9406	0.9245	**0.9494**	**0.9523**	5.3064	5.1295	5.3986	5.4356
		50	0.9428	0.9339	**0.9471**	**0.9529**	4.0265	3.9460	4.0678	4.0766
		100	**0.9489**	**0.9462**	**0.9505**	**0.9487**	2.8142	2.7861	2.8285	2.8320
		300	**0.9471**	**0.9462**	**0.9479**	**0.9491**	1.6076	1.6023	1.6103	1.6115

หมายเหตุ ตัวพิมพ์หนาหมายถึงค่าความน่าจะเป็นคุ้มรวมที่อยู่ในช่วงความเชื่อมั่น $[0.9457, 0.9543]$

ตารางที่ 4.5 ค่าความน่าจะเป็นคุ้มรวมและค่าเฉลี่ยความกว้างของช่วงความเชื่อมั่น กรณี $\mu = 5$

			CP				EL			
μ	σ	n	S	S.mle	S.hat	S	S.mle	S.hat	S	S.mle
5	0.5	10	0.9244	0.8846	0.9387	**0.9468**	0.5127	0.4614	0.5419	0.5545
		20	0.9390	0.9159	**0.9465**	**0.9515**	0.3332	0.3165	0.3421	0.3450
		30	0.9385	0.9224	**0.9461**	**0.9510**	0.2645	0.2556	0.2691	0.2712
		50	**0.9472**	0.9375	**0.9512**	**0.9509**	0.2015	0.1975	0.2036	0.2043
		100	0.9448	0.9398	**0.9471**	**0.9532**	0.1405	0.1391	0.1412	0.1417
		300	**0.9522**	**0.9506**	**0.9531**	**0.9504**	0.0804	0.0801	0.0805	0.0806
	1	10	0.9209	0.8793	0.9392	**0.9510**	1.0233	0.9209	1.0816	1.1008
		20	0.9342	0.9111	0.9431	**0.9527**	0.6654	0.6322	0.6832	0.6896
		30	0.9433	0.9274	**0.9506**	**0.9494**	0.5306	0.5129	0.5398	0.5434
		50	0.9429	0.9332	**0.9466**	**0.9477**	0.4023	0.3943	0.4064	0.4092
		100	**0.9466**	0.9412	**0.9484**	**0.9486**	0.2806	0.2778	0.2820	0.2826
		300	**0.9483**	**0.9475**	**0.9490**	**0.9504**	0.1609	0.1603	0.1611	0.1611
	2	10	0.9175	0.8799	0.9326	**0.9491**	2.0447	1.8403	2.1613	2.2143
		20	0.9310	0.9080	0.9403	**0.9482**	1.3273	1.2610	1.3627	1.3830
		30	0.9411	0.9231	**0.9487**	**0.9482**	1.0602	1.0248	1.0786	1.0869
		50	0.9441	0.9346	**0.9491**	**0.9499**	0.8059	0.7898	0.8141	0.8193
		100	0.9453	0.9408	**0.9478**	0.9453	0.5617	0.5561	0.5646	0.5663
		300	**0.9499**	**0.9484**	**0.9506**	0.9446	0.3214	0.3203	0.3219	0.3223
	5	10	0.9168	0.8762	0.9330	**0.9480**	5.1007	4.5907	5.3915	5.5462
		20	0.9356	0.9110	0.9430	**0.9479**	3.3294	3.1629	3.4181	3.4540
		30	0.9418	0.9252	**0.9480**	**0.9512**	2.6495	2.5612	2.6956	2.7171
		50	**0.9498**	0.9402	**0.9537**	**0.9468**	2.0165	1.9762	2.0372	2.0456
		100	**0.9476**	0.9432	**0.9497**	**0.9480**	1.4057	1.3916	1.4128	1.4137
		300	**0.9484**	**0.9477**	**0.9497**	**0.9494**	0.8038	0.8011	0.8051	0.8060
	10	10	0.9233	0.8818	0.9385	**0.9499**	10.1996	9.1797	10.7811	11.0864
		20	0.9351	0.9119	0.9449	**0.9491**	6.6754	6.3416	6.8533	6.9219
		30	0.9359	0.9206	0.9421	**0.9532**	5.3024	5.1256	5.3946	5.4318
		50	0.9409	0.9314	**0.9457**	**0.9479**	4.0270	3.9465	4.0683	4.0831
		100	0.9454	0.9396	**0.9479**	**0.9525**	2.8074	2.7794	2.8216	2.8330
		300	**0.9499**	**0.9496**	**0.9502**	**0.9493**	1.6078	1.6025	1.6105	1.6097

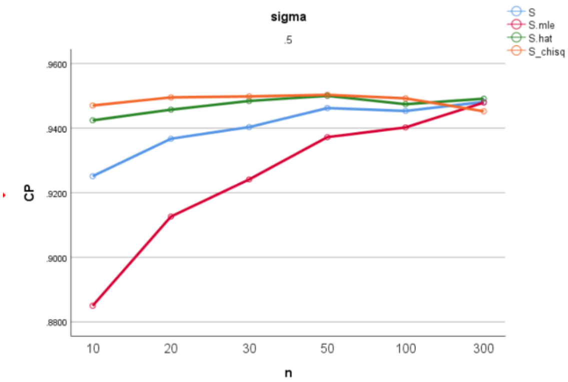
หมายเหตุ ตัวพิมพ์หนาหมายถึงค่าความน่าจะเป็นคุ้มรวมที่อยู่ในช่วงความเชื่อมั่น $[0.9457, 0.9543]$

ตารางที่ 4.6 ค่าความน่าจะเป็นคุ้มรวมและค่าเฉลี่ยความกว้างของช่วงความเชื่อมั่น กรณี $\mu = -5$

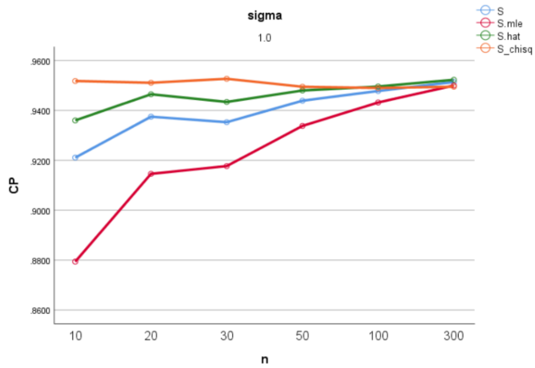
			CP				EL			
μ	σ	n	S	S.mle	S.hat	S	S.mle	S.hat	S	S.mle
-5	0.5	10	0.9185	0.8776	0.9343	**0.9495**	0.5102	0.4592	0.5393	0.5520
		20	0.9392	0.9164	**0.9477**	**0.9497**	0.3335	0.3168	0.3424	0.3452
		30	0.9440	0.9273	**0.9514**	**0.9524**	0.2658	0.2569	0.2704	0.2720
		50	0.9454	0.9342	**0.9499**	**0.9504**	0.2016	0.1976	0.2037	0.2045
		100	0.9443	0.9393	**0.9470**	**0.9465**	0.1405	0.1391	0.1412	0.1415
		300	0.9443	0.9414	0.9444	**0.9466**	0.0804	0.0801	0.0805	0.0806
	1	10	0.9264	0.8861	0.9410	**0.9487**	1.0273	0.9245	1.0858	1.1044
		20	0.9337	0.9103	0.9423	**0.9484**	0.6669	0.6335	0.6847	0.6919
		30	0.9377	0.9210	0.9446	**0.9513**	0.5304	0.5127	0.5396	0.5428
		50	**0.9459**	0.9345	**0.9495**	**0.9477**	0.4031	0.3950	0.4072	0.4089
		100	**0.9466**	0.9402	**0.9484**	**0.9516**	0.2806	0.2778	0.2820	0.2829
		300	**0.9520**	**0.9499**	**0.9525**	0.9544	0.1609	0.1603	0.1611	0.1611
	2	10	0.9203	0.8773	0.9371	0.9452	2.0365	1.8329	2.1526	2.2122
		20	0.9318	0.9086	0.9412	**0.9495**	1.3322	1.2656	1.3677	1.3818
		30	0.9351	0.9215	0.9428	**0.9487**	1.0576	1.0224	1.0760	1.0857
		50	0.9444	0.9334	**0.9481**	**0.9497**	0.8064	0.7903	0.8147	0.8173
		100	0.9427	0.9365	**0.9466**	**0.9530**	0.5617	0.5561	0.5645	0.5651
		300	**0.9464**	0.9428	**0.9467**	**0.9477**	0.3214	0.3203	0.3220	0.3224
	5	10	0.9183	0.8765	0.9361	**0.9498**	5.0964	4.5867	5.3869	5.5437
		20	0.9340	0.9081	0.9425	**0.9503**	3.3199	3.1539	3.4083	3.4418
		30	0.9412	0.9255	**0.9485**	**0.9486**	2.6492	2.5609	2.6953	2.7136
		50	0.9415	0.9315	**0.9461**	**0.9532**	2.0134	1.9731	2.0340	2.0432
		100	**0.9493**	0.9449	**0.9520**	**0.9498**	1.4060	1.3920	1.4131	1.4153
		300	**0.9495**	**0.9475**	**0.9499**	**0.9520**	0.8034	0.8007	0.8048	0.8058
	10	10	0.9152	0.8758	0.9331	**0.9493**	10.2296	9.2066	10.8128	11.0729
		20	0.9369	0.9144	0.9455	**0.9496**	6.6482	6.3157	6.8253	6.9187
		30	0.9380	0.9210	0.9434	**0.9505**	5.2944	5.1179	5.3865	5.4343
		50	0.9409	0.9318	0.9456	**0.9474**	4.0274	3.9469	4.0687	4.0908
		100	**0.9463**	0.9407	**0.9481**	**0.9512**	2.8094	2.7813	2.8236	2.8308
		300	0.9449	0.9431	**0.9462**	**0.9480**	1.6065	1.6011	1.6092	1.6118

หมายเหตุ ตัวพิมพ์หนาหมายถึงค่าความน่าจะเป็นคุ้มรวมที่อยู่ในช่วงความเชื่อมั่น $[0.9457, 0.9543]$

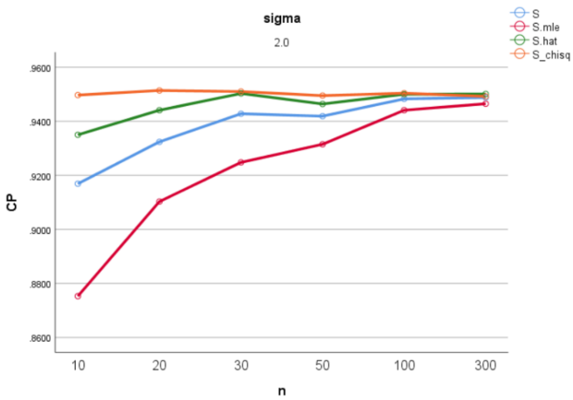
กราฟแสดงค่าความน่าจะเป็นคุ้มรวม กรณี $\mu = 0$



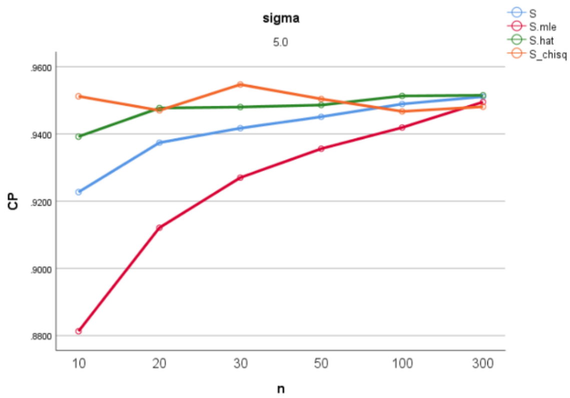
กราฟที่ 4.16 CP กรณีที่ $\sigma = 0.5$



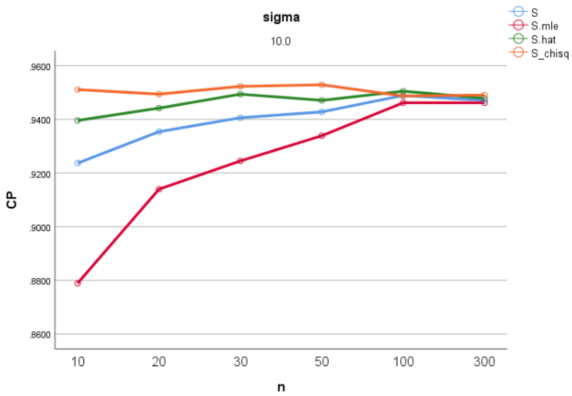
กราฟที่ 4.17 CP กรณีที่ $\sigma = 1$



กราฟที่ 4.18 CP กรณีที่ $\sigma = 2$

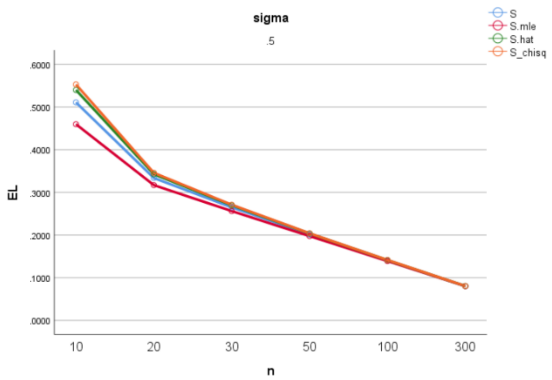


กราฟที่ 4.19 CP กรณีที่ $\sigma = 5$

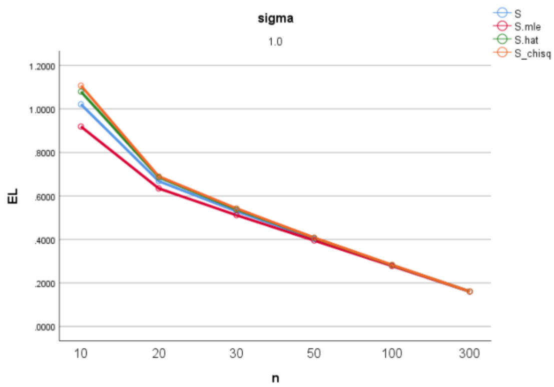


กราฟที่ 4.20 CP กรณีที่ $\sigma = 10$

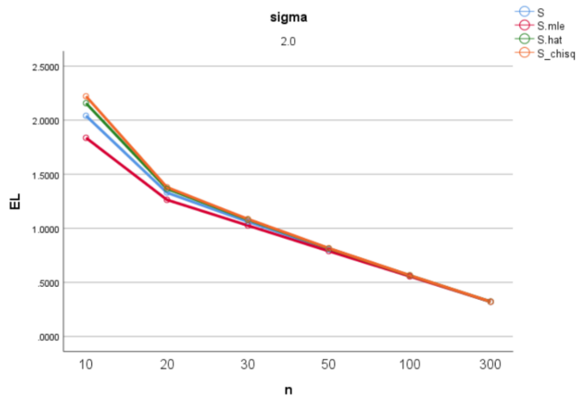
กราฟแสดงค่าเฉลี่ยความกว้างของช่วงความเชื่อมั่นช่วงกรณี $\mu = 0$



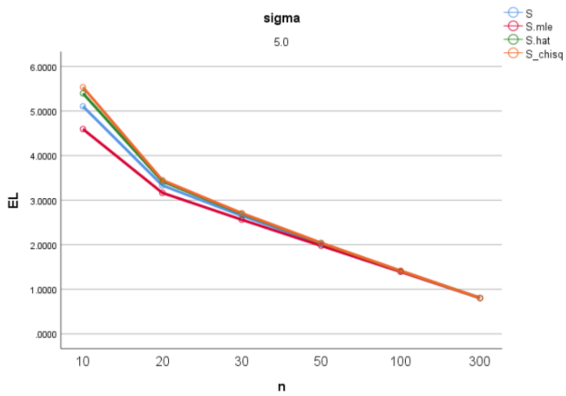
กราฟที่ 4.21 EL กรณีที่ $\sigma = 0.5$



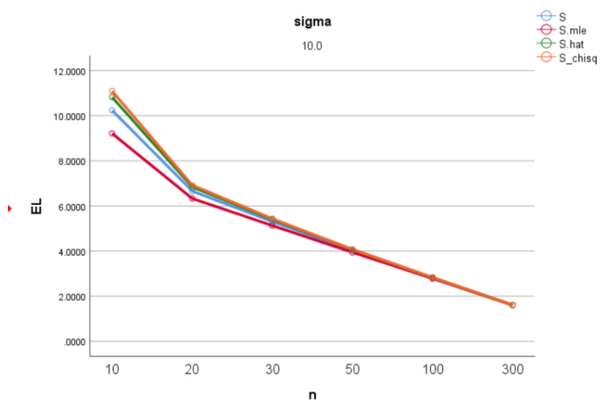
กราฟที่ 4.22 EL กรณีที่ $\sigma = 1$



กราฟที่ 4.23 EL กรณีที่ $\sigma = 2$



กราฟที่ 4.24 EL กรณีที่ $\sigma = 5$



กราฟที่ 4.25 EL กรณีที่ $\sigma = 10$

จากตารางและกราฟกรณี $\mu = 0$ ที่แสดงข้างต้น

ผลที่ได้พบว่า

ค่าความน่าจะเป็นคุ้มรวมของตัวประมาณ $\hat{S}$ ให้ค่าที่สูงกว่าวิธีของตัวประมาณ S และ S_{MLE} ในทุกกรณี โดยเฉพาะอย่างยิ่งค่าความน่าจะเป็นคุ้มรวมของตัวประมาณ S และ S_{MLE} โดยส่วนใหญ่มีค่าที่ห่างจาก 0.95 อย่างมากทำให้ส่งผลต่อการพิจารณาค่าความกว้างของช่วงความเชื่อมั่น

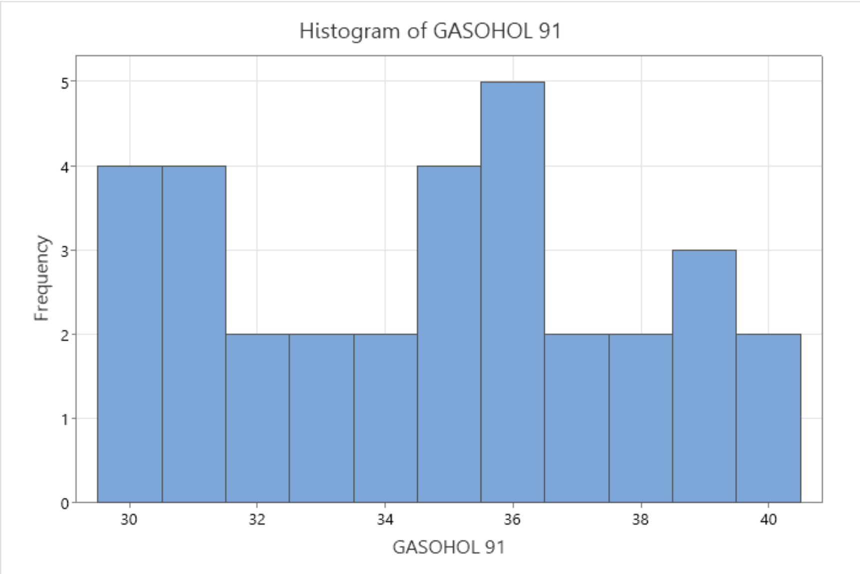
ถ้าหากพิจารณาโดยใช้ช่วงความเชื่อมั่นของค่าความน่าจะเป็นคุ้มรวม ผู้วิจัยพบว่าตัวประมาณ S ที่ใช้ตัวสถิติไคกำลังสอง ให้ค่าความน่าจะเป็นคุ้มรวมตกอยู่ในช่วง $[0.9457, 0.9543]$ เป็นส่วนใหญ่มากกว่าตัวประมาณ $\hat{S}$ แต่ตัวประมาณ $\hat{S}$ ก็ยังให้ค่าความน่าจะเป็นคุ้มรวมตกอยู่ในช่วง $[0.9457, 0.9543]$ เป็นส่วนใหญ่มากกว่าตัวประมาณ S และ S_{MLE}

เนื่องจากค่าความน่าจะเป็นคุ้มรวมของช่วงความเชื่อมั่นของตัวประมาณ S และ S_{MLE} ต่ำกว่าที่ควรจะเป็นจึงจะไม่นำค่าความกว้างของช่วงมาเปรียบเทียบ จากนั้นเปรียบเทียบวิธีที่เหลืออยู่ได้ข้อสรุปว่าค่าความกว้างของช่วงความเชื่อมั่นของตัวประมาณ S ที่ใช้ตัวสถิติไคกำลังสอง มีค่ามากกว่า ตัวประมาณ $\hat{S}$ ในทุกกรณี

ผู้วิจัยพบว่าผลลัพธ์ กรณี $\mu = 5$ และ $\mu = -5$ ให้ข้อสรุปของผลลัพธ์สอดคล้องกันกับกรณี $\mu = 0$ จากผลลัพธ์และข้อสรุปใน (4.1) และ (4.2) จึงกล่าวได้ว่าวิธีที่นำเสนอตัวประมาณแบบจุด $\hat{S}$ มีประสิทธิภาพที่ดีกว่าและตัวประมาณแบบช่วง S ที่ใช้ตัวสถิติไคกำลังสอง ซึ่งนิยมใช้มีประสิทธิภาพที่ดีกว่า

4.3 ตัวอย่างการคำนวณจากข้อมูลจริง

จากข้อมูลจริงของราคาขายปลีกรายวันของน้ำมันแก๊สโซฮอล 91 (มีหน่วยเป็น บาท/ลิตร) ซึ่งนับจากจำนวนวันที่ราคาน้ำมันมีการเปลี่ยนแปลง ในกรุงเทพมหานคร และปริมณฑลในเดือนธันวาคม พ.ศ.2564 ถึงมีนาคม พ.ศ.2565 จากเว็บไซต์บริษัท ปตท. น้ำมันและการค้าปลีก จำกัด (มหาชน) รวมทั้งสิ้นจำนวน 32 วัน ดังแสดงในกราฟที่ 4.26 ผู้วิจัยเลือกข้อมูลชุดนี้มาศึกษาเนื่องจากในช่วงเวลาดังกล่าวมีเหตุการณ์ที่กระทบต่อราคาน้ำมันในตลาดโลก การคำนวณค่าส่วนเบี่ยงเบนมาตรฐานจะทำให้ทราบถึงความผันผวนของข้อมูลเบื้องต้นได้



กราฟที่ 4.26 กราฟแสดงแผนภูมิข้อมูลจริงของราคาน้ำมันแก๊สโซฮอล 91

เมื่อทำการทดสอบการแจกแจงของข้อมูลโดยใช้สถิติทดสอบจาก Anderson-Darling ดังกราฟที่ 4.2 สรุปได้ว่า ข้อมูลมีการแจกแจงแบบปรกติ เนื่องจาก $p-value$ มีค่าเท่ากับ 0.184 พบว่า $p-value$ มีค่ามากกว่า ระดับนัยสำคัญ 0.05 ซึ่งให้ผลสอดคล้องกับแผนภูมิซึ่งมีลักษณะสมมาตร เมื่อทำการหาส่วนเบี่ยงเบนมาตรฐาน คำนวณได้ค่าดังนี้

$S = 3.2009$ $\qquad S_{MLE} = 3.1505$ $\qquad \hat{S} = 3.2269$ $\qquad$ (หน่วย : บาทต่อลิตร)

จากนั้นคำนวณช่วงความเชื่อมั่นได้ค่าดังนี้

ช่วงความเชื่อมั่น 95% ของ σ โดยใช้ตัวประมาณ S คือ $(2.3752, 4.0266)$

ค่าความกว้างของช่วง คือ 1.6514

ช่วงความเชื่อมั่น 95% ของ σ โดยใช้ตัวประมาณ S_{MLE} คือ $(2.3506, 3.9504)$

ค่าความกว้างของช่วง คือ 1.5998

ช่วงความเชื่อมั่น 95% ของ σ โดยใช้ตัวประมาณ $\hat{S}$ คือ $(2.3877, 4.0660)$

ค่าความกว้างของช่วง คือ 1.6782

ช่วงความเชื่อมั่น 95% ของ σ โดยใช้ตัวประมาณ S ที่ใช้สถิติไคกำลังสอง คือ $(2.5662, 4.2556)$

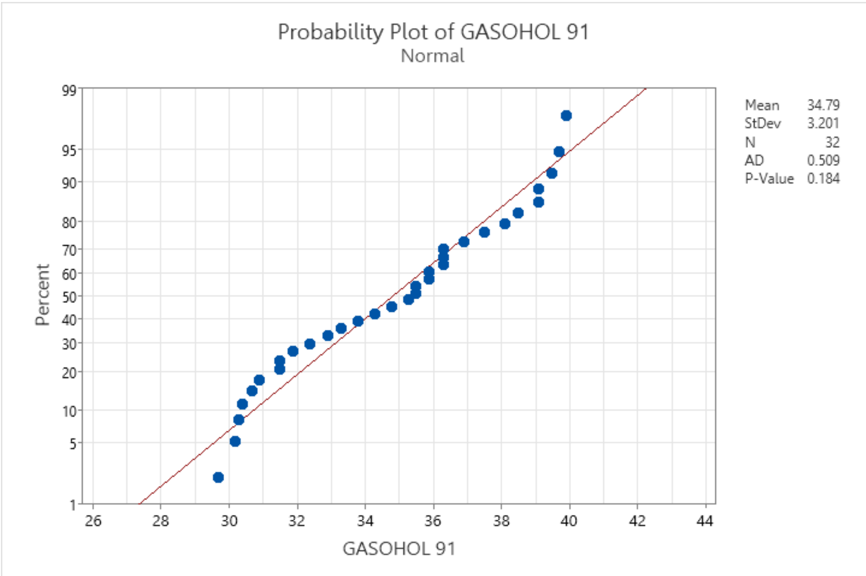
ค่าความกว้างของช่วง คือ 1.6894

ผลจากการคำนวณด้วยข้อมูลจริงข้างต้นพบว่าสอดคล้องกับผลจากการจำลองข้อมูล

เมื่อทำการค่าเฉลี่ยของข้อมูลคำนวณได้ค่าดังนี้ $\bar{x} = 34.7894$ (หน่วย : บาทต่อลิตร)

จากนั้นคำนวณสัมประสิทธิ์ของการแปรผันจาก $CV = \frac{\hat{S}}{\bar{x}} \times 100\% = \frac{3.2269}{34.7894} \times 100\% = 9.2755\%$

จะเห็นได้ว่า ราคาน้ำมันในช่วงเดือน ธันวาคม-มีนาคม มีความผันผวนปานกลาง



กราฟที่ 4.27 กราฟแสดงการทดสอบการแจกแจงแบบปรกติ

บทที่ 5

สรุปและอภิปรายผลการศึกษา

5.1 สรุปผลการวิจัย

จากการศึกษาค่าประมาณแบบจุดสำหรับส่วนเบี่ยงเบนมาตรฐานในการแจกแจงปรกติ โดยวิธีที่นำเสนอ คือ ตัวประมาณ $\hat{S}$ เขียนโปรแกรมในการจำลองข้อมูลและตรวจสอบข้อมูลที่จำลอง มีการแจกแจงปรกติ คำนวณค่าความเอนเอียง ค่าความแปรปรวน และค่าคลาดเคลื่อนกำลังสองเฉลี่ย พบว่าค่าความเอนเอียงน้อยกว่าวิธีพื้นฐานและวิธีภาวะน่าจะเป็นสูงสุด แต่ค่าความแปรปรวนและค่าคลาดเคลื่อนกำลังสองเฉลี่ย มีค่าใกล้เคียงกัน เมื่อขนาดตัวอย่างมากขึ้น จากนั้นคำนวณหาค่าความน่าจะเป็นคุ้มรวมและความกว้างของช่วงความเชื่อมั่นพบว่า ค่าความน่าจะเป็นคุ้มรวมมีค่ามากกว่าวิธีพื้นฐานและวิธีภาวะน่าจะเป็นสูงสุด แต่ความกว้างของช่วงความเชื่อมั่นมีค่าสูงกว่าวิธีพื้นฐานและวิธีภาวะน่าจะเป็นสูงสุด เมื่อนำข้อมูลจริงที่มีการแจกแจงปรกติมาคำนวณด้วยโปรแกรม ก็ได้ผลลัพธ์สอดคล้องกันกับการจำลองข้อมูล

จากการวิจัยในครั้งนี้จึงสามารถสรุปได้ว่าตัวประมาณ $\hat{S}$ มีความเหมาะสมและสามารถนำเสนอมาใช้เป็นทางเลือกในการประมาณค่าส่วนเบี่ยงเบนมาตรฐานในการแจกแจงปรกติได้

5.2 ข้อเสนอแนะ

1. ศึกษาประสิทธิภาพสัมประสิทธิ์ของการแปรผันที่ใช้ส่วนเบี่ยงเบนมาตรฐานแบบที่นำเสนอ

2. ศึกษาความเป็นไปได้ในการใช้ตัวประมาณส่วนเบี่ยงเบนมาตรฐานแบบที่นำเสนอ เพื่อประมาณค่าความแปรปรวนประชากร

3. ในวิธีการสถิติเชิงอนุมาน อาจใช้ส่วนเบี่ยงเบนมาตรฐานแบบที่นำเสนอแทนแบบดั้งเดิมในการทดสอบสมมติฐาน

บรรณานุกรม

กมล บุษบา. (2564). *คณิตสถิติศาสตร์ 1.* กรุงเทพฯ: สำนักพิมพ์มหาวิทยาลัยธรรมศาสตร์.

จิรัชย์ สุขะเกตุ. (2548). *ความน่าจะเป็นและทฤษฎีสถิติเบื้องต้น.* กรุงเทพฯ: สำนักพิมพ์มหาวิทยาลัยเกษตรศาสตร์.

ดวงตา สิงหกลางพล. (2561). *การเปรียบเทียบวิธีการประมาณค่าแบบช่วงสำหรับค่าเฉลี่ยประชากรที่มีการแจกแจงปรกติของข้อมูลที่ถูกปัดเศษ.* doi:10.14457/TU.the.2018.698

ภทรวรรณ แสงนวกิจ. (2563). *เทคนิคการชักตัวอย่างเบื้องต้นและการสำรวจตัวอย่างเบื้องต้น.* มหาวิทยาลัยธรรมศาสตร์.

วรพจน์ แซ่หลี. (2554). การประมาณช่วงความเชื่อมั่นของสัมประสิทธิ์การแปรผันใหม่. *วารสารวิทยาศาสตร์ลาดกระบัง, 20,* 61-71.

วราฤทธิ์ พานิชกิจโกศลกุล. (2558). *คณิตสถิติศาสตร์: การทดสอบสมมติฐานเชิงสถิติ.* กรุงเทพฯ: สำนักพิมพ์

สุชาดา กีระนันทน์. (2545). *ทฤษฎีความน่าจะเป็น.* กรุงเทพฯ: โรงพิมพ์มหาวิทยาลัยศิลปากร.

Casella, G., & Berger, R. L. (2015). *Statistical Inference.* Pacific Grove: Duxbury.

Donner, A., & Zou, G. Y. (2010). Closed-form confidence intervals for functions of the normal mean and standard deviation. *Statistical Methods in Medical Research, 21,* 347-359.

Krithikadatta, J. (2014). *Normal Distribution. Journal of conservative dentistry : JCD,* 96-97. doi:10.4103/0972-0707.124171

Mahmoudvand, R., & Hassani, H. (2009). Two new confidence intervals for the coefficient of variation in a normal distribution. *Journal of Applied Statistics, 36,* 429-442.

Muraleedharan, K., & Raval, N. (2012). Estimation of Process Standard deviation. *International journal of Computational Mathematics and Numerical Simulations, 5,* 179-186.

Sangnawakij, P., & Niwitpong, S.-A. (2020). Interval Estimation forthe Common Coefficient of Variation of Gamma Distributions. *Thailand Statistician, 18*(3), 340-353.

Nadarajah, S. (2005). A generalized normal distribution. *Journal of Applied Statistics, 32,* 685-694. doi:10.1080/02664760500079464

ภาคผนวก

การตรวจสอบการแจกแจงของข้อมูลจริง

การจำลองข้อมูลในงานวิจัยครั้งนี้ ได้จำลองสถานการณ์ต่าง ๆ ที่เป็นการแจกแจงปรกติ ผู้วิจัยยังคงมีข้อสงสัยว่า ข้อมูลจริงมีการแจกแจงปรกติหรือไม่ ดังนั้นผู้วิจัยจึงทำการตรวจสอบข้อสงสัยดังกล่าวในหัวข้อนี้

การตรวจสอบการแจกแจงของข้อมูล จะใช้การทดสอบ Goodness of fit Test ซึ่งมีสมมติฐานที่พิจารณาดังนี้

$H_0:$ การแจกแจงข้อมูลจริงเป็นการแจกแจงปรกติ

$H_1:$ การแจกแจงข้อมูลจริงไม่เป็นการแจกแจงปรกติ

กำหนดระดับนัยสำคัญ α

โดยใช้สถิติทดสอบคือ $\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$

เมื่อ O_i คือ ค่าความถี่ที่ได้จากการสังเกต

E_i คือ ค่าความถี่ที่คาดหวัง

k คือ จำนวนกลุ่ม

สถิติทดสอบ χ^2 ดังกล่าว มีการแจกแจงไคกำลังสอง ที่มีองศาเสรีเท่ากับ $k-1$

การตัดสินใจในการทดสอบสมมติฐาน คือ จะทำการปฏิเสธ H_0 เมื่อ χ^2 ที่คำนวณได้มีค่ามากกว่าค่าวิกฤต $\chi^2_{\alpha, df=k-1}$ หรือในทางปฏิบัติอาจพิจารณาการปฏิเสธ H_0 ได้จากค่า $p-value < \alpha$

ผลลัพธ์ที่ได้เพิ่มเติม

ตารางแสดงค่าความน่าจะเป็นคุ้มรวมและความกว้างของช่วงความเชื่อมั่นกรณี $\mu = 0$ ที่ใช้ตัวสถิติ z ที่มาจากช่วงความเชื่อมั่น $(1-\alpha)100\%$ ที่อยู่ในรูปแบบ

$$\hat{\theta} \pm z_{\alpha/2}\sqrt{Var(\hat{\theta})}$$

เมื่อ $\hat{\theta}$ คือ ตัวประมาณแบบจุดใด ๆ ของส่วนเบี่ยงเบนมาตรฐานประชากร $Var(\hat{\theta})$ คือ ตัวประมาณค่าความแปรปรวนของ $\hat{\theta}$ และ $z_{\alpha/2}$ คือ เปอร์เซ็นไทล์ที่ $(\alpha/2)100\%$ ของการแจกแจงปรกติมาตรฐาน

			CP			EL		
μ	σ	n	S	S.mle	S.hat	S	S.mle	S.hat
5	0.5	10	0.8991	0.8484	0.9171	0.4442	0.3998	0.4695
		20	0.9273	0.9027	0.9374	0.3120	0.2964	0.3203
		30	0.9291	0.9133	0.9379	0.2534	0.2450	0.2578
		50	0.9427	0.9322	0.9473	0.1965	0.1926	0.1985
		100	0.9419	0.9373	0.9446	0.1388	0.1374	0.1395
		300	0.9514	0.9499	0.9526	0.0800	0.0798	0.0802
	1	10	0.8915	0.8409	0.9132	0.8860	0.7974	0.9366
		20	0.9231	0.8975	0.9341	0.6239	0.5927	0.6405
		30	0.9322	0.9172	0.9406	0.5083	0.4914	0.5172
		50	0.9396	0.9285	0.9443	0.3936	0.3858	0.3977
		100	0.9470	0.9425	0.9478	0.2777	0.2749	0.2791
		300	0.9510	0.9500	0.9516	0.1602	0.1596	0.1604
	2	10	0.8971	0.8501	0.9152	1.7757	1.5981	1.8769
		20	0.9183	0.8954	0.9296	1.2466	1.1843	1.2798
		30	0.9314	0.9153	0.9383	1.0156	0.9818	1.0333
		50	0.9352	0.9241	0.9393	0.7851	0.7694	0.7931
		100	0.9433	0.9364	0.9441	0.5549	0.5493	0.5577
		300	0.9461	0.9450	0.9477	0.3201	0.3191	0.3207
	5	10	0.8982	0.8495	0.9175	4.4366	3.9929	4.6895
		20	0.9189	0.8945	0.9297	3.1072	2.9518	3.1900
		30	0.9312	0.9108	0.9382	2.5390	2.4544	2.5831
		50	0.9412	0.9305	0.9446	1.9663	1.9269	1.9864
		100	0.9464	0.9428	0.9476	1.3884	1.3745	1.3955
		300	0.9490	0.9473	0.9507	0.8004	0.7978	0.8018
	10	10	0.8943	0.8440	0.9140	8.8559	7.9704	9.3608

			CP			EL		
μ	σ	n	S	S.mle	S.hat	S	S.mle	S.hat
5	10	20	0.9181	0.8929	0.9281	6.2246	5.9134	6.3905
		30	0.9288	0.9106	0.9357	5.0749	4.9057	5.1631
		50	0.9332	0.9230	0.9398	3.9237	3.8452	3.9639
		100	0.9462	0.9403	0.9474	2.7729	2.7452	2.7870
		300	0.9460	0.9437	0.9473	1.6019	1.5966	1.6046

ตารางที่ 8 ค่าความน่าจะเป็นคุ้มรวมและค่าเฉลี่ยความกว้างของช่วงความเชื่อมั่นกรณี $\mu = 5$

จากตาราง พบว่าค่าความน่าจะเป็นคุ้มรวมต่ำกว่าที่ควรจะเป็น

โปรแกรมที่ใช้ในการวิจัย

งานวิจัยครั้งนี้ใช้โปรแกรม R ในการจำลองข้อมูลเพื่อศึกษาประสิทธิภาพของตัวประมาณค่าส่วนเบี่ยงเบนมาตรฐานในการแจกแจงปรกติ ซึ่งต้องมีการกำหนดจำนวนรอบของการทำซ้ำ ขนาดของประชากร ขนาดตัวอย่าง นอกจากนั้นการจำลองข้อมูลยังคำนวณค่าความเอนเอียง ค่าความแปรปรวน ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย ค่าความน่าจะเป็นคุ้มรวมและค่าเฉลี่ยความกว้างของช่วงความเชื่อมั่น เพื่อใช้เป็นเกณฑ์ในการเปรียบเทียบประสิทธิภาพของตัวประมาณส่วนเบี่ยงเบนมาตรฐานที่นำเสนอกับตัวประมาณส่วนเบี่ยงเบนมาตรฐานโดยวิธีดั้งเดิมและวิธีภาวะน่าจะเป็นสูงสุด

ตัวอย่างฟังก์ชันสำหรับการจำลองข้อมูลเพื่อเปรียบเทียบประสิทธิภาพของตัวประมาณส่วนเบี่ยงเบนมาตรฐานที่นำเสนอกับตัวประมาณส่วนเบี่ยงเบนมาตรฐานโดยวิธีดั้งเดิมและวิธีภาวะน่าจะเป็นสูงสุด มีดังนี้

```
########## Part 1 ###########

fix(sdnorm)


##### Part 2:  run in R Editor #####

function (n,mu,sigma, alpha, b)

{

s = rep(0,b)

bias.s = rep(0,b)
```

```
var.s = rep(0,b)

s.mle = rep(0,b)

bias.s.mle = rep(0,b)

var.s.mle = rep(0,b)

s.hat = rep(0,b)

bias.s.hat = rep(0,b)

var.s.hat = rep(0,b)

lower.1 = rep(0,b)

upper.1 = rep(0,b)

length.1 = rep(0,b)

cp.1 = rep(0,b)


lower.2 = rep(0,b)

upper.2 = rep(0,b)

length.2 = rep(0,b)

cp.2 = rep(0,b)

lower.3 = rep(0,b)

upper.3 = rep(0,b)

length.3 = rep(0,b)

cp.3 = rep(0,b)

lower.4 = rep(0,b)

upper.4 = rep(0,b)

length.4 = rep(0,b)
```

```
cp.4 = rep(0,b)

t = qt(1-alpha/2,n-1)

x1 = qchisq(alpha/2,n-1)

x2 = qchisq(1-alpha/2,n-1)

n1 = gamma((n-1)/2)

n2 = gamma(n/2)


for (i in 1:b) {   ##start for loop

## Gen data for normal

x = rnorm(n, mu, sigma)

## Calculate point estimators

s[i] = sd(x)

s.mle[i]=((sqrt(n-1))/sqrt(n))*sd(x)

s.hat[i]=sqrt((n-1)/2)*(n1/n2)* s[i]

## Calculate theoretical bias and variance of estimators

bias.s[i] = (sqrt(2/(n-1)) * (n2/n1)-1) * s[i]

bias.s.mle[i] = (sqrt(2/n) * (n2/n1)-1)* s.mle[i]

bias.s.hat[i] = 0

var.s[i] = (1-2/(n-1)*(n2/n1)^2)*s[i]^2

var.s.mle[i] = ((1-(1/n))-(2/n)*(n2/n1)^2)*s.mle[i]^2

var.s.hat[i]=(((n-1)/2)*(n1/n2)^2-1)* s.hat[i]^2

### CP/EL/CI

lower.1[i] = s[i] - t*sqrt(var.s[i])
```

```
upper.1[i] = s[i] + t*sqrt(var.s[i])

length.1[i] = upper.1[i]-lower.1[i]

## Find CP

if(sigma >= lower.1[i] & sigma <= upper.1[i] ) {cp.1[i] = 1}

else {cp.1[i] = 0 }

lower.2[i] = s.mle[i] - t*sqrt(var.s.mle[i])

upper.2[i] = s.mle[i] + t*sqrt(var.s.mle[i])

length.2[i] = upper.2[i]-lower.2[i]

## Find CP

if(sigma >= lower.2[i] & sigma <= upper.2[i] ) {cp.2[i] = 1}

else {cp.2[i] = 0 }

lower.3[i] = s.hat[i] - t*sqrt(var.s.hat[i])

upper.3[i] = s.hat[i] + t*sqrt(var.s.hat[i])

length.3[i] = upper.3[i]-lower.3[i]

## Find CP

if(sigma >= lower.3[i] & sigma <= upper.3[i] ) {cp.3[i] = 1}

else {cp.3[i] = 0 }

### Calculate CI,CP,EL for S using chi square statistics

lower.4[i] = s[i]*sqrt(n-1)/sqrt(x2)

upper.4[i] = s[i]*sqrt(n-1)/sqrt(x1)

length.4[i] = upper.4[i]-lower.4[i]

## Find CP

if(sigma >= lower.4[i] & sigma <= upper.4[i] ) {cp.4[i] = 1}
```

```
else {cp.4[i] = 0 }

} ##end for loop


## Average values from b simulation runs

ave.s = mean(s)

ave.bias.s = mean(bias.s)

ave.var.s = mean(var.s)

ave.s.mle = mean(s.mle)

ave.bias.s.mle = mean(bias.s.mle)

ave.var.s.mle = mean(var.s.mle)

ave.s.hat = mean(s.hat)

ave.bias.s.hat = mean(bias.s.hat)

ave.var.s.hat = mean(var.s.hat)

## empirical value for bias and variance of estimator

em.ave.bias.s = mean(s)-sigma

em.ave.var.s = var(s)

em.ave.bias.s.mle = mean(s.mle)-sigma

em.ave.var.s.mle = var(s.mle)

em.ave.bias.s.hat = mean(s.hat)-sigma

em.ave.var.s.hat = var(s.hat)

## Calculate MSE

mse.s = mean(var.s) + (mean(bias.s))^2

mse.s.mle = mean(var.s.mle) + (mean(bias.s.mle))^2
```

```
mse.s.hat = mean(var.s.hat) + (mean(bias.s.hat))^2

## Calculate Empirical MSE

em.mse.s = em.ave.var.s + (em.ave.bias.s)^2

em.mse.s.mle = em.ave.var.s.mle + (em.ave.bias.s.mle)^2

em.mse.s.hat = em.ave.var.s.hat + (em.ave.bias.s.hat)^2

## Average values of CP/EL from b simulation runs

ave.cp.1 = mean(cp.1)

ave.length.1 = mean(length.1)

ave.cp.2 = mean(cp.2)

ave.length.2 = mean(length.2)

ave.cp.3 = mean(cp.3)

ave.length.3 = mean(length.3)

ave.cp.4 = mean(cp.4)

ave.length.4 = mean(length.4)

ave.cp.5 = mean(cp.5)

ave.length.5 = mean(length.5)


cat(n, '\t', mu, '\t', sigma, '\n',

ave.s, '\t', ave.s.mle , '\t', ave.s.hat ,'\n' ,

ave.bias.s, '\t',ave.bias.s.mle , '\t',ave.bias.s.hat ,'\n',

em.ave.bias.s, '\t',em.ave.bias.s.mle , '\t' , em.ave.bias.s.hat ,'\n' ,

ave.var.s ,'\t',ave.var.s.mle,  '\t', ave.var.s.hat,'\n',

em.ave.var.s,'\t', em.ave.var.s.mle, '\t', em.ave.var.s.hat,'\n',
```

```
mse.s,  '\t', mse.s.mle ,  '\t', mse.s.hat,'\n',

ave.cp.1, '\t', ave.cp.2, '\t', ave.cp.3, '\t', ave.cp.4 ,'\n',

ave.length.1, '\t' , ave.length.2, '\t' ,ave.length.3, '\t', ave.length.4 ,'\n' )

}



##### Part 3:  run in R console  #####

## Case mu=0

sdnorm(n=10 ,mu=0 ,sigma=0.5 ,alpha=0.05 ,b=10000)

sdnorm(n=20 ,mu=0 ,sigma=0.5 ,alpha=0.05 ,b=10000)

sdnorm(n=30 ,mu=0 ,sigma=0.5 ,alpha=0.05 ,b=10000)

sdnorm(n=50 ,mu=0 ,sigma=0.5 ,alpha=0.05 ,b=10000)

sdnorm(n=100 ,mu=0 ,sigma=0.5 ,alpha=0.05 ,b=10000)

sdnorm(n=300 ,mu=0 ,sigma=0.5 ,alpha=0.05 ,b=10000)


sdnorm(n=10 ,mu=0 ,sigma=1 ,alpha=0.05 ,b=10000)

sdnorm(n=20 ,mu=0 ,sigma=1 ,alpha=0.05 ,b=10000)

sdnorm(n=30 ,mu=0 ,sigma=1 ,alpha=0.05 ,b=10000)

sdnorm(n=50 ,mu=0 ,sigma=1 ,alpha=0.05 ,b=10000)

sdnorm(n=100 ,mu=0 ,sigma=1 ,alpha=0.05 ,b=10000)

sdnorm(n=300 ,mu=0 ,sigma=1 ,alpha=0.05 ,b=10000)


sdnorm(n=10 ,mu=0 ,sigma=2 ,alpha=0.05 ,b=10000)
```

```
sdnorm(n=20 ,mu=0 ,sigma=2 ,alpha=0.05 ,b=10000)

sdnorm(n=30 ,mu=0 ,sigma=2 ,alpha=0.05 ,b=10000)

sdnorm(n=50 ,mu=0 ,sigma=2 ,alpha=0.05 ,b=10000)

sdnorm(n=100 ,mu=0 ,sigma=2 ,alpha=0.05 ,b=10000)

sdnorm(n=300 ,mu=0 ,sigma=2 ,alpha=0.05 ,b=10000)


sdnorm(n=10 ,mu=0 ,sigma=5 ,alpha=0.05 ,b=10000)

sdnorm(n=20 ,mu=0 ,sigma=5 ,alpha=0.05 ,b=10000)

sdnorm(n=30 ,mu=0 ,sigma=5 ,alpha=0.05 ,b=10000)

sdnorm(n=50 ,mu=0 ,sigma=5 ,alpha=0.05 ,b=10000)

sdnorm(n=100 ,mu=0 ,sigma=5 ,alpha=0.05 ,b=10000)

sdnorm(n=300 ,mu=0 ,sigma=5 ,alpha=0.05 ,b=10000)


sdnorm(n=10 ,mu=0 ,sigma=10 ,alpha=0.05 ,b=10000)

sdnorm(n=20 ,mu=0 ,sigma=10 ,alpha=0.05 ,b=10000)

sdnorm(n=30 ,mu=0 ,sigma=10 ,alpha=0.05 ,b=10000)

sdnorm(n=50 ,mu=0 ,sigma=10 ,alpha=0.05 ,b=10000)

sdnorm(n=100 ,mu=0 ,sigma=10 ,alpha=0.05 ,b=10000)

sdnorm(n=300 ,mu=0 ,sigma=10 ,alpha=0.05 ,b=10000)
```

ตัวอย่างฟังก์ชันที่ใช้ในข้อมูลจริง

```
########## Part 1 ###########

fix(sd)


##### Part 2:  run in R Editor #####

function (alpha)

{

x=c(39.68,39.08,38.48,39.48,39.88,39.08,38.08,37.48,36.88,36.28,36.28,36.28,35.48,35.48,35.88,35.88,35.28,34.78,34.28,33.78,33.28,32.88,32.38,31.88,31.48,31.48,30.88,30.38,30.68,30.28,29.68,30.18)

n = length(x)

t = qt(1-alpha/2,n-1)

x1 = qchisq(alpha/2,n-1)

x2 = qchisq(1-alpha/2,n-1)

n1 = gamma((n-1)/2)

n2 = gamma(n/2)

## Calculate mean

xbar = mean(x)

## Calculate point estimators

s = sd(x)

s.mle=((sqrt(n-1))/sqrt(n))*sd(x)

s.hat=sqrt((n-1)/2)*(n1/n2)* sd(x)
```

```
## Calculate variance of estimators

var.s = (1-2/(n-1)*(n2/n1)^2)*s^2

var.s.mle = ((1-(1/n))-(2/n)*(n2/n1)^2)*s.mle^2

var.s.hat =(((n-1)/2)*(n1/n2)^2-1)* s.hat^2

## Calculate CV

CV.s = (s/mean(x))*100

CV.s.mle = (s.mle/mean(x))*100

CV.s.hat = (s.hat/mean(x))*100

## Calculate lower and upper limit of CI and Length

lower.1 = s - t*sqrt(var.s)

upper.1 = s + t*sqrt(var.s)

length.1 = upper.1-lower.1

lower.2 = s.mle - t*sqrt(var.s.mle)

upper.2 = s.mle + t*sqrt(var.s.mle)

length.2 = upper.2-lower.2

lower.3 = s.hat - t*sqrt(var.s.hat)

upper.3 = s.hat + t*sqrt(var.s.hat)

length.3 = upper.3-lower.3

lower.4 = s*sqrt(n-1)/sqrt(x2)

upper.4 = s*sqrt(n-1)/sqrt(x1)

length.4 = upper.4-lower.4


cat(n,'\t', xbar ,'\t', alpha , '\n',
```

```
s,'\t', s.mle ,'\t', s.hat ,'\n' ,

CV.s,'\t', CV.s.mle ,'\t', CV.s.hat ,'\n' ,

var.s ,'\t',var.s.mle,  '\t', var.s.hat,'\n',

lower.1, '\t', upper.1, '\t', length.1,'\n',

lower.2, '\t', upper.2, '\t', length.2,'\n',

lower.3, '\t', upper.3, '\t', length.3,'\n',

lower.4, '\t', upper.4, '\t', length.4,'\n')

}
```

```
##### Part 3:  run in R console  #####

   sd(0.05)
```