การศึกษาเปรียบเทียบการประมาณพารามิเตอร์โดยตัวประมาณแบบกราฟด้วยข้อมูลบางส่วนจากข้อมูลที่ถูกตัดปลายทางขวา

นางสาวขวัญรัตน์ ตั้งพิษฐานสกุล

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญามหาบัณฑิต
สาขาวิชาสถิติ ภาควิชาสถิติ
คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2554

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

บทคัดย่อและแฟ้มข้อมูลฉบับเต็มของวิทยานิพนธ์ตั้งแต่ปีการศึกษา 2554 ที่ให้บริการในคลังปัญญาจุฬาฯ (CUIR)
เป็นแฟ้มข้อมูลของนิสิตเจ้าของวิทยานิพนธ์ที่ส่งผ่านทางบัณฑิตวิทยาลัย
The abstract and full text of theses from the academic year 2011 in Chulalongkorn University Intellectual Repository(CUIR)
are the thesis authors' files submitted through the Graduate School.

A COMPARATIVE STUDY ON PARAMETER ESTIMATION BY GRAPHICAL ESTIMATOR WITH PARTIAL DATA FROM RIGHT-CENSORED DATA

Miss Kwanrat Tangpittansakul

A Thesis Submitted in Partial Fulfillment of the Requirements

for the Degree of Master of Science Program in Statistics

Department of Statistics

Faculty of Commerce and Accountancy

Chulalongkorn University

Academic Year 2011

Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์	การศึกษาเปรียบเทียบการประมาณพารามิเตอร์โดยตัวประมาณแบบกราฟด้วยข้อมูลบางส่วนจากข้อมูลที่ถูกตัดปลายทางขวา
โดย	นางสาวขวัญรัตน์ ตั้งพิษฐานสกุล
สาขาวิชา	สถิติ
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก	อาจารย์ ดร.อนุภาพ สมบูรณ์สวัสดี

คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญามหาบัณฑิต

..... คณบดีคณะพาณิชยศาสตร์และการบัญชี
(รองศาสตราจารย์ ดร.พสุ เดชะรินทร์)

คณะกรรมการสอบวิทยานิพนธ์

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

..... อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก
(อาจารย์ ดร.อนุภาพ สมบูรณ์สวัสดี)

..... กรรมการ
(อาจารย์ ดร.อนันตณัฐ กันต์ธัญญรัตน์)

..... กรรมการภายนอกมหาวิทยาลัย
(อาจารย์ ดร.อัครินทร์ ไพบูลย์พานิช)

ขวัญรัตน์ ตั้งพิษฐานสกุล : การศึกษาเปรียบเทียบการประมาณพารามิเตอร์โดยตัวประมาณแบบกราฟด้วยข้อมูลบางส่วนจากข้อมูลที่ถูกตัดปลายทางขวา. (A COMPARATIVE STUDY ON PARAMETER ESTIMATION BY GRAPHICAL ESTIMATOR WITH PARTIAL DATA FROM RIGHT-CENSORED DATA) อ.ที่ปรึกษาวิทยานิพนธ์หลัก: อ.ดร.อนุภาพ สมบูรณ์สวัสดี, 103 หน้า.

งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาเปรียบเทียบวิธีการประมาณพารามิเตอร์ของการแจกแจงของข้อมูลที่ถูกตัดปลายทางขวาประเภทที่ 2 ด้วยวิธีภาวะน่าจะเป็นสูงสุด (Maximum Likelihood Estimation Method; MLE Method) วิธีการประมาณแบบกราฟ (Graphical Estimation Method; GE Method) และวิธีการประมาณแบบกราฟด้วยข้อมูลบางส่วน (Graphical Estimation with Partial Data Method; GEPD Method) โดยการแจกแจงที่สนใจศึกษา ได้แก่ การแจกแจงปกติ (Normal Distribution; NOR) การแจกแจงโลจิสติค(Logistic Distribution; LOG) การแจกแจงค่าต่ำสุดขีด (Smallest Extreme Value Distribution; SEV) และการแจกแจงค่าสูงสุดขีด (Largest Extreme Value Distribution; LEV) การเปรียบเทียบกระทำภายใต้สถานการณ์ของขนาดตัวอย่าง (Sample Size; n) เท่ากับ 20, 40, 80 และ 120 ด้วยสัดส่วนของข้อมูลที่ถูกตัดปลาย (Censoring Proportion; p) เป็น 10%, 20% และ 30% และสร้างสถานการณ์ต่างๆ ในการทดลองโดยใช้โปรแกรม R เวอร์ชัน 2.9.2 ทำการทดลองซ้ำๆ กัน 5,000 ครั้ง สำหรับแต่ละสถานการณ์

ผลการวิจัยสรุปได้ดังนี้

1. การประมาณพารามิเตอร์ด้วยวิธีภาวะน่าจะเป็นสูงสุด (MLE) เป็นวิธีที่มีประสิทธิภาพมากที่สุดเมื่อเทียบกับวิธีการประมาณแบบกราฟ (GE) และวิธีการประมาณแบบกราฟด้วยข้อมูลบางส่วน (GEPD) ในทุกสถานการณ์ที่ทำการศึกษา โดยประสิทธิภาพจะมากยิ่งขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น

2. การประมาณพารามิเตอร์ด้วยวิธี GEPD ผลปรากฏว่า ในกรณีส่วนใหญ่มีประสิทธิภาพมากกว่าการประมาณด้วยวิธี GE และเมื่อพิจารณาเปรียบเทียบวิธี GEPD กับวิธี GE ภายใต้การแจกแจงต่างๆ พบว่า

2.1 การประมาณค่าพารามิเตอร์ μ (location parameter) ด้วยวิธี GEPD จะมีประสิทธิภาพมากกว่าวิธี GE ในทุกๆ การแจกแจงที่ทำการศึกษา โดยการประมาณด้วยวิธี GEPD แบบ K Cluster-Mean นั้นเหมาะสำหรับข้อมูลที่ไม่ถูกตัดปลาย (p =0) มากกว่าข้อมูลที่ถูกตัดปลายทางขวา ($p \neq 0$) แต่ถ้าข้อมูลถูกตัดปลายทางขวาแล้ว ควรประมาณด้วยวิธี GEPD แบบ Trimmed q% & K Cluster-Mean

2.2 การประมาณค่าพารามิเตอร์ σ (scale parameter) ด้วยวิธี GEPD จะมีประสิทธิภาพมากกว่าวิธี GE เมื่อข้อมูลมีการแจกแจงแบบ SEV หรือ LOG โดยเมื่อ p =0 ควรประมาณแบบ K Cluster-Mean และเมื่อ p =0.1, 0.2, 0.3 ควรประมาณแบบ Trimmed q% & K Cluster-Mean อย่างไรก็ตามถ้าหากข้อมูลมีการแจกแจงแบบ NOR หรือ LEV แล้ว ควรจะประมาณด้วยวิธี GE

ภาควิชา สถิติ ลายมือชื่อนิสิต .....

สาขาวิชา สถิติ ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์หลัก .....

ปีการศึกษา 2554

5281769926 : MAJOR STATISTICS

KEYWORDS : Location-Scale Distribution / Maximum Likelihood / Graphical Estimator / Right-Censored Data

KWANRAT TANGPITTANSAKUL: A COMPARATIVE STUDY ON PARAMETER ESTIMATION BY GRAPHICAL ESTIMATOR WITH PARTIAL DATA FROM RIGHT-CENSORED DATA. ADVISOR: ANUPAP SOMBOONSAVATDEE, Ph.D. 103 pp.

The objective of this study is to compare parameter estimations of type-II right-censored data by Maximum Likelihood Estimation (MLE), Graphical Estimation (GE) and Graphical Estimation with Partial Data (GEPD). The distributions of data under considerations in this study are Normal (NOR), Logistic (LOG), Smallest Extreme Value (SEV) and Largest Extreme Value (LEV) distributions. The comparison was done under conditions of sample sizes n = 20, 40, 80 and 120 with the censoring proportion p = 10%, 20% and 30%. The data for this experiment was generated and analyzed data by R program version 2.9.2. The experiment was repeated 5,000 times under each condition.

Results of the study are as follows:-

1. Under all scenarios in this study, the MLE was the most efficient when compared with the GE and the GEPD. And its efficiency increased as sample size increased.

2. In most scenarios, parameter estimations by the GEPD were more efficient than by the GE. When the GEPD and the GE were compared under distribution of data, it found that

2.1 the parameter estimations for μ (location parameter) by the GEPD were more efficient than the GE for each distributions in this study. For uncensored data (p =0), the GEPD —— type K Cluster-Mean was more preferred. On the other hand, the GEPD —— type Trimmed q% & K Cluster-Mean was more preferred for right-censored data ($p \neq 0$), and

2.2 the parameter estimations for σ (scale parameter) by the GEPD were more efficient than the GE when data have SEV or LOG distributions. For p =0, the GEPD —— type K Cluster-Mean was more preferred. On the other hand, the GEPD —— type Trimmed q% & K Cluster-Mean was more preferred for p = 0.1, 0.2, 0.3. However, the GE was more preferred when distributions of data were NOR or LEV.

Department : Statistics Student's Signature .....

Field of Study : Statistics Advisor's Signature .....

Academic Year : 2011

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สามารถสำเร็จลุล่วงไปได้ด้วยความช่วยเหลืออย่างดียิ่งของ อาจารย์ ดร.อนุภาพ สมบูรณ์สวัสดี อาจารย์ที่ปรึกษาวิทยานิพนธ์ ที่กรุณาให้คำแนะนำสั่งสอน และให้ข้อคิดเห็นต่างๆ ตลอดจนช่วยเหลือแก้ไขข้อบกพร่องต่างๆ จนกระทั่งวิทยานิพนธ์เสร็จสมบูรณ์ ผู้วิจัยใคร่ขอกราบขอบพระคุณเป็นอย่างสูง

ผู้วิจัยขอกราบขอบพระคุณคณะกรรมการสอบวิทยานิพนธ์ซึ่งประกอบไปด้วย รองศาสตราจารย์ ดร.สุพล ดุรงค์วัฒนา ประธานกรรมการ อาจารย์ ดร.อนันตณัฐ กันต์ธัญญรัตน์ และอาจารย์ ดร.อัครินทร์ ไพบูลย์พานิช ผู้เป็นกรรมการ ที่ช่วยตรวจสอบและแก้ไขวิทยานิพนธ์ฉบับนี้ให้สมบูรณ์ยิ่งขึ้น ขอกราบขอบพระคุณคณาจารย์ทุกท่านที่ได้ให้ความรู้แก่ผู้วิจัยจนกระทั่งสำเร็จการศึกษา

ผู้วิจัยขอกราบขอบพระคุณนายจิณณภัทร พิบูลวิทิตธำรง ผู้อำนวยการโรงเรียนมัธยมวัดธาตุทอง ที่ให้โอกาสในการลาศึกษาต่อเพื่อให้ผู้วิจัยได้พัฒนาความรู้ความสามารถ และขอขอบคุณครูอาจารย์โรงเรียนมัธยมวัดธาตุทองทุกท่านสำหรับความรักและความห่วงใย

ท้ายนี้ ผู้วิจัยใคร่ขอน้อมรำลึกถึงพระคุณครูอาจารย์ทุกท่านที่ได้ให้ความรู้แก่ผู้วิจัย และกราบขอบพระคุณ บิดา มารดา ตลอดจนทุกคนในครอบครัวซึ่งสนับสนุนและให้กำลังใจเสมอมา จนสำเร็จการศึกษา และขอขอบคุณนายกรกฎ วัฒนวีร์ เพื่อนที่ให้คำปรึกษาทางด้านโปรแกรม R, นางสาวสาวิตรี บุญพัชรนนท์ เพื่อนที่ให้ความช่วยเหลือด้านการเรียน รวมถึงเพื่อนๆ ทุกคนสำหรับมิตรภาพและกำลังใจที่มีให้ตลอดมา

สารบัญ

หน้า

บทคัดย่อภาษาไทย..... ง

บทคัดย่อภาษาอังกฤษ..... จ

กิตติกรรมประกาศ..... ฉ

สารบัญ..... ช

สารบัญตาราง..... ฌ

สารบัญภาพ..... ด

บทที่ 1 บทนำ..... 1

1.1 ความเป็นมาและความสำคัญของปัญหา..... 1

1.2 วัตถุประสงค์ของการวิจัย..... 2

1.3 คำจำกัดความที่ใช้ในการวิจัย..... 2

1.4 ขอบเขตของการวิจัย..... 3

1.5 วิธีดำเนินการวิจัย..... 6

1.6 ลำดับขั้นตอนในการเสนอผลการวิจัย..... 6

1.7 ประโยชน์ที่คาดว่าจะได้รับ..... 7

บทที่ 2 ทฤษฎีและตัวสถิติที่เกี่ยวข้อง..... 8

2.1 ทฤษฎีพื้นฐาน..... 8

2.2 การแจกแจงที่ใช้ในการวิจัย..... 10

2.3 การประมาณค่าพารามิเตอร์..... 14

2.4 เกณฑ์ที่ใช้ในการพิจารณาประสิทธิภาพของการประมาณ..... 16

บทที่ 3 วิธีดำเนินการวิจัย..... 18

3.1 แผนการดำเนินการวิจัย..... 18

3.2 ขั้นตอนในการดำเนินการวิจัย..... 23

3.3 ขั้นตอนการทำงานของโปรแกรม..... 25

หน้า

บทที่ 4 ผลการวิเคราะห์ข้อมูล..... 27
4.1 สรุปผลการศึกษาเปรียบเทียบการประมาณค่าพารามิเตอร์ทั้ง 3 วิธี..... 28
4.1.1 ผลการศึกษาเมื่อข้อมูลแจกแจงแบบค่าต่ำสุดขีด..... 31
4.1.2 ผลการศึกษาเมื่อข้อมูลแจกแจงแบบค่าสูงสุดขีด..... 46
4.1.3 ผลการศึกษาเมื่อข้อมูลแจกแจงแบบโลจิสติค..... 53
4.1.4 ผลการศึกษาเมื่อข้อมูลแจกแจงแบบปกติ..... 62

บทที่ 5 สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ..... 70
5.1 สรุปผลการวิจัย..... 70
5.2 ข้อเสนอแนะ..... 78
5.2.1 ด้านการนำไปใช้ประโยชน์..... 78
5.2.1 ด้านการศึกษาวิจัย..... 79

รายการอ้างอิง..... 80
ภาคผนวก..... 81
ประวัติผู้เขียนวิทยานิพนธ์..... 103

สารบัญตาราง

ตารางที่		หน้า
3.1	แสดงเกณฑ์ในการ Trim ข้อมูล และแบ่ง Cluster สำหรับการประมาณพารามิเตอร์ μ และ σ ด้วยวิธีการประมาณแบบใช้กราฟ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา.....	21
3.1	(ต่อ)แสดงเกณฑ์ในการ Trim ข้อมูล และแบ่ง Cluster สำหรับการประมาณพารามิเตอร์ μ และ σ ด้วยวิธีการประมาณแบบใช้กราฟ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา.....	22
4.1	แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงแบบค่าต่ำสุดขีด ที่พารามิเตอร์ $\mu = 0$ และ $\sigma = 1$ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา.....	31
4.2	แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงแบบค่าต่ำสุดขีด ที่พารามิเตอร์ $\mu = 0$ และ $\sigma = 1$ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา.....	33
4.3	แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงแบบค่าต่ำสุดขีด ที่พารามิเตอร์ $\mu = 0$ และ $\sigma = 1$ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— 4 Cluster-Mean และ 8 Cluster-Mean ตามลำดับ.....	34
4.4	แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงแบบค่าต่ำสุดขีด ที่พารามิเตอร์ $\mu = 0$ และ $\sigma = 1$ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— 4 Cluster-Mean และ 8 Cluster-Mean ตามลำดับ.....	36
4.5	แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงแบบค่าต่ำสุดขีด ที่พารามิเตอร์ $\mu = 0$ และ $\sigma = 1$ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— Trimmed 5% และ Trimmed 10% ตามลำดับ.....	37

ตารางที่ หน้า

4.6 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงแบบค่าต่ำสุดขีด ที่พารามิเตอร์ μ = 0และ σ = 1 โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— Trimmed 5% และ Trimmed 10% ตามลำดับ..... 38

4.7 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงแบบค่าต่ำสุดขีด ที่พารามิเตอร์ μ = 0และ σ = 1 โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— Trimmed 5% & 4 Cluster-MeanและTrimmed 5% & 8 Cluster-Mean ตามลำดับ…….. 40

4.8 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงแบบค่าต่ำสุดขีด ที่พารามิเตอร์ μ = 0และ σ = 1 โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— Trimmed 5% & 4 Cluster-MeanและTrimmed 5% & 8 Cluster-Mean ตามลำดับ…….. 41

4.9 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงแบบค่าต่ำสุดขีด ที่พารามิเตอร์ μ = 0และ σ = 1 โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— Trimmed 10% & 4 Cluster-Mean และ Trimmed 10% & 8 Cluster-Mean ตามลำดับ……. 43

4.10 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงแบบค่าต่ำสุดขีด ที่พารามิเตอร์ μ = 0และ σ = 1 โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— Trimmed 10% & 4 Cluster-Mean และ Trimmed 10% & 8 Cluster-Mean ตามลำดับ……. 44

4.11 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงแบบค่าสูงสุดขีด ที่พารามิเตอร์ μ = 0 และ σ = 1 โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา…….. 46

ตารางที่ หน้า

4.12 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงแบบค่าสูงสุดขีด ที่พารามิเตอร์ $\mu = 0$ และ $\sigma = 1$ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา..... 47

4.13 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงแบบค่าสูงสุดขีด ที่พารามิเตอร์ $\mu = 0$ และ $\sigma = 1$ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— 4 Cluster-Mean และ 8 Cluster-Mean ตามลำดับ.. 48

4.14 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงแบบค่าสูงสุดขีด ที่พารามิเตอร์ $\mu = 0$ และ $\sigma = 1$ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— 4 Cluster-Mean และ 8 Cluster-Mean ตามลำดับ... 48

4.15 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงแบบค่าสูงสุดขีด ที่พารามิเตอร์ $\mu = 0$ และ $\sigma = 1$ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— Trimmed 5% และ Trimmed 10% ตามลำดับ..... 49

4.16 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงแบบค่าสูงสุดขีด ที่พารามิเตอร์ $\mu = 0$ และ $\sigma = 1$ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— Trimmed 5% และ Trimmed 10% ตามลำดับ..... 50

4.17 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงแบบค่าสูงสุดขีด ที่พารามิเตอร์ $\mu = 0$ และ $\sigma = 1$ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— Trimmed 5% & 4 Cluster-Mean และTrimmed 5% & 8 Cluster-Mean ตามลำดับ..... 50

ตารางที่ หน้า

4.18 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงแบบค่าสูงสุดขีด ที่พารามิเตอร์ $\mu = 0$ และ $\sigma = 1$ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— Trimmed 5% & 4 Cluster-Mean และ Trimmed 5% & 8 Cluster-Mean ตามลำดับ..... 51

4.19 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงแบบค่าสูงสุดขีด ที่พารามิเตอร์ $\mu = 0$ และ $\sigma = 1$ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— Trimmed 10% & 4 Cluster-Mean และTrimmed 10% & 8 Cluster-Mean ตามลำดับ..... 52

4.20 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงแบบค่าสูงสุดขีด ที่พารามิเตอร์ $\mu = 0$ และ $\sigma = 1$ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— Trimmed 10% & 4 Cluster-Mean และTrimmed 10% & 8 Cluster-Mean ตามลำดับ..... 52

4.21 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงแบบโลจิสติค ที่พารามิเตอร์ $\mu = 0$ และ $\sigma = 1$ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา..... 53

4.22 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงแบบโลจิสติค ที่พารามิเตอร์ $\mu = 0$ และ $\sigma = 1$ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา..... 54

4.23 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงแบบโลจิสติค ที่พารามิเตอร์ $\mu = 0$ และ $\sigma = 1$ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— 4 Cluster-Mean และ 8 Cluster-Mean ตามลำดับ.. 55

ตารางที่ หน้า

4.24 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงแบบโลจิสติค ที่พารามิเตอร์ $\mu = 0$ และ $\sigma = 1$ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— 4 Cluster-Mean และ 8 Cluster-Mean ตามลำดับ.. 56

4.25 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงแบบโลจิสติค ที่พารามิเตอร์ $\mu = 0$ และ $\sigma = 1$ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— Trimmed 5% และ Trimmed 10% ตามลำดับ..... 57

4.26 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงแบบโลจิสติค ที่พารามิเตอร์ $\mu = 0$ และ $\sigma = 1$ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— Trimmed 5% และ Trimmed 10% ตามลำดับ..... 57

4.27 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงแบบโลจิสติค ที่พารามิเตอร์ $\mu = 0$ และ $\sigma = 1$ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— Trimmed 5% & 4 Cluster-Mean และTrimmed 5% & 8 Cluster-Mean ตามลำดับ..... 58

4.28 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงแบบโลจิสติค ที่พารามิเตอร์ $\mu = 0$ และ $\sigma = 1$ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— Trimmed 5% & 4 Cluster-Mean และTrimmed 5% & 8 Cluster-Mean ตามลำดับ..... 59

4.29 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงแบบโลจิสติค ที่พารามิเตอร์ $\mu = 0$ และ $\sigma = 1$ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— Trimmed 10% & 4 Cluster-Mean และTrimmed 10% & 8 Cluster-Mean ตามลำดับ..... 60

ตารางที่ หน้า

4.30 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงแบบโลจิสติค ที่พารามิเตอร์ μ = 0 และ σ = 1 โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— Trimmed 10% & 4 Cluster-Mean และTrimmed 10% & 8 Cluster-Mean ตามลำดับ..... 60

4.31 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงแบบปกติ ที่พารามิเตอร์ μ = 0 และ σ = 1 โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา..... 62

4.32 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงแบบปกติ ที่พารามิเตอร์ μ = 0 และ σ = 1 โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา..... 63

4.33 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงแบบปกติ ที่พารามิเตอร์ μ = 0 และ σ = 1 โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— 4 Cluster-Mean และ 8 Cluster-Mean ตามลำดับ.. 64

4.34 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงแบบปกติ ที่พารามิเตอร์ μ = 0 และ σ = 1 โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— 4 Cluster-Mean และ 8 Cluster-Mean ตามลำดับ.. 64

4.35 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงแบบปกติ ที่พารามิเตอร์ μ = 0 และ σ = 1 โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— Trimmed 5% และ Trimmed 10% ตามลำดับ..... 65

ตารางที่		หน้า
4.36	แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงแบบปกติ ที่พารามิเตอร์ μ = 0 และ σ = 1 โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— Trimmed 5% และ Trimmed 10% ตามลำดับ…….	66
4.37	แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงแบบปกติ ที่พารามิเตอร์ μ = 0 และ σ = 1 โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— Trimmed 5% & 4 Cluster-MeanและTrimmed 5% & 8 Cluster-Mean ตามลำดับ……..	67
4.38	แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงแบบปกติ ที่พารามิเตอร์ μ = 0 และ σ = 1 โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— Trimmed 5% & 4 Cluster-MeanและTrimmed 5% & 8 Cluster-Mean ตามลำดับ.....	67
4.39	แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงแบบปกติ ที่พารามิเตอร์ μ = 0 และ σ = 1 โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— Trimmed 10% & 4 Cluster-Mean และTrimmed 10% & 8 Cluster-Mean ตามลำดับ……..……..…….	68
4.40	แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงแบบปกติ ที่พารามิเตอร์ μ = 0 และ σ = 1 โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— Trimmed 10% & 4 Cluster-Mean และTrimmed 10% & 8 Cluster-Mean ตามลำดับ……..……	69
5.1	แสดงการเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์ μ ระหว่างวิธี GE และวิธี GEPD ที่มีประสิทธิภาพมากที่สุด สำหรับการแจกแจงแบบต่างๆ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา……..	72

ตารางที่ หน้า

5.2 แสดงการเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์ σ ระหว่างวิธี GE และวิธี GEPD ที่มีประสิทธิภาพมากที่สุด สำหรับการแจกแจงแบบต่างๆ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา..... 73

5.3 แสดงการพิจารณาเลือกใช้วิธีการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงค่าต่ำสุดขีด โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา 74

5.4 แสดงการพิจารณาเลือกใช้วิธีการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงค่าต่ำสุดขีด โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา 74

5.5 แสดงการพิจารณาเลือกใช้วิธีการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงค่าสูงสุดขีด โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา 75

5.6 แสดงการพิจารณาเลือกใช้วิธีการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงค่าสูงสุดขีด โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา 75

5.7 แสดงการพิจารณาเลือกใช้วิธีการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงโลจิสติค โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา 76

5.8 แสดงการพิจารณาเลือกใช้วิธีการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงโลจิสติค โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา 76

5.9 แสดงการพิจารณาเลือกใช้วิธีการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงปกติ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา..... 77

5.10 แสดงการพิจารณาเลือกใช้วิธีการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงปกติ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา..... 77

สารบัญภาพ

ภาพที่		หน้า
1.1	แสดงการประมาณพารามิเตอร์โดยใช้กราฟแบบ All points.....	4
1.2	แสดงการประมาณพารามิเตอร์โดยใช้กราฟด้วยวิธี Trimmed q%.....	4
1.3	แสดงการประมาณพารามิเตอร์โดยใช้กราฟด้วยวิธี K Cluster-Mean.....	5
1.4	แสดงการประมาณพารามิเตอร์โดยใช้กราฟด้วยวิธี Trimmed q% & K Cluster-Mean	5
2.1	แสดงฟังก์ชันความหนาแน่นน่าจะเป็นของการแจกแจง NOR, LOGIS, SEV และ LEV ที่พารามิเตอร์ μ = 0 และ σ = 1.....	13
2.2	แสดงการประมาณด้วยวิธีแบบใช้กราฟ.....	15
3.1	แสดงแผนผังขั้นตอนการวิจัย.....	26
4.1	แสดงการเปรียบเทียบค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงแบบค่าต่ำสุดขีดที่พารามิเตอร์ μ = 0 และ σ = 1 ด้วยวิธี GE กับวิธี MLE.....	32
4.2	แสดงการเปรียบเทียบค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงแบบค่าต่ำสุดขีดที่พารามิเตอร์ μ = 0 และ σ = 1 ด้วยวิธี GE กับวิธี MLE.....	33
4.3	แสดงการเปรียบเทียบค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงแบบค่าต่ำสุดขีดที่พารามิเตอร์ μ = 0 และ σ = 1 ด้วยวิธี GE กับวิธี GEPD ── 4 Cluster-Mean และ 8 Cluster-Mean ตามลำดับ.....	35
4.4	แสดงการเปรียบเทียบค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงแบบค่าต่ำสุดขีดที่พารามิเตอร์ μ = 0 และ σ = 1 ด้วยวิธี GE กับวิธี GEPD ── 4 Cluster-Mean และ 8 Cluster-Mean ตามลำดับ.....	36
4.5	แสดงการเปรียบเทียบค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงแบบค่าต่ำสุดขีดที่พารามิเตอร์ μ = 0 และ σ = 1 ด้วยวิธี GE กับวิธี GEPD ── Trimmed 5% และ Trimmed 10% ตามลำดับ.....	37

ภาพที่ หน้า

4.6 แสดงการเปรียบเทียบค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงแบบค่าต่ำสุดขีดที่พารามิเตอร์ μ = 0 และ σ = 1 ด้วยวิธี GE กับวิธี GEPD ── Trimmed 5% และ Trimmed 10% ตามลำดับ…….. 39

4.7 แสดงการเปรียบเทียบค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงแบบค่าต่ำสุดขีดที่พารามิเตอร์ μ = 0 และ σ = 1 ด้วยวิธี GE กับวิธี GEPD ── Trimmed 5% & 4 Cluster-Mean และ Trimmed 5% & 8 Cluster-Mean ตามลำดับ……..... 40

4.8 แสดงการเปรียบเทียบค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงแบบค่าต่ำสุดขีดที่พารามิเตอร์ μ = 0 และ σ = 1 ด้วยวิธี GE กับวิธี GEPD ── Trimmed 5% & 4 Cluster-Mean และ Trimmed 5% & 8 Cluster-Mean ตามลำดับ…..…... 42

4.9 แสดงการเปรียบเทียบค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงแบบค่าต่ำสุดขีดที่พารามิเตอร์ μ = 0 และ σ = 1 ด้วยวิธี GE กับวิธี GEPD ── Trimmed 10% & 4 Cluster-Mean และ Trimmed 10% & 8 Cluster-Mean ตามลำดับ…….. 43

4.10 แสดงการเปรียบเทียบค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงแบบค่าต่ำสุดขีดที่พารามิเตอร์ μ = 0 และ σ = 1 ด้วยวิธี GE กับวิธี GEPD ── Trimmed 10% & 4 Cluster-Mean และ Trimmed 10% & 8 Cluster-Mean ตามลำดับ…….. 45

บทที่ 1
บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

การอนุมานเชิงสถิติ (Statistical Inference) เป็นระเบียบวิธีทางสถิติที่ใช้ในการศึกษาวิจัยดำเนินงานต่างๆ เพื่อหาข้อสรุปเกี่ยวกับประชากร ประกอบด้วย การประมาณค่าพารามิเตอร์ (Parameter Estimation) และการทดสอบสมมติฐาน (Hypothesis Testing)

ในการศึกษาค่าที่บอกถึงลักษณะประชากร ซึ่งเรียกว่าค่าพารามิเตอร์นั้น จะทราบได้ก็ต่อเมื่อจะต้องทำการแจงนับจากทุกๆ หน่วยในประชากร สำหรับการสำรวจหรือการทดลองต่างๆ ที่เป็นการศึกษาเกี่ยวกับระยะเวลาการรอคอยจนกระทั่งเหตุการณ์ที่สนใจหรือความล้มเหลวจะเกิดขึ้น บ่อยครั้งมักจะเสียทั้งเวลาและค่าใช้จ่ายในการรอคอยจนกระทั่งเกิดความล้มเหลวครบทุกค่าสังเกต เนื่องจากผู้ทดลองไม่สามารถคาดเดาได้ว่าระยะเวลาการทดลองจะสิ้นสุดเมื่อใด เพื่อแก้ปัญหาเหล่านี้ผู้ที่ทำการสำรวจหรือทดลองอาจจะออกแบบการทดลองโดยกำหนดเวลาที่สิ้นสุดการเก็บข้อมูลไว้ล่วงหน้า (Fixed Censoring Time) หรือกำหนดจำนวนค่าสังเกตที่จะเกิดความล้มเหลวไว้ล่วงหน้า (Fixed Number of Uncensored Failure) โดยไม่ต้องทำการทดลองจนครบตามขนาดตัวอย่าง ซึ่งมีผลทำให้ข้อมูลที่เก็บรวบรวมมาได้นั้นเป็นข้อมูลที่ไม่สมบูรณ์ (Incomplete Data) โดยจะพบว่าข้อมูลที่ได้เป็นข้อมูลที่ถูกตัดปลายทางขวา (Right-Censored Data) ข้อมูลลักษณะนี้จำเป็นสำหรับงานวิจัยด้านต่างๆ เช่น การศึกษาทางการแพทย์ การศึกษาทางด้านประชากรศาสตร์ การศึกษาทางด้านอุตสาหกรรม และทางด้านธุรกิจประกันภัย เป็นต้น

เนื่องจากในงานวิจัยส่วนใหญ่ มักพบว่าการประมาณค่าพารามิเตอร์ด้วยวิธีภาวะน่าจะเป็นสูงสุดเป็นวิธีที่มีประสิทธิภาพดีกว่าวิธีการประมาณอื่นๆ แต่ในทางปฏิบัติการวิเคราะห์ข้อมูลด้วยวิธีภาวะน่าจะเป็นสูงสุดนั้นทำได้ยาก เพราะวิธีภาวะน่าจะเป็นสูงสุดไม่มีรูปแบบสมการที่แน่นอน ทำให้คำนวณได้ยากและใช้เวลานาน โดย ศศิประภา โมรากุล (2553) ได้ศึกษาเปรียบเทียบการประมาณพารามิเตอร์จากข้อมูลที่ถูกเซ็นเซอร์แบบช่วง ด้วยวิธีภาวะน่าจะเป็นสูงสุดและวิธีการประมาณแบบใช้กราฟที่มีความสะดวกและง่ายต่อการคำนวณมากกว่าวิธีภาวะน่าจะเป็นสูงสุด และทำการเปรียบเทียบว่าวิธีการประมาณแบบใดมีประสิทธิภาพมากกว่าและควรจะใช้การประมาณแบบใดในแต่ละสถานการณ์ นอกจากนี้ยังศึกษาวิธีการประมาณแบบใช้กราฟที่ปรับค่าเอนเอียงว่าจะมีประสิทธิภาพดีกว่าวิธีแบบใช้กราฟหรือไม่ เพื่อนำมาเป็นแนวทางในการปฏิบัติจริง

จากปัญหาดังกล่าว ผู้วิจัยสนใจที่จะทำการศึกษาเปรียบเทียบการประมาณพารามิเตอร์จากข้อมูลที่ถูกตัดปลายทางขวาประเภทที่ 2 ซึ่งจะกำหนดจำนวนค่าสังเกตที่จะเกิดความล้มเหลวไว้ล่วงหน้า ด้วยวิธีภาวะน่าจะเป็นสูงสุด วิธีการประมาณแบบกราฟและวิธีการประมาณแบบกราฟด้วยข้อมูลบางส่วน

1.2 วัตถุประสงค์ของการวิจัย

1.2.1 เพื่อเปรียบเทียบการประมาณค่าพารามิเตอร์ของการแจกแจงของข้อมูลที่ถูกตัดปลายทางขวาประเภทที่ 2 ด้วยวิธีภาวะน่าจะเป็นสูงสุด (Maximum Likelihood Estimation Method; MLE Method) และวิธีการประมาณแบบกราฟ (Graphical Estimation Method; GE Method) โดยการแจกแจงที่สนใจศึกษา ได้แก่ การแจกแจงปกติ (Normal Distribution) การแจกแจงโลจิสติค (Logistic Distribution) การแจกแจงค่าต่ำสุดขีด (Smallest Extreme Value Distribution) และการแจกแจงค่าสูงสุดขีด (Largest Extreme Value Distribution)

1.2.2 เพื่อเปรียบเทียบประสิทธิภาพของวิธีการประมาณแบบกราฟด้วยข้อมูลบางส่วน (Graphical Estimation with Partial Data Method; GEPD Method) กับวิธีการประมาณแบบกราฟ (Graphical Estimation Method; GE Method)

1.3 คำจำกัดความที่ใช้ในการวิจัย

1.3.1 ข้อมูลถูกตัดปลายทางขวาประเภทที่ 2 (Type-II Right Censoring หรือ Fixed Number of Uncensored Failure) คือ ข้อมูลที่มีการตัดปลายข้อมูลโดยการกำหนดจำนวนความล้มเหลวไว้ล่วงหน้า ผู้ทดลองจะทำการทดลองจนกว่าจะครบตามจำนวนที่กำหนดแล้วจะหยุดทำการทดลองทันทีโดยไม่ต้องทำการทดลองให้ครบตามขนาดตัวอย่าง

1.3.2 ความเอนเอียง (Bias) คือ ค่าที่ใช้วัดว่าค่าเฉลี่ยของตัวสถิติที่ได้ห่างจากฟังก์ชันพารามิเตอร์ที่สนใจมากน้อยเพียงใด นอกจากนั้นยังบอกทิศทางได้ด้วยว่า ตัวสถิติที่ได้ให้ค่าสูงหรือต่ำกว่าพารามิเตอร์ ; $Bias(\mu) = E(\hat{\mu}) - \mu$ เมื่อ $\hat{\mu}$ เป็นค่าประมาณของ μ และ $Bias(\sigma) = E(\hat{\sigma}) - \sigma$ เมื่อ $\hat{\sigma}$ เป็นค่าประมาณของ σ

1.3.3 ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) คือ ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของค่าประมาณกับค่าจริง ; $MSE(\mu) = E[\hat{\mu} - \mu]^2$ เมื่อ $\hat{\mu}$ เป็นค่าประมาณของ μ และ $MSE(\sigma) = E[\hat{\sigma} - \sigma]^2$ เมื่อ $\hat{\sigma}$ เป็นค่าประมาณของ σ

1.4 ขอบเขตของการวิจัย

1.4.1 การแจกแจงที่นำมาศึกษาเป็นการแจกแจงแบบ location-scale โดยที่ μ เป็น location parameter และ σ เป็น scale parameter ซึ่งทำการศึกษาทั้งหมด 4 การแจกแจง คือ

1.4.1.1 การแจกแจงปกติ (Normal Distribution; NOR) โดยกำหนด μ =0 และσ =1

1.4.1.2 การแจกแจงโลจิสติค (Logistic Distribution; LOGIS) โดยกำหนด μ =0 และσ =1

1.4.1.3 การแจกแจงค่าต่ำสุดขีด (Smallest Extreme Value Distribution; SEV) โดยกำหนด μ =0 และσ =1

1.4.1.4 การแจกแจงค่าสูงสุดขีด (Largest Extreme Value Distribution; LEV) โดยกำหนด μ =0 และσ =1

1.4.2 กำหนดขนาดตัวอย่าง (Sample Size; n) ที่ใช้ในการศึกษาเท่ากับ 20, 40, 80 และ 120

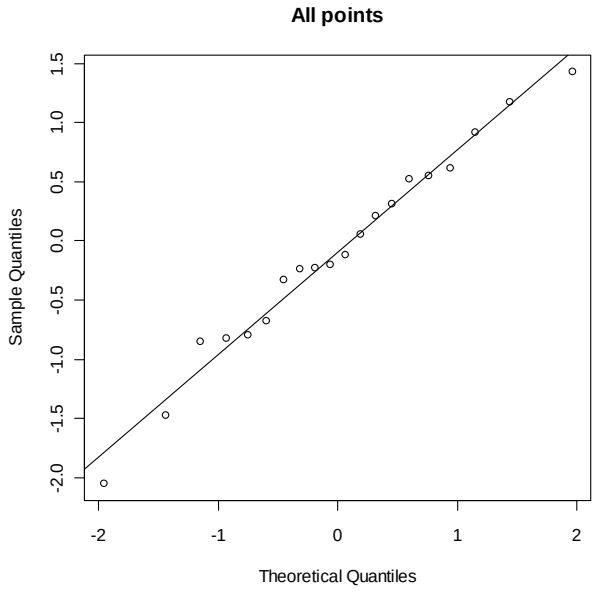
1.4.3 ข้อมูลที่ใช้ในการศึกษาเป็นข้อมูลที่ถูกตัดปลายทางขวาประเภทที่ 2 (Type-II Right Censoring) ซึ่งมีสัดส่วนของการตัดปลาย (Censoring Proportion; p) เท่ากับ 10%, 20% และ 30%

1.4.4 ศึกษาวิธีการประมาณค่าพารามิเตอร์ 3 วิธี คือ

1.4.4.1 การประมาณโดยวิธีภาวะน่าจะเป็นสูงสุด (MLE Method)

1.4.4.2 การประมาณแบบกราฟ (GE Method) เป็นการใช้ข้อมูลจากทุกส่วนหรือทุกจุด ซึ่งเรียกว่าแบบ All points เช่น n = 20

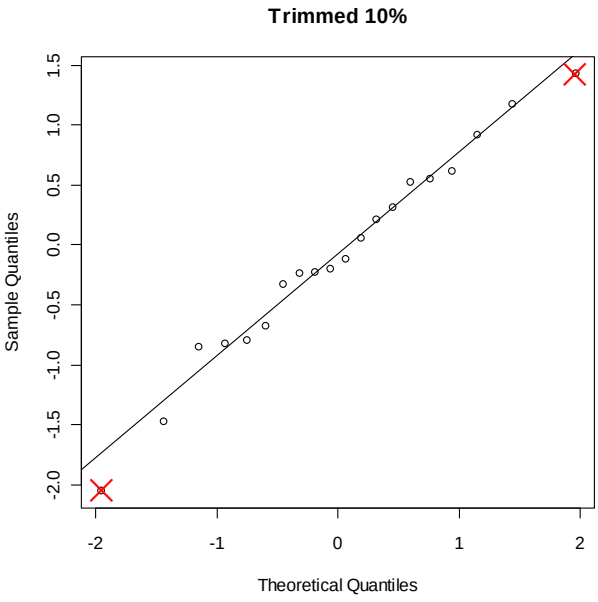
ภาพที่ 1.1 แสดงการประมาณพารามิเตอร์โดยใช้กราฟแบบ All points



1.4.4.3 การประมาณแบบกราฟด้วยข้อมูลบางส่วน (GEPD Method) ประกอบไปด้วย

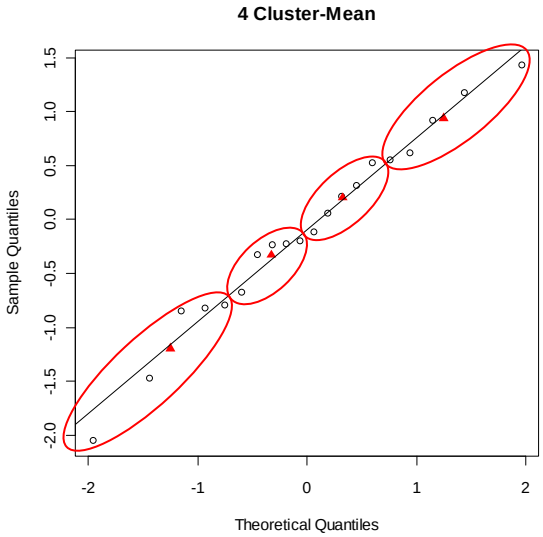
(1) วิธี Trimmed q% เมื่อ q = 5% และ 10% สามารถทำได้โดยการตัดข้อมูลที่มีค่ามากสุดและน้อยสุดออกเป็นจำนวนเท่าๆ กัน เช่น n = 20, q = 10%

ภาพที่ 1.2 แสดงการประมาณพารามิเตอร์โดยใช้กราฟด้วยวิธี Trimmed q%



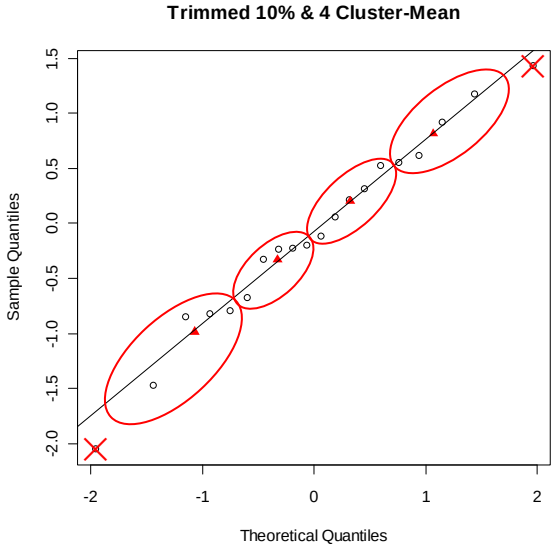
(2) วิธี K Cluster-Mean เมื่อ K = 4 และ 8 สามารถทำได้โดยแบ่งข้อมูลออกเป็นกลุ่มๆ ซึ่งแต่ละกลุ่มจะต้องนำมาหาค่าเฉลี่ย เพื่อให้เหลือข้อมูลเพียงจุดเดียวในแต่ละกลุ่ม เช่น n = 20, K = 4

ภาพที่ 1.3 แสดงการประมาณพารามิเตอร์โดยใช้กราฟด้วยวิธี K Cluster-Mean



(3) วิธี Trimmed q% & K Cluster-Mean เมื่อ q = 5%, 10% และ K = 4, 8 คือ การนำวิธีจากข้อ (2) และข้อ (3) มารวมกัน สามารถทำโดยการตัดข้อมูลที่มีค่ามากสุดและน้อยสุดออกเป็นจำนวนเท่าๆ กันแล้วแบ่งข้อมูลออกเป็นกลุ่มๆ เพื่อหาค่าเฉลี่ยของแต่ละกลุ่ม เช่น n = 20, q = 10%, K = 4

ภาพที่ 1.4 แสดงการประมาณพารามิเตอร์โดยใช้กราฟแบบ Trimmed q% & K Cluster-Mean



1.4.5 ในการศึกษาครั้งนี้ประมวลผลโดยโปรแกรม R เวอร์ชัน 2.9.2 (R Development Core Team, 2009) โดยการจำลองในแต่ละสถานการณ์ของการทดลองจะกระทำซ้ำ 5,000 รอบ

1.5 วิธีดำเนินการวิจัย

1.5.1 การเปรียบเทียบประสิทธิภาพของวิธีการประมาณค่าพารามิเตอร์

1.5.1.1 ศึกษาค้นคว้าเอกสารและข้อมูลเกี่ยวกับการประมาณค่าพารามิเตอร์จากข้อมูลที่ถูกตัดปลายทางขวา

1.5.1.2 สร้างข้อมูลที่มีการแจกแจง ลักษณะข้อมูล และขนาดตัวอย่าง (n) ตามที่กำหนด

1.5.1.3 คำนวณค่าพารามิเตอร์โดยวิธีการประมาณแบบภาวะน่าจะเป็นสูงสุดกับวิธีการประมาณแบบใช้กราฟ

1.5.2 การเลือกใช้ข้อมูลเพียงบางส่วนสำหรับวิธีการประมาณแบบกราฟ ซึ่งประกอบไปด้วย วิธี K Cluster-Mean, วิธี Trimmed q% และวิธี Trimmed q% & K Cluster-Mean

1.5.3 ทำซ้ำข้อ 1.5.1 – 1.5.2 จำนวน 5,000 รอบ

1.5.4 คำนวณค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ย

1.5.5 สรุปผลการศึกษา

1.6 ลำดับขั้นตอนในการเสนอผลการวิจัย

การวิจัยครั้งนี้ นำเสนอผลการวิจัยโดยเสนอค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ย ในแต่ละสถานการณ์ นำเสนอในรูปแบบกราฟและตาราง โดยการนำเสนอผลการวิจัยแบ่งออกเป็น 2 ส่วน

ส่วนที่ 1 นำเสนอผลการเปรียบเทียบประสิทธิภาพวิธีการประมาณพารามิเตอร์ 2 วิธี คือ วิธีการประมาณแบบภาวะน่าจะเป็นสูงสุด และวิธีการประมาณแบบกราฟ

ส่วนที่ 2 นำเสนอผลการเปรียบเทียบประสิทธิภาพวิธีการประมาณพารามิเตอร์ 2 วิธี คือ วิธีการประมาณแบบกราฟ และวิธีการประมาณแบบกราฟด้วยข้อมูลบางส่วน

1.7 ประโยชน์ที่คาดว่าจะได้รับ

1.7.1 เพื่อเปรียบเทียบประสิทธิภาพของตัวประมาณด้วยวิธีภาวะน่าจะเป็นสูงสุด (MLE) และตัวประมาณแบบกราฟ (GE) สำหรับข้อมูลที่ถูกตัดปลายทางขวาประเภทที่ 2

1.7.2 เพื่อเปรียบเทียบประสิทธิภาพของตัวประมาณแบบกราฟด้วยข้อมูลบางส่วน (GEPD) และตัวประมาณแบบกราฟ (GE) สำหรับข้อมูลที่ถูกตัดปลายทางขวาประเภทที่ 2

1.7.3 เพื่อเป็นแนวทางในการเลือกใช้วิธีการประมาณค่าพารามิเตอร์ได้อย่างเหมาะสม สำหรับข้อมูลที่ถูกตัดปลายทางขวาประเภทที่ 2 ซึ่งมีการแจกแจงปกติ การแจกแจงโลจิสติค การแจกแจงค่าต่ำสุดขีด และการแจกแจงค่าสูงสุดขีด

1.7.4 เพื่อเป็นแนวทางในการศึกษาเปรียบเทียบเกณฑ์ทางสถิติอื่นๆ ที่เกี่ยวข้องต่อไป

บทที่ 2
ทฤษฎีและตัวสถิติที่เกี่ยวข้อง

ในบทนี้กล่าวถึงทฤษฎีพื้นฐานต่างๆ ที่เกี่ยวข้อง และวิธีการประมาณค่าพารามิเตอร์จากข้อมูลที่ถูกตัดปลายทางขวาประเภทที่ 2 ด้วย 3 วิธีคือ วิธีภาวะน่าจะเป็นสูงสุด วิธีการประมาณแบบกราฟ และวิธีการประมาณแบบกราฟด้วยข้อมูลบางส่วน โดยการแจกแจงที่สนใจศึกษา ได้แก่ การแจกแจงปกติ การแจกแจงโลจิสติค การแจกแจงค่าต่ำสุดขีด และการแจกแจงค่าสูงสุดขีด

2.1 ทฤษฎีพื้นฐาน

2.1.1 ประเภทของการถูกตัด (Type of Censoring)

ก. การถูกตัดประเภทที่ 1 (Type I Censoring)

ข้อมูลที่ถูกตัดประเภทที่ 1 เกิดขึ้นเนื่องจากมีการกำหนดเวลาการสิ้นสุดการเก็บข้อมูลไว้ล่วงหน้า (Fixed Censoring Time; T_c) การกำหนดลักษณะการศึกษาแบบนี้เป็นเพราะด้วยเหตุของเวลาหรือค่าใช้จ่ายที่มีจำกัดจึงต้องกำหนดเวลาที่ชัดเจน ตัวอย่างเช่น การทดลองฉีดยาให้กับหนูทดลองซึ่งกำหนดระยะเวลาไว้ 6 เดือน ถ้าหนูทดลองตัวใดตายภายใน 6 เดือนจะถือว่าเป็นข้อมูลที่ไม่ถูกตัด และบันทึกเวลาค่าสังเกตตามจริง แต่ถ้าครบ 6 เดือนแล้วหนูทดลองตัวใดไม่ตายจะถือว่าเป็นข้อมูลที่ถูกตัดทิ้ง และบันทึกเวลาค่าสังเกตไว้เพียง 6 เดือนเท่านั้น สำหรับข้อมูลทางด้านการประกันภัย เช่น การประกันภัยสุขภาพ บริษัทจะจ่ายค่ารักษาพยาบาลให้ตามที่ผู้เอาประกันภัยจ่ายไปจริงแต่ไม่เกินผลประโยชน์สูงสุดที่กำหนดไว้

ให้ T_c เป็นเวลาที่กำหนดไว้ล่วงหน้า และให้ $T_1, T_2, ..., T_n$ เป็นค่าสังเกตที่ไม่ถูกตัดซึ่งมีการแจกแจงเหมือนกันและเป็นอิสระกัน จะได้ตัวแปรสุ่มของค่าสังเกต $X_1, X_2, ..., X_n$ ซึ่ง

$$X_i = \begin{cases} T_i & ; \ T_i \le T_c \\ T_c & ; \ T_i > T_c \end{cases}$$

โดยมีฟังก์ชันภาวะน่าจะเป็น (Likelihood Function) ดังนี้

$$L(x_i) = \begin{cases} f(x_i) & ; \text{ ถ้าค่าสังเกตไม่ถูกตัด} \\ P(T_i > T_c) = S(T_c) & ; \text{ ถ้าค่าสังเกตถูกตัด} \end{cases}$$

และมีฟังก์ชันภาวะน่าจะเป็นรวมดังนี้

$$L = \prod_{i \in u} f(x_i) \cdot S(T_c)^{\sum_{i=1}^{n} 1\{x_i > T_c\}} \quad ; \ i \in u \quad \text{หมายถึงค่าสังเกตที่ไม่ถูกตัดทิ้ง}$$

โดยที่ $\sum_{i=1}^{n} 1\{x_i > T_c\} = \begin{cases} 1 & ; \ x_i > T_c \\ 0 & ; \ x_i \le T_c \end{cases}$ เป็น indicator function

ข. การถูกตัดประเภทที่ 2 (Type II Censoring)

ในบางกรณีที่ไม่อาจกำหนดเวลา หรือค่าสูงสุดของการตัดข้อมูลที่เหมาะสมได้ ดังนั้น จะกำหนดจำนวนค่าสังเกตที่ไม่ถูกตัดทิ้ง หรือจำนวนค่าสังเกตความล้มเหลวที่กำหนดไว้ล่วงหน้า (Fixed Number of Uncensored Failure; r) แทน นั่นคือจะหยุดการทดลองเมื่อได้จำนวนค่าสังเกตตามกำหนด โดยไม่ต้องทำการทดลองจนครบตามขนาดตัวอย่าง เช่น การทดสอบอายุการใช้งานของหลอดไฟ จะกำหนดจำนวนหลอดไฟที่เสื่อมสภาพไว้ล่วงหน้า เริ่มทดลองโดยเปิดหลอดไฟให้ทำงานทั้งหมดแล้วเริ่มบันทึกเวลาและนับจำนวนหลอดไฟที่เสื่อมสภาพ เมื่อได้จำนวนหลอดไฟที่เสื่อมสภาพครบแล้วก็จะหยุดทำการทดลอง

ถ้า n คือจำนวนข้อมูลทั้งหมด และกำหนด r คือจำนวนค่าสังเกตที่ไม่ถูกตัดทิ้ง โดยที่ $r \le n$ ให้ $T_{(1)} \le T_{(2)} \le \ldots \le T_{(r)}$ เป็นค่าสังเกตที่ไม่ถูกตัดทิ้ง และ $T_{(r+1)} \le T_{(r+2)} \le \ldots \le T_{(n)}$ เป็นค่าสังเกตที่ถูกตัดทิ้ง ซึ่ง $T_{(i)} \ge T_{(r)}$; $i = r+1, r+2, \ldots, n$ ไม่ทราบค่าที่แท้จริงของค่าสังเกต ดังนั้น X_i เป็นตัวแปรสุ่มของค่าสังเกต ซึ่ง

$$X_i = \begin{cases} T_{(i)} & \text{เมื่อ } i \le r \\ T_{(r)} & \text{เมื่อ } i = r+1, r+2, \ldots, n \end{cases} .$$

และมีฟังก์ชันภาวะน่าจะเป็นรวมดังนี้

$$L = \prod_{i=1}^{r} f(x_i) \cdot S(T_{(r)})^{n-r}$$

2.1.2 ฟังก์ชันการอยู่รอด (Survival Function)

ให้ T เป็นตัวแปรสุ่มต่อเนื่อง

$f(t)$ เป็นฟังก์ชันความหนาแน่นของ T (Probability density function)

$F(t)$ เป็นฟังก์ชันการแจกแจงสะสมของ T (Distribution Function)

$S(t)$ เป็นฟังก์ชันการอยู่รอดของ T (Survival Function)

นิยามฟังก์ชัน $S(t)$ คือความน่าจะเป็นที่ตัวแปรสุ่ม T จะมีค่ามากกว่า t

$$S(t) = P(T > t)$$

$$= 1 - F(t) = \int_{t}^{\infty} f(u)du$$

โดยที่ $S(t)$ มีคุณสมบัติดังนี้

1. $S(t)$ เป็นฟังก์ชันไม่เพิ่ม (Nonincreasing Function)
2. $S(t)$ เป็นฟังก์ชันต่อเนื่องของ t
3. $S(t) = 1$ เมื่อ $t = 0$
4. $S(t) = 0$ เมื่อ $t = \infty$

2.2 การแจกแจงที่ใช้ในการวิจัย

สำหรับการวิจัยในครั้งนี้ ศึกษาจากตัวอย่างสุ่มที่มาจากประชากรที่มีการแจกแจงแบบต่างๆ คือ ที่มีลักษณะสมมาตร (การแจกแจงปกติ, การแจกแจงโลจิสติค) ที่มีลักษณะเบ้ซ้าย (การแจกแจงค่าต่ำสุดขีด) และที่มีลักษณะเบ้ขวา (การแจกแจงค่าสูงสุดขีด) ดังนี้

2.2.1 การแจกแจงปกติ (Normal Distribution)

ให้ X เป็นตัวแปรสุ่มต่อเนื่องที่มีการแจกแจงปกติ ด้วยพารามิเตอร์ μ และ σ เขียนแทนด้วย $X \sim NOR(\mu, \sigma)$ แล้วฟังก์ชันความหนาแน่นน่าจะเป็น (Probability Density Function; PDF) อยู่ในรูปของ

$$f(x;\mu,\sigma)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] \qquad ,-\infty<x<\infty.$$

เมื่อ $\sigma>0$ และ $-\infty<\mu<\infty$ โดยที่ μ เป็นพารามิเตอร์แสดงตำแหน่ง (location parameter) และ σ เป็นพารามิเตอร์แสดงขนาด (scale parameter) ของฟังก์ชันความหนาแน่น และมีฟังก์ชันการแจกแจงสะสม (Cumulative Distribution Function ; CDF) อยู่ในรูปของ

$$F(x;\mu,\sigma)=\Phi\left(\frac{x-\mu}{\sigma}\right).$$

เมื่อ $\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\exp\left[\frac{-t^2}{2}\right]dt$ เป็น Standard Normal CDF

ซึ่งมีค่าเฉลี่ยและค่าความแปรปรวน คือ

$$E(X)=\mu \quad \text{และ} \quad Var(X)=\sigma^2 \quad \text{ตามลำดับ}$$

2.2.2 การแจกแจงโลจิสติค (Logistic Distribution)

ให้ X เป็นตัวแปรสุ่มต่อเนื่องที่มีการแจกแจงโลจิสติค ด้วยพารามิเตอร์ μ และ σ เขียนแทนด้วย $X\sim LOGIS(\mu,\sigma)$ แล้วฟังก์ชันความหนาแน่นน่าจะเป็น (PDF) อยู่ในรูปของ

$$f(x;\mu,\sigma)=\frac{1}{\sigma}\frac{\exp\left[\frac{x-\mu}{\sigma}\right]}{\left[1+\exp\left(\frac{x-\mu}{\sigma}\right)\right]^2} \qquad ,-\infty<x<\infty\ .$$

เมื่อ $\sigma>0$ และ $-\infty<\mu<\infty$ โดยที่ μ เป็นพารามิเตอร์แสดงตำแหน่ง (location parameter) และ σ เป็นพารามิเตอร์แสดงขนาด (scale parameter) ของฟังก์ชันความหนาแน่น และมีฟังก์ชันการแจกแจงสะสม (CDF) อยู่ในรูปของ

$$F(x;\mu,\sigma)=\frac{\exp\left[\frac{x-\mu}{\sigma}\right]}{\left[1+\exp\left(\frac{x-\mu}{\sigma}\right)\right]}\ .$$

ซึ่งมีค่าเฉลี่ยและค่าความแปรปรวน คือ

$$E(X)=\mu \quad \text{และ} \quad Var(X)=\sigma^2\pi^2/3 \quad \text{ตามลำดับ}$$

2.2.3 การแจกแจงค่าต่ำสุดขีด (Smallest Extreme Value Distribution)

ให้ X เป็นตัวแปรสุ่มต่อเนื่องที่มีการแจกแจงค่าต่ำสุดขีด ด้วยพารามิเตอร์ μ และ σ เขียนแทนด้วย $X \sim SEV(\mu,\sigma)$ แล้วฟังก์ชันความหนาแน่นน่าจะเป็น (PDF) อยู่ในรูปของ

$$f(x;\mu,\sigma)=\frac{1}{\sigma}\exp\left[\frac{x-\mu}{\sigma}-\exp\left(\frac{x-\mu}{\sigma}\right)\right] \qquad , -\infty < x < \infty .$$

เมื่อ $\sigma > 0$ และ $-\infty < \mu < \infty$ โดยที่ μ เป็นพารามิเตอร์แสดงตำแหน่ง (location parameter) และ σ เป็นพารามิเตอร์แสดงขนาด (scale parameter) ของฟังก์ชันความหนาแน่น และมีฟังก์ชันการแจกแจงสะสม (CDF) อยู่ในรูปของ

$$F(x;\mu,\sigma)=1-\exp\left[-\exp\left(\frac{x-\mu}{\sigma}\right)\right].$$

ซึ่งมีค่าเฉลี่ยและค่าความแปรปรวน คือ

$$E(X)=\mu-\sigma\gamma \quad \text{และ} \quad Var(X)=\sigma^2\pi^2/6 \quad \text{ตามลำดับ}$$

และ $\gamma \approx 0.5772$ คือค่าคงที่ของออยเลอร์ (Euler's constant)

2.2.4 การแจกแจงค่าสูงสุดขีด (Largest Extreme Value Distribution)

ให้ X เป็นตัวแปรสุ่มต่อเนื่องที่มีการแจกแจงค่าสูงสุดขีด ด้วยพารามิเตอร์ μ และ σ เขียนแทนด้วย $X \sim LEV(\mu,\sigma)$ แล้วฟังก์ชันความหนาแน่นน่าจะเป็น (PDF) อยู่ในรูปของ

$$f(x;\mu,\sigma)=\frac{1}{\sigma}\exp\left[-\left(\frac{x-\mu}{\sigma}\right)-\exp\left\{-\left(\frac{x-\mu}{\sigma}\right)\right\}\right] \ , -\infty < x < \infty .$$

เมื่อ $\sigma > 0$ และ $-\infty < \mu < \infty$ โดยที่ μ เป็นพารามิเตอร์แสดงตำแหน่ง (location parameter) และ σ เป็นพารามิเตอร์แสดงขนาด (scale parameter) ของฟังก์ชันความหนาแน่น และมีฟังก์ชันการแจกแจงสะสม (CDF) อยู่ในรูปของ

$$F(x;\mu,\sigma) = \exp\left[-\exp\left\{-\left(\frac{x-\mu}{\sigma}\right)\right\}\right].$$

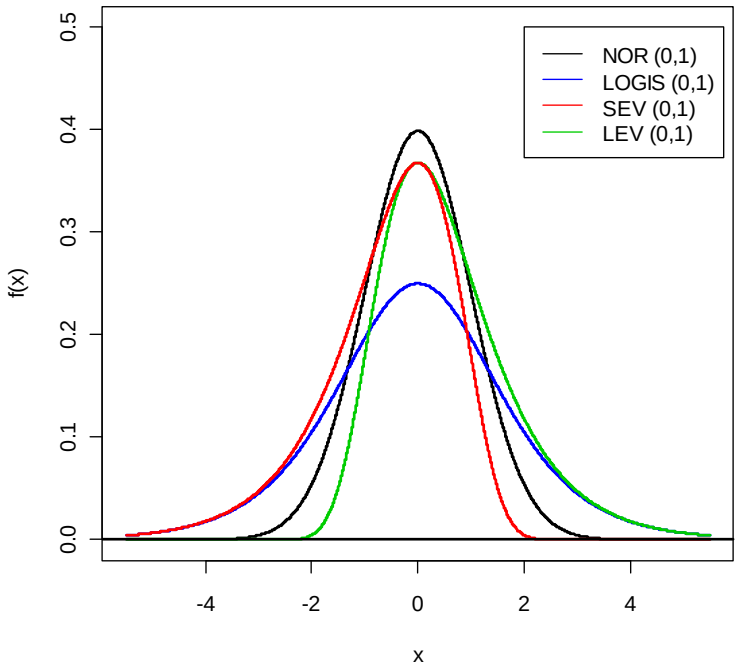
ซึ่งมีค่าเฉลี่ยและค่าความแปรปรวน คือ

$$E(X) = \mu + \sigma\gamma \quad \text{และ} \quad Var(X) = \sigma^2\pi^2/6 \quad \text{ตามลำดับ}$$

และ $\gamma \approx 0.5772$ คือค่าคงที่ของออยเลอร์ (Euler's constant)

<u>หมายเหตุ</u>:- สามารถแสดงความสัมพันธ์ระหว่าง LEV และ SEV ได้ดังนี้ ถ้า $X \sim LEV(\mu,\sigma)$ แล้ว $-X \sim SEV(-\mu,\sigma)$

ภาพที่ 2.1 แสดงฟังก์ชันความหนาแน่นน่าจะเป็นของการแจกแจง NOR, LOGIS, SEV และ LEV ที่พารามิเตอร์ μ =0 และ σ =1



2.3 การประมาณค่าพารามิเตอร์

2.3.1 วิธีภาวะน่าจะเป็นสูงสุด (Maximum Likelihood Estimation Method; MLE Method)

ให้ $X_1,...,X_n$ เป็นตัวอย่างสุ่มขนาด n จากประชากรที่มีฟังก์ชันความหนาแน่นน่าจะเป็น $f(x;\theta)$

ฟังก์ชันภาวะน่าจะเป็น (Likelihood Function) ของตัวอย่างสุ่ม ได้แก่ ฟังก์ชันความหนาแน่นน่าจะเป็นร่วมของ $X_1,...,X_n$ โดยถือว่าเป็นฟังก์ชันของพารามิเตอร์ θ ซึ่งเรามักแทนฟังก์ชันภาวะน่าจะเป็นด้วย L หรือ $L(\theta;x_1,...,x_n)$ หรือ $L(\theta)$ นั่นคือ $L=f(x_1;\theta),...,f(x_n;\theta)$ โดยถือว่า L เป็นฟังก์ชันของ θ

การหาค่าของพารามิเตอร์ θ ที่ทำให้ $L(\theta;x_1,...,x_n)$ มีค่ามากที่สุด (ธีระพร, 2536)

$$\hat{\theta}=\arg\max_{\theta} L(\theta;x_1,...,x_n)$$

2.3.2 วิธีการประมาณแบบกราฟ (Graphical Estimation Method; GE Method)

ให้ X เป็นตัวแปรสุ่มจากฟังก์ชันการแจกแจง F โดยที่

$$P(X\le x;\mu,\sigma)=F_{\mu,\sigma}(x)$$

และ F มาจากการแจกแจงแบบ location-scale family (Klein & Moesschberger, 1997)

ก็ต่อเมื่อ $$F_{\mu,\sigma}(x)=F_{0,1}\left(\frac{x-\mu}{\sigma}\right)$$

โดยจะเรียก μ ว่า location parameter เมื่อ $-\infty<\mu<\infty$

σ ว่า scale parameter เมื่อ $\sigma>0$

เนื่องจาก $F_{\mu,\sigma}(x)=F_{0,1}\left(\frac{x-\mu}{\sigma}\right)$ ทำให้

$$x_p=\mu+\sigma F_{0,1}^{-1}(p)$$

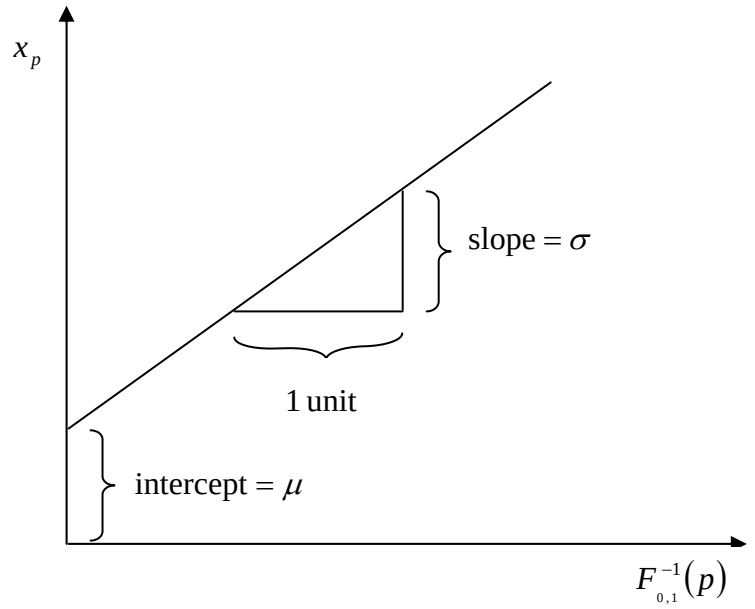
โดยที่ $F_{0,1}^{-1}(p)$ คือ ฟังก์ชันควอนไทล์ (quantile function) ของ $F_{0,1}$

และ x_p คือ ควอนไทล์ตัวที่ p ของ X

จะเห็นว่า ควอนไทล์ของ X ($X \sim F_{\mu,\sigma}$) เป็นฟังก์ชันเชิงเส้นของ ควอนไทล์จาก $F_{0,1}$

ถ้าเราพล็อตกราฟระหว่าง x_p กับ $F_{0,1}^{-1}(p)$ โดยพล็อต x_p บนแกน y และพล็อต $F_{0,1}^{-1}(p)$ บนแกน x โดยที่ μ และ σ เป็นค่าจุดตัดบนแกน y (intercept) และค่าความชัน (slope) ของเส้นตรง ตามลำดับ

ภาพที่ 2.2 แสดงการประมาณด้วยวิธีแบบใช้กราฟ



จากความสัมพันธ์ของ x_p และ $F_{0,1}^{-1}(p)$ เราสามารถนำมาใช้ในการประมาณโดยใช้กราฟ ด้วยการพล็อตกราฟระหว่าง $x_{(i)}$ กับ $F_{0,1}^{-1}(p_{(i)})$ โดยให้ $x_{(i)}$ เป็น sample quantile และ $p_{(i)}$ เป็นลำดับควอนไทล์ของ $x_{(i)}$ ดังนั้น ค่าประมาณของ intercept และ slope จากกราฟดังกล่าว จึงสามารถนำมาใช้ในการประมาณค่า μ และ σ ได้ (Lawless, 2003; Somboonsavatdee & Nair, 2007)

2.3.3 วิธีการประมาณแบบกราฟด้วยข้อมูลบางส่วน (Graphical Estimation with Partial Data Method; GEPD Method)

การหาค่าพารามิเตอร์ θ^* ทำแบบเดียวกับการประมาณด้วยวิธีแบบใช้กราฟ กล่าวคือ เป็นการพล็อตกราฟระหว่าง $x^*_{(i)}$ กับ $F^{-1}_{0,1}\left(p^*_{(i)}\right)$ โดยให้ $x^*_{(i)}$ เป็นข้อมูลบางส่วนที่เหลืออยู่จากวิธี Trimmed q%, K Cluster-Mean, หรือ Trimmed q% & K Cluster-Mean ซึ่ง $x^*_{(i)}$ นั้นเป็น sample quantile และ $p^*_{(i)}$ เป็นลำดับควอนไทล์ของ $x^*_{(i)}$ ดังนั้น ค่าประมาณของ intercept และ slope จากกราฟ จึงนำมาใช้ในการประมาณค่า μ และ σ ได้เช่นเดียวกัน

2.4 เกณฑ์ที่ใช้ในการพิจารณาประสิทธิภาพของการประมาณ

2.4.1 การประมาณค่าแบบจุด จะพิจารณาเปรียบเทียบค่าเฉลี่ยของความคลาดเคลื่อนกำลังสอง (Mean Square Error; MSE)

กำหนดให้ μ เป็นพารามิเตอร์แสดงตำแหน่ง (location parameter)

σ เป็นพารามิเตอร์แสดงขนาด (scale parameter)

$\hat{\mu}_i$ เป็นค่าประมาณของพารามิเตอร์ μ จากตัวอย่างที่ i

$\hat{\sigma}_i$ เป็นค่าประมาณของพารามิเตอร์ σ จากตัวอย่างที่ i

$$MSE(\mu) = \frac{\sum_{i=1}^{M}(\hat{\mu}_i - \mu)^2}{M} \quad \text{และ} \quad MSE(\sigma) = \frac{\sum_{i=1}^{M}(\hat{\sigma}_i - \sigma)^2}{M}$$

เมื่อ M = จำนวนครั้งของการทดลองในแต่ละสถานการณ์ของการทดลอง

โดยถ้าวิธีการประมาณค่าแบบใด ที่ทำให้ได้ค่า MSE ต่ำกว่าจะถือว่าการประมาณค่าจากวิธีนั้น เป็นวิธีที่มีประสิทธิภาพมากกว่า ทั้งนี้จะใช้ค่าความเอนเอียงสัมบูรณ์ (Absolute Bias; $|\text{Bias}|$) และค่าความแปรปรวน (Variance) ประกอบการพิจารณาด้วย

2.4.2 ประสิทธิภาพสัมพัทธ์ (Relative Efficiency; RE) ของตัวประมาณพารามิเตอร์ μ และ σ จะพิจารณาเปรียบเทียบใน 2 รูปแบบ คือ

<u>รูปแบบที่ 1</u>: ค่าประสิทธิภาพสัมพัทธ์ที่คิดจากค่าสัดส่วนของ Variances ของตัวประมาณพารามิเตอร์ สำหรับ 2 วิธีที่ต้องการเปรียบเทียบ

กำหนดให้ RE1 เป็นค่าประสิทธิภาพสัมพัทธ์ที่คิดจากค่าสัดส่วนของ Variances

$\mathrm{Var}_{\mathrm{MLE}}$ เป็นค่าความแปรปรวนของตัวประมาณด้วยวิธี MLE

$\mathrm{Var}_{\mathrm{GE}}$ เป็นค่าความแปรปรวนของตัวประมาณด้วยวิธี GE

$\mathrm{Var}_{\mathrm{GEPD}}$ เป็นค่าความแปรปรวนของตัวประมาณด้วยวิธี GEPD

$$\mathrm{RE1}(\mathrm{GE},\mathrm{MLE}) = \frac{\mathrm{Var}_{\mathrm{MLE}}}{\mathrm{Var}_{\mathrm{GE}}} \quad \text{และ} \quad \mathrm{RE1}(\mathrm{GEPD},\mathrm{GE}) = \frac{\mathrm{Var}_{\mathrm{GE}}}{\mathrm{Var}_{\mathrm{GEPD}}}$$

<u>รูปแบบที่ 2</u>: ค่าประสิทธิภาพสัมพัทธ์ที่คิดจากค่าสัดส่วนของ MSEs ของตัวประมาณพารามิเตอร์ สำหรับ 2 วิธีที่ต้องการเปรียบเทียบ

กำหนดให้ RE2 เป็นค่าประสิทธิภาพสัมพัทธ์ที่คิดจากค่าสัดส่วนของ MSE

$\mathrm{MSE}_{\mathrm{MLE}}$ เป็นค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณด้วยวิธี MLE

$\mathrm{MSE}_{\mathrm{GE}}$ เป็นค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณด้วยวิธี GE

$\mathrm{MSE}_{\mathrm{GEPD}}$ เป็นค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณด้วยวิธี GEPD

$$\mathrm{RE2}(\mathrm{GE},\mathrm{MLE}) = \frac{\mathrm{MSE}_{\mathrm{MLE}}}{\mathrm{MSE}_{\mathrm{GE}}} \quad \text{และ} \quad \mathrm{RE2}(\mathrm{GEPD},\mathrm{GE}) = \frac{\mathrm{MSE}_{\mathrm{GE}}}{\mathrm{MSE}_{\mathrm{GEPD}}}$$

ซึ่งถ้า $\mathrm{RE}(\mathrm{GE},\mathrm{MLE})$ มีค่ามากกว่า 1 หมายความว่า ตัวประมาณ GE มีประสิทธิภาพมากกว่า MLE, ถ้า RE มีค่าน้อยกว่า 1 หมายความว่า ตัวประมาณ GE มีประสิทธิภาพน้อยกว่า MLE, และถ้า RE มีค่าเท่ากับ 1 หมายความว่า ตัวประมาณ GE และ MLE มีประสิทธิภาพเท่ากัน

และถ้า $\mathrm{RE}(\mathrm{GEPD},\mathrm{GE})$ มีค่ามากกว่า 1 หมายความว่า ตัวประมาณ GEPD มีประสิทธิภาพมากกว่า GE, ถ้า RE มีค่าน้อยกว่า 1 หมายความว่า ตัวประมาณ GEPD มีประสิทธิภาพน้อยกว่า GE, และถ้า RE มีค่าเท่ากับ 1 หมายความว่า ตัวประมาณ GEPD และ GE มีประสิทธิภาพเท่ากัน

บทที่ 3

วิธีดำเนินการวิจัย

ในการศึกษาครั้งนี้สิ่งที่ต้องการคือเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์สำหรับข้อมูลที่มีค่าถูกตัดทิ้งประเภทที่ 2 ซึ่งได้เสนอไว้ 3 วิธีด้วยกันคือ วิธีภาวะน่าจะเป็นสูงสุด วิธีการประมาณแบบกราฟ และวิธีการประมาณแบบกราฟด้วยข้อมูลบางส่วน โดยทำการศึกษาสำหรับข้อมูลที่มีการแจกแจง 4 แบบ คือแบบปกติ แบบโลจิสติค แบบค่าต่ำสุดขีด และแบบค่าสูงสุดขีด ที่มีขนาดตัวอย่างเป็น 20, 40, 80 และ 120 ด้วยสัดส่วนของข้อมูลที่ถูกตัดทิ้งเป็น 10%, 20% และ 30% จะพิจารณาประสิทธิภาพของการประมาณด้วยค่าประสิทธิภาพสัมพัทธ์ของตัวประมาณค่าพารามิเตอร์ μ และ σ โดยจะใช้ค่าความเอนเอียงสัมบูรณ์ ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ย ประกอบการพิจารณาด้วย สำหรับการวิจัยครั้งนี้ใช้โปรแกรม R เวอร์ชัน 2.9.2 (R Development Core Team, 2009) ในการจำลองข้อมูล

3.1 แผนการดำเนินการวิจัย

การวิจัยครั้งนี้ได้กำหนดสถานการณ์ต่างๆ ในการศึกษา ดังนี้

3.1.1 เมื่อข้อมูลมีการแจกแจงปกติ กำหนดให้ข้อมูลมีค่าพามิเตอร์ μ =0 และ σ =1

3.1.2 เมื่อข้อมูลมีการแจกแจงโลจิสติค กำหนดให้ข้อมูลมีค่าพามิเตอร์ μ =0 และ σ =1

3.1.3 เมื่อข้อมูลมีการแจกแจงค่าต่ำสุดขีด กำหนดให้ข้อมูลมีค่าพามิเตอร์ μ =0 และ σ =1

3.1.4 เมื่อข้อมูลมีการแจกแจงค่าสูงสุดขีด กำหนดให้ข้อมูลมีค่าพามิเตอร์ μ =0 และ σ =1

3.1.5 วิธีการประมาณค่าพารามิเตอร์ที่นำมาใช้เปรียบเทียบ มี 3 วิธี ดังนี้

- วิธีภาวะน่าจะเป็นสูงสุด (MLE)

- วิธีประมาณแบบกราฟ (GE)

- วิธีประมาณแบบกราฟด้วยข้อมูลบางส่วน (GEPD) ซึ่งได้กำหนดเกณฑ์การเลือกใช้ข้อมูลเพียงบางส่วนสำหรับข้อมูลที่ไม่ถูกตัดปลายทางขวา ($p = 0$) และข้อมูลที่ถูกตัดปลายทางขวา ($p \neq 0$) ไว้ดังตารางที่ 3.1 สามารถอธิบายได้ดังนี้

(1) วิธี K Cluster-Mean ; K = 4, 8 สามารถทำได้โดยการแบ่งข้อมูลที่มีอยู่ทั้งหมดออกเป็นกลุ่มๆ ละเท่าๆ กันตามจำนวน Cluster ที่ต้องการ ต่อจากนั้นจึงหาค่าเฉลี่ยของข้อมูลภายในกลุ่ม เพื่อให้เหลือข้อมูลเพียงจุดเดียวสำหรับแต่ละกลุ่ม ซึ่งงานวิจัยชิ้นนี้สนใจศึกษาการแบ่ง Cluster ที่ 2 ระดับคือ 4 กลุ่ม และ 8 กลุ่ม ตัวอย่างเช่น

เมื่อ n = 40; วิธี 4 Cluster ทำได้โดยแบ่งข้อมูลออกเป็น 10, 10, 10, 10

วิธี 8 Cluster ทำได้โดยแบ่งข้อมูลออกเป็น 5, 5, 5, 5, 5, 5, 5, 5

และถ้าข้อมูลมีสัดส่วนตัดปลายทางขวาที่ขนาด p = 0.1 สำหรับ n = 40 จะเหลือข้อมูลอยู่ 36 ตัว ซึ่งวิธี 8 Cluster จะทำให้ข้อมูลในแต่ละกลุ่มมีจำนวนไม่เท่ากัน ในกรณีนี้เราจะให้ความสำคัญกับข้อมูลที่อยู่ช่วงตรงกลาง โดยจะแบ่งข้อมูลออกเป็น 4, 4, 5, 5, 5, 5, 4, 4

<u>หมายเหตุ</u> จะไม่ใช้วิธี 8 Cluster ที่ n = 20 เนื่องจากมีกลุ่มตัวอย่างน้อยเกินไป

(2) วิธี Trimmed q% ; q = 5%, 10% สามารถทำได้โดยการตัดส่วนปลายข้อมูลทั้งสองข้างออกเป็นจำนวนเท่าๆ กัน ซึ่งงานวิจัยชิ้นนี้สนใจศึกษาการตัดข้อมูลที่ 2 ระดับคือ 5% และ 10% ตัวอย่างเช่น

เมื่อ n = 40; วิธี Trimmed 5% ทำได้โดยการตัดข้อมูลออก 2 ตัว ซึ่งจะแบ่งตัดส่วนปลายออกข้างละ 1 ตัว เป็น ~~1~~, 38, ~~1~~

วิธี Trimmed 10% สามารถทำได้โดยการตัดข้อมูลออก 4 ตัว ซึ่งจะแบ่งตัดส่วนปลายออกข้างละ 2 ตัว เป็น ~~2~~, 36, ~~2~~

และถ้าข้อมูลมีสัดส่วนตัดปลายทางขวาที่ขนาด p = 0.1 สำหรับ n = 40 จะเหลือข้อมูลอยู่ 36 ตัว ซึ่งวิธี Trimmed 5% นั้นจะต้องตัดข้อมูลส่วนปลายออกข้างละ 1 ตัว แต่เนื่องจากข้อมูลส่วนปลายทางขวาถูกตัดออกไปแล้วจากการเซ็นเซอร์เป็นจำนวน 4 ตัว ดังนั้นการตัดปลายข้อมูลจึงเหลือข้างซ้ายเพียงข้างเดียวจำนวน 1 ตัว เป็น ~~1~~, 35

<u>หมายเหตุ</u> จะไม่ใช้วิธี Trimmed 5% ที่ n = 20 เนื่องจากไม่สามารถแบ่งตัดข้อมูลทั้งสองข้างได้

(3) วิธี Trimmed q% & K Cluster-Mean ; q = 5%, 10% และ K = 4, 8 เกิดจากการนำวิธีในข้อ (1) และ ข้อ (2) มารวมกัน สามารถทำได้โดยการตัดข้อมูลที่มากสุดและน้อยสุดออกเป็นจำนวนเท่าๆ กันแล้วแบ่งข้อมูลออกเป็นกลุ่มๆ เพื่อหาค่าเฉลี่ยของแต่ละกลุ่ม โดย

งานวิจัยชิ้นนี้สนใจศึกษาวิธี Trimmed q% & K Cluster-Mean อยู่ 4 แบบคือ Trimmed 5% & 4 Cluster-Mean, Trimmed 5% & 8 Cluster-Mean, Trimmed 10% & 4 Cluster-Mean และ Trimmed10% & 8 Cluster-Mean ตัวอย่างเช่น

เมื่อ n = 40; วิธี Trimmed 5% & 4 Cluster-Mean ทำได้โดยการตัดข้อมูลออก 2 ตัว ซึ่งจะแบ่งตัดส่วนปลายออกข้างละ 1 ตัว เป็น ~~1~~, 38, ~~1~~ จะเหลือข้อมูลอยู่ 38 ตัวเพื่อทำการแบ่ง 4 Cluster เป็น 9, 10, 10, 9

วิธี Trimmed 10% & 4 Cluster-Mean ก็สามารถทำได้โดยการตัดข้อมูลออก 4 ตัว ซึ่งจะแบ่งตัดส่วนปลายออกข้างละ 2 ตัว เป็น ~~2~~, 36, ~~2~~ จะเหลือข้อมูลอยู่ 36 ตัวเพื่อทำการแบ่ง 4 Cluster เป็น 9, 9, 9, 9

และถ้าข้อมูลมีสัดส่วนตัดปลายทางขวาที่ขนาด p = 0.1 สำหรับ n = 40 จะเหลือข้อมูลอยู่ 36 ตัว ซึ่งวิธี Trimmed 5% & 4 Cluster-Mean นั้นจะต้องตัดข้อมูลส่วนปลายออกข้างละ 1 ตัว แต่เนื่องจากข้อมูลส่วนปลายทางขวาถูกตัดออกไปแล้วจากการเซ็นเซอร์เป็นจำนวน 4 ตัว ดังนั้นการตัดปลายข้อมูลจึงเหลือข้างซ้ายเพียงข้างเดียวจำนวน 1 ตัว เป็น ~~1~~, 35 จะเหลือข้อมูลอยู่ 35 ตัว เพื่อทำการแบ่ง 4 Cluster เป็น 9, 9, 9, 8

<u>หมายเหตุ</u> การแบ่ง Cluster หากแบ่งไม่ลงตัว จะให้ความสำคัญกับข้อมูลที่อยู่ตรงกลางทางซ้ายก่อนเป็นอันดับแรก

3.1.6 ขนาดตัวอย่าง (Sample Size; n) ที่ใช้ในการศึกษาเท่ากับ 20, 40, 80 และ 120

3.1.7 กำหนดสัดส่วนของข้อมูลที่ถูกตัดทิ้ง (Censoring Proportion; p) เป็น 3 ระดับ คือ 10%, 20% และ 30%

3.1.8 การจำลองข้อมูลในแต่ละสถานการณ์จะกระทำซ้ำ 5,000 รอบ

ตารางที่ 3.1 แสดงเกณฑ์ในการ Trim ข้อมูล และแบ่ง Cluster สำหรับการประมาณพารามิเตอร์ μ และ σ ด้วยวิธีการประมาณแบบใช้กราฟ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา

Censoring Proportion p	Sample Size n	GE	GEPD							
			4 Cluster-Mean	8 Cluster-Mean	Trimmed 5%	Trimmed 10%	Trimmed 5% & 4 Cluster-Mean	Trimmed 5% & 8 Cluster-Mean	Trimmed 10% & 4 Cluster-Mean	Trimmed 10% & 8 Cluster-Mean
0	20	20	5, 5, 5, 5	-	-	~~1~~, 18, ~~1~~	-	-	4, 5, 5, 4	-
	40	40	10, 10, 10, 10	5, 5, 5, 5, 5, 5, 5, 5	~~1~~, 38, ~~1~~	~~2~~, 36, ~~2~~	9, 10, 10, 9	4, 5, 5, 5, 5, 5, 5, 4	9, 9, 9, 9	4, 4, 5, 5, 5, 5, 4, 4
	80	80	20, 20, 20, 20	10, 10, 10, 10, 10, 10, 10, 10	~~2~~, 76, ~~2~~	~~4~~, 72, ~~4~~	19, 19, 19, 19	9, 9, 10, 10, 10, 10, 9, 9	18, 18, 18, 18	9, 9, 9, 9, 9, 9, 9, 9
	120	120	30, 30, 30, 30	15, 15, 15, 15, 15, 15, 15, 15	~~3~~, 114, ~~3~~	~~6~~, 108, ~~6~~	28, 29, 29, 28	14, 14, 14, 15, 15, 14, 14, 14	27, 27, 27, 27	13, 13, 14, 14, 14, 14, 13, 13
0.1	20	~~18, 2~~	4, 5, 5, 4	-	-	~~1~~, 17	-	-	4, 5, 4, 4	-
	40	~~36, 4~~	9, 9, 9, 9	4, 4, 5, 5, 5, 5, 4, 4	~~1~~, 35	~~2~~, 34	9, 9, 9, 8	4, 4, 5, 5, 5, 4, 4, 4	8, 9, 9, 8	4, 4, 4, 5, 5, 4, 4, 4
	80	~~72, 8~~	18, 18, 18, 18	9, 9, 9, 9, 9, 9, 9, 9	~~2~~, 70	~~4~~, 68	17, 18, 18, 17	8, 9, 9, 9, 9, 9, 9, 8	17, 17, 17, 17	8, 8, 9, 9, 9, 9, 8, 8
	120	~~108, 12~~	27, 27, 27, 27	13, 13, 14, 14, 14, 14, 13, 13	~~3~~, 105	~~6~~, 102	26, 27, 26, 26	13, 13, 13, 14, 13, 13, 13, 13	25, 26, 26, 25	12, 13, 13, 13, 13, 13, 13, 12

ตารางที่ 3.1 (ต่อ) แสดงเกณฑ์ในการ Trim ข้อมูล และแบ่ง Cluster สำหรับการประมาณพารามิเตอร์ μ และ σ ด้วยวิธีการประมาณแบบใช้กราฟ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา

Censoring Proportion p	Sample Size n	GE	GEPD							
			4 Cluster-Mean	8 Cluster-Mean	Trimmed 5%	Trimmed 10%	Trimmed 5% & 4 Cluster-Mean	Trimmed 5% & 8 Cluster-Mean	Trimmed 10% & 4 Cluster-Mean	Trimmed 10% & 8 Cluster-Mean
0.2	20	16, 4	4, 4, 4, 4	-	-	1, 15	-	-	4, 4, 4, 3	-
	40	32, 8	8, 8, 8, 8	4, 4, 4, 4, 4, 4, 4, 4	1, 31	2, 30	8, 8, 8, 7	4, 4, 4, 4, 4, 4, 4, 3	7, 8, 8, 7	3, 4, 4, 4, 4, 4, 4, 3
	80	64, 16	16, 16, 16, 16	8, 8, 8, 8, 8, 8, 8, 8	2, 62	4, 60	15, 16, 16, 15	7, 8, 8, 8, 8, 8, 8, 7	15, 15, 15, 15	7, 7, 8, 8, 8, 8, 7, 7
	120	96, 24	24, 24, 24, 24	12, 12, 12, 12, 12, 12, 12, 12	3, 93	6, 90	23, 24, 23, 23	11, 12, 12, 12, 12, 12, 11, 11	22, 23, 23, 22	11, 11, 11, 12, 12, 11, 11, 11
0.3	20	14, 6	3, 4, 4, 3	-	-	1, 13	-	-	3, 4, 3, 3	-
	40	28, 12	7, 7, 7, 7	3, 3, 4, 4, 4, 4, 3, 3	1, 27	2, 26	7, 7, 7, 6	3, 3, 4, 4, 4, 3, 3	6, 7, 7, 6	3, 3, 3, 4, 4, 3, 3, 3
	80	56, 24	14, 14, 14, 14	7, 7, 7, 7, 7, 7, 7, 7	2, 54	4, 52	13, 14, 14, 13	6, 7, 7, 7, 7, 7, 7, 6	13, 13, 13, 13	6, 6, 7, 7, 7, 7, 6, 6
	120	84, 36	21, 21, 21, 21	10, 10, 11, 11, 11, 11, 10, 10	3, 81	6, 78	20, 21, 20, 20	10, 10, 10, 11, 10, 10, 10, 10	19, 20, 20, 19	9, 10, 10, 10, 10, 10, 10, 9

3.2 ขั้นตอนในการดำเนินการวิจัย

3.2.1 จำลองข้อมูลที่มีการแจงแจงแบบ $NOR(0,1)$, $LOGIS(0,1)$, $SEV(0,1)$ และ $LEV(0,1)$ ภายใต้ขนาดตัวอย่าง (n) เท่ากับ 20, 40, 80 และ 120

3.2.2 กำหนดสัดส่วนของข้อมูลที่ถูกตัดทิ้ง (p) เป็น 10%, 20% และ 30% เพื่อใช้ในการกำหนดจำนวนความล้มเหลว (Fixed Number of Uncensored Failure; r)

3.2.3 นำข้อมูลที่ได้ในข้อ 3.2.1 ไปเปรียบเทียบกับจำนวนความล้มเหลวในข้อ 3.2.2 เพื่อสร้างชุดข้อมูลที่ถูกตัดปลายทางขวาประเภทที่ 2 (Type-II Right Censoring) โดยกำหนด $T_{(i)} = T_{(r)}$ เมื่อ $T_{(i)} > T_{(r)}$ ในการวิจัยครั้งนี้จะกำหนดจำนวนค่าสังเกตที่จะเกิดความล้มเหลวไว้ล่วงหน้า ดังนั้น จะหยุดการทดลองเมื่อได้จำนวนค่าสังเกตตามกำหนด โดยไม่ต้องทำการทดลองจนครบตามขนาดตัวอย่าง เช่น ถ้าสุ่มข้อมูลที่มีการแจกแจงแบบ $NOR(0,1)$ ด้วยขนาดตัวอย่าง (n) เท่ากับ 10 และมีสัดส่วนของการตัดปลาย (p) เท่ากับ 20% แล้วจะสามารถสร้างชุดข้อมูลที่ถูกตัดปลายทางขวาประเภทที่ 2 ได้ดังนี้

(1) สุ่มข้อมูลที่มีการแจกแจงแบบ $NOR(0,1)$ มา 10 ค่าสังเกต ได้แก่ T_1 = -0.935 , T_2 = -0.531 , T_3 = -0.007 , T_4 = 0.636 , T_5 = -0.159 , T_6 = 0.725 , T_7 = -0.752 , T_8 = 1.320 , T_9 = 2.109 , T_{10} = 0.369

(2) นำข้อมูลที่สุ่มได้มาเรียงลำดับจากน้อยไปมาก โดยกำหนด $T_{(1)}$ = -0.935 , $T_{(2)}$ = -0.752 , $T_{(3)}$ = -0.531 , $T_{(4)}$ = -0.159 , $T_{(5)}$ = -0.007 , $T_{(6)}$ = 0.369 , $T_{(7)}$ = 0.636 , $T_{(8)}$ = 0.725 , $T_{(9)}$ = 1.320 , $T_{(10)}$ = 2.109 จะได้ $T_{(1)} \le T_{(2)} \le ... \le T_{(10)}$ เป็นสถิติเชิงอันดับ

(3) ด้วยสัดส่วนของการตัดปลาย p = 0.2 และ n = 10 ดังนั้นจะเหลือข้อมูลที่ไม่ถูกตัดทิ้งทั้งหมด 8 ค่า ($r = 8$) นั่นคือ $T_{(1)} \le T_{(2)} \le ... \le T_{(8)}$ เป็นค่าสังเกตที่ไม่ถูกตัดทิ้ง และ $T_{(9)} \le T_{(10)}$ เป็นค่าสังเกตที่ถูกตัดทิ้ง

(4) เมื่อ $T_{(i)} > T_{(8)}$; $i = 9,10$ แต่ในความเป็นจริงนั้น เราจะไม่ทราบค่าที่แท้จริงของค่าสังเกต ดังนั้นให้ X_i ; $i = 1,2,...,10$ เป็นตัวแปรสุ่มของค่าสังเกตซึ่ง

$$X_i = \begin{cases} T_{(i)} & \text{เมื่อ } i \le 8 \\ T_{(8)} & \text{เมื่อ } i > 8 \end{cases}$$

จะได้ X_1 = -0.935, X_2 = -0.752, X_3 = -0.531, X_4 = -0.159, X_5 = -0.007, X_6 = 0.369, X_7 = 0.636, X_8 = 0.725, X_9 = 0.725, X_{10} = 0.725

3.2.4 จากข้อ 3.2.1 และข้อ 3.2.3 คำนวณค่าประมาณพารามิเตอร์ μ และ σ ด้วยวิธีการประมาณแบบภาวะน่าจะเป็นสูงสุด (MLE) วิธีการประมาณแบบกราฟ (GE) และวิธีการประมาณแบบกราฟด้วยข้อมูลบางส่วน (GEPD)

(1) วิธี MLE โดยการหาค่าของพารามิเตอร์ θ ที่ทำให้ $L(\theta\,;x_1,...,x_n)$ มีค่ามากที่สุด นั่นคือ $\hat{\theta} = \arg\max_{\theta} L(\theta\,;x_1,...,x_n)$ และ $L = \prod_{i=1}^{r} f(x_i)\cdot S(T_{(r)})^{n-r}$

โดยที่ n คือ ขนาดตัวอย่างทั้งหมด ($i = 1,2,...,n$)

r คือ จำนวนค่าสังเกตที่ไม่ถูกตัดทิ้ง ($r \le n$)

$T_{(r)}$ คือ ค่าสังเกตตัวที่ r ที่ไม่ถูกตัดทิ้ง

(2) วิธี GE สามารถทำได้ด้วยการพล็อตกราฟระหว่าง $x_{(i)}$ กับ $F_{0,1}^{-1}(p_{(i)})$ โดยพล็อต $x_{(i)}$ บนแกน y และพล็อต $F_{0,1}^{-1}(p_{(i)})$ บนแกน x เมื่อกำหนดให้ $x_{(i)}$ เป็น sample quantile และ $p_{(i)}$ เป็นลำดับควอนไทล์ของ $x_{(i)}$ ดังนั้น ค่าประมาณของ intercept และ slope จากกราฟ จึงสามารถนำมาใช้ในการประมาณค่า μ และ σ ได้ตามลำดับ

เนื่องจาก $F_{\mu,\sigma}(x) = F_{0,1}\left(\frac{x-\mu}{\sigma}\right)$ ทำให้ $x_p = \mu + \sigma F_{0,1}^{-1}(p)$

โดยที่ $F_{0,1}^{-1}(p)$ คือ ฟังก์ชันควอนไทล์ (quantile function) ของ $F_{0,1}$

x_p คือ ควอนไทล์ตัวที่ p ของ X

โดยจะเรียก μ ว่า location parameter เมื่อ $-\infty < \mu < \infty$

σ ว่า scale parameter เมื่อ $\sigma > 0$

(3) วิธี GEPD สามารถทำได้แบบเดียวกับการประมาณด้วยวิธีแบบใช้กราฟ คือการพล็อตกราฟระหว่าง $x_{(i)}^{*}$ กับ $F_{0,1}^{-1}(p_{(i)}^{*})$ โดยพล็อต $x_{(i)}^{*}$ บนแกน y และพล็อต $F_{0,1}^{-1}(p_{(i)}^{*})$ บนแกน x เมื่อกำหนดให้ $x_{(i)}^{*}$ เป็นข้อมูลบางส่วนที่เหลืออยู่จากวิธี Trimmed q%, K Cluster-Mean,

หรือ Trimmed q% & K Cluster-Mean ซึ่ง $x^*_{(i)}$ นั้นเป็น sample quantile และ $p^*_{(i)}$ เป็นลำดับควอนไทล์ของ $x^*_{(i)}$ ดังนั้น ค่าประมาณของ intercept และ slope จากกราฟ จึงนำมาใช้ในการประมาณค่า μ และ σ ได้เช่นเดียวกัน

3.2.5 ทำซ้ำข้อ 3.2.1-3.2.4 จำนวน 5,000 รอบในแต่ละสถานการณ์

3.2.6 คำนวณค่าความเอนเอียง (Bias) ค่าความแปรปรวน (Variance) และค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) สำหรับค่าประมาณพารามิเตอร์ μ และ σ

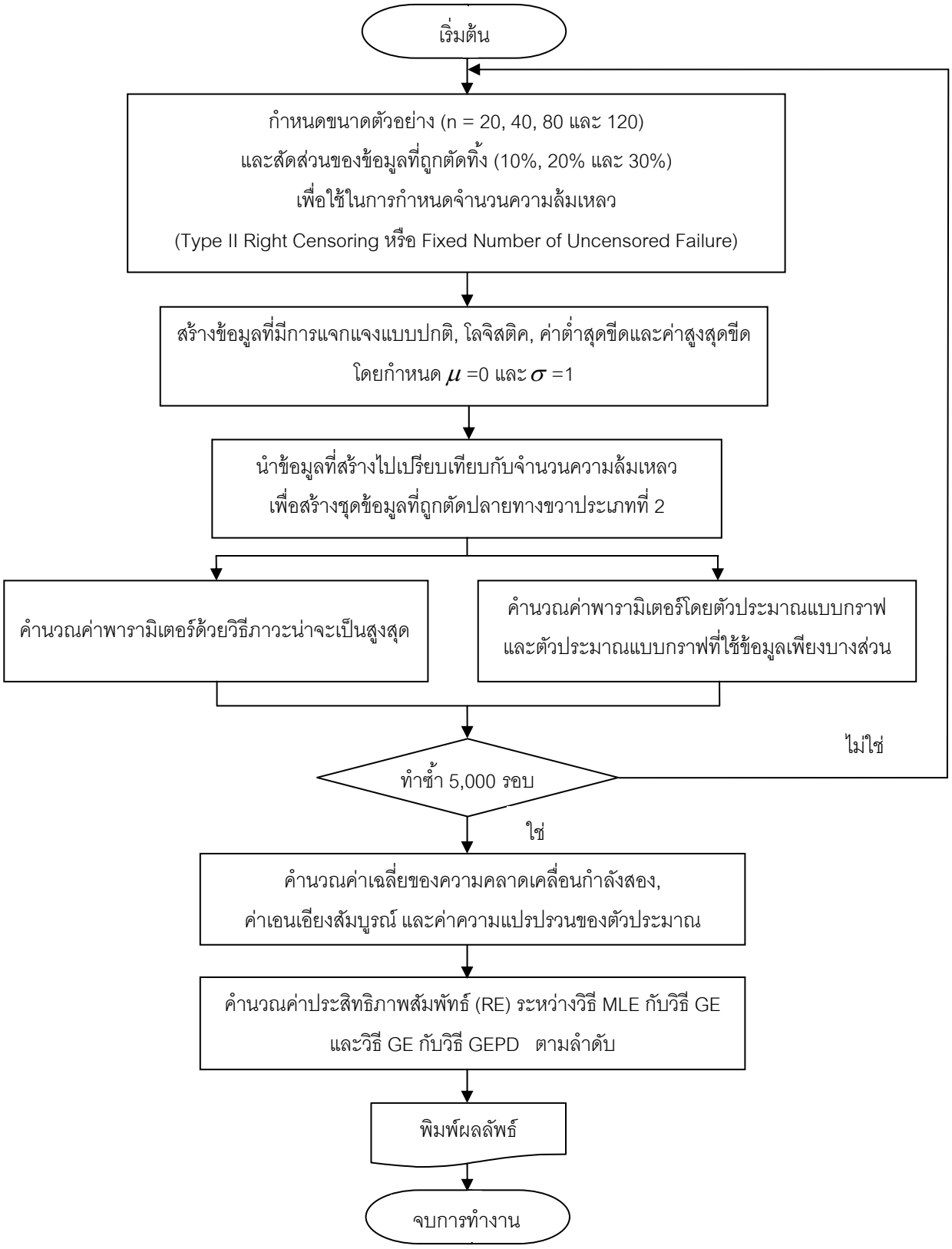
3.2.7 เปรียบเทียบประสิทธิภาพของการประมาณแต่ละวิธี โดยคำนวณค่าประสิทธิภาพสัมพัทธ์ (Relative Efficiency; RE) ของค่าประมาณพารามิเตอร์ μ และ σ โดยแบ่งออกเป็น (1) ค่า RE ระหว่างการประมาณแบบภาวะน่าจะเป็นสูงสุด กับการประมาณแบบใช้กราฟ และ(2) ค่า RE ระหว่างการประมาณแบบใช้กราฟ กับการประมาณแบบใช้กราฟที่ใช้ข้อมูลเพียงบางส่วน

3.2.8 สรุปผลการวิจัยในแต่ละสถานการณ์

3.3 ขั้นตอนการทำงานของโปรแกรม

โปรแกรมที่ใช้ในการศึกษาวิจัยครั้งนี้ เขียนด้วยโปรแกรม R เวอร์ชัน 2.9.2 ซึ่งในแต่ละสถานการณ์ของการทดลองจะกระทำซ้ำ 5,000 รอบ สำหรับขั้นตอนการทำงานของโปรแกรมที่ใช้ในการประมาณค่าพารามิเตอร์ μ และ σ ด้วยวิธีภาวะน่าจะเป็นสูงสุด (MLE), วิธีประมาณแบบกราฟ (GE) และวิธีประมาณแบบกราฟด้วยข้อมูลบางส่วน (GEPD) สามารถสรุปเป็นรายละเอียดของผังงานแสดงการทำงานของโปรแกรมได้ดังนี้

ภาพที่ 3.1 แสดงแผนผังขั้นตอนการวิจัย



บทที่ 4
ผลการวิเคราะห์ข้อมูล

จากการศึกษาการประมาณค่าพารามิเตอร์ด้วยวิธีภาวะน่าจะเป็นสูงสุด วิธีการประมาณแบบกราฟ และวิธีการประมาณแบบกราฟด้วยข้อมูลบางส่วน สำหรับข้อมูลที่ถูกตัดปลายทางขวาประเภทที่ 2 โดยศึกษาเมื่อข้อมูลมี 4 การแจกแจง คือ การแจกแจงปกติ การแจกแจงโลจิสติค การแจกแจงค่าต่ำสุดขีด และการแจกแจงค่าสูงสุดขีด ขนาดตัวอย่างเท่ากับ 20, 40, 80 และ 120 และสัดส่วนของข้อมูลที่ถูกตัดทิ้งเท่ากับ 10%, 20% และ 30% ของขนาดตัวอย่าง จากการศึกษาในครั้งนี้เกณฑ์ในการพิจารณาเปรียบเทียบคือ ค่าประสิทธิภาพสัมพัทธ์ (Relative Efficiency; RE) นอกจากนี้แล้วยังได้นำค่าความเอนเอียงสัมบูรณ์ ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยมาประกอบการพิจารณาอีกด้วย โดยในการสรุปผลการวิจัยแสดงผลในรูปของตาราง และเพื่อความสะดวกในการสรุปผลจึงกำหนดสัญลักษณ์แทนความหมายต่างๆ ในตาราง ดังต่อไปนี้

วิธีการประมาณค่าพารามิเตอร์

MLE หมายถึง วิธีภาวะน่าจะเป็นสูงสุด

GE หมายถึง วิธีการประมาณแบบกราฟ

GEPD หมายถึง วิธีการประมาณแบบกราฟด้วยข้อมูลบางส่วน

RE1 หมายถึง ค่าประสิทธิภาพสัมพัทธ์ของวิธีที่ใช้ในการประมาณค่าพารามิเตอร์ โดยคิดจากสัดส่วนของ Variances

RE2 หมายถึง ค่าประสิทธิภาพสัมพัทธ์ของวิธีที่ใช้ในการประมาณค่าพารามิเตอร์ โดยคิดจากสัดส่วนของ MSEs

|Bias| หมายถึง ค่าความเอนเอียงสัมบูรณ์

n หมายถึง ขนาดตัวอย่างทั้งหมดทั้งที่ถูกตัดทิ้งและไม่ถูกตัดทิ้ง

p หมายถึง สัดส่วนของข้อมูลที่ถูกตัดทิ้ง

ในงานวิจัยครั้งนี้ แบ่งการศึกษาออกเป็น 2 ส่วนด้วยกัน คือ

ส่วนที่ 1 เป็นการเปรียบเทียบประสิทธิภาพของวิธีการประมาณค่าพารามิเตอร์ μ และ σ ด้วยวิธีการประมาณ MLE กับวิธีการประมาณ GE ภายใต้การแจกแจงแบบค่าต่ำสุดขีด ค่าสูงสุดขีด โลจิสติค และปกติ ของข้อมูลที่ถูกตัดปลายทางขวาประเภทที่ 2

สำหรับการคำนวณค่าประสิทธิภาพสัมพัทธ์ (RE) ที่แสดงในตารางนั้น มี 2 รูปแบบ คือ $\mathrm{RE1} = \frac{\mathrm{Var}_{\mathrm{MLE}}}{\mathrm{Var}_{\mathrm{GE}}}$ และ $\mathrm{RE2} = \frac{\mathrm{MSE}_{\mathrm{MLE}}}{\mathrm{MSE}_{\mathrm{GE}}}$ โดย RE1 และ RE2 คือ ค่าประสิทธิภาพสัมพัทธ์ของตัวประมาณ GE เทียบกับ MLE ที่คิดจากค่าสัดส่วนของ Variances และสัดส่วนของ MSEs ตามลำดับ ซึ่งถ้า RE มีค่ามากกว่า 1 หมายความว่า ตัวประมาณ GE มีประสิทธิภาพมากกว่า MLE, ถ้า RE มีค่าน้อยกว่า 1 หมายความว่า ตัวประมาณ GE มีประสิทธิภาพน้อยกว่า MLE, และถ้า RE มีค่าเท่ากับ 1 หมายความว่า ตัวประมาณ GE และ MLE มีประสิทธิภาพเท่ากัน

ส่วนที่ 2 เป็นการศึกษาว่าถ้าเลือกใช้ข้อมูลเพียงบางส่วนจากข้อมูลทั้งหมดด้วยวิธีการประมาณGEPDแล้ว ตัวประมาณพารามิเตอร์ μ และ σ ที่ได้จะมีประสิทธิภาพดีกว่าหรือไม่

สำหรับการคำนวณค่าประสิทธิภาพสัมพัทธ์ (RE) ที่แสดงในตารางนั้น มี 2 รูปแบบ คือ $\mathrm{RE1} = \frac{\mathrm{Var}_{\mathrm{GE}}}{\mathrm{Var}_{\mathrm{GEPD}}}$ และ $\mathrm{RE2} = \frac{\mathrm{MSE}_{\mathrm{GE}}}{\mathrm{MSE}_{\mathrm{GEPD}}}$ โดย RE1 และ RE2 คือ ค่าประสิทธิภาพสัมพัทธ์ของตัวประมาณ GEPD เทียบกับ GE ที่คิดจากค่าสัดส่วนของ Variances และสัดส่วนของ MSEs ตามลำดับ ซึ่งถ้า RE มีค่ามากกว่า 1 หมายความว่า ตัวประมาณ GEPD มีประสิทธิภาพมากกว่า GE, ถ้า RE มีค่าน้อยกว่า 1 หมายความว่า ตัวประมาณ GEPD มีประสิทธิภาพน้อยกว่า GE, และถ้า RE มีค่าเท่ากับ 1 หมายความว่า ตัวประมาณ GEPD และ GE มีประสิทธิภาพเท่ากัน

4.1 สรุปผลการศึกษาเปรียบเทียบการประมาณค่าพารามิเตอร์ทั้ง 3 วิธี

จากผลการศึกษาพบว่า การประมาณค่าพารามิเตอร์ μ และ σ ด้วยวิธีภาวะน่าจะเป็นสูงสุด (MLE) เป็นวิธีที่มีประสิทธิภาพมากที่สุด เมื่อเทียบกับวิธีการประมาณแบบกราฟ และการประมาณแบบกราฟด้วยข้อมูลบางส่วน ภายใต้ขอบเขตของการศึกษานี้ให้ความสนใจการประมาณพารามิเตอร์โดยใช้กราฟแต่เลือกใช้ข้อมูลเพียงบางส่วนเท่านั้น ซึ่งในกรณีส่วนใหญ่พบว่า การประมาณแบบกราฟด้วยข้อมูลเพียงบางส่วนนั้นจะมีประสิทธิภาพมากกว่าการประมาณแบบกราฟที่ใช้ข้อมูลทุกส่วน โดยสามารถสรุปผลภายใต้การแจกแจงของข้อมูล ได้ดังต่อไปนี้

<u>กรณีข้อมูลมีการแจกแจงแบบ SEV (0, 1)</u>

- การประมาณค่าพารามิเตอร์ μ

เมื่อ $p = 0$; พบว่า วิธี GEPD แบบ K Cluster-Mean จะมีประสิทธิภาพมากที่สุด โดยแบบ 4 Cluster-Mean มีประสิทธิภาพมากกว่าแบบ 8 Cluster-Mean

เมื่อ $p \neq 0$; พบว่า วิธี GEPD แบบ Trimmed 10% & K Cluster-Mean จะมีประสิทธิภาพมากที่สุด โดยแบบ Trimmed 10% & 8 Cluster-Mean มีประสิทธิภาพมากกว่าแบบ Trimmed 10% & 4 Cluster-Mean (ยกเว้น $p = 0.2$)

- การประมาณค่าพารามิเตอร์ σ

เมื่อ $p = 0$; พบว่า วิธี GEPD แบบ 4 Cluster-Mean จะมีประสิทธิภาพมากที่สุด รองลงมาคือ แบบ Trimmed 5% & 4 Cluster-Mean

เมื่อ $p \neq 0$; พบว่า วิธี GEPD แบบ Trimmed 5% & 4 Cluster-Mean จะมีประสิทธิภาพมากที่สุด

<u>กรณีข้อมูลมีการแจกแจงแบบ LEV (0, 1)</u>

- การประมาณค่าพารามิเตอร์ μ

เมื่อ $p = 0$ หรือ 0.1; พบว่า วิธี GEPD แบบ K Cluster-Mean จะมีประสิทธิภาพมากที่สุด โดยแบบ 4 Cluster-Mean มีประสิทธิภาพมากกว่าแบบ 8 Cluster-Mean

เมื่อ $p = 0.2$ หรือ 0.3; ในกรณีส่วนใหญ่พบว่า วิธี GEPD แบบ Trimmed 5% & 8 Cluster-Mean จะมีประสิทธิภาพมากที่สุด

- การประมาณค่าพารามิเตอร์ σ

เมื่อ $p = 0$; พบว่า วิธี GEPD แบบ 4 Cluster-Mean จะมีประสิทธิภาพมากที่สุด รองลงมาคือ แบบ Trimmed 5% & 4 Cluster-Mean

เมื่อ $p \neq 0$; พบว่า วิธี GE จะมีประสิทธิภาพมากกว่าวิธี GEPD ในทุกๆ กรณี

<u>กรณีข้อมูลมีการแจกแจงแบบ NOR (0, 1)</u>

- การประมาณค่าพารามิเตอร์ μ

เมื่อ p = 0; พบว่า วิธี GE, วิธี GEPD แบบ 4 Cluster-Mean และแบบ 8 Cluster-Mean ให้ค่าประมาณพารามิเตอร์ที่เท่ากัน และมีประสิทธิภาพมากที่สุด

เมื่อ p = 0.1; ในกรณีส่วนใหญ่พบว่า วิธี GEPD แบบ Trimmed 5% & 8 Cluster-Mean จะมีประสิทธิภาพมากที่สุด

เมื่อ p = 0.2; ในกรณีส่วนใหญ่พบว่า วิธี GEPD แบบ Trimmed 5% & 4 Cluster-Mean จะมีประสิทธิภาพมากที่สุด

เมื่อ p = 0.3; ในกรณีส่วนใหญ่พบว่า วิธี GEPD แบบ Trimmed 10% & 8 Cluster-Mean จะมีประสิทธิภาพมากที่สุด

- การประมาณค่าพารามิเตอร์ σ ในกรณีส่วนใหญ่พบว่า วิธี GE จะมีประสิทธิภาพมากกว่าวิธี GEPD

<u>กรณีข้อมูลมีการแจกแจงแบบ LOG (0, 1)</u>

- การประมาณค่าพารามิเตอร์ μ

เมื่อ p = 0; พบว่า วิธี GEPD แบบ Trimmed 10% จะมีประสิทธิภาพมากที่สุด รองลงมาคือ แบบ Trimmed 10% & 4 Cluster-Mean

เมื่อ p = 0.1 หรือ 0.2; พบว่า วิธี GEPD แบบ Trimmed q% จะมีประสิทธิภาพมากที่สุด โดยแบบ Trimmed 10% มีประสิทธิภาพมากกว่าแบบ Trimmed 5% (ยกเว้น p = 0.2 และ $n \geq 80$)

เมื่อ p = 0.3; พบว่า วิธี GEPD แบบ Trimmed 10% & 8 Cluster-Mean จะมีประสิทธิภาพมากที่สุด

- การประมาณค่าพารามิเตอร์ σ

เมื่อ p = 0; พบว่า วิธี GEPD แบบ K Cluster-Mean จะมีประสิทธิภาพมากที่สุด โดยแบบ 4 Cluster-Mean มีประสิทธิภาพมากกว่าแบบ 8 Cluster-Mean

เมื่อ p = 0.1; ในกรณีส่วนใหญ่พบว่า วิธี GEPD แบบ Trimmed 5% & 8 Cluster-Mean จะมีประสิทธิภาพมากที่สุด

เมื่อ $p \geq$ 0.2; ในกรณีส่วนใหญ่พบว่า วิธี GEPD แบบ Trimmed 5% & 4 Cluster-Mean จะมีประสิทธิภาพมากที่สุด

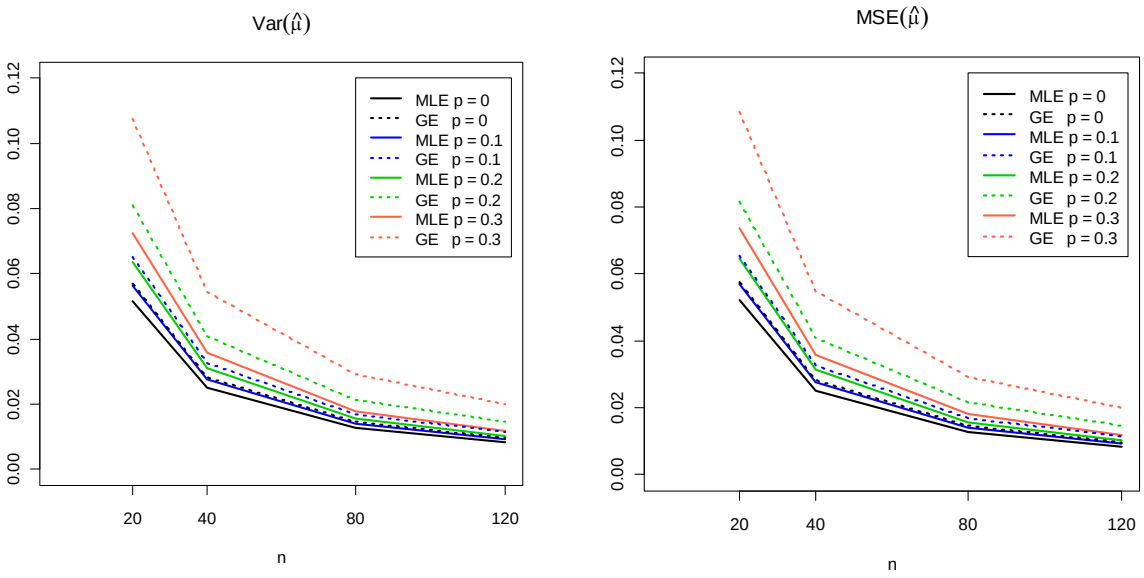
4.1.1 ผลการศึกษาเมื่อข้อมูลแจกแจงแบบค่าต่ำสุดขีด

ส่วนที่ 1 การเปรียบเทียบประสิทธิภาพของวิธีการประมาณค่าพารามิเตอร์แบบภาวะน่าจะเป็นสูงสุด (MLE) กับแบบใช้กราฟ (GE)

ตารางที่ 4.1 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงแบบค่าต่ำสุดขีด ที่พารามิเตอร์ μ = 0 และ σ = 1 โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา

Censoring	Sample	($\hat{\mu}$) → MLE			($\hat{\mu}$) → GE				
Proportion p	Size n	Bias	Var	MSE	Bias	Var	MSE	RE1	RE2
0	20	-0.02458	0.05169	0.05229	-0.02488	0.05712	0.05772	0.905	0.906
	40	-0.01002	0.02503	0.02513	-0.01116	0.02827	0.02838	0.886	0.885
	80	-0.00369	0.01273	0.01274	-0.00483	0.01459	0.01461	0.872	0.872
	120	-0.00157	0.00836	0.00836	-0.00304	0.00956	0.00956	0.875	0.875
0.1	20	-0.02551	0.05641	0.05705	-0.02220	0.06517	0.06565	0.866	0.869
	40	-0.01007	0.02757	0.02767	-0.01029	0.03264	0.03274	0.845	0.845
	80	-0.00402	0.01403	0.01405	-0.00468	0.01694	0.01696	0.828	0.828
	120	-0.00275	0.00926	0.00926	-0.00340	0.01136	0.01137	0.815	0.815
0.2	20	-0.02879	0.06357	0.06439	-0.02562	0.08086	0.08150	0.786	0.790
	40	-0.01400	0.03114	0.03133	-0.01273	0.04086	0.04101	0.762	0.764
	80	-0.00474	0.01573	0.01575	-0.00652	0.02140	0.02144	0.735	0.734
	120	-0.00263	0.01036	0.01036	-0.00482	0.01460	0.01462	0.710	0.709
0.3	20	-0.03271	0.07256	0.07361	-0.03113	0.10744	0.10839	0.675	0.679
	40	-0.01334	0.03574	0.03591	-0.01597	0.05459	0.05483	0.655	0.655
	80	-0.00522	0.01794	0.01797	-0.00913	0.02904	0.02912	0.618	0.617
	120	-0.00336	0.01189	0.01190	-0.00702	0.02001	0.02006	0.594	0.593

ภาพที่ 4.1 แสดงการเปรียบเทียบค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงแบบค่าต่ำสุดขีด ที่พารามิเตอร์ μ = 0 และ σ = 1 ด้วยวิธี GE กับวิธี MLE

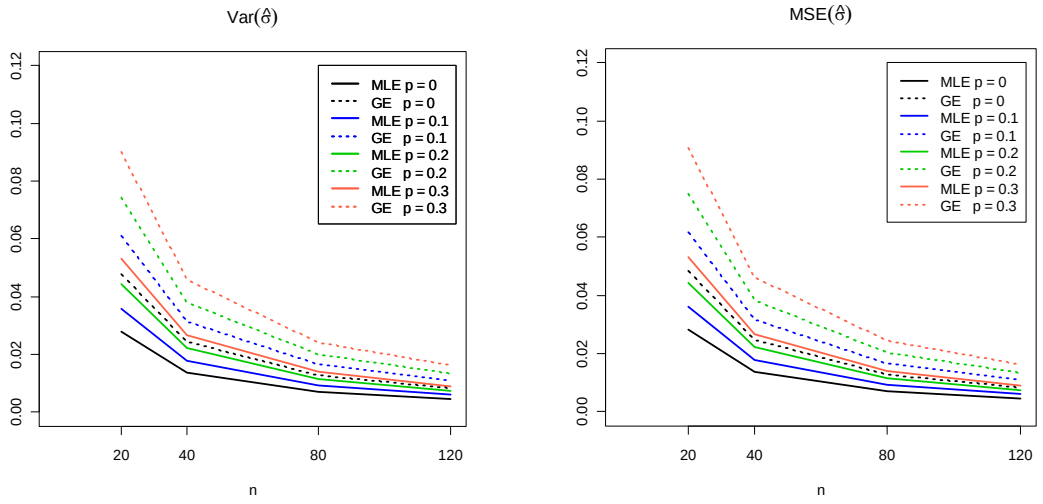


จากตารางและภาพที่ 4.1 พบว่า การประมาณค่าพารามิเตอร์ μ ด้วยวิธี MLE มีประสิทธิภาพมากกว่าวิธี GE ในทุกกรณี โดยพิจารณาจากค่า |Biases| และค่า Variances ของค่าประมาณพารามิเตอร์ μ ซึ่งวิธี GE มีค่าสูงกว่าวิธี MLE เป็นส่วนใหญ่ ส่งผลให้ค่า MSEs มีค่าสูงขึ้นตามไปด้วย ถ้ากำหนดให้สัดส่วนของการตัดปลาย (p) มีขนาดคงที่พบว่า ค่า RE1 และ RE2 จะมีค่าลดลง เมื่อขนาดตัวอย่าง (n) เพิ่มขึ้น เนื่องจาก MLE จะใช้ได้ดีกับข้อมูลที่มีขนาดตัวอย่าง (n) มาก แต่ถ้ากำหนดให้ขนาดตัวอย่าง (n) คงที่พบว่า ค่า RE1 และ RE2 จะมีค่าลดลง เมื่อสัดส่วนของการตัดปลาย (p) เพิ่มขึ้น

ตารางที่ 4.2 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงแบบค่าต่ำสุดขีด ที่พารามิเตอร์ $\mu = 0$ และ $\sigma = 1$ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา

Censoring	Sample	$(\hat{\sigma}) \rightarrow$ MLE			$(\hat{\sigma}) \rightarrow$ GE				
Proportion p	Size n	Bias	Var	MSE	Bias	Var	MSE	RE1	RE2
0	20	-0.01548	0.02798	0.02822	-0.02433	0.04783	0.04842	0.585	0.583
	40	-0.00895	0.01365	0.01372	-0.01557	0.02443	0.02467	0.559	0.556
	80	-0.00472	0.00692	0.00694	-0.01005	0.01269	0.01279	0.545	0.542
	120	-0.00244	0.00450	0.00451	-0.00707	0.00835	0.00840	0.539	0.537
0.1	20	-0.01021	0.03589	0.03599	-0.02231	0.06117	0.06166	0.587	0.584
	40	-0.00577	0.01781	0.01784	-0.01483	0.03141	0.03162	0.567	0.564
	80	-0.00389	0.00918	0.00919	-0.00986	0.01646	0.01655	0.558	0.555
	120	-0.00302	0.00599	0.00599	-0.00724	0.01088	0.01093	0.550	0.548
0.2	20	-0.00744	0.04437	0.04441	-0.02436	0.07442	0.07500	0.596	0.592
	40	-0.00718	0.02214	0.02219	-0.01627	0.03814	0.03839	0.581	0.578
	80	-0.00346	0.01150	0.01151	-0.01091	0.02007	0.02019	0.573	0.570
	120	-0.00205	0.00747	0.00748	-0.00804	0.01335	0.01341	0.560	0.558
0.3	20	-0.00552	0.05306	0.05308	-0.02735	0.08998	0.09071	0.590	0.585
	40	-0.00401	0.02672	0.02674	-0.01795	0.04589	0.04621	0.582	0.579
	80	-0.00298	0.01394	0.01395	-0.01224	0.02424	0.02438	0.575	0.572
	120	-0.00193	0.00890	0.00890	-0.00916	0.01621	0.01629	0.549	0.546

ภาพที่ 4.2 แสดงการเปรียบเทียบค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงแบบค่าต่ำสุดขีด ที่พารามิเตอร์ $\mu = 0$ และ $\sigma = 1$ ด้วยวิธี GE กับวิธี MLE



จากตารางและภาพที่ 4.2 พบว่า การประมาณค่าพารามิเตอร์ σ ด้วยวิธี MLE มีประสิทธิภาพมากกว่าวิธี GE ในทุกกรณี โดยพิจารณาจากค่า |Biases| และค่า Variances ของค่าประมาณพารามิเตอร์ σ ซึ่งวิธี GE มีค่าสูงกว่าวิธี MLE ในทุกกรณี ส่งผลให้ค่า MSEs มีค่าสูงขึ้นตามไปด้วย ถ้ากำหนดให้สัดส่วนของการตัดปลาย (p) มีขนาดคงที่พบว่า ค่า RE1 และ RE2 จะมีค่าลดลง เมื่อขนาดตัวอย่าง (n) เพิ่มขึ้น เนื่องจาก MLE จะใช้ได้ดีกับข้อมูลที่มีขนาดตัวอย่าง (n) มาก แต่ถ้ากำหนดให้ขนาดตัวอย่าง (n) คงที่ พบว่า ค่า RE1 และRE2 จะมีค่าเพิ่มขึ้น เมื่อสัดส่วนของการตัดปลาย (p) เพิ่มขึ้น (ยกเว้นกรณี p =0.3)

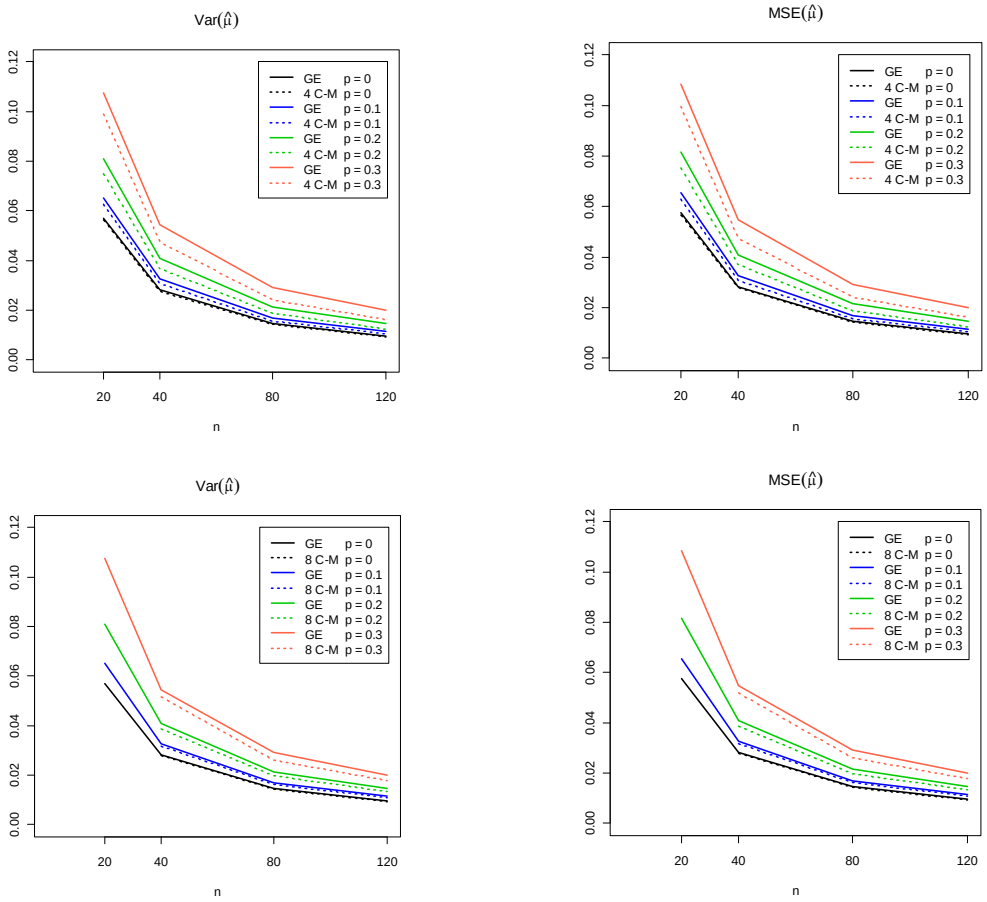
จากตารางที่ 4.1 และ 4.2 เมื่อกำหนดให้ขนาดตัวอย่าง (n) คงที่และสัดส่วนของการตัดปลาย (p) เพิ่มขึ้นพบว่า ค่า RE1 และRE2 ที่ได้ไม่สอดคล้องกัน โดยน่าจะเป็นผลมาจาก (1) ขนาดตัวอย่างที่ลดลง (2) สัดส่วนของการตัดปลายที่เพิ่มขึ้น และ (3) Outlier ที่ถูกตัดออกไป ซึ่งทั้ง 3 ปัจจัยนี้น่าจะส่งผลต่อการเพิ่มขึ้นหรือลดลงของประสิทธิภาพการประมาณ

ส่วนที่ 2 การเปรียบเทียบประสิทธิภาพของวิธีการประมาณค่าพารามิเตอร์แบบใช้กราฟ (GE) กับแบบใช้กราฟที่ใช้ข้อมูลเพียงบางส่วน (GEPD)

ตารางที่ 4.3 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงแบบค่าต่ำสุดขีด ที่พารามิเตอร์ μ = 0 และ σ = 1 โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— 4 Cluster-Mean และ 8 Cluster-Mean ตามลำดับ

p	n	($\hat{\mu}$) → GE			($\hat{\mu}$) → 4 Cluster-Mean					($\hat{\mu}$) → 8 Cluster-Mean				
		Bias	Var	MSE	Bias	Var	MSE	RE1	RE2	Bias	Var	MSE	RE1	RE2
0	20	-0.02488	0.05712	0.05772	-0.02205	0.05629	0.05677	1.015	1.017	NA	NA	NA	NA	NA
	40	-0.01116	0.02827	0.02838	-0.00882	0.02771	0.02778	1.020	1.022	-0.00970	0.02792	0.02801	1.012	1.013
	80	-0.00483	0.01459	0.01461	-0.00301	0.01419	0.01419	1.028	1.029	-0.00358	0.01429	0.01430	1.021	1.022
	120	-0.00304	0.00956	0.00956	-0.00155	0.00925	0.00925	1.033	1.034	-0.00194	0.00933	0.00933	1.024	1.025
0.1	20	-0.02220	0.06517	0.06565	-0.01863	0.06268	0.06301	1.040	1.042	NA	NA	NA	NA	NA
	40	-0.01029	0.03264	0.03274	-0.00683	0.03076	0.03080	1.061	1.063	-0.00845	0.03170	0.03176	1.030	1.031
	80	-0.00468	0.01694	0.01696	-0.00193	0.01566	0.01566	1.082	1.083	-0.00282	0.01610	0.01611	1.052	1.053
	120	-0.00340	0.01136	0.01137	-0.00117	0.01039	0.01039	1.093	1.094	-0.00176	0.01072	0.01072	1.059	1.060
0.2	20	-0.02562	0.08086	0.08150	-0.02020	0.07505	0.07544	1.077	1.080	NA	NA	NA	NA	NA
	40	-0.01273	0.04086	0.04101	-0.00796	0.03689	0.03694	1.108	1.110	-0.00995	0.03869	0.03879	1.056	1.057
	80	-0.00652	0.02140	0.02144	-0.00274	0.01872	0.01873	1.143	1.145	-0.00397	0.01974	0.01975	1.085	1.086
	120	-0.00482	0.01460	0.01462	-0.00174	0.01258	0.01258	1.160	1.162	-0.00251	0.01329	0.01329	1.099	1.100
0.3	20	-0.03113	0.10744	0.10839	-0.02603	0.09895	0.09961	1.086	1.088	NA	NA	NA	NA	NA
	40	-0.01597	0.05459	0.05483	-0.00989	0.04751	0.04760	1.149	1.152	-0.01301	0.05167	0.05183	1.057	1.058
	80	-0.00913	0.02904	0.02912	-0.00419	0.02420	0.02421	1.200	1.203	-0.00579	0.02611	0.02614	1.112	1.114
	120	-0.00702	0.02001	0.02006	-0.00286	0.01637	0.01637	1.223	1.225	-0.00405	0.01781	0.01782	1.124	1.125

ภาพที่ 4.3 แสดงการเปรียบเทียบค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงแบบค่าต่ำสุดขีด ที่พารามิเตอร์ μ = 0 และ σ = 1 ด้วยวิธี GE กับวิธี GEPD —— 4 Cluster-Mean และ 8 Cluster-Mean ตามลำดับ

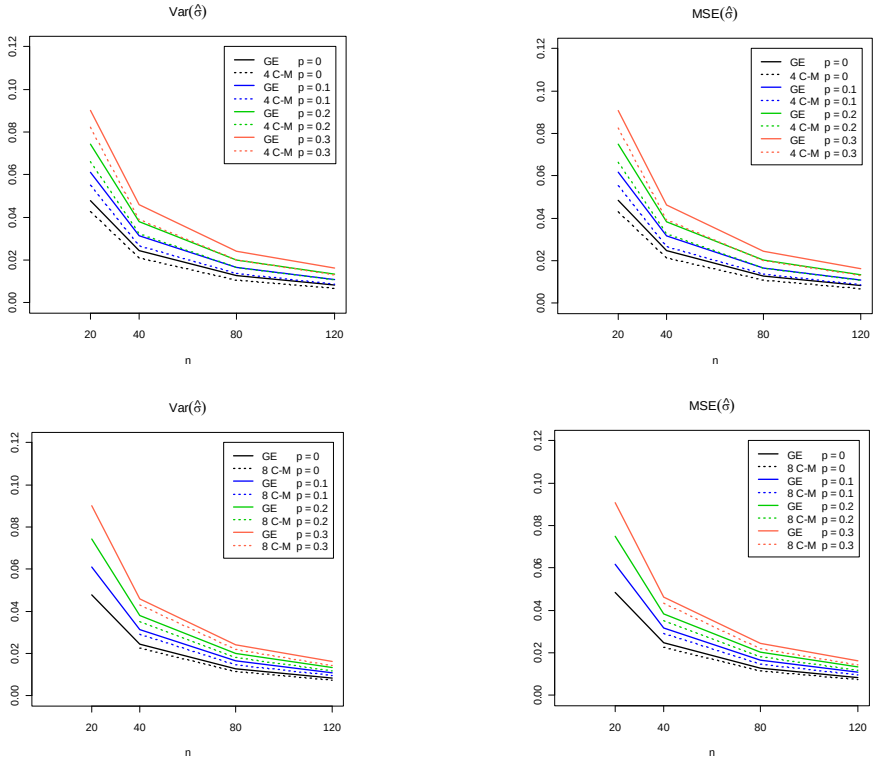


จากตารางและภาพที่ 4.3 พบว่า การประมาณค่าพารามิเตอร์ μ ด้วยวิธี GEPD แบบ 4 Cluster-Mean มีประสิทธิภาพมากกว่าวิธี GE ในทุกกรณี เพราะวิธี K Cluster-Mean นั้นเป็นการปรับข้อมูลหลายๆ จุดที่กวัดแกว่งไปมาให้เหลือเพียงจุดเดียวในแต่ละกลุ่มด้วยการหาค่าเฉลี่ยภายในกลุ่ม โดยพิจารณาจากค่า |Biases| และค่า Variances ของค่าประมาณพารามิเตอร์ μ ซึ่งวิธี GE มีค่าสูงกว่า 4 Cluster-Mean ในทุกกรณี ส่งผลให้ค่า MSEs มีค่าสูงขึ้นตามไปด้วย ถ้ากำหนดให้สัดส่วนของการตัดปลาย (p) มีขนาดคงที่พบว่า ค่า RE1 และ RE2 จะมีค่าเพิ่มขึ้น เมื่อขนาดตัวอย่าง (n) เพิ่มขึ้น เนื่องจากอาจเป็นไปได้ว่าขนาดตัวอย่าง (n) มากทำให้การประมาณค่าใกล้เคียงกับประชากรมากยิ่งขึ้น แต่ถ้ากำหนดให้ขนาดตัวอย่าง (n) คงที่พบว่าค่าRE1 และRE2จะมีค่าเพิ่มขึ้น เมื่อสัดส่วนของการตัดปลาย(p) เพิ่มขึ้น นอกจากนี้ยังพบว่าวิธีประมาณแบบ 4 Cluster-Mean มีประสิทธิภาพมากกว่าวิธีประมาณแบบ 8 Cluster-Mean ในทุกกรณีอีกด้วย

ตารางที่ 4.4 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงแบบค่าต่ำสุดขีด ที่พารามิเตอร์ $\mu = 0$ และ $\sigma = 1$ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— 4 Cluster-Mean และ 8 Cluster-Mean ตามลำดับ

p	n	($\hat{\sigma}$) → GE			($\hat{\sigma}$) → 4 Cluster-Mean					($\hat{\sigma}$) → 8 Cluster-Mean				
		Bias	Var	MSE	Bias	Var	MSE	RE1	RE2	Bias	Var	MSE	RE1	RE2
0	20	-0.02433	0.04783	0.04842	-0.01930	0.04266	0.04302	1.121	1.125	NA	NA	NA	NA	NA
	40	-0.01557	0.02443	0.02467	-0.01146	0.02102	0.02115	1.162	1.166	-0.01301	0.02248	0.02264	1.087	1.089
	80	-0.01005	0.01269	0.01279	-0.00688	0.01067	0.01071	1.190	1.194	-0.00787	0.01143	0.01149	1.110	1.113
	120	-0.00707	0.00835	0.00840	-0.00447	0.00688	0.00690	1.213	1.216	-0.00514	0.00741	0.00744	1.127	1.129
0.1	20	-0.02231	0.06117	0.06166	-0.01756	0.05504	0.05534	1.111	1.114	NA	NA	NA	NA	NA
	40	-0.01483	0.03141	0.03162	-0.01028	0.02668	0.02678	1.177	1.181	-0.01241	0.02914	0.02929	1.078	1.079
	80	-0.00986	0.01646	0.01655	-0.00627	0.01362	0.01366	1.208	1.212	-0.00743	0.01475	0.01481	1.116	1.118
	120	-0.00724	0.01088	0.01093	-0.00433	0.00879	0.00880	1.238	1.242	-0.00510	0.00962	0.00964	1.131	1.133
0.2	20	-0.02436	0.07442	0.07500	-0.01852	0.06599	0.06632	1.128	1.131	NA	NA	NA	NA	NA
	40	-0.01627	0.03814	0.03839	-0.01117	0.03238	0.03250	1.178	1.181	-0.01329	0.03507	0.03524	1.088	1.090
	80	-0.01091	0.02007	0.02019	-0.00689	0.01659	0.01663	1.210	1.214	-0.00819	0.01802	0.01808	1.114	1.116
	120	-0.00804	0.01335	0.01341	-0.00478	0.01075	0.01077	1.242	1.245	-0.00560	0.01175	0.01178	1.136	1.138
0.3	20	-0.02735	0.08998	0.09071	-0.02273	0.08209	0.08259	1.096	1.098	NA	NA	NA	NA	NA
	40	-0.01795	0.04589	0.04621	-0.01250	0.03908	0.03923	1.174	1.178	-0.01530	0.04308	0.04330	1.065	1.067
	80	-0.01224	0.02424	0.02438	-0.00784	0.02005	0.02010	1.209	1.213	-0.00927	0.02180	0.02188	1.112	1.114
	120	-0.00916	0.01621	0.01629	-0.00547	0.01308	0.01310	1.240	1.243	-0.00653	0.01441	0.01445	1.125	1.128

ภาพที่ 4.4 แสดงการเปรียบเทียบค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงแบบค่าต่ำสุดขีด ที่พารามิเตอร์ $\mu =$ 0 และ $\sigma = 1$ ด้วยวิธี GE กับวิธี GEPD —— 4 Cluster-Mean และ 8 Cluster-Mean ตามลำดับ

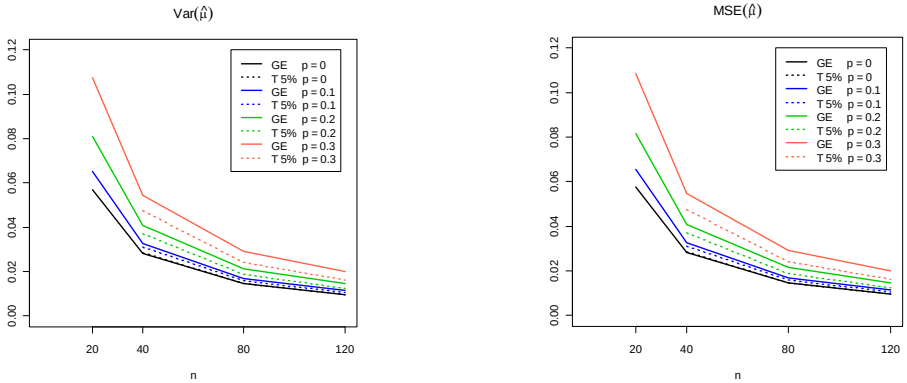


จากตารางและภาพที่ 4.4 พบว่า การประมาณค่าพารามิเตอร์ σ ด้วยวิธี GEPD แบบ K Cluster-Mean มีประสิทธิภาพมากกว่าวิธี GE ถ้ากำหนดให้สัดส่วนของการตัดปลาย (p) มีขนาดคงที่ พบว่าค่า RE1 และ RE2 จะมีค่าเพิ่มขึ้น เมื่อขนาดตัวอย่าง (n) เพิ่มขึ้น แต่ถ้ากำหนดให้ขนาดตัวอย่าง (n) คงที่ พบว่าผลค่อนข้างไม่ชัดเจน นอกจากนี้ยังพบว่าแบบ 4 Cluster-Mean มีประสิทธิภาพมากกว่าแบบ 8 Cluster-Mean ในทุกกรณี

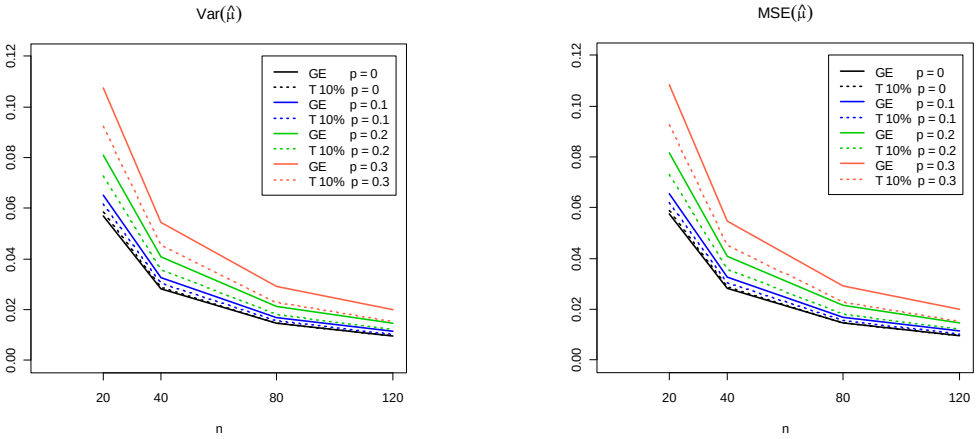
ตารางที่ 4.5 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงแบบค่าต่ำสุดขีด ที่พารามิเตอร์ μ = 0 และ σ = 1 โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— Trimmed 5% และ Trimmed 10% ตามลำดับ

p	n	($\hat{\mu}$) → GE			($\hat{\mu}$) → Trimmed 5%					($\hat{\mu}$) → Trimmed 10%				
		Bias	Var	MSE	Bias	Var	MSE	RE1	RE2	Bias	Var	MSE	RE1	RE2
0	20	-0.02488	0.05712	0.05772	NA	NA	NA	NA	NA	-0.01919	0.05846	0.05882	0.977	0.981
	40	-0.01116	0.02827	0.02838	-0.00794	0.02842	0.02848	0.995	0.997	-0.00713	0.02888	0.02892	0.979	0.981
	80	-0.00483	0.01459	0.01461	-0.00256	0.01454	0.01454	1.004	1.005	-0.00217	0.01473	0.01473	0.991	0.992
	120	-0.00304	0.00956	0.00956	-0.00137	0.00951	0.00951	1.004	1.005	-0.00119	0.00968	0.00968	0.987	0.988
0.1	20	-0.02220	0.06517	0.06565	NA	NA	NA	NA	NA	-0.01798	0.06179	0.06210	1.055	1.057
	40	-0.01029	0.03264	0.03274	-0.00695	0.03103	0.03107	1.052	1.054	-0.00661	0.03051	0.03055	1.070	1.072
	80	-0.00468	0.01694	0.01696	-0.00197	0.01578	0.01578	1.074	1.075	-0.00180	0.01551	0.01551	1.092	1.093
	120	-0.00340	0.01136	0.01137	-0.00114	0.01046	0.01046	1.086	1.087	-0.00099	0.01028	0.01028	1.105	1.106
0.2	20	-0.02562	0.08086	0.08150	NA	NA	NA	NA	NA	-0.01773	0.07289	0.07319	1.109	1.113
	40	-0.01273	0.04086	0.04101	-0.00695	0.03716	0.03720	1.099	1.102	-0.00624	0.03592	0.03595	1.137	1.141
	80	-0.00652	0.02140	0.02144	-0.00198	0.01878	0.01878	1.140	1.142	-0.00162	0.01812	0.01811	1.182	1.184
	120	-0.00482	0.01460	0.01462	-0.00104	0.01259	0.01259	1.160	1.161	-0.00074	0.01216	0.01216	1.201	1.203
0.3	20	-0.03113	0.10744	0.10839	NA	NA	NA	NA	NA	-0.01883	0.09225	0.09259	1.165	1.171
	40	-0.01597	0.05459	0.05483	-0.00708	0.04761	0.04765	1.147	1.151	-0.00585	0.04527	0.04529	1.206	1.211
	80	-0.00913	0.02904	0.02912	-0.00229	0.02402	0.02402	1.209	1.212	-0.00169	0.02276	0.02275	1.276	1.280
	120	-0.00702	0.02001	0.02006	-0.00138	0.01621	0.01621	1.235	1.238	-0.00089	0.01539	0.01539	1.300	1.303

ภาพที่ 4.5 แสดงการเปรียบเทียบค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงแบบค่าต่ำสุดขีด ที่พารามิเตอร์ μ = 0 และ σ = 1 ด้วยวิธี GE กับวิธี GEPD —— Trimmed 5% และ Trimmed 10% ตามลำดับ



ภาพที่ 4.5 **(ต่อ)**

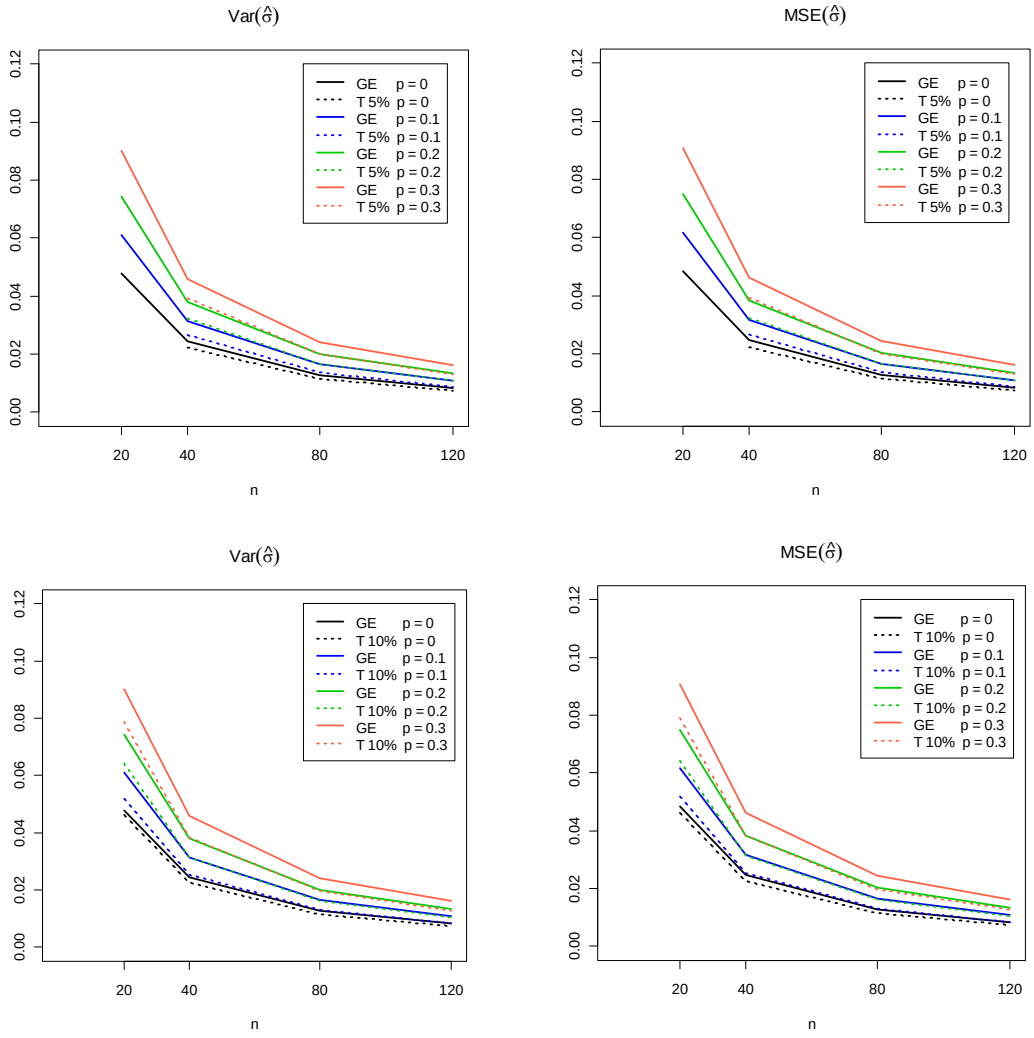


จากตารางและภาพที่ 4.5 พบว่า การประมาณค่าพารามิเตอร์ μ ด้วยวิธี GEPD แบบ Trimmed q% มีประสิทธิภาพมากกว่าวิธี GE ในทุกกรณี (ยกเว้น p = 0) เพราะวิธี Trimmed นั้นน่าจะช่วยตัดข้อมูลที่เป็น Outlier ออกไป ถ้ากำหนดให้สัดส่วนของการตัดปลาย (p) มีขนาดคงที่พบว่า ค่า RE1 และ RE2 จะมีค่าเพิ่มขึ้น เมื่อขนาดตัวอย่าง (n) เพิ่มขึ้น แต่ถ้ากำหนดให้ขนาดตัวอย่าง (n) คงที่พบว่า ค่าRE1 และRE2 จะมีค่าเพิ่มขึ้น เมื่อสัดส่วนของการตัดปลาย (p) เพิ่มขึ้น นอกจากนี้ยังพบว่าวิธีประมาณแบบ Trimmed 10% มีประสิทธิภาพมากกว่าวิธีประมาณแบบ Trimmed 5% ในทุกกรณี (ยกเว้น p = 0)

ตารางที่ 4.6 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงแบบค่าต่ำสุดขีด ที่พารามิเตอร์ μ = 0 และ σ = 1 โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— Trimmed 5% และ Trimmed 10% ตามลำดับ

p	n	($\hat{\sigma}$) → GE			($\hat{\sigma}$) → Trimmed 5%					($\hat{\sigma}$) → Trimmed 10%				
		Bias	Var	MSE	Bias	Var	MSE	RE1	RE2	Bias	Var	MSE	RE1	RE2
0	20	-0.02433	0.04783	0.04842	NA	NA	NA	NA	NA	-0.00999	0.04617	0.04626	1.036	1.047
	40	-0.01557	0.02443	0.02467	-0.00702	0.02221	0.02226	1.100	1.108	-0.00544	0.02267	0.02270	1.078	1.087
	80	-0.01005	0.01269	0.01279	-0.00404	0.01134	0.01135	1.119	1.126	-0.00329	0.01163	0.01164	1.091	1.099
	120	-0.00707	0.00835	0.00840	-0.00237	0.00734	0.00734	1.138	1.143	-0.00186	0.00748	0.00749	1.115	1.122
0.1	20	-0.02231	0.06117	0.06166	NA	NA	NA	NA	NA	-0.00890	0.05196	0.05203	1.177	1.185
	40	-0.01483	0.03141	0.03162	-0.00621	0.02650	0.02653	1.185	1.192	-0.00497	0.02546	0.02548	1.234	1.241
	80	-0.00986	0.01646	0.01655	-0.00357	0.01362	0.01363	1.209	1.215	-0.00295	0.01312	0.01313	1.255	1.261
	120	-0.00724	0.01088	0.01093	-0.00218	0.00880	0.00880	1.237	1.242	-0.00168	0.00843	0.00843	1.291	1.297
0.2	20	-0.02436	0.07442	0.07500	NA	NA	NA	NA	NA	-0.00869	0.06416	0.06422	1.160	1.168
	40	-0.01627	0.03814	0.03839	-0.00622	0.03237	0.03240	1.178	1.185	-0.00468	0.03134	0.03136	1.217	1.224
	80	-0.01091	0.02007	0.02019	-0.00357	0.01665	0.01666	1.206	1.212	-0.00280	0.01617	0.01617	1.242	1.249
	120	-0.00804	0.01335	0.01341	-0.00211	0.01079	0.01079	1.237	1.243	-0.00147	0.01041	0.01041	1.282	1.288
0.3	20	-0.02735	0.08998	0.09071	NA	NA	NA	NA	NA	-0.00949	0.07890	0.07898	1.140	1.149
	40	-0.01795	0.04589	0.04621	-0.00629	0.03928	0.03931	1.169	1.176	-0.00439	0.03835	0.03837	1.197	1.204
	80	-0.01224	0.02424	0.02438	-0.00376	0.02016	0.02017	1.202	1.209	-0.00285	0.01972	0.01973	1.229	1.236
	120	-0.00916	0.01621	0.01629	-0.00232	0.01313	0.01314	1.234	1.240	-0.00158	0.01278	0.01278	1.269	1.275

ภาพที่ 4.6 แสดงการเปรียบเทียบค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงแบบค่าต่ำสุดขีด ที่พารามิเตอร์ μ = 0 และ σ = 1 ด้วยวิธี GE กับวิธี GEPD —— Trimmed 5% และ Trimmed 10% ตามลำดับ

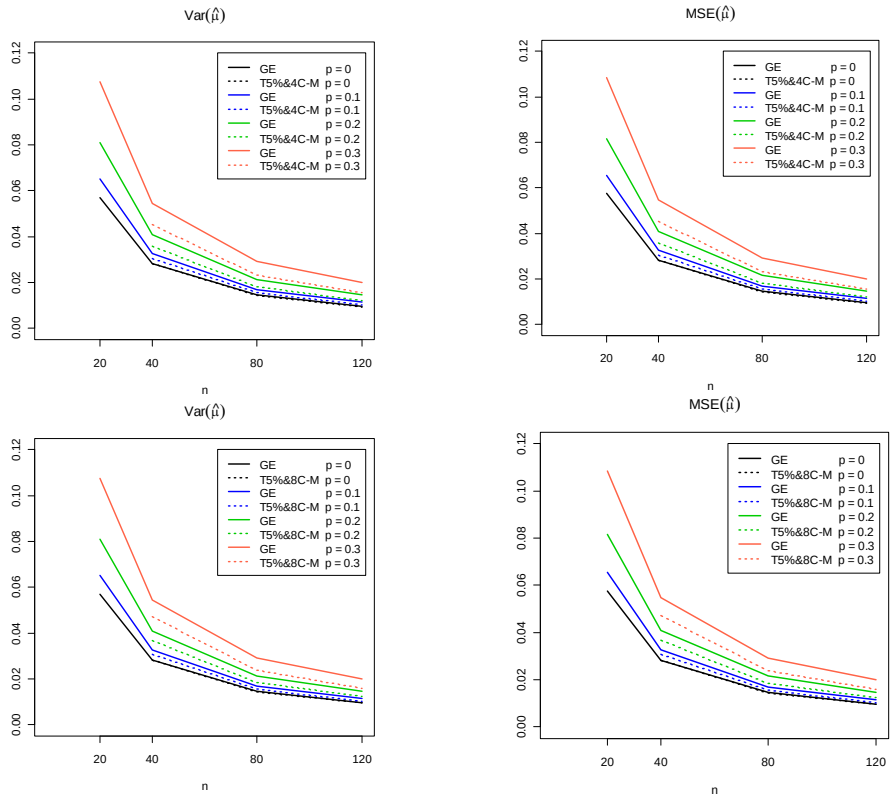


จากตารางและภาพที่ 4.6 พบว่า การประมาณค่าพารามิเตอร์ σ ด้วยวิธี GEPD แบบ Trimmed q% มีประสิทธิภาพมากกว่าวิธี GE ในทุกกรณี ถ้ากำหนดให้สัดส่วนของการตัดปลาย (p) มีขนาดคงที่พบว่า ค่า RE1 และ RE2 จะมีค่าเพิ่มขึ้น เมื่อขนาดตัวอย่าง (n) เพิ่มขึ้น และถ้ากำหนดให้ขนาดตัวอย่าง (n) คงที่และ $p \leq 0.1$ พบว่า ค่า RE1 และ RE2 จะมีค่าเพิ่มขึ้นเมื่อสัดส่วนของการตัดปลาย (p) เพิ่มขึ้น (แต่ถ้า $p > 0.1$ แล้ว ค่า RE1 และ RE2 จะมีค่าลดลงเมื่อสัดส่วนของการตัดปลาย (p) เพิ่มขึ้น) ทั้งนี้ยังพบว่าวิธีประมาณแบบ Trimmed 10% มีประสิทธิภาพมากกว่าวิธีประมาณแบบ Trimmed 5% ในทุกกรณี (ยกเว้น p = 0)

ตารางที่ 4.7 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงแบบค่าต่ำสุดขีด ที่พารามิเตอร์ μ = 0 และ σ = 1 โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— Trimmed 5% & 4 Cluster-Mean และ Trimmed 5% & 8 Cluster-Mean ตามลำดับ

p	n	($\hat{\mu}$) → GE			($\hat{\mu}$) → Trimmed 5% & 4 Cluster-Mean					($\hat{\mu}$) → Trimmed 5% & 8 Cluster-Mean				
		Bias	Var	MSE	Bias	Var	MSE	RE1	RE2	Bias	Var	MSE	RE1	RE2
0	20	-0.02488	0.05712	0.05772	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	40	-0.01116	0.02827	0.02838	-0.00763	0.02798	0.02803	1.010	1.013	-0.00795	0.02804	0.02809	1.008	1.010
	80	-0.00483	0.01459	0.01461	-0.00237	0.01434	0.01434	1.018	1.019	-0.00259	0.01436	0.01436	1.016	1.018
	120	-0.00304	0.00956	0.00956	-0.00124	0.00937	0.00937	1.020	1.021	-0.00136	0.00937	0.00937	1.020	1.020
0.1	20	-0.02220	0.06517	0.06565	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	40	-0.01029	0.03264	0.03274	-0.00667	0.02922	0.02925	1.117	1.119	-0.00653	0.02892	0.02896	1.129	1.131
	80	-0.00468	0.01694	0.01696	-0.00175	0.01494	0.01494	1.134	1.136	-0.00201	0.01476	0.01476	1.148	1.149
	120	-0.00340	0.01136	0.01137	-0.00095	0.00983	0.00983	1.155	1.156	-0.00130	0.00970	0.00970	1.171	1.172
0.2	20	-0.02562	0.08086	0.08150	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	40	-0.01273	0.04086	0.04101	-0.00697	0.03165	0.03169	1.291	1.294	-0.00705	0.03383	0.03388	1.208	1.211
	80	-0.00652	0.02140	0.02144	-0.00164	0.01606	0.01606	1.333	1.335	-0.00187	0.01711	0.01711	1.251	1.253
	120	-0.00482	0.01460	0.01462	-0.00054	0.01065	0.01065	1.371	1.373	-0.00078	0.01143	0.01142	1.278	1.280
0.3	20	-0.03113	0.10744	0.10839	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	40	-0.01597	0.05459	0.05483	-0.00538	0.04342	0.04345	1.257	1.262	-0.00642	0.04373	0.04376	1.248	1.253
	80	-0.00913	0.02904	0.02912	-0.00144	0.02189	0.02189	1.327	1.330	-0.00202	0.02201	0.02201	1.319	1.323
	120	-0.00702	0.02001	0.02006	-0.00081	0.01477	0.01477	1.355	1.358	-0.00108	0.01487	0.01487	1.346	1.349

ภาพที่ 4.7 แสดงการเปรียบเทียบค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงแบบค่าต่ำสุดขีด ที่พารามิเตอร์ μ = 0 และ σ = 1 ด้วยวิธี GE กับวิธี GEPD——Trimmed 5% & 4 Cluster-Mean และ Trimmed 5% & 8 Cluster-Mean ตามลำดับ

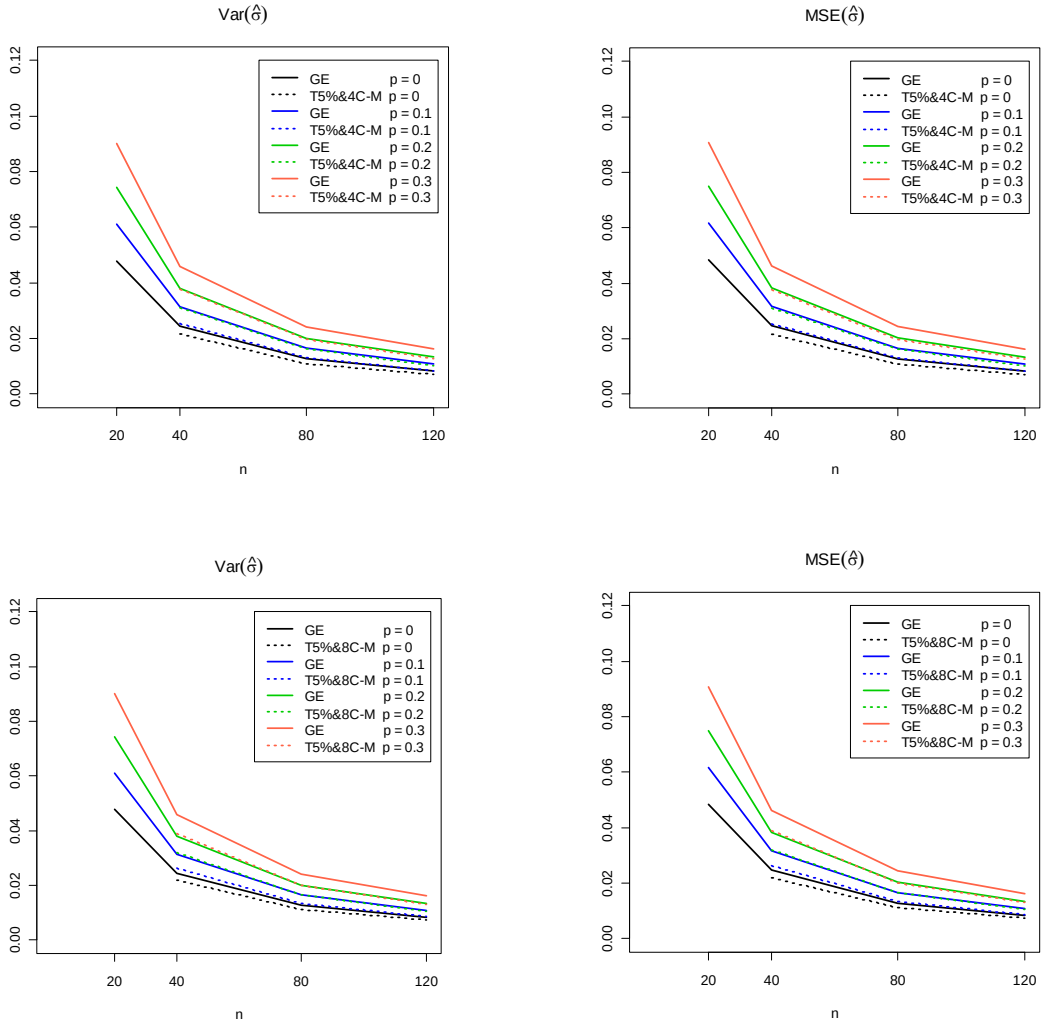


จากตารางที่ 4.7 พบว่า การประมาณค่าพารามิเตอร์ μ ด้วยวิธี GEPD แบบ Trimmed 5% & K Cluster-Mean มีประสิทธิภาพมากกว่าวิธี GE โดยจะมีประสิทธิภาพมากยิ่งขึ้นเมื่อสัดส่วนของการตัดปลาย (p) และขนาดตัวอย่าง (n) มีค่าเพิ่มขึ้น (ยกเว้น p = 0.3) เมื่อ $p \neq 0.1$ แล้วจะพบว่าวิธีประมาณแบบ Trimmed 5% & 4 Cluster-Mean มีประสิทธิภาพมากกว่าแบบ Trimmed 5% & 8 Cluster-Mean ในทุกกรณี

ตารางที่ 4.8 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงแบบค่าต่ำสุดขีด ที่พารามิเตอร์ μ = 0 และ σ = 1 โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— Trimmed 5% & 4 Cluster-Mean และ Trimmed 5% & 8 Cluster-Mean ตามลำดับ

p	n	($\hat{\sigma}$) → GE			($\hat{\sigma}$) → Trimmed 5% & 4 Cluster-Mean					($\hat{\sigma}$) → Trimmed 5% & 8 Cluster-Mean				
		Bias	Var	MSE	Bias	Var	MSE	RE1	RE2	Bias	Var	MSE	RE1	RE2
0	20	-0.02433	0.04783	0.04842	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	40	-0.01557	0.02443	0.02467	-0.00633	0.02148	0.02151	1.137	1.147	-0.00704	0.02223	0.02228	1.099	1.107
	80	-0.01005	0.01269	0.01279	-0.00364	0.01099	0.01100	1.155	1.163	-0.00409	0.01135	0.01136	1.119	1.126
	120	-0.00707	0.00835	0.00840	-0.00209	0.00711	0.00711	1.174	1.181	-0.00231	0.00735	0.00735	1.136	1.142
0.1	20	-0.02231	0.06117	0.06166	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	40	-0.01483	0.03141	0.03162	-0.00561	0.02388	0.02391	1.315	1.323	-0.00600	0.02395	0.02398	1.311	1.318
	80	-0.00986	0.01646	0.01655	-0.00318	0.01229	0.01230	1.339	1.346	-0.00365	0.01230	0.01231	1.338	1.345
	120	-0.00724	0.01088	0.01093	-0.00188	0.00793	0.00793	1.373	1.379	-0.00225	0.00794	0.00794	1.370	1.376
0.2	20	-0.02436	0.07442	0.07500	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	40	-0.01627	0.03814	0.03839	-0.00578	0.02669	0.02672	1.429	1.437	-0.00634	0.03033	0.03037	1.257	1.264
	80	-0.01091	0.02007	0.02019	-0.00310	0.01376	0.01377	1.459	1.467	-0.00355	0.01562	0.01563	1.285	1.292
	120	-0.00804	0.01335	0.01341	-0.00160	0.00887	0.00887	1.504	1.511	-0.00190	0.01011	0.01011	1.321	1.327
0.3	20	-0.02735	0.08998	0.09071	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	40	-0.01795	0.04589	0.04621	-0.00473	0.03661	0.03663	1.253	1.262	-0.00597	0.03781	0.03784	1.214	1.221
	80	-0.01224	0.02424	0.02438	-0.00297	0.01883	0.01883	1.287	1.295	-0.00364	0.01943	0.01944	1.247	1.254
	120	-0.00916	0.01621	0.01629	-0.00179	0.01225	0.01225	1.323	1.330	-0.00209	0.01266	0.01266	1.281	1.287

ภาพที่ 4.8 แสดงการเปรียบเทียบค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงแบบค่าต่ำสุดขีด ที่พารามิเตอร์ $\mu = 0$ และ $\sigma = 1$ ด้วยวิธี GE กับวิธี GEPD——Trimmed 5% & 4 Cluster-Mean และ Trimmed 5% & 8 Cluster-Mean ตามลำดับ

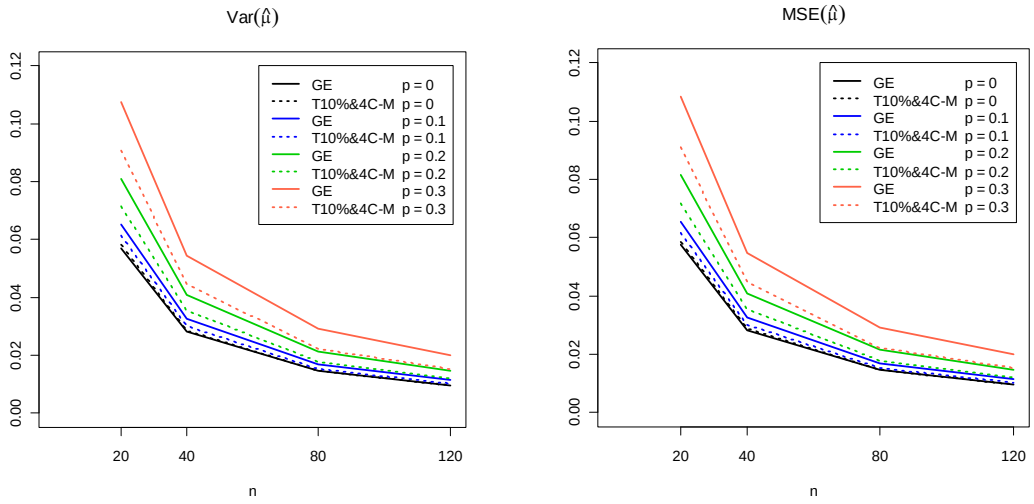


จากตารางและภาพที่ 4.8 พบว่า การประมาณค่าพารามิเตอร์ σ ด้วยวิธี GEPD แบบ Trimmed 5% & K Cluster-Mean มีประสิทธิภาพมากกว่าวิธี GE ในทุกกรณี โดยวิธีประมาณแบบ Trimmed 5% & 4 Cluster-Mean พบว่าค่า RE1 และ RE2 จะเพิ่มขึ้นเมื่อขนาดตัวอย่าง (n) และสัดส่วนของการตัดปลายมากขึ้น (ยกเว้น $p = 0.3$) และเมื่อพิจารณาวิธีประมาณแบบ Trimmed 5% & 8 Cluster-Mean พบว่าค่า RE1 และ RE2 จะเพิ่มขึ้นเมื่อขนาดตัวอย่าง (n) มากขึ้นและสัดส่วนของการตัดปลาย $p \leq 0.1$ นอกจากนี้ยังพบว่าวิธีประมาณแบบ Trimmed 5% & 4 Cluster-Mean มีประสิทธิภาพมากกว่าวิธีประมาณแบบ Trimmed 5% & 8 Cluster-Mean ในทุกกรณี

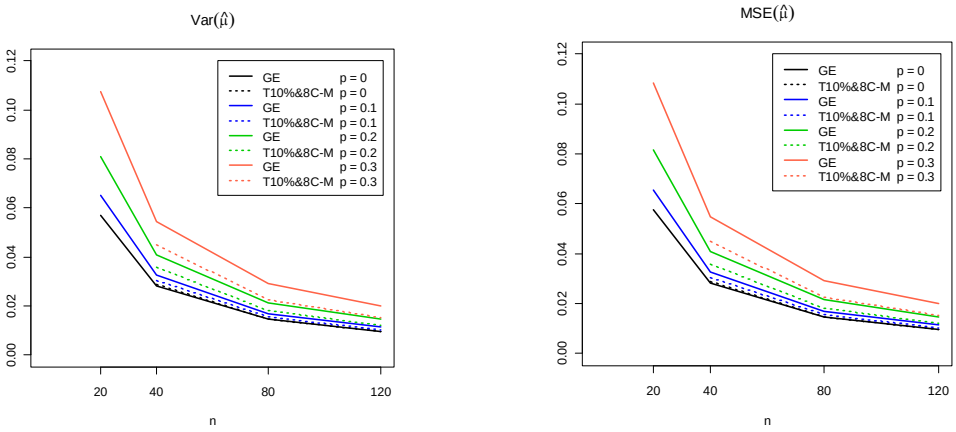
ตารางที่ 4.9 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงแบบค่าต่ำสุดขีด ที่พารามิเตอร์ μ = 0 และ σ = 1 โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— Trimmed 10% & 4 Cluster-Mean และ Trimmed 10% & 8 Cluster-Mean ตามลำดับ

p	n	($\hat{\mu}$) → GE			($\hat{\mu}$) → Trimmed 10% & 4 Cluster-Mean					($\hat{\mu}$) → Trimmed 10% & 8 Cluster-Mean				
		Bias	Var	MSE	Bias	Var	MSE	RE1	RE2	Bias	Var	MSE	RE1	RE2
0	20	-0.02488	0.05712	0.05772	-0.01890	0.05747	0.05782	0.994	0.998	NA	NA	NA	NA	NA
	40	-0.01116	0.02827	0.02838	-0.00713	0.02835	0.02839	0.997	1.000	-0.00730	0.02825	0.02830	1.001	1.003
	80	-0.00483	0.01459	0.01461	-0.00213	0.01451	0.01452	1.005	1.007	-0.00234	0.01445	0.01445	1.010	1.011
	120	-0.00304	0.00956	0.00956	-0.00116	0.00950	0.00950	1.006	1.007	-0.00131	0.00945	0.00945	1.011	1.012
0.1	20	-0.02220	0.06517	0.06565	-0.01755	0.05922	0.05952	1.100	1.103	NA	NA	NA	NA	NA
	40	-0.01029	0.03264	0.03274	-0.00673	0.02919	0.02923	1.118	1.120	-0.00649	0.02884	0.02888	1.132	1.134
	80	-0.00468	0.01694	0.01696	-0.00177	0.01492	0.01492	1.135	1.136	-0.00199	0.01473	0.01473	1.150	1.152
	120	-0.00340	0.01136	0.01137	-0.00098	0.00982	0.00982	1.156	1.157	-0.00130	0.00967	0.00967	1.174	1.175
0.2	20	-0.02562	0.08086	0.08150	-0.01635	0.06435	0.06461	1.257	1.261	NA	NA	NA	NA	NA
	40	-0.01273	0.04086	0.04101	-0.00696	0.03160	0.03164	1.293	1.296	-0.00676	0.03347	0.03351	1.221	1.224
	80	-0.00652	0.02140	0.02144	-0.00163	0.01603	0.01603	1.335	1.337	-0.00173	0.01691	0.01691	1.266	1.268
	120	-0.00482	0.01460	0.01462	-0.00053	0.01063	0.01063	1.374	1.376	-0.00065	0.01130	0.01130	1.292	1.294
0.3	20	-0.03113	0.10744	0.10839	-0.01736	0.08805	0.08833	1.220	1.227	NA	NA	NA	NA	NA
	40	-0.01597	0.05459	0.05483	-0.00491	0.04315	0.04316	1.265	1.270	-0.00577	0.04299	0.04302	1.270	1.275
	80	-0.00913	0.02904	0.02912	-0.00122	0.02175	0.02175	1.335	1.339	-0.00172	0.02160	0.02160	1.345	1.348
	120	-0.00702	0.02001	0.02006	-0.00065	0.01468	0.01468	1.363	1.366	-0.00081	0.01461	0.01461	1.370	1.373

ภาพที่ 4.9 แสดงการเปรียบเทียบค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงแบบค่าต่ำสุดขีด ที่พารามิเตอร์ μ = 0 และ σ = 1 ด้วยวิธี GE กับวิธี GEPD —— Trimmed 10% & 4 Cluster-Mean และ Trimmed 10% & 8 Cluster-Mean ตามลำดับ



ภาพที่ 4.9 **(ต่อ)**

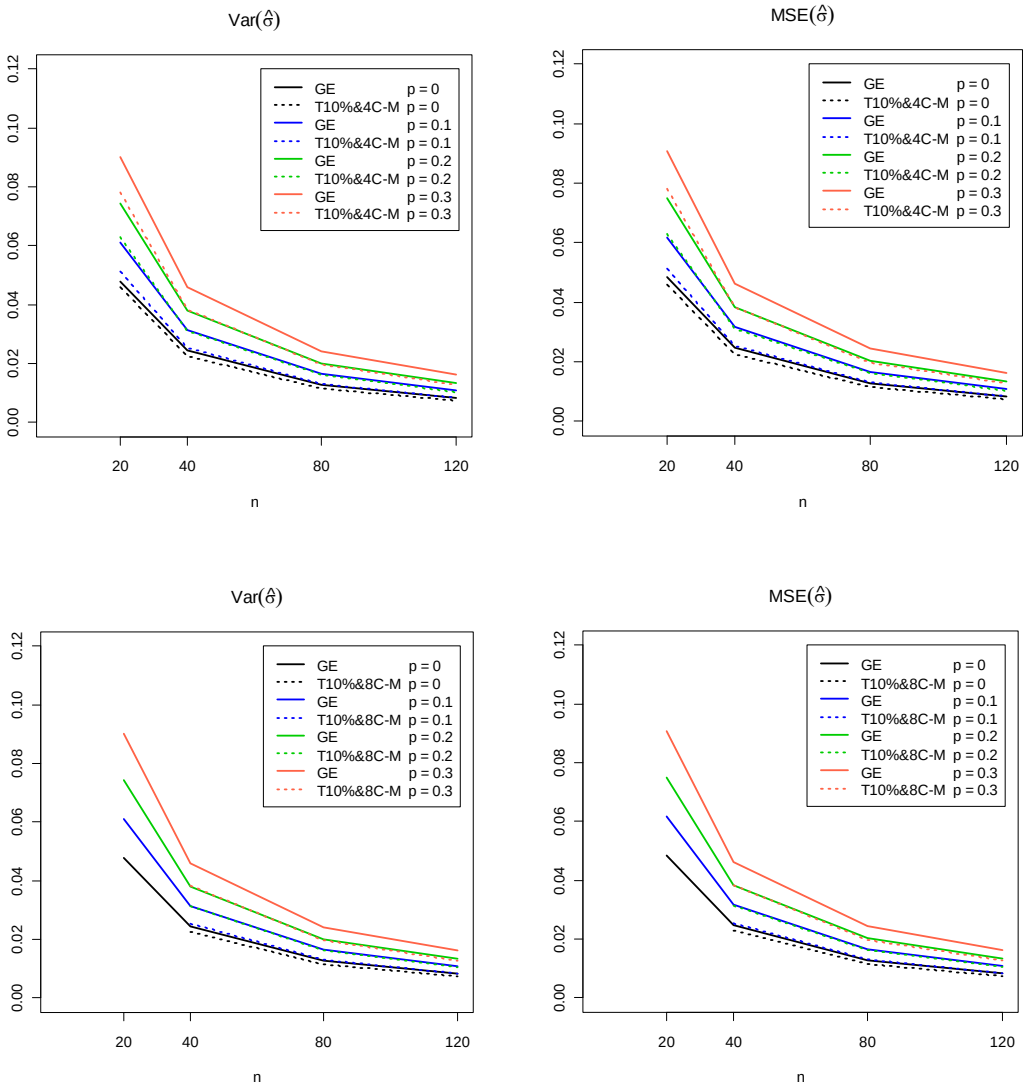


จากตารางและภาพที่ 4.9 พบว่า การประมาณค่าพารามิเตอร์ μ ด้วยวิธี GEPD แบบ Trimmed 10% & 4 Cluster-Mean มีประสิทธิภาพมากกว่าวิธี GE ในทุกกรณี (ยกเว้น p = 0 และ $n \leq 40$) ถ้ากำหนดให้สัดส่วนของการตัดปลาย (p) มีขนาดคงที่พบว่า ค่า RE1 และ RE2 จะมีค่าเพิ่มขึ้น เมื่อขนาดตัวอย่าง (n) เพิ่มขึ้น นอกจากนี้ยังพบว่าในกรณีส่วนใหญ่วิธีประมาณแบบ Trimmed 10% & 8 Cluster-Mean มีประสิทธิภาพมากกว่าวิธีประมาณแบบ Trimmed 10% & 4 Cluster-Mean (ยกเว้น p = 0.2)

ตารางที่ 4.10 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงแบบค่าต่ำสุดขีด ที่พารามิเตอร์ μ = 0 และ σ = 1 โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— Trimmed 10% & 4 Cluster-Mean และ Trimmed 10% & 8 Cluster-Mean ตามลำดับ

p	n	($\hat{\sigma}$) → GE			($\hat{\sigma}$) → Trimmed 10% & 4 Cluster-Mean					($\hat{\sigma}$) → Trimmed 10% & 8 Cluster-Mean				
		Bias	Var	MSE	Bias	Var	MSE	RE1	RE2	Bias	Var	MSE	RE1	RE2
0	20	-0.02433	0.04783	0.04842	-0.00926	0.04580	0.04588	1.044	1.055	NA	NA	NA	NA	NA
	40	-0.01557	0.02443	0.02467	-0.00521	0.02253	0.02256	1.084	1.094	-0.00583	0.02292	0.02295	1.066	1.075
	80	-0.01005	0.01269	0.01279	-0.00310	0.01157	0.01158	1.097	1.104	-0.00358	0.01170	0.01171	1.085	1.092
	120	-0.00707	0.00835	0.00840	-0.00174	0.00747	0.00747	1.117	1.124	-0.00193	0.00755	0.00755	1.106	1.112
0.1	20	-0.02231	0.06117	0.06166	-0.00820	0.04927	0.04932	1.242	1.250	NA	NA	NA	NA	NA
	40	-0.01483	0.03141	0.03162	-0.00487	0.02418	0.02420	1.299	1.306	-0.00520	0.02405	0.02407	1.306	1.314
	80	-0.00986	0.01646	0.01655	-0.00281	0.01247	0.01248	1.320	1.326	-0.00330	0.01234	0.01234	1.334	1.341
	120	-0.00724	0.01088	0.01093	-0.00159	0.00804	0.00804	1.354	1.360	-0.00191	0.00794	0.00794	1.371	1.377
0.2	20	-0.02436	0.07442	0.07500	-0.00731	0.05570	0.05574	1.336	1.345	NA	NA	NA	NA	NA
	40	-0.01627	0.03814	0.03839	-0.00499	0.02726	0.02728	1.399	1.407	-0.00539	0.03067	0.03069	1.243	1.251
	80	-0.01091	0.02007	0.02019	-0.00269	0.01408	0.01408	1.426	1.433	-0.00311	0.01578	0.01578	1.272	1.279
	120	-0.00804	0.01335	0.01341	-0.00126	0.00907	0.00907	1.472	1.479	-0.00144	0.01017	0.01017	1.313	1.319
0.3	20	-0.02735	0.08998	0.09071	-0.00808	0.07750	0.07755	1.161	1.170	NA	NA	NA	NA	NA
	40	-0.01795	0.04589	0.04621	-0.00353	0.03793	0.03794	1.210	1.218	-0.00474	0.03859	0.03861	1.189	1.197
	80	-0.01224	0.02424	0.02438	-0.00240	0.01954	0.01954	1.240	1.248	-0.00310	0.01980	0.01980	1.224	1.231
	120	-0.00916	0.01621	0.01629	-0.00135	0.01270	0.01270	1.277	1.283	-0.00155	0.01285	0.01285	1.261	1.268

ภาพที่ 4.10 แสดงการเปรียบเทียบค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงแบบค่าต่ำสุดขีด ที่พารามิเตอร์ $\mu = 0$ และ $\sigma = 1$ ด้วยวิธี GE กับวิธี GEPD——Trimmed 10% & 4 Cluster-Mean และ Trimmed 10% & 8 Cluster-Mean ตามลำดับ



จากตารางที่ 4.10 พบว่า การประมาณค่าพารามิเตอร์ σ ด้วยวิธี GEPD แบบ Trimmed 10% & K Cluster-Mean มีประสิทธิภาพมากกว่าวิธีประมาณแบบ GE ในทุกกรณี โดยวิธีประมาณแบบ Trimmed 10% & 4 Cluster-Mean มีประสิทธิภาพมากกว่าวิธีประมาณแบบ Trimmed 10% & 8 Cluster-Mean ในทุกกรณี (ยกเว้น p =0.1)

4.1.2 ผลการศึกษาเมื่อข้อมูลแจกแจงแบบค่าสูงสุดขีด

ส่วนที่ 1 การเปรียบเทียบประสิทธิภาพของวิธีการประมาณค่าพารามิเตอร์แบบภาวะน่าจะเป็นสูงสุด (MLE) กับแบบใช้กราฟ (GE)

ตารางที่ 4.11 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงแบบค่าสูงสุดขีด ที่พารามิเตอร์ $\mu = 0$ และ $\sigma = 1$ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา

Censoring	Sample	$(\hat{\mu}) \rightarrow$ MLE			$(\hat{\mu}) \rightarrow$ GE				
Proportion p	Size n	Bias	Var	MSE	Bias	Var	MSE	RE1	RE2
0	20	0.02458	0.05169	0.05229	0.02488	0.05712	0.05772	0.905	0.906
	40	0.01002	0.02503	0.02513	0.01116	0.02827	0.02838	0.886	0.885
	80	0.00369	0.01273	0.01274	0.00483	0.01459	0.01461	0.872	0.872
	120	0.00157	0.00836	0.00836	0.00304	0.00956	0.00956	0.875	0.875
0.1	20	0.02465	0.05174	0.05234	0.02407	0.05680	0.05736	0.911	0.912
	40	0.01008	0.02503	0.02512	0.01006	0.02797	0.02807	0.895	0.895
	80	0.00373	0.01272	0.01273	0.00380	0.01433	0.01434	0.888	0.888
	120	0.00159	0.00837	0.00837	0.00210	0.00935	0.00935	0.895	0.895
0.2	20	0.02450	0.05178	0.05237	0.02326	0.05797	0.05849	0.893	0.895
	40	0.01016	0.02504	0.02513	0.00959	0.02852	0.02861	0.878	0.878
	80	0.00369	0.01274	0.01275	0.00353	0.01459	0.01460	0.873	0.874
	120	0.00161	0.00839	0.00839	0.00197	0.00957	0.00957	0.876	0.876
0.3	20	0.02463	0.05207	0.05267	0.02180	0.06017	0.06063	0.865	0.869
	40	0.01003	0.02511	0.02521	0.00848	0.02961	0.02968	0.848	0.849
	80	0.00364	0.01276	0.01277	0.00287	0.01504	0.01505	0.848	0.849
	120	0.00160	0.00839	0.00839	0.00162	0.00991	0.00991	0.847	0.847

จากตารางที่ 4.11 พบว่า การประมาณค่าพารามิเตอร์ μ ด้วยวิธี MLE มีประสิทธิภาพมากกว่าวิธี GE ในทุกกรณี โดยพิจารณาจากค่า |Biases| และค่า Variances ของค่าประมาณพารามิเตอร์ μ ซึ่งวิธี GE มีค่าสูงกว่าวิธี MLE เป็นส่วนใหญ่ ส่งผลให้ค่า MSEs มีค่าสูงขึ้นตามไปด้วย เมื่อขนาดตัวอย่าง $n \neq 120$, ถ้ากำหนดให้สัดส่วนของการตัดปลาย (p) มีขนาดคงที่พบว่า ค่า RE1 และ RE2 จะมีค่าลดลง เมื่อขนาดตัวอย่าง (n) เพิ่มขึ้น แต่ถ้ากำหนดให้ขนาดตัวอย่าง (n) คงที่พบว่า ค่า RE1 และ RE2 จะมีค่าเพิ่มขึ้น เมื่อสัดส่วนของการตัดปลาย $p \leq 0.1$ และจะมีค่าลดลงเมื่อ $p \geq 0.2$

ตารางที่ 4.12 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงแบบค่าสูงสุดขีด ที่พารามิเตอร์ $\mu = 0$ และ $\sigma = 1$ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา

Censoring	Sample	$(\hat{\sigma}) \rightarrow$ MLE			$(\hat{\sigma}) \rightarrow$ GE				
Proportion p	Size n	Bias	Var	MSE	Bias	Var	MSE	RE1	RE2
0	20	-0.01548	0.02798	0.02822	-0.02433	0.04783	0.04842	0.585	0.583
	40	-0.00895	0.01365	0.01372	-0.01557	0.02443	0.02467	0.559	0.556
	80	-0.00472	0.00692	0.00694	-0.01005	0.01269	0.01279	0.545	0.542
	120	-0.00244	0.00450	0.00451	-0.00707	0.00835	0.00840	0.539	0.537
0.1	20	-0.01693	0.02962	0.02990	-0.01410	0.03881	0.03900	0.763	0.767
	40	-0.00974	0.01443	0.01452	-0.00854	0.01879	0.01886	0.768	0.770
	80	-0.00507	0.00734	0.00737	-0.00506	0.00956	0.00958	0.768	0.769
	120	-0.00248	0.00485	0.00486	-0.00314	0.00623	0.00624	0.778	0.778
0.2	20	-0.02045	0.03240	0.03281	-0.01763	0.04094	0.04124	0.791	0.795
	40	-0.01082	0.01563	0.01574	-0.01044	0.01965	0.01976	0.795	0.797
	80	-0.00526	0.00799	0.00802	-0.00615	0.01006	0.01010	0.794	0.794
	120	-0.00274	0.00524	0.00524	-0.00367	0.00657	0.00658	0.798	0.797
0.3	20	-0.02389	0.03532	0.03588	-0.02144	0.04548	0.04593	0.777	0.781
	40	-0.01309	0.01702	0.01719	-0.01328	0.02195	0.02212	0.775	0.777
	80	-0.00664	0.00882	0.00886	-0.00783	0.01114	0.01120	0.791	0.791
	120	-0.00330	0.00576	0.00577	-0.00457	0.00729	0.00731	0.790	0.789

จากตารางที่ 4.12 พบว่า การประมาณค่าพารามิเตอร์ σ ด้วยวิธี MLE มีประสิทธิภาพมากกว่าวิธี GE ในทุกกรณี โดยพิจารณาจากค่า |Biases| และค่า Variances ของค่าประมาณพารามิเตอร์ σ ซึ่งวิธี GE มีค่าสูงกว่าวิธี MLE เป็นส่วนใหญ่ ส่งผลให้ค่า MSEs มีค่าสูงขึ้นตามไปด้วย เมื่อกำหนดให้สัดส่วนของการตัดปลาย (p) คงที่ พบว่ากรณีที่ข้อมูลไม่ถูกตัดปลายทางขวา (p =0) ค่า RE1 และ RE2 จะมีค่าลดลง เมื่อขนาดตัวอย่าง (n) เพิ่มขึ้น ส่วนกรณีที่ข้อมูลถูกตัดปลายทางขวา ($p \neq 0$) พบว่า ค่า RE1 และ RE2 มีแนวโน้มเพิ่มขึ้น เมื่อขนาดตัวอย่าง (n) เพิ่มขึ้น แต่ถ้ากำหนดให้ขนาดตัวอย่าง (n) คงที่พบว่า ค่า RE1 และ RE2 จะมีค่าเพิ่มขึ้น เมื่อสัดส่วนของการตัดปลาย $p \leq 0.2$ และจะมีค่าลดลงเมื่อ $p > 0.2$

ส่วนที่ 2 การเปรียบเทียบประสิทธิภาพของวิธีการประมาณค่าพารามิเตอร์แบบใช้กราฟ (GE) กับแบบใช้กราฟที่ใช้ข้อมูลเพียงบางส่วน (GEPD)

ตารางที่ 4.13 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงแบบค่าสูงสุดขีด ที่พารามิเตอร์ $\mu = 0$ และ $\sigma = 1$ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา ── 4 Cluster-Mean และ 8 Cluster-Mean ตามลำดับ

p	n	($\hat{\mu}$) → GE			($\hat{\mu}$) → 4 Cluster-Mean					($\hat{\mu}$) → 8 Cluster-Mean				
		Bias	Var	MSE	Bias	Var	MSE	RE1	RE2	Bias	Var	MSE	RE1	RE2
0	20	0.02488	0.05712	0.05772	0.02205	0.05629	0.05677	1.015	1.017	NA	NA	NA	NA	NA
	40	0.01116	0.02827	0.02838	0.00882	0.02771	0.02778	1.020	1.022	0.00970	0.02792	0.02801	1.012	1.013
	80	0.00483	0.01459	0.01461	0.00301	0.01419	0.01419	1.028	1.029	0.00358	0.01429	0.01430	1.021	1.022
	120	0.00304	0.00956	0.00956	0.00155	0.00925	0.00925	1.033	1.034	0.00194	0.00933	0.00933	1.024	1.025
0.1	20	0.02407	0.05680	0.05736	0.02377	0.05672	0.05727	1.001	1.002	NA	NA	NA	NA	NA
	40	0.01006	0.02797	0.02807	0.00977	0.02791	0.02800	1.002	1.002	0.00990	0.02795	0.02805	1.001	1.001
	80	0.00380	0.01433	0.01434	0.00361	0.01430	0.01431	1.002	1.002	0.00368	0.01432	0.01433	1.001	1.001
	120	0.00210	0.00935	0.00935	0.00197	0.00932	0.00932	1.003	1.003	0.00203	0.00934	0.00934	1.001	1.001
0.2	20	0.02326	0.05797	0.05849	0.02307	0.05792	0.05845	1.001	1.001	NA	NA	NA	NA	NA
	40	0.00959	0.02852	0.02861	0.00946	0.02851	0.02859	1.001	1.001	0.00952	0.02852	0.02860	1.000	1.000
	80	0.00353	0.01459	0.01460	0.00345	0.01458	0.01459	1.001	1.001	0.00348	0.01458	0.01459	1.000	1.000
	120	0.00197	0.00957	0.00957	0.00193	0.00956	0.00957	1.001	1.001	0.00194	0.00957	0.00957	1.000	1.000
0.3	20	0.02180	0.06017	0.06063	0.02199	0.06020	0.06067	1.000	0.999	NA	NA	NA	NA	NA
	40	0.00848	0.02961	0.02968	0.00866	0.02963	0.02970	0.999	0.999	0.00855	0.02962	0.02968	1.000	1.000
	80	0.00287	0.01504	0.01505	0.00298	0.01505	0.01505	0.999	0.999	0.00293	0.01504	0.01505	1.000	1.000
	120	0.00162	0.00991	0.00991	0.00169	0.00992	0.00992	0.999	0.999	0.00166	0.00991	0.00991	1.000	1.000

ตารางที่ 4.14 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงแบบค่าสูงสุดขีด ที่พารามิเตอร์ $\mu = 0$ และ $\sigma = 1$ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา ── 4 Cluster-Mean และ 8 Cluster-Mean ตามลำดับ

p	n	($\hat{\sigma}$) → GE			($\hat{\sigma}$) → 4 Cluster-Mean					($\hat{\sigma}$) → 8 Cluster-Mean				
		Bias	Var	MSE	Bias	Var	MSE	RE1	RE2	Bias	Var	MSE	RE1	RE2
0	20	-0.02433	0.04783	0.04842	-0.01930	0.04266	0.04302	1.121	1.125	NA	NA	NA	NA	NA
	40	-0.01557	0.02443	0.02467	-0.01146	0.02102	0.02115	1.162	1.166	-0.01301	0.02248	0.02264	1.087	1.089
	80	-0.01005	0.01269	0.01279	-0.00688	0.01067	0.01071	1.190	1.194	-0.00787	0.01143	0.01149	1.110	1.113
	120	-0.00707	0.00835	0.00840	-0.00447	0.00688	0.00690	1.213	1.216	-0.00514	0.00741	0.00744	1.127	1.129
0.1	20	-0.01410	0.03881	0.03900	-0.01302	0.03977	0.03993	0.976	0.977	NA	NA	NA	NA	NA
	40	-0.00854	0.01879	0.01886	-0.00752	0.01951	0.01956	0.963	0.964	-0.00798	0.01902	0.01908	0.988	0.988
	80	-0.00506	0.00956	0.00958	-0.00438	0.01000	0.01001	0.956	0.957	-0.00464	0.00974	0.00976	0.982	0.982
	120	-0.00314	0.00623	0.00624	-0.00266	0.00653	0.00653	0.955	0.956	-0.00288	0.00635	0.00636	0.982	0.982
0.2	20	-0.01763	0.04094	0.04124	-0.01538	0.04291	0.04313	0.954	0.956	NA	NA	NA	NA	NA
	40	-0.01044	0.01965	0.01976	-0.00889	0.02075	0.02083	0.947	0.949	-0.00967	0.02003	0.02012	0.981	0.982
	80	-0.00615	0.01006	0.01010	-0.00514	0.01068	0.01070	0.942	0.944	-0.00559	0.01029	0.01032	0.978	0.978
	120	-0.00367	0.00657	0.00658	-0.00314	0.00699	0.00699	0.940	0.940	-0.00329	0.00673	0.00674	0.976	0.976
0.3	20	-0.02144	0.04548	0.04593	-0.01918	0.04688	0.04723	0.970	0.972	NA	NA	NA	NA	NA
	40	-0.01328	0.02195	0.02212	-0.01110	0.02334	0.02346	0.940	0.943	-0.01249	0.02227	0.02242	0.986	0.987
	80	-0.00783	0.01114	0.01120	-0.00654	0.01188	0.01192	0.938	0.940	-0.00709	0.01141	0.01146	0.977	0.978
	120	-0.00457	0.00729	0.00731	-0.00376	0.00781	0.00782	0.934	0.935	-0.00407	0.00748	0.00749	0.975	0.976

จากตารางที่ 4.13 พบว่า การประมาณค่าพารามิเตอร์ μ ด้วยวิธี GEPD แบบ 4 Cluster-Mean มีประสิทธิภาพมากกว่าวิธี GE ในทุกกรณี (ยกเว้น p = 0.3) ส่วนวิธีประมาณแบบ 8 Cluster-Mean นั้น จะมีประสิทธิภาพมากกว่าวิธี GE เมื่อ $p \leq 0.1$ ถ้ากำหนดให้ขนาดตัวอย่าง (n) คงที่แล้ว ค่า RE1 และ RE2 จะมีค่าลดลง เมื่อสัดส่วนของการตัดปลาย (p) มีขนาดเพิ่มขึ้น นอกจากนี้ยังพบว่าวิธีประมาณแบบ 4 Cluster-Mean มีประสิทธิภาพมากกว่าวิธีประมาณแบบ 8 Cluster-Mean ในทุกกรณี (ยกเว้น p = 0.3)

จากตารางที่ 4.14 พบว่า โดยรวมการประมาณค่าพารามิเตอร์ σ ด้วยวิธี GE มีประสิทธิภาพมากกว่าวิธี GEPD แบบ 4 Cluster-Mean และ 8 Cluster-Mean ในทุกกรณี (ยกเว้น p = 0) แต่ถ้า p = 0 วิธี GEPD แบบ K Cluster-Mean จะมีประสิทธิภาพมากกว่าวิธี GE โดยพบว่าแบบ 4 Cluster-Mean มีประสิทธิภาพมากกว่าแบบ 8 Cluster-Mean

ตารางที่ 4.15 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงแบบค่าสูงสุดขีด ที่พารามิเตอร์ μ = 0 และ σ = 1 โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— Trimmed 5% และ Trimmed 10% ตามลำดับ

p	n	($\hat{\mu}$) → GE			($\hat{\mu}$) → Trimmed 5%					($\hat{\mu}$) → Trimmed 10%				
		Bias	Var	MSE	Bias	Var	MSE	RE1	RE2	Bias	Var	MSE	RE1	RE2
0	20	0.02488	0.05712	0.05772	NA	NA	NA	NA	NA	0.01919	0.05846	0.05882	0.977	0.981
	40	0.01116	0.02827	0.02838	0.00794	0.02842	0.02848	0.995	0.997	0.00713	0.02888	0.02892	0.979	0.981
	80	0.00483	0.01459	0.01461	0.00256	0.01454	0.01454	1.004	1.005	0.00217	0.01473	0.01473	0.991	0.992
	120	0.00304	0.00956	0.00956	0.00137	0.00951	0.00951	1.004	1.005	0.00119	0.00968	0.00968	0.987	0.988
0.1	20	0.02407	0.05680	0.05736	NA	NA	NA	NA	NA	0.01906	0.05821	0.05856	0.976	0.980
	40	0.01006	0.02797	0.02807	0.00784	0.02826	0.02832	0.990	0.991	0.00709	0.02874	0.02878	0.973	0.975
	80	0.00380	0.01433	0.01434	0.00251	0.01446	0.01446	0.991	0.992	0.00215	0.01467	0.01467	0.977	0.977
	120	0.00210	0.00935	0.00935	0.00132	0.00946	0.00946	0.988	0.988	0.00119	0.00964	0.00963	0.970	0.970
0.2	20	0.02326	0.05797	0.05849	NA	NA	NA	NA	NA	0.01895	0.05881	0.05916	0.986	0.989
	40	0.00959	0.02852	0.02861	0.00770	0.02868	0.02874	0.994	0.996	0.00702	0.02901	0.02905	0.983	0.985
	80	0.00353	0.01459	0.01460	0.00244	0.01466	0.01466	0.995	0.996	0.00212	0.01480	0.01480	0.986	0.986
	120	0.00197	0.00957	0.00957	0.00131	0.00962	0.00962	0.994	0.994	0.00118	0.00974	0.00974	0.983	0.983
0.3	20	0.02180	0.06017	0.06063	NA	NA	NA	NA	NA	0.01878	0.06056	0.06090	0.994	0.996
	40	0.00848	0.02961	0.02968	0.00726	0.02967	0.02972	0.998	0.999	0.00678	0.02986	0.02990	0.992	0.993
	80	0.00287	0.01504	0.01505	0.00219	0.01507	0.01508	0.998	0.998	0.00197	0.01516	0.01516	0.992	0.992
	120	0.00162	0.00991	0.00991	0.00121	0.00993	0.00993	0.998	0.998	0.00113	0.01000	0.01000	0.991	0.991

จากตารางที่ 4.15 พบว่า การประมาณค่าพารามิเตอร์ μ ด้วยวิธี GEPD แบบ Trimmed 5% จะมีประสิทธิภาพมากกว่าวิธี GE และวิธี GEPD แบบ Trimmed 10% เมื่อสัดส่วนของการตัดปลาย p = 0 และขนาดตัวอย่าง $n \geq 80$ สำหรับกรณีอื่นๆ นั้นการประมาณแบบ GE จะมีประสิทธิภาพมากกว่า

ตารางที่ 4.16 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงแบบค่าสูงสุดขีด ที่พารามิเตอร์ $\mu = 0$ และ $\sigma = 1$ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา——Trimmed 5% และTrimmed 10% ตามลำดับ

p	n	($\hat{\sigma}$) → GE			($\hat{\sigma}$) → Trimmed 5%					($\hat{\sigma}$) → Trimmed 10%				
		Bias	Var	MSE	Bias	Var	MSE	RE1	RE2	Bias	Var	MSE	RE1	RE2
0	20	-0.02433	0.04783	0.04842	NA	NA	NA	NA	NA	-0.00999	0.04617	0.04626	1.036	1.047
	40	-0.01557	0.02443	0.02467	-0.00702	0.02221	0.02226	1.100	1.108	-0.00544	0.02267	0.02270	1.078	1.087
	80	-0.01005	0.01269	0.01279	-0.00404	0.01134	0.01135	1.119	1.126	-0.00329	0.01163	0.01164	1.091	1.099
	120	-0.00707	0.00835	0.00840	-0.00237	0.00734	0.00734	1.138	1.143	-0.00186	0.00748	0.00749	1.115	1.122
0.1	20	-0.01410	0.03881	0.03900	NA	NA	NA	NA	NA	-0.00769	0.04541	0.04546	0.855	0.858
	40	-0.00854	0.01879	0.01886	-0.00553	0.02071	0.02074	0.907	0.909	-0.00461	0.02226	0.02228	0.844	0.846
	80	-0.00506	0.00956	0.00958	-0.00326	0.01058	0.01059	0.903	0.905	-0.00282	0.01142	0.01143	0.837	0.838
	120	-0.00314	0.00623	0.00624	-0.00203	0.00691	0.00692	0.902	0.903	-0.00187	0.00744	0.00744	0.838	0.839
0.2	20	-0.01763	0.04094	0.04124	NA	NA	NA	NA	NA	-0.00869	0.04879	0.04886	0.839	0.844
	40	-0.01044	0.01965	0.01976	-0.00619	0.02202	0.02206	0.892	0.896	-0.00486	0.02386	0.02388	0.824	0.827
	80	-0.00615	0.01006	0.01010	-0.00359	0.01131	0.01132	0.889	0.891	-0.00295	0.01232	0.01233	0.817	0.819
	120	-0.00367	0.00657	0.00658	-0.00206	0.00742	0.00742	0.885	0.887	-0.00182	0.00806	0.00806	0.815	0.816
0.3	20	-0.02144	0.04548	0.04593	NA	NA	NA	NA	NA	-0.00943	0.05549	0.05557	0.820	0.827
	40	-0.01328	0.02195	0.02212	-0.00764	0.02498	0.02503	0.879	0.884	-0.00589	0.02729	0.02732	0.804	0.810
	80	-0.00783	0.01114	0.01120	-0.00441	0.01271	0.01273	0.877	0.880	-0.00359	0.01397	0.01398	0.798	0.801
	120	-0.00457	0.00729	0.00731	-0.00240	0.00839	0.00840	0.869	0.871	-0.00209	0.00920	0.00921	0.793	0.795

จากตารางที่ 4.16 พบว่า การประมาณค่าพารามิเตอร์ σ ด้วยวิธี GEPD แบบ Trimmed q% จะมีประสิทธิภาพมากกว่าวิธี GE เมื่อ $p = 0$ โดยแบบ Trimmed 5% จะมีประสิทธิภาพมากกว่าแบบ Trimmed 10% ในทุกกรณี

ตารางที่ 4.17 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงแบบค่าสูงสุดขีด ที่พารามิเตอร์ $\mu = 0$ และ $\sigma = 1$ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— Trimmed 5% & 4 Cluster-Mean และ Trimmed 5% & 8 Cluster-Mean ตามลำดับ

p	n	($\hat{\mu}$) → GE			($\hat{\mu}$) → Trimmed 5% & 4 Cluster-Mean					($\hat{\mu}$) → Trimmed 5% & 8 Cluster-Mean				
		Bias	Var	MSE	Bias	Var	MSE	RE1	RE2	Bias	Var	MSE	RE1	RE2
0	20	0.02488	0.05712	0.05772	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	40	0.01116	0.02827	0.02838	0.00763	0.02798	0.02803	1.010	1.013	0.00795	0.02804	0.02809	1.008	1.010
	80	0.00483	0.01459	0.01461	0.00237	0.01434	0.01434	1.018	1.019	0.00259	0.01436	0.01436	1.016	1.018
	120	0.00304	0.00956	0.00956	0.00124	0.00937	0.00937	1.020	1.021	0.00136	0.00937	0.00937	1.020	1.020
0.1	20	0.02407	0.05680	0.05736	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	40	0.01006	0.02797	0.02807	0.00789	0.02808	0.02814	0.996	0.997	0.00798	0.02799	0.02804	1.000	1.001
	80	0.00380	0.01433	0.01434	0.00251	0.01439	0.01440	0.995	0.996	0.00260	0.01434	0.01434	0.999	1.000
	120	0.00210	0.00935	0.00935	0.00129	0.00939	0.00939	0.995	0.995	0.00137	0.00934	0.00934	1.001	1.001
0.2	20	0.02326	0.05797	0.05849	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	40	0.00959	0.02852	0.02861	0.00796	0.02836	0.02842	1.006	1.007	0.00799	0.02839	0.02845	1.005	1.006
	80	0.00353	0.01459	0.01460	0.00258	0.01456	0.01457	1.002	1.002	0.00263	0.01455	0.01455	1.002	1.003
	120	0.00197	0.00957	0.00957	0.00138	0.00953	0.00953	1.004	1.005	0.00143	0.00952	0.00952	1.005	1.005
0.3	20	0.02180	0.06017	0.06063	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	40	0.00848	0.02961	0.02968	0.00748	0.02956	0.02961	1.002	1.002	0.00762	0.02937	0.02942	1.008	1.009
	80	0.00287	0.01504	0.01505	0.00229	0.01502	0.01502	1.001	1.001	0.00244	0.01496	0.01496	1.006	1.006
	120	0.00162	0.00991	0.00991	0.00127	0.00989	0.00989	1.002	1.002	0.00130	0.00983	0.00983	1.008	1.008

จากตารางที่ 4.17 พบว่า ในกรณีส่วนใหญ่การประมาณค่าพารามิเตอร์ μ ด้วยวิธี GEPD แบบTrimmed 5% & K Cluster-Mean มีประสิทธิภาพมากกว่าวิธี GE (ยกเว้น p =0.1) โดยเมื่อ p = 0 วิธี GEPD แบบ Trimmed 5% & 4 Cluster-Mean จะมีประสิทธิภาพมากกว่าแบบ Trimmed 5% & 8 Cluster-Mean และเมื่อ $p \neq 0$ แล้ว วิธี GEPD แบบ Trimmed 5% & 8 Cluster-Mean จะมีประสิทธิภาพมากกว่าแบบ Trimmed 5% & 4 Cluster-Mean เป็นส่วนใหญ่

ตารางที่ 4.18 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงแบบค่าสูงสุดขีด ที่พารามิเตอร์ μ = 0 และ σ = 1 โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— Trimmed 5% & 4 Cluster-Mean และ Trimmed 5% & 8 Cluster-Mean ตามลำดับ

p	n	($\hat{\sigma}$) → GE			($\hat{\sigma}$) → Trimmed 5% & 4 Cluster-Mean					($\hat{\sigma}$) → Trimmed 5% & 8 Cluster-Mean				
		Bias	Var	MSE	Bias	Var	MSE	RE1	RE2	Bias	Var	MSE	RE1	RE2
0	20	-0.02433	0.04783	0.04842	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	40	-0.01557	0.02443	0.02467	-0.00633	0.02148	0.02151	1.137	1.147	-0.00704	0.02223	0.02228	1.099	1.107
	80	-0.01005	0.01269	0.01279	-0.00364	0.01099	0.01100	1.155	1.163	-0.00409	0.01135	0.01136	1.119	1.126
	120	-0.00707	0.00835	0.00840	-0.00209	0.00711	0.00711	1.174	1.181	-0.00231	0.00735	0.00735	1.136	1.142
0.1	20	-0.01410	0.03881	0.03900	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	40	-0.00854	0.01879	0.01886	-0.00510	0.02219	0.02221	0.847	0.849	-0.00578	0.02278	0.02281	0.825	0.827
	80	-0.00506	0.00956	0.00958	-0.00298	0.01137	0.01137	0.841	0.842	-0.00352	0.01146	0.01147	0.834	0.835
	120	-0.00314	0.00623	0.00624	-0.00183	0.00743	0.00743	0.839	0.840	-0.00205	0.00749	0.00750	0.832	0.833
0.2	20	-0.01763	0.04094	0.04124	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	40	-0.01044	0.01965	0.01976	-0.00514	0.02427	0.02429	0.810	0.813	-0.00582	0.02212	0.02215	0.888	0.892
	80	-0.00615	0.01006	0.01010	-0.00288	0.01245	0.01246	0.808	0.810	-0.00337	0.01138	0.01139	0.884	0.887
	120	-0.00367	0.00657	0.00658	-0.00152	0.00820	0.00820	0.801	0.802	-0.00190	0.00747	0.00747	0.879	0.881
0.3	20	-0.02144	0.04548	0.04593	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	40	-0.01328	0.02195	0.02212	-0.00701	0.02671	0.02675	0.822	0.827	-0.00715	0.02473	0.02478	0.888	0.893
	80	-0.00783	0.01114	0.01120	-0.00399	0.01359	0.01360	0.820	0.824	-0.00412	0.01263	0.01265	0.882	0.886
	120	-0.00457	0.00729	0.00731	-0.00205	0.00899	0.00899	0.812	0.813	-0.00239	0.00832	0.00833	0.877	0.878

จากตารางที่ 4.18 พบว่า โดยรวมการประมาณค่าพารามิเตอร์ σ ด้วยวิธี GE มีประสิทธิภาพมากกว่าวิธี GEPD แบบ Trimmed 5% & K Cluster-Mean (ยกเว้น p = 0) โดยเมื่อ p = 0 แล้ว วิธีประมาณแบบ Trimmed 5% & 4 Cluster-Mean จะมีประสิทธิภาพมากที่สุด

เมื่อ $p \leq 0.1$ แล้ว วิธี GEPD แบบ Trimmed 5% & 4 Cluster-Mean จะมีประสิทธิภาพมากกว่าแบบ Trimmed 5% & 8 Cluster-Mean แต่ถ้า $p > 0.1$ แล้ว แบบ Trimmed 5% & 8 Cluster-Mean จะมีประสิทธิภาพมากกว่าแบบ แบบ Trimmed 5% & 4 Cluster-Mean

ตารางที่ 4.19 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงแบบค่าสูงสุดขีด ที่พารามิเตอร์ μ = 0 และ σ = 1 โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— Trimmed 10% & 4 Cluster-Mean และ Trimmed 10% & 8 Cluster-Mean ตามลำดับ

p	n	($\hat{\mu}$) → GE			($\hat{\mu}$) → Trimmed 10% & 4 Cluster-Mean					($\hat{\mu}$) → Trimmed 10% & 8 Cluster-Mean				
		Bias	Var	MSE	Bias	Var	MSE	RE1	RE2	Bias	Var	MSE	RE1	RE2
0	20	0.02488	0.05712	0.05772	0.01890	0.05747	0.05782	0.994	0.998	NA	NA	NA	NA	NA
	40	0.01116	0.02827	0.02838	0.00713	0.02835	0.02839	0.997	1.000	0.00730	0.02825	0.02830	1.001	1.003
	80	0.00483	0.01459	0.01461	0.00213	0.01451	0.01452	1.005	1.007	0.00234	0.01445	0.01445	1.010	1.011
	120	0.00304	0.00956	0.00956	0.00116	0.00950	0.00950	1.006	1.007	0.00131	0.00945	0.00945	1.011	1.012
0.1	20	0.02407	0.05680	0.05736	0.01920	0.05756	0.05791	0.987	0.991	NA	NA	NA	NA	NA
	40	0.01006	0.02797	0.02807	0.00725	0.02840	0.02845	0.985	0.987	0.00731	0.02822	0.02827	0.991	0.993
	80	0.00380	0.01433	0.01434	0.00220	0.01454	0.01454	0.985	0.986	0.00234	0.01444	0.01444	0.992	0.993
	120	0.00210	0.00935	0.00935	0.00117	0.00952	0.00952	0.982	0.982	0.00130	0.00943	0.00943	0.991	0.992
0.2	20	0.02326	0.05797	0.05849	0.01913	0.05800	0.05835	0.999	1.002	NA	NA	NA	NA	NA
	40	0.00959	0.02852	0.02861	0.00732	0.02861	0.02866	0.997	0.998	0.00735	0.02857	0.02862	0.998	1.000
	80	0.00353	0.01459	0.01460	0.00227	0.01467	0.01468	0.994	0.995	0.00238	0.01462	0.01463	0.997	0.998
	120	0.00197	0.00957	0.00957	0.00125	0.00962	0.00962	0.995	0.995	0.00137	0.00959	0.00959	0.998	0.999
0.3	20	0.02180	0.06017	0.06063	0.01902	0.06023	0.06058	0.999	1.001	NA	NA	NA	NA	NA
	40	0.00848	0.02961	0.02968	0.00701	0.02969	0.02973	0.997	0.998	0.00711	0.02947	0.02952	1.005	1.005
	80	0.00287	0.01504	0.01505	0.00206	0.01509	0.01509	0.997	0.997	0.00224	0.01500	0.01500	1.003	1.003
	120	0.00162	0.00991	0.00991	0.00118	0.00994	0.00994	0.997	0.997	0.00125	0.00986	0.00986	1.005	1.005

ตารางที่ 4.20 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงแบบค่าสูงสุดขีด ที่พารามิเตอร์ μ = 0 และ σ = 1 โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— Trimmed 10% & 4 Cluster-Mean และ Trimmed 10% & 8 Cluster-Mean ตามลำดับ

p	n	($\hat{\sigma}$) → GE			($\hat{\sigma}$) → Trimmed 10% & 4 Cluster-Mean					($\hat{\sigma}$) → Trimmed 10% & 8 Cluster-Mean				
		Bias	Var	MSE	Bias	Var	MSE	RE1	RE2	Bias	Var	MSE	RE1	RE2
0	20	-0.02433	0.04783	0.04842	-0.00926	0.04580	0.04588	1.044	1.055	NA	NA	NA	NA	NA
	40	-0.01557	0.02443	0.02467	-0.00521	0.02253	0.02256	1.084	1.094	-0.00583	0.02292	0.02295	1.066	1.075
	80	-0.01005	0.01269	0.01279	-0.00310	0.01157	0.01158	1.097	1.104	-0.00358	0.01170	0.01171	1.085	1.092
	120	-0.00707	0.00835	0.00840	-0.00174	0.00747	0.00747	1.117	1.124	-0.00193	0.00755	0.00755	1.106	1.112
0.1	20	-0.01410	0.03881	0.03900	-0.00741	0.04737	0.04742	0.819	0.822	NA	NA	NA	NA	NA
	40	-0.00854	0.01879	0.01886	-0.00451	0.02326	0.02327	0.808	0.810	-0.00518	0.02359	0.02362	0.796	0.798
	80	-0.00506	0.00956	0.00958	-0.00269	0.01194	0.01195	0.800	0.802	-0.00328	0.01190	0.01190	0.803	0.805
	120	-0.00314	0.00623	0.00624	-0.00172	0.00779	0.00779	0.801	0.802	-0.00198	0.00776	0.00776	0.804	0.804
0.2	20	-0.01763	0.04094	0.04124	-0.00805	0.05297	0.05303	0.773	0.778	NA	NA	NA	NA	NA
	40	-0.01044	0.01965	0.01976	-0.00443	0.02561	0.02563	0.767	0.771	-0.00487	0.02321	0.02323	0.847	0.851
	80	-0.00615	0.01006	0.01010	-0.00253	0.01317	0.01318	0.764	0.766	-0.00298	0.01198	0.01199	0.840	0.842
	120	-0.00367	0.00657	0.00658	-0.00137	0.00866	0.00866	0.758	0.759	-0.00179	0.00785	0.00785	0.837	0.838
0.3	20	-0.02144	0.04548	0.04593	-0.00886	0.05796	0.05802	0.785	0.792	NA	NA	NA	NA	NA
	40	-0.01328	0.02195	0.02212	-0.00589	0.02872	0.02875	0.764	0.770	-0.00583	0.02629	0.02632	0.835	0.841
	80	-0.00783	0.01114	0.01120	-0.00345	0.01465	0.01466	0.761	0.764	-0.00359	0.01351	0.01352	0.825	0.829
	120	-0.00457	0.00729	0.00731	-0.00183	0.00968	0.00968	0.753	0.755	-0.00225	0.00888	0.00888	0.822	0.824

จากตารางที่ 4.19 พบว่า การประมาณค่าพารามิเตอร์ μ ด้วยวิธี GEPD แบบ Trimmed 10% & 8 Cluster-Mean มีประสิทธิภาพมากกว่าแบบ Trimmed 10% & 4 Cluster-Mean ในทุกกรณี และจะมีประสิทธิภาพมากกว่าวิธี GE เมื่อ p = 0 หรือ p = 0.3

จากตารางที่ 4.20 พบว่า การประมาณค่าพารามิเตอร์ σ ด้วยวิธี GEPD แบบ Trimmed 10% & 4 Cluster-Mean มีประสิทธิภาพมากกว่าวิธี GE และวิธี GEPD แบบ Trimmed 10% & 4 Cluster-Mean เมื่อ p = 0

4.1.3 ผลการศึกษาเมื่อข้อมูลแจกแจงแบบโลจิสติค

ส่วนที่ 1 การเปรียบเทียบประสิทธิภาพของวิธีการประมาณค่าพารามิเตอร์แบบภาวะน่าจะเป็นสูงสุด (MLE) กับแบบใช้กราฟ (GE)

ตารางที่ 4.21 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงแบบโลจิสติค ที่พารามิเตอร์ μ = 0 และ σ = 1 โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา

Censoring	Sample	($\hat{\mu}$) → MLE			($\hat{\mu}$) → GE				
Proportion p	Size n	Bias	Var	MSE	Bias	Var	MSE	RE1	RE2
0	20	-0.00293	0.14540	0.14538	-0.00329	0.16082	0.16080	0.904	0.904
	40	-0.00539	0.07349	0.07351	-0.00626	0.08058	0.08060	0.912	0.912
	80	-0.00530	0.03782	0.03784	-0.00626	0.04198	0.04201	0.901	0.901
	120	-0.00382	0.02461	0.02462	-0.00518	0.02691	0.02693	0.914	0.914
0.1	20	-0.00277	0.14573	0.14571	0.00106	0.15067	0.15064	0.967	0.967
	40	-0.00539	0.07355	0.07357	-0.00379	0.07490	0.07490	0.982	0.982
	80	-0.00527	0.03782	0.03784	-0.00461	0.03874	0.03875	0.976	0.977
	120	-0.00384	0.02463	0.02464	-0.00373	0.02503	0.02504	0.984	0.984
0.2	20	-0.00427	0.14686	0.14685	-0.00262	0.15030	0.15028	0.977	0.977
	40	-0.00586	0.07408	0.07410	-0.00602	0.07454	0.07456	0.994	0.994
	80	-0.00538	0.03817	0.03819	-0.00585	0.03841	0.03844	0.994	0.994
	120	-0.00393	0.02494	0.02496	-0.00412	0.02515	0.02516	0.992	0.992
0.3	20	-0.00678	0.15046	0.15048	-0.00747	0.15995	0.15997	0.941	0.941
	40	-0.00763	0.07580	0.07584	-0.00956	0.07979	0.07987	0.950	0.950
	80	-0.00618	0.03885	0.03888	-0.00788	0.04100	0.04105	0.948	0.947
	120	-0.00448	0.02552	0.02554	-0.00496	0.02707	0.02709	0.943	0.943

จากตารางที่ 4.21 พบว่า การประมาณค่าพารามิเตอร์ μ ด้วยวิธี MLE มีประสิทธิภาพมากกว่าวิธี GE ในทุกกรณี โดยพิจารณาจากค่า |Biases| และค่า Variances ของค่าประมาณพารามิเตอร์ μ ซึ่งวิธี GE มีค่าสูงกว่าวิธี MLE เป็นส่วนใหญ่ ส่งผลให้ค่า MSEs มีค่าสูงขึ้นตามไปด้วย ถ้ากำหนดให้ขนาดตัวอย่าง (n) คงที่พบว่า ค่า RE1 และ RE2 จะมีค่าเพิ่มขึ้นเมื่อสัดส่วนของการตัดปลาย (p) เพิ่มขึ้น (ยกเว้น p =0.3)

ตารางที่ 4.22 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงแบบโลจิสติค ที่พารามิเตอร์ μ = 0 และ σ = 1 โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา

Censoring	Sample	($\hat{\sigma}$) → MLE			($\hat{\sigma}$) → GE				
Proportion p	Size n	Bias	Var	MSE	Bias	Var	MSE	RE1	RE2
0	20	-0.00414	0.03543	0.03544	-0.01837	0.03856	0.03889	0.919	0.911
	40	-0.00256	0.01723	0.01723	-0.01133	0.01889	0.01901	0.912	0.906
	80	-0.00145	0.00871	0.00871	-0.00681	0.00955	0.00959	0.912	0.908
	120	-0.00040	0.00577	0.00577	-0.00399	0.00636	0.00638	0.907	0.905
0.1	20	-0.00426	0.03860	0.03861	-0.01293	0.04607	0.04623	0.838	0.835
	40	-0.00272	0.01870	0.01871	-0.00771	0.02269	0.02274	0.824	0.822
	80	-0.00123	0.00950	0.00950	-0.00415	0.01173	0.01174	0.810	0.809
	120	-0.00021	0.00638	0.00638	-0.00161	0.00788	0.00789	0.809	0.809
0.2	20	-0.00578	0.04404	0.04406	-0.01573	0.05801	0.05825	0.759	0.756
	40	-0.00249	0.02122	0.02122	-0.00934	0.02881	0.02889	0.736	0.734
	80	-0.00075	0.01089	0.01089	-0.00504	0.01517	0.01519	0.718	0.717
	120	-0.00002	0.00723	0.00723	-0.00190	0.01010	0.01010	0.716	0.716
0.3	20	-0.00548	0.05050	0.05052	-0.01880	0.07244	0.07278	0.697	0.694
	40	-0.00278	0.02445	0.02445	-0.01151	0.03656	0.03668	0.669	0.667
	80	-0.00094	0.01273	0.01273	-0.00626	0.01938	0.01942	0.657	0.656
	120	0.00021	0.00840	0.00839	-0.00240	0.01289	0.01290	0.651	0.651

จากตารางที่ 4.22 พบว่า การประมาณค่าพารามิเตอร์ σ ด้วยวิธี MLE มีประสิทธิภาพมากกว่าวิธี GE ในทุกกรณี โดยพิจารณาจากค่า |Biases| และค่า Variances ของค่าประมาณพารามิเตอร์ σ ซึ่งวิธี GE มีค่าสูงกว่าวิธี MLE ในทุกกรณี ส่งผลให้ค่า MSEs มีค่าสูงขึ้นตามไปด้วย ถ้ากำหนดให้สัดส่วนของการตัดปลาย (p) มีขนาดคงที่พบว่า ค่า RE1 และ RE2 จะมีค่าลดลง เมื่อขนาดตัวอย่าง (n) เพิ่มขึ้น และถ้ากำหนดให้ขนาดตัวอย่าง (n) คงที่พบว่า ค่า RE1 และRE2 จะมีค่าลดลง เมื่อสัดส่วนของการตัดปลาย (p) เพิ่มขึ้น

จากตารางที่ 4.21 และ 4.22 เมื่อกำหนดให้ขนาดตัวอย่าง (n) คงที่และสัดส่วนของการตัดปลาย (p) เพิ่มขึ้นพบว่า ค่า RE1 และRE2 ที่ได้ไม่สอดคล้องกัน โดยน่าจะเป็นผลมาจาก (1) ขนาดตัวอย่างที่ลดลง (2) สัดส่วนของการตัดปลายที่เพิ่มขึ้น และ (3) Outlier ที่ถูกตัดออกไป ซึ่งทั้ง 3 ปัจจัยนี้น่าจะส่งผลต่อการเพิ่มขึ้นหรือลดลงของประสิทธิภาพการประมาณ

ส่วนที่ 2 การเปรียบเทียบประสิทธิภาพของวิธีการประมาณค่าพารามิเตอร์แบบใช้กราฟ (GE) กับแบบใช้กราฟที่ใช้ข้อมูลเพียงบางส่วน (GEPD)

ตารางที่ 4.23 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงแบบโลจิสติค ที่พารามิเตอร์ $\mu = 0$ และ $\sigma = 1$ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา ── 4 Cluster-Mean และ 8 Cluster-Mean ตามลำดับ

p	n	($\hat{\mu}$) → GE			($\hat{\mu}$) → 4 Cluster-Mean					($\hat{\mu}$) → 8 Cluster-Mean				
		Bias	Var	MSE	Bias	Var	MSE	RE1	RE2	Bias	Var	MSE	RE1	RE2
0	20	-0.00329	0.16082	0.16080	-0.00329	0.16082	0.16080	1.000	1.000	NA	NA	NA	NA	NA
	40	-0.00626	0.08058	0.08060	-0.00626	0.08058	0.08060	1.000	1.000	-0.00626	0.08058	0.08060	1.000	1.000
	80	-0.00626	0.04198	0.04201	-0.00626	0.04198	0.04201	1.000	1.000	-0.00626	0.04198	0.04201	1.000	1.000
	120	-0.00518	0.02691	0.02693	-0.00518	0.02691	0.02693	1.000	1.000	-0.00518	0.02691	0.02693	1.000	1.000
0.1	20	0.00106	0.15067	0.15064	0.00222	0.15141	0.15139	0.995	0.995	NA	NA	NA	NA	NA
	40	-0.00379	0.07490	0.07490	-0.00279	0.07553	0.07553	0.992	0.992	-0.00317	0.07510	0.07510	0.997	0.997
	80	-0.00461	0.03874	0.03875	-0.00394	0.03911	0.03912	0.990	0.991	-0.00413	0.03890	0.03891	0.996	0.996
	120	-0.00373	0.02503	0.02504	-0.00335	0.02527	0.02528	0.990	0.990	-0.00343	0.02514	0.02514	0.996	0.996
0.2	20	-0.00262	0.15030	0.15028	0.00031	0.15110	0.15107	0.995	0.995	NA	NA	NA	NA	NA
	40	-0.00602	0.07454	0.07456	-0.00373	0.07500	0.07500	0.994	0.994	-0.00461	0.07459	0.07460	0.999	1.000
	80	-0.00585	0.03841	0.03844	-0.00437	0.03861	0.03862	0.995	0.995	-0.00484	0.03840	0.03842	1.000	1.001
	120	-0.00412	0.02515	0.02516	-0.00336	0.02527	0.02527	0.995	0.996	-0.00346	0.02515	0.02516	1.000	1.000
0.3	20	-0.00747	0.15995	0.15997	-0.00352	0.15890	0.15888	1.007	1.007	NA	NA	NA	NA	NA
	40	-0.00956	0.07979	0.07987	-0.00557	0.07886	0.07888	1.012	1.013	-0.00767	0.07927	0.07931	1.007	1.007
	80	-0.00788	0.04100	0.04105	-0.00544	0.04017	0.04019	1.021	1.022	-0.00622	0.04036	0.04039	1.016	1.016
	120	-0.00496	0.02707	0.02709	-0.00364	0.02648	0.02648	1.022	1.023	-0.00385	0.02663	0.02664	1.016	1.017

จากตารางที่ 4.23 พบว่า การประมาณค่าพารามิเตอร์ μ ด้วยวิธี GEPD แบบ K Cluster-Mean มีประสิทธิภาพมากกว่าวิธี GE เมื่อสัดส่วนของการตัดปลาย p =0.3 เป็นที่น่าสังเกตว่าเมื่อข้อมูลไม่ถูกตัดปลาย (p =0) วิธี GEPD แบบ 4 Cluster-Mean, แบบ 8 Cluster-Mean และวิธี GE มีประสิทธิภาพเท่ากัน ทั้งนี้น่าจะเป็นผลมาจากการแจกแจงแบบโลจิสติคเป็นการแจกแจงที่มีลักษณะสมมาตร

ตารางที่ 4.24 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงแบบโลจิสติค ที่พารามิเตอร์ $\mu = 0$ และ $\sigma = 1$ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— 4 Cluster-Mean และ 8 Cluster-Mean ตามลำดับ

p	n	($\hat{\sigma}$) → GE			($\hat{\sigma}$) → 4 Cluster-Mean					($\hat{\sigma}$) → 8 Cluster-Mean				
		Bias	Var	MSE	Bias	Var	MSE	RE1	RE2	Bias	Var	MSE	RE1	RE2
0	20	-0.01837	0.03856	0.03889	-0.01341	0.03672	0.03689	1.050	1.054	NA	NA	NA	NA	NA
	40	-0.01133	0.01889	0.01901	-0.00761	0.01756	0.01762	1.075	1.079	-0.00881	0.01809	0.01817	1.044	1.046
	80	-0.00681	0.00955	0.00959	-0.00412	0.00875	0.00877	1.091	1.094	-0.00491	0.00899	0.00901	1.062	1.064
	120	-0.00399	0.00636	0.00638	-0.00208	0.00579	0.00580	1.098	1.100	-0.00255	0.00595	0.00596	1.069	1.071
0.1	20	-0.01293	0.04607	0.04623	-0.00955	0.04387	0.04395	1.050	1.052	NA	NA	NA	NA	NA
	40	-0.00771	0.02269	0.02274	-0.00486	0.02082	0.02084	1.090	1.091	-0.00595	0.02182	0.02185	1.040	1.041
	80	-0.00415	0.01173	0.01174	-0.00227	0.01056	0.01056	1.111	1.112	-0.00282	0.01095	0.01096	1.071	1.072
	120	-0.00161	0.00788	0.00789	-0.00057	0.00704	0.00704	1.120	1.120	-0.00080	0.00732	0.00732	1.076	1.077
0.2	20	-0.01573	0.05801	0.05825	-0.01088	0.05382	0.05392	1.078	1.080	NA	NA	NA	NA	NA
	40	-0.00934	0.02881	0.02889	-0.00562	0.02575	0.02578	1.119	1.121	-0.00705	0.02726	0.02731	1.057	1.058
	80	-0.00504	0.01517	0.01519	-0.00266	0.01319	0.01319	1.150	1.151	-0.00341	0.01396	0.01397	1.086	1.087
	120	-0.00190	0.01010	0.01010	-0.00067	0.00871	0.00871	1.160	1.160	-0.00083	0.00921	0.00921	1.097	1.097
0.3	20	-0.01880	0.07244	0.07278	-0.01414	0.06785	0.06803	1.068	1.070	NA	NA	NA	NA	NA
	40	-0.01151	0.03656	0.03668	-0.00687	0.03241	0.03245	1.128	1.130	-0.00932	0.03502	0.03510	1.044	1.045
	80	-0.00626	0.01938	0.01942	-0.00345	0.01658	0.01659	1.169	1.170	-0.00435	0.01773	0.01775	1.093	1.094
	120	-0.00240	0.01289	0.01290	-0.00088	0.01089	0.01089	1.184	1.185	-0.00112	0.01171	0.01171	1.101	1.101

จากตารางที่ 4.24 พบว่า โดยรวมแล้วการประมาณค่าพารามิเตอร์ σ ด้วยวิธี GEPD แบบ K Cluster-Mean มีประสิทธิภาพมากกว่าวิธี GE ในทุกกรณี โดยวิธี GEPD แบบ 4 Cluster-Mean ก็มีประสิทธิภาพมากกว่าแบบ 8 Cluster-Mean และถ้ากำหนดให้สัดส่วนของการตัดปลาย (p) มีขนาดคงที่แล้ว ค่า RE1 และ RE2 จะมีค่าเพิ่มขึ้น เมื่อขนาดตัวอย่าง (n) มากขึ้น

ตารางที่ 4.25 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงแบบโลจิสติค ที่พารามิเตอร์ $\mu = 0$ และ $\sigma = 1$ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— Trimmed 5% และ Trimmed 10% ตามลำดับ

p	n	($\hat{\mu}$) → GE			($\hat{\mu}$) → Trimmed 5%					($\hat{\mu}$) → Trimmed 10%				
		Bias	Var	MSE	Bias	Var	MSE	RE1	RE2	Bias	Var	MSE	RE1	RE2
0	20	-0.00329	0.16082	0.16080	NA	NA	NA	NA	NA	-0.00343	0.15086	0.15084	1.066	1.066
	40	-0.00626	0.08058	0.08060	-0.00605	0.07659	0.07661	1.052	1.052	-0.00581	0.07494	0.07495	1.075	1.075
	80	-0.00626	0.04198	0.04201	-0.00570	0.03981	0.03983	1.054	1.055	-0.00548	0.03884	0.03887	1.081	1.081
	120	-0.00518	0.02691	0.02693	-0.00426	0.02558	0.02559	1.052	1.052	-0.00388	0.02505	0.02506	1.074	1.074
0.1	20	0.00106	0.15067	0.15064	NA	NA	NA	NA	NA	-0.00201	0.14941	0.14938	1.008	1.008
	40	-0.00379	0.07490	0.07490	-0.00491	0.07457	0.07458	1.004	1.004	-0.00517	0.07424	0.07425	1.009	1.009
	80	-0.00461	0.03874	0.03875	-0.00504	0.03856	0.03858	1.005	1.004	-0.00511	0.03835	0.03837	1.010	1.010
	120	-0.00373	0.02503	0.02504	-0.00385	0.02493	0.02494	1.004	1.004	-0.00377	0.02484	0.02485	1.007	1.007
0.2	20	-0.00262	0.15030	0.15028	NA	NA	NA	NA	NA	-0.00180	0.15032	0.15029	1.000	1.000
	40	-0.00602	0.07454	0.07456	-0.00485	0.07455	0.07456	1.000	1.000	-0.00484	0.07455	0.07456	1.000	1.000
	80	-0.00585	0.03841	0.03844	-0.00493	0.03838	0.03840	1.001	1.001	-0.00493	0.03839	0.03840	1.001	1.001
	120	-0.00412	0.02515	0.02516	-0.00350	0.02512	0.02513	1.001	1.001	-0.00350	0.02512	0.02513	1.001	1.001
0.3	20	-0.00747	0.15995	0.15997	NA	NA	NA	NA	NA	-0.00148	0.15820	0.15818	1.011	1.011
	40	-0.00956	0.07979	0.07987	-0.00537	0.07876	0.07877	1.013	1.014	-0.00498	0.07848	0.07849	1.017	1.018
	80	-0.00788	0.04100	0.04105	-0.00517	0.04013	0.04014	1.022	1.023	-0.00507	0.04003	0.04004	1.024	1.025
	120	-0.00496	0.02707	0.02709	-0.00331	0.02645	0.02646	1.023	1.024	-0.00341	0.02636	0.02637	1.027	1.027

ตารางที่ 4.26 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงแบบโลจิสติค ที่พารามิเตอร์ $\mu = 0$ และ $\sigma = 1$ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— Trimmed 5% และ Trimmed 10% ตามลำดับ

p	n	($\hat{\sigma}$) → GE			($\hat{\sigma}$) → Trimmed 5%					($\hat{\sigma}$) → Trimmed 10%				
		Bias	Var	MSE	Bias	Var	MSE	RE1	RE2	Bias	Var	MSE	RE1	RE2
0	20	-0.01837	0.03856	0.03889	NA	NA	NA	NA	NA	-0.00366	0.04016	0.04016	0.960	0.968
	40	-0.01133	0.01889	0.01901	-0.00283	0.01850	0.01851	1.021	1.027	-0.00147	0.01929	0.01929	0.979	0.986
	80	-0.00681	0.00955	0.00959	-0.00137	0.00917	0.00917	1.041	1.046	-0.00084	0.00970	0.00970	0.985	0.990
	120	-0.00399	0.00636	0.00638	-0.00027	0.00610	0.00610	1.043	1.046	-0.00013	0.00642	0.00642	0.991	0.994
0.1	20	-0.01293	0.04607	0.04623	NA	NA	NA	NA	NA	-0.00163	0.04347	0.04346	1.060	1.064
	40	-0.00771	0.02269	0.02274	-0.00135	0.02095	0.02095	1.083	1.086	-0.00057	0.02082	0.02081	1.090	1.093
	80	-0.00415	0.01173	0.01174	-0.00052	0.01044	0.01044	1.124	1.125	-0.00032	0.01051	0.01051	1.116	1.117
	120	-0.00161	0.00788	0.00789	0.00029	0.00699	0.00699	1.127	1.128	0.00004	0.00701	0.00701	1.125	1.125
0.2	20	-0.01573	0.05801	0.05825	NA	NA	NA	NA	NA	-0.00138	0.05330	0.05329	1.088	1.093
	40	-0.00934	0.02881	0.02889	-0.00125	0.02605	0.02605	1.106	1.109	-0.00017	0.02546	0.02546	1.131	1.135
	80	-0.00504	0.01517	0.01519	-0.00039	0.01310	0.01309	1.158	1.160	-0.00010	0.01301	0.01301	1.165	1.167
	120	-0.00190	0.01010	0.01010	0.00062	0.00867	0.00867	1.165	1.165	0.00033	0.00861	0.00861	1.173	1.174
0.3	20	-0.01880	0.07244	0.07278	NA	NA	NA	NA	NA	-0.00109	0.06624	0.06622	1.094	1.099
	40	-0.01151	0.03656	0.03668	-0.00164	0.03282	0.03281	1.114	1.118	-0.00028	0.03187	0.03187	1.147	1.151
	80	-0.00626	0.01938	0.01942	-0.00057	0.01647	0.01647	1.177	1.179	-0.00022	0.01630	0.01629	1.189	1.192
	120	-0.00240	0.01289	0.01290	0.00076	0.01089	0.01088	1.184	1.185	0.00041	0.01077	0.01077	1.197	1.197

จากตารางที่ 4.25 พบว่า โดยรวมการประมาณค่าพารามิเตอร์ μ ด้วยวิธี GEPD แบบ Trimmed q% มีประสิทธิภาพมากกว่าวิธี GE ในทุกกรณี โดยวิธี GEPD แบบ Trimmed 10% จะมีประสิทธิภาพมากกว่าแบบ Trimmed 5% เมื่อสัดส่วนของการตัดปลายมีขนาดเล็ก ($p \leq 0.1$)

จากตารางที่ 4.26 พบว่า การประมาณค่าพารามิเตอร์ σ ด้วยวิธี GEPD แบบ Trimmed 5% มีประสิทธิภาพมากกว่าวิธี GE ในทุกกรณี โดยวิธี GEPD แบบ Trimmed 5% จะมีประสิทธิภาพมากกว่าการ Trimmed 10% เมื่อข้อมูลไม่ถูกตัดปลายทางขวา (p =0) แต่ถ้าข้อมูลถูกตัดปลายทางขวาด้วยสัดส่วนของการตัดปลาย $p \geq 0.1$ แล้ว การประมาณด้วยวิธี GEPD แบบ Trimmed 10% จะมีประสิทธิภาพมากกว่าแบบ Trimmed 5% และวิธี GE ทั้งนี้ประสิทธิภาพจะมากยิ่งขึ้นเมื่อสัดส่วนของการตัดปลาย (p) และขนาดของตัวอย่าง (n) เพิ่มมากขึ้น

ตารางที่ 4.27 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงแบบโลจิสติค ที่พารามิเตอร์ μ = 0 และ σ = 1 โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— Trimmed 5% & 4 Cluster-Mean และ Trimmed 5% & 8 Cluster-Mean ตามลำดับ

p	n	($\hat{\mu}$) → GE			($\hat{\mu}$) → Trimmed 5% & 4 Cluster-Mean					($\hat{\mu}$) → Trimmed 5% & 8 Cluster-Mean				
		Bias	Var	MSE	Bias	Var	MSE	RE1	RE2	Bias	Var	MSE	RE1	RE2
0	20	-0.00329	0.16082	0.16080	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	40	-0.00626	0.08058	0.08060	-0.00607	0.07729	0.07732	1.042	1.042	-0.00628	0.07815	0.07818	1.031	1.031
	80	-0.00626	0.04198	0.04201	-0.00573	0.04021	0.04024	1.044	1.044	-0.00581	0.04068	0.04071	1.032	1.032
	120	-0.00518	0.02691	0.02693	-0.00430	0.02581	0.02582	1.043	1.043	-0.00435	0.02608	0.02609	1.032	1.032
0.1	20	0.00106	0.15067	0.15064	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	40	-0.00379	0.07490	0.07490	-0.00483	0.07626	0.07626	0.982	0.982	-0.00533	0.07765	0.07766	0.965	0.965
	80	-0.00461	0.03874	0.03875	-0.00502	0.03954	0.03956	0.980	0.980	-0.00535	0.04028	0.04031	0.962	0.961
	120	-0.00373	0.02503	0.02504	-0.00393	0.02544	0.02545	0.984	0.984	-0.00408	0.02577	0.02578	0.971	0.971
0.2	20	-0.00262	0.15030	0.15028	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	40	-0.00602	0.07454	0.07456	-0.00447	0.07656	0.07657	0.974	0.974	-0.00458	0.07497	0.07497	0.994	0.994
	80	-0.00585	0.03841	0.03844	-0.00470	0.03972	0.03974	0.967	0.967	-0.00479	0.03878	0.03879	0.991	0.991
	120	-0.00412	0.02515	0.02516	-0.00351	0.02579	0.02579	0.975	0.975	-0.00352	0.02525	0.02526	0.996	0.996
0.3	20	-0.00747	0.15995	0.15997	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	40	-0.00956	0.07979	0.07987	-0.00498	0.07845	0.07846	1.017	1.018	-0.00479	0.07742	0.07742	1.031	1.032
	80	-0.00788	0.04100	0.04105	-0.00490	0.03992	0.03994	1.027	1.028	-0.00487	0.03959	0.03961	1.036	1.036
	120	-0.00496	0.02707	0.02709	-0.00327	0.02628	0.02629	1.030	1.030	-0.00343	0.02601	0.02602	1.041	1.041

จากตารางที่ 4.27 พบว่า การประมาณค่าพารามิเตอร์ μ ด้วยวิธี GEPD แบบ Trimmed 5% & K Cluster-Mean มีประสิทธิภาพมากกว่าวิธี GE เมื่อสัดส่วนของการตัดปลาย p =0 หรือ p =0.3 โดยหากพิจารณาที่ p =0 จะพบว่า วิธี Trimmed 5% & 4 Cluster-Mean มีประสิทธิภาพมากกว่าวิธี Trimmed 5% & 8 Cluster-Mean แต่ถ้าพิจารณาที่ p =0.3 จะพบว่า วิธี Trimmed 5% & 8 Cluster-Mean มีประสิทธิภาพมากกว่าวิธี Trimmed 5% & 4 Cluster-Mean

ตารางที่ 4.28 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงแบบโลจิสติค ที่พารามิเตอร์ $\mu = 0$ และ $\sigma = 1$ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— Trimmed 5% & 4 Cluster-Mean และ Trimmed 5% & 8 Cluster-Mean ตามลำดับ

p	n	($\hat{\sigma}$) → GE			($\hat{\sigma}$) → Trimmed 5% & 4 Cluster-Mean					($\hat{\sigma}$) → Trimmed 5% & 8 Cluster-Mean				
		Bias	Var	MSE	Bias	Var	MSE	RE1	RE2	Bias	Var	MSE	RE1	RE2
0	20	-0.01837	0.03856	0.03889	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	40	-0.01133	0.01889	0.01901	-0.00239	0.01846	0.01846	1.023	1.030	-0.00273	0.01865	0.01866	1.013	1.019
	80	-0.00681	0.00955	0.00959	-0.00108	0.00924	0.00924	1.034	1.039	-0.00140	0.00925	0.00925	1.032	1.037
	120	-0.00399	0.00636	0.00638	-0.00014	0.00614	0.00614	1.036	1.038	-0.00028	0.00615	0.00615	1.035	1.037
0.1	20	-0.01293	0.04607	0.04623	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	40	-0.00771	0.02269	0.02274	-0.00121	0.02027	0.02027	1.119	1.122	-0.00162	0.02014	0.02014	1.127	1.129
	80	-0.00415	0.01173	0.01174	-0.00039	0.01019	0.01019	1.151	1.152	-0.00083	0.00991	0.00991	1.183	1.185
	120	-0.00161	0.00788	0.00789	0.00024	0.00683	0.00683	1.154	1.155	0.00006	0.00668	0.00668	1.180	1.180
0.2	20	-0.01573	0.05801	0.05825	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	40	-0.00934	0.02881	0.02889	-0.00092	0.02300	0.02299	1.253	1.256	-0.00093	0.02460	0.02459	1.171	1.175
	80	-0.00504	0.01517	0.01519	-0.00013	0.01164	0.01163	1.303	1.305	-0.00032	0.01238	0.01237	1.225	1.227
	120	-0.00190	0.01010	0.01010	0.00061	0.00776	0.00776	1.301	1.301	0.00054	0.00820	0.00820	1.231	1.232
0.3	20	-0.01880	0.07244	0.07278	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	40	-0.01151	0.03656	0.03668	-0.00138	0.03114	0.03114	1.174	1.178	-0.00111	0.03156	0.03156	1.158	1.162
	80	-0.00626	0.01938	0.01942	-0.00031	0.01584	0.01583	1.224	1.226	-0.00041	0.01590	0.01590	1.219	1.222
	120	-0.00240	0.01289	0.01290	0.00076	0.01051	0.01051	1.227	1.227	0.00060	0.01049	0.01049	1.229	1.229

จากตารางที่ 4.28 พบว่า โดยรวมการประมาณค่าพารามิเตอร์ σ ด้วยวิธี GEPD แบบ Trimmed 5% & K Cluster-Mean มีประสิทธิภาพมากกว่าวิธี GE ในทุกกรณี และเมื่อ p =0.1 การประมาณด้วยวิธี GEPD แบบ Trimmed 5% & 8 Cluster-Mean มีประสิทธิภาพมากกว่าแบบ Trimmed 5% & 4 Cluster-Mean ในทุกขนาดตัวอย่าง (n) ซึ่งสังเกตได้จากค่า RE1 และ RE2 ที่สูงกว่า

ตารางที่ 4.29 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงแบบโลจิสติค ที่พารามิเตอร์ $\mu = 0$ และ $\sigma = 1$ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— Trimmed 10% & 4 Cluster-Mean และ Trimmed 10% & 8 Cluster-Mean ตามลำดับ

p	n	($\hat{\mu}$) → GE			($\hat{\mu}$) → Trimmed 10% & 4 Cluster-Mean					($\hat{\mu}$) → Trimmed 10% & 8 Cluster-Mean				
		Bias	Var	MSE	Bias	Var	MSE	RE1	RE2	Bias	Var	MSE	RE1	RE2
0	20	-0.00329	0.16082	0.16080	-0.00365	0.15302	0.15300	1.051	1.051	NA	NA	NA	NA	NA
	40	-0.00626	0.08058	0.08060	-0.00583	0.07599	0.07601	1.060	1.060	-0.00631	0.07730	0.07732	1.042	1.042
	80	-0.00626	0.04198	0.04201	-0.00552	0.03947	0.03949	1.063	1.064	-0.00567	0.04019	0.04021	1.045	1.045
	120	-0.00518	0.02691	0.02693	-0.00392	0.02539	0.02541	1.060	1.060	-0.00395	0.02578	0.02579	1.044	1.044
0.1	20	0.00106	0.15067	0.15064	-0.00208	0.15219	0.15216	0.990	0.990	NA	NA	NA	NA	NA
	40	-0.00379	0.07490	0.07490	-0.00516	0.07564	0.07566	0.990	0.990	-0.00582	0.07722	0.07724	0.970	0.970
	80	-0.00461	0.03874	0.03875	-0.00511	0.03919	0.03921	0.988	0.988	-0.00543	0.04003	0.04005	0.968	0.968
	120	-0.00373	0.02503	0.02504	-0.00382	0.02528	0.02529	0.990	0.990	-0.00392	0.02566	0.02567	0.975	0.975
0.2	20	-0.00262	0.15030	0.15028	-0.00200	0.15340	0.15337	0.980	0.980	NA	NA	NA	NA	NA
	40	-0.00602	0.07454	0.07456	-0.00476	0.07606	0.07607	0.980	0.980	-0.00486	0.07493	0.07493	0.995	0.995
	80	-0.00585	0.03841	0.03844	-0.00479	0.03943	0.03945	0.974	0.974	-0.00482	0.03873	0.03874	0.992	0.992
	120	-0.00412	0.02515	0.02516	-0.00342	0.02564	0.02564	0.981	0.981	-0.00344	0.02524	0.02524	0.996	0.997
0.3	20	-0.00747	0.15995	0.15997	-0.00080	0.15775	0.15772	1.014	1.014	NA	NA	NA	NA	NA
	40	-0.00956	0.07979	0.07987	-0.00498	0.07836	0.07837	1.018	1.019	-0.00473	0.07739	0.07740	1.031	1.032
	80	-0.00788	0.04100	0.04105	-0.00489	0.03987	0.03989	1.028	1.029	-0.00483	0.03958	0.03960	1.036	1.037
	120	-0.00496	0.02707	0.02709	-0.00325	0.02625	0.02626	1.031	1.032	-0.00344	0.02600	0.02600	1.041	1.042

ตารางที่ 4.30 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงแบบโลจิสติค ที่พารามิเตอร์ $\mu = 0$ และ $\sigma = 1$ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— Trimmed 10% & 4 Cluster-Mean และ Trimmed 10% & 8 Cluster-Mean ตามลำดับ

p	n	($\hat{\sigma}$) → GE			($\hat{\sigma}$) → Trimmed 10% & 4 Cluster-Mean					($\hat{\sigma}$) → Trimmed 10% & 8 Cluster-Mean				
		Bias	Var	MSE	Bias	Var	MSE	RE1	RE2	Bias	Var	MSE	RE1	RE2
0	20	-0.01837	0.03856	0.03889	-0.00303	0.04069	0.04069	0.948	0.956	NA	NA	NA	NA	NA
	40	-0.01133	0.01889	0.01901	-0.00146	0.01954	0.01954	0.967	0.973	-0.00162	0.01948	0.01948	0.970	0.976
	80	-0.00681	0.00955	0.00959	-0.00069	0.00988	0.00988	0.967	0.971	-0.00104	0.00973	0.00973	0.981	0.986
	120	-0.00399	0.00636	0.00638	-0.00002	0.00655	0.00655	0.971	0.974	-0.00019	0.00644	0.00644	0.987	0.990
0.1	20	-0.01293	0.04607	0.04623	-0.00139	0.04345	0.04344	1.060	1.064	NA	NA	NA	NA	NA
	40	-0.00771	0.02269	0.02274	-0.00078	0.02080	0.02080	1.091	1.094	-0.00100	0.02050	0.02050	1.107	1.109
	80	-0.00415	0.01173	0.01174	-0.00027	0.01055	0.01054	1.112	1.114	-0.00073	0.01017	0.01017	1.153	1.155
	120	-0.00161	0.00788	0.00789	0.00010	0.00704	0.00704	1.120	1.120	-0.00015	0.00682	0.00682	1.156	1.156
0.2	20	-0.01573	0.05801	0.05825	-0.00127	0.04993	0.04992	1.162	1.167	NA	NA	NA	NA	NA
	40	-0.00934	0.02881	0.02889	-0.00040	0.02368	0.02368	1.217	1.220	-0.00003	0.02483	0.02483	1.160	1.164
	80	-0.00504	0.01517	0.01519	0.00003	0.01208	0.01208	1.255	1.258	-0.00013	0.01263	0.01263	1.200	1.202
	120	-0.00190	0.01010	0.01010	0.00048	0.00805	0.00805	1.255	1.255	0.00030	0.00835	0.00835	1.209	1.210
0.3	20	-0.01880	0.07244	0.07278	-0.00030	0.06652	0.06651	1.089	1.094	NA	NA	NA	NA	NA
	40	-0.01151	0.03656	0.03668	-0.00063	0.03202	0.03201	1.142	1.146	0.00004	0.03186	0.03186	1.147	1.151
	80	-0.00626	0.01938	0.01942	-0.00008	0.01644	0.01644	1.179	1.181	-0.00016	0.01626	0.01626	1.192	1.194
	120	-0.00240	0.01289	0.01290	0.00059	0.01091	0.01091	1.181	1.182	0.00030	0.01071	0.01071	1.204	1.204

จากตารางที่ 4.29 พบว่า โดยรวมแล้วการประมาณค่าพารามิเตอร์ μ ด้วยวิธี GEPD แบบ Trimmed 10% & K Cluster-Mean มีประสิทธิภาพมากกว่าวิธี GE เมื่อสัดส่วนของการตัดปลาย $p = 0$ หรือ $p = 0.3$ โดยหากพิจารณาที่ $p = 0$ จะพบว่า วิธี GEPD แบบ Trimmed 10% & 4 Cluster-Mean มีประสิทธิภาพมากกว่าแบบ Trimmed 10% & 8 Cluster-Mean แต่ถ้าพิจารณาที่ $p = 0.3$ จะพบว่า วิธี GEPD แบบ Trimmed 10% & 8 Cluster-Mean มีประสิทธิภาพมากกว่าแบบ Trimmed 10% & 4 Cluster-Mean

จากตารางที่ 4.30 พบว่า การประมาณค่าพารามิเตอร์ σ ด้วยวิธี GEPD แบบ Trimmed 10% & K Cluster-Mean มีประสิทธิภาพมากกว่าวิธี GE ในทุกกรณี (ยกเว้น $p = 0$) โดยวิธี GEPD แบบ Trimmed 10% & 4 Cluster-Mean จะมีประสิทธิภาพมากกว่าแบบ Trimmed 10% & 8 Cluster-Mean เมื่อ $p = 0.2$

4.1.4 ผลการศึกษาเมื่อข้อมูลแจกแจงแบบปกติ

ส่วนที่ 1 การเปรียบเทียบประสิทธิภาพของวิธีการประมาณค่าพารามิเตอร์แบบภาวะน่าจะเป็นสูงสุด (MLE) กับแบบใช้กราฟ (GE)

ตารางที่ 4.31 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงแบบปกติ ที่พารามิเตอร์ μ = 0 และ σ = 1 โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา

Censoring	Sample	$(\hat{\mu}) \rightarrow$ MLE			$(\hat{\mu}) \rightarrow$ GE				
Proportion p	Size n	Bias	Var	MSE	Bias	Var	MSE	RE1	RE2
0	20	-0.00046	0.04845	0.04844	-0.00046	0.04845	0.04844	1.000	1.000
	40	0.00021	0.02465	0.02464	0.00021	0.02465	0.02464	1.000	1.000
	80	0.00025	0.01225	0.01225	0.00025	0.01225	0.01225	1.000	1.000
	120	0.00062	0.00806	0.00806	0.00062	0.00806	0.00806	1.000	1.000
0.1	20	0.00010	0.04945	0.04944	0.00263	0.05014	0.05014	0.986	0.986
	40	0.00035	0.02527	0.02527	0.00197	0.02545	0.02545	0.993	0.993
	80	0.00027	0.01247	0.01246	0.00148	0.01264	0.01264	0.986	0.986
	120	0.00076	0.00819	0.00819	0.00134	0.00830	0.00830	0.987	0.986
0.2	20	0.00030	0.05105	0.05104	0.00106	0.05348	0.05347	0.955	0.954
	40	0.00059	0.02581	0.02581	0.00107	0.02694	0.02693	0.958	0.958
	80	0.00015	0.01277	0.01276	0.00124	0.01338	0.01338	0.954	0.954
	120	0.00044	0.00845	0.00845	0.00099	0.00879	0.00879	0.962	0.962
0.3	20	-0.00008	0.05335	0.05334	-0.00239	0.05933	0.05933	0.899	0.899
	40	0.00038	0.02677	0.02677	-0.00078	0.02985	0.02985	0.897	0.897
	80	0.00040	0.01317	0.01316	0.00082	0.01479	0.01479	0.890	0.890
	120	0.00044	0.00867	0.00867	0.00040	0.00970	0.00969	0.894	0.894

จากตารางที่ 4.31 พบว่า การประมาณค่าพารามิเตอร์ μ ด้วยวิธี MLE มีประสิทธิภาพมากกว่าวิธี GE ในทุกกรณี โดยพิจารณาจากค่า |Biases| และค่า Variances ของค่าประมาณพารามิเตอร์ μ ซึ่งวิธี GE มีค่าสูงกว่าวิธี MLE แทบทุกกรณี ส่งผลให้ค่า MSEs มีค่าสูงขึ้นตามไปด้วย เมื่อกำหนดให้ขนาดตัวอย่าง (n) คงที่ พบว่า ค่า RE1 และRE2 จะมีค่าลดลงเมื่อสัดส่วนของการตัดปลาย (p) เพิ่มขึ้น ทั้งนี้เป็นที่น่าสังเกตว่า สำหรับข้อมูลที่ไม่ถูกตัดปลายทางขวาหรือสัดส่วนของการตัดปลาย p = 0 นั้น การประมาณด้วยวิธี MLE และ GE จะมีประสิทธิภาพเท่ากัน เนื่องจากการประมาณค่าพารามิเตอร์ μ ด้วยวิธี MLE และวิธี GE ต่างก็มีค่าเท่ากับ $\bar{x}$ เหมือนกันนั่นเอง

ตารางที่ 4.32 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงแบบปกติ ที่พารามิเตอร์ μ = 0 และ σ = 1 โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา

Censoring	Sample	$(\hat{\sigma}) \rightarrow$ MLE			$(\hat{\sigma}) \rightarrow$ GE				
Proportion p	Size n	Bias	Var	MSE	Bias	Var	MSE	RE1	RE2
0	20	-0.04108	0.02481	0.02650	-0.03268	0.02563	0.02670	0.968	0.992
	40	-0.02067	0.01215	0.01258	-0.01844	0.01229	0.01263	0.989	0.996
	80	-0.00825	0.00617	0.00624	-0.00796	0.00620	0.00626	0.996	0.997
	120	-0.00496	0.00416	0.00418	-0.00503	0.00416	0.00419	0.999	0.999
0.1	20	-0.01619	0.02896	0.02922	-0.02637	0.03209	0.03278	0.902	0.891
	40	-0.00833	0.01399	0.01406	-0.01453	0.01545	0.01565	0.906	0.898
	80	-0.00205	0.00726	0.00726	-0.00508	0.00792	0.00795	0.916	0.914
	120	-0.00065	0.00483	0.00483	-0.00322	0.00529	0.00530	0.912	0.910
0.2	20	-0.01699	0.03314	0.03343	-0.02848	0.03913	0.03994	0.847	0.837
	40	-0.00879	0.01644	0.01651	-0.01571	0.01935	0.01959	0.850	0.843
	80	-0.00232	0.00854	0.00854	-0.00539	0.00978	0.00981	0.873	0.871
	120	-0.00089	0.00571	0.00571	-0.00369	0.00657	0.00658	0.869	0.867
0.3	20	-0.01870	0.03767	0.03801	-0.03242	0.04818	0.04922	0.782	0.772
	40	-0.00983	0.01908	0.01918	-0.01778	0.02425	0.02456	0.787	0.781
	80	-0.00188	0.00968	0.00968	-0.00586	0.01204	0.01207	0.804	0.802
	120	-0.00119	0.00659	0.00659	-0.00434	0.00806	0.00808	0.817	0.815

จากตารางที่ 4.32 พบว่า การประมาณค่าพารามิเตอร์ σ ด้วยวิธี MLE มีประสิทธิภาพมากกว่าวิธี GE ในทุกกรณี โดยพิจารณาจากค่า |Biases| และค่า Variances ของค่าประมาณพารามิเตอร์ σ ซึ่งวิธี GE มีค่าสูงกว่าวิธี MLE เกือบทุกกรณี ส่งผลให้ค่า MSEs มีค่าสูงขึ้นตามไปด้วย ถ้ากำหนดให้สัดส่วนของการตัดปลาย (p) มีขนาดคงที่พบว่า ค่า RE1 และ RE2 จะมีค่าเพิ่มขึ้น เมื่อขนาดตัวอย่าง (n) เพิ่มขึ้น (ยกเว้น n = 120 และ p = 0.1, 0.2) แต่ถ้ากำหนดให้ขนาดตัวอย่าง (n) คงที่ พบว่า ค่า RE1 และRE2 จะมีค่าลดลง เมื่อสัดส่วนของการตัดปลาย (p) เพิ่มขึ้น

ส่วนที่ 2 การเปรียบเทียบประสิทธิภาพของวิธีการประมาณค่าพารามิเตอร์แบบใช้กราฟ (GE) กับแบบใช้กราฟที่ใช้ข้อมูลเพียงบางส่วน (GEPD)

ตารางที่ 4.33 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงแบบปกติ ที่พารามิเตอร์ $\mu = 0$ และ $\sigma = 1$ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— 4 Cluster-Mean และ 8 Cluster-Mean ตามลำดับ

p	n	($\hat{\mu}$) → GE			($\hat{\mu}$) → 4 Cluster-Mean					($\hat{\mu}$) → 8 Cluster-Mean				
		Bias	Var	MSE	Bias	Var	MSE	RE1	RE2	Bias	Var	MSE	RE1	RE2
0	20	-0.00046	0.04845	0.04844	-0.00046	0.04845	0.04844	1.000	1.000	NA	NA	NA	NA	NA
	40	0.00021	0.02465	0.02464	0.00021	0.02465	0.02464	1.000	1.000	0.00021	0.02465	0.02464	1.000	1.000
	80	0.00025	0.01225	0.01225	0.00025	0.01225	0.01225	1.000	1.000	0.00025	0.01225	0.01225	1.000	1.000
	120	0.00062	0.00806	0.00806	0.00062	0.00806	0.00806	1.000	1.000	0.00062	0.00806	0.00806	1.000	1.000
0.1	20	0.00263	0.05014	0.05014	0.00306	0.05015	0.05015	1.000	1.000	NA	NA	NA	NA	NA
	40	0.00197	0.02545	0.02545	0.00238	0.02547	0.02547	0.999	0.999	0.00218	0.02544	0.02544	1.000	1.000
	80	0.00148	0.01264	0.01264	0.00170	0.01265	0.01265	0.999	0.999	0.00165	0.01264	0.01264	1.000	1.000
	120	0.00134	0.00830	0.00830	0.00151	0.00832	0.00832	0.998	0.998	0.00146	0.00831	0.00831	0.999	0.999
0.2	20	0.00106	0.05348	0.05347	0.00213	0.05345	0.05344	1.001	1.001	NA	NA	NA	NA	NA
	40	0.00107	0.02694	0.02693	0.00196	0.02689	0.02689	1.002	1.001	0.00158	0.02690	0.02689	1.001	1.001
	80	0.00124	0.01338	0.01338	0.00175	0.01336	0.01336	1.002	1.001	0.00161	0.01337	0.01337	1.001	1.001
	120	0.00099	0.00879	0.00879	0.00136	0.00880	0.00880	0.999	0.999	0.00127	0.00879	0.00879	1.000	1.000
0.3	20	-0.00239	0.05933	0.05933	-0.00088	0.05915	0.05914	1.003	1.003	NA	NA	NA	NA	NA
	40	-0.00078	0.02985	0.02985	0.00074	0.02969	0.02968	1.006	1.006	-0.00012	0.02977	0.02976	1.003	1.003
	80	0.00082	0.01479	0.01479	0.00173	0.01470	0.01470	1.006	1.006	0.00144	0.01474	0.01474	1.004	1.003
	120	0.00040	0.00970	0.00969	0.00105	0.00968	0.00968	1.002	1.002	0.00089	0.00968	0.00968	1.001	1.001

ตารางที่ 4.34 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงแบบปกติ ที่พารามิเตอร์ $\mu = 0$ และ $\sigma = 1$ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— 4 Cluster-Mean และ 8 Cluster-Mean ตามลำดับ

p	n	($\hat{\sigma}$) → GE			($\hat{\sigma}$) → 4 Cluster-Mean					($\hat{\sigma}$) → 8 Cluster-Mean				
		Bias	Var	MSE	Bias	Var	MSE	RE1	RE2	Bias	Var	MSE	RE1	RE2
0	20	-0.03268	0.02563	0.02670	-0.02863	0.02666	0.02748	0.961	0.972	NA	NA	NA	NA	NA
	40	-0.01844	0.01229	0.01263	-0.01544	0.01277	0.01301	0.962	0.970	-0.01672	0.01243	0.01271	0.988	0.993
	80	-0.00796	0.00620	0.00626	-0.00604	0.00649	0.00653	0.954	0.959	-0.00669	0.00630	0.00635	0.983	0.986
	120	-0.00503	0.00416	0.00419	-0.00359	0.00436	0.00438	0.954	0.957	-0.00402	0.00424	0.00426	0.981	0.983
0.1	20	-0.02637	0.03209	0.03278	-0.02409	0.03283	0.03341	0.978	0.981	NA	NA	NA	NA	NA
	40	-0.01453	0.01545	0.01565	-0.01241	0.01589	0.01604	0.972	0.976	-0.01344	0.01552	0.01570	0.995	0.997
	80	-0.00508	0.00792	0.00795	-0.00395	0.00819	0.00820	0.967	0.969	-0.00425	0.00802	0.00804	0.988	0.989
	120	-0.00322	0.00529	0.00530	-0.00236	0.00548	0.00548	0.967	0.968	-0.00260	0.00535	0.00535	0.990	0.991
0.2	20	-0.02848	0.03913	0.03994	-0.02536	0.03984	0.04047	0.982	0.987	NA	NA	NA	NA	NA
	40	-0.01571	0.01935	0.01959	-0.01314	0.01957	0.01973	0.989	0.993	-0.01424	0.01935	0.01955	1.000	1.002
	80	-0.00539	0.00978	0.00981	-0.00394	0.00993	0.00995	0.985	0.986	-0.00434	0.00984	0.00986	0.994	0.995
	120	-0.00369	0.00657	0.00658	-0.00264	0.00668	0.00668	0.984	0.985	-0.00288	0.00659	0.00660	0.997	0.998
0.3	20	-0.03242	0.04818	0.04922	-0.02932	0.04853	0.04938	0.993	0.997	NA	NA	NA	NA	NA
	40	-0.01778	0.02425	0.02456	-0.01469	0.02429	0.02450	0.998	1.002	-0.01644	0.02417	0.02444	1.003	1.005
	80	-0.00586	0.01204	0.01207	-0.00401	0.01212	0.01213	0.993	0.995	-0.00460	0.01206	0.01208	0.998	0.999
	120	-0.00434	0.00806	0.00808	-0.00302	0.00810	0.00811	0.996	0.997	-0.00336	0.00805	0.00806	1.002	1.003

จากตารางที่ 4.33 พบว่า การประมาณค่าพารามิเตอร์ μ ด้วยวิธี GEPD แบบ 4 Cluster-Mean มีประสิทธิภาพมากกว่าวิธี GE และแบบ 8 Cluster-Mean ในทุกขนาดตัวอย่าง (n) เมื่อ p = 0.3 นอกจากนี้ยังพบว่า เมื่อ p = 0 การประมาณด้วยวิธี 4 Cluster-Mean, 8 Cluster-Mean และวิธี GE นั้นมีประสิทธิภาพเท่ากัน ทั้งนี้น่าจะเป็นผลมาจากการแจกแจงแบบปกติมีลักษณะที่สมมาตร

จากตารางที่ 4.34 พบว่า ในกรณีส่วนใหญ่การประมาณค่าพารามิเตอร์ σ ด้วยวิธี GE มีประสิทธิภาพมากกว่าวิธี GEPD แบบ K Cluster-Mean โดยแบบ 8 Cluster-Mean จะมีประสิทธิภาพมากกว่าแบบ 4 Cluster-Mean ซึ่งจะเห็นว่าผลที่ได้มีความสอดคล้องกันเพราะแบบ 8 Cluster-Mean จะมีจำนวนจุดมากกว่าแบบ 4 Cluster-Mean การประมาณค่าจึงดีขึ้น และหากเพิ่มจำนวนจุดต่อไปอีกเรื่อยๆ การประมาณค่าก็จะลู่เข้าสู่วิธีประมาณแบบ GE นั่นเอง

ตารางที่ 4.35 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงแบบปกติ ที่พารามิเตอร์ μ = 0 และ σ = 1 โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— Trimmed 5% และ Trimmed 10% ตามลำดับ

p	n	($\hat{\mu}$) → GE			($\hat{\mu}$) → Trimmed 5%					($\hat{\mu}$) → Trimmed 10%				
		Bias	Var	MSE	Bias	Var	MSE	RE1	RE2	Bias	Var	MSE	RE1	RE2
0	20	-0.00046	0.04845	0.04844	NA	NA	NA	NA	NA	-2.52E-05	4.96E-02	4.95E-02	0.978	0.978
	40	0.00021	0.02465	0.02464	0.00045	0.02491	0.02490	0.990	0.990	6.26E-04	2.52E-02	2.52E-02	0.977	0.977
	80	0.00025	0.01225	0.01225	0.00071	0.01238	0.01238	0.990	0.989	8.55E-04	1.25E-02	1.25E-02	0.977	0.977
	120	0.00062	0.00806	0.00806	0.00087	0.00810	0.00810	0.995	0.995	1.04E-03	8.20E-03	8.20E-03	0.983	0.983
0.1	20	0.00263	0.05014	0.05014	NA	NA	NA	NA	NA	8.86E-04	5.05E-02	5.05E-02	0.993	0.993
	40	0.00197	0.02545	0.02545	0.00129	0.02552	0.02552	0.997	0.997	1.08E-03	2.56E-02	2.56E-02	0.993	0.993
	80	0.00148	0.01264	0.01264	0.00119	0.01268	0.01268	0.997	0.997	1.16E-03	1.27E-02	1.27E-02	0.993	0.993
	120	0.00134	0.00830	0.00830	0.00115	0.00831	0.00831	0.999	0.999	1.16E-03	8.34E-03	8.34E-03	0.996	0.996
0.2	20	0.00106	0.05348	0.05347	NA	NA	NA	NA	NA	1.10E-03	5.35E-02	5.35E-02	1.000	1.000
	40	0.00107	0.02694	0.02693	0.00135	0.02690	0.02689	1.002	1.002	1.32E-03	2.69E-02	2.69E-02	1.001	1.001
	80	0.00124	0.01338	0.01338	0.00148	0.01336	0.01336	1.002	1.002	1.48E-03	1.34E-02	1.34E-02	1.001	1.001
	120	0.00099	0.00879	0.00879	0.00120	0.00878	0.00878	1.001	1.001	1.21E-03	8.78E-03	8.78E-03	1.000	1.000
0.3	20	-0.00239	0.05933	0.05933	NA	NA	NA	NA	NA	9.03E-05	5.88E-02	5.88E-02	1.010	1.010
	40	-0.00078	0.02985	0.02985	0.00085	0.02956	0.02956	1.010	1.010	1.08E-03	2.94E-02	2.94E-02	1.014	1.014
	80	0.00082	0.01479	0.01479	0.00185	0.01464	0.01464	1.010	1.010	1.92E-03	1.46E-02	1.46E-02	1.015	1.015
	120	0.00040	0.00970	0.00969	0.00120	0.00964	0.00964	1.006	1.006	1.20E-03	9.60E-03	9.60E-03	1.010	1.009

จากตารางที่ 4.35 พบว่า การประมาณค่าพารามิเตอร์ μ ด้วยวิธี GEPD แบบ Trimmed q% จะมีประสิทธิภาพมากกว่าวิธี GE เมื่อสัดส่วนของการตัดปลายมีขนาด $p \geq 0.2$ และเมื่อพิจารณาประสิทธิภาพของการประมาณด้วยวิธี GEPD แบบ Trimmed 5% กับแบบ Trimmed 10% จะพบว่า ที่ p = 0.2 Trimmed 5% มีประสิทธิภาพมากกว่า Trimmed 10% และที่ p = 0.3 Trimmed 10% มีประสิทธิภาพมากกว่า Trimmed 5%

ตารางที่ 4.36 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงแบบปกติ ที่พารามิเตอร์ $\mu = 0$ และ $\sigma = 1$ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— Trimmed 5% และ Trimmed 10% ตามลำดับ

p	n	($\hat{\sigma}$) → GE			($\hat{\sigma}$) → Trimmed 5%					($\hat{\sigma}$) → Trimmed 10%				
		Bias	Var	MSE	Bias	Var	MSE	RE1	RE2	Bias	Var	MSE	RE1	RE2
0	20	-0.03268	0.02563	0.02670	NA	NA	NA	NA	NA	-0.01952	0.03193	0.03231	0.803	0.826
	40	-0.01844	0.01229	0.01263	-0.01172	0.01370	0.01384	0.897	0.912	-0.01008	0.01541	0.01551	0.797	0.814
	80	-0.00796	0.00620	0.00626	-0.00365	0.00705	0.00706	0.879	0.887	-0.00316	0.00789	0.00790	0.786	0.793
	120	-0.00503	0.00416	0.00419	-0.00195	0.00476	0.00477	0.874	0.878	-0.00165	0.00534	0.00534	0.780	0.784
0.1	20	-0.02637	0.03209	0.03278	NA	NA	NA	NA	NA	-0.01739	0.03584	0.03614	0.895	0.907
	40	-0.01453	0.01545	0.01565	-0.00990	0.01613	0.01622	0.958	0.965	-0.00900	0.01713	0.01720	0.902	0.910
	80	-0.00508	0.00792	0.00795	-0.00262	0.00843	0.00844	0.939	0.942	-0.00246	0.00886	0.00887	0.894	0.896
	120	-0.00322	0.00529	0.00530	-0.00134	0.00562	0.00562	0.942	0.944	-0.00137	0.00595	0.00595	0.890	0.892
0.2	20	-0.02848	0.03913	0.03994	NA	NA	NA	NA	NA	-0.01697	0.04385	0.04412	0.893	0.905
	40	-0.01571	0.01935	0.01959	-0.00977	0.02006	0.02015	0.964	0.972	-0.00852	0.02126	0.02133	0.910	0.919
	80	-0.00539	0.00978	0.00981	-0.00215	0.01040	0.01040	0.941	0.943	-0.00186	0.01094	0.01094	0.894	0.896
	120	-0.00369	0.00657	0.00658	-0.00126	0.00695	0.00695	0.945	0.947	-0.00129	0.00739	0.00739	0.889	0.891
0.3	20	-0.03242	0.04818	0.04922	NA	NA	NA	NA	NA	-0.01848	0.05454	0.05488	0.883	0.897
	40	-0.01778	0.02425	0.02456	-0.01043	0.02513	0.02523	0.965	0.973	-0.00887	0.02674	0.02681	0.907	0.916
	80	-0.00586	0.01204	0.01207	-0.00167	0.01279	0.01279	0.941	0.944	-0.00119	0.01353	0.01353	0.890	0.892
	120	-0.00434	0.00806	0.00808	-0.00126	0.00852	0.00852	0.946	0.948	-0.00130	0.00910	0.00910	0.886	0.888

จากตารางที่ 4.36 พบว่า การประมาณค่าพารามิเตอร์ σ ด้วยวิธี GE มีประสิทธิภาพมากกว่าวิธี GEPD แบบ Trimmed q% ในทุกกรณี และวิธี GEPD แบบ Trimmed 5% ก็มีประสิทธิภาพมากกว่าแบบ Trimmed 10% ในทุกกรณีเช่นกัน

สำหรับการประมาณค่าพารามิเตอร์ σ ด้วยวิธี GEPD แบบ Trimmed 5% เมื่อกำหนดให้ขนาดตัวอย่าง (n) คงที่ พบว่าค่า RE1 และ RE2 จะมีค่าเพิ่มขึ้น เมื่อสัดส่วนของการตัดปลาย (p) เพิ่มขึ้น ส่วนวิธี GEPD แบบ Trimmed 10% นั้นพบว่า ผลค่อนข้างไม่ชัดเจน แต่เมื่อกำหนดให้สัดส่วนของการตัดปลาย (p) คงที่ พบว่า ทั้งแบบ Trimmed 5% และแบบ Trimmed 10% นั้น ผลค่อนข้างไม่ชัดเจน

ตารางที่ 4.37 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงแบบปกติ ที่พารามิเตอร์ μ = 0 และ σ = 1 โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— Trimmed 5% & 4 Cluster-Mean และ Trimmed 5% & 8 Cluster-Mean ตามลำดับ

p	n	($\hat{\mu}$) → GE			($\hat{\mu}$) → Trimmed 5% & 4 Cluster-Mean					($\hat{\mu}$) → Trimmed 5% & 8 Cluster-Mean				
		Bias	Var	MSE	Bias	Var	MSE	RE1	RE2	Bias	Var	MSE	RE1	RE2
0	20	-0.00046	0.04845	0.04844	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	40	0.00021	0.02465	0.02464	0.00042	0.02491	0.02491	0.989	0.989	0.00034	0.02491	0.02491	0.989	0.989
	80	0.00025	0.01225	0.01225	0.00065	0.01239	0.01239	0.989	0.989	0.00059	0.01238	0.01238	0.990	0.990
	120	0.00062	0.00806	0.00806	0.00086	0.00810	0.00810	0.995	0.995	0.00084	0.00810	0.00809	0.996	0.995
0.1	20	0.00263	0.05014	0.05014	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	40	0.00197	0.02545	0.02545	0.00135	0.02546	0.02545	1.000	1.000	0.00107	0.02542	0.02541	1.001	1.001
	80	0.00148	0.01264	0.01264	0.00107	0.01265	0.01265	1.000	1.000	0.00092	0.01256	0.01256	1.007	1.007
	120	0.00134	0.00830	0.00830	0.00115	0.00828	0.00828	1.003	1.003	0.00120	0.00821	0.00821	1.011	1.011
0.2	20	0.00106	0.05348	0.05347	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	40	0.00107	0.02694	0.02693	0.00169	0.02628	0.02628	1.025	1.025	0.00158	0.02635	0.02635	1.022	1.022
	80	0.00124	0.01338	0.01338	0.00111	0.01307	0.01307	1.024	1.024	0.00129	0.01309	0.01309	1.022	1.022
	120	0.00099	0.00879	0.00879	0.00122	0.00860	0.00860	1.022	1.022	0.00121	0.00861	0.00861	1.020	1.020
0.3	20	-0.00239	0.05933	0.05933	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	40	-0.00078	0.02985	0.02985	0.00134	0.02932	0.02932	1.018	1.018	0.00124	0.02887	0.02887	1.034	1.034
	80	0.00082	0.01479	0.01479	0.00208	0.01447	0.01447	1.022	1.022	0.00196	0.01425	0.01425	1.039	1.038
	120	0.00040	0.00970	0.00969	0.00135	0.00955	0.00955	1.016	1.015	0.00129	0.00939	0.00939	1.033	1.032

ตารางที่ 4.38 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงแบบปกติ ที่พารามิเตอร์ μ = 0 และ σ = 1 โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— Trimmed 5% & 4 Cluster-Mean และ Trimmed 5% & 8 Cluster-Mean ตามลำดับ

p	n	($\hat{\sigma}$) → GE			($\hat{\sigma}$) → Trimmed 5% & 4 Cluster-Mean					($\hat{\sigma}$) → Trimmed 5% & 8 Cluster-Mean				
		Bias	Var	MSE	Bias	Var	MSE	RE1	RE2	Bias	Var	MSE	RE1	RE2
0	20	-0.03268	0.02563	0.02670	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	40	-0.01844	0.01229	0.01263	-0.01101	0.01433	0.01445	0.858	0.874	-0.01178	0.01373	0.01387	0.895	0.911
	80	-0.00796	0.00620	0.00626	-0.00342	0.00737	0.00738	0.840	0.848	-0.00369	0.00707	0.00708	0.876	0.884
	120	-0.00503	0.00416	0.00419	-0.00179	0.00497	0.00498	0.837	0.841	-0.00195	0.00478	0.00478	0.871	0.875
0.1	20	-0.02637	0.03209	0.03278	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	40	-0.01453	0.01545	0.01565	-0.00950	0.01638	0.01647	0.943	0.951	-0.01038	0.01530	0.01541	1.009	1.016
	80	-0.00508	0.00792	0.00795	-0.00269	0.00854	0.00854	0.928	0.930	-0.00302	0.00800	0.00800	0.991	0.993
	120	-0.00322	0.00529	0.00530	-0.00132	0.00569	0.00569	0.930	0.932	-0.00131	0.00531	0.00531	0.998	0.999
0.2	20	-0.02848	0.03913	0.03994	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	40	-0.01571	0.01935	0.01959	-0.00901	0.01906	0.01913	1.015	1.024	-0.00959	0.01913	0.01921	1.012	1.020
	80	-0.00539	0.00978	0.00981	-0.00255	0.00993	0.00994	0.985	0.987	-0.00238	0.00999	0.00999	0.980	0.982
	120	-0.00369	0.00657	0.00658	-0.00121	0.00668	0.00668	0.984	0.986	-0.00130	0.00668	0.00668	0.984	0.986
0.3	20	-0.03242	0.04818	0.04922	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	40	-0.01778	0.02425	0.02456	-0.00952	0.02587	0.02596	0.937	0.946	-0.01006	0.02450	0.02460	0.990	0.998
	80	-0.00586	0.01204	0.01207	-0.00127	0.01313	0.01313	0.917	0.919	-0.00156	0.01250	0.01250	0.963	0.965
	120	-0.00434	0.00806	0.00808	-0.00107	0.00879	0.00879	0.917	0.919	-0.00122	0.00835	0.00835	0.965	0.967

จากตารางที่ 4.37 พบว่า การประมาณค่าพารามิเตอร์ μ ด้วยวิธี GEPD แบบ Trimmed 5% & K Cluster-Mean จะมีประสิทธิภาพมากกว่าวิธี GE เมื่อสัดส่วนของการตัดปลายมีขนาด $p \geq 0.2$ สำหรับแบบ Trimmed 5% & 4 Cluster-Mean และ $p \geq 0.1$ สำหรับแบบ Trimmed 5% & 8 Cluster-Mean นอกจากนี้ยังพบว่าวิธี GEPD แบบ Trimmed 5% & 8 Cluster-Mean จะมีประสิทธิภาพมากกว่าแบบ Trimmed 5% & 4 Cluster-Mean ในทุกขนาดตัวอย่าง n เมื่อสัดส่วนของการตัดปลาย $p \neq 0.2$

จากตารางที่ 4.38 พบว่า ในกรณีส่วนใหญ่การประมาณค่าพารามิเตอร์ σ ด้วยวิธี GE มีประสิทธิภาพมากกว่าวิธี GEPD แบบ Trimmed 5% & K Cluster-Mean และวิธี GEPD แบบ Trimmed 5% & 8 Cluster-Mean มีประสิทธิภาพมากกว่าวิธีประมาณแบบ Trimmed 5% & 4 Cluster-Mean ในทุกกรณี (ยกเว้น p = 0.2)

ตารางที่ 4.39 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงแบบปกติ ที่พารามิเตอร์ μ = 0 และ σ = 1 โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— Trimmed 10% & 4 Cluster-Mean และ Trimmed 10% & 8 Cluster-Mean ตามลำดับ

p	n	($\hat{\mu}$) → GE			($\hat{\mu}$) → Trimmed 10% & 4 Cluster-Mean					($\hat{\mu}$) → Trimmed 10% & 8 Cluster-Mean				
		Bias	Var	MSE	Bias	Var	MSE	RE1	RE2	Bias	Var	MSE	RE1	RE2
0	20	-0.00046	0.04845	0.04844	0.00004	0.04941	0.04940	0.980	0.980	NA	NA	NA	NA	NA
	40	0.00021	0.02465	0.02464	0.00058	0.02520	0.02519	0.978	0.978	0.00043	0.02518	0.02518	0.979	0.979
	80	0.00025	0.01225	0.01225	0.00076	0.01253	0.01253	0.978	0.978	0.00063	0.01250	0.01249	0.981	0.981
	120	0.00062	0.00806	0.00806	0.00104	0.00818	0.00818	0.985	0.985	0.00102	0.00815	0.00815	0.988	0.988
0.1	20	0.00263	0.05014	0.05014	0.00111	0.05016	0.05015	1.000	1.000	NA	NA	NA	NA	NA
	40	0.00197	0.02545	0.02545	0.00107	0.02553	0.02553	0.997	0.997	0.00080	0.02550	0.02550	0.998	0.998
	80	0.00148	0.01264	0.01264	0.00101	0.01270	0.01269	0.996	0.996	0.00085	0.01260	0.01260	1.003	1.003
	120	0.00134	0.00830	0.00830	0.00117	0.00830	0.00830	1.000	1.000	0.00123	0.00823	0.00823	1.009	1.009
0.2	20	0.00106	0.05348	0.05347	0.00198	0.05217	0.05217	1.025	1.025	NA	NA	NA	NA	NA
	40	0.00107	0.02694	0.02693	0.00144	0.02633	0.02632	1.023	1.023	0.00144	0.02639	0.02639	1.021	1.021
	80	0.00124	0.01338	0.01338	0.00106	0.01310	0.01310	1.021	1.022	0.00124	0.01311	0.01311	1.021	1.021
	120	0.00099	0.00879	0.00879	0.00123	0.00860	0.00860	1.021	1.021	0.00124	0.00861	0.00861	1.020	1.020
0.3	20	-0.00239	0.05933	0.05933	0.00069	0.05841	0.05839	1.016	1.016	NA	NA	NA	NA	NA
	40	-0.00078	0.02985	0.02985	0.00132	0.02929	0.02929	1.019	1.019	0.00130	0.02883	0.02883	1.035	1.035
	80	0.00082	0.01479	0.01479	0.00208	0.01445	0.01445	1.024	1.023	0.00197	0.01422	0.01422	1.040	1.040
	120	0.00040	0.00970	0.00969	0.00136	0.00953	0.00953	1.017	1.017	0.00130	0.00937	0.00937	1.034	1.034

ตารางที่ 4.40 แสดงค่าความเอนเอียง ค่าความแปรปรวน และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงแบบปกติ ที่พารามิเตอร์ $\mu = 0$ และ $\sigma = 1$ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา —— Trimmed 10% & 4 Cluster-Mean และ Trimmed 10% & 8 Cluster-Mean ตามลำดับ

p	n	($\hat{\sigma}$) → GE			($\hat{\sigma}$) → Trimmed 10% & 4 Cluster-Mean					($\hat{\sigma}$) → Trimmed 10% & 8 Cluster-Mean				
		Bias	Var	MSE	Bias	Var	MSE	RE1	RE2	Bias	Var	MSE	RE1	RE2
0	20	-0.03268	0.02563	0.02670	-0.01912	0.03286	0.03322	0.780	0.804	NA	NA	NA	NA	NA
	40	-0.01844	0.01229	0.01263	-0.00980	0.01594	0.01603	0.771	0.788	-0.01054	0.01505	0.01516	0.817	0.833
	80	-0.00796	0.00620	0.00626	-0.00307	0.00816	0.00817	0.759	0.766	-0.00335	0.00771	0.00772	0.804	0.811
	120	-0.00503	0.00416	0.00419	-0.00157	0.00551	0.00551	0.755	0.760	-0.00175	0.00522	0.00522	0.798	0.802
0.1	20	-0.02637	0.03209	0.03278	-0.01731	0.03617	0.03646	0.887	0.899	NA	NA	NA	NA	NA
	40	-0.01453	0.01545	0.01565	-0.00893	0.01739	0.01746	0.888	0.896	-0.00979	0.01613	0.01622	0.958	0.965
	80	-0.00508	0.00792	0.00795	-0.00259	0.00898	0.00899	0.882	0.884	-0.00289	0.00836	0.00836	0.948	0.950
	120	-0.00322	0.00529	0.00530	-0.00135	0.00602	0.00602	0.880	0.882	-0.00137	0.00557	0.00557	0.951	0.953
0.2	20	-0.02848	0.03913	0.03994	-0.01599	0.04228	0.04252	0.926	0.939	NA	NA	NA	NA	NA
	40	-0.01571	0.01935	0.01959	-0.00833	0.02031	0.02038	0.952	0.961	-0.00872	0.02025	0.02032	0.955	0.964
	80	-0.00539	0.00978	0.00981	-0.00243	0.01051	0.01051	0.931	0.933	-0.00216	0.01049	0.01049	0.933	0.935
	120	-0.00369	0.00657	0.00658	-0.00123	0.00711	0.00711	0.924	0.926	-0.00138	0.00706	0.00706	0.931	0.933
0.3	20	-0.03242	0.04818	0.04922	-0.01807	0.05648	0.05679	0.853	0.867	NA	NA	NA	NA	NA
	40	-0.01778	0.02425	0.02456	-0.00853	0.02781	0.02788	0.872	0.881	-0.00894	0.02613	0.02621	0.928	0.937
	80	-0.00586	0.01204	0.01207	-0.00097	0.01407	0.01407	0.856	0.858	-0.00120	0.01326	0.01326	0.908	0.910
	120	-0.00434	0.00806	0.00808	-0.00109	0.00948	0.00948	0.851	0.853	-0.00132	0.00891	0.00891	0.905	0.907

จากตารางที่ 4.39 พบว่า เมื่อสัดส่วนของการตัดปลายมีขนาดเล็ก ($p \leq 0.1$) การประมาณค่าพารามิเตอร์ μ ด้วยวิธี GE จะมีประสิทธิภาพมากกว่าวิธี GEPD แบบ Trimmed 10% & K Cluster-Mean เป็นส่วนใหญ่ แต่ถ้าสัดส่วนของการตัดปลายมีขนาดใหญ่ ($p \geq 0.2$) แล้ว จะพบว่าวิธี GEPD แบบ Trimmed 10% & K Cluster-Mean มีประสิทธิภาพมากกว่าวิธี GE ในทุกกรณี และหากพิจารณาต่อไปอีกจะพบว่า ที่ $p = 0.2$ แบบ Trimmed 10% & 4 Cluster-Mean มีประสิทธิภาพมากกว่าแบบ Trimmed 10% & 8 Cluster-Mean และที่ $p = 0.3$ Trimmed 10% & 8 Cluster-Mean มีประสิทธิภาพมากกว่า Trimmed 10% & 4 Cluster-Mean

จากตารางที่ 4.40 พบว่า การประมาณค่าพารามิเตอร์ σ ด้วยวิธี GE มีประสิทธิภาพมากกว่าวิธี GEPD แบบ Trimmed 10% & K Cluster-Mean ในทุกกรณี และวิธี GEPD แบบ Trimmed 10% & 8 Cluster-Mean ก็มีประสิทธิภาพมากกว่า Trimmed 10% & 4 Cluster-Mean ในทุกกรณีเช่นกัน

บทที่ 5

สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ

จากการศึกษาถึงวิธีการประมาณค่าพารามิเตอร์ด้วยวิธีภาวะน่าจะเป็นสูงสุด วิธีการประมาณแบบกราฟ และวิธีการประมาณแบบกราฟด้วยข้อมูลบางส่วน พบว่าการประมาณค่าพารามิเตอร์ด้วยวิธีภาวะน่าจะเป็นสูงสุดมีประสิทธิภาพมากกว่าวิธีประมาณแบบกราฟในทุกๆ กรณี แต่เนื่องจากวิธีภาวะน่าจะเป็นสูงสุดนั้นค่อนข้างยุ่งยาก เพราะไม่มีรูปแบบสมการที่แน่นอน ต่างจากการประมาณค่าพารามิเตอร์โดยวิธีใช้กราฟ ซึ่งคำนวณค่าประมาณพารามิเตอร์ได้สะดวกและง่ายกว่า ดังนั้นผู้วิจัยจึงสนใจที่จะประมาณค่าพารามิเตอร์โดยวิธีใช้กราฟ แต่ได้ลองปรับเปลี่ยนจากการใช้ข้อมูลจากทุกส่วนหรือทุกจุด เป็นการใช้ข้อมูลเพียงบางส่วน และจากงานวิจัยนี้พบว่า ในกรณีส่วนใหญ่การประมาณค่าพารามิเตอร์แบบกราฟด้วยข้อมูลบางส่วนมีประสิทธิภาพมากกว่าการประมาณแบบกราฟที่ใช้ข้อมูลทั้งหมด ซึ่งสามารถสรุปได้ดังนี้

5.1 สรุปผลการวิจัย

จากผลการวิจัยสามารถสรุปได้ดังนี้

ส่วนที่ 1: การประมาณค่าพารามิเตอร์ μ

- เมื่อพิจารณาเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์ μ ระหว่างวิธี GE และวิธี GEPD ภายใต้สัดส่วนของการตัดปลายทางขวา (p) สำหรับการแจกแจงแบบต่างๆ พบว่า

(1) ข้อมูลที่ไม่ถูกตัดปลายทางขวาหรือสัดส่วนของการตัดปลายมีขนาด p = 0 และมีการแจกแจงแบบ SEV, LEV และ NOR การประมาณค่าพารามิเตอร์ด้วยวิธี 4 Cluster-Mean จะมีประสิทธิภาพมากที่สุด ส่วนการแจกแจงแบบ LOG นั้น พบว่า การประมาณค่าพารามิเตอร์ด้วยวิธี Trimmed 10% จะมีประสิทธิภาพมากที่สุด ทั้งนี้อาจเป็นเพราะว่าการแจกแจงแบบ LOG ถึงแม้จะมีลักษณะที่สมมาตรเหมือนกับ NOR แต่มีหางยาวกว่ามาก การTrim ส่วนหางทั้งสองข้างน่าจะมีผลทำให้ข้อมูลที่เป็น outlier ถูกตัดออกไป

(2) เมื่อสัดส่วนของการตัดปลายทางขวามีขนาด p = 0.3 และข้อมูลมาจากการแจกแจงแบบ SEV, NOR และ LOG การประมาณพารามิเตอร์ด้วยวิธี Trimmed 10% & 8 Cluster-Mean จะมีประสิทธิภาพมากที่สุด ส่วนการแจกแจงแบบ LEV นั้น พบว่า การประมาณพารามิเตอร์ด้วยวิธี Trimmed 5% & 8 Cluster-Mean จะมีประสิทธิภาพมากที่สุด ทั้งนี้อาจเป็น

เพราะว่าการแจกแจงแบบ LEV มีลักษณะเบ้ขวา การ Trim ส่วนหัว (ตัดข้อมูลส่วนปลายทางซ้าย) น่าจะทำให้ข้อมูลส่วนที่ให้รายละเอียดได้ดีถูกตัดทิ้งออกไป

- เมื่อพิจารณาเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์ μ ระหว่างวิธี GE และวิธี GEPD ภายใต้วิธี Trimmed q% สำหรับการแจกแจงแบบต่างๆ พบว่า

(1) ภายใต้ขอบเขตของการวิจัยนี้ วิธี Trimmed q% จะตัดส่วนปลายของทั้งสองข้างออกเป็นจำนวนเท่าๆ กัน แต่ถ้าหากข้อมูลชุดนั้นถูกตัดปลายทางขวามาแล้วไม่น้อยกว่าจำนวนข้อมูลที่จะต้องทำการ Trim ก็ไม่จำเป็นต้อง Trim ข้อมูลในส่วนนั้นอีก คงเหลือแต่การ Trim ข้อมูลตรงส่วนปลายทางซ้าย จากหลักเกณฑ์นี้น่าจะส่งผลให้การประมาณค่าพารามิเตอร์สำหรับการแจกแจงแบบ LEV ด้วยวิธี GEPD แบบ Trimmed q% มีประสิทธิภาพน้อย เนื่องจากข้อมูลของ LEV มีลักษณะเบ้ขวา การ Trim ส่วนหัวทางซ้ายจึงน่าจะเป็นการตัดข้อมูลส่วนที่ให้รายละเอียดได้ดีทิ้งไป ยิ่ง Trim มากขึ้น ข้อมูลส่วนที่ให้รายละเอียดได้ดีก็น่าจะหายไปมากขึ้นเช่นกัน

(2) สำหรับข้อมูลที่มีการแจกแจงแบบ SEV, NOR และ LOG ในกรณีส่วนใหญ่พบว่า การประมาณค่าพารามิเตอร์ที่มีวิธี Trimmed q% ร่วมอยู่ด้วยจะให้ประสิทธิภาพมาก โดยเมื่อสัดส่วนของการตัดปลาย (p) มีขนาดเพิ่มขึ้น ก็ควรที่จะ Trim ข้อมูลให้มีขนาดมากขึ้น

<u>ส่วนที่ 2: การประมาณค่าพารามิเตอร์ σ</u>

- เมื่อพิจารณาเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์ σ ระหว่างวิธี GE และวิธี GEPD ภายใต้การแจกแจงแบบต่างๆ พบว่า

(1) ข้อมูลที่ไม่ถูกตัดปลายทางขวาหรือสัดส่วนของการตัดปลายมีขนาด p = 0 และมาจากการแจกแจงแบบ SEV, LEV และ LOG การประมาณค่าพารามิเตอร์ด้วยวิธี 4 Cluster-Mean จะมีประสิทธิภาพมากที่สุด ต่างจากข้อมูลที่แจกแจงแบบ NOR โดยจะเห็นว่าข้อมูลที่มาจากการแจงแจงแบบ NOR นั้น ในกรณีส่วนใหญ่พบว่า วิธี GE จะเป็นวิธีที่มีประสิทธิภาพมากที่สุด

(2) สำหรับข้อมูลที่มีการแจกแจงแบบ SEV และ LOG พบว่า การประมาณค่าพารามิเตอร์ด้วยวิธี GEPD ที่มีการ Trim ข้อมูลอยู่ด้วย จะมีประสิทธิภาพมาก (ยกเว้น p = 0) ทั้งนี้อาจจะเป็นผลมาจาก การแจกแจงแบบ SEV มีลักษณะเบ้ซ้าย ข้อมูลส่วนที่ให้รายละเอียดได้ดีมักจะอยู่ทางขวา การ Trim ข้อมูลที่อยู่ทางซ้ายน่าจะช่วยตัดข้อมูลที่เป็น outlier ทิ้งไป ในทำนองเดียวกันข้อมูลที่มาจากการแจกแจงแบบ LOG ถึงแม้จะมีลักษณะสมมาตร แต่ก็มีหางที่ยาวผิดปกติ

การ Trim ข้อมูลส่วนปลายทั้งสองข้างก็น่าจะช่วยตัดข้อมูลที่เป็น outlier ทิ้งไปด้วยเช่นกัน

(3) เมื่อสัดส่วนของการตัดปลาย $p \neq 0$ สำหรับข้อมูลที่มีการแจกแจงแบบ LEV การประมาณค่าพารามิเตอร์ด้วยวิธี GE จะมีประสิทธิภาพมากที่สุด ส่วนข้อมูลที่มีการแจกแจงแบบ SEV และ LOG ควรประมาณด้วยวิธี GEPD แบบ Trimmed q% & K Cluster-Mean

ตารางที่ 5.1 แสดงการเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์ μ ระหว่างวิธี GE และวิธี GEPD ที่มีประสิทธิภาพมากที่สุด สำหรับการแจกแจงแบบต่างๆ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา

Censoring Proportion p	Sample Size n	SEV	LEV	NOR	LOG
0	20	4 C-M	4 C-M	GE, 4 C-M	T 10%
	40	4 C-M	4 C-M	GE, 4 C-M, 8 C-M	T 10%
	80	4 C-M	4 C-M	GE, 4 C-M, 8 C-M	T 10%
	120	4 C-M	4 C-M	GE, 4 C-M, 8 C-M	T 10%
0.1	20	T 10% & 4 C-M	4 C-M	4 C-M	T 10%
	40	T 10% & 8 C-M	4 C-M	T 5% & 8 C-M	T 10%
	80	T 10% & 8 C-M	4 C-M	T 5% & 8 C-M	T 10%
	120	T 10% & 8 C-M	4 C-M	T 5% & 8 C-M	T 10%
0.2	20	T 10% & 4 C-M	4 C-M	T 10% & 4 C-M	T 10%
	40	T 10% & 4 C-M	T 5% & 4 C-M	T 5% & 4 C-M	T 10%
	80	T 10% & 4 C-M	T 5% & 8 C-M	T 5% & 4 C-M	T 5%
	120	T 10% & 4 C-M	T 5% & 8 C-M	T 5% & 4 C-M	T 5%
0.3	20	T 10% & 4 C-M	4 C-M	T 10% & 4 C-M	T 10% & 4 C-M
	40	T 10% & 8 C-M	T 5% & 8 C-M	T 10% & 8 C-M	T 10% & 8 C-M
	80	T 10% & 8 C-M	T 5% & 8 C-M	T 10% & 8 C-M	T 10% & 8 C-M
	120	T 10% & 8 C-M	T 5% & 8 C-M	T 10% & 8 C-M	T 10% & 8 C-M

GE คือ วิธี Graphical Estimation

GEPD คือ วิธี Graphical Estimation with Partial Data ซึ่งประกอบไปด้วย

- 4 C-M คือ วิธี 4 Cluster-Mean
- 8 C-M คือ วิธี 8 Cluster-Mean
- T 5% คือ วิธี Trimmed 5%
- T 10% คือ วิธี Trimmed 10%
- T 5% & 4 C-M คือ วิธี Trimmed 5% & 4 Cluster-Mean
- T 5% & 8 C-M คือ วิธี Trimmed 5% & 8 Cluster-Mean
- T 10% & 4 C-M คือ วิธี Trimmed 10% & 4 Cluster-Mean
- T 10% & 8 C-M คือ วิธี Trimmed 10% & 8 Cluster-Mean

<u>หมายเหตุ</u> ภายใต้ขอบเขตของการวิจัยครั้งนี้ วิธี GEPD ศึกษาการแบ่ง Cluster อยู่ 2 ระดับคือ 4 Cluster, 8 Cluster และศึกษาการ Trim ข้อมูลอยู่ 2 ระดับคือ Trimmed 5%, Trimmed 10%

ตารางที่ 5.2 แสดงการเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์ σ ระหว่างวิธี GE และวิธี GEPD ที่มีประสิทธิภาพมากที่สุด สำหรับการแจกแจงแบบต่างๆ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา

Censoring Proportion p	Sample Size n	SEV	LEV	NOR	LOG
0	20	4 C-M	4 C-M	GE	4 C-M
	40	4 C-M	4 C-M	GE	4 C-M
	80	4 C-M	4 C-M	GE	4 C-M
	120	4 C-M	4 C-M	GE	4 C-M
0.1	20	T 10% & 4 C-M	GE	GE	T 10% & 4 C-M
	40	T 5% & 4 C-M	GE	T 5% & 8 C-M	T 5% & 8 C-M
	80	T 5% & 4 C-M	GE	GE	T 5% & 8 C-M
	120	T 5% & 4 C-M	GE	GE	T 5% & 8 C-M
0.2	20	T 10% & 4 C-M	GE	GE	T 10% & 4 C-M
	40	T 5% & 4 C-M	GE	T 5% & 4 C-M	T 5% & 4 C-M
	80	T 5% & 4 C-M	GE	GE	T 5% & 4 C-M
	120	T 5% & 4 C-M	GE	GE	T 5% & 4 C-M
0.3	20	T 10% & 4 C-M	GE	GE	T 10%
	40	T 5% & 4 C-M	GE	8 C-M	T 5% & 4 C-M
	80	T 5% & 4 C-M	GE	GE	T 5% & 4 C-M
	120	T 5% & 4 C-M	GE	8 C-M	T 5% & 8 C-M

GE คือ วิธี Graphical Estimation

GEPD คือ วิธี Graphical Estimation with Partial Data ซึ่งประกอบไปด้วย

- 4 C-M คือ วิธี 4 Cluster-Mean
- 8 C-M คือ วิธี 8 Cluster-Mean
- T 5% คือ วิธี Trimmed 5%
- T 10% คือ วิธี Trimmed 10%
- T 5% & 4 C-M คือ วิธี Trimmed 5% & 4 Cluster-Mean
- T 5% & 8 C-M คือ วิธี Trimmed 5% & 8 Cluster-Mean
- T 10% & 4 C-M คือ วิธี Trimmed 10% & 4 Cluster-Mean
- T 10% & 8 C-M คือ วิธี Trimmed 10% & 8 Cluster-Mean

<u>หมายเหตุ</u> ภายใต้ขอบเขตของการวิจัยครั้งนี้ วิธี GEPD ศึกษาการแบ่ง Cluster อยู่ 2 ระดับคือ 4 Cluster, 8 Cluster และศึกษาการ Trim ข้อมูลอยู่ 2 ระดับคือ Trimmed 5%, Trimmed 10%

ตารางที่ 5.3 แสดงการพิจารณาเลือกใช้วิธีการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงค่าต่ำสุดขีด โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา

Censoring Proportion p	Sample Size n	($\hat{\mu}$)→GE	($\hat{\mu}$)→GEPD							
			4 Cluster-Mean	8 Cluster-Mean	Trimmed 5%	Trimmed 10%	Trimmed 5% & 4 Cluster-Mean	Trimmed 5% & 8 Cluster-Mean	Trimmed 10% & 4 Cluster-Mean	Trimmed 10% & 8 Cluster-Mean
0	20		✓							
	40		✓							
	80		✓							
	120		✓							
0.1	20								✓	
	40									✓
	80									✓
	120									✓
0.2	20								✓	
	40								✓	
	80								✓	
	120								✓	
0.3	20								✓	
	40									✓
	80									✓
	120									✓

ตารางที่ 5.4 แสดงการพิจารณาเลือกใช้วิธีการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงค่าต่ำสุดขีด โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา

Censoring Proportion p	Sample Size n	($\hat{\sigma}$)→GE	($\hat{\sigma}$)→GEPD							
			4 Cluster-Mean	8 Cluster-Mean	Trimmed 5%	Trimmed 10%	Trimmed 5% & 4 Cluster-Mean	Trimmed 5% & 8 Cluster-Mean	Trimmed 10% & 4 Cluster-Mean	Trimmed 10% & 8 Cluster-Mean
0	20		✓							
	40		✓							
	80		✓							
	120		✓							
0.1	20								✓	
	40						✓			
	80						✓			
	120						✓			
0.2	20								✓	
	40						✓			
	80						✓			
	120						✓			
0.3	20								✓	
	40						✓			
	80						✓			
	120						✓			

ตารางที่ 5.5 แสดงการพิจารณาเลือกใช้วิธีการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงค่าสูงสุดขีด โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา

Censoring Proportion p	Sample Size n	($\hat{\mu}$)→GE	($\hat{\mu}$)→GEPD							
			4 Cluster-Mean	8 Cluster-Mean	Trimmed 5%	Trimmed 10%	Trimmed 5% & 4 Cluster-Mean	Trimmed 5% & 8 Cluster-Mean	Trimmed 10% & 4 Cluster-Mean	Trimmed 10% & 8 Cluster-Mean
0	20		✓							
	40		✓							
	80		✓							
	120		✓							
0.1	20		✓							
	40		✓							
	80		✓							
	120		✓							
0.2	20		✓							
	40						✓			
	80							✓		
	120							✓		
0.3	20		✓							
	40							✓		
	80							✓		
	120							✓		

ตารางที่ 5.6 แสดงการพิจารณาเลือกใช้วิธีการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงค่าสูงสุดขีด โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา

Censoring Proportion p	Sample Size n	($\hat{\sigma}$)→GE	($\hat{\sigma}$)→GEPD							
			4 Cluster-Mean	8 Cluster-Mean	Trimmed 5%	Trimmed 10%	Trimmed 5% & 4 Cluster-Mean	Trimmed 5% & 8 Cluster-Mean	Trimmed 10% & 4 Cluster-Mean	Trimmed 10% & 8 Cluster-Mean
0	20		✓							
	40		✓							
	80		✓							
	120		✓							
0.1	20	✓								
	40	✓								
	80	✓								
	120	✓								
0.2	20	✓								
	40	✓								
	80	✓								
	120	✓								
0.3	20	✓								
	40	✓								
	80	✓								
	120	✓								

ตารางที่ 5.7 แสดงการพิจารณาเลือกใช้วิธีการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงโลจิสติค โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา

Censoring Proportion p	Sample Size n	$(\hat{\mu})\rightarrow$GE	$(\hat{\mu})\rightarrow$GEPD							
			4 Cluster-Mean	8 Cluster-Mean	Trimmed 5%	Trimmed 10%	Trimmed 5% & 4 Cluster-Mean	Trimmed 5% & 8 Cluster-Mean	Trimmed 10% & 4 Cluster-Mean	Trimmed 10% & 8 Cluster-Mean
0	20					✓				
	40					✓				
	80					✓				
	120					✓				
0.1	20					✓				
	40					✓				
	80					✓				
	120					✓				
0.2	20					✓				
	40					✓				
	80				✓					
	120				✓					
0.3	20								✓	
	40									✓
	80									✓
	120									✓

ตารางที่ 5.8 แสดงการพิจารณาเลือกใช้วิธีการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงโลจิสติค โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา

Censoring Proportion p	Sample Size n	$(\hat{\sigma})\rightarrow$GE	$(\hat{\sigma})\rightarrow$GEPD							
			4 Cluster-Mean	8 Cluster-Mean	Trimmed 5%	Trimmed 10%	Trimmed 5% & 4 Cluster-Mean	Trimmed 5% & 8 Cluster-Mean	Trimmed 10% & 4 Cluster-Mean	Trimmed 10% & 8 Cluster-Mean
0	20		✓							
	40		✓							
	80		✓							
	120		✓							
0.1	20								✓	
	40							✓		
	80							✓		
	120							✓		
0.2	20								✓	
	40						✓			
	80						✓			
	120						✓			
0.3	20					✓				
	40						✓			
	80						✓			
	120							✓		

ตารางที่ 5.9 แสดงการพิจารณาเลือกใช้วิธีการประมาณค่าพารามิเตอร์ μ สำหรับการแจกแจงปกติ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา

Censoring Proportion p	Sample Size n	$(\hat{\mu})\rightarrow$GE	$(\hat{\mu})\rightarrow$GEPD							
			4 Cluster-Mean	8 Cluster-Mean	Trimmed 5%	Trimmed 10%	Trimmed 5% & 4 Cluster-Mean	Trimmed 5% & 8 Cluster-Mean	Trimmed 10% & 4 Cluster-Mean	Trimmed 10% & 8 Cluster-Mean
0	20	✓*	✓*							
	40	✓*	✓*	✓*						
	80	✓*	✓*	✓*						
	120	✓*	✓*	✓*						
0.1	20		✓							
	40							✓		
	80							✓		
	120							✓		
0.2	20								✓	
	40						✓			
	80						✓			
	120						✓			
0.3	20								✓	
	40									✓
	80									✓
	120									✓

หมายแหตุ * หมายถึง ค่าประมาณพารามิเตอร์ที่ได้มีประสิทธิภาพเท่ากัน

ตารางที่ 5.10 แสดงการพิจารณาเลือกใช้วิธีการประมาณค่าพารามิเตอร์ σ สำหรับการแจกแจงปกติ โดยจำแนกตามขนาดตัวอย่าง และสัดส่วนของการตัดปลายทางขวา

Censoring Proportion p	Sample Size n	$(\hat{\sigma})\rightarrow$GE	$(\hat{\sigma})\rightarrow$GEPD							
			4 Cluster-Mean	8 Cluster-Mean	Trimmed 5%	Trimmed 10%	Trimmed 5% & 4 Cluster-Mean	Trimmed 5% & 8 Cluster-Mean	Trimmed 10% & 4 Cluster-Mean	Trimmed 10% & 8 Cluster-Mean
0	20	✓								
	40	✓								
	80	✓								
	120	✓								
0.1	20	✓								
	40							✓		
	80	✓								
	120	✓								
0.2	20	✓								
	40						✓			
	80	✓								
	120	✓								
0.3	20	✓								
	40			✓						
	80	✓								
	120			✓						

5.2 ข้อเสนอแนะ

ผลการวิจัยแบ่งข้อเสนอแนะเป็น 2 ด้าน ดังนี้

5.2.1 ด้านการนำไปใช้ประโยชน์

จากการวิจัยสามารถสรุปขั้นตอนการประมาณค่าพารามิเตอร์ ได้ดังนี้

(1) พิจารณาว่าข้อมูลที่ต้องการประมาณค่าพารามิเตอร์เป็นข้อมูลที่ถูกตัดปลายทางขวาประเภทที่ 2 ซึ่งมาจากการแจกแจงแบบ location-scale แบบใดแบบหนึ่งใน 4 แบบ นี้คือ การแจกแจงแบบค่าต่ำสุดขีด แบบค่าสูงสุดขีด แบบโลจิสติค หรือแบบปกติ

(2) การประมาณค่าพารามิเตอร์ภายใต้การศึกษานี้ เหมาะสำหรับข้อมูลที่มีขนาดตัวอย่างเล็ก n = 20, 40, 80 และ 120 ซึ่งมีสัดส่วนของการตัดปลาย p = 0.1, 0.2 และ 0.3

(3) ใช้วิธีพล็อตกราฟในการประมาณค่าพารามิเตอร์ μ และ σ โดยใช้วิธีกำลังสองต่ำสุดในการลากเส้นกราฟ ซึ่งจะได้ส่วนตัดแกน y (intercept) เป็นค่าประมาณพารามิเตอร์ μ และความชัน (slope) เป็นค่าประมาณพารามิเตอร์ σ

(4) ภายใต้ขอบเขตของการศึกษาวิจัยนี้พบว่า วิธีการประมาณแบบกราฟด้วยข้อมูลบางส่วนเหมาะสำหรับข้อมูลที่มีการแจกแจงไม่ปกติ คือ เป็นข้อมูลที่มีลักษณะเบ้ซ้าย (SEV) หรือสมมาตรแต่มีหางยาวกว่าปกติ (LOG) ส่วนข้อมูลที่มีลักษณะเบ้ขวา (LEV) หรือมีการแจกแจงแบบปกติ (NOR) นั้นควรใช้วิธีการประมาณแบบกราฟที่ใช้ข้อมูลทุกจุด เพราะในกรณีส่วนใหญ่แล้วจะให้ประสิทธิภาพในการประมาณที่ดีกว่าการประมาณแบบกราฟด้วยข้อมูลบางส่วน

(5) จากงานวิจัยนี้พบว่า วิธีการประมาณพารามิเตอร์ด้วยวิธีภาวะน่าจะเป็นสูงสุดยังคงเป็นวิธีการประมาณที่มีประสิทธิภาพมากที่สุด เมื่อเทียบกับการประมาณแบบกราฟและการประมาณแบบกราฟด้วยข้อมูลบางส่วน

5.2.2 ด้านการศึกษาวิจัย

การวิจัยครั้งนี้ได้ทำการศึกษาภายใต้ขอบเขตของข้อมูลที่มาการแจกแจงแบบค่าต่ำสุดขีด แบบค่าสูงสุดขีด แบบโลจิสติค และแบบปกติสำหรับข้อมูลที่ถูกตัดปลายทางขวาประเภทที่ 2 ด้วยสัดส่วนของการตัดปลาย (p) เป็น 10%, 20% และ 30% ที่ขนาดตัวอย่าง (n) เท่ากับ 20, 40, 80 และ 120 จึงเป็นที่น่าสนใจว่า

(1) ถ้าสัดส่วนของการตัดปลายมีขนาดเพิ่มขึ้นการประมาณค่าพารามิเตอร์จะให้ผลที่ดีขึ้นหรือไม่ โดยอาจจะหาขนาดของการตัดปลายที่เหมาะสมกับจำนวน Cluster และขนาดตัวอย่าง

(2) ถ้าเปลี่ยนการแจกแจงในการศึกษาเป็นแบบอื่น ซึ่งเป็นการแจกแจงแบบ location-scale แล้ว โดยรวมการประมาณค่าแบบ GEPD ยังจะให้ผลดีกว่าแบบ GE หรือไม่

(3) การประมาณค่าด้วยวิธี GEPD แบบ Trimmed q% หากสัดส่วนของการ Trim ข้อมูลมากขึ้น การประมาณค่าพารามิเตอร์จะให้ผลที่ดีขึ้นหรือไม่ โดยอาจจะหาสัดส่วนของวิธี Trimmed q% ที่เหมาะสมสำหรับสัดส่วนของการตัดปลายทางขวา

(4) ศึกษาการประมาณค่าพารามิเตอร์ด้วยวิธี GEPD กับข้อมูลที่ถูกเซ็นเซอร์ประเภทอื่นๆ เช่น ข้อมูลที่ถูกเซ็นเซอร์แบบช่วง ข้อมูลที่ถูกเซ็นเซอร์ทางซ้าย รวมถึงข้อมูลที่ถูกเซ็นเซอร์ทางขวาประเภทที่ 1 ต่อไปอีก เป็นต้น

รายการอ้างอิง

ภาษาไทย

จำเนียน จำนงค์รักษ์. 2539. การพยากรณ์ในสมการถดถอยเชิงเส้นพหุเมื่อค่าตัวแปรตามถูกตัดปลายทางขวา. วิทยานิพนธ์ปริญญามหาบัณฑิต, ภาควิชาสถิติ คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย.

ทศพร แถลงธรรม. 2539. การเปรียบเทียบการประมาณค่าพารามิเตอร์ของการแจกแจงความเสียหายที่มีข้อมูลเป็นแบบกลุ่มถูกตัดปลายทางซ้ายและขวา. วิทยานิพนธ์ปริญญามหาบัณฑิต, ภาควิชาสถิติ คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย.

ธีระพร วีระถาวร. 2536. การอนุมานเชิงสถิติขั้นกลาง : โครงสร้างและความหมาย. พิมพ์ครั้งที่ 2. กรุงเทพมหานคร : สำนักพิมพ์จุฬาลงกรณ์มหาวิทยาลัย.

นันทพร อารยะสกุลวงศ์. 2539. การประมาณค่าฟังก์ชันการอยู่รอดสำหรับข้อมูลที่มีค่าถูกตัดทิ้งทางขวา. วิทยานิพนธ์ปริญญามหาบัณฑิต, ภาควิชาสถิติ คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย.

ศศิประภา โมรากุล. 2553. การศึกษาเปรียบเทียบการประมาณพารามิเตอร์จากข้อมูลที่ถูกเซ็นเซอร์แบบช่วง. วิทยานิพนธ์ปริญญามหาบัณฑิต, ภาควิชาสถิติ คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย.

ภาษาอังกฤษ

Genschel, U., and Meeker, W.Q. 2010. A Comparison of Maximum Likelihood and Median-Rank Regression for Weibull Estimation. Quality Engineering 20: 236-255.

Klein, J.P., Moesschberger, M. L. 1997. Survival Analysis Techniques for Censored and Truncated Data. New York : Springer.

Lawless, J.F. 2003. Statistical Models and Methods for Lifetime Data. Second Edition. New York : John Wiley.

Meeker, W.Q., Escobar, L.A. 1998. Statistical Methods for Reliability Data. New York : John Wiley.

R Development Core Team. 2009. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0, URL http://www.R-project.org.

Somboonsavatdee, A. and Nair, V. 2007. Graphical Estimation from Probability Plots with Right Censored Data. Technometrics 19 : 420 - 429.

ภาคผนวก

โปรแกรมที่ใช้สำหรับการแจกแจงแบบค่าต่ำสุดขีด

*** กรณีข้อมูลมีการแจกแจงแบบค่าต่ำสุดขีด μ = 0 และ σ =1 ด้วยสัดส่วนของการตัดปลาย p = 0.1 ***

```
##############################################################################################
######################### SEV Distribution (Type II censoring 10%) ###########################
##############################################################################################
mu<-0;
sigma<-1;
Nsim<-5000;
nn<-c(20,40,80,120);

rsev<-function(n,mu=0,sigma=1){
log(rweibull(n,1/sigma,exp(mu)))}

psev<-function(q,mu=0,sigma=1){
1-exp(-exp((q-mu)/sigma))}

dsev<-function(x,mu=0,sigma=1){
(1/sigma)*exp(((x-mu)/sigma)-exp((x-mu)/sigma))}

qsev<-function(p,mu=0,sigma=1){
log(qweibull(p,1/sigma,exp(mu)))}

##############################################################################################
########################################## MLE Method ########################################
##############################################################################################
out1<-c();
lksev<-function(sampcens,mu=0,sigma=1){
if(length(mu)==1){
out1<-prod(dsev(sampcens,mu,sigma),(1-psev(sampcens[Fc],mu,sigma))^(n-Fc))
```

```
}
if(length(mu)>1){
out1<-c();
for(i in 1:length(mu)){
out1<-c(out1,prod(dsev(sampcens,mu[i],sigma),(1-psev(sampcens[Fc],mu[i],sigma))^(n-Fc)))
}
}
return(out1)
}

out2<-c();
lksev.sd<-function(sampcens,mu=0,sigma=1){
if(length(sigma)==1){
out2<-prod(dsev(sampcens,mu,sigma),(1-psev(sampcens[Fc],mu,sigma))^(n-Fc))
}
if(length(sigma)>1){
out2<-c();
for(i in 1:length(sigma)){
out2<-c(out2,prod(dsev(sampcens,mu,sigma[i]),(1-psev(sampcens[Fc],mu,sigma[i]))^(n-Fc)))
}
}
return(out2)
}

bbias1<-c();
bbias2<-c();
vvar1<-c();
vvar2<-c();
mmse1<-c();
mmse2<-c();
```

```
for(j in 1:length(nn)){

n<-nn[j];
Fc<-n-(0.1*n);
cens<-c(rep(0,Fc),rep(1,(n-Fc)))

muhat<-c();
sigmahat<-c();
set.seed(106);
for(i in 1:Nsim){

samp<-rsev(n,mu,sigma);
samp<-sort(samp);
samp[cens==1]<-samp[Fc];
data<-cbind(samp,cens);
sampcens<-data[data[,2]==0,1];
kk<-seq(-1,1,0.01);
ll<-seq(0.01,2,0.01);
muhat[i]<-kk[lksev(sampcens,kk,sigma)==max(lksev(sampcens,kk,sigma))];
sigmahat[i]<-ll[lksev.sd(sampcens,mu,ll)==max(lksev.sd(sampcens,mu,ll))];
}

write.csv(cbind(muhat,sigmahat),paste("SEVcen10.mesh",n,".csv",sep=""));
bbias1[j]<-mean(muhat)-mu;
bbias2[j]<-mean(sigmahat)-sigma;
vvar1[j]<-var(muhat);
vvar2[j]<-var(sigmahat);
mmse1[j]<-mean((muhat-mu)^2);
mmse2[j]<-mean((sigmahat-sigma)^2);
}
```

```
dum<-cbind(bbias1,vvar1,mmse1,bbias2,vvar2,mmse2);
colnames(dum)<-c("Bias1","Var1","MSE1","Bias2","Var2","MSE2");
rownames(dum)<-c("n=20","n=40","n=80","n=120");
write.csv(dum,file="SumSEVcen10.mesh.csv");

##############################################################################################
############################### Graphical (All points) Method ################################
##############################################################################################
bbias1<-c();
bbias2<-c();
vvar1<-c();
vvar2<-c();
mmse1<-c();
mmse2<-c();
for(j in 1:length(nn)){

n<-nn[j];
Fc<-n-(0.1*n);

muhat<-c();
sigmahat<-c();
set.seed(106);
for(i in 1:Nsim){

samp<-rsev(n,mu,sigma);
sortx<-sort(samp);
p<-((1:n)-0.5)/n;
x<-cbind(sortx[1:Fc],p[1:Fc]);
ols<-coef(lm(x[,1]~qsev(x[,2])));
```

```
muhat[i]<-ols[1];
sigmahat[i]<-ols[2];
}

write.csv(cbind(muhat,sigmahat),paste("SEVcen10.aps",n,".csv",sep=""));
bbias1[j]<-mean(muhat)-mu;
bbias2[j]<-mean(sigmahat)-sigma;
vvar1[j]<-var(muhat);
vvar2[j]<-var(sigmahat);
mmse1[j]<-mean((muhat-mu)^2);
mmse2[j]<-mean((sigmahat-sigma)^2);
}

dum<-cbind(bbias1,vvar1,mmse1,bbias2,vvar2,mmse2);
colnames(dum)<-c("Bias1","Var1","MSE1","Bias2","Var2","MSE2");
rownames(dum)<-c("n=20","n=40","n=80","n=120");
write.csv(dum,file="SumSEVcen10.aps.csv");

##############################################################################
########################## Graphical (4 Cluster-Mean) Method ##################
##############################################################################
estim<-function(samp){
sortx<-sort(samp);
N<-length(samp);
p<-((1:N)-0.5)/N;
Fc<-N-(0.1*N);
ncluster<-4;
m<-floor(Fc/ncluster);
x<-c();
begin<-0;
```

```
end<-0;

for(i in 1:ncluster){
ifelse((N==20)&(i==2|i==3),end<-1,end<-0);
xx<-cbind(sortx[begin+1:(m+end)],p[begin+1:(m+end)]);
xx[,2]<-qsev(xx[,2]);
x<-rbind(x,apply(xx,2,mean));
begin<-begin+m+end;
}

ifelse(N==20,end<-1,end<-0);
lin<-lm(x[,1]~x[,2],weight=c(m,m+end,m+end,m));
ols<-lin$coef;

return(ols)
}

bbias1<-c();
bbias2<-c();
vvar1<-c();
vvar2<-c();
mmse1<-c();
mmse2<-c();
for(j in 1:length(nn)){

n<-nn[j];

muhat<-c();
sigmahat<-c();
set.seed(106);
```

```
for(i in 1:Nsim){
samp<-rsev(n,mu,sigma);
muhat[i]<-estim(samp)[1];
sigmahat[i]<-estim(samp)[2];
}

write.csv(cbind(muhat,sigmahat),paste("SEVcen10.4cmean",n,".csv",sep=""));
bbias1[j]<-mean(muhat)-mu;
bbias2[j]<-mean(sigmahat)-sigma;
vvar1[j]<-var(muhat);
vvar2[j]<-var(sigmahat);
mmse1[j]<-mean((muhat-mu)^2);
mmse2[j]<-mean((sigmahat-sigma)^2);
}

dum<-cbind(bbias1,vvar1,mmse1,bbias2,vvar2,mmse2);
colnames(dum)<-c("Bias1","Var1","MSE1","Bias2","Var2","MSE2");
rownames(dum)<-c("n=20","n=40","n=80","n=120");
write.csv(dum,file="SumSEVcen10.4cmean.csv");

##############################################################################
########################## Graphical (8 Cluster-Mean) Method ##################
##############################################################################
estim<-function(samp){
sortx<-sort(samp);
N<-length(samp);
p<-((1:N)-0.5)/N;
Fc<-N-(0.1*N);
ncluster<-8;
m<-floor(Fc/ncluster);
```

```
x<-c();
begin<-0;
end<-0;

for(i in 1:ncluster){
ifelse((N==40|N==120)&(i==3|i==4|i==5|i==6),end<-1,end<-0);
xx<-cbind(sortx[begin+1:(m+end)],p[begin+1:(m+end)]);
xx[,2]<-qsev(xx[,2]);
x<-rbind(x,apply(xx,2,mean));
begin<-begin+m+end;
}

ifelse((N==40|N==120),end<-1,end<-0);
lin<-lm(x[,1]~x[,2],weight=c(m,m,m+end,m+end,m+end,m+end,m,m));
ols<-lin$coef;

return(ols)
}

bbias1<-c();
bbias2<-c();
vvar1<-c();
vvar2<-c();
mmse1<-c();
mmse2<-c();
for(j in 2:length(nn)){

n<-nn[j];

muhat<-c();
```

```
sigmahat<-c();
set.seed(106);
for(i in 1:Nsim){
samp<-rsev(n,mu,sigma);
muhat[i]<-estim(samp)[1];
sigmahat[i]<-estim(samp)[2];
}

write.csv(cbind(muhat,sigmahat),paste("SEVcen10.8cmean",n,".csv",sep=""));
bbias1[j]<-mean(muhat)-mu;
bbias2[j]<-mean(sigmahat)-sigma;
vvar1[j]<-var(muhat);
vvar2[j]<-var(sigmahat);
mmse1[j]<-mean((muhat-mu)^2);
mmse2[j]<-mean((sigmahat-sigma)^2);
}

dum<-cbind(bbias1,vvar1,mmse1,bbias2,vvar2,mmse2);
colnames(dum)<-c("Bias1","Var1","MSE1","Bias2","Var2","MSE2");
rownames(dum)<-c("n=20","n=40","n=80","n=120");
write.csv(dum,file="SumSEVcen10.8cmean.csv");

##############################################################################
############################ Graphical (Trimmed 5%) Method ###################
##############################################################################
estim<-function(samp){
sortx<-sort(samp);
N<-length(samp);
p<-((1:N)-0.5)/N;
Fc<-N-(0.1*N);
```

```
t<-(0.05*N)/2
x<-cbind(sortx[(t+1):Fc],p[(t+1):Fc])

lin<-lm(x[,1]~qsev(x[,2]));
ols<-lin$coef;

return(ols)
}

bbias1<-c();
bbias2<-c();
vvar1<-c();
vvar2<-c();
mmse1<-c();
mmse2<-c();
for(j in 2:length(nn)){

n<-nn[j];

muhat<-c();
sigmahat<-c();
set.seed(106);
for(i in 1:Nsim){
samp<-rsev(n,mu,sigma);
muhat[i]<-estim(samp)[1];
sigmahat[i]<-estim(samp)[2];
}

write.csv(cbind(muhat,sigmahat),paste("SEVcen10.tmd5%",n,".csv",sep=""));
bbias1[j]<-mean(muhat)-mu;
```

```
bbias2[j]<-mean(sigmahat)-sigma;
vvar1[j]<-var(muhat);
vvar2[j]<-var(sigmahat);
mmse1[j]<-mean((muhat-mu)^2);
mmse2[j]<-mean((sigmahat-sigma)^2);
}

dum<-cbind(bbias1,vvar1,mmse1,bbias2,vvar2,mmse2);
colnames(dum)<-c("Bias1","Var1","MSE1","Bias2","Var2","MSE2");
rownames(dum)<-c("n=20","n=40","n=80","n=120");
write.csv(dum,file="SumSEVcen10.tmd5%.csv");

##############################################################################
########################## Graphical (Trimmed 10%) Method ####################
##############################################################################
estim<-function(samp){
sortx<-sort(samp);
N<-length(samp);
p<-((1:N)-0.5)/N;
Fc<-N-(0.1*N);
t<-(0.1*N)/2
x<-cbind(sortx[(t+1):Fc],p[(t+1):Fc])

lin<-lm(x[,1]~qsev(x[,2]));
ols<-lin$coef;

return(ols)
}

bbias1<-c();
```

```
bbias2<-c();
vvar1<-c();
vvar2<-c();
mmse1<-c();
mmse2<-c();
for(j in 1:length(nn)){

n<-nn[j];

muhat<-c();
sigmahat<-c();
set.seed(106);
for(i in 1:Nsim){
samp<-rsev(n,mu,sigma);
muhat[i]<-estim(samp)[1];
sigmahat[i]<-estim(samp)[2];
}

write.csv(cbind(muhat,sigmahat),paste("SEVcen10.tmd10%",n,".csv",sep=""));
bbias1[j]<-mean(muhat)-mu;
bbias2[j]<-mean(sigmahat)-sigma;
vvar1[j]<-var(muhat);
vvar2[j]<-var(sigmahat);
mmse1[j]<-mean((muhat-mu)^2);
mmse2[j]<-mean((sigmahat-sigma)^2);
mmse<-(mmse1+mmse2)/2;
}

dum<-cbind(bbias1,vvar1,mmse1,bbias2,vvar2,mmse2);
colnames(dum)<-c("Bias1","Var1","MSE1","Bias2","Var2","MSE2");
```

```
rownames(dum)<-c("n=20","n=40","n=80","n=120");
write.csv(dum,file="SumSEVcen10.tmd10%.csv");

##########################################################################################
################### Graphical (Trimmed 5% & 4 Cluster-Mean) Method ######################
##########################################################################################
estim<-function(samp){
sortx<-sort(samp);
N<-length(samp);
p<-((1:N)-0.5)/N;
Fc<-N-(0.1*N);
t<-(0.05*N)/2;
ncluster<-4;
m<-floor((Fc-t)/ncluster);
x<-c();
begin<-0;
end<-0;

for(i in 1:ncluster){
end<-0;
if(N==40&(i==1|i==2|i==3))end<-1;
if(N==80&(i==2|i==3))end<-1;
if(N==120&i==2)end<-1;
xx<-cbind(sortx[begin+t+1:(m+end)],p[begin+t+1:(m+end)]);
xx[,2]<-qsev(xx[,2]);
x<-rbind(x,apply(xx,2,mean));
begin<-begin+m+end;
}

if(N==40)lin<-lm(x[,1]~x[,2],weight=c(m+1,m+1,m+1,m));
```

```
if(N==80)lin<-lm(x[,1]~x[,2],weight=c(m,m+1,m+1,m));
if(N==120)lin<-lm(x[,1]~x[,2],weight=c(m,m+1,m,m));
ols<-lin$coef;

return(ols)
}

bbias1<-c();
bbias2<-c();
vvar1<-c();
vvar2<-c();
mmse1<-c();
mmse2<-c();
for(j in 2:length(nn)){

n<-nn[j];

muhat<-c();
sigmahat<-c();
set.seed(106);
for(i in 1:Nsim){
samp<-rsev(n,mu,sigma);
muhat[i]<-estim(samp)[1];
sigmahat[i]<-estim(samp)[2];
}

write.csv(cbind(muhat,sigmahat),paste("SEVcen10.tmd5%&4cmean",n,".csv",sep=""));
bbias1[j]<-mean(muhat)-mu;
bbias2[j]<-mean(sigmahat)-sigma;
vvar1[j]<-var(muhat);
```

```
vvar2[j]<-var(sigmahat);
mmse1[j]<-mean((muhat-mu)^2);
mmse2[j]<-mean((sigmahat-sigma)^2);
}

dum<-cbind(bbias1,vvar1,mmse1,bbias2,vvar2,mmse2);
colnames(dum)<-c("Bias1","Var1","MSE1","Bias2","Var2","MSE2");
rownames(dum)<-c("n=20","n=40","n=80","n=120");
write.csv(dum,file="SumSEVcen10.tmd5%&4cmean.csv");

##########################################################################################
#################### Graphical (Trimmed 5% & 8 Cluster-Mean) Method ######################
##########################################################################################
estim<-function(samp){
sortx<-sort(samp);
N<-length(samp);
p<-((1:N)-0.5)/N;
Fc<-N-(0.1*N);
t<-(0.05*N)/2;
ncluster<-8;
m<-floor((Fc-t)/ncluster);
x<-c();
begin<-0;
end<-0;

for(i in 1:ncluster){
end<-0;
if(N==40&(i==3|i==4|i==5))end<-1;
if(N==80&(i==2|i==3|i==4|i==5|i==6|i==7))end<-1;
if(N==120&i==4)end<-1;
```

```
xx<-cbind(sortx[begin+t+1:(m+end)],p[begin+t+1:(m+end)]);
xx[,2]<-qsev(xx[,2]);
x<-rbind(x,apply(xx,2,mean));
begin<-begin+m+end;
}

if(N==40)lin<-lm(x[,1]~x[,2],weight=c(m,m,m+1,m+1,m+1,m,m,m));
if(N==80)lin<-lm(x[,1]~x[,2],weight=c(m,m+1,m+1,m+1,m+1,m+1,m+1,m));
if(N==120)lin<-lm(x[,1]~x[,2],weight=c(m,m,m,m+1,m,m,m,m));
ols<-lin$coef;

return(ols)
}

bbias1<-c();
bbias2<-c();
vvar1<-c();
vvar2<-c();
mmse1<-c();
mmse2<-c();
for(j in 2:length(nn)){

n<-nn[j];

muhat<-c();
sigmahat<-c();
set.seed(106);
for(i in 1:Nsim){
samp<-rsev(n,mu,sigma);
muhat[i]<-estim(samp)[1];
```

```
sigmahat[i]<-estim(samp)[2];
}

write.csv(cbind(muhat,sigmahat),paste("SEVcen10.tmd5%&8cmean",n,".csv",sep=""));
bbias1[j]<-mean(muhat)-mu;
bbias2[j]<-mean(sigmahat)-sigma;
vvar1[j]<-var(muhat);
vvar2[j]<-var(sigmahat);
mmse1[j]<-mean((muhat-mu)^2);
mmse2[j]<-mean((sigmahat-sigma)^2);
}

dum<-cbind(bbias1,vvar1,mmse1,bbias2,vvar2,mmse2);
colnames(dum)<-c("Bias1","Var1","MSE1","Bias2","Var2","MSE2");
rownames(dum)<-c("n=20","n=40","n=80","n=120");
write.csv(dum,file="SumSEVcen10.tmd5%&8cmean.csv");

##############################################################################
#################### Graphical (Trimmed 10% & 4 Cluster-Mean) Method ##################
##############################################################################
estim<-function(samp){
sortx<-sort(samp);
N<-length(samp);
p<-((1:N)-0.5)/N;
Fc<-N-(0.1*N);
t<-(0.1*N)/2;
ncluster<-4;
m<-floor((Fc-t)/ncluster);
x<-c();
begin<-0;
```

```
end<-0;

for(i in 1:ncluster){
end<-0;
if(N==20&i==2)end<-1;
if((N==40|N==120)&(i==2|i==3))end<-1;
xx<-cbind(sortx[begin+t+1:(m+end)],p[begin+t+1:(m+end)]);
xx[,2]<-qsev(xx[,2]);
x<-rbind(x,apply(xx,2,mean));
begin<-begin+m+end;
}

if(N==20)lin<-lm(x[,1]~x[,2],weight=c(m,m+1,m,m));
if(N==40|N==120)lin<-lm(x[,1]~x[,2],weight=c(m,m+1,m+1,m));
if(N==80)lin<-lm(x[,1]~x[,2],weight=c(m,m,m,m));
ols<-lin$coef;

return(ols)
}

bbias1<-c();
bbias2<-c();
vvar1<-c();
vvar2<-c();
mmse1<-c();
mmse2<-c();
for(j in 1:length(nn)){

n<-nn[j];
```

```
muhat<-c();
sigmahat<-c();
set.seed(106);
for(i in 1:Nsim){
samp<-rsev(n,mu,sigma);
muhat[i]<-estim(samp)[1];
sigmahat[i]<-estim(samp)[2];
}

write.csv(cbind(muhat,sigmahat),paste("SEVcen10.tmd10%&4cmean",n,".csv",sep=""));
bbias1[j]<-mean(muhat)-mu;
bbias2[j]<-mean(sigmahat)-sigma;
vvar1[j]<-var(muhat);
vvar2[j]<-var(sigmahat);
mmse1[j]<-mean((muhat-mu)^2);
mmse2[j]<-mean((sigmahat-sigma)^2);
}

dum<-cbind(bbias1,vvar1,mmse1,bbias2,vvar2,mmse2);
colnames(dum)<-c("Bias1","Var1","MSE1","Bias2","Var2","MSE2");
rownames(dum)<-c("n=20","n=40","n=80","n=120");
write.csv(dum,file="SumSEVcen10.tmd10%&4cmean.csv");

##############################################################################
#################### Graphical (Trimmed 10% & 8 Cluster-Mean) Method ##################
##############################################################################
estim<-function(samp){
sortx<-sort(samp);
N<-length(samp);
p<-((1:N)-0.5)/N;
```

```
Fc<-N-(0.1*N);
t<-(0.1*N)/2;
ncluster<-8;
m<-floor((Fc-t)/ncluster);
x<-c();
begin<-0;
end<-0;

for(i in 1:ncluster){
end<-0;
if(N==40&(i==4|i==5))end<-1;
if(N==80&(i==3|i==4|i==5|i==6))end<-1;
if(N==120&(i==2|i==3|i==4|i==5|i==6|i==7))end<-1;
xx<-cbind(sortx[begin+t+1:(m+end)],p[begin+t+1:(m+end)]);
xx[,2]<-qsev(xx[,2]);
x<-rbind(x,apply(xx,2,mean));
begin<-begin+m+end;
}

if(N==40)lin<-lm(x[,1]~x[,2],weight=c(m,m,m,m+1,m+1,m,m,m));
if(N==80)lin<-lm(x[,1]~x[,2],weight=c(m,m,m+1,m+1,m+1,m+1,m,m));
if(N==120)lin<-lm(x[,1]~x[,2],weight=c(m,m+1,m+1,m+1,m+1,m+1,m+1,m));
ols<-lin$coef;

return(ols)
}

bbias1<-c();
bbias2<-c();
vvar1<-c();
```

```
vvar2<-c();
mmse1<-c();
mmse2<-c();
for(j in 2:length(nn)){

n<-nn[j];

muhat<-c();
sigmahat<-c();
set.seed(106);
for(i in 1:Nsim){
samp<-rsev(n,mu,sigma);
muhat[i]<-estim(samp)[1];
sigmahat[i]<-estim(samp)[2];
}

write.csv(cbind(muhat,sigmahat),paste("SEVcen10.tmd10%&8cmean",n,".csv",sep=""));
bbias1[j]<-mean(muhat)-mu;
bbias2[j]<-mean(sigmahat)-sigma;
vvar1[j]<-var(muhat);
vvar2[j]<-var(sigmahat);
mmse1[j]<-mean((muhat-mu)^2);
mmse2[j]<-mean((sigmahat-sigma)^2);
}

dum<-cbind(bbias1,vvar1,mmse1,bbias2,vvar2,mmse2);
colnames(dum)<-c("Bias1","Var1","MSE1","Bias2","Var2","MSE2");
rownames(dum)<-c("n=20","n=40","n=80","n=120");
write.csv(dum,file="SumSEVcen10.tmd10%&8cmean.csv");
```

ประวัติผู้เขียนวิทยานิพนธ์

นางสาวขวัญรัตน์ ตั้งพิษฐานสกุล เกิดวันที่ 19 พฤศจิกายน พ.ศ.2523 ที่อำเภอจอมทอง จังหวัดเชียงใหม่ สำเร็จการศึกษาปริญญาวิทยาศาสตรบัณฑิต (วท.บ.) ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ มหาวิทยาลัยเชียงใหม่ ในปีการศึกษา 2545 และเข้าศึกษาต่อในหลักสูตรสถิติศาสตรมหาบัณฑิต (สต.ม.) สาขาวิชาสถิติ ภาควิชาสถิติ คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย ในปีการศึกษา 2552