

ต้นฉบับ หน้าขาดหาย

A COMPARISON OF PARAMETER ESTIMATION FOR LOSS DISTRIBUTIONS  
WITH LEFT AND RIGHT TRUNCATED GROUP DATA

Miss Tosaporn Talangtam

A Thesis submitted in Partial fulfilment of the Requirements

for the Degree of Master of Science

Department of Statistics

Graduate School

Chulalongkorn University

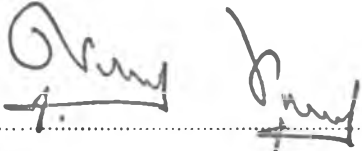
Academic Year 1996

ISBN 974-636-147-3


หัวข้อวิทยานิพนธ์      การเปรียบเทียบการประมาณค่าพารามิเตอร์ของการแจกแจงความเสียหาย  
ที่มีข้อมูลเป็นแบบกลุ่มถูกตัดปลายทางซ้ายและขวา  
โดย                      นางสาวทศพร แดงธรรม  
ภาควิชา                  สถิติ  
อาจารย์ที่ปรึกษา      ผู้ช่วยศาสตราจารย์ ร้อยเอก มานพ วราภักดิ์

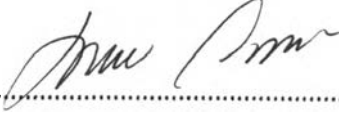
---

บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้บัณฑิตวิทยาลัยนี้เป็นส่วนหนึ่ง  
ของการศึกษาตามหลักสูตรปริญญามหาบัณฑิต


  
..... คณบดีบัณฑิตวิทยาลัย  
( ศาสตราจารย์ นายแพทย์ ศุภวัฒน์ สุติวงศ์ )

คณะกรรมการสอบวิทยานิพนธ์

  
..... ประธานกรรมการ  
( รองศาสตราจารย์ มัลลิกา นูนาค )

  
..... อาจารย์ที่ปรึกษา  
( ผู้ช่วยศาสตราจารย์ ร้อยเอก มานพ วราภักดิ์ )

  
..... กรรมการ  
( รองศาสตราจารย์ จีสิทธิ์ โกลากุล )

  
..... กรรมการ  
( ผู้ช่วยศาสตราจารย์ เสาวรส ใหญ่สว่าง )

พิมพ์ต้นฉบับบทคัดย่อวิทยานิพนธ์ภายในกรอบสี่เหลี่ยมนี้เพียงแผ่นเดียว

ทศพร แดงธรรม : การเปรียบเทียบการประมาณค่าพารามิเตอร์ของการแจกแจงความเสียหาย ที่มี  
ข้อมูลเป็นแบบกลุ่มถูกตัดปลายทางซ้ายและขวา (A COMPARISON OF PARAMETER ESTIMATION FOR  
LOSS DISTRIBUTIONS WITH LEFT AND RIGHT TRUNCATED GROUP DATA)

อาจารย์ที่ปรึกษา : ผศ. ร.อ. มานพ วรวิทย์, 171 หน้า. ISBN 974-636-147-3.



การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อศึกษาและเปรียบเทียบการประมาณค่าพารามิเตอร์ สำหรับข้อมูลแบบกลุ่มที่ถูก  
ตัดปลายทางซ้ายและขวา ซึ่งเปรียบเทียบค่าความคลาดเคลื่อนในรูปรากที่สองของค่าคลาดเคลื่อนกำลังสองเฉลี่ย (RMSE)  
การเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์ทางสถิติ 3 วิธี คือ วิธีกำลังสองน้อยที่สุด วิธีภาชนะน่าจะเป็นสูงสุด และวิธีโค-  
สแควร์ต่ำสุด การแจกแจงที่ศึกษามีลักษณะเบ้ขวา คือ การแจกแจงไวบูลล์ และการแจกแจงลอกนอร์มอล

ขนาดตัวอย่างเป็น 100 300 500 700 และ 1,000 ตามลำดับ ค่าตัดปลายทางซ้ายเป็น 1,000 และ 2,000 ค่า  
ตัดปลายทางขวา 100,000 150,000 และ 200,000 เปอร์เซ็นต์การตัดปลายทางขวาของข้อมูลเท่ากับ 5% 10% 15% 20%  
25% และ 30% ของขนาดตัวอย่าง สำหรับข้อมูลลักษณะต่าง ๆ ที่ใช้ในการทดลองนี้ จำลองด้วยโปรแกรมคอมพิวเตอร์โดย  
ใช้เทคนิคมอนติคาร์โล และกระทำซ้ำกัน 1,000 ครั้งในแต่ละกรณี

ผลการศึกษากการเปรียบเทียบการประมาณค่าพารามิเตอร์สำหรับข้อมูลแบบกลุ่มที่ถูกตัดปลายทางซ้ายและขวา  
สรุปผลได้ดังนี้

สำหรับข้อมูลที่มีการแจกแจงแบบไวบูลล์และลอกนอร์มอล ทุก ๆ ขนาดตัวอย่าง ทุกระดับค่าตัดปลายทางซ้าย  
ค่าตัดปลายทางขวา และทุก ๆ เปอร์เซ็นต์การตัดปลายทางขวาของข้อมูล วิธีภาชนะน่าจะเป็นสูงสุด จะให้ค่า RMSE ต่ำสุด  
รองลงมาคือวิธี โค-สแควร์ต่ำสุด และวิธีกำลังสองน้อยที่สุด ตามลำดับ ซึ่งการแจกแจงแบบไวบูลล์ วิธีโค-สแควร์ต่ำสุดจะให้ค่า  
RMSE ที่ใกล้เคียงกับวิธีภาชนะน่าจะเป็นสูงสุด

ภาควิชา ..... สถิติ  
สาขาวิชา ..... การประกันภัย  
ปีการศึกษา 2539.....

ลายมือชื่อนักศึกษา ทศพร แดงธรรม  
ลายมือชื่ออาจารย์ที่ปรึกษา .....  
ลายมือชื่ออาจารย์ที่ปรึกษาร่วม .....

## C724068 : MAJOR INSURANCE

KEY WORD: LOSS DISTRIBUTIONS / LEFT AND RIGHT TRUNCATED / GROUP DATA

TOSAPORN TALANGTAM: A COMPARISON OF PARAMETER ESTIMATION FOR LOSS DISTRIBUTIONS WITH LEFT AND RIGHT TRUNCATED GROUP DATA.

THESIS ADVISOR : ASST.PROF.CAPT. MANOP VARAPHAUDI, 171 pp. ISBN 974-636-147-3.

The objective of this study is to compare the parameter estimations for loss distributions with left and right truncated group data by comparing RMSE (square root of mean square error). The methods of parameter estimation are Least Squares, Maximum Likelihood, and Minimum Chi-Square. The distributions are used in case of skew distribution functions. They are Weibull Distribution and Lognormal Distribution.

Sample sizes are 100, 300, 500, 700 and 1,000. Left truncated points are 1,000 and 2,000, right truncated points are 100,000, 150,000, and 200,000. Percentages of right truncated data are 5% 10% 15% 20% 25% and 30% of sample sizes. Data are obtained through simulation using Monte Carlo Technique, repeat 1,000 times.

The results of the study are as follows:

For Weibull and Lognormal Distributions, any sample size, each left truncated point, each right truncated point, and each percentage of right truncated data, the RMSE of Maximum Likelihood Method is the lowest, following by Minimum Chi-Square Method and Least-Squares Method respectively. In Weibull Distribution, the RMSE of Minimum Chi-Square and Maximum Likelihood Method are rather the same.

ภาควิชา.....สถิติ

สาขาวิชา.....การประกันภัย

ปีการศึกษา.....2539

ลายมือชื่อนิสิต..... ทศพร แภคทองกุล

ลายมือชื่ออาจารย์ที่ปรึกษา.....

ลายมือชื่ออาจารย์ที่ปรึกษาร่วม..... -

## กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปได้ด้วยความช่วยเหลืออย่างดียิ่งของ ผู้ช่วยศาสตราจารย์ ร้อยเอก มานพ วรภักดิ์ อาจารย์ที่ปรึกษาวิทยานิพนธ์ ซึ่งท่านได้ให้คำแนะนำ และข้อคิดเห็นต่าง ๆ ของการวิจัยมาด้วยดีตลอด จึงขอกราบขอบพระคุณอาจารย์มา ณ ที่นี้ด้วย

ผู้วิจัยขอกราบขอบพระคุณ รองศาสตราจารย์ มัลลิกา บุญนาค ในฐานะประธานกรรมการสอบวิทยานิพนธ์ รองศาสตราจารย์ จสิทธิ์ โกลากุล และ ผู้ช่วยศาสตราจารย์ เสาวรส ใหญ่สว่าง กรรมการสอบวิทยานิพนธ์ที่กรุณาตรวจแก้วิทยานิพนธ์ฉบับนี้ให้สมบูรณ์ยิ่งขึ้น

ท้ายนี้ขอกราบขอบพระคุณ คุณพ่อพันโททองพันธ์ และคุณแม่เพ็ญศรี แดลงธรรม ที่ส่งเสริมและสนับสนุนในการศึกษาผู้วิจัยเสมอมา และขอกราบขอบพระคุณอาจารย์ทุกท่านที่ประสิทธิ์ประสาทวิชาความรู้ให้โดยตลอดจนสำเร็จการศึกษา

ทศพร แดลงธรรม

## สารบัญ

|   | หน้า |
|---|------|
| บทคัดย่อภาษาไทย .....                             | ๔    |
| บทคัดย่อภาษาอังกฤษ .....                          | ๑    |
| กิตติกรรมประกาศ .....                             | ๖    |
| สารบัญตาราง .....                                 | ๗    |
| สารบัญรูป .....                                   | ๗    |
| สัญลักษณ์ต่าง ๆ .....                             | ๘    |
| <br>  |      |
| บทที่ 1 บทนำ .....                                | 1    |
| ที่มาและความสำคัญของปัญหา .....                   | 1    |
| วัตถุประสงค์ของการวิจัย .....                     | 3    |
| สมมติฐานของการวิจัย .....                         | 3    |
| ข้อตกลงเบื้องต้น .....                            | 3    |
| ขอบเขตการวิจัย .....                              | 4    |
| เกณฑ์การตัดสินใจ .....                            | 6    |
| ประโยชน์ที่คาดว่าจะได้รับ .....                   | 6    |
| <br>  |      |
| บทที่ 2 สถิติที่ใช้ในการวิจัย .....               | 7    |
| การแจกแจงที่ใช้ในการวิจัย .....                   | 7    |
| วิธีการประมาณค่าพารามิเตอร์ที่ใช้ในการวิจัย ..... | 10   |
| <br>  |      |
| บทที่ 3 วิธีการดำเนินการวิจัย .....               | 19   |
| ลักษณะข้อมูล .....                                | 19   |
| วิธีการจัดกลุ่มข้อมูล .....                       | 19   |
| วิธีการวิจัย .....                                | 21   |
| การเปรียบเทียบวิธีการประมาณค่า .....              | 34   |
| วิธีวิเคราะห์ข้อมูล .....                         | 35   |

## สารบัญ (ต่อ)

|   | หน้า |
|---|------|
| โปรแกรมที่ใช้ในการวิจัย .....             | 35   |
| บทที่ 4 ผลการวิจัย .....                  | 39   |
| ข้อมูลมีการแจกแจงแบบไวบูลล์ .....         | 35   |
| ข้อมูลมีการแจกแจงแบบลอกนอรัมอล .....      | 81   |
| บทที่ 5 สรุปผลการวิจัยและข้อเสนอแนะ ..... | 122  |
| สรุปผลการวิจัย .....                      | 122  |
| ข้อเสนอแนะ .....                          | 123  |
| รายการอ้างอิง .....                       | 125  |
| ภาคผนวก .....                             | 126  |
| ภาคผนวก ก .....                           | 126  |
| ภาคผนวก ข .....                           | 131  |
| ภาคผนวก ค .....                           | 134  |
| ภาคผนวก ง .....                           | 142  |
| ประวัติผู้วิจัย .....                     | 171  |





## สารบัญรูป

|   | หน้า |
|---|------|
| รูปที่ 1.1 แสดงฟังก์ชันการแจกแจงแบบไวบูลล์ที่มีค่าเฉลี่ยของการแจกแจงเท่ากับ 5 .....   | 5    |
| รูปที่ 1.2 แสดงฟังก์ชันการแจกแจงแบบลอกนอรัมอลที่มีค่าเฉลี่ยของการแจกแจงเท่ากับ 5 .....  | 6    |
| รูปภาพ 2.1 ฟังก์ชันการแจกแจงไวบูลล์ .....   | 8    |
| รูปภาพ 2.1 ฟังก์ชันการแจกแจงลอกนอรัมอล .....  | 9    |
| ผังงานของโปรแกรมที่ใช้ในการวิจัย .....  | 36   |
| รูปที่ 4.1 แสดงค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (RMSE) กับ<br>จำนวนตัวอย่าง ( $n$ ) สำหรับการแจกแจงแบบไวบูลล์ $d=0.1$ $w=10.0$ .....       | 48   |
| รูปที่ 4.2 แสดงค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (RMSE) กับ<br>จำนวนตัวอย่าง ( $n$ ) สำหรับการแจกแจงแบบไวบูลล์ $d=0.1$ $w=15.0$ .....       | 51   |
| รูปที่ 4.3 แสดงค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (RMSE) กับ<br>จำนวนตัวอย่าง ( $n$ ) สำหรับการแจกแจงแบบไวบูลล์ $d=0.1$ $w=20.0$ .....       | 54   |
| รูปที่ 4.4 แสดงค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (RMSE) กับ<br>จำนวนตัวอย่าง ( $n$ ) สำหรับการแจกแจงแบบไวบูลล์ $d=0.2$ $w=10.0$ .....       | 57   |
| รูปที่ 4.5 แสดงค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (RMSE) กับ<br>จำนวนตัวอย่าง ( $n$ ) สำหรับการแจกแจงแบบไวบูลล์ $d=0.2$ $w=15.0$ .....       | 60   |
| รูปที่ 4.6 แสดงค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (RMSE) กับ<br>จำนวนตัวอย่าง ( $n$ ) สำหรับการแจกแจงแบบไวบูลล์ $d=0.2$ $w=20.0$ .....       | 63   |
| รูปที่ 4.7 แสดงค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย(RMSE) กับเปอร์เซ็นต์<br>การตัดปลายหางขวา สำหรับการแจกแจงแบบไวบูลล์ $d=0.1$ $w=10.0$ .....  | 66   |
| รูปที่ 4.8 แสดงค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย(RMSE) กับเปอร์เซ็นต์<br>การตัดปลายหางขวา สำหรับการแจกแจงแบบไวบูลล์ $d=0.1$ $w=15.0$ .....  | 68   |
| รูปที่ 4.9 แสดงค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย(RMSE) กับเปอร์เซ็นต์<br>การตัดปลายหางขวา สำหรับการแจกแจงแบบไวบูลล์ $d=0.1$ $w=20.0$ .....  | 71   |
| รูปที่ 4.10 แสดงค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย(RMSE) กับเปอร์เซ็นต์<br>การตัดปลายหางขวา สำหรับการแจกแจงแบบไวบูลล์ $d=0.2$ $w=10.0$ ..... | 73   |



## สารบัญรูป (ต่อ)

|  | หน้า |
|--|------|
| รูปที่ 5.1 แผนผังแสดงลำดับขั้นตอนของการนำไปใช้ประโยชน์ ..... | 124  |

## สัญลักษณ์ต่าง ๆ

|                          |   |
|--------------------------|---|
| $n$                      | จำนวนตัวอย่าง (จำนวนกรมธรรม์)   |
| $d$                      | ค่าตัดปลายทางซ้าย (ค่ารับผิดส่วนแรก)  |
| $w$                      | ค่าตัดปลายทางขวา (ค่ารับผิดสูงสุด)  |
| $k$                      | จำนวนชั้นข้อมูล (จำนวนช่วงความเสียหาย)  |
| $F(x)$                   | ฟังก์ชันการแจกแจง (สะสม)  |
| $F_n(x)$                 | การแจกแจงเชิงการทดลอง (Empirical Distribution)  |
| $f(x)$                   | ฟังก์ชัน (ความน่าจะเป็น) ความหนาแน่น  |
| $f(x \mid d < x \leq w)$ | ฟังก์ชันความหนาแน่นที่มีเงื่อนไข เมื่อกำหนด $d < x \leq w$                                    |
| $Y$                      | ตัวแปรสุ่มของความเสียหายที่มีการบันทึก  |
| $X$                      | ตัวแปรสุ่มของความเสียหายที่เกิดขึ้นจริง   |
| $F_Y(x)$                 | ฟังก์ชันความหนาแน่นสะสมของ $Y$  |
| $F_X(x)$                 | ฟังก์ชันความหนาแน่นสะสมของ $X$  |
| $f_Y(x)$                 | ฟังก์ชันความหนาแน่นของ $Y$  |
| $f_X(x)$                 | ฟังก์ชันความหนาแน่นของ $X$  |
| $(c_{i-1}, c_i]$         | ช่วงความเสียหาย $i = 1, 2, \dots, k$  |
| $f_i$                    | จำนวนกรมธรรม์ในกลุ่มชั้นที่ $i$   |
| $E[X]$                   | ค่าเฉลี่ยหรือค่าคาดหวังของ $X$  |
| $Var[X]$                 | ความแปรปรวนของ $X$  |
| $\phi(x)$                | หมายถึง ฟังก์ชันความหนาแน่น $\exp\left(\frac{-x^2}{2}\right) / \sqrt{2\pi}$                   |
| $\Phi(x)$                | หมายถึง ฟังก์ชันการแจกแจง $\int_{-\infty}^x \exp\left(\frac{-y^2}{2}\right) dy / \sqrt{2\pi}$ |
| $U(a,b)$                 | การแจกแจงสม่ำเสมอในช่วงเปิด $(a,b)$   |
| $\Gamma(\alpha)$         | ฟังก์ชันแกมมา $\Gamma(\alpha) = (\alpha - 1)!$ เมื่อ $\alpha$ เป็นจำนวนเต็ม $\geq 1$          |
| $\alpha!$                | $(\alpha)(\alpha - 1)\dots\dots\dots(1)$  |
| $N(\mu, \sigma^2)$       | การแจกแจงปกติซึ่งมีค่าเฉลี่ย $\mu$ และความแปรปรวน $\sigma^2$                                  |