บทที่ 1

บทนำ

1.1 ความสำคัญและความเป็นมาของปัญหา

ในการประมาณค่าเพื่อคาดคะเนเหตุการณ์ล่วงหน้าหรือการพยากรณ์ ผู้วิจัยมักเลือกวิธีการวิเคราะห์ความถดถอย (regression analysis) การวิเคราะห์ความถดถอยพหุเป็นกรณีหนึ่งของการวิเคราะห์ความถดถอยเชิงเส้น เมื่อตัวแปรตามที่ใช้ในการศึกษามีความสัมพันธ์กับปัจจัยอื่นๆ ซึ่งเราเรียกว่า**ตัวแปรอิสระ** และสามารถเขียนในรูปของตัวแบบทั่วไปของการวิเคราะห์ความถดถอยพหุเชิงเส้นดังนี้

(1.1.1) $$\underset{\sim}{y} = X\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$$

เมื่อ $\underset{\sim}{y}$ เป็นเวกเตอร์ของตัวแปรตามขนาด n x 1 โดยที่ n เป็นจำนวนค่าสังเกต

X เป็นเมตริกซ์ของตัวแปรอิสระขนาด n x p $(p < n)$ และมี full rank p

$\underset{\sim}{\beta}$ เป็นเวกเตอร์ของพารามิเตอร์ที่ไม่ทราบค่า ขนาด p x 1

และ $\underset{\sim}{\varepsilon}$ เป็นเวกเตอร์ของความผิดพลาด ขนาด n x 1

ภายใต้ข้อสมมติดังนี้

$$E(\underset{\sim}{\varepsilon}) = \underset{\sim}{0}$$

$$E(\underset{\sim}{\varepsilon}\underset{\sim}{\varepsilon}') = \sigma^2 I_n$$

วิธีที่นิยมใช้ในการประมาณค่าสัมประสิทธิ์การถดถอยพหุ คือวิธีกำลังสองน้อยที่สุด (least square method) ซึ่งตัวประมาณกำลังสองน้อยที่สุดอยู่ในรูปของ

(1.1.2) $$\hat{\underset{\sim}{\beta}} = (X'X)^{-1}X'\underset{\sim}{y}$$

โดยที่ $\hat{\underset{\sim}{\beta}}$ จะเป็นตัวประมาณที่ไม่เอนเอียง (unbniased estimator) ซึ่งมีความแปรปรวนต่ำสุดในบรรดาตัวประมาณที่ไม่เอนเอียง

ในกรณีที่ค่าสุ่มตัวอย่างที่ได้มาจากประชากรที่มีการแจกแจงแบบเบ้ และการแจกแจงแบบหางยาวกว่าปกติ (skewed distribution and long tailed distribution) วิธีกำลังสองน้อยที่สุดอาจจะไม่เหมาะสม เพราะวิธีนี้มีความไวต่อข้อมูลที่ผิดปกติ และสูญเสียประสิทธิภาพไปเมื่อการแจกแจงของความผิดพลาดไม่เป็นแบบปกติ

วิธีแก้ปัญหาที่เกิดขึ้นนี้วิธีหนึ่งที่รู้จักกันดีคือ M-estimator ซึ่งอยู่ในรูปของค่าน้อยที่สุดของผลรวมของฟังก์ชันความผิดพลาดซึ่งสามารถเขียนในรูปของ

(1.1.3) $$\min \rho(\varepsilon_i/S) = \min \Sigma\rho\{(y_i - X_i\beta)/S\}$$

เมื่อ ρ คือฟังก์ชันที่เหมาะสมซึ่งถูกเลือกขึ้นมา และ S เป็นค่าประมาณค่าหนึ่งของสเกล (scale) ของความคลาดเคลื่อน

ในปี ค.ศ. 1977 Huber ได้เสนอรูปแบบของค่าน้อยที่สุดของฟังก์ชันความผิดพลาดสำหรับ $\underset{\sim}{\beta}$ และสเกล ซึ่งอยู่ในรูปของ

(1.1.4) $$Q(\underset{\sim}{\beta}, \sigma) = (1/n) \sum_{i}^{n} \rho\{(y_i - \underset{\sim}{X}_i'\underset{\sim}{\beta})/\sigma\}\sigma$$

เมื่อ ρ คือฟังก์ชันที่เหมาะสม โดยที่ $\rho > 0$ และ $\rho(0) = 0$

ถ้าเราหาอนุพันธ์บางส่วนเทียบกับ β และ σ จะได้สมการดังนี้

(1.1.5) $$\Sigma x_{ij}\, \psi\{(y_i - \underset{\sim}{X}_i'\underset{\sim}{\beta})/S\} = 0$$

(1.1.6) $$\Sigma \chi\{(y_i - \underset{\sim}{X}_i'\underset{\sim}{\beta})/S\} = 0$$

เมื่อกำหนด $\psi = \rho'$ และ $\chi = \varepsilon_i\psi - \rho$ โดยที่ χ มีค่าสมบูรณ์ (absolute) น้อยที่สุดถ้า $\varepsilon_i = 0$

การประมาณ $\underset{\sim}{\beta}$ และสเกลโดยการ iteration สมการที่ (1.1.5) และ (1.1.6) พร้อมๆ กัน (simultaneous estimation and scale) จะกระทำได้ยากในทางปฏิบัติ

โดยทั่วไปค่า S ที่ประมาณขึ้นมาจะถูกเลือกเพื่อไม่ให้มีอิทธิพลกับความคลาดเคลื่อนที่มีขนาดใหญ่ ดังนั้นฟังก์ชันความผิดพลาดต่างๆ จะมีค่าความคลาดเคลื่อนน้อยที่สุดเมื่อกำหนดค่า S คงที่ และเหตุผลในการประมาณ S อีกข้อหนึ่งคือ ถ้ามีตัวประมาณสำหรับ y_i ที่คงเส้นคงวาจะสามารถประมาณสเกล S ที่คงเส้นคงวาและเหมาะสมกับ ε_i และสามารถนำค่าไปใช้ในสมการ (1.1.5) สำหรับประมาณค่า $\underset{\sim}{\beta}$ ที่เหมาะสม

ในการประมาณค่า $\underset{\sim}{\beta}$ เมื่อค่าสเกลคงที่ Andrew et.at. (1972) ได้เสนอตัวประมาณสเกลที่ใช้ค่าสมบูรณ์ของความคลาดเคลื่อน (the median absolute deviation(MAD)) ซึ่งอยู่ในรูปของ

(1.1.7) $$S = \text{med}|\varepsilon_i - \text{med}(\varepsilon_i)| \;/\; 0.6745$$

แต่ค่าประมาณ $\underset{\sim}{\beta}$ ที่ได้ยังไม่ดีพอ เหตุผลเหล่านี้สนับสนุนโดยค่าจริงที่ได้จาก MAD ที่เป็นค่าที่มากที่สุดในกลุ่มที่มีความเอนเอียงต่ำ และมี breakdown point* ที่เป็นไปได้สูงเท่ากับ 0.5

ในปี ค.ศ. 1981 Huber ได้เสนอข้อคิดเห็นเมื่อค่าสังเกตมีการกระจายแบบปลอมปน และค่าสังเกตที่ได้จากกรณี asymmetric distribution โดยที่การเลือกตัวประมาณ S ที่เหมาะสมขึ้นกับ

1. ขอบเขตของ breakdown point ที่มีขนาดใหญ่

---

*breakdown point หรือ breakdown bound (Hampel,1968) ของตัวประมาณคือส่วนหนึ่งของค่าสังเกตที่ใหญ่ที่สุดที่เป็นไปได้ โดยที่ค่าสังเกตนั้นมีขอบเขตอยู่บนค่าประมาณที่เปลี่ยนแปลงได้ เมื่อส่วนหนึ่งของตัวอย่างนั้นเปลี่ยนแปลงไปโดยไม่มีข้อจำกัด และค่า breakdown bound ของตัวประมาณควรมีค่าไม่มากกว่า 0.5

2. มีความเอนเอียง (bias) ไม่มาก

3. มีความแปรปรวน (variance) ไม่มาก

นอกจากนี้ในกรณีที่ค่าสังเกตมีการแจกแจงแบบสมมาตร (symmetric distribution) ตัวประมาณสเกลจะมีผลเมื่อค่าสังเกตมีการปลอมปนสูง หรือกรณีที่มีหางหนา (heavy-tailed)

ในปี ค.ศ. 1983 Hoaglis D.C. ได้เปรียบเทียบค่าสเกล 4 ชนิดได้แก่ standard deviation fourth-spread MAD และ mean absolute deviation from sample median (AD) โดยพิจารณาจากประสิทธิภาพเมื่อค่าสังเกตมีการกระจายแบบสมมาตรและหางหนา (heavy-tailed symmetric) และสรุปว่าค่าสเกลจะช่วยให้ได้ค่าประมาณตำแหน่ง (location) ที่ดีขึ้น

และในปี ค.ศ. 1985 David A. Lax ได้ศึกษาเกณฑ์ความแกร่งของสเกล 17 ชนิด เมื่อค่าสังเกตมาจากประชากรเดี่ยว ซึ่งมีการแจกแจงแบบสมมาตรหางยาว (long-tailed symmetric distribution) และได้เสนอแนะว่า ตัวสเกลที่ดีจะเป็นพื้นฐานที่ดีในการหาค่าประมาณตำแหน่ง

1.2 <u>วัตถุประสงค์ของการวิจัย</u>

1.2.1 เพื่อเปรียบเทียบประสิทธิภาพของตัวประมาณที่ได้จากการประมาณค่าด้วยวิธีกำลังสองน้อยที่สุดกับวิธี M-estimator ในกรณีที่วิธี M-estimator มีการเปลี่ยนค่าประมาณสเกล

1.3 <u>ข้อตกลงเบื้องต้น</u>

1.3.1 ค่าผิดพลาดเป็นตัวแปรสุ่มที่มีการแจกแจงเป็นอิสระซึ่งกันและกัน (identically independent distribution)

1.3.2 การวิจัยครั้งนี้ใช้เกณฑ์การประมาณค่าสัมประสิทธิ์การถดถอยพหุภายใต้ค่าผิดพลาดที่มีการแจกแจงแบบเบ้และการแจกแจงที่มีหางยาวกว่าการแจกแจงปกติ ซึ่งให้ค่าอัตราส่วน-

ผลต่างของค่าเฉลี่ยความผิดพลาดกำลังสอง (ratio of different average mean square error (RDAMSE)) และอัตราส่วนผลต่างของความผิดพลาดกำลังสอง (ratio of different mean square error (RDMSE)) ต่ำสุดจะเป็นวิธีที่เหมาะสมสำหรับแต่ละสถานการณ์

1.4 ขอบเขตการวิจัย

1.4.1 ลักษณะการแจกแจงของความผิดพลาดที่ศึกษามีดังนี้

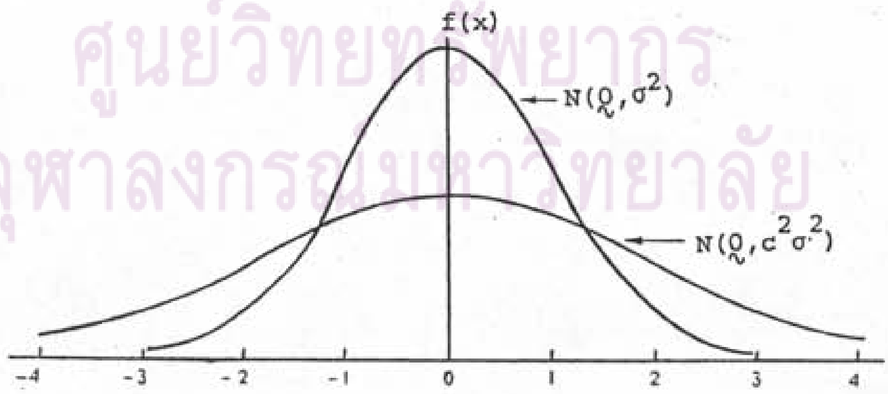
1.4.1.1 การแจกแจงแบบปกติปลอมปน (scale-contaminated normal distribution)

ฟังก์ชันการแจกแจงอยู่ในรูปของ

$$F = (1 - p)\, N(0\ ,\sigma^2) + p\, N(0\ ,c^2\sigma^2)$$

เมื่อ c คือสเกลแฟกเตอร์ (scale factor) ซึ่งถ้ามีค่าสูงจะทำให้ค่าสังเกตที่ผิดปกติมีค่าสูงด้วย ในที่นี้ใช้ c = 5 10 และ 15

และ p คือเปอร์เซนต์การปลอมปน (percent of contamination) ในที่นี้ใช้ p = 5 10 20 และ 30



รูปที่ 1.4.1 แสดงเส้นโค้งของการแจกแจงแบบปกติปลอมปน ณ p และ c

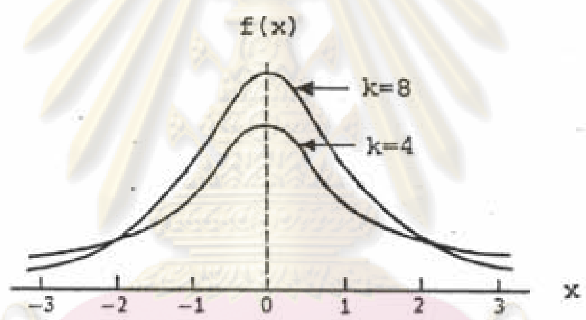
1.4.1.2 <u>การแจกแจงแบบที</u> (t-distribution)

ฟังก์ชันการแจกแจงอยู่ในรูปของ

$$T = X/(\sqrt{Y/k})$$

เมื่อ X มีการแจกแจงแบบปกติมาตรฐาน (N(0,1)) และ Y มีการแจกแจงแบบไคสแควร์ โดยที่ X และ Y เป็นอิสระซึ่งกันและกัน

k เป็นระดับความเป็นอิสระ (ร.ส.) และจะศึกษาในกรณีที่ k = 4 และ 8 โดยที่ n = 20 เท่านั้น



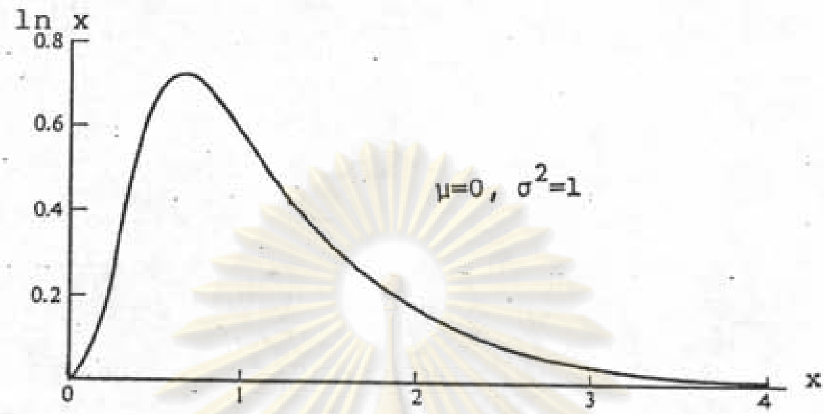
รูปที่ 1.4.2 แสดงเส้นโค้งของการแจกแจงแบบที ณ DF = 4 และ 8

1.4.1.3 <u>การแจกแจงแบบลอกนอร์มอล</u> (lognormal distribution)

ฟังก์ชันความหนาแน่นอยู่ในรูปของ

$$f(x) = \begin{cases} \dfrac{1}{x\sigma\sqrt{2\pi}} \exp -\{ \frac{1}{2}(\ln x-\mu)^2/\sigma^2 \} & ; x>0, \sigma>0, \mu \in \mathbb{R} \\ 0 & ; \text{อื่น ๆ} \end{cases}$$

เมื่อ μ และ σ^2 เป็นค่าเฉลี่ยและความแปรปรวนของ Y โดยที่ Y = ln X และ Y มีการแจกแจงแบบปกติ ในที่นี้พิจารณาใช้ $\mu = 0, \sigma^2 = 1$



รูปที่ 1.4.3 แสดงเส้นโค้งของการแจกแจงแบบลอกนอร์มอล

1.4.1.4 <u>การแจกแจงแบบแกมมา</u> (gamma - distribution) ฟังก์ชันความหนาแน่นอยู่ในรูปของ

$$f(x) = \begin{cases} \dfrac{x^{\alpha-1}\exp\{-x/\beta\}}{\beta^{\alpha}\,\Gamma(\alpha)} & ;\ x>0, \alpha>0, \beta>0 \\ 0 & ;\ \text{อื่น ๆ} \end{cases}$$

เมื่อ β เป็น scale parameter

และ α เป็น shape parameter

จะได้ว่า $E(X) = \beta\alpha$

$Var(X) = \beta^2\alpha$

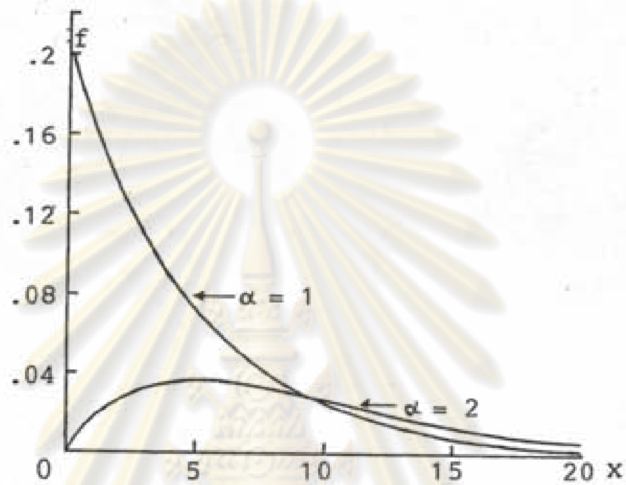
Coefficient of Variance (C.V.) = $1/\sqrt{\alpha}$

ในการวิจัยครั้งนี้จะศึกษา ณ ค่า β = 5, 10 และ 150 เมื่อ α = 1 และ 2 กล่าวคือ

C.V.(X) = 100% (β = 5 10 และ 150, α = 1)

C.V.(X) = 70% (β = 5 10 และ 150, α = 2)

และสาเหตุที่เลือกใช้ค่า C.V. = 100% และ 70% โดยที่ไม่เลือกค่า C.V. ที่มีค่าต่ำกว่านี้เพราะจากการพิจารณากราฟที่แสดงเส้นโค้งของการแจกแจงแบบแกมมา ถ้าค่า C.V. มีค่าต่ำกว่านี้กราฟของการแจกแจงจะลู่เข้าสู่การแจกแจงแบบปกติมากขึ้น



รูปที่ 1.4.4 แสดงเส้นโค้งของการแจกแจงแบบแกมมา ณ β = 5 และ α = 1 และ 2

1.4.1.5 <u>การแจกแจงแบบไวบูลล์</u> (weibull distribution)

ฟังก์ชันความหนาแน่นอยู่ในรูปของ

$$f(x) = \begin{cases} \dfrac{x^{\alpha-1} \exp\{-x/_{\beta}\}^2}{\beta^{\alpha}} & ; \; x>0, \alpha>0, \beta>0 \\ 0 & ; \; \text{อื่น ๆ} \end{cases}$$

เมื่อ β เป็น scale parameter

และ α เป็น shape parameter

จะได้ว่า $E(X) = \beta$

$Var(X) = \beta^2$

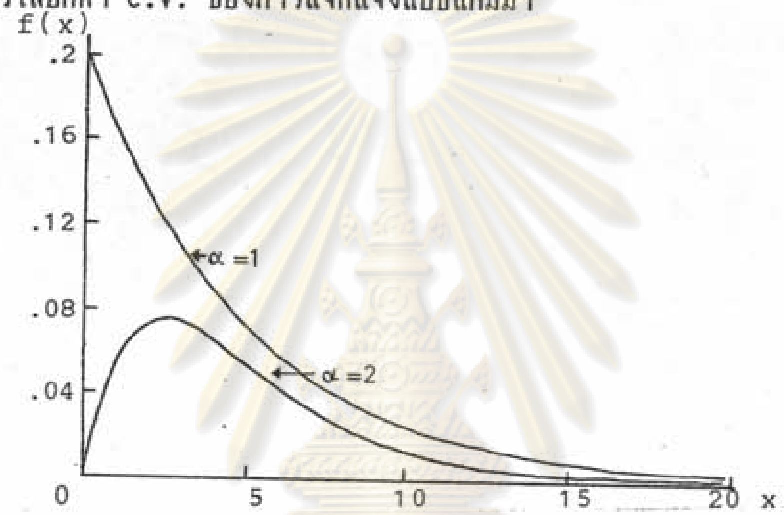
$$C.V.(X) = \left[\frac{\Gamma(1+2/\alpha)}{\Gamma^2(1+1/\alpha)} - 1\right]^{1/2}$$

ในการวิจัยครั้งนี้จะพิจารณา ณ.ค่า $\beta = 5$ 10 และ 150, $\alpha = 1$ และ 2 กล่าวคือ

C.V.(X) = 100% ($\beta = 5$ 10 และ 150, $\alpha = 1$)

C.V.(X) = 52% ($\beta = 5$ 10 และ 150, $\alpha = 2$)

และในการเลือกใช้ค่า C.V. = 100% และ 52% โดยที่ไม่เลือกค่า C.V. ที่มีค่าต่ำกว่านี้ด้วยเหตุผลเดียวกับการเลือกค่า C.V. ของการแจกแจงแบบแกมมา



รูปที่ 1.4.5 แสดงเส้นโค้งของการแจกแจงแบบไวบูลล์ ณ $\beta = 5$ และ $\alpha = 1$ และ 2

1.4.2 <u>จำนวนตัวแปรอิสระและขนาดตัวอย่าง</u>

กรณีที่ความผิดพลาดมีการแจกแจงที่ไม่ได้อยู่ในรูปการแจกแจงแบบที จำนวนตัวแปรอิสระและขนาดตัวอย่างจะถูกกำหนดดังนี้

1.4.2.1 เมื่อจำนวนตัวแปรอิสระ $p = 3$ จะใช้ขนาดตัวอย่าง $n = 20$ ซึ่งเป็นขนาดตัวอย่างที่พบมากในงานวิจัยที่เกี่ยวกับเรื่องนี้

1.4.2.2 เมื่อจำนวนตัวแปรอิสระ $p = 5$ และ 10 จะใช้ขนาดตัวอย่าง $n = 50$ 100 และ 150

สำหรับกรณีที่ความผิดพลาดมีการแจกแจงแบบที จะใช้ขนาดตัวอย่าง $n = 20$

ซึ่งเป็นขนาดตัวอย่างที่เหมาะสมกับการแจกแจงแบบที และจะใช้จำนวนตัวแปรอิสระ $p = 3$ 5 และ 10 ตามลำดับ

1.4.3 <u>การจำลองประชากรที่ศึกษาจากตัวแบบเชิงเส้น</u>

1.4.3.1 เมื่อการแจกแจงของความผิดปกติเป็นแบบปกติปลอมปนและที่จะจำลองประชากรจากตัวแบบ (1.1.1) เมตริกซ์ของ X จะคงที่ในตัวแบบสำหรับการแจกแจงของความผิดพลาดที่กำลังศึกษาทั้งหมดโดยเมตริกซ์ X เป็นเมตริกซ์ของตัวแปรอิสระขนาด $n \times p$ การสร้างเมตริกซ์ X จะอาศัยการจำลองตัวแปรสุ่มที่มีการแจกแจงแบบปกติ $N(\mu, \sigma^2)$ เพื่อให้คล้ายกับข้อมูลตามธรรมชาติ $\underset{\sim}{\beta}$ เป็นเวกเตอร์ของสัมประสิทธิ์การถดถอยพหุของประชากรที่นำมาจากสมการถดถอยเชิงเส้นพหุที่ทำให้ค่าสหสัมพันธ์ร่วมของ X และ $\underset{\sim}{y}$ สูง $\underset{\sim}{\varepsilon}$ เป็นเวกเตอร์ของความผิดพลาดโดยมีการแจกแจงตามที่ต้องการศึกษา

1.4.3.2 เมื่อการแจกแจงของความผิดพลาดเป็นแบบเบ้ ตัวแปร $\underset{\sim}{y}$ จะต้องสร้างให้มีการแจกแจงเป็นแบบเบ้โดยตรง เมตริกซ์ X และเวกเตอร์ของ $\underset{\sim}{\beta}$ จะสร้างในทำนองเดียวกับข้อ 1.4.3.1

1.5 <u>ประโยชน์ที่คาดว่าจะได้รับ</u>

1.5.1 ผลการศึกษากระบวนการถดถอยที่แกร่งทั้งตำแหน่งและสเกล เมื่อเลือกค่าสเกลด้วยวิธีใดวิธีหนึ่งให้คงที่จะเป็นแนวทางในการศึกษากระบวนการที่แกร่งทั้งตำแหน่งและสเกล (robust location and scale estimation) ในรูปแบบอื่นๆต่อไป

1.5.2 ผลจากการศึกษาเปรียบเทียบสามารถแสดงได้ว่า ถ้ามีข้อมูลชุดหนึ่งที่มีการแจกแจงไม่เป็นปกติผู้วิจัยควรจะเลือกใช้วิธีใดในการประมาณค่าสัมประสิทธิ์การถดถอยพหุจึงจะทำให้ผลการประมาณค่าผิดพลาดน้อยที่สุด

1.6 <u>วิธีดำเนินการวิจัย</u>

1.6.1 ศึกษาวิธีการประมาณค่าสัมประสิทธิ์การถดถอยพหุในกระบวนการถดถอยพหุ และเขียนโปรแกรมจำลองค่าสังเกตของตัวแปรในตัวแบบที่ต้องการศึกษา รวมทั้งโปรแกรมสำหรับคำนวณค่าสัมประสิทธิ์การถดถอยพหุของวิธีการแต่ละวิธีดังนี้

1.6.1.1 วิธีกำลังสองน้อยที่สุด

1.6.1.2 วิธี M-estimator ที่ใช้เกณฑ์ความแกร่งของ Huber เมื่อพิจารณาตัวประมาณสเกลด้วยวิธี

ก) The Standard Deviation of Location (STD)

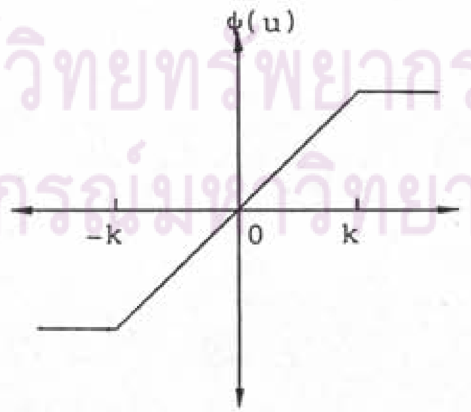
ข) The Median Absolute Deviation (MAD)

ค) The Modified Biweight A-estimator (MBA)

1.6.2 ศึกษาเปรียบเทียบวิธีการประมาณค่าสัมประสิทธิ์การถดถอยพหุด้วย วิธีกำลังสองน้อยที่สุดกับวิธี M-estimator ด้วยเกณฑ์สเกลแบบต่างๆ ข้อมูลที่ใช้ในการวิจัยครั้งนี้ได้จากการจำลองโดยใช้เทคนิค Monte Carlo Simulation และจะกระทำซ้ำ 200 ครั้งในแต่ละสถานการณ์ ยกเว้นกรณีเมื่อค่าผิดพลาดมีการแจกแจงแบบ เบ้จะกระทำซ้ำ 100 ครั้ง

1.7 ศัพท์ต่างๆที่ใช้ในการวิจัย

เกณฑ์ความแกร่งของ Huber คือสมการเชิงเส้นที่จุดกำเนิดและมีค่าคงที่ที่ปลายทั้งสองข้างพิจารณาจากรูปที่ 1.7.1



รูปที่ 1.7.1 แสดงกราฟของเกณฑ์ความแกร่ง Huber ณ k
เมื่อ k คือจุดเปลี่ยนเว้า

และมีฟังก์ชันความแกร่งดังนี้

$$\rho(r) = \begin{cases} -r^2/2 & ;\ \text{ถ้า } |r| < k \\ k|r| - k^2/2 & ;\ \text{อื่นๆ} \end{cases}$$

ซึ่งค่าประมาณของ Huber เป็นค่า maximum-likelihood estimate (MLE) ที่ดีสำหรับการแจกแจงแบบ ε-contaminate Gaussion และทำให้ค่าประมาณสเกลเป็นเงื่อนไขที่สำคัญของค่าสังเกตในการแปลงข้อมูลจากกำลังสองในรูปสมการเชิงเส้นพิจารณาจากตารางที่ 1.7.1

ตารางที่ 1.7.1 แสดงฟังก์ชันความแกร่งและตัวประมาณสเกลที่สามารถใช้ได้

Estimator with Tuning Constant(S)	Objective Function ρ	ψ - Function	ψ'	Weight Function w	Range of u	Commonly used Denominator of u
Mean	$\frac{1}{2}u^2$	u	1	1	$-\infty$ to $+\infty$	none
Median	$\|u\|$	sgn(u)	$\delta(u)^2$	$\frac{\text{sgn}(u)}{u}$	$-\infty$ to $+\infty$	none
Huber (k) 0 < k	$\frac{1}{2}u^2$	u	1	1	$\|u\| < k$	normalized F-spread or
	$k\|u\| - \frac{1}{2}k^2$	k sgn(u)	0	$\frac{k\,\text{sgn}(u)}{u}$	$\|u\| > k$	MAD
Redescending Estimators (Finite Rejection Point)						
Hampel (a,b,c) (three-part redescending) 0 < u < b < c	$\frac{1}{2}u^2$	u	1	1	$\|u\| < a$	MAD
	$a\|u\| - \frac{1}{2}a^2$	a sgn(u)	0	$\frac{a\,\text{sgn}(u)}{u}$	$a < \|u\| < b$	
	$ab - \frac{1}{2}a^2 + (c-b)\frac{a}{2}\left[1 - \left(\frac{c-\|u\|}{c-b}\right)^2\right]$	$a\frac{c-\|u\|}{c-b}\text{sgn}(u)$	$-\frac{a}{c-b}\text{sgn}(u)$	$a\frac{c-\|u\|}{c-b}\cdot\frac{\text{sgn}(u)}{u}$	$b < \|u\| < c$	
	$ab - \frac{1}{2}a^2 + (c-b)\frac{a}{2}$	0	0	0	$\|u\| > c$	
Andrews' wave(c) 0 < c	$\frac{1}{\pi^2}(1 - \cos\pi u)$	$\frac{1}{\pi}\sin\pi u$	$\cos\pi u$	$\frac{1}{\pi u}\sin\pi u$	$\|u\| < 1$	cMAD
	$\frac{2}{\pi^2}$	0	0	0		
Tukey's biweight(c) 0 < c	$\frac{1}{6}[1-(1-u^2)^3]$	$u(1-u^2)^2$	$(1-u^2)(1-5u^2)$	$(1-u^2)^2$	$\|u\| < 1$	cMAD
	$\frac{1}{6}$	0	0	0	$\|u\| > 1$	

ในขณะที่ Ramsay (ค.ศ.1977) ได้พิจารณาฟังก์ชันของ ρ และ ψ สำหรับชุดของ M-esimator ที่อยู่ในรูปของ

$$\rho(\varepsilon_i/S) = a^{-2}[1-\exp(-a|\varepsilon_i|/S) \times (1+a|\varepsilon_i|/S]$$

และ

$$\psi(\varepsilon_i/S) = (\varepsilon_i/S)\exp(-a|\varepsilon_i|/S$$

โดยที่ a = 0.3 และ E_a จะมีขอบเขตและมีค่าลู่เข้าสู่ขอบเขตเมื่อ $|\varepsilon_i|/S = 1/a$ ซึ่งทำให้ค่าผิดปกติที่มีค่ามากถูกตัดออกจากตัวอย่างและทำให้ค่าประมาณสเกลที่คำนวณได้ไม่ได้มาจากทุกค่าของค่าสังเกตุ

ค่าอัตราส่วนผลต่างของค่าเฉลี่ยความผิดพลาดกำลังสอง (ratio of different average mean square error (RDAMSE)) คือผลต่างของวิธีที่ให้ค่าเฉลี่ยความผิดพลาดกำลังสองต่ำที่สุดกับวิธีที่เหลืออีก 4 วิธี หารด้วยค่าเฉลี่ยความผิดพลาดกำลังสองที่มีค่าต่ำสุด

$$\text{RDAMSE} = \frac{\left[\text{AMSE}_{nom} - \text{AMSE}_{min}\right]}{\text{AMSE}_{min}} \times 100$$

โดยที่ AMSE_{min} แทนวิธีที่ให้ค่า AMSE ต่ำที่สุด

และ AMSE_{nom} แทนวิธีที่ให้ค่า AMSE โดยทั่วไปที่ไม่ใช่ค่าต่ำที่สุด

ค่าอัตราส่วนผลต่างของความผิดพลาดกำลังสอง (ratio of different mean square error (RDMSE)) คือผลต่างของวิธีที่ให้ค่าความผิดพลาดกำลังสองต่ำที่สุดกับวิธีที่เหลืออีก 4 วิธี หารด้วยค่าความผิดพลาดกำลังสองที่มีค่าต่ำสุด หลังจากนั้นหาผลรวมของการกระทำซ้ำแล้วหารด้วยจำนวนครั้งของแต่ละวิธีที่ให้ค่าความผิดพลาดกำลังสองต่ำที่สุด

$$\text{RDMSE} = \frac{\sum^{\text{TIMES}} \left[\frac{\text{MSE}_{nom} - \text{MSE}_{min}}{\text{MSE}_{min}}\right]}{\text{TIMES}} \times 100$$

โดยที่ MSE_{min} แทนวิธีที่ให้ค่า MSE ต่ำที่สุด

MSE_{nom} แทนวิธีที่ให้ค่า MSE โดยทั่วไปที่ไม่ใช่ค่าต่ำที่สุด

และ TIMES คือจำนวนครั้งของวิธีที่ให้ค่า MSE ต่ำที่สุด