บทที่ 2

สถิติทดสอบและผลงานวิจัยที่เกี่ยวข้อง

ในกรณีที่ผู้วิจัยต้องการตรวจสอบดูว่าความคลาดเคลื่อนมีความแปรปรวนคงที่หรือไม่ในการวิเคราะห์ความถดถอยนั้น สถิติที่ใช้ในการทดสอบมีอยู่หลายวิธี วิธีที่ใช้ในการวิจัยนี้ก็คือ การทดสอบโกลฟิลด์และควอนท์ (Goldfeld-Quandt test) การทดสอบสโรเตอร์ (Szroeter test) การทดสอบบรูสและพาแกน (Breusch-Pagan test) และการทดสอบ BAMSET (Bartlelt's M Specification Error test) และเนื่องจากตัวสถิติทดสอบของการทดสอบดังกล่าว ประมาณการแจกแจงได้ด้วยการแจกแจงแบบเอฟ (F-distribution) การแจกแจงแบบปกติ (normal distribution) และการแจกแจงแบบไคสแควร์ (Chi-square distribution) ดังนั้นในบทนี้จะกล่าวถึงการแจกแจงเอฟ การแจกแจงปกติ และการแจกแจงไคสแควร์ก่อน แล้วจึงจะกล่าวถึงรายละเอียดของการทดสอบแต่ละวิธี พร้อมทั้งตัวอย่างของการคำนวณ ส่วนในตอนท้ายของบทนี้จะนำเสนอผลงานวิจัยที่เกี่ยวข้องพอเป็นสังเขป ซึ่งรายละเอียดต่าง ๆ เป็นดังนี้

2.1 <u>การแจกแจงที่เกี่ยวข้องในการศึกษา</u>

2.1.1 <u>การแจกแจงเอฟ</u> (F-distribution)

Sir Ronald A. Fisher เป็นบุคคลแรกที่เสนอการแจกแจงเอฟในปี ค.ศ. 1924 เดิมใช้ชื่อว่าการแจกแจงซี (Z-distribution) โดยมีสูตรเป็น $Z = \frac{1}{2} \ln \frac{S_1^2}{S_2^2}$ ซึ่งค่าลอกการิทึม (logarithm) ทำให้ยุ่งยากในการคำนวณ W. Snedecor จึงได้ทำการปรับปรุงใหม่โดยใช้สูตร $F = \frac{S_1^2}{S_2^2}$ ค่า S_1^2 และ S_2^2 คือค่าความแปรปรวนของกลุ่มตัวอย่างจากประชากรที่ 1 และกลุ่มตัวอย่างจากประชากรที่ 2 ตามลำดับ Snedecor ได้ตั้งชื่อการแจกแจงนี้ใหม่ว่าการแจกแจงเอฟ (F-distribution) เพื่อเป็นเกียรติแก่ Ronald A. Fisher แต่ในบางครั้งจะเรียกว่าการแจกแจงเอฟของสเนคดีเคอร์ (Snedecor's F distributionX เพื่อเป็นเกียรติแก่ W. Snedecor

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ฟังก์ชั่นความน่าจะเป็นของการแจกแจง

ตัวแปรสุ่มที่มีการแจกแจงแบบเอฟ จะได้จากอัตราส่วนระหว่างตัวแปรสุ่มไคส-แควร์ (Chi-Square) 2 ตัว ซึ่งเป็นอิสระต่อกัน และต่างหารด้วยจำนวนชั้นของความเป็นอิสระ (degrees of fredom) ดังนั้นฟังก์ชั่นความน่าจะเป็นของการแจกแจงเอฟแสดงได้เป็น

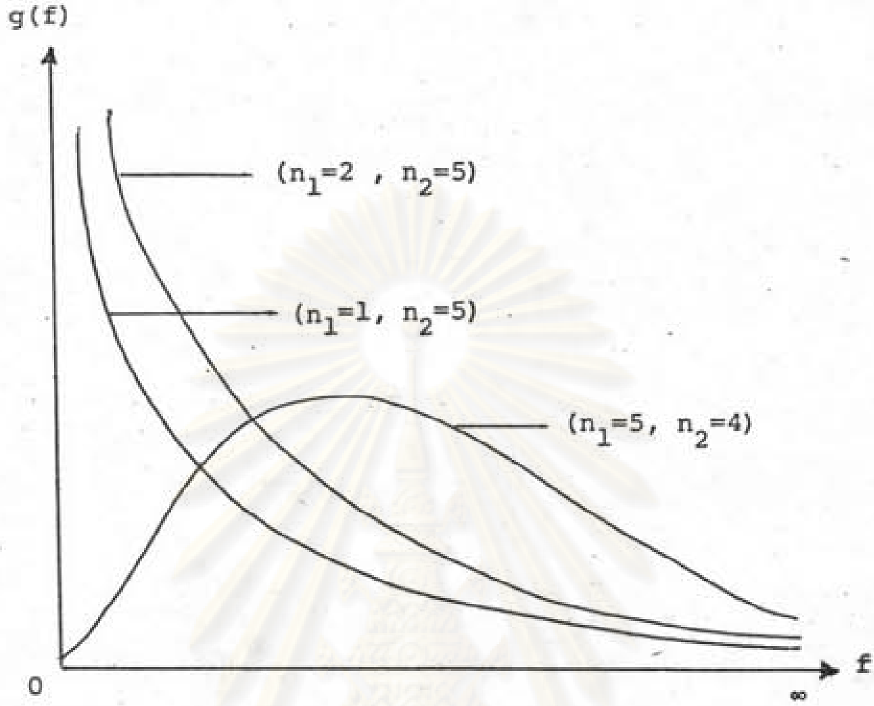
$$g(f) = \frac{\Gamma\left(\frac{n_1+n_2}{2}\right)\left(\frac{n_1}{n_2}\right)^{n_1/2} f^{\frac{n_1}{2}-1}}{\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)\left(1+\frac{n_1}{n_2}f\right)^{\frac{n_1+n_2}{2}}} \quad \text{เมื่อ } 0<f<\infty$$

$$= \quad 0 \quad \text{เมื่อ } f \text{ มีค่าอื่น ๆ}$$

ค่า n_1 และ n_2 เป็นจำนวนชั้นแห่งความเป็นอิสระของตัวแปรสุ่มแบบไคสแควร์ที่เป็นเศษและส่วนตามลำดับ

จากฟังก์ชั่นความน่าจะเป็นของการแจกแจงเอฟ จะเห็นว่าการแจกแจงเอฟมีค่าพารามิเตอร์ 2 ตัวคือ n_1 และ n_2 ซึ่งจะเป็นค่าที่กำหนดรูปแบบ (Shape) ของการแจกแจง ซึ่งลักษณะของการแจกแจงเอฟไม่สมมาตร (Asymmetric) โดยมีลักษณะเบ้ไปทางขวา (Positive Skewness) เช่นเมื่อ $n_1 = 1,2$ ลักษณะการแจกแจงจะเป็นรูปของอักษร J กลับข้าง และถ้า n_1 มากกว่า 3 การแจกแจงจะเบ้ไปทางขวา ซึ่งแสดงได้ดังรูป 2.1 ในกรณีที่ n_1 และ n_2 มีค่าใกล้อนันต์ (infinity) การแจกแจงเอฟ จะมีลักษณะเข้าใกล้ความสมมาตร

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย



<u>รูปที่ 2.1</u> แสดงการแจกแจงเอฟเมื่อ ($n_1 = 1, n_2 = 5$) ($n_1 = 2, n_2 = 5$) และ ($n_1 = 5, n_2 = 4$)

<u>คุณสมบัติของการแจกแจง</u>

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

1. ค่าคาดหวัง (Expected Value) และความแปรปรวนหาได้ดังนี้

$$\text{ค่าคาดหวัง} = \frac{n_2}{(n_2 - 2)} \quad \text{สำหรับ} \quad n_2 > 2$$

$$\text{ความแปรปรวน} = \frac{2\, n_2^2\, (n_1 + n_2 - 2)}{(n_2 - 4)\, n_1\, (n_2 - 2)^2} \quad \text{สำหรับ} \quad n_2 > 4$$

2. การแจกแจงเอฟ จะเบ้ทางขวาและความเบ้จะลดลง ถ้าจำนวนชั้นแห่งความเป็นอิสระเพิ่มขึ้น

3. การแจกแจงเอฟมีคุณสมบัติของส่วนกลับได้คือ ถ้า $F(n_1, n_2)$ มีการแจกแจงเอฟซึ่งมีขั้นแห่งความเป็นอิสระ n_1 และ n_2 แล้ว $F(n_2, n_1) = \frac{1}{F(n_1, n_2)}$ จะมีการแจกแจงเอฟด้วยขั้นแห่งความเป็นอิสระ n_2 และ n_1

4. การแจกแจงทีกับเอฟ มีความสัมพันธ์กันดังนี้คือ ถ้า x มีการแจกแจงทีที่มีขั้นแห่งความเป็นอิสระ n แล้ว x^2 จะมีการแจกแจงเอฟที่มีขั้นแห่งความเป็นอิสระ (1 , n)

5. ค่าเอฟ จะไม่มีทางเป็นค่าลบ จะมีค่าเป็นบวกเสมอ ดังนั้นการแจกแจงเอฟจึงไม่สมมาตรกันรอบจุดศูนย์

ความสำคัญของการแจกแจง

การแจกแจงเอฟ มีความสำคัญในงานวิจัยต่าง ๆ เป็นอันมาก เช่นใช้ในการประมาณช่วง (Interval estimate) และการทดสอบสมมติฐานทางสถิติ ยกตัวอย่างเช่น ใช้ทดสอบสมมติฐานเกี่ยวกับการเท่ากันของความแปรปรวนระหว่างประชากร 2 กลุ่ม การทดสอบความแตกต่างระหว่างมัชฌิมเลขคณิตของประชากรหลายกลุ่ม เป็นต้น

2.1.2 การแจกแจงปกติ (normal distribution)

การแจกแจงความน่าจะเป็นของตัวแปรสุ่มที่สำคัญ และมีการใช้มากที่สุดในวิชาสถิติก็คือ การแจกแจงปกติ ซึ่งเป็นการแจกแจงแบบต่อเนื่องอย่างหนึ่ง ลักษณะที่เห็นเด่นชัดคือ มีลักษณะเป็นรูปโค้งสมมาตรคล้ายระฆัง (Bell-Shaped) ซึ่งเรียกว่า โค้งปกติ (Normal Curve) ลักษณะการแจกแจงแบบนี้มักมีปรากฏอยู่ตามธรรมชาติ เช่น ความสูงของคน คะแนนการทดสอบทางสติปัญญา ฯลฯ ในปี 1733 เดอมัวร์ (De Moivre) ได้สร้างสมการของโค้งปกติ ซึ่งได้กลายมาเป็นรากฐานของทฤษฎีต่าง ๆ ในวิชาสถิติและต่อมา เกาส์ (Gauss) ก็ได้สมการนี้จากการศึกษาเรื่องความผิดพลาดในการวัดปริมาณเดียวกันหลาย ๆ ครั้ง การแจกแจงปกตินี้บางทีจึงถูกเรียกว่า "การแจกแจงแบบเกาส์ (Gaussian distribution) เพื่อเป็นเกียรติแก่ Gauss นักคณิตศาสตร์ผู้เกี่ยวข้องกับการคิดค้นการแจกแจงนี้

ฟังก์ชั่นความน่าจะเป็นของการแจกแจง

ฟังก์ชั่นความน่าจะเป็นของการแจกแจงปกติ มีรูปแบบเป็น

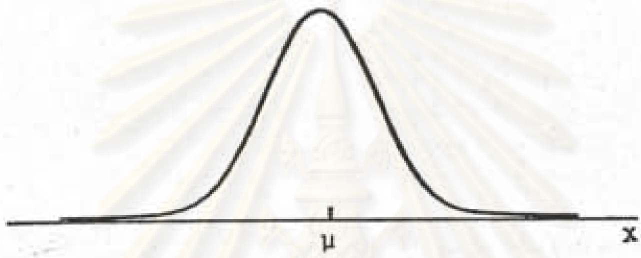
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} \quad ; -\infty < x < \infty$$

016010

โดยที่ $\pi = 3.14159 \ldots$

$e = 2.71828$ และ

μ , σ เป็นพารามิเตอร์ ซึ่ง $-\infty < \mu < \infty$, $\sigma > 0$ ลักษณะของเส้นโค้งของการแจกแจงแบบนี้ มีลักษณะคล้ายระฆังคว่ำดังรูป 2.2

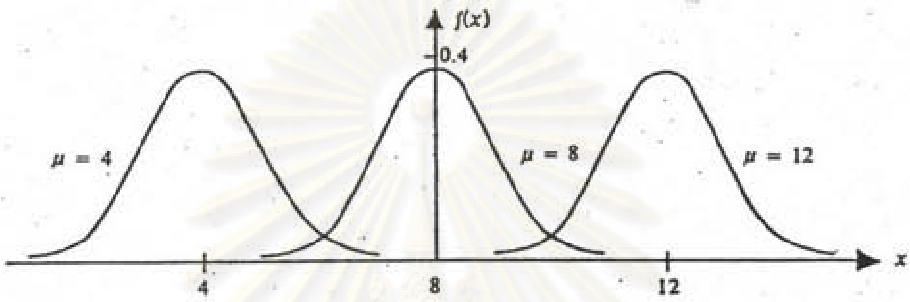


รูปที่ 2.2 แสดงเส้นโค้งการแจกแจงแบบปกติ

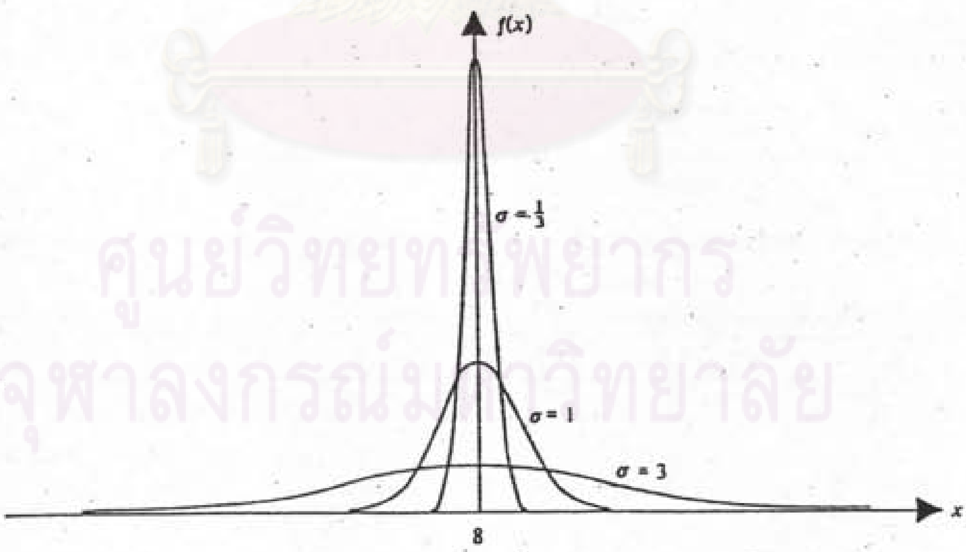
คุณสมบัติของการแจกแจง

1. โค้งปกติมีฐานนิยมอยู่ที่ $x = \mu$ (โค้งสูงที่สุด)
2. โค้งมีสมมาตรกับแกนตั้งที่ลากผ่าน μ
3. โค้งปกติมีจุดเปลี่ยนเว้าที่ $x = \mu \pm \sigma$
4. ปลายโค้งเข้าใกล้แกน x เมื่อ x มีค่าห่างจาก μ ออกไปทุกที
5. พื้นที่ทั้งหมดที่อยู่ใต้เส้นโค้ง และอยู่เหนือแกน x มีค่าเป็น 1
6. ค่าเฉลี่ยของ x คือ μ และความแปรปรวนของ x คือ σ^2

พารามิเตอร์ μ และ σ จะเป็นตัวกำหนดตำแหน่งที่ตั้งของเส้นโค้งปกติและความโค้งหรือแบบของเส้นโค้งปกติ โค้งจะอยู่ทางซ้ายหรือขวาของแกน $y = f(x)$ ขึ้นอยู่กับค่าของ μ ลักษณะของโค้งปกติจะดูโด่งหรือแบนราบขึ้นอยู่กับค่าของ σ ถ้า σ มีค่ามากโค้งจะมีลักษณะแบบราบ ถ้า σ มีค่าน้อยโค้งจะมีลักษณะโด่ง ดังรูปที่ 2.3 และ 2.4 ซึ่งแสดงการเปรียบเทียบเมื่อ μ และ σ มีค่าต่าง ๆ กัน



รูปที่ 2.3 แสดงกราฟฟังก์ชั่นความน่าจะเป็นของตัวแปรสุ่มแบบปกติที่มี μ ต่างกันแต่ค่า σ เท่ากัน



รูปที่ 2.4 แสดงกราฟฟังก์ชั่นความน่าจะเป็นของตัวแปรสุ่มปกติที่มี μ = 8 และ σ ต่างกัน

ความสำคัญของการแจกแจง

เหตุผลที่การแจกแจงปกติ เป็นการแจกแจงที่สำคัญที่สุดเนื่องจาก

1. มีตัวแปรสุ่มหลายอย่างจากการสังเกต หรือจากการทดลองมีการแจกแจงปกติ หรือใกล้เป็นการแจกแจงปกติ (approximatly normal distribution)

2. บางทีตัวแปรสุ่มบางอย่าง อาจมีการแจกแจงไม่เป็นการแจกแจงปกติ และไม่ใกล้กับการแจกแจงปกติ แต่สามารถที่จะแปลงให้เป็นตัวแปรสุ่มใหม่ที่มีการแจกแจงใกล้เป็นการแจกแจงปกติได้โดยสมการง่าย ๆ

3. มีการแจกแจงที่สำคัญหลายอย่างที่หาค่าได้ยาก ก็สามารถประมาณค่าโดยการแจกแจงแบบปกติได้ เช่นการแจกแจงทวินาม

4. มีตัวสถิติหลายตัว ที่ใช้รากฐานของการแจกแจงของประชากรเป็นแบบปกติ

การแจกแจงแบบปกติในรูปมาตรฐาน (Standard Normal Distribution or The Unit Normal)

ถ้าการแจกแจงปกติมี $\mu = 0$ และ $\sigma = 1$ เราจะเรียกการแจกแจงนี้ว่าการแจกแจงปกติมาตรฐาน (Standard normal distribution) ใช้สัญลักษณ์ N(0 , 1) ประโยชน์ของการแจกแจงปกติมาตรฐานจะใช้ในการเปิดตาราง เพื่อหาพื้นที่ เพราะตารางสถิติทั่วไปจะแสดงค่าพื้นที่ของเส้นโค้งปกติมาตรฐานเท่านั้น วิธีการเปลี่ยนการแจกแจงปกติทั่วไปให้เป็นการแจกแจงปกติมาตรฐาน สามารถกระทำได้โดยใช้สูตรแปลงค่าดังนี้

$$Z = \frac{x - \mu}{\sigma}$$

และฟังก์ชั่นความน่าจะเป็นของการแจกแจงแบบปกติมาตรฐานคือ

$$f(Z) = \frac{1}{\sqrt{2\pi}} \cdot e^{-Z^2/2} \quad ; \quad -\infty < Z < \infty$$

และเนื่องจากการแจกแจงแบบนี้ เป็นกรณีเฉพาะของการแจกแจงปกติ ดังนั้นคุณสมบัติต่าง ๆ ก็เป็นเช่นเดียวกัน

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

2.1.3 การแจกแจงไคส์แควร์ (Chi-square distribution)

ตัวแปรสุ่มไคส์แควร์ คือตัวแปรสุ่มที่พัฒนาขึ้น มาจากฟังก์ชั่นของตัวแปรสุ่มปกติ กล่าวคือ ถ้า $x_1, x_2, \ldots, x_n$ เป็นตัวแปรสุ่มที่มีการแจกแจงปกติโดยมี $E(x_i) = \mu$ และ $V(x_i) = \sigma^2$; $i = 1,2,\ldots, n$ ถ้าสุ่มตัวอย่างขึ้นมา 1 ตัว แล้วแปลงเป็น การแจกแจงปกติมาตรฐาน Z ดังนี้

$$Z = \frac{x - \mu}{\sigma} \sim N(0, 1)$$

ถ้ายกกำลังสอง $Z^2 = (\frac{x - \mu}{\sigma})^2 = W, W \sim \chi^2_{(1)}$

จะได้ว่า Z^2 จะมีการแจกแจงแบบไคส์แควร์ ที่มีขั้นแห่งความเป็นอิสระเท่ากับ 1 ในทำนองเดียวกัน ถ้าสุ่มกลุ่มตัวอย่างขนาด n ซึ่งเป็นอิสระต่อกันจะได้ว่าผลรวมของ Z^2 จะมีการแจกแจงแบบไคส์แควร์ ที่มีขั้นแห่งความเป็นอิสระเท่ากับ n ดังนี้ $W = Z_1^2 + Z_2^2 + Z_3^2 + \ldots + Z_n^2$; $W \sim \chi^2_{(n)}$

ฟังก์ชั่นความน่าจะเป็นของการแจกแจง

ฟังก์ชั่นความน่าจะเป็นของการแจกแจงไคส์แควร์ มีรูปแบบเป็น

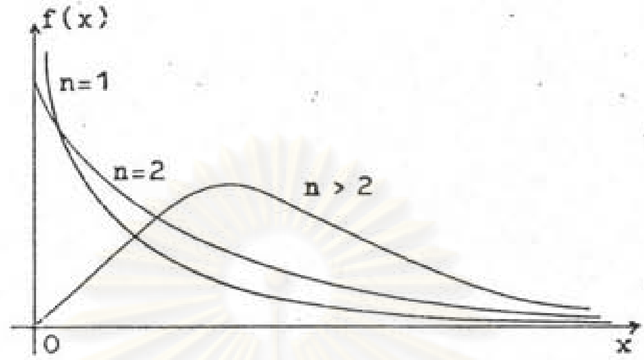
$$f(x) = \frac{1}{2^{n/2}\,(n/2 - 1)!} \cdot x^{n/2 - 1}\, e^{-x/2} \quad ; x > 0$$

n เป็นจำนวนขั้นแห่งความเป็นอิสระ

e เป็นจำนวนอตักยะ (irrational) ที่มีค่าเท่ากับ 2.71828...

จากฟังก์ชั่นความน่าจะเป็นของการแจกแจงไคส์แควร์ จะเห็นว่าการแจกแจงไคส์แควร์ขึ้นอยู่กับจำนวนขั้นแห่งความเป็นอิสระ การแจกแจงไคส์แควร์จะเบ้ขวา เมื่อจำนวนขั้นแห่งความเป็นอิสระมีค่าน้อย และความเบ้จะลดลงไปเมื่อจำนวนขั้นแห่งความเป็นอิสระมีค่าเพิ่มขึ้นและถ้าจำนวนขั้นแห่งความเป็นอิสระมีค่ามาก (ประมาณ 20) การแจกแจงค่อนข้างจะสมมาตร รูปของการแจกแจงไคส์แควร์สำหรับ n หลาย ๆ ค่าที่แตกต่างกันจะมีลักษณะดังรูปที่

2.5 จากการสังเกตจากรูปการแจกแจงไคสแควร์ สำหรับ n = 1 และ n = 2 จะพบว่าแตกต่างจากรูปทั่วไปของการแจกแจงสำหรับ n > 2



รูปที่ 2.5 แสดงลักษณะกราฟของการแจกแจงไคสแควร์ เมื่อจำนวนชั้นแห่งความเป็นอิสระแตกต่างกันเป็น n = 1, n = 2 และ n > 2

คุณสมบัติของการแจกแจง

1. การแจกแจงไคสแควร์ เป็นการแจกแจงแบบต่อเนื่อง
2. ค่าไคสแควร์มีค่าระหว่าง 0 ถึงค่าอนันต์ (infinity) นั่นคือค่าไคสแควร์เป็นค่าบวกเสมอ
3. การแจกแจงไคสแควร์มีค่าเฉลี่ยเท่ากับ n ค่าความแปรปรวนเท่ากับ 2 n
4. ลักษณะของส่วนโค้งจะเบ้ไปทางขวามือ
5. เป็นส่วนโค้งแบบเพียร์สันชนิดที่สาม (Pearson Type III Curve)

ความสำคัญของการแจกแจง

การแจกแจงไคสแควร์ เป็นการแจกแจงชนิดต่อเนื่อง ใช้ได้กับการทดสอบที่ใช้พารามิเตอร์ (Parametric test) และที่ไม่ใช้พารามิเตอร์ (nonparametric test) เช่น ใช้ในการหาช่วงความเชื่อมั่น (Confidence interval) ของความแปรปรวนของประชากร การทดสอบภาวะสารูปสนิทดี (test of goodness of fit) การทดสอบความเป็นอิสระ (test of independence) เป็นต้น

2.2 การทดลอบที่ใช้ในการศึกษา

2.2.1 การทดลอบโกลฟิลด์และควอนท์ (Goldfeld-Quandt test)

โกลฟิลด์และควอนท์ (Goldfeld and Quandt 1965:539-547) ได้เล่นอวิธีการทดลอบปัญหาความคลาดเคลื่อนมีความแปรปรวนไม่คงที่ ในการวิเคราะห์ความถดถอยอื่น โดยมีขั้นตอนการดำเนินงานทดลอบดังนี้

1. จากค่าลังเกต (X_t, Y_t) ; $t = 1,2,\ldots, T$ ให้จัดเรียงลำดับค่าลังเกตเหล่านี้ (จากน้อยไปมาก หรือมากไปน้อยก็ได้) ตามการเพิ่มขึ้นของความแปรปรวน ซึ่งในการวิจัยครั้งนี้จะทำการเรียงลำดับค่าลังเกตตามค่าของ X_t จากน้อยไปมาก เนื่องจากความแปรปรวนเป็นลัดส่วนโดยตรงกับตัวแปรอิสระ X_t เมื่อ X_t มีค่าเพิ่มขึ้น ขนาดของความแปรปรวนก็มีค่าเพิ่มขึ้นด้วย การเรียงลำดับค่าลังเกตตามค่าของ X_t จากน้อยไปมาก จึงเป็นการเรียงลำดับตามการเพิ่มขึ้นของความแปรปรวน

2. จำแนกค่าลังเกตที่จัดเรียงลำดับแล้วในขั้นตอนที่ 1 ออกเป็น 2 ล่วนที่เป็นอิสระต่อกัน การจำแนกให้กระทำโดยตัดค่าลังเกตที่อยู่กลางลำดับ (Sequence) ทิ้งไป c หน่วย โดยหลักปฏิบัตินิยมตัดทิ้ง ประมาณ 1 ใน 4 ของค่าลังเกตทั้งหมด นั่นคือ $c \approx \frac{T}{4}$ ภายหลังเมื่อตัดค่าลังเกตกลางลำดับ (central Observation) ทิ้งแล้วก็จะทำให้เหลือกลุ่มตัวอย่างอิสระ 2 กลุ่ม ๆ ละ $\frac{T-c}{2}$ หน่วยเท่า ๆ กัน คือกลุ่มที่ล่อดคล้องกับค่าลังเกตที่ความแปรปรวนมีค่าน้อย กับกลุ่มที่ล่อดคล้องกับค่าลังเกตที่ความแปรปรวนมีค่ามาก

3. จากกลุ่มตัวอย่างย่อยในขั้นตอนที่ 2 ให้แยกวิเคราะห์ล่มการถดถอยโดยวิธี OLS โดยถือเล่มือนว่าเป็นแบบจำลองคนละชุด แล้วทำการคำนวณหาผลบวกของค่าคลาดเคลื่อนกำลังล่อง หรือ SS (Residual) จาก

$$\sum_{t}^{T} \hat{u}_t^2 = \sum_{t}^{T} (y_t - \hat{\beta}_o - \hat{\beta}_1 X_t)^2 = \sum_{t}^{T} (y_t - \hat{y}_t)^2$$

ก็จะได้ S_1 และ S_2 เมื่อ S_1 = SS (Residual) ของกลุ่มค่าลังเกตที่ความแปรปรวนมีค่าน้อย

S_2 = SS (Residual) ของกลุ่มค่าลังเกตที่ความแปรปรวนมีค่ามาก

4. คำนวณค่าสถิติทดสอบ

$$R = \frac{S_2}{S_1}$$

ภายใต้สมมติฐานหลัก (null hypothesis) ตัวสถิติ R จะมีการแจกแจงแบบเอฟ (F-distribution) ซึ่งมีจำนวนชั้นแห่งความเป็นอิสระ $(T-c-2p)/2$ และ $(t-c-2p)/2$ ซึ่งการวิจัยครั้งนี้กำหนดให้ $c = 4$ และ 10 เมื่อขนาดตัวอย่างเป็น 20 และ 50 ตามลำดับ และ $p = 2$ (จำนวนพารามิเตอร์)

5. <u>เกณฑ์การตัดสินใจ</u> จะตัดสินใจปฏิเสธสมมติฐานหลัก $H_o : \sigma_1^2 = \sigma_2^2$ เทียบกับ $H_1 : \sigma_1^2 \neq \sigma_2^2$ ณ ระดับนัยสำคัญ α เมื่อ

$$R > F_\alpha \left[(T-c-4)/2 \;,\; (T-c-4)/2\right]$$

F_α เป็นค่าที่ได้จากตารางการแจกแจงเอฟ ซึ่งมีจำนวนชั้นแห่งความเป็นอิสระคือ $(T-c-4)/2$ และ $(T-c-4)/2$ ซึ่งจะยืนยันว่าภายใต้ความเป็นจริงจากข้อมูลที่มีอยู่ เราสามารถเชื่อได้ถึง $(1 - \alpha)$ 100% ว่าขณะนี้ได้เกิดปัญหาความคลาดเคลื่อนมีความแปรปรวนไม่คงที่ขึ้นแล้ว

<u>ตัวอย่างวิธีการคำนวณหาค่าสถิติทดสอบ</u>

<u>ตัวอย่างที่ 1</u> ข้อมูลในตารางที่ 2.1 นี้เป็นข้อมูลที่สร้างขึ้นจากการทดลองโดยความแปรปรวนมีรูปแบบของการบวกเป็น $\sigma_t^2 = (1 + 0.05X_t)^2$ และขนาดตัวอย่างเท่ากับ 20 แสดงได้ดังนี้

ตารางที่ 2.1 แสดงรายละเอียดข้อมูลที่สร้างขึ้นจากการทดลองโดยความแปรปรวนมีรูปแบบของการบวก และขนาดตัวอย่างเท่ากับ 20

X_t	Y_t
69.985	76.266
77.220	83.014
84.683	97.949
71.383	79.812
89.142	102.434
73.610	85.455
79.177	86.658
81.863	91.345
83.124	92.983
86.108	92.314
88.183	98.121
95.992	109.823
89.353	93.282
91.279	105.073
94.458	106.679
97.835	114.258
98.343	112.070
99.832	106.052
102.375	104.390
106.661	111.357

จากข้อมูลในตารางที่ 2.1 เราสามารถนำมาคำนวณหาค่าสถิติทดสอบโกลด์ฟิลด์และควอนท์ได้ดังนี้

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

<u>วิธีทำ</u> จัดเรียงลำดับค่าสังเกตทั้งหมดตามค่าของตัวแปรอิสระ X_t จากน้อยไปมาก แล้วจำแนกค่าสังเกตที่จัดเรียงลำดับแล้วนี้ออกเป็น 2 ส่วน โดยการตัดค่าสังเกตที่อยู่กลางลำดับทิ้งไป 4 ค่า และจากกลุ่มตัวอย่างย่อยนี้ จะทำการแยกวิเคราะห์สมการถดถอยโดยวิธี OLS แล้วคำนวณหา SS (Residual) แสดงได้ดังนี้

	X_t	Y_t	$\hat{Y}_t$	$\hat{u}_t$	$\hat{u}_t^2$
	69.985	76.216	77.297	-1.081	1.1685
	71.383	79.812	79.014	0.798	0.6368
	73.610	85.455	81.749	3.706	13.7344
สมการถดถอยสำหรับ	77.220	83.014	86.182	-3.168	10.0362
กลุ่มตัวอย่างที่ 1	79.177	86.658	88.585	-1.927	3.7133
$\hat{Y}_t = -8.644+1.228X_t$	81.863	91.345	91.583	-0.538	0.2894
	83.124	92.983	93.432	-0.449	0.2016
	84.683	97.949	95.346	2.603	6.775
				S_1 =	36.555
	91.279	105.073	111.191	-6.118	37.429
สมการถดถอยสำหรับ	94.458	106.679	111.768	-5.071	25.715
กลุ่มตัวอย่างที่ 2	95.992	109.823	112.047	-2.224	4.9416
$\hat{Y}_t = 94.615+0.181X_t$	97.835	114.258	112.381	1.877	3.5231
	98.343	112.070	112.070	-0.404	0.1632
	99.832	106.052	112.744	-6.692	44.7828
	102.375	104.390	113.206	-8.816	77.7218
	106.661	111.357	113.984	-2.627	6.9011
				S_2 =	201.182

ตัวสถิติทดสอบ $$R = \frac{S_2}{S_1} = \frac{201.182}{36.555} = 5.50354$$

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ซึ่ง $R = 5.50354 > F_{0.05}(6, 6) = 4.280$

ดังนั้น จึงปฏิเสธสมมติฐานหลัก ที่ระดับนัยสำคัญ 0.05

2.2.2 การทดสอบสโรเตอร์ (Szroeter test)

Szroeter (1978:1311-1327) เป็นผู้เสนอการทดสอบนี้ขึ้นโดยสมมติว่า ความคลาดเคลื่อนมีการแจกแจงแบบปกติ และค่าสังเกตสามารถถูกนำมาจัดลำดับในลักษณะ $\sigma^2_{t-1} \leq \sigma^2_t$; $t = 2,3,\ldots, T$ นั่นคือค่าสังเกตสามารถนำมาจัดลำดับตามการเพิ่มขึ้นของความแปรปรวนได้นั่นเอง ซึ่งการวิจัยนี้ดำเนินการได้โดย ทำการจัดลำดับค่าสังเกตตามค่าของตัวแปรอิสระ X_t จากน้อยไปมาก เหมือนกับ การทดสอบโกลด์ฟิลด์และควอนท์ จากนั้นทำการวิเคราะห์สมการถดถอยโดยวิธี OLS แล้วทำการคำนวณหาเวกเตอร์ของความคลาดเคลื่อน (Residual Vector) $\hat{\underset{\sim}{u}} = (\hat{u}_1, \hat{u}_2, \ldots, \hat{u}_T)'$ เมื่อ
$\hat{u}_t = y_t - \hat{\beta}_o - \hat{\beta}_1 X_t$; $t = 1,2,\ldots, T$

<u>สถิติทดสอบ</u>

$$Q = \frac{T(\tilde{h} - \bar{h})}{\left[2 \sum_{t=1}^{T} (h_t - \bar{h})^2\right]^{\frac{1}{2}}}$$

โดยที่ $$\tilde{h} = \sum_{t}^{T} W_t h_t$$

$$W_t = \hat{u}_t^2 \Big/ \sum_{t}^{T} \hat{u}_t^2$$

$$\bar{h} = T^{-1} \sum_{t}^{T} h_t$$

เมื่อ $\hat{u}_t$ คือ OLS-Residual ของค่าสังเกตที่ t และ

$$h_t = t$$

ภายใต้สมมติฐานหลัก ตัวสถิติ Q หรือตัวสถิติของสโรเตอร์จะมีการแจกแจงเข้าใกล้การแจกแจงปกติที่มีค่าเฉลี่ยเป็น 0 และความแปรปรวนเป็น 1 (Standard normal distribution) กล่าวคือ $Q \sim N(0, 1)$

เกณฑ์การตัดสินใจ

จะตัดสินใจปฏิเสธสมมติฐานหลัก $H_o : \sigma_t^2 = \sigma_{t-1}^2 ; t = 2,3,...,T$ เทียบกับ $H_1 : \sigma_t^2 > \sigma_{t-1}^2$ ที่ระดับนัยสำคัญ α เมื่อ $Q > \Psi_\alpha$ โดยที่ Ψ_α คือจุด 100% ด้านบนของการแจกแจงปกติมาตรฐาน ซึ่งจะยืนยันว่าภายใต้ความเป็นจริงของข้อมูลที่มีอยู่ เราสามารถเชื่อได้ถึง (1- α) 100% ว่าขณะนี้ได้เกิดปัญหาความคลาดเคลื่อนมีความแปรปรวนไม่คงที่ขึ้นแล้ว

ตัวอย่างวิธีการคำนวณหาค่าสถิติทดสอบ

ตัวอย่างที่ 2 จากข้อมูลในตารางที่ 2.1 ของตัวอย่างที่ 1 สามารถนำมาทำการตรวจสอบปัญหาความคลาดเคลื่อนมีความแปรปรวนไม่คงที่ โดยการทดสอบสโรเตอร์ได้ดังนี้

วิธีทำ ทำการจัดลำดับค่าสังเกตทั้งหมดตามค่าของ X_t จากน้อยไปมากเหมือนกับการทดสอบโกลฟิลด์และควอนท์ แล้วทำการวิเคราะห์สมการถดถอย โดยวิธี OLS ได้สมการถดถอยเป็น

$$\hat{Y}_t = 8.198 + 1.014X_t \quad ; \quad t = 1,2,..., 20$$

จากสมการถดถอยที่ได้ นำมาคำนวณหาเวกเตอร์ของความคลาดเคลื่อน (Residual Vector) $\hat{\underset{\sim}{u}} = (\hat{u}_1 , \hat{u}_2,..., \hat{u}_{20})'$ และ $w_t = \hat{u}_t^2 / \sum_t^T \hat{u}_t^2$ ได้ดังนี้

X_t	Y_t	$\hat{Y}_t$	$\hat{u}_t$	$\hat{u}_t^2$	w_t
69.985	76.266	79.163	-2.897	8.3926	0.0272
71.383	79.812	81.580	-1.768	3.1258	0.0101
73.610	85.455	82.839	2.616	6.8434	0.0222
77.220	83.014	86.499	-3.485	12.1452	0.0394
79.177	86.658	88.483	-1.825	3.3306	0.0108
81.863	91.345	91.207	0.138	0.0190	0.00006
83.124	92.983	92.485	0.4985	0.2480	0.0008
84.683	97.949	94.067	3.882	15.0690	0.0489
86.108	92.314	95.511	-3.197	10.2200	0.0332
88.183	98.121	97.615	0.506	0.2560	0.0008
89.142	102.434	98.587	3.847	14.7994	0.0480
89.353	93.282	98.801	-5.519	30.4593	0.0989
91.279	105.073	100.755	4.318	18.6450	0.0605
94.458	106.679	103.978	2.701	7.2950	0.0236
95.992	109.823	105.533	4.290	18.4040	0.0597
97.835	114.258	107.403	6.855	46.9910	0.1526
98.343	112.070	107.917	4.153	17.2471	0.0560
99.832	106.052	109.427	-3.375	11.3906	0.0370
102.375	104.390	112.006	-7.616	58.0030	0.1884
106.661	111.357	116.351	-4.994	24.9400	0.0810
				307.822	

คำนวณหาค่าของ $\tilde{h} = \sum_{t}^{T} w_t h_t \;;\; h_t = t \quad t = 1,2,\ldots, 20$

$\therefore \tilde{h} = (0.0272)(1) + (0.0101)(2) + \ldots + (0.0810)(20)$

$= 14.021391$

$$\text{และ} \quad \bar{h} = T^{-1} \sum_{t}^{T} h_t = \frac{1 + 2 + 3 + \ldots + 20}{20}$$

$$= 10.5$$

$$\sum_{t}^{T} (h_t - \bar{h})^2 = (1 - 10.5)^2 + (2 - 10.5)^2 + \ldots + (20 - 10.5)^2$$

$$= 665$$

$$Q = \frac{T(\tilde{h} - \bar{h})}{[2 \sum_{t}^{T} (h_t - \bar{h})^2]^{\frac{1}{2}}}$$

$$Q = \frac{20(14.021 - 10.50)}{\sqrt{2(665)}} = \frac{70.42}{36.469}$$

$$Q = 1.9309$$

$$Q = 1.9309 > \Psi_{0.05} = 1.645$$

ดังนั้น จึงปฏิเสธสมมติฐานหลัก ที่ระดับนัยสำคัญ 0.05

2.2.3 การทดสอบบรูสและพาแกน (Breusch-Pagan test)

การทดสอบนี้มีพื้นฐานมาจากการทดสอบ LM (Lagrangian Multiplier test) ของ Silvey (1959 :389-407) และได้ถูกพัฒนาขึ้นโดย Godfrey (1978:227-336) และ Breusch และ Pagan (1979:1287-1294) เป็นการทดสอบที่ใช้ในการตรวจสอบปัญหาความคลาดเคลื่อนมีความแปรปรวนไม่คงที่ 3 รูปแบบในลักษณะที่ $V(u_t)$ เป็นฟังก์ชันเชิงเส้น (Linear function) ของตัวแปรอิสระ และลักษณะที่ $\ell n \ V(u_t)$ เป็นฟังก์ชันเชิงเส้นของตัวแปรอิสระดังนี้

$$1. \quad E(u_t^2) = \sigma_t^2 = \alpha_1 + \alpha_2 Z_{2t} + \ldots + \alpha_s Z_{st} \; ; \; t = 1,2,\ldots, T$$

$$\text{หรือเขียนได้เป็น} \quad \sigma_t^2 = (\underset{\sim}{Z}_t' \underset{\sim}{\alpha})$$

2. $E(u_t^2) = \sigma_t^2 = (\alpha_1 + \alpha_2 Z_{2t} + \ldots + \alpha_s Z_{st})^2 \; ; \; t = 1,2,\ldots,T$

หรือเขียนได้เป็น $\sigma_t^2 = (\underset{\sim}{Z}_t' \underset{\sim}{\alpha})^2$

3. $\ell n \, V(u_t) = \ell n \, \sigma_t^2 = \underset{\sim}{Z}_t' \underset{\sim}{\alpha}$ หรือ $\sigma_t^2 = \exp(\underset{\sim}{Z}_t' \underset{\sim}{\alpha})$

เมื่อ $\underset{\sim}{Z}_t = (1, Z_{2t}, Z_{3t}, \ldots, Z_{st})'$ โดยที่ $Z_j \; ; \; j = 2,3,\ldots,S$ คือตัวแปรอิสระ X_j หรือฟังก์ชั่นของตัวแปรอิสระตัวใดตัวหนึ่งเพียงตัวเดียวเท่านั้น และ $\underset{\sim}{\alpha}$ เป็นเวกเตอร์ของพารามิเตอร์ที่ไม่ทราบค่า ขนาด $S \times 1$ คือ

$$\underset{\sim}{\alpha} = (\alpha_1 , \alpha_2, \ldots, \alpha_s)'$$

ดังนั้นวิธีการทดสอบนี้ จึงถือรูปแบบของความแปรปรวนเป็น $\sigma_t^2 = h\,(\underset{\sim}{Z}_t' \underset{\sim}{\alpha})$ หรือ σ_t^2 เป็นฟังก์ชั่นของ $\underset{\sim}{Z}_t' \underset{\sim}{\alpha}$ ทั้งนี้ h จะเป็นฟังก์ชั่นที่ไม่เกี่ยวข้องกับ t แต่ประการใด ซึ่งแสดงว่า $\sigma_t^2 = h\,(\underset{\sim}{Z}_t' \underset{\sim}{\alpha})$ สามารถยืดหยุ่นให้รับกับรูปแบบของความแปรปรวน ถึง 3 รูปแบบข้างต้น ได้ตามความต้องการของนักวิจัย สมมติฐานที่ต้องการจะทำการทดสอบก็คือ

$$H_o = \alpha_2 = \alpha_3 = \alpha_4 = \ldots = \alpha_s = 0$$

H_1 : มี α อย่างน้อย 1 ตัวที่ $\neq 0$

ซึ่งภายใต้สมมติฐานหลัก (H_o) จะได้ว่า $\underset{\sim}{Z}_t' \underset{\sim}{\alpha} = \alpha_1$ ซึ่งทำให้ $\sigma_t^2 = h\,(\underset{\sim}{Z}_t' \underset{\sim}{\alpha}) = h\,(\alpha_1) = \sigma^2$ นั่นคือความแปรปรวนมีค่าคงที่นั่นเอง

จะเห็นว่าความแปรปรวนในรูปแบบที่ 2 และ 3 ดังกล่าวข้างต้นก็คือรูปแบบของความแปรปรวน ที่ใช้ในการวิจัยครั้งนี้นั่นเอง กล่าวคือ

ความแปรปรวนรูปแบบของการคูณ (Multiplicative model) ที่มีรูปแบบเป็น $V(u_t) = \sigma_t^2 = kX_t^2$ ก็คือรูปแบบเฉพาะรูปแบบหนึ่งของ $\sigma_t^2 = \exp(\underset{\sim}{Z}_t' \underset{\sim}{\alpha})$ เมื่อ $\underset{\sim}{Z}_t = (1, \ell nX_t)'$ และ $\underset{\sim}{\alpha} = (\ell nk , r)'$ ส่วนความแปรปรวนรูปแบบของการบวก

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

(additive model) ที่มีรูปแบบเป็น $V(u_t) = \sigma_t^2 = k^2\ (1+\lambda X_t)^2$ ก็คือรูปแบบเฉพาะรูปแบบหนึ่งของ $\sigma_t^2 = (\underset{\sim}{z}_t' \underset{\sim}{\alpha})^2$ เมื่อ $\underset{\sim}{z}_t = (1,\ X_t)'$ และ $\underset{\sim}{\alpha} = (1,\lambda)'$, $k = 1$

การทดสอบนี้จึงสามารถที่จะนำมาใช้ได้ในการวิจัยครั้งนี้ โดยมีขั้นตอนดำเนินงานทดสอบดังนี้

1. จากสมการ $Y_t = \beta_o + \beta_1 X_t + u_t \quad ; t = 1,2,\ldots, T$ ให้ประมาณค่าพารามิเตอร์ β_o, β_1 โดยวิธี OLS เพื่อสร้างสมการถดถอย

2. คำนวณหาค่าของ $\hat{u}_t$ เมื่อ $\hat{u}_t$ คือ OLS-Residual ของค่าสังเกตที่ t ที่ได้จากสมการในขั้นตอนที่ 1

3. จากเวกเตอร์ของความคลาดเคลื่อน (Residual Vector) $\hat{\underset{\sim}{u}} = (\hat{u}_1\ ,\ \hat{u}_2, \ldots,\ \hat{u}_T)'$ ที่ได้จากขั้นตอนที่ 2 ให้คำนวณหาตัวประมาณค่าความแปรปรวนของความคลาดเคลื่อน $\hat{\sigma}^2$ จาก $\hat{\sigma}^2 = T^{-1} \sum_t^T \hat{u}_t^2$

4. คำนวณค่าเวกเตอร์ $\underset{\sim}{g}$ โดยสมาชิกในเวกเตอร์ $\underset{\sim}{g}$ คำนวณได้จาก

$$g_t = \frac{\hat{u}_t^2}{\hat{\sigma}^2} - 1 \qquad ; t = 1,2,\ldots, T$$

5. <u>สถิติทดสอบ</u>

$$LM = \frac{1}{2}\ \underset{\sim}{g}' Z\ (Z' Z)^{-1}\ Z' \underset{\sim}{g}$$

เมื่อ Z เป็นเมตริกซ์ขนาด (T x 2) ซึ่งมีสมาชิกในแถวที่ t ใช้สัญลักษณ์เป็น $Z_t = (1,\ \ell n\ X_t)$ สำหรับความแปรปรวนรูปแบบของการคูณ และ

$Z_t = (1,\ X_t)$ สำหรับความแปรปรวนรูปแบบของการบวก

ภายใต้สมมติฐานหลักตัวสถิติ LM หรือตัวสถิติของ Breusch และ Pagan จะมีการแจกแจงเข้าใกล้การแจกแจงแบบไคสแควร์ มีจำนวนชั้นของความเป็นอิสระเป็น 1 (เมื่อ Z_t มีขนาดเป็น p ตัวสถิติ $LM \sim \chi^2_{(p-1)}$)

6. <u>เกณฑ์การตัดสินใจ</u> จะตัดสินใจปฏิเสธสมมติฐาน $H_o : \alpha_2 = 0$ เทียบกับ $H_1 : \alpha_2 \neq 0$ ณ ระดับนัยสำคัญ α เมื่อ

$$LM > \chi^2_{(1),\ \alpha}$$

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ซึ่งจะยืนยันว่าภายใต้ความเป็นจริง จากข้อมูลที่มีอยู่เราสามารถเชื่อได้ถึง $(1 - \alpha)$ 100% ว่าขณะนี้ได้เกิดปัญหาความคลาดเคลื่อนมีความแปรปรวนไม่คงที่ขึ้นแล้ว

ตัวอย่างวิธีการคำนวณหาค่าสถิติทดสอบ

ตัวอย่างที่ 3 จากข้อมูลในตารางที่ 2.1 ของตัวอย่างที่ 1 สามารถนำมาทำการตรวจสอบปัญหาความคลาดเคลื่อนมีความแปรปรวนไม่คงที่ โดยการทดสอบบรูสและพาแกนได้ดังนี้

วิธีทำ ทำการวิเคราะห์สมการถดถอยโดยวิธี OLS เพื่อใช้ในการหาค่าของเวกเตอร์ของความคลาดเคลื่อน $\hat{\underset{\sim}{u}} = (\hat{u}_1, \hat{u}_2, \ldots, \hat{u}_{20})'$

$\hat{\sigma}^2 = T^{-1} \sum_{t}^{T} \hat{u}_t^2$ และเวกเตอร์ $\underset{\sim}{g}$ ซึ่งมีสมาชิกเป็น $g_t = \frac{\hat{u}_t^2}{\hat{\sigma}^2} - 1 \quad ; t = 1,2,\ldots,20$

แสดงได้ดังนี้

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

X_t	Y_t	Y_t	u_t	u_t^2	$g_t = \frac{\hat{u}_t^2}{15.391} - 1$
69.985	76.266	79.163	-2.897	8.3926	-0.4540
71.383	79.812	81.580	-1.768	3.1258	-0.7969
73.610	85.455	82.839	2.616	6.8454	-0.5553
77.220	83.014	86.499	-3.485	12.1452	-0.2100
79.177	86.658	88.483	-1.825	3.3306	-0.7830
81.863	91.345	91.207	0.138	0.0190	-0.9980
83.124	92.983	92.485	0.498	0.2480	-0.9830
84.683	97.949	94.067	3.882	15.0690	-0.0209
86.108	92.314	95.511	-3.197	10.2200	-0.3359
88.183	98.121	97.615	0.502	0.2560	-0.4833
89.142	102.434	98.587	3.847	14.7994	-0.0384
89.353	93.282	98.801	-5.519	30.4593	0.9790
91.279	105.073	100.755	4.318	18.6450	0.2114
94.458	106.679	103.978	2.701	7.2950	-0.5260
95.992	109.823	105.533	4.290	18.4040	0.1957
97.835	114.258	107.403	6.855	46.9910	2.0531
98.343	112.070	107.917	4.153	17.2471	0.1205
99.832	106.052	109.427	-3.375	11.3906	-0.2599
102.375	104.390	112.006	-7.616	58.0030	2.7686
106.661	111.357	116.351	-4.994	24.9400	0.6204
				307.822	

$$\hat{\sigma}^2 = T^{-1} \sum_{t}^{T} \hat{u}_t^2 = 307.822 / 20 = 15.391$$

คำนวณค่าของตัวสถิติ $LM = \frac{1}{2} g' Z (Z' Z)^{-1} Z' g$

เมื่อ Z เป็นเมตริกซ์ขนาด (20 x 2) ซึ่งมีสมาชิกในแถวที่ t ใช้สัญลักษณ์เป็น

$Z_t = (1, X_t)$

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

แทนค่า

$$z'z = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 69.985 & 71.383 & \cdots & 106.661 \end{bmatrix} \begin{bmatrix} 1 & 69.985 \\ 1 & 71.383 \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ 1 & 106.661 \end{bmatrix} = \begin{bmatrix} 20 & 1760.606 \\ 1760.606 & 157082.539 \end{bmatrix}$$

$$(z'z)^{-1} = \frac{1}{41917.293} \begin{bmatrix} 157082.539 & -1760.606 \\ -1760.606 & 20 \end{bmatrix}$$

$$= \begin{bmatrix} 3.7474 & -0.042 \\ -0.042 & 0.000477 \end{bmatrix}$$

$$z'g = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 69.985 & 71.383 & \cdots & 106.661 \end{bmatrix} \begin{bmatrix} -0.454 \\ -0.796 \\ \cdot \\ \cdot \\ \cdot \\ 0.620 \end{bmatrix} = \begin{bmatrix} 0.008 \\ 143.8217 \end{bmatrix}$$

$$g'z = \begin{bmatrix} -0.454 & -0.796 & \cdots & 0.620 \end{bmatrix} \begin{bmatrix} 1 & 69.985 \\ 1 & 71.383 \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ 1 & 106.661 \end{bmatrix} = \begin{bmatrix} 0.008 & 143.8217 \end{bmatrix}$$

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

$$LM = \frac{1}{2}\left\{\begin{bmatrix} 0.008 & 143.8217 \end{bmatrix}\begin{bmatrix} 3.7474 & -0.042 \\ -0.042 & 0.000477 \end{bmatrix}\begin{bmatrix} 0.008 \\ 143.8217 \end{bmatrix}\right\}$$

$$= 9.818$$

ซึ่ง $LM = 9.818 > \chi^2_{(1),0.05} = 3.846$

ดังนั้นจึงปฏิเสธสมมติฐานหลัก ที่ระดับนัยสำคัญ 0.05

2.2.4 <u>การทดสอบ BAMSET</u> (Bartlett's M Specification Error test)

การทดสอบนี้ถูกเสนอขึ้นโดย Ramsey (1969 :350-370) เป็นการทดสอบที่มุ่งทำการทดสอบดูว่าประชากรกลุ่มต่าง ๆ m กลุ่ม จะมีลักษณะของความแปรปรวนเท่ากันหรือไม่ สมมติฐานที่มุ่งทดสอบคือ

$$H_o : \sigma_1^2 = \sigma_2^2 = \ldots = \sigma_m^2$$

H_1 : มี σ^2 อย่างน้อย 1 คู่ ที่แตกต่างกัน

โดยมีขั้นตอนในการดำเนินการทดสอบดังนี้

1. ค่าสังเกตจะถูกนำมาจัดลำดับตามการเพิ่มขึ้นของความแปรปรวน เช่นเดียวกับการทดสอบอื่น ๆ

2. จำแนกค่าสังเกต (X_t , Y_t) ; $t = 1,2, \ldots, T$ ออกเป็น m กลุ่ม กลุ่มละเท่า ๆ กัน การจำแนกจะใช้วิธีจำแนกใน ลักษณะให้ข้อมูลมีความคล้ายคลึงกันในกลุ่มเดียวกัน ใช้ตัวแปรสุ่มมี่ หรือใช้ตัวแปรจำแนกใด ๆ ก็ได้ เกี่ยวกับจำนวนกลุ่ม (m) นั้น Ramsey แนะนำว่าให้ใช้ m = 3 จะเป็นจำนวนที่พอเหมาะ ในการวิจัยครั้งนี้จึงใช้ m = 3 โดยเซ็ตของกลุ่มตัวอย่าง 3 กลุ่ม ปรากฏดังนี้

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

เมื่อขนาดตัวอย่าง (T) = 20, $S_1 = \{(X_t, Y_t) ; t = 1,2,..., 7\}$

$S_2 = \{(X_t, Y_t) ; t = 8,9,..., 13\}$

$S_3 = \{(X_t, Y_t) ; t = 14,15,...,20\}$

เมื่อขนาดตัวอย่าง (T) = 50, $S_1 = \{(X_t, Y_t) ; t = 1,2,..., 17\}$

$S_2 = \{(X_t, Y_t) ; t = 18,19,..., 33\}$

$S_3 = \{(X_t, Y_t) ; t = 34,35,...,50\}$

3. จากแต่ละกลุ่มตัวอย่างย่อยให้คำนวณหาเวกเตอร์ของ OLS-Residual ในขณะเดียวกันก็คำนวณหา เวกเตอร์ของ OLS-Residual ของค่าสังเกตเดิมก่อนการแบ่งกลุ่มขนาด T ด้วย

4. <u>สถิติทดสอบ</u>

$$BS = (T - 2) \ln \hat{\sigma}^2 - \frac{T-2}{3} \sum_{i}^{3} \ln s_i^2$$

เมื่อ

$$\hat{\sigma}^2 = (T - 2)^{-1} \sum_{t}^{T} \hat{u}_t^2$$

$$s_i^2 = (3/(T - 2)) \sum_{t \in S_i} \hat{u}_t^2$$

โดยที่ $\hat{u}_t$ คือ OLS-Residual ของค่าสังเกตที่ t

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

I10304198

การทดสอบนี้ขึ้นกับจำนวนกลุ่มและการเลือกใช้วิธีการในการประมาณค่าความคลาดเคลื่อน (Residual) เช่น Theil (1971) ใช้ BLUS และ Harvey และ Phillips (1974:307-316) ใช้ Recursive Residual แทน OLS-Residual การวิจัยนี้เลือกใช้ OLS-Residual เนื่องจากการทดลองของ Ramsey และ Gilbert (1972 :180-186) และ Ali และ Giaccotto (1984:355-373) พบว่าเมื่อใช้ OLS-Residual ตัวสถิติจะมีอำนาจการทดสอบสูงกว่าเมื่อใช้ BLUS หรือRecursive Residual

ภายใต้สมมติฐานหลัก ตัวสถิติ BS จะมีการแจกแจงเข้าใกล้การแจกแจงไคสแควร์ที่มีจำนวนชั้นของความเป็นอิสระเท่ากับ 2

5. เกณฑ์การตัดสินใจ จะตัดสินใจปฏิเสธสมมติฐานหลัก ณ ระดับนัยสำคัญ α เมื่อ

$$BS > \chi^2_{(2),\,\alpha}$$

โดย $\chi^2_{(2),\alpha}$ เป็นค่าวิกฤตที่ได้จากตารางการแจกแจงไคสแควร์ที่มีจำนวนชั้นแห่งความเป็นอิสระเท่ากับ 2

ตัวอย่างวิธีการคำนวณหาค่าสถิติทดสอบ

ตัวอย่างที่ 4 จากข้อมูลในตารางที่ 2.1 ของตัวอย่างที่ 1 สามารถนำมาทำการตรวจสอบปัญหาความคลาดเคลื่อนมีความแปรปรวนไม่คงที่ โดยการทดสอบ BAMSET ได้ดังนี้

วิธีทำ จัดลำดับค่าสังเกตตามการเพิ่มขึ้นของความแปรปรวน เช่นเดียวกับการทดสอบอื่น ๆ แล้วจำแนกค่าสังเกตออกเป็น 3 กลุ่ม กลุ่มละเท่า ๆ กันแล้ว คำนวณหาเวกเตอร์ของ OLS-Residual ของแต่ละกลุ่มตัวอย่างย่อย แสดงได้ดังนี้

	X_t	Y_t	$\hat{Y}_t$	$\hat{u}_t$	$\hat{u}_t^2$
	69.985	76.266	77.813	-1.547	2.393
	71.383	79.812	79.343	0.469	0.219
สมการถดถอยสำหรับ	73.610	88.455	81.779	3.676	13.513
กลุ่มตัวอย่างที่ 1	77.220	83.017	85.729	-2.715	7.371
$\hat{Y}_t = 1.243 + 1.094X_t$	79.177	86.658	87.870	-1.212	1.468
	81.863	91.345	90.809	0.536	0.287
	83.124	92.983	92.188	0.795	0.632
					25.882
	84.683	97.949	98.543	-0.594	0.352
	86.108	92.314	100.182	-7.868	61.905
สมการถดถอยสำหรับ	88.183	98.121	102.569	-4.448	19.784
กลุ่มตัวอย่างที่ 2	89.142	102.434	103.672	-1.238	1.532
$\hat{Y}_t = -3.163 + 1.150X_t$	89.353	93.282	103.914	-10.632	113.039
	91.279	105.073	106.129	-1.056	1.115
					197.728
	94.458	106.679	109.831	- 3.152	9.935
	95.992	109.823	109.265	0.558	0.311
	97.835	114.258	109.256	5.002	25.020
สมการถดถอยสำหรับ	98.343	112.070	109.244	2.826	7.986
กลุ่มตัวอย่างที่ 3	99.832	106.052	109.232	-3.180	10.112
$\hat{Y}_t = 109.831 - 0.006X_t$	102.375	104.390	109.217	-4.827	23.299
	106.661	111.357	109.192	2.165	4.687
					81.350

และจากค่าสังเกตทั้งหมดจำนวน 20 ค่านี้ ทำการวิเคราะห์สมการถดถอย โดยวิธี OLS แล้วหาผลบวกกำลังสองของความคลาดเคลื่อนได้เป็น

$$\sum_{t=1}^{20} \hat{u}_t^2 = 307.822$$

ดังนั้น $$\hat{\sigma}^2 = (T - 2)^{-1} \sum_{t}^{T} \hat{u}_t^2 = 307.822/18$$

$$= 17.1012$$

และจาก $$s_i^2 = (3/(T - 2)) \sum_{t \epsilon S_i} \hat{u}_t^2$$

$$s_1^2 = 0.167\ (25.882) = 4.314 \quad , \quad \ell n\ 4.314 = 1.4618$$

$$s_1^2 = 0.167\ (197.728) = 33.020 \quad , \quad \ell n\ 33.020 = 3.497$$

$$s_3^2 = 0.167\ (81.350) = 13.585 \quad , \quad \ell n\ 13.585 = 2.608$$

แทนค่า

$$BS = (T - 2)\ \ell n\hat{\sigma}^2 - \frac{T - 2}{3} \sum_{i}^{3} \ell n\ s_i^2$$

$$= 18\ (2.8391) - 6\ (1.4618 + 3.4970 + 2.6080)$$

$$= 5.7038$$

ซึ่ง $$BS = 5.7038 < \chi^2_{(2),0.05} = 5.991$$

ดังนั้นยอมรับสมมติฐานหลัก ที่ระดับนัยสำคัญ 0.05

2.3 ผลงานวิจัยที่เกี่ยวข้อง

สำหรับผลงานวิจัยที่เกี่ยวข้องนั้น ปรากฏว่ามีนักวิจัยเป็นจำนวนมากได้ทำการศึกษาอำนาจของการทดสอบ และความแกร่งของสถิติทดสอบที่ใช้ในการตรวจสอบปัญหาความคลาดเคลื่อนมีความแปรปรวนไม่คงที่ ซึ่งในส่วนนี้จะเสนอเฉพาะผลงานวิจัยที่สำคัญพอเป็นสังเขปเท่านั้น

Adesi และ Talwar ศึกษาเปรียบเทียบอำนาจการทดสอบและความแกร่งของการทดสอบแบบนอนพาราเมตริกคือ การทดสอบพีค (peak test) การทดสอบบิคเคล (Bickel's test) การทดสอบของจอนส์ตัน (Johnston) การทดสอบแคนดอล (Kandall test) เทียบกับ การทดสอบแบบพาราเมตริกคือ การทดสอบโกลฟีลด์และควอนท์ การทดสอบเกลเซอร์ (Glyser test) และการทดสอบ บาร์ตเลต (Bartlett test) โดยกำหนดให้ความแปรปรวนของความคลาดเคลื่อนมีรูปแบบเป็น $\sigma_t^2 = 2X_t$, $\sigma_t^2 = X_t^2$ และ $\sigma_t^2 = \ln X_t$; $t = 1,2,\ldots, T$ เมื่อความคลาดเคลื่อนมีการแจกแจงแบบต่าง ๆ ได้ข้อสรุปดังนี้

1. การทดสอบโกลฟีลด์และควอนท์ และการทดสอบบาร์ตเลตไม่แกร่ง (nonrobust) เมื่อความคลาดเคลื่อนมีการแจกแจงแบบเบ้ (skewes) หรือโด่งมาก ๆ (large krutosis)

2. การทดสอบเกลเซอร์ เป็นการทดสอบที่แกร่ง แต่เป็นการทดสอบที่มีอำนาจการทดสอบต่ำ เมื่อเทียบกับการทดสอบโกลฟีลด์และควอนท์และการทดสอบบาร์ตเลต

3. เมื่อความคลาดเคลื่อนมีการแจกแจงแบบปกติ การทดสอบโกลฟีลด์และควอนท์เป็นการทดสอบที่มีอำนาจการทดสอบสูงที่สุด

4. การทดสอบบิคเคล เมื่อ $\alpha = 1$ มีความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1 และมีอำนาจการทดสอบมากกว่าเมื่อ $\alpha = 1.5$ และ 2

5. เมื่อความคลาดเคลื่อนมีการแจกแจงแบบไม่ปกติ การทดสอบของจอนห์ตัน (Johnston) และการทดสอบบิคเคล ($\alpha = 1$) ยังคงแกร่งและมีอำนาจการทดสอบพอ ๆ กัน แต่มีอำนาจการทดสอบสูงเมื่อเทียบกับการทดสอบพีคและการทดสอบแคนดอล

6. การทดสอบทั้งหมด มีอำนาจการทดสอบลดลงไม่ว่าความคลาดเคลื่อนจะมีการแจกแจงแบบใด เมื่อความแปรปรวนมีรูปแบบเป็น $\sigma_t^2 = \ln X_t$

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Ali และ Giaccotto (1984:355-373) ได้ทำการศึกษาอำนาจการทดสอบและความแกร่ง ในกลุ่มของการทดสอบแบบนอนพาราเมตริก ร่วมกับการทดสอบแบบพาราเมตริกคือ การทดสอบโกลด์ฟิลด์และควอนท์ การทดสอบของแรมเซย์ (Ramsey) การทดสอบของทีลล์ (Theil) การทดสอบของ ฮาร์วี และฟิลลิปป์ (Harvey and Phillips) การทดสอบสโรเตอร์ (Szroeter test) การทดสอบบรูสและพาแกน (Breusch and Pagan test) และการทดสอบของแฮร์สันและ เมคเคลป (Harrison and McCabe) โดยสนใจที่จะศึกษาอิทธิพลของวิธีที่ใช้ในการประมาณค่าความคลาดเคลื่อน การแจกแจงและการมีสห-สัมพันธ์ต่อกัน ของความคลาดเคลื่อน ขนาดตัวอย่าง รูปแบบของความแปรปรวน ที่มีผลต่อความแกร่งและอำนาจของการทดสอบ โดยกำหนดให้ความแปรปรวนมีรูปแบบเป็น

$\sigma_t^2 = \sigma^2 X_t^2$, $\sigma_t^2 = \sigma^2 X_t$, $\sigma_t^2 = \sigma^2 (E(Y_t))^2$ $\sigma_t^2 = \sigma^2 (E(Y_t))$;

$t = 1,2,\ldots, T$ ที่ขนาดตัวอย่าง 10 25 40 ได้ข้อสรุปดังนี้

1. การทดสอบทุกการทดสอบจะแกร่งสำหรับการประมาณค่าความคลาดเคลื่อนโดยใช้ OLS-Residual และไม่ว่าจะเป็นการทดสอบแบบนอนพาราเมตริกหรือการทดสอบแบบพาราเมตริก การทดสอบที่ประมาณค่าความคลาดเคลื่อนโดยใช้ OLS-Residual จะมีอำนาจของการทดสอบสูงกว่าการทดสอบที่ประมาณค่าความคลาดเคลื่อนโดยใช้ BLUS (Best Linear unbiased Scalar variance-Covariance matrix) หรือ Recursive-Residual

2. ความแกร่งของการทดสอบสโรเตอร์ การทดสอบของแรมเซย์ และการทดสอบบรูสและพาแกน มีผลกระทบเนื่องมาจากขนาดตัวอย่าง

3. อำนาจของการทดสอบจะมีค่าลดต่ำลง ถ้าความคลาดเคลื่อนไม่ได้มีการแจกแจงแบบปกติ หรือมีสหสัมพันธ์ต่อกัน และในกรณีที่ความคลาดเคลื่อนมีการแจกแจงแบบเบ้ (skewed) หรือหางยาวมาก ๆ (long tailed) การทดสอบแบบพาราเมตริกมีแนวโน้ม(trend) ที่ไม่ถูกต้องมากกว่าการทดสอบแบบนอนพาราเมตริก

4. การทดสอบทั้งหมดจะมีอำนาจการทดสอบสูงขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น

5. รูปแบบของความแปรปรวน มีผลต่อค่าอำนาจของการทดสอบกล่าวคือ การทดสอบจะเหมาะสมที่สุด สำหรับเฉพาะรูปแบบของความแปรปรวน

Evan และ King (1985:163-178) ใช้วิธีการมอนติคาร์โล ศึกษาเปรียบเทียบอำนาจของการทดสอบสโรเตอร์ การทดสอบบรูสและพาแกน การทดสอบโกลด์ฟิลด์และควอนท์

การทดสอบแอริสันและเมคเคลป เมื่อความแปรปรวนมีรูปแบบของการบวก (additive model) เป็น $\sigma_t^2 = \sigma^2(1 + Z_t)^k$; $t = 1,2,\ldots, T$; $k = 1,2$ Z_t แทนตัวแปรอิสระหรือฟังก์ชั่นของตัวแปรอิสระ X ตัวใดตัวหนึ่ง λ เป็นพารามิเตอร์ ทำการทดลองภายใต้ตัวแปรอิสระที่มีการแจกแจงแบบยูนิฟอร์ม และล็อคนอมอลโดยสนใจเปรียบเทียบอำนาจการทดสอบเมื่อ C.V.(σ_t^2) มีค่าต่าง ๆ ซึ่งถูกควบคุมโดยการแปรค่าของพารามิเตอร์ λ ซึ่งผลปรากฏว่า เมื่อ C.V. (σ_t^2) มีค่าไม่มากนักการทดสอบสโรเตอร์จะมีอำนาจการทดสอบสูงสุด ทุกการแจกแจงของตัวแปรอิสระ ยกเว้นกรณีที่ $k = 1$, $Z_t = X_t^2$ และตัวแปรอิสระมีการแจกแจงแบบล็อคนอมอลแล้ว การทดสอบบรูสและพาแกน จะมีอำนาจการทดสอบสูงกว่าการทดสอบสโรเตอร์ และการทดสอบจะมีอำนาจการทดสอบสูงขึ้น เมื่อ C.V. (σ_t^2) และขนาดตัวอย่างมีค่าเพิ่มขึ้น นอกจากนี้ยังพบว่า เมื่อ C = 4 สำหรับขนาดตัวอย่าง 20 และ c = 8 สำหรับขนาดตัวอย่าง 40 การทดสอบโกลฟิลด์และควอนท์จะมีอำนาจการทดสอบสูงกว่า เมื่อ c = 0 สำหรับขนาดตัวอย่าง 20 และc = 0,4 สำหรับขนาดตัวอย่าง 40

ความสัมพันธ์ระหว่างงานวิจัยนี้กับผลงานวิจัยที่เกี่ยวข้อง

สำหรับการวิจัยนี้ได้กระทำการวางแผนการทดลองโดยอาศัยแบบแผนการทดลองของ Evan และ King เป็นหลัก และใช้ผลงานวิจัยของอีกหลาย ๆ ท่านในการพิจารณาความเหมาะสม ในการกำหนดค่าพารามิเตอร์ต่าง ๆ ดังนั้น การวิจัยนี้จึงมีลักษณะคล้ายคลึงกับแผนการทดลองของ Evan และ King เป็นส่วนใหญ่ แต่อย่างไรก็ตาม Evan และ King ก็ได้ทำการศึกษาเฉพาะกรณีที่สามารถจัดลำดับค่าสังเกตตามการเพิ่มขึ้นของความแปรปรวนได้เท่านั้น ซึ่งในการวิจัยนี้ได้ทำการศึกษาในกรณีที่ไม่สามารถจัดลำดับค่าสังเกตตามการเพิ่มขึ้นของความแปรปรวนด้วย

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย