บทที่ 2

สถิติทดสอบและผลงานวิจัยที่เกี่ยวข้อง

ในบทนี้จะกล่าวถึงรายละเอียดของการทดสอบแต่ละวิธีและรายละเอียดของการแจกแจงที่สำคัญที่ใช้ในงานวิจัยครั้งนี้ รวมทั้งนำเสนอผลงานวิจัยที่เกี่ยวข้องดังรายละเอียดต่อไปนี้ ในการวิจัยครั้งนี้ใช้รูปแบบของสมการถดถอยเชิงเส้นอย่างง่ายที่มีรูปแบบดังนี้

$$Y_t = \beta_o + \beta_1 X_t + U_t$$

และรูปแบบของอัตตสหสัมพันธ์เป็นแบบ AR(1) คือ

$$U_t = \rho U_{t-1} + V_t$$

สมมติฐานของการทดสอบอัตตสหสัมพันธ์ทางบวกคือ

$$H_o : \rho = 0$$

$$H_1 : \rho > 0$$

ตัวสถิติต่าง ๆ ที่ใช้ในการทดสอบสมมติฐานข้างต้นมีรายละเอียดดังนี้

2.1 สถิติทดสอบเดอร์บินและวัตสัน

เดอร์บินและวัตสัน เป็นผู้เสนอสถิติทดสอบนี้ขึ้นในปี ค.ศ. 1950 DW นับเป็นเครื่องมือทดสอบอัตตสหสัมพันธ์ใน AR(1) ที่มีผู้นิยมใช้ค่อนข้างกว้างขวางแม้จะได้พัฒนามาตั้งแต่ปี ค.ศ. 1950 สถิติทดสอบของเดอร์บินและวัตสันมีขั้นตอนที่ใช้ในการทดสอบดังนี้

2.1.1 วิเคราะห์สมการถดถอย $Y = X\beta + U$ ด้วยวิธี OLS ได้ $\hat{\underset{\sim}{\beta}} = (X'X)^{-1}X'\underset{\sim}{Y}$ Residual Vector คือ $\underset{\sim}{e} = \underset{\sim}{Y} - X\hat{\underset{\sim}{\beta}}$

2.1.2 จากเวคเตอร์ $\underset{\sim}{e}(e_1\ e_2, \ldots, e_n)'$ นำมาคำนวณโดยที่สถิติทดสอบเดอร์บินและวัตสันมีรูปแบบดังนี้

$$d = \frac{\sum_{t=2}^{n} (e_t - e_{t-1})^2}{\sum_{t=1}^{n} e_t^2}$$

2.1.3 เกณฑ์การตัดสินใจ

จะปฏิเสธ H_o เมื่อ $d < d_L$

จะยอมรับ H_o เมื่อ $d > d_U$

ไม่อาจตัดสินใจได้เมื่อ $d_L < d < d_U$

โดยที่ d_L และ d_U เป็นค่าวิกฤตที่เปิดได้จากตารางค่าของเดอร์บินและวัตสัน ที่ระดับนัยสำคัญ α และขนาดตัวอย่าง n $(n \geqslant 15)$

2.2 <u>สถิติทดสอบเบเรนบรูตและเวบบ์</u>

ขั้นตอนการทดสอบที่สำคัญมีดังนี้

2.2.1 วิเคราะห์สมการถดถอย $Y = X\beta + U$ ด้วยวิธี OLS ได้ $\hat{\underset{\sim}{\beta}} = (X\ X)^{-1} X\underset{\sim}{Y}$ และ Residual Vector คือ $\underset{\sim}{e} = \underset{\sim}{Y} - X\hat{\underset{\sim}{\beta}}$

2.2.2 จากสมการ $CY = (CX)\beta + CU$ คือสมการถดถอย First Difference นั่นเองคือ

$$\begin{bmatrix} Y_1 \\ Y_2 - Y_1 \\ Y_3 - Y_2 \\ \vdots \\ \cdot \\ Y_n - Y_{n-1} \end{bmatrix} = \begin{bmatrix} 1 & X_{21} \\ 0 & X_{22} - X_{21} \\ 0 & X_{23} - X_{22} \\ & \vdots \\ & \cdot \\ 0 & X_{2n} - X_{2,n-1} \end{bmatrix} \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} + \begin{bmatrix} U_1 \\ U_2 - U_1 \\ U_3 - U_2 \\ \vdots \\ \cdot \\ U_n - U_{n-1} \end{bmatrix}$$

ตัดค่าสังเกต $Z_1 = (Y_1, X_{21})$ ทิ้ง แล้วหา Residual Sum Square จากสมการนี้นั่นก็คือ Residual Sum Square จากสมการ

$$Y_t - Y_{t-1} = \beta_2(X_{2t} - X_{2,t-1}) \quad ; \quad t = 2,3,\ldots, n$$

2.2.3 จากเวคเตอร์ $e\ (e_1,e_2, \ldots , e_n)'$ ในข้อ 2.2.1 และ Residual Sum Square จากข้อ 2.2.2 นำมาหาค่าสถิติทดสอบแบเรนบรูตและเวบบ์ที่มีรูปแบบดังนี้

$$g = \frac{RSS^*}{RSS}$$

โดยที่ RSS^* คือ Residual Sum Square จากข้อ 2.2.2

และ RSS คือ Residual Sum Square จากข้อ 2.2.1

2.2.4 เกณฑ์การตัดสินใจ

จะปฏิเสธ H_o เมื่อ $g < d_L$

จะยอมรับ H_o เมื่อ $g > d_U$

ไม่อาจตัดสินใจได้เมื่อ $d_L > g > d_U$

โดยที่ d_L และ d_U เป็นค่าวิกฤตที่เปิดได้จากตารางค่าของเดอร์บินและวัตสัน ที่ระดับนัยสำคัญ α และขนาดตัวอย่าง $n\ (n \geqslant 15)$

2.3 <u>สถิติทดสอบเกียรี่</u>

ขั้นตอนการทดสอบที่สำคัญมีดังนี้

2.3.1 วิเคราะห์สมการถดถอย $Y = X\beta + U$ ด้วยวิธี OLS ได้ $\hat{\beta} = (X'X)^{-1} X'Y$ และ Residual Vector คือ $e = Y - X\hat{\beta}$

2.3.2 จากเวคเตอร์ $e\ (e_1, e_2, \ldots, e_n)'$ นำมาหาค่าสถิติทดสอบเกียรี่ที่มีรูปแบบดังนี้

τ = จำนวนครั้งทั้งหมดของการเปลี่ยนแปลงเครื่องหมายของความคลาดเคลื่อน

2.3.3 เกณฑ์การตัดสินใจ

จะยอมรับ H_o เมื่อ Min $\tau \leqslant \tau \leqslant$ Max τ นอกนั้นปฏิเสธ H_o

โดยที่ Min τ และ Max τ เป็นค่าวิกฤตที่เปิดได้จากตารางค่าของเกียรี่ที่ระดับนัยสำคัญ α และขนาดตัวอย่าง n

2.4 วิธีคำนวณค่าสถิติที่ใช้ในการทดสอบอัตตสหสัมพันธ์ทางบวก

ในการแสดงวิธีการคำนวณค่าสถิติดังกล่าวจะอาศัยข้อมูลของตัวอย่างสุ่มขนาดเท่ากับ 15 เมื่อ $\rho = 0.0$ ดังต่อไปนี้

X	Y
207.530	321.845
194.988	303.142
207.292	321.587
229.613	352.422
210.830	327.268
191.626	295.061
193.191	299.994
196.563	303.448
192.587	298.936
192.500	298.035
189.144	293.327
218.800	337.334
203.651	315.460
224.866	347.923
206.312	319.695

จากข้อมูลข้างต้นสามารถแสดงวิธีการคำนวณค่าสถิติทั้ง 3 วิธีได้ดังนี้

2.4.1 วิธีการคำนวณค่าสถิติทดสอบเดอร์บินและวัตสัน

สถิติทดสอบ

$$d = \frac{\sum_{t=2}^{n} (e_t - e_{t-1})^2}{\sum_{t=1}^{n} e_t^2}$$

การคำนวณ

สามารถทำได้ 2 วิธี คือ

<u>วิธีที่</u> <u>1</u> นำข้อมูลที่มีตัวแปร 2 ตัวคือ ตัวแปร X และตัวแปร Y มาวิเคราะห์หาสมการถดถอยเชิงเส้นอย่างง่ายได้สมการถดถอยออกมาคือ

$$\hat{Y} = 10.717 + 1.496X \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (1)$$

นำค่า X แทนลงในสมการที่ (1) ทั้ง 15 ค่าได้ค่า $\hat{Y}$ ออกมา 15 ค่า

หาค่า $e = Y - \hat{Y}$ ออกมาทั้ง 15 ค่า ผลที่ได้เป็นดังนี้

$$e_1 = 0.663$$

$$e_2 = 0.723$$

$$e_3 = 0.761$$

$$e_4 = -1.769$$

$$e_5 = 1.149$$

$$e_6 = -2.329$$

$$e_7 = 0.263$$

$$e_8 = 0.673$$

$$e_9 = 0.109$$

$$e_{10} = -0.662$$

$$e_{11} = -0.349$$

$$e_{12} = -0.708$$

$$e_{13} = 0.081$$

$$e_{14} = 0.807$$

$$e_{15} = 0.335$$

$$\sum_{t=1}^{15} e_t^2 = (0.663)^2 + (0.723)^2 + \ldots + (0.335)^2$$

$$= 13.872831$$

$$\sum_{t=2}^{15} (e_t - e_{t-1})^2 = (0.723 - 0.663)^2 + (0.761 - 0.723)^2 + \ldots + (0.335 - 0.807)^2$$

$$= 36.711134$$

$$d = \frac{36.711134}{13.872831}$$

$$= 2.646$$

วิธีที่ 2

ใช้โปรแกรมสำเร็จรูปซึ่งมีคำสั่งที่ให้ค่าของ $d = 2.646$

2.4.2 วิธีการคำนวณค่าสถิติทดสอบเบเรนบรูตและเวบบ์

สถิติทดสอบ

$$g = \frac{RSS^*}{RSS}$$

การคำนวณ

จากข้อ 2.4.1 ได้ค่า $\sum_{t=1}^{15} e_t^2$ ซึ่งในที่นี้ก็คือค่า RSS นั่นเองส่วนค่า RSS^* นั้นมีวิธีการคำนวณดังนี้

นำค่า X และ Y มาหาค่า X^* และ Y^* ซึ่งได้จากผลต่างของ X และ Y ตามลำดับโดยที่ข้อมูลจะลดเหลือ n = 14 ดังต่อไปนี้

X^*	Y^*
$X_2 - X_1 = -12.542$	$Y_2 - Y_1 = -18.703$
$X_3 - X_2 = 12.304$	$Y_3 - Y_2 = 18.445$
$X_4 - X_3 = 22.321$	$Y_4 - Y_3 = 30.835$
$X_5 - X_4 = -18.783$	$Y_5 \ Y_4 = -25.154$
$X_6 - X_5 = -19.204$	$Y_6 - Y_5 = -32.207$
$X_7 - X_6 = 1.565$	$Y_7 - Y_6 = 4.933$
$X_8 - X_7 = 3.372$	$Y_8 - Y_7 = 5.454$
$X_9 - X_8 = -3.976$	$Y_9 - Y_8 = -6.512$
$X_{10} - X_9 = -0.087$	$Y_{10} - Y_9 = -0.901$
$X_{11} - X_{10} = -3.356$	$Y_{11} - Y_{10} = -4.708$
$X_{12} - X_{11} = 29.656$	$Y_{12} - Y_{11} = 44.007$
$X_{13} - X_{12} = -15.149$	$Y_{13} - Y_{12} = -21.784$
$X_{14} - X_{13} = 21.215$	$Y_{14} - Y_{13} = 32.463$
$X_{15} - X_{14} = -18.554$	$Y_{15} - Y_{14} = -28.228$

นำค่า X^* และ Y^* มาวิเคราะห์หาสมการถดถอยเชิงเส้นอย่างง่ายที่ไม่ต้องมีค่าจุดตัดบนแกน $Y(\beta_0)$ กล่าวคือ

$$\hat{Y}^* = 1.485X^* \qquad \ldots\ldots\ldots\ldots\ldots\ldots(2)$$

นำค่า X^* แทนลงในสมการที่ (2) ทั้ง 14 ค่าได้ค่า $\hat{Y}^*$ 14 ค่าแล้ว

หาค่า $e^* = Y^* - \hat{Y}^*$ ผลที่ได้เป็นดังนี้

$$e_1^* = -0.078$$

$$e_2^* = 0.174$$

$$e_3^* = -2.312$$

$$e_4^* = 2.739$$

$$e_5^* = -3.689$$

$$e_6^* = 2.609$$

$$e_7^* = 0.447$$

$$e_8^* = -0.608$$

$$e_9^* = -0.772$$

$$e_{10}^* = 0.276$$

$$e_{11}^* = -0.032$$

$$e_{12}^* = 0.712$$

$$e_{13}^* = 0.959$$

$$e_{14}^* = -0.675$$

$$RSS^* = \sum_{t=1}^{14} e_t^{*2} = (-0.078)^2 + (0.174)^2 + \ldots + (-0.675)^2$$

$$= 36.424334$$

$$g = \frac{RSS^*}{RSS} = \frac{36.424334}{13.872831}$$

$$= 2.626$$

2.4.3 วิธีการคำนวณค่าสถิติทดสอบเกียรี่

สถิติทดสอบ

τ = จำนวนครั้งทั้งหมดของการเปลี่ยนแปลงเครื่องหมายของความคลาดเคลื่อน

การคำนวณ

จากข้อ 2.4.1 ได้ค่า $e(e_1, e_2, \ldots, e_{15})'$ แล้วนับจำนวนครั้งของการเปลี่ยนเครื่องหมาย ผลที่ได้คือ

$$\tau = 6$$

ดังนั้นค่าสถิติทดสอบเกียรี่เท่ากับ 6

เนื่องจากในการวิจัยครั้งนี้สนใจที่จะศึกษาถึงความแกร่งและอำนาจของการทดสอบของตัวสถิติดังกล่าวทั้ง 3 เมื่อข้อมูลที่มีอยู่นั้นความคลาดเคลื่อนมีการแจกแจงแบบสมมาตรหางยาว ดังนั้นจึงเลือกศึกษากับข้อมูลที่สร้างจากการแจกแจงแบบโลจิสติค แบบดับเบิ้ลเอ็กซ์-โปเนนเชียล นอกจากนี้ยังได้ศึกษาตัวสถิติเหล่านี้กับข้อมูลที่สร้างจากการแจกแจงแบบปกติด้วยเพื่อใช้ทำการเปรียบเทียบผล ซึ่งรายละเอียดและคุณสมบัติต่าง ๆ เกี่ยวกับการแจกแจงดังกล่าวข้างต้นเป็นดังนี้

2.5 การแจกแจงแบบปกติ

การแจกแจงแบบปกติอาจกล่าวได้ว่าเป็นการแจกแจงความน่าจะเป็นแบบต่อเนื่องที่สำคัญที่สุด และเป็นการแจกแจงซึ่งใช้ในการอธิบายข้อมูลที่เกิดขึ้นโดยทั่วไปทางธรรมชาติ เช่น น้ำหนัก ความสูงเป็นต้น ในการวิเคราะห์ทางสถิติโดยใช้พารามิเตอร์มักจะตั้งอยู่บนข้อสมมติว่าการแจกแจงของประชากรมีการแจกแจงแบบปกติ โดยมีฟังก์ชันความหนาแน่น (Probability

Density Function) ดังนี้

$$f(x) \quad = \quad \frac{1}{\sqrt{2\pi\sigma}} \exp \quad \left[- (x - \mu)^2 / 2\sigma^2\right] \quad ; \quad - \alpha < x < \alpha$$

โดยที่ $f(x)$ = แทนความสูงของโค้งที่วัดจากแกนนอน ณ จุดใด ๆ ทุกจุด (x)

σ = ความเบี่ยงเบนมาตรฐานของประชากร

σ^2 = ความแปรปรวนของประชากร

μ = ค่าเฉลี่ยของประชากร

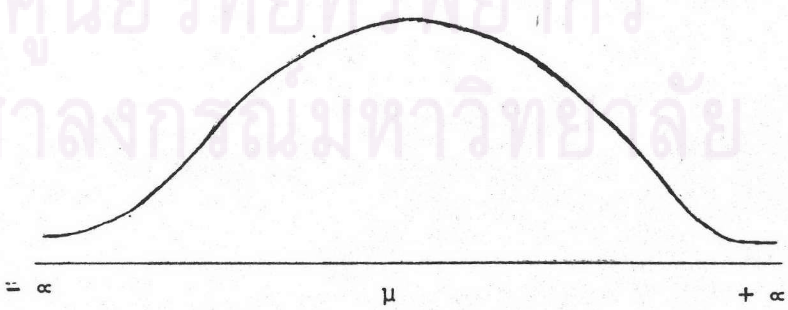
π = 3.14159

e = 2.71828

x = ค่าของข้อมูลจากกลุ่มตัวอย่าง

และค่า μ ,σ เป็นพารามิเตอร์ที่บอกถึงลักษณะของประชากรว่าประชากรนั้นมีตำแหน่งอยู่ที่ใด และมีการกระจายมากน้อยเพียงใด

$$; - \alpha < \mu < \alpha \; ; \; \sigma^2 > 0$$



$$E(X) = \mu$$

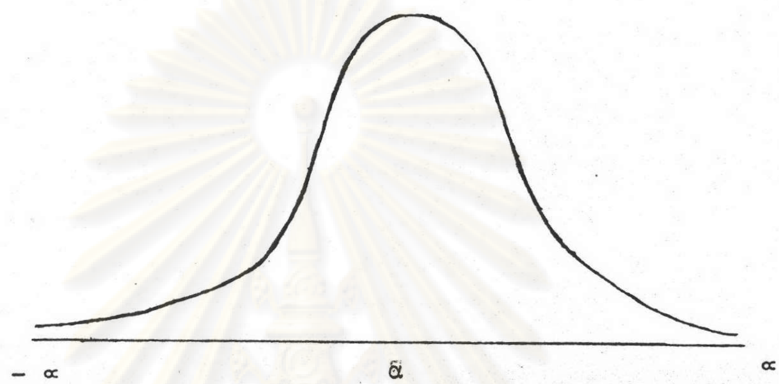
$$V(X) = \sigma^2$$

$$\text{Kurtosis} = 3.0$$

2.6 การแจกแจงแบบโลจิสติค

เป็นการแจกแจงที่ต่อเนื่องที่คล้ายการแจกแจงแบบปกติ โดยมีฟังก์ชันความหนาแน่นดังนี้

$$f(X) = \frac{1}{\beta} \frac{\exp\left[-(X-\alpha)/\beta\right]}{\left[1+\exp\left[-(X-\alpha)/\beta\right]\right]^2} \quad ; \; -\alpha < x < \alpha$$



$$E(X) = \alpha$$

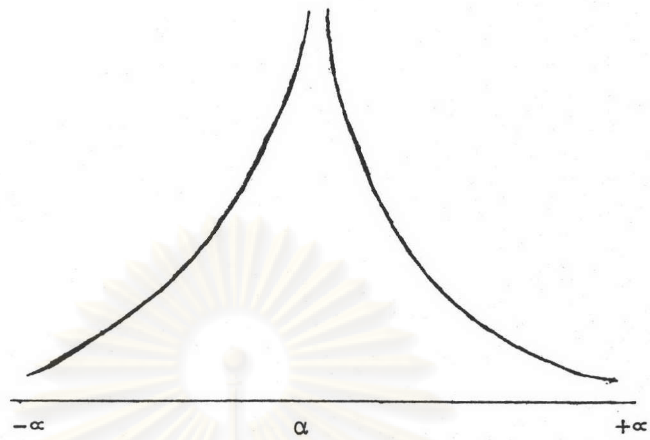
$$V(X) = \frac{1}{3}(\pi\beta)^2$$

$$\text{Kurtosis} = 4.2$$

2.7 การแจกแจงแบบดับเบิ้ลเอ็กซ์โปเนนเชียล

เป็นการแจกแจงแบบต่อเนื่องที่สมมาตร ซึ่งมีฟังก์ชันความหนาแน่นคือ

$$f(x) = \frac{1}{2\beta} \exp\left[- |X - \alpha| / \beta\right] \quad ; - \alpha < x < \alpha$$



$$E(X) = \alpha$$

$$V(X) = 2\beta^2$$

2.8 <u>ผลงานวิจัยที่เกี่ยวข้อง</u>

สำหรับผลงานวิจัยที่เกี่ยวข้องในการศึกษาเกี่ยวกับการทดสอบอัตตสหสัมพันธ์ ในการวิเคราะห์ความถดถอยเชิงเส้นอย่างง่าย (Simple Linear Regression Analysis) นั้น มีนักสถิติหลายท่านได้ทำการศึกษาไว้ ซึ่งแต่ละงานวิจัยนั้นล้วนแล้วแต่ใช้วิธีการมอนติคาร์โล (Monte Carlo Method) ทั้งสิ้น ส่วนผลงานวิจัยที่จะนำเสนอต่อไปนี้เป็นการศึกษาเปรียบ-เทียบตัวสถิติต่าง ๆ โดยอาศัยค่าอำนาจของการทดสอบเป็นเกณฑ์ดังมีรายละเอียดต่อไปนี้

อามิดและจอห์น (Hamid and John:1971: 179-185) ได้ศึกษาถึงอำนาจของการทดสอบ (Power of the test) ระหว่างการทดสอบเดอร์บินและวัตสัน(DW) การทดสอบเกียรี่ (Geory) เมื่อขนาดตัวอย่างเท่ากับ 15, 17,..., 51 ภายใต้ลักษณะการแจกแจงของความคลาดเคลื่อนเป็นแบบปกติ ณ ระดับนัยสำคัญ 0.05 รูปแบบอัตตสหสัมพันธ์เป็นแบบ AR(1) และเป็นอัตตสหสัมพันธ์ทางบวก (Positive Autocorrelation) ซึ่งได้ข้อสรุปดังนี้

1. เมื่อสัมประสิทธิ์สหสัมพันธ์ (ρ) เพิ่มขึ้น การทดสอบเดอร์บินและวัตสัน กับการทดสอบเกียร์ มีอำนาจของการทดสอบเพิ่มขึ้น

2. การทดสอบเดอร์บินและวัตสันมีอำนาจการทดสอบสูงกว่าการทดสอบเกียร์

3. เมื่อขนาดตัวอย่างใหญ่ ๆ การทดสอบเกียร์มีอำนาจของการทดสอบค่อนข้างพอ ๆ กับการทดสอบเดอร์บินและวัตสัน

นอกจากนี้ทั้ง 2 ท่านยังได้เสนอว่าเมื่อขนาดตัวอย่างมากกว่า 50 การทดสอบเกียร์ให้ผลดีพอ ๆ กับการทดสอบเดอร์บินและวัตสัน และการคำนวณหาค่าสถิติทดสอบก็ง่ายกว่าวิธีการของเดอร์บินและวัตสันด้วย ดังนั้นจึงเหมาะสมที่จะนำไปใช้ในทางปฏิบัติเช่นเดียวกับวิธีการของเดอร์บินและวัตสัน