บทที่ 2
ทฤษฎีของแบบจำลองโลจิต

เมื่อเข้าใจแนวคิดพื้นฐานของแบบจำลองโลจิตมาบ้างแล้ว ลำดับต่อไปเป็นการทำความเข้าใจอย่างลึกซึ้งกับตัวทฤษฎี ซึ่งแน่นนอนว่าต้องอ้างอิงผลงานของปรมาจารย์คือ George Judge และ William Greene ดังนั้น เพื่อไม่ให้เนื้อหาขาดตกบกพร่องหรือบิดเบือน และเพื่อประโยชน์ในการค้นคว้าอ้างอิง ผู้เขียนจึงได้แปลและเรียบเรียงคำอธิบายเชิงทฤษฎีที่ George Judge, et al (1988) และ Greene (2003) ได้กล่าวถึงแบบจำลองโลจิตไว้ในบทนี้ โดยยกตัวอย่างการประยุกต์ใช้ในเรื่องการท่องเที่ยว ท่านผู้อ่านสามารถอ่านต้นฉบับภาษาอังกฤษได้จากหนังสือของปรมาจารย์ทั้งสองท่านตามที่ระบุไว้ในรายการเอกสารอ้างอิงท้ายบท

2.1 ทฤษฎีตามคำอธิบายของ George Judge, et al

เมื่อกำหนดให้ P_i แทนโอกาสที่นักท่องเที่ยวจะมาเที่ยว หรือ โอกาสที่ $y_i = 1$

และ $1 - P_i$ แทนโอกาสที่นักท่องเที่ยวจะไม่มาเที่ยว หรือ โอกาสที่ $y_i = 0$

จะสามารถเขียน probability function ได้ดังนี้

$$f(y_i) = P_i^{y_i}(1 - P_i)^{1-y_i} \qquad y_i = 0,1 \qquad \text{…………..(2.1)}$$

ในการศึกษาวิจัย ผู้วิจัยสนใจที่จะวิเคราะห์ถึงปัจจัยที่มีอิทธิพลต่อ P_i ซึ่งแยกได้เป็นสองส่วนคือ

1. ลักษณะของทางเลือก (Attributes of the choice) ให้แทนด้วย z_i
2. ลักษณะของผู้เลือก (Attributes of the individual) ให้แทนด้วย w_i

และเมื่อกำหนดให้ U_{i0} และ U_{i1} เป็นอรรถประโยชน์ของบุคคลที่ i สำหรับทางเลือกที่ไม่ได้เลือก และ ทางเลือกที่ตัดสินใจเลือก ตามลำดับ และ $\overline{U}$ เป็นอรรถประโยชน์เฉลี่ยของบุคคลคน แล้ว จะเขียนได้ดังนี้ (Judge et al, 1988)

$$U_{i0} = \overline{U}_{i0} + e_{i0} \quad = z'_{i0}\delta + w'_i\gamma_0 + e_{i0} \qquad \text{.............(2.2)}$$

$$U_{i1} = \overline{U}_{i1} + e_{i1} \quad = z'_{i1}\delta + w'_i\gamma_1 + e_{i1} \qquad \text{.............(2.3)}$$

บุคคลหนึ่ง ๆ จะเลือกทางเลือกใด ๆ ก็ต่อเมื่อ $U_{i1} > U_{i0}$ หรือหากสร้างตัวแปร latent ขึ้นมา คือ $y_i^* = U_{i1} - U_{i0}$ แล้วจะพบผลลัพธ์ดังนี้

$$y_i = \begin{cases} 1 & if \quad y_i^* > 0 \\ 0 & if \quad y_i^* \le 0 \end{cases}$$

ซึ่งตัวแปรดังกล่าวสามารถเขียนใหม่ได้ว่า

$$y_i^* = (z_{i1} - z_{i0})'\delta + w'_i(\gamma_1 - \gamma_0) + (e_{i1} - e_{i0}) \qquad \text{.............(2.4)}$$

$$= \left[(z_{i1} - z_{i0})'\delta, w'_i\right] \begin{bmatrix} \delta \\ \gamma_1 - \gamma_0 \end{bmatrix} + e_i^* \qquad \text{.............(2.5)}$$

$$= x'_i\beta + e_i^* \qquad \text{.............(2.6)}$$

ซึ่ง x_i คือ เมตริกซ์ของตัวแปรต้น (explanatory variables)

β คือ เวกเตอร์ของค่าสัมประสิทธิ์จากการประมาณค่า

e_i^* คือ เวกเตอร์ของค่าคลาดเคลื่อน

แล้วโอกาสที่ $y_i = 1$ จะมีค่าดังต่อไปนี้

$$P_i = \Pr[y_i = 1] = \Pr[y_i^* > 0] = \Pr[e_i^* > -x_i'\beta] \quad \text{…………..(2.7)}$$

ในการนี้ ค่าคลาดเคลื่อนอาจจะมีการกระจายในหลายรูปแบบ หากค่าคลาดเคลื่อนมีการกระจายในลักษณะของ normal distribution แล้ว จะทำให้ cumulative probability function (c.d.f.) เขียนได้ว่า

$$F(t) = \int_{-\infty}^{t} (2\pi)^{-1/2} \exp\{-x^2/2\} dx \quad \text{…………..(2.8)}$$

ซึ่งจะต้องทำการประมาณค่าด้วยแบบจำลอง Probit

อีกรูปแบบหนึ่งของการกระจายของค่าคลาดเคลื่อนคือการกระจายแบบ logistic distribution ซึ่งจะทำให้cumulative probability function (c.d.f.) เขียนได้ว่า

$$F(t) = \frac{1}{1 + \exp(-t)} \quad \text{…………..(2.9)}$$

ซึ่งจะต้องทำการประมาณค่าด้วยแบบจำลอง Logit

อย่างไรก็ตาม ไม่ว่าค่าคลาดเคลื่อนจะกระจายตัวแบบใดก็ตาม คุณสมบัติที่ว่า $F(-t) = 1 - F(t)$ ยังคงเหมือนกัน ซึ่งจะทำให้เขียนโอกาสของการเกิด $y_i = 1$ ได้ใหม่ว่า

$$P_i = \Pr[e_i^* > -x_i'\beta] \quad \text{…………..(2.10)}$$

$$= 1 - \Pr[e_i^* \le -x_i'\beta] = 1 - F(-x_i'\beta) \quad \text{…………..(2.11)}$$

ดังนั้น $$P_i = F(x_i'\beta) \quad \text{…………..(2.12)}$$

3.4.11 การประมาณค่าด้วยแบบจำลอง Logit เมื่อมี repeated observations

Judge, et al (1988) แนะนำว่าในการประมาณค่าด้วยแบบจำลอง Logit มีสิ่งที่ต้องสังเกตบางประการก่อนเลือกวิธีการในการประมาณค่าคือ มี repeated observations หรือไม่

Repeated observations คือ การที่บุคคลคนหนึ่งทำการเลือกในเรื่องเดียวกันจำนวนหลาย ๆ รอบ (n_i) เช่น ในคำถามที่ว่า นักท่องเที่ยวจะมาเที่ยวจุดขายใหม่ของจังหวัดลำปางหรือไม่ แล้วมีจุดขายจำนวน n_i แห่งให้เลือกตอบ นักวิจัยก็จะสามารถสังเกตพฤติกรรมของนักท่องเที่ยวคนนี้ได้หลายครั้ง และจะบันทึกค่า p_i ซึ่งก็คือสัดส่วนของการตอบว่ามาเที่ยว หรือ $y_i = 1$ ในจำนวนทั้งหมด n_i ครั้ง

เมื่อมีค่า p_i แล้ว จะสามารถกล่าวได้ว่า p_i เป็นตัวประมาณค่า (estimator) ของ p_i ความเป็นไปได้ที่นักท่องเที่ยวผู้นั้นจะมาเที่ยวสถานที่ท่องเที่ยวที่เป็นจุดขายใหม่ของลำปางโดยรวม ซึ่งเขียนได้ว่า

$$p_i = P_i + e_i = F(x_i'\beta) + e_i \qquad \text{…………..(2.13)}$$

โอกาสที่นักท่องเที่ยวทุกคนที่เป็นประชากรของการศึกษาจะมาเที่ยวจุดขายใหม่ของจังหวัดลำปาง เขียนได้ในแบบจำลอง Logit ดังนี้

$$P_i = F(x_i'\beta) = \frac{1}{1+\exp(-x_i'\beta)} \qquad \text{…………..(2.14)}$$

และเมื่อทำการสุ่มตัวอย่าง จะได้โอกาสดังกล่าวของนักท่องเที่ยวแต่ละคน คือ p_i ซึ่งเขียน odd ratio หรืออัตราส่วนของโอกาสที่จะมาเที่ยวต่อโอกาสที่จะไม่มาได้ว่า $\frac{p_i}{1-p_i}$

ค่า natural logarithm ของ odd ratio จะให้ผลลัพธ์เป็นแบบจำลองเส้นตรง ดังนี้

$$\ln\left(\frac{p_i}{1-p_i}\right) = \ln\left(\frac{P_i}{1-P_i}\right) + \frac{e_i}{P_i(1-P_i)} = x_i'\beta + u_i \qquad \text{…………..(2.15)}$$

ทั้งนี้ odd ratio ของประชากร คือ

$$\ln\left(\frac{P_i}{1-P_i}\right) = x_i'\beta \qquad \text{…………..(2.16)}$$

และเมื่อพิจารณา reverse function ของ โอกาสการเกิด $y_i = 1$ จะได้เป็นแบบจำลองเส้นตรงเช่นกัน ดังนี้

$$v_i = F^{-1}(P_i) + u_i = x_i'\beta + u_i \qquad \text{…………..(2.17)}$$

ซึ่งแบบจำลองเส้นตรงดังกล่าวจะมีค่าความแปรปรวนของค่าคลาดเคลื่อน ดังนี้

$$Var(u_i) = \frac{1}{n_i P_i (1-P_i)} \qquad \text{…………..(2.18)}$$

ในที่นี้หากเขียน reverse function ดังกล่าวในรูปของเมตริกซ์ จะได้ว่า

$$v = X\beta + u \qquad \text{…………..(2.19)}$$

ค่าเฉลี่ยของค่าคลาดเคลื่อนจากแบบจำลองดังกล่าวจะมีค่าเท่ากับศูนย์ $E[u] = 0$ และเมื่อกำหนดให้ $Cov(u) = \Phi$ ซึ่ง Φ เป็นเมตริกซ์แบบ diagonal matrix ซึ่งมีค่าบนเส้นทแยงมุมที่คำนวณได้จากสมการ (4.2-18)

ตัวประมาณค่า Estimated generalized least squares (EGLS) จะหาได้ดังนี้

$$\hat{\hat{\beta}} = \left(X'\hat{\Phi}^{-1}X\right)^{-1} X'\hat{\Phi}^{-1}v \qquad \text{…………..(2.20)}$$

เมื่อ $\hat{\Phi}$ สร้างขึ้นมาจาก p_i

ทั้งนี้ ตัวประมาณค่าแบบ EGLS นี้มีคุณสมบัติดังนี้

$$\hat{\hat{\beta}} \sim N\left[\beta, \left(X'\hat{\Phi}^{-1}X\right)^{-1}\right] \qquad \text{…………..(2.21)}$$

อนึ่ง ในการแปลความหมายค่าที่ได้จากแบบจำลอง Logit สามารถใช้แนวทางของ Marginal effect คือ เมื่อตัวแปร j มีค่าเปลี่ยนไป 1 หน่วยแล้ว จะทำให้โอกาสการเกิด $y_i = 1$ เปลี่ยนแปลงไปเท่าใด สามารถคำนวณได้ ดังนี้

$$\frac{\partial P_i}{\partial x_{ij}} = f\left(x_i'\beta\right) \cdot \beta_j \qquad \text{…………..(2.22)}$$

หรือ เขียนได้ว่า

$$\frac{\partial P_i}{\partial x_{ij}} = \frac{\beta_j \cdot \exp(-x_i'\beta)}{\left[1 + \exp(-x_i'\beta)\right]^2} \qquad \text{…………..(2.23)}$$

3.4.2 การประมาณค่าด้วยแบบจำลอง Logit เมื่อไม่มี repeated observations

ในการประมาณค่าด้วยแบบจำลอง Logit หากไม่มี repeated observations หรือ มีค่าสังเกตพฤติกรรมการเลือกของนักท่องเที่ยวคนหนึ่งเพียงครั้งเดียวหรือจำนวนเพียงไม่กี่ครั้ง เช่น นักท่องเที่ยวตอบว่าจะมาเที่ยวจังหวัดลำปางซ้ำอีกในรอบปีต่อไป ซึ่งเป็นเพียงคำตอบเดียวในเรื่องการเที่ยวซ้ำ หากเป็นเช่นนี้แล้ว จะไม่สามารถใช้การประมาณค่าด้วยวิธี EGLS ได้ แต่ต้องใช้วิธี maximum likelihood แทน

เมื่อพิจารณานักท่องเที่ยวจำนวน T คน จะสร้าง likelihood function ได้ดังนี้ (Judge et al, 1988)

$$L = \prod_{i=1}^{T} f(y_i) = \prod_{i=1}^{T} P_i^{y_i}(1 - P_i)^{(1-y_i)} \qquad \text{…………..(2.24)}$$

$$= \prod_{i=1}^{T} F(x_i'\beta)^{y_i}\left[1 - F(x_i'\beta)\right]^{(1-y_i)} \qquad \text{…………….(2.25)}$$

likelihood function คือ ความเป็นไปได้ที่ตัวประมาณค่าจะให้คำตอบที่ถูกต้องของโอกาสการมาเที่ยวซ้ำของนักท่องเที่ยวทุก ๆ คนโดยรวม

ในที่นี้ $F(\bullet)$ สามารถเป็นได้ทั้ง normal c.d.f. และ logistic c.d.f. และ $y_i = 1$ ถ้านักท่องเที่ยวเลือกที่จะมาเที่ยวซ้ำ และ $y_i = 0$ หากนักท่องเที่ยวจะไม่มาเที่ยวซ้ำ

เมื่อคำนวณค่า natural logarithm ของ likelihood function จะได้ดังนี้

$$\ln L = \sum_{i=1}^{T} y_i \ln F(x_i'\beta) + \sum_{i=1}^{T} (1-y_i)\ln\left[1-F(x_i'\beta)\right] \qquad \text{…………..(2.26)}$$

การใส่ natural logarithm เข้าไปเป็น monotonic transformation ที่ไม่ทำให้การเรียงลำดับของค่าฟังก์ชั่นมีการเปลี่ยนแปลง ดังนั้น การหาจุดสูงสุดของความเป็นไปได้ที่ตัวประมาณค่าจะให้คำตอบที่ถูกต้องของโอกาสการมาเที่ยวซ้ำของนักท่องเที่ยวทุก ๆ คนโดยรวม สามารถกระทำผ่านฟังก์ชั่น log-likelihood ได้ ดังนี้

$$\frac{\partial \ln L}{\partial \beta} = \sum_{i=1}^{T} y_i \frac{f}{F} x_i - \sum_{i=1}^{T} \left(1-y_i\right) \frac{f}{1-F} x_i \qquad \text{…………..(2.27)}$$

ในที่นี้ F คือ cumulative probability function (c.d.f.) และ f คือ คือ density function ของ $x_i'\beta$ ซึ่งเป็นฟังก์ชั่น nonlinear ทำให้ไม่สามารถหาคำตอบได้โดยตรง

วิธีการหาค่าสูงสุดของ nonlinear เช่นนี้กระทำได้โดย Newton-Raphson method ซึ่งทำการคำนวณซ้ำหลาย ๆ รอบ (iterative procedure) ซึ่งจะได้ผลลัพธ์จากการคำนวณค่าในรอบที่ t+1 ดังนี้

$$\hat{\beta}_{t+1} = \hat{\beta}_t - \left[\left. \frac{\partial^2 \ln L}{\partial \beta \quad \partial \beta'} \right|_{\hat{\beta}_t} \right]^{-1} \left[\left. \frac{\partial \ln L}{\partial \beta} \right|_{\hat{\beta}_t} \right] \qquad \text{…………..(2.28)}$$

เมื่อค่าในวงเล็บแรก เป็นเมตริกซ์ขนาด K× K (ซึ่ง K คือ จำนวนของ attributes ทั้งของทางเลือกและของผู้เลือก) ที่ได้จากหา second partials ของฟังก์ชั่น log-likelihood ในรอบที่ t

วิธีการประมาณค่าเช่นนี้จะให้ค่าสูงสุดแบบ global maximum เสมอ ไม่ว่าจะเริ่มต้นด้วยค่า β_0 ใด ๆ และไม่ว่าจะใช้ c.d.f แบบใดก็ตาม นอกจากนั้น วิธีการ maximum likelihood ยังให้ตัวประมาณค่าที่มีคุณสมบัติดังนี้อีกด้วยคือ

1.) consistent

2.) asymptotically efficient

3.) asymptotically normal distributed

ทั้งนี้ asymptotic covariance matrix ที่จะนำไปใช้ในการคำนวณช่วงความเชื่อมั่น (confidence interval) คือ

$$-\left[\frac{\partial^2 \ln L}{\partial\beta \quad \partial\beta'}\right]^{-1}$$

ซึ่งหามาได้จากตัวประมาณค่าจากการคำนวณซ้ำในรอบสุดท้าย

อย่างไรก็ตาม สำหรับแบบจำลอง Logit ซึ่งกำหนดรูปแบบตั้งต้นดังนี้

$$F(t)=\frac{1}{1+e^{-t}} \quad , \quad f(t)=\frac{e^{-t}}{\left(1+e^{-t}\right)^2}$$

$$1-F(t)=\frac{e^{-t}}{1+e^{-t}} \quad , \quad \frac{f(t)}{F(t)}=1-F(t)$$

$$f'(t)=-f(t)\cdot F(t)\cdot(1-e^{-t})$$

จะได้ว่า

$$\frac{\partial^2 \ln L}{\partial\beta \quad \partial\beta'}=-\sum f(x_i'\beta)x_i x_i' \qquad \text{…………..(2.29)}$$

ซึ่งมีประโยชน์ในการนำไปแทนค่าเพื่อคำนวณหา asymptotic covariance matrix ต่อไป

2.2 ทฤษฎีตามคำอธิบายของ William Greene

Greene (2003) อธิบายที่มาของแบบจำลองโลจิตโดยเริ่มจากลักษณะของข้อมูล แล้วโยงไปจนถึงฟังก์ชั่นโลจิต ซึ่งผลลัพธ์สุดท้ายออกมาเป็นฟังก์ชั่นเส้นตรงของ log odds โดยใช้แนวคิดเรื่อง repeated observations ซึ่งสอดคล้องกับการอธิบายของ Judge et al (1988) คำอธิบายของ William Greene มีรายละเอียดดังต่อไปนี้

ข้อมูลที่ใช้ในการวิเคราะห์มีอยู่สองแบบ แบบแรกคือข้อมูลรายบุคคล (individuals) ซึ่งข้อมูลแต่ละรายการประกอบไปด้วยค่าของตัวแปรต้นและตัวแปรตาม อีกแบบหนึ่งคือแบบกลุ่ม (grouped data) ซึ่งส่วนใหญ่จะประกอบไปด้วยข้อมูลที่เป็นจำนวนนับหรือไม่ก็เป็นแบบสัดส่วน ข้อมูลแบบกลุ่มหาได้จากการสังเกตการณ์ตอบสนองของบุคคลจำนวน n_i คน โดยทุก ๆ คนมีตัวแปรต้นที่เหมือนกัน ข้อมูลนี้จะบอกสัดส่วน P_i ของคนที่ตอบว่า $y_{ij} = 1$ ดังนั้นข้อมูลจะประกอบด้วยข้อมูลดังนี้ $[n_i, P_i, x_i], i = 1,..., N$ ยกตัวอย่างเช่น ข้อมูลกลุ่มที่ i ประกอบไปด้วยคนจำนวน n_i คน แต่ละคนได้รับสิ่งเร้า x_i เหมือน ๆ กัน เช่น ให้ทานไอศกรีมรสแรกเหมือน ๆ กัน แล้วพบว่าร้อยละ 80 ของคนกลุ่มนี้ชอบไอศกรีมรสนี้ นั่นคือ $P_i = 0.8$ หรือ คนที่ตอบว่า $y_{ij} = 1$ มีจำนวนเท่ากับ $0.8 \times n_i$

ต่อมาในอีกกลุ่มหนึ่ง ให้ทานไอศกรีมรถที่สอง แล้วบันทึกจำนวนคนที่ชอบไอศกรีมรสนั้น เมื่อมีไอศกรีมให้ทดลองจำนวน j รส ก็จะได้ P_i ออกมาจำนวน j ค่า จำนวนตัวอย่างในการวิเคราะห์ทั้งหมดก็จะมีจำนวน j ตัวอย่าง

เป้าหมายในการวิเคราะห์ข้อมูลที่เป็นกลุ่มนี้คือการหาความสัมพันธ์ระหว่าง P_i และ x_i คือสัดส่วนของการตอบสนองว่าอร่อยกับรสชาติไอศกรีมรสต่าง ๆ (ส่วนผสมของไอศกรีมที่ต่างกัน คือ x_i) โดยสามารถใช้ regression หรือ maximum likelihood ก็ได้

สัดส่วน P_i จะเป็นตัวแทนของความรู้สึกของประชากร โดยที่โอกาสที่คนทั่วไปจะชอบไอศกรีมรสนี้จะมีค่าเท่ากับ $\pi_i = F(x'\beta)$ ถ้าหากเราคิดว่าปัญหานี้เป็นกรณีอย่างง่ายของการสุ่มตัวอย่างแบบเบอร์นูลี (การสุ่มตัวอย่างแบบไบโนเมียลโดยไม่มีการใส่กลับคืน) เราจะสร้างสมการทางสถิติพื้นฐานได้ว่า

$$P_i = F(x'\beta) + \varepsilon_i = \pi_i + \varepsilon_i$$

หรือ

$$F(x'\beta) = \pi_i = P_i - \varepsilon_i$$

ซึ่ง $E[\varepsilon_i]=0$ และ $Var[\varepsilon_i]=\dfrac{\pi_i(1-\pi_i)}{n_i}$

รูปแบบของ heteroscedasticity regression แบบนี้สามารถประมาณค่าได้ด้วย nonlinear weighted least squares regression แต่มีวิธีการที่ง่ายกว่า ทั้งนี้เนื่องจากฟังก์ชั่น $\pi_i = F(x'\beta)$ มีคุณสมบัติเป็น strictly monotonic คือ เป็นฟังก์ชั่นที่เพิ่มขึ้นอย่างเดียวไม่มีลดหรือคงที่ (เพราะโอกาสสะสมหรือ cumulative density function ย่อมเพิ่มขึ้นเรื่อย ๆ ไม่มีลด) ดังนั้นมันจึงมีค่า inverse อย่างแน่นอน เมื่อเป็นเช่นนั้นสามารถใช้ Taylor series approximation ได้รอบ ๆ จุด $\varepsilon_i = 0$ หรือจุด $P_i = \pi_i$ นั่นเอง

$$\frac{1}{F(P_i)} = \frac{1}{F(\pi_i+\varepsilon_i)} \approx \frac{1}{F(\pi_i)} + \left[\frac{d\left(\frac{1}{F(\pi_i)}\right)}{d\pi_i}\right](P_i - \pi_i)$$

แต่ $\dfrac{1}{F(\pi_i)} = x'\beta$ และ

$$\frac{d\left(\frac{1}{F(\pi_i)}\right)}{d\pi_i} = \frac{1}{F'\left[\frac{1}{F(\pi_i)}\right]} = \frac{1}{f(\pi_i)}$$

ดังนั้น

$$\frac{1}{F(P_i)} \approx x_i{}'\beta + \frac{1}{f(\pi_i)}\varepsilon_i$$

สมการนี้ก่อให้เกิด heteroscedasticity linear regression

$$\frac{1}{F(P_i)} = z_i = x_i{}'\beta + u_i$$

เมื่อ

$$E[u_i|x_i] = 0 \quad \text{และ} \quad Var[u_i|x_i] = \frac{F(\pi_i)[1-F(\pi_i)]}{n_i[f(\pi_i)]^2}$$

ฟังก์ชั่น inverse คือ $\frac{1}{F(P_i)}$ สำหรับ logistic model สามารถหาได้โดยง่าย หากว่า π_i มีการกระจายแบบ logistics

$$\pi_i = \frac{e^{x'\beta}}{1+e^{x'\beta}} = \frac{1}{1+e^{-x'\beta}}$$

สาเหตุที่เราต้องหาฟังก์ชั่น inverse นี้ก็เพราะว่าเราจะหาค่าพารามิเตอร์ β ที่ทำให้ค่าของสมการ log-likelihood มีค่ามากที่สุดได้เมื่อทราบฟังก์ชั่น inverse นี้ ดังนี้

หากฟังก์ชั่น log-likelihood มีหน้าตาดังสมการต่อไปนี้

$$\ln L = \sum_{i=1}^{N} n_i \left\{ P_i \ln F\left(x_i{}'\beta\right) + (1-P_i)\ln\left[1-F\left(x_i{}'\beta\right)\right]\right\}$$

แล้วค่าพารามิเตอร์ β ที่ทำให้ค่าของสมการ log-likelihood มีค่ามากที่สุดสามารถหาได้จากสมการต่อไปนี้

$$\frac{\partial \ln L}{\partial \beta} = \sum_{i=1}^{N} n_i \left[P_i \frac{f\left(x_i{}'\beta\right)}{F\left(x_i{}'\beta\right)} - (1-P_i)\frac{f\left(x_i{}'\beta\right)}{1-F\left(x_i{}'\beta\right)} \right] x_i = 0$$

สังเกตว่า $\frac{1}{F\left(x_i{}'\beta\right)}$ ก็คือ $\frac{1}{\pi_i}$ นั่นเอง

และเมื่อ $$\pi_i = \frac{e^{x'\beta}}{1+e^{x'\beta}} = \frac{1}{1+e^{-x'\beta}}$$

ดังนั้นจึงหาส่วนกลับของฟังก์ชั่น π_i ได้ ดังนี้

$$\frac{1}{F\left(x_i'\beta\right)} = \frac{1}{\pi_i} = 1 + e^{-x'\beta}$$

แล้ว

$$\ln\left(\frac{\pi_i}{1-\pi_i}\right) = x_i'\beta$$

ฟังก์ชั่นนี้เรียกว่า **logit** ของ π_i ซึ่งเป็นที่มาของชื่อแบบจำลองโลจิต หาก π_i มีการกระจายแบบปกติ (normal distribution) แล้วฟังก์ชั่น inverse คือ $\frac{1}{\Phi(\pi_i)}$ จะเรียกใหม่ว่า normit ของ π_i

เอกสารอ้างอิง

Greene, William H. 2003. **Econometric Analysis**. 5th ed. New Jersey: Prentice Hall.

Judge, George, et al. 1988. **Introduction to the Theory and Practice of Econometrics**. 2nd ed. New York: John Wiley and Sons

---

ขอความกรุณาท่านผู้อ่านที่อ้างอิงเนื้อหาจากเอกสารฉบับนี้ ช่วยเขียนในบรรณานุกรมของท่านดังนี้ ขอบคุณมากครับ

คมสัน สุริยะ. 2552. **แบบจำลองโลจิต: ทฤษฎีและการประยุกต์ใช้ในการวิจัยทางเศรษฐศาสตร์**. เชียงใหม่: ศูนย์การวิเคราะห์เชิงปริมาณ คณะเศรษฐศาสตร์ มหาวิทยาลัยเชียงใหม่. [online] http://www.tourismlogistics.com