บทที่ 2
แนวคิดและทฤษฎี

การศึกษาการคัดเลือกตัวแบบเป็นการพิจารณาคัดเลือกตัวแบบที่ให้ค่าพยากรณ์ที่ถูกต้องและแม่นยำมากที่สุด ในการวิเคราะห์การถดถอยสามารถจำแนกตัวแบบออกได้เป็น ตัวแบบติดกลุ่ม(Nested Model) และตัวแบบไม่ติดกลุ่ม(Non-nested Model) ในปัจจุบันมีแนวคิดเบื้องต้นในการคัดเลือกตัวแบบ 2 แนวคิด ดังนี้

1. แนวคิดของการทดสอบด้วยอัตราส่วนภาวะน่าจะเป็น(Likelihood Ratio Test : LRT)
2. แนวคิดของวงศ์เกณฑ์ข้อสนเทศ(Information Criterion Family)

เนื่องจากการคัดเลือกตัวแบบโดยใช้แนวคิดของการทดสอบด้วยอัตราส่วนภาวะน่าจะเป็นมีข้อจำกัดว่า ตัวแบบที่นำมาศึกษาต้องเป็นตัวแบบติดกลุ่ม ส่วนการคัดเลือกตัวแบบโดยใช้แนวคิดของวงศ์เกณฑ์ข้อสนเทศสามารถใช้ศึกษาได้ทั้งในตัวแบบติดกลุ่มและไม่ติดกลุ่ม ในการวิจัยครั้งนี้จะศึกษาเกณฑ์การคัดเลือกตัวแบบและเปรียบเทียบความถูกต้องของการคัดเลือกตัวแบบในการวิเคราะห์การถดถอยของตัวแบบการถดถอยเชิงเส้นแบบไม่ติดกลุ่ม ดังนั้นเกณฑ์การคัดเลือกตัวแบบในการศึกษาครั้งนี้ประกอบด้วย เกณฑ์การคัดเลือกตัวแบบโดยข้อสนเทศน์ของอาไคเคะ(Akaike 's Information Criterion : AIC) โดยอาไคเคะ(Akaike) นำเสนอไว้ในปี ค.ศ. 1973 เป็นเกณฑ์ที่ใช้ในการคัดเลือกตัวแบบ โดยใช้แนวคิดของ คูลส์แบล็ค-ไลท์เบอร์(Kullback Leibler :1951)) มาใช้ในการพิจารณาหาตัวแบบที่เหมาะสมเพื่อนำไปใช้ในการพยากรณ์โดยตัวแบบที่ให้ค่า AIC ต่ำสุด จะเป็นตัวแบบที่ดีที่สุด

เกณฑ์การคัดเลือกตัวแบบโดยข้อสนเทศของเบส์(Bayesian Information Criterion : BIC) ซึ่งชวาร์ซ(Schwarz) นำเสนอไว้ในปี ค.ศ.1978 โดยการสมมติว่าความน่าจะเป็นก่อน (Prior Probability)ของทุกตัวแบบเหมือนกัน ในการพิจารณาการคัดเลือกตัวแบบที่เหมาะสมจะพิจารณาจากความน่าจะเป็นภายหลัง(Posterior Probability) โดยการพิจารณาจากค่าสูงสุดของความน่าจะเป็นภายหลัง หรือพิจารณาจากตัวแบบที่ให้ค่า BIC ต่ำสุดจะเป็นตัวแบบที่ดีที่สุด

2.1 การวิเคราะห์การถดถอยเชิงเส้น (Linear Regression Analysis)

ตัวแบบที่ใช้สำหรับการวิจัยครั้งนี้เป็นตัวแบบการถดถอยเชิงเส้น นั่นคือ เป็นเชิงเส้นในพารามิเตอร์และในตัวแปรอิสระ โดยมีรูปแบบทั่วไปดังนี้

$$\underset{\sim}{Y} = X\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$$

เมื่อ	$\underset{\sim}{Y}$	แทน	เวกเตอร์ของตัวแปรตามขนาด $(n \times 1)$
	X	แทน	เมทริกซ์ของตัวแปรอิสระที่ควบคุมให้คงที่ขนาด $(n \times p)$
	$\underset{\sim}{\beta}$	แทน	เวกเตอร์ของพารามิเตอร์ของสัมประสิทธิ์การถดถอย ขนาด $(p \times 1)$
	$\underset{\sim}{\varepsilon}$	แทน	เวกเตอร์ของความคลาดเคลื่อนขนาด $(n \times 1)$
	n	แทน	ขนาดตัวอย่าง
	p	แทน	จำนวนพารามิเตอร์ของสัมประสิทธิ์การถดถอย

โดยมีข้อสมมติของรูปแบบดังนี้ ความคลาดเคลื่อนเป็นตัวแปรสุ่มที่มีการแจกแจงปกติ $N(0, \sigma^2)$ และเป็นอิสระซึ่งกันและกัน

สำหรับการวิจัยครั้งนี้ได้ศึกษาถึงตัวแบบการถดถอยเชิงเส้นแบบไม่ติดกลุ่ม โดยมีจำนวนตัวแปรอิสระเริ่มต้นเป็น 2 3 และ 4 ตัว ซึ่งสามารถพิจารณาเป็นตัวแบบการถดถอยเชิงเส้นแบบไม่ติดกลุ่มได้ตามขอบเขตของการวิจัยดังได้กล่าวไว้แล้วในบทที่ 1

เนื่องจากตัวแบบที่ศึกษาเป็นตัวแบบการถดถอยเชิงเส้นและค่าความคลาดเคลื่อนมีการแจกแจงแบบปกติ ดังนั้นในการประมาณค่าพารามิเตอร์ของสัมประสิทธิ์การถดถอยด้วยวิธีการภาวะน่าจะเป็นสูงสุด(Maximum Likelihood Estimation) และวิธีการกำลังสองน้อยที่สุด (Least Square Estimation) จะได้ตัวประมาณตัวเดียวกัน ซึ่งจะได้ว่า

ตัวประมาณค่าภาวะน่าจะเป็นสูงสุด(Maximum Likelihood Estimator : $\hat{\beta}_{\sim M}$) คือ

$$\hat{\beta}_{\sim M} = \left(X'X\right)^{-1} X' \underset{\sim}{Y}$$

ซึ่งเป็นตัวประมาณที่ใช้สำหรับ เกณฑ์การคัดเลือกตัวแบบโดยข้อสนเทศของอาไคเคะ(AIC) และเกณฑ์การคัดเลือกตัวแบบโดยข้อสนเทศของเบส์(BIC) สำหรับการประมาณด้วยวิธีการภาวะน่าจะเป็นสูงสุด จะกล่าวรายละเอียดในหัวข้อต่อไป

2.2 การประมาณด้วยภาวะน่าจะเป็นสูงสุด (Maximum Likelihood Estimation)

ตัวแบบซึ่งมีรูปแบบดังสมการ

$$\underset{\sim}{Y} = X\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$$

มีฟังก์ชันภาวะน่าจะเป็นดังนี้

$$\begin{aligned} ML = L\left(\beta_{\sim M}, \sigma^2 \mid \underset{\sim}{Y}\right) &= f(y_1 \mid \beta_1, \sigma^2) \times f(y_2 \mid \beta_2, \sigma^2) \times \cdots \times f(y_n \mid \beta_n, \sigma^2) \\ &= (2\pi\sigma^2)^{\frac{-n}{2}} \exp\left[\frac{-1}{2\sigma^2}\left(\underset{\sim}{Y} - X\beta_{\sim M}\right)'\left(\underset{\sim}{Y} - X\beta_{\sim M}\right)\right] \end{aligned}$$

จะได้

$$\begin{aligned} \log L\left(\beta_{\sim M}, \sigma^2 \mid \underset{\sim}{Y}\right) &= -\frac{n}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\left(\underset{\sim}{Y} - X\beta_{\sim M}\right)'\left(\underset{\sim}{Y} - X\beta_{\sim M}\right) \\ &= -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\left(\underset{\sim}{Y} - X\beta_{\sim M}\right)'\left(\underset{\sim}{Y} - X\beta_{\sim M}\right) \end{aligned}$$

จาก $\dfrac{\partial}{\partial \beta_{\sim M}} \log L\left(\beta_{\sim M}, \sigma^2 \mid \underset{\sim}{Y}\right) = 0$

และ $\dfrac{\partial}{\partial \sigma^2} \log L\left(\beta_{\sim M}, \sigma^2 \mid \underset{\sim}{Y}\right) = 0$

ดังนั้น $\hat{\beta}_{\sim M} = \left(X'X\right)^{-1} X' \underset{\sim}{Y}$

$$\hat{\sigma}^2 = \frac{1}{n}\left(\underset{\sim}{Y} - X\hat{\beta}_{\sim M}\right)'\left(\underset{\sim}{Y} - X\hat{\beta}_{\sim M}\right)$$

เพราะฉะนั้น

$$\begin{aligned} \log L\left(\beta_{\sim M}, \sigma^2 \mid \underset{\sim}{Y}\right) &= -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\hat{\sigma}^2) - \frac{n}{2} \\ &= -\frac{n}{2}(\log(2\pi) + 1) - \frac{n}{2}\log(\hat{\sigma}^2) \end{aligned}$$

2.3 การพิจารณาความเหมาะสมของตัวแบบ

ค่าสถิติที่ใช้วัดว่าตัวแบบใดจะมีความถูกต้องและเหมาะสมที่สุด สำหรับการพยากรณ์ จะพิจารณาจาก ค่าคลาดเคลื่อนกำลังสองเฉลี่ย(Mean Square Error) ซึ่งมีรูปแบบดังนี้

$$\mathrm{MSE} = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n-p}$$

เมื่อ y_i แทน ค่าสังเกตที่ i

$\hat{y}_i$ แทน ค่าพยากรณ์ที่ i

p แทน จำนวนพารามิเตอร์

n แทน ขนาดตัวอย่าง

ตัวแบบที่มีค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำสุดเป็นตัวแบบที่เหมาะสมที่สุด

2.4 เกณฑ์ที่นำมาใช้ในการคัดเลือกตัวแบบ

เกณฑ์ในการคัดเลือกตัวแบบการถดถอยเชิงเส้นแบบไม่ติดกลุ่มที่ใช้ในการศึกษาครั้งนี้ประกอบด้วย 2 เกณฑ์ ดังนี้

2.4.1 เกณฑ์การคัดเลือกตัวแบบโดยข้อสนเทศของอาไคเคะ (Akaike 's Information Criterion : AIC)

2.4.2 เกณฑ์การคัดเลือกตัวแบบโดยข้อสนเทศของเบส์ (Bayesian Information Criterion : BIC)

2.4.1 เกณฑ์การคัดเลือกตัวแบบโดยข้อสนเทศของอาไคเคะ (Akaike 's Information Criterion : AIC)

ในปี ค.ศ. 1973 อาไคเคะได้เสนอเกณฑ์ในการคัดเลือกตัวแบบ นั่นคือ เกณฑ์การคัดเลือกตัวแบบโดยข้อสนเทศของอาไคเคะ(AIC) **ซึ่งเป็นเกณฑ์ที่พิจารณาจากการประมาณ** ความคลาดเคลื่อนรวมเข้ากับข้อสนเทศ (Information) ของค่าสังเกต และใช้แนวคิดค่าต่ำสุดของ

คูลแบ็ค-ไลท์เบอร์ (Kullback-Leiber(1951)) เพื่อนำมาใช้ในปรับค่าประมาณของการพยากรณ์ให้มีความแม่นยำมากขึ้น

เกณฑ์การคัดเลือกตัวแบบโดยข้อสนเทศของอาไคเคะ (AIC) คือ

$$\text{AIC} = -2\log\text{ML} + 2p \tag{2.1}$$

เมื่อ ML คือ ฟังก์ชันภาวะน่าจะเป็นสูงสุด(Maximum Likelihood Function) ของแต่ละตัวแบบ

p คือ จำนวนพารามิเตอร์

ดังนั้น AIC ในสมการ (2.1) สามารถเขียนได้ใหม่ดังนี้

$$\begin{aligned}
\text{AIC} &= -2\log\text{ML}+2p \\
&= -2\left[-\frac{n}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\left(\underset{\sim}{Y} - X\underset{\sim M}{\beta}\right)'\left(\underset{\sim}{Y} - X\underset{\sim M}{\beta}\right)\right] + 2p \\
&= -2\left[-\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\left(\underset{\sim}{Y} - X\underset{\sim M}{\beta}\right)'\left(\underset{\sim}{Y} - X\underset{\sim M}{\beta}\right)\right] + 2p \\
&= -2\left[-\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \frac{n}{2}\right] + 2p \\
&= -2\left[-\frac{n}{2}(\log(2\pi)+1) + \frac{n}{2}\log(\hat{\sigma}^2)\right] + 2p \\
&= n(\log(2\pi)+1) + n\log(\hat{\sigma}^2) + 2p
\end{aligned}$$

เนื่องจาก เทอมของ $n(\log(2\pi)+1)$ เป็นค่าคงที่สำหรับทุกตัวแบบที่นำมาพิจารณา ดังนั้นจึงเขียน AIC ได้ดังนี้

$$\text{AIC} = n\log(\hat{\sigma}^2) + 2p \tag{2.2}$$

2.4.1.1 วิธีการคัดเลือกตัวแบบของเกณฑ์การคัดเลือกตัวแบบโดยข้อสนเทศของอาไคเคะ (Akaike 's Information Criterion : AIC)

1. กำหนดตัวแบบการถดถอยเชิงเส้นแบบไม่ติดกลุ่ม เป็นตามขอบเขตของการวิจัยดังที่กำหนดไว้ข้างต้น

2. คำนวณค่า AIC ของแต่ละตัวแบบที่พิจารณา จากสูตรในสมการที่(2.2)

3. เลือกตัวแบบที่มีค่า AIC ต่ำสุด

4. พิจารณาค่าคลาดเคลื่อนกำลังสองเฉลี่ย ของตัวแบบที่ได้จากเกณฑ์การคัดเลือกตัวแบบโดยข้อสนเทศของอาไคเคะ (AIC) ดังนี้

4.1 ตัวแบบที่ได้มีค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำสุด นั่นแสดงว่าเกณฑ์การคัดเลือกตัวแบบโดยข้อสนเทศของอาไคเคะ (AIC) เลือกตัวแบบได้ถูกต้อง

4.2 ตัวแบบที่ได้ไม่ได้มีค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำสุด นั่นแสดงว่าเกณฑ์การคัดเลือกตัวแบบโดยข้อสนเทศของอาไคเคะ (AIC) เลือกตัวแบบผิด

2.4.2 เกณฑ์การคัดเลือกตัวแบบโดยข้อสนเทศของเบส์ (Bayesian Information Criterion : BIC)

ในปี ค.ศ. 1978 ชวาร์ซได้เสนอเกณฑ์การคัดเลือกตัวแบบ ซึ่งใช้เป็นเกณฑ์ในการคัดเลือกตัวแบบทางสถิติที่ใช้กันอย่างกว้างขวาง นั่นคือ เกณฑ์การคัดเลือกตัวแบบโดยข้อสนเทศของชวาร์ซ (Schwarz Information Criterion : (SIC,BIC,SBC))

เกณฑ์การคัดเลือกตัวแบบโดยข้อสนเทศของชวาร์ซ สามารถนำไปพัฒนาโดยใช้แนวคิดของเบส์(Bayesian Approach)ซึ่งเรียกว่า เกณฑ์การคัดเลือกตัวแบบโดยข้อสนเทศของเบส์ (BIC) ได้ดังนี้

กำหนดให้

$p\left(X \mid M_k, \underset{\sim k}{\beta}\right)$ แทน ความน่าจะเป็นของ X เมื่อกำหนด M_k และ $\underset{\sim k}{\beta}$

$p(M_k)$ แทน ความน่าจะเป็นก่อนของตัวแบบ $M_1,\ldots,M_L$ เมื่อ $1 \le k \le L$

$p\left(\underset{\sim}{\beta} \mid M_k\right)$ แทน ความน่าจะเป็นก่อนบน $\underset{\sim}{\beta}$ เมื่อกำหนด M_k

โดยการประยุกต์ใช้ทฤษฎีของเบส์(Bayes's Theorem) สามารถเขียนความหนาแน่นร่วมภายหลังของ M_k และ $\underset{\sim}{\beta}_k$ ได้ดังนี้

$$p\left(M_k, \underset{\sim}{\beta}_k \mid X\right) = \frac{p(M_k)\,p\left(\underset{\sim}{\beta}_k \mid M_k\right)p\left(\underset{\sim}{X} \mid M_k, \underset{\sim}{\beta}_k\right)}{p(X)} \tag{2.3}$$

เมื่อ $p(X)$ แทน การแจกแจงส่วนริม(Marginal Distribution) ของ X

ในการคัดเลือกตัวแบบแบบเบส์ จะเลือกตัวแบบ M_k ซึ่งมีความน่าจะเป็นภายหลัง มากที่สุด ซึ่งการแจกแจงภายหลัง(Posterior Distribution) ของตัวแบบ M_k คือ

$$p(M_k \mid X) = \int p\left(M_k, \underset{\sim}{\beta}_k \mid X\right) d\underset{\sim}{\beta}_k \tag{2.4}$$

เนื่องจาก $p\left(M_k, \underset{\sim}{\beta}_k \mid X\right) = L\left(\underset{\sim}{\beta}_k \mid X\right)$ ดังนั้นจาก (2.3) และ (2.4) สามารถเขียนใหม่ได้ดังนี้

$$p(M_k \mid X) = p(\underset{\sim}{X})^{-1} p(M_k) \int L\left(\underset{\sim}{\beta}_k \mid X\right) p\left(\underset{\sim}{\beta}_k \mid M_k\right) d\underset{\sim}{\beta}_k \tag{2.5}$$

เนื่องจากสามารถหาค่าต่ำสุดของ $-2\log p(M_k \mid X)$ แทนการหาค่าสูงสุดของ $p(M_k \mid X)$ ได้ ดังนั้นจาก (2.5) จึงสามารถเขียนใหม่ได้ดังนี้

$$\begin{aligned} -2\log p(M_k \mid X) &= 2\log[p(X)] - 2\log[p(M_k)] \\ &\quad - 2\log\left[\int L\left(\underset{\sim}{\beta}_k \mid X\right) p\left(\underset{\sim}{\beta}_k \mid M_k\right) d\underset{\sim}{\beta}_k\right] \end{aligned} \tag{2.6}$$

$$-2\log p(M_k \mid X) - 2\log[p(X)] \approx -2\log L\left(\underset{\sim}{\beta}_k \mid X\right) + D_k \log(n)$$

ดังนั้นเกณฑ์การคัดเลือกตัวแบบโดยข้อสนเทศของเบส์ (BIC) สามารถเขียนได้ดังนี้

$$BIC = -2\log ML + \log(n)p \tag{2.7}$$

เมื่อ ML คือ ฟังก์ชันภาวะน่าจะเป็นสูงสุด(Maximum Likelihood Function) ของแต่ละตัวแบบ

p คือ จำนวนพารามิเตอร์

n คือ ขนาดตัวอย่าง

ดังนั้น BIC ในสมการ (2.7) สามารถเขียนได้ใหม่ดังนี้

$$
\begin{aligned}
BIC &= -2\log ML+\log(n)p \\
&= -2\left[\frac{-n}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\left(\underset{\sim}{Y} - X\underset{\sim M}{\beta}\right)'\left(\underset{\sim}{Y} - X\underset{\sim M}{\beta}\right)\right] + \log(n)p \\
&= -2\left[\frac{-n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\left(\underset{\sim}{Y} - X\underset{\sim M}{\beta}\right)'\left(\underset{\sim}{Y} - X\underset{\sim M}{\beta}\right)\right] + \log(n)p \\
&= -2\left[\frac{-n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \frac{n}{2}\right] + \log(n)p \\
&= -2\left[\frac{-n}{2}(\log(2\pi)+1) + \frac{n}{2}\log(\hat{\sigma}^2)\right] + \log(n)p \\
&= n(\log(2\pi)+1) + n\log(\hat{\sigma}^2) + \log(n)p
\end{aligned}
$$

เนื่องจาก พจน์ $n(\log(2\pi)+1)$ เป็นค่าคงที่สำหรับทุกตัวแบบที่นำมาพิจารณา ดังนั้น จึงเขียน BIC ได้ดังนี้

$$BIC = n\log(\hat{\sigma}^2) + \log(n)p \tag{2.8}$$

2.4.2.1 วิธีการคัดเลือกตัวแบบของเกณฑ์การคัดเลือกตัวแบบโดยข้อสนเทศของเบส์ (Bayesian Information Criterion : BIC)

1. กำหนดตัวแบบการถดถอยเชิงเส้นแบบไม่ติดกลุ่มตามขอบเขตของการวิจัยดังที่กำหนดไว้ข้างต้น

2. คำนวณค่า BIC ของแต่ละตัวแบบที่พิจารณา จากสูตรในสมการที่(2.8)

3. เลือกตัวแบบที่มีค่า BIC ต่ำสุด

4. พิจารณาค่าคลาดเคลื่อนกำลังสองเฉลี่ยของตัวแบบที่ได้จากเกณฑ์การคัดเลือกตัวแบบโดยข้อสนเทศของเบส์(BIC) ดังนี้

4.1 ตัวแบบที่ได้มีค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำสุด นั่นแสดงว่าเกณฑ์การคัดเลือกตัวแบบโดยข้อสนเทศของเบส์(BIC) เลือกตัวแบบได้ถูกต้อง

4.2 ตัวแบบที่ได้ไม่ได้มีค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำสุด นั่นแสดงว่าเกณฑ์การคัดเลือกตัวแบบโดยข้อสนเทศของเบส์(BIC) เลือกตัวแบบผิด

2.5 การทดสอบสมมติฐานเกี่ยวกับค่าสัดส่วน(Proportion Test)

ในการวิจัยครั้งนี้จะพิจารณาความถูกต้องในการคัดเลือกตัวแบบของเกณฑ์การคัดเลือกตัวแบบโดยข้อสนเทศของอาไคเคะ (AIC) และเกณฑ์การคัดเลือกตัวแบบโดยข้อสนเทศของเบส์ (BIC) โดยใช้วิธีการทดสอบสมมติฐานค่าสัดส่วนซึ่งแบ่งการทดสอบสมมติฐานเป็นดังนี้

2.5.1 การทดสอบสมมติฐานเปรียบเทียบระหว่างค่าสัดส่วนของการคัดเลือกตัวแบบผิดจากเกณฑ์การคัดเลือกตัวแบบ กับค่าสัดส่วนของการคัดเลือกตัวแบบผิดที่กำหนดขึ้น

สมมติฐานที่ใช้ในการทดสอบคือ

$H_0 : P_i \geq P_0$

$H_1 : P_i < P_0$

ตัวสถิติทดสอบคือ

$$z = \frac{\hat{P}_i - P_0}{\sqrt{\frac{P_0(1-P_0)}{n*}}}$$

เมื่อ $i = A, B$

โดยที่ $\hat{P}_A$ คือ ค่าสัดส่วนของการเลือกตัวแบบผิด $\left(\frac{n_A}{n*}\right)$ ที่คำนวณได้จากเกณฑ์การคัดเลือกตัวแบบโดยข้อสนเทศของอาไคเคะ(AIC)

$\hat{P}_B$ คือ ค่าสัดส่วนของการเลือกตัวแบบผิด$\left(\frac{n_B}{n^*}\right)$ที่คำนวณได้จากเกณฑ์การคัดเลือกตัวแบบโดยข้อสนเทศของเบส์ (BIC)

n_A คือ จำนวนครั้งทั้งหมดที่เลือกตัวแบบผิดโดยใช้เกณฑ์การคัดเลือกตัวแบบโดยข้อสนเทศของอาไคเคะ(AIC)

n_B คือ จำนวนครั้งทั้งหมดที่เลือกตัวแบบผิดโดยใช้เกณฑ์การคัดเลือกตัวแบบโดยข้อสนเทศของเบส์ (BIC)

P_0 คือ ค่าสัดส่วนของการเลือกผิดที่กำหนดขึ้นเท่ากับ 1% 5% และ 10%

n^* คือ จำนวนครั้งที่ทำการทดลอง

α คือ ระดับนัยสำคัญทางสถิติที่กำหนดคือ 0.01 และ 0.05

เกณฑ์ในการตัดสินใจ

ถ้า $z_{cal} < -z_{\alpha}$ จะปฏิเสธ H_0 นั่นคือ ค่าสัดส่วนของการคัดเลือกตัวแบบผิดที่ได้จากเกณฑ์การคัดเลือกแบบมีค่าต่ำกว่าค่าสัดส่วนของการคัดเลือกตัวแบบผิดที่กำหนดขึ้น

หรือพิจารณาจากค่า p-value ถ้า ค่า p-value(one-tailed) < α และ $z_{cal} < 0$ จะปฏิเสธ H_0

2.5.2 การทดสอบสมมติฐานเปรียบเทียบระหว่างค่าสัดส่วนของการคัดเลือกตัวแบบผิดจากเกณฑ์การคัดเลือกตัวแบบโดยข้อสนเทศของอาไคเคะ (Akaike's Information Criterion : AIC) กับเกณฑ์การคัดเลือกตัวแบบโดยข้อสนเทศของเบส์ (Bayesian Information Criterion : BIC)

สมมติฐานที่ใช้ในการทดสอบคือ

$H_0 : P_B \geq P_A$

$H_1 : P_B < P_A$

ตัวสถิติทดสอบคือ

$$z = \frac{\hat{P}_B - \hat{P}_A}{\sqrt{\hat{p}\hat{q}\left(\frac{1}{n^*_A} + \frac{1}{n^*_B}\right)}} \qquad \text{เมื่อ } \hat{p} = \frac{n_A + n_B}{n^*_A + n^*_B}$$

เกณฑ์ในการตัดสินใจ

ถ้า $z_{cal} < -z_{\alpha}$ จะปฏิเสธ H_0 นั่นคือ เกณฑ์ในการคัดเลือกตัวแบบโดยข้อสนเทศของเบส์(BIC) มีค่าสัดส่วนของการคัดเลือกตัวแบบผิดต่ำกว่าเกณฑ์การคัดเลือกตัวแบบโดยข้อสนเทศของอาไคเคะ(AIC) ซึ่งแสดงว่าเกณฑ์การคัดเลือกตัวแบบโดยข้อสนเทศของเบส์(BIC) มีความถูกต้องในการคัดเลือกตัวแบบมากกว่าเกณฑ์การคัดเลือกตัวแบบโดยข้อสนเทศของอาไคเคะ(AIC)

หรือพิจารณาจากค่า p-value ถ้า ค่า p-value(one-tailed) < α และ $z_{cal} < 0$ จะปฏิเสธ H_0