การเปรียบเทียบวิธีการประมาณค่าแบบช่วงสำหรับสัมประสิทธิ์การถดถอยของตัวแบบโลจิสติกเมื่อใช้การประมาณค่าพารามิเตอร์ด้วยวิธีภาวะน่าจะเป็นสูงสุดและวิธีฟังก์ชันจำแนกประเภท

A Comparison of Interval Estimation Methods for Logistic Regression Model When Using Estimation of Parameters by Maximum Likelihood Method and Discriminant Function Method

ขวัญชนก หงษ์ชูเกียรติ*, อำไพ ทองธีรภาพ และมีนา ปทุมสูตร

ภาควิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์

แขวงลาดยาว เขตจตุจักร กรุงเทพมหานคร 10900

Kwanchanok Hongchukiet*, Ampai Thongteeraparp and Mena Patummasut

Department of Statistics, Faculty of Science, Kasetsart University,

Ladyao, Chatuchak, Bangkok 10900

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของวิธีการประมาณค่าแบบช่วงสำหรับสัมประสิทธิ์การถดถอยของตัวแบบโลจิสติก 3 วิธี คือ วิธีคลาสสิค วิธีปริมาณหมุน และวิธีแบบเบส์ เมื่อใช้วิธีการประมาณค่าพารามิเตอร์ด้วยวิธีภาวะน่าจะเป็นสูงสุด และวิธีฟังก์ชันจำแนกประเภท เกณฑ์ในการเปรียบเทียบ คือ ค่าสัมประสิทธิ์ความเชื่อมั่นและความกว้างเฉลี่ยของช่วงความเชื่อมั่น กำหนดจำนวนตัวแปรอิสระเท่ากับ 3, 5 และ 7 ตัวแปร โดยตัวแปรอิสระมีการแจกแจงแบบปกติมาตรฐาน ส่วนตัวแปรตามมีการแจกแจงแบบเบอร์นูลี ค่าพารามิเตอร์เริ่มต้น $(\beta_j \ ; \ j = 0,1,...,p)$ เท่ากับ 1 ขนาดตัวอย่าง (n) เท่ากับ 20, 30, 50 และ 100 กำหนดระดับสัมประสิทธิ์ความเชื่อมั่นเท่ากับ 0.90, 0.95 และ 0.99 จำลองข้อมูลภายใต้สถานการณ์ต่าง ๆ ด้วยเทคนิคมอนติคาร์โล ทำซ้ำในแต่ละสถานการณ์จำนวน 2,000 รอบ ผลการศึกษาพบว่าการประมาณค่าทั้ง 3 วิธี ให้ค่าสัมประสิทธิ์ความเชื่อมั่นไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนดในทุกสถานการณ์ และการประมาณค่าแบบช่วงด้วยวิธีปริมาณหมุนให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นแคบที่สุด ส่วนวิธีคลาสสิคและวิธีแบบเบส์ให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นใกล้เคียงกันในทุกสถานการณ์ แต่เมื่อขนาดตัวอย่างเพิ่มขึ้นการประมาณค่าทั้ง 3 วิธี ที่ได้จากการประมาณค่าพารามิเตอร์ด้วยวิธีฟังก์ชันจำแนกประเภทจะให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นแคบกว่าวิธีภาวะน่าจะเป็นสูงสุดในทุกสถานการณ์

***ผู้รับผิดชอบบทความ :** kwanchanok.h@hotmail.com

คำสำคัญ : การถดถอยโลจิสติก; การประมาณแบบช่วง; สัมประสิทธิ์การถดถอย; วิธีภาวะน่าจะเป็นสูงสุด; วิธีฟังก์ชันจำแนกประเภท

Abstract

The objective of this research is to compare three intervals estimation methods for the coefficient of logistic regression. The estimation methods are classical method, pivotal quantity method and Bayesian method which the coefficients are estimated by the maximum likelihood method and the discriminant function method. The comparison criterion is based on the confidence coefficient and the average widths of the confidence intervals. The study is performed using: the independent variables equal to 3, 5 and 7 which are standard normal distribution, the dependent variable is Bernoulli distribution, the initial value $\beta_j = 1$; $j = 0,1,..., p$ and the sample sizes (n) are 20, 30, 50 and 100. The confidence coefficients are 0.90, 0.95 and 0.99. Monte Carlo simulations are conducted with 2,000 replications for each case. The results show that all three methods have the coverage probability not lower than the nominal level for all situations. Pivotal quantity method has the smallest average widths. In terms of average widths, classical method is as good as Bayesian method for all situations. Moreover, all three methods which the coefficient are estimated by the discriminant function method have smaller average widths than the maximum likelihood method for all situations.

Keywords: logistic regression; interval estimation; regression coefficient; maximum likelihood method; discriminant function method

1. บทนำ

การวิเคราะห์การถดถอย (regression analysis) เป็นวิธีการวิเคราะห์ทางสถิติที่ใช้อธิบายความสัมพันธ์ระหว่างตัวแปรตาม (dependent variable) และตัวแปรอิสระ (independent variable) โดยที่ตัวแปรอิสระอาจจะเป็นตัวแปรเชิงปริมาณเพียงอย่างเดียว หรืออาจมีทั้งตัวแปรเชิงปริมาณและตัวแปรเชิงคุณภาพก็ได้ ในขณะที่ตัวแปรตามจะต้องเป็นตัวแปรเชิงปริมาณ แต่งานวิจัยหลายด้านมีตัวแปรตามเป็นตัวแปรเชิงคุณภาพหรือตัวแปรเชิงกลุ่มที่มีค่าเป็นไปได้ 2 ค่า หรือมากกว่า เช่น การวิจัยทางด้านการแพทย์ (โอกาสการเกิดโรค : เกิดโรค-ไม่เกิดโรค) ทางด้านการศึกษา (ผลการสอบ : ผ่าน-ไม่ผ่าน) และทางเศรษฐศาสตร์ (การจัดกลุ่มผลตอบ- แทนของสหกรณ์ : กำไร-ขาดทุน) เป็นต้น ตัวแปรตามที่มีค่าเป็นไปได้ 2 ค่า เรียกว่า ตัวแปรทวิ (binary variable) เนื่องจากตัวแปรตามเป็นข้อมูลเชิงคุณภาพทำให้ค่าความคลาดเคลื่อนไม่มีการแจกแจงแบบปกติ ซึ่งไม่สอดคล้องกับข้อกำหนดในการวิเคราะห์การถดถอย จึงจำเป็นต้องใช้การวิเคราะห์การถดถอยโลจิสติก (logistic regression analysis) เข้ามาวิเคราะห์ความสัมพันธ์ของข้อมูล ซึ่งวัตถุประสงค์และแนวคิดยังคงเหมือนกับการวิเคราะห์การถดถอย คือ เพื่อศึกษาความสัมพันธ์ระหว่างตัวแปรตาม และตัวแปรอิสระ

การประมาณค่าสัมประสิทธิ์การถดถอยของตัวแบบโลจิสติก $\beta_0, \beta_1, ..., \beta_p$ มี 2 ประเภท คือ การประมาณค่าแบบจุด (point estimation) และการประมาณค่าแบบช่วง (interval estimation) โดยการประมาณค่าแบบจุด เป็นการประมาณค่าพารามิเตอร์ของประชากรด้วยตัวเลขใดตัวเลขหนึ่ง ซึ่งอาจมีค่าความคลาดเคลื่อนเนื่องมาจากตัวอย่างที่ใช้หรือเกิดจากเทคนิคการสุ่ม ส่วนการประมาณค่าแบบช่วงเป็นการประมาณค่าพารามิเตอร์ของประชากรว่าจะอยู่ในช่วงใดช่วงหนึ่ง ซึ่งการประมาณค่าแบบช่วงมีโอกาสความคลาดเคลื่อนน้อยกว่าการประมาณค่าแบบจุด ความกว้างของช่วงจะขึ้นอยู่กับระดับความเชื่อมั่นที่กำหนด และช่วงประมาณยิ่งแคบจะทำให้ได้ค่าใกล้เคียงกับค่าพารามิเตอร์มากยิ่งขึ้น

การประมาณค่าแบบจุดในสัมประสิทธิ์การถดถอยของตัวแบบโลจิสติกมีหลายวิธี เช่น วิธีภาวะน่าจะเป็นสูงสุด (maximum likelihood method) วิธีฟังก์ชันจำแนกประเภท (discriminant function method) วิธีการถ่วงน้ำหนัก (weighting method) วิธีปรับแก้เบื้องต้น (prior correction method) และวิธีแบบริดจ์ (ridge method) เป็นต้น ซึ่งพบว่ามีผู้วิจัยหลายท่านได้ศึกษาและเปรียบเทียบการประมาณค่าแบบจุดวิธีต่าง ๆ ไว้ เช่น ในงานวิจัยของกาญจนา [1] ได้ศึกษาเปรียบเทียบระหว่างวิธีภาวะน่าจะเป็นสูงสุดและวิธีฟังก์ชันจำแนกประเภท ทัศนาพร [2] ศึกษาเปรียบเทียบการประมาณพารามิเตอร์ด้วยวิธีภาวะน่าจะเป็นสูงสุด วิธีการถ่วงน้ำหนัก และวิธีปรับแก้เบื้องต้น เรวดี [3] เปรียบเทียบการประมาณด้วยวิธีแบบริดจ์ วิธีภาวะน่าจะเป็นสูงสุด และวิธีฟังก์ชันจำแนกประเภท Efron [4] เปรียบเทียบด้วยวิธีภาวะน่าจะเป็นสูงสุดกับการวิเคราะห์จำแนกประเภท เมื่อตัวแปรอิสระมีการแจกแจงแบบปกติ Press และ Wilson [5] เปรียบเทียบความแตกต่างของประสิทธิภาพในการจำแนกประเภทของสมการถดถอยโลจิสติก และการวิเคราะห์จำแนกประเภทด้วยภาวะน่าจะเป็นสูงสุด เมื่อตัวแปรอิสระไม่มีการแจกแจงแบบปกติ เป็นต้น ส่วนการประมาณค่าแบบช่วงในสัมประสิทธิ์การถดถอยของตัวแบบโลจิสติกมีหลายวิธี เช่น วิธีคลาสสิค (classical method) วิธีปริมาณหมุน (pivotal quantity method) และวิธีแบบเบส์ (Bayesian method) เป็นต้น จากการศึกษาพบว่ามีงานวิจัยที่ศึกษาเปรียบเทียบการประมาณค่าแบบช่วงตัวอย่าง เช่น ในงานวิจัยของวนิดา [6] ได้ศึกษาการประมาณด้วยวิธีคลาสสิค ราตรี [7] ศึกษาการประมาณค่าแบบช่วงด้วยวิธีคลาสสิค วิธีปริมาณหมุน และวิธีแบบเบส์ ตัวแปรอิสระ 1 ตัวแปร วราภรณ์ [8] เปรียบเทียบวิธีการประมาณเช่นเดียวกับงานวิจัยของราตรี แต่ศึกษาที่ตัวแปรอิสระ 2 ตัวแปร Wilson และ Langenberg [9] ศึกษาวิธีการประมาณด้วยวิธีแบบคลาสสิค วิธีปริมาณหมุน และวิธีแบบเบส์ โดยที่ตัวแปรอิสระมีการแจกแจงแบบปกติ เป็นต้น

จากที่กล่าวมาข้างต้นผู้วิจัยสนใจที่จะทำการเปรียบเทียบวิธีการประมาณช่วงความเชื่อมั่นสำหรับสัมประสิทธิ์การถดถอยของตัวแบบโลจิสติก 3 วิธี คือ วิธีคลาสสิค วิธีปริมาณหมุน และวิธีแบบเบส์ โดยจะเปรียบเทียบช่วงการประมาณที่ได้จากวิธีการประมาณค่าแบบจุดที่แตกต่างกัน คือ วิธีภาวะน่าเป็นสูงสูด และวิธีฟังก์ชันจำแนกประเภท เนื่องจากการประมาณค่าแบบช่วงนั้นจะขึ้นอยู่กับตัวประมาณค่าแบบจุดด้วย

2. วัตถุประสงค์ในการวิจัย

เพื่อเปรียบเทียบประสิทธิภาพของวิธีการประมาณค่าแบบช่วง 3 วิธี คือ วิธีคลาสสิค วิธีปริมาณหมุน และวิธีแบบเบส์ เมื่อใช้วิธีการประมาณค่าพารามิเตอร์ในสมการถดถอยโลจิสติกด้วยวิธีภาวะน่าจะเป็นสูงสุด และวิธีฟังก์ชันจำแนกประเภท

3. ขอบเขตการวิจัย

การวิจัยครั้งนี้กำหนดรายละเอียดขอบเขตที่จะศึกษาดังนี้

3.1 ตัวแปรตาม (Y) เป็นข้อมูลเชิงคุณภาพที่มีลักษณะที่เป็นไปได้ 2 ค่า คือ 0 และ 1

3.2 กำหนดจำนวนตัวแปรอิสระ (X) เท่ากับ 3, 5 และ 7 ตัวแปร โดยตัวแปรอิสระเป็นข้อมูลเชิงปริมาณที่มีการแจกแจงแบบปกติมาตรฐาน มีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ 1

3.3 ขนาดตัวอย่าง (n) เท่ากับ 20, 30, 50 และ 100

3.4 กำหนดค่าพารามิเตอร์เริ่มต้น $\underline{\beta} = 1$

3.5 กำหนดระดับสัมประสิทธิ์ความเชื่อมั่น $(1 - \alpha)$ เท่ากับ 0.90, 0.95 และ 0.99

4. เกณฑ์ที่ใช้ในการศึกษา

4.1 พิจารณาการประมาณช่วงความเชื่อมั่นที่ให้ค่าสัมประสิทธิ์ความเชื่อมั่นไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด

4.2 การเปรียบเทียบความกว้างเฉลี่ยของช่วงความเชื่อมั่นในแต่ละวิธี

5. วิธีการวิจัย

5.1 สร้างข้อมูลของตัวแปรอิสระให้เป็นไปตามลักษณะการแจกแจงที่กำหนดในขอบเขตการวิจัย

5.2 กำหนดค่าพารามิเตอร์ $(\underline{\beta})$ ขนาดตัวอย่าง และสัมประสิทธิ์ความเชื่อมั่นตามที่กำหนดในขอบเขตการวิจัย

5.3 สร้างข้อมูลของตัวแปรตาม จากรูปแบบความสัมพันธ์ [8]

$$Y_i = \begin{cases} 1 \text{ ถ้า } \pi(x_i) \geq R_i \\ 0 \text{ ถ้า } \pi(x_i) < R_i \end{cases}$$

โดย ได้จากการจำลองเลขสุ่ม $R \sim U(0,1)$ และ

$$\pi(x_i) = \frac{e^{\beta_0+\beta_1 x_{1i}+\ldots+\beta_p x_{pi}}}{1+e^{\beta_0+\beta_1 x_{1i}+\ldots+\beta_p x_{pi}}}$$

5.4 ประมาณค่าสัมประสิทธิ์การถดถอย $(\hat{\underline{\beta}})$ โดยใช้วิธีภาวะน่าจะเป็นสูงสุด และฟังก์ชันจำแนกประเภท

5.4.1 วิธีภาวะน่าจะเป็นสูงสุด

พิจารณาข้อมูลที่อยู่ในรูปของ (x_i, y_i) ; $i=1,2,\ldots,n$ โดยตัวแปรตามมีการแจกแจงแบบเบอร์นูลี (Bernoulli Distribution) $[y_i \sim b(1, \pi(x_i))]$ ซึ่ง x_i แทนค่าสังเกตความน่าจะเป็นที่จะเกิดเหตุการณ์ที่สนใจ คือ $\pi(x_i) = P\ (y_i = 1 \mid x_i)$ นั่นคือ

$$Y_i = \begin{cases} 1 \text{ ด้วยความน่าจะเป็น } \pi(x_i) \\ 0 \text{ ด้วยความน่าจะเป็น } 1 - \pi(x_i) \end{cases}$$

ให้ฟังก์ชันการแปลงโลจิสติกของ $\pi(x_i)$ อยู่ในรูป $\text{logit}\,(\pi(x_i))$ ซึ่งเป็นการแปลงค่าความน่าจะเป็นที่เกิดเหตุการณ์ที่สนใจ $\pi(x_i)$ ที่มีค่าอยู่ในช่วง (0,1) ให้เป็นค่าของ $\text{logit}\,(\pi(x_i))$ ที่อยู่ในช่วง $(-\infty, \infty)$ ดังนั้นตัวแบบเชิงเส้นโลจิสติก (logistic linear model) สำหรับ $\pi(x_i)$ [10] คือ $\text{logit}\,(\pi(x_i))$ = $\ln\left(\frac{\pi(x_i)}{1 - \pi(x_i)}\right) = \beta_0 + \beta_1 x_{1i} + \ldots + \beta_p x_{pi}$ จะได้ $\pi(x_i)$ $= \frac{e^{\beta_0+\beta_1 x_{1i}+\ldots+\beta_p x_{pi}}}{1+e^{\beta_0+\beta_1 x_{1i}+\ldots+\beta_p x_{pi}}}$ ซึ่งประมาณค่าพารามิเตอร์ได้ด้วยวิธี Newton-Raphson [11] จากสมการ $\hat{\underline{B}}_{r+1}$ = $\hat{\underline{B}}_r - H^{-1}(\underline{B}_r)\ \underline{U}(\underline{B}_r)$ ถ้าผลต่างระหว่าง $\hat{\underline{B}}$ ในรอบที่ r กับรอบที่ r+1 มีค่าน้อยมากจนถือว่าไม่แตกต่างกัน ซึ่งในที่นี้กำหนดเกณฑ์ว่า $\left|\hat{\underline{B}}_{r+1} - \hat{\underline{B}}\right| < 0.000001$ ค่า $\hat{\underline{B}}_{r+1}$ นั้นจะเป็นค่าที่ยอมรับได้

5.4.2 วิธีฟังก์ชันจำแนกประเภท

ฟังก์ชันจำแนกประเภทเชิงเส้นของฟิชเชอร์ (Fisher's linear discriminant function) แสดงไว้ดังนี้ [4] $\lambda(X) = \ln\left(\frac{\pi(x_i)}{1 - \pi(x_i)}\right) = \beta_0 + \underline{B}'X$

เมื่อ $\beta_0 = \ln\left(\frac{\pi_1}{1-\pi_1}\right) - \frac{1}{2}\left(\underline{U}_1 - \underline{U}_0\right)' \Sigma^{-1}\left(\underline{U}_1 + \underline{U}_0\right)$

$\underline{B}' = \underline{B}' = \left(\underline{U}_1 - \underline{U}_0\right)' \Sigma^{-1}$ ซึ่ง X จะอยู่ในประชากรกลุ่มย่อย 1 ถ้า $\lambda(X) > 0$ และ X จะอยู่ในประชากรกลุ่มย่อย 0 ถ้า $\lambda(X) > 0$ ในทางปฏิบัติพารามิเตอร์ $\pi_1, \underline{U}_1, \underline{U}_0$ และ Σ เป็นพารามิเตอร์ที่ไม่ทราบค่า แต่กลุ่มตัวอย่าง $(y_1, x_1), (y_2, x_2), ..., (y_n, x_n)$ ทำให้หาตัวประมาณค่าของพารามิเตอร์เหล่านี้ได้ โดย y_i จะชี้ว่า x_i มาจากประชากรกลุ่มย่อยกลุ่มใด และ $X_i \mid Y_i \sim N(U_{yi}, \Sigma)$ โดยสมมติให้ค่าสังเกต (y_i, x_i) เป็นอิสระซึ่งกันและกัน ซึ่งทำให้ตัวประมาณค่าเหล่านี้หาได้จาก

$\hat{\pi}_1 = \frac{n_1}{n}$, $\hat{\pi}_0 = \frac{n_0}{n}$, $\hat{\underline{U}}_1 = \bar{X}_1 = \frac{1}{n_1}\sum_{yi=1} x_i$, $\hat{\underline{U}}_0 = \bar{X}_0 = \frac{1}{n_0}\sum_{yi=0} x_i$ เมื่อ n_1 และ n_0 เป็นจำนวนค่าสังเกตที่ $y = 1$ และ $y = 0$ ตามลำดับ และ $\hat{\Sigma} = \frac{1}{n}\left\{\sum_{yi=1}(x_i - \bar{X}_1)(x_i - \bar{X}_1)' + \sum_{yi=0}(x_i - \bar{X}_0)(x_i - \bar{X}_0)'\right\}$

5.5 ประมาณค่าสัมประสิทธิ์ช่วงความเชื่อมั่นในสมการถดถอยโลจิสติก ที่มีตัวแปรอิสระ 3, 5 และ 7 ตัวแปรด้วย 3 วิธี คือ วิธีคลาสสิค วิธีปริมาณหมุน และวิธีแบบเบส์ ดังนี้

5.5.1 วิธีการประมาณช่วงความเชื่อมั่นด้วยวิธีคลาสสิค (classical method)

$\hat{\beta}_i \sim N(\beta_i, \hat{\sigma}_i^2)\ ;\ i = 0,1,...,p$ ได้จากการประมาณด้วยวิธีภาวะน่าจะเป็นสูงสุดเมื่อตัวอย่างขนาดใหญ่ ทำให้ $\frac{\hat{\beta}_i - \beta_i}{\hat{\sigma}_i} \sim N(0,1)$ [12] ดังนั้นจะได้ช่วงความเชื่อมั่น $(1-\alpha)100$ % ด้วยวิธีคลาสสิค ดังนี้

ขีดจำกัดล่างของช่วงความเชื่อมั่น คือ $\exp\left(\hat{\beta}_i - Z_{1-\frac{\alpha}{2}}\hat{\sigma}_i\right)$

ขีดจำกัดบนของช่วงความเชื่อมั่น คือ $\exp\left(\hat{\beta}_i + Z_{1-\frac{\alpha}{2}}\hat{\sigma}_i\right)$

5.5.2 วิธีการประมาณช่วงความเชื่อมั่นด้วยวิธีปริมาณหมุน (pivotal quantity method) จะหาช่วงความเชื่อมั่น $(1-\alpha)100$ % ของ $g(\theta)$ ได้โดย [13]

(1) หาปริมาณหมุน $Q_i(x_1, x_2, ..., x_n; \theta)$

(2) หาค่าคงที่ q_1 และ q_2 ที่ทำให้ $P\left[q_1 \le Q_i \le q_2\right] = 1-\alpha$ ซึ่งจะต้องหาฟังก์ชันความหนาแน่นของ Q_i มาก่อน

(3) เปลี่ยน $P\left[q_1 \le Q_i \le q_2\right] = 1-\alpha$ เป็น $P\left[L(x_1, x_2, ..., x_n) \le g(\theta) \le U(x_1, x_2, ..., x_n)\right] = 1-\alpha$

จาก $\hat{\beta}_i \sim N(\beta_i, \hat{\sigma}_i^2)\ ;\ i = 0,1,...,p$ ได้จากการประมาณด้วยวิธีภาวะน่าจะเป็นสูงสุดของตัวอย่างขนาดใหญ่ ดังนั้น β_i จะมีปริมาณหมุนเท่ากับ $Q_i(\hat{\beta}_i; \beta_i) = \frac{\hat{\beta}_i - \beta_i}{\hat{\sigma}_i}$ ซึ่งจะได้ช่วงความเชื่อมั่น $(1-\alpha)100$ % ด้วยวิธีปริมาณหมุน ดังนี้

ขีดจำกัดล่างของช่วงความเชื่อมั่น คือ $\exp(\hat{\beta}_i - q_{2i}\hat{\sigma}_i)$

ขีดจำกัดบนของช่วงความเชื่อมั่น คือ $\exp(\hat{\beta}_i - q_{1i}\hat{\sigma}_i)$

เมื่อค่า q_1 และ q_2 หาได้จากช่วงของ $\exp(-q_{1i}\hat{\sigma}_i) - \exp(-q_{2i}\hat{\sigma}_i)$ ที่มีค่าน้อยที่สุดภายใต้ข้อจำกัดที่ $P(Z \le q_{2i}) - P(Z \le q_{1i}) = 1-\alpha$

5.5.3 วิธีการประมาณช่วงความเชื่อมั่นด้วยวิธีแบบเบส์ (Bayesian method)

David และ Patricia (1999) กำหนดให้ $\beta_i\ ;\ i = 0,1,...,p$ มีการแจกแจงเบื้องต้น (Prior Distribution) แบบยูนิฟอร์ม $\beta_i \sim U(0,1)$ จากการประมาณด้วยวิธีภาวะน่าจะเป็นสูงสุดของตัวอย่างขนาดใหญ่ ดังนั้นได้การแจกแจงภายหลัง (Posterior

Distribution) ของ β_i มีการแจกแจงแบบปกติ $\beta_i \sim N(\beta_i, \hat{\sigma}_i^2)$ มีฟังก์ชันความหนาแน่นน่าจะเป็น คือ $f(\beta_i) = \frac{1}{\sqrt{2\pi}\hat{\sigma}_i} \exp\left\{-\frac{1}{2\hat{\sigma}_i^2}\left(\hat{\beta}_i - \beta_i\right)^2\right\}$ จาก ความสัมพันธ์ระหว่างการแจงแจงแบบล็อกนอร์มอล กับการแจกแจงแบบปกติ

ให้ $Y_i = \beta_i - \hat{\beta}_i$ เมื่อ $Y_i \sim N(0, \hat{\sigma}_i^2)$ และ $U_i = \exp(Y_i)$ เมื่อ $U_i \sim LN(0, \hat{\sigma}_i^2)$ ดังนั้น U_i มีการแจกแจงแบบล็อกนอร์มอลมีฟังก์ชันความหนาแน่นน่าจะเป็น คือ $g(U_i) = \frac{1}{u_i \hat{\sigma}_i \sqrt{2\pi}} \exp\left\{-\frac{1}{2\hat{\sigma}_i^2}(\ln u_i)^2\right\}$

ดังนั้นจะได้ช่วงความเชื่อมั่น $(1-\alpha)100$ % ด้วยวิธีคลาสสิค ดังนี้

ขีดจำกัดล่างของช่วงความเชื่อมั่น คือ $\exp(\hat{\beta}_i)u_{1i}$

ขีดจำกัดบนของช่วงความเชื่อมั่น คือ $\exp(\hat{\beta}_i)u_{2i}$

เมื่อค่า u_{1i} และ u_{2i} กำหนดให้หาได้จากความสัมพันธ์ $\ln(u_{2i}) / \hat{\sigma}_i = -\ln(u_{1i}) / \hat{\sigma}_i - 2\hat{\sigma}_i$ ภายใต้ข้อตกลง $P(Z \le \ln(u_{2i}) / \hat{\sigma}_i) - P(Z \le \ln(u_{1i}) / \hat{\sigma}_i) = 1-\alpha$ เมื่อ $Z \sim N(0,1)$ [9]

5.6 คำนวณค่าสัมประสิทธิ์ความเชื่อมั่น

5.6.1 คำนวณช่วงความเชื่อมั่นของวิธีการประมาณแต่ละวิธี

5.6.2 พิจารณาว่าช่วงความเชื่อมั่นที่คำนวณได้นั้นคลุมค่าพารามิเตอร์เริ่มต้นที่กำหนดไว้หรือไม่ ถ้าช่วงความเชื่อมั่นที่ได้คลุมค่าพารามิเตอร์เริ่มต้น จะทำการนับจำนวนครั้งและบวกสะสมไว้แล้วนำมาหารด้วยจำนวนครั้งในการกระทำซ้ำ ซึ่งค่าที่ได้ คือ ค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการทดลองของช่วงความเชื่อมั่นที่คำนวณได้จากแต่ละวิธีการประมาณ

5.7 คำนวณค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่น

โดยการคำนวณหาผลต่างระหว่างขีดจำกัดบน และขีดจำกัดล่างของช่วงความเชื่อมั่น และบวกสะสมผลต่างนั้นเอาไว้ ทำเช่นนี้จนครบ 2,000 รอบ หารด้วย 2,000 ในแต่ละสถานการณ์

5.8 เปรียบเทียบค่าความกว้างเฉลี่ยของแต่ละวิธีการประมาณ

การพิจารณาว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการทดลองมีค่าไม่ต่ำกว่าค่าระดับสัมประสิทธิ์ความเชื่อมั่นที่กำหนดจะใช้การทดสอบสมมติฐานด้วยตัวสถิติ Z [8] ดังนี้

สมมติฐานของการทดสอบ คือ

$H_0 : p \ge p_0$

$H_1 : p < p_0$

ช่วงในการยอมรับสมมติฐานหลัก คือ

$$p_0 - Z_\alpha \sqrt{\frac{p_0(1-p_0)}{r}} \le p \le 1$$

โดย p_0 แทนค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด

p แทนค่าสัมประสิทธิ์ความเชื่อมั่นที่คำนวณได้จากแต่ละวิธีการประมาณ

r แทนจำนวนการทดลองซ้ำ

α แทนความน่าจะเป็นของความผิดพลาดประเภทที่ 1 ที่กำหนดในการทดสอบสมมติฐาน

ดังนั้นที่ระดับสัมประสิทธิ์ความเชื่อมั่นที่กำหนด 0.90, 0.95 และ 0.99 หากวิธีการประมาณใดให้ค่าสัมประสิทธิ์ความเชื่อมั่นไม่ต่ำกว่าค่า 0.8914, 0.9420 และ 0.9848 ตามลำดับ จะถือว่าวิธีการประมาณนั้นให้ค่าสัมประสิทธิ์ความเชื่อมั่นไม่ต่ำกว่าค่าระดับสัมประสิทธิ์ความเชื่อมั่นที่กำหนด และถ้าวิธีการใดมีค่าสัมประสิทธิ์ความเชื่อมั่นไม่ต่ำกว่าค่าระดับสัมประสิทธิ์ความเชื่อมั่นที่กำหนดจะนำไปคำนวณหาค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นต่อไป

6. ผลการวิจัย

6.1 การพิจารณาการประมาณช่วงความเชื่อมั่นที่ให้ค่าสัมประสิทธิ์ความเชื่อมั่นไม่ต่ำกว่าที่กำหนด

ทุก ๆ สัมประสิทธิ์การถดถอยของจำนวนตัวแปรอิสระ 3, 5 และ 7 ตัวแปร ที่ระดับสัมประสิทธิ์ความเชื่อมั่น 0.90, 0.95 และ 0.99 การประมาณค่าแบบช่วงทั้ง 3 วิธี คือ วิธีคลาสสิค วิธีปริมาณหมุน และวิธีแบบเบส์ ที่ได้มาจากวิธีการประมาณค่าพารามิเตอร์ในสมการถดถอยโลจิสติกด้วยวิธีภาวะน่าจะเป็นสูงสุด และวิธีฟังก์ชันจำแนกประเภท ให้ค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการทดลองไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด คือ 0.8914, 0.9420 และ 0.9848 ตามลำดับ และพบว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการทดลองของวิธีการประมาณค่าแบบช่วงด้วยวิธีแบบเบส์ ให้ค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการทดลองไม่ต่ำกว่าวิธีปริมาณหมุนในทุกสถานการณ์ ค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการทดลอง เมื่อจำนวนตัวแปรอิสระเท่ากับ 3 ที่ระดับสัมประสิทธิ์ความเชื่อมั่น 0.90 แสดงดังตารางที่ 1

ตารางที่ 1 การเปรียบเทียบค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการทดลองของวิธีการประมาณช่วงความเชื่อมั่น 3 วิธี เมื่อใช้วิธีการประมาณค่าพารามิเตอร์ในสมการถดถอยโลจิสติกด้วยวิธีภาวะน่าจะเป็นสูงสุด และวิธีฟังก์ชันจำแนกประเภท ที่ระดับสัมประสิทธิ์ความเชื่อมั่น 0.90 จำแนกตามขนาดตัวอย่าง

	n	Classical		Pivotal		Bayesian	
		MLE	DF	MLE	DF	MLE	DF
β_1	20	0.9490	0.9490	0.8950	0.8965	0.9495	0.9500
	30	0.9495	0.9555	0.9190	0.9175	0.9530	0.9580
	50	0.9390	0.9350	0.9340	0.9275	0.9635	0.9530
	100	0.9120	0.9170	0.9300	0.9205	0.9575	0.9445
β_2	20	0.9465	0.9455	0.9090	0.9070	0.9465	0.9465
	30	0.9430	0.9470	0.9135	0.9105	0.9470	0.9500
	50	0.9320	0.9350	0.9275	0.9205	0.9520	0.9490
	100	0.9110	0.9200	0.9295	0.9165	0.9565	0.9440
β_3	20	0.9485	0.9460	0.9030	0.9080	0.9495	0.9480
	30	0.9525	0.9495	0.9135	0.9090	0.9565	0.9515
	50	0.9310	0.9295	0.9265	0.9120	0.9515	0.9440
	100	0.9060	0.9125	0.9320	0.9215	0.9590	0.9480

ตารางที่ 1 พบว่าในกรณีตัวแปรอิสระเท่ากับ 3 ค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการทดลองของวิธีการประมาณช่วงความเชื่อมั่น 3 วิธี คือ วิธีคลาสสิค วิธีปริมาณหมุน และวิธีแบบเบส์ ที่ได้มาจากวิธีการประมาณค่าพารามิเตอร์ในสมการถดถอยโลจิสติกด้วยวิธีภาวะน่าจะเป็นสูงสุด และวิธีฟังก์ชันจำแนกประเภท เมื่อขนาดตัวอย่าง 20, 30, 50 และ 100 ที่ระดับสัมประสิทธิ์ความเชื่อมั่น 0.90 ให้ค่า

สัมประสิทธิ์ความเชื่อมั่นไม่ต่ำกว่าที่กำหนด คือ 0.8914 เช่นเดียวกับที่ระดับสัมประสิทธิ์ความเชื่อมั่น 0.95 และ 0.99 ค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการทดลองไม่ต่ำกว่าที่กำหนดคือ 0.9420 และ 0.9848 ตามลำดับ นอกจากนี้กรณีตัวแปรอิสระเท่ากับ 5 และ 7 เมื่อขนาดตัวอย่าง 20, 30, 50 และ 100 ที่ระดับสัมประสิทธิ์ความเชื่อมั่น 0.90, 0.95 และ 0.99 ให้ค่าสัมประสิทธิ์ความเชื่อมั่นไม่ต่ำกว่าที่กำหนดในทุกสถานการณ์

6.2 การพิจารณาความกว้างเฉลี่ย

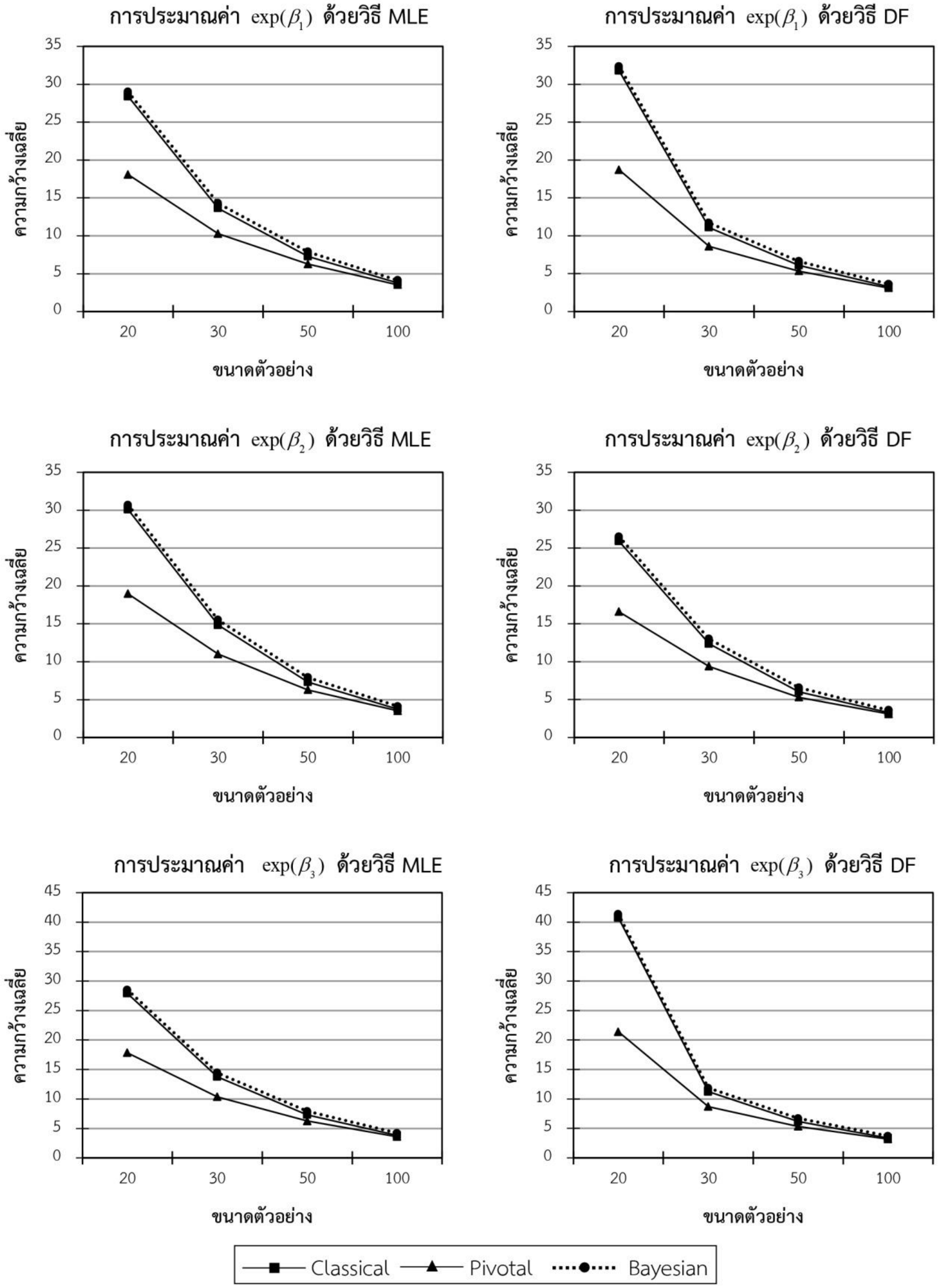
การประมาณค่าแบบช่วงด้วยวิธีปริมาณหมุนทั้งที่มาจากการประมาณค่าพารามิเตอร์ด้วยวิธีภาวะน่าจะเป็นสูงสุด และวิธีฟังก์ชันจำแนกประเภท ให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นแคบที่สุด ส่วนวิธีคลาสสิค และวิธีแบบเบส์ให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นใกล้เคียงกันในทุกสถานการณ์

เมื่อขนาดตัวอย่างเท่ากับ 50 และ 100 ตัวอย่าง การประมาณค่าแบบช่วงทั้ง 3 วิธี ที่ได้มาจากวิธีการประมาณด้วยวิธีฟังก์ชันจำแนกประเภทจะให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นแคบกว่าวิธีภาวะน่าจะเป็นสูงสุดในทุก ๆ สถานการณ์ และยังพบว่าค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นจะแคบลงและมีค่าใกล้เคียงกันเมื่อขนาดตัวอย่างเพิ่มขึ้นในทุกสถานการณ์

การเปรียบเทียบค่าความกว้างเฉลี่ยของวิธีการประมาณช่วงความเชื่อมั่นทั้ง 3 วิธี เมื่อจำนวนตัวแปรอิสระเท่ากับ 3 ที่ระดับสัมประสิทธิ์ความเชื่อมั่น 0.90 แสดงดังรูปที่ 1

รูปที่ 1 พบว่าในกรณีตัวแปรอิสระเท่ากับ 3 ค่าความกว้างเฉลี่ยของวิธีการประมาณช่วงความเชื่อมั่น 3 วิธี คือวิธีคลาสสิค วิธีปริมาณหมุน และวิธีแบบเบส์ ที่ได้มาจากวิธีการประมาณค่าพารามิเตอร์ในสมการถดถอยโลจิสติกด้วยวิธีภาวะน่าจะเป็นสูงสุดและวิธีฟังก์ชันจำแนกประเภท เมื่อขนาดตัวอย่างเท่ากับ 20, 30, 50 และ 100 ที่ระดับสัมประสิทธิ์ความเชื่อมั่น 0.90 การประมาณค่าแบบช่วงด้วยวิธีปริมาณหมุนทั้งที่มาจากการประมาณค่าพารามิเตอร์ด้วยวิธีภาวะน่าจะเป็นสูงสุด และวิธีฟังก์ชันจำแนกประเภท ให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นแคบที่สุด ส่วนวิธีคลาสสิคและวิธีแบบเบส์ให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นใกล้เคียงกัน เช่นเดียวกับที่ระดับสัมประสิทธิ์ความเชื่อมั่น 0.95 และ 0.99 วิธีปริมาณหมุนทั้งที่มาจากการประมาณค่าพารามิเตอร์ด้วยวิธีภาวะน่าจะเป็นสูงสุด และวิธีฟังก์ชันจำแนกประเภท ให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นแคบที่สุด ส่วนวิธีคลาสสิคและวิธีแบบเบส์ให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นใกล้เคียงกัน นอกจากนี้กรณีตัวแปรอิสระเท่ากับ 5 และ 7 ที่ระดับสัมประสิทธิ์ความเชื่อมั่น 0.90, 0.95 และ 0.99 เมื่อขนาดตัวอย่างเท่ากับ 20, 30, 50 และ 100 การประมาณค่าแบบช่วงด้วยวิธีปริมาณหมุนทั้งที่มาจากการประมาณค่าพารามิเตอร์ด้วยวิธีภาวะน่าจะเป็นสูงสุด และวิธีฟังก์ชันจำแนกประเภท ให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นแคบที่สุด ส่วนวิธีคลาสสิคและวิธีแบบเบส์ให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นใกล้เคียงกันในทุกสถานการณ์

กรณีตัวแปรอิสระเท่ากับ 3 เมื่อขนาดตัวอย่างเท่ากับ 50 และ 100 ที่ระดับสัมประสิทธิ์ความเชื่อมั่น 0.90 การประมาณค่าแบบช่วงทั้ง 3 วิธี ที่ได้มาจากการประมาณด้วยวิธีฟังก์ชันจำแนกประเภทจะให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นแคบกว่าวิธีภาวะน่าจะเป็นสูงสุดในทุกสถานการณ์ และพบว่าค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นจะแคบลงและมีค่าใกล้เคียงกันเมื่อขนาดตัวอย่างเพิ่มขึ้น เช่นเดียวกับที่ระดับ 0.95 และ 0.99 การประมาณค่าแบบช่วงทั้ง 3 วิธี ที่ได้มาจากการประมาณด้วยวิธีฟังก์ชันจำแนก



รูปที่ 1 ค่าความกว้างเฉลี่ยของวิธีการประมาณช่วงความเชื่อมั่นทั้ง 3 วิธี เมื่อใช้การประมาณค่าพารามิเตอร์ด้วยวิธีภาวะน่าจะเป็นสูงสุด และวิธีฟังก์ชันจำแนกประเภท ที่ระดับสัมประสิทธิ์ความเชื่อมั่น 0.90

ประเภทจะให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นแคบกว่าวิธีภาวะน่าจะเป็นสูงสุดในทุกสถานการณ์และพบว่าค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นจะแคบลงและมีค่าใกล้เคียงกันเมื่อขนาดตัวอย่างเพิ่มขึ้น นอกจากนี้กรณีตัวแปรอิสระเท่ากับ 5 และ 7 ที่ระดับสัมประสิทธิ์ความเชื่อมั่น 0.90, 0.95 และ 0.99 เมื่อขนาดตัวอย่างเท่ากับ 50 และ 100 การประมาณค่าแบบช่วงทั้ง 3 วิธี ที่ได้มาจากการประมาณด้วยวิธีฟังก์ชันจำแนกประเภทจะให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นแคบกว่าวิธีภาวะน่าจะเป็นสูงสุดในทุกสถานการณ์ และพบว่าค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นจะแคบลงและมีค่าใกล้เคียงกันเมื่อขนาดตัวอย่างเพิ่มขึ้นในทุกสถานการณ์

7. สรุป

ทุกสัมประสิทธิ์การถดถอยของจำนวนตัวแปรอิสระ 3, 5 และ 7 ตัวแปร ที่ระดับสัมประสิทธิ์ความเชื่อมั่น 0.90, 0.95 และ 0.99 การประมาณค่าแบบช่วงทั้ง 3 วิธี คือ วิธีคลาสสิค วิธีปริมาณหมุน และวิธีแบบเบส์ ที่ได้มาจากวิธีการประมาณค่าพารามิเตอร์ในสมการถดถอยโลจิสติกด้วยวิธีภาวะน่าจะเป็นสูงสุด และวิธีฟังก์ชันจำแนกประเภท ให้ค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการทดลองไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนดในทุกสถานการณ์ และการประมาณค่าแบบช่วงด้วยวิธีปริมาณหมุนทั้งที่มาจากการประมาณค่าพารามิเตอร์ด้วยวิธีภาวะน่าจะเป็นสูงสุด และวิธีฟังก์ชันจำแนกประเภท ให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นแคบที่สุด ส่วนวิธีคลาสสิคและวิธีแบบเบส์ให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นใกล้เคียงกันในทุกสถานการณ์ แต่เมื่อขนาดตัวอย่างเท่ากับ 50 และ 100 ตัวอย่าง การประมาณค่าแบบช่วงทั้ง 3 วิธี ที่ได้มาจากวิธีการประมาณค่าพารามิเตอร์ด้วยวิธีฟังก์ชันจำแนกประเภทจะให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นแคบกว่าวิธีภาวะน่าจะเป็นสูงสุดในทุกสถานการณ์ และพบว่าค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นจะแคบลงและมีค่าใกล้เคียงกันเมื่อขนาดตัวอย่างเพิ่มขึ้น

8. วิจารณ์

การเปรียบเทียบประสิทธิภาพของวิธีการประมาณค่าแบบช่วง 3 วิธี คือ วิธีคลาสสิค วิธีปริมาณหมุน และวิธีแบบเบส์ พบว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการทดลองไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด และการประมาณค่าแบบช่วงด้วยวิธีปริมาณหมุนให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นแคบที่สุด ส่วนวิธีคลาสสิค และวิธีแบบเบส์ให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นใกล้เคียงกันในทุกสถานการณ์ ซึ่งสอดคล้องกับงานวิจัยของราตรี [7] ศึกษาช่วงความเชื่อมั่นสำหรับอัตราส่วนความน่าจะเป็นในตัวแบบการถดถอยโลจิสติกพบว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการทดลองของการประมาณค่าแบบช่วงวิธีคลาสสิค วิธีปริมาณหมุน และวิธีแบบเบส์ให้ค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการทดลองไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด และการประมาณค่าด้วยวิธีปริมาณหมุน ให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นแคบที่สุดในทุกๆสถานการณ์ เช่นเดียวกับงานวิจัยของวราภรณ์ [8] ศึกษาการเปรียบเทียบวิธีการประมาณช่วงความเชื่อมั่น สำหรับอัตราส่วนออดส์ในการถดถอยโลจิสติกพหุคูณ พบว่าการประมาณค่าแบบช่วงวิธีคลาสสิค วิธีปริมาณหมุน และวิธีแบบเบส์ ให้ค่าสัมประสิทธิ์ความเชื่อมั่นของพารามิเตอร์อัตราส่วนออดส์ในตัวแปรอิสระทั้ง 2 ตัวแปร ไม่ต่ำกว่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด และการประมาณค่าด้วยวิธีปริมาณหมุน ให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นแคบที่สุด ส่วนการประมาณค่าด้วยวิธีคลาสสิค และวิธีแบบเบส์ให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นใกล้เคียงกันในทุกสถานการณ์

9. ข้อเสนอแนะ

9.1 ควรศึกษาและเปรียบเทียบวิธีการประมาณแบบช่วงในสัมประสิทธิ์การถดถอยของตัวแบบโลจิสติก ภายใต้สถานการณ์อื่นๆต่อไป เช่น กำหนดขนาดตัวอย่างมากกว่า 100 หรือกำหนดตัวแปรตามที่มีค่าเป็นไปได้มากกว่า 2 ค่า เป็นต้น

9.2 ในงานวิจัยครั้งนี้ กำหนดค่าพารามิเตอร์เริ่มต้นเท่ากับ 1 เพียงอย่างเดียว ในการศึกษาครั้งต่อไปอาจกำหนดค่าพารามิเตอร์เริ่มต้นอื่น ๆ

9.3 การศึกษาครั้งต่อไปอาจมีการเปรียบเทียบการประมาณค่าแบบช่วงทั้ง 3 วิธีนี้ โดยศึกษาการประมาณค่าพารามิเตอร์ของสัมประสิทธิ์การถดถอยของตัวแบบโลจิสติกด้วยวิธีอื่น

10. รายการอ้างอิง

[1] กาญจนา พานิชการ, 2539, การประมาณค่าพารามิเตอร์ในสมการถดถอยโลจิสติกด้วยภาวะน่าจะเป็นสูงสุดและฟังก์ชันจำแนกประเภท, วิทยานิพนธ์ปริญญาโท, จุฬาลงกรณ์มหาวิทยาลัย, กรุงเทพฯ.

[2] ทัศนาพร จงเกตุกรณ์, 2546, การประมาณค่าพารามิเตอร์ของตัวแบบถดถอยโลจิสติคทวินาม, วิทยานิพนธ์ปริญญาโท, จุฬาลงกรณ์มหาวิทยาลัย, กรุงเทพฯ.

[3] เรวดี เรืองอยู่, 2547, การเปรียบเทียบการประมาณค่าพารามิเตอร์ด้วยการประมาณแบบริดจ์ภาวะน่าจะเป็นสูงสุดและฟังก์ชันจำแนกประเภทในสมการถดถอยโลจิสติค, วิทยานิพนธ์ปริญญาโท, สถาบันเทคโนโลยีพระจอมเกล้าคุณทหารลาดกระบัง, กรุงเทพฯ.

[4] Efron, B., 1975, The efficiency of logistic regression compared to normal discriminant analysis, J. Amer. Stat. Assoc. 70: 892-898.

[5] Press, J.S. and Wilson, S., 1978, Choosing between logistic regression and discriminant analysis, J. Amer. Stat. Assoc. 73: 699-705.

[6] วนิดา เลิศพิพัฒนานนท์, 2540, การประมาณช่วงความเชื่อมั่นสำหรับสัมประสิทธิ์การถดถอยของตัวแบบถดถอยโลจิสติก, วิทยานิพนธ์ปริญญาโท, จุฬาลงกรณ์มหาวิทยาลัย, กรุงเทพฯ.

[7] ราตรี จรัสมาธุสร, 2547, ช่วงความเชื่อมั่นสำหรับอัตราส่วนความน่าจะเป็นในตัวแบบการถดถอยโลจิสติก, วิทยานิพนธ์ปริญญาโท, จุฬาลงกรณ์มหาวิทยาลัย, กรุงเทพฯ.

[8] วราภรณ์ กรงทอง, 2551, การเปรียบเทียบวิธีการประมาณช่วงความเชื่อมั่นสำหรับอัตราส่วนออดส์ในการถดถอยโลจิสติกพหุคูณ, วิทยานิพนธ์ปริญญาโท, มหาวิทยาลัยเชียงใหม่, เชียงใหม่.

[9] Wilson, P.D. and Langenberg, P., 1999, Usual and shortest confidence intervals on odds ratios from logistic regression, Amer. Stat. 53: 332-335.

[10] Hosmer, D., Lemeshow, S. and Sturdivant, R., 2013, Applied logistic regression, 3rd Ed., John Wiley & Sons, New Jersey.

[11] Millar, R.B., 2011, Maximum likelihood estimation and inference, John Wiley & Sons, England.

[12] Demidenko, E., 2012, The shortest width confidence interval for odds ratio in logistic regression, Open J. Stat. 2: 305-308.

[13] สายชล สินสมบูรณ์ทอง, 2554, สถิติคณิตศาสตร์ 1, พิมพ์ครั้งที่ 5, จามจุรีโปรดักท์, กรุงเทพฯ, 486 น.