

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์ในตัวแบบถดถอยโลจิสติกทวินาม โดยวิธีการประมาณค่าพารามิเตอร์ที่ใช้ในการวิจัยครั้งนี้คือ วิธีความควรจะเป็นสูงสุด (Maximum Likelihood Method : ML) วิธีการถ่วงน้ำหนัก (Weighting Method : WE) และ วิธีปรับแก้เบื้องต้น (Prior Correction Method : PC) เมื่อตัวแบบถดถอยโลจิสติกมีรูปแบบดังนี้ $\pi(x_i) = \frac{e^{\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p}}{1 + e^{\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p}}$ โดยที่ $\pi(x_i)$ แทนความน่าจะเป็นเมื่อเกิดเหตุการณ์ที่สนใจของตัวแปรตาม (Y) X_1, X_2, \dots, X_p แทนตัวแปรอิสระ $\beta_0, \beta_1, \dots, \beta_p$ แทนค่าสัมประสิทธิ์ของการถดถอย p แทนจำนวนตัวแปรอิสระ สำหรับข้อมูลตัวแปรตามที่ใช้ในการวิจัยครั้งนี้มีการแจกแจงแบบทวินามด้วยพารามิเตอร์ $n_i = n$ และ $\pi(x_i)$ การเปรียบเทียบกระทำภายใต้สถานการณ์ของจำนวนตัวแปรอิสระแต่ละตัวแบบเท่ากับ 3 5 และ 7 ตัว จำนวนกลุ่ม (m) เท่ากับ 30 90 150 และ 210 ค่าพารามิเตอร์ n เท่ากับ 10 20 และ 30 ค่าเฉลี่ยความน่าจะเป็นของเหตุการณ์ที่สนใจของประชากรเท่ากับ 0.1 0.3 0.5 และ 0.8 เกณฑ์ที่ใช้ในการเปรียบเทียบคือ ค่าระยะทางมาหาโลบิสเฉลี่ย (Average Mahalanobis distance : AMH) ข้อมูลที่ใช้ในวิจัยครั้งนี้ได้จากการจำลองด้วยเทคนิคมอนติคาร์โลโดยการกระทำซ้ำ 500 ครั้ง ในแต่ละสถานการณ์โดยใช้โปรแกรม S-Plus 2000 ผลการวิจัยสรุปดังนี้

เมื่อค่าเฉลี่ยความน่าจะเป็นของเหตุการณ์ที่สนใจของประชากรเท่ากับ 0.1 และ 0.3 จะพบว่า วิธีความควรจะเป็นสูงสุดให้ค่า AMH ต่ำที่สุด รองลงมาคือ วิธีปรับแก้เบื้องต้น และ วิธีการถ่วงน้ำหนัก ตามลำดับ สำหรับตัวอย่างทุกจำนวนกลุ่มและทุก n แต่ในกรณีที่ค่าเฉลี่ยความน่าจะเป็นของเหตุการณ์ที่สนใจของประชากรเท่ากับ 0.5 และ 0.8 จะพบว่า วิธีปรับแก้เบื้องต้น ให้ค่า AMH ต่ำที่สุด รองลงมาคือ วิธีการถ่วงน้ำหนัก และ วิธีความควรจะเป็นสูงสุด ตามลำดับ สำหรับตัวอย่างทุกจำนวนกลุ่มและทุก n จะเห็นได้ว่ากรณีค่าเฉลี่ยความน่าจะเป็นของเหตุการณ์ที่สนใจระหว่างประชากรและตัวอย่างใกล้เคียงกันมากขึ้นวิธีการถ่วงน้ำหนักและวิธีการปรับแก้เบื้องต้นจะทำให้ประมาณค่าพารามิเตอร์มีประสิทธิภาพมากขึ้น

The objective of this research is to compare the parameter estimation methods of binomial logistic regression model. The methods of estimating parameter under consideration in this research are Maximum Likelihood Method (ML), Weighting Method (WE) and Prior Correction Method (PC). The model for logistic regression is as follows:
$$\pi(x_i) = \frac{e^{\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p}}{1 + e^{\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p}}$$
 where $\pi(x_i)$ is the probability of interested events of dependent variable; X_1, X_2, \dots, X_p are the independent variable; $\beta_0, \beta_1, \dots, \beta_p$ are the regression coefficients; p is the number of independent variable. In addition the data of dependent variable of this research has been binomial distribution with $n_i = n$ and $\pi(x_i)$ parameters. The comparison is done under conditions of the numbers of independent variable are equal to 3, 5 and 7, sample group are equal to 30, 90, 150 and 210, the determination of n parameter value is 10, 20 and 30 and the average probability of success in population are equal to 0.1, 0.3, 0.5 and 0.8. The criteria employed for the comparison are Average Mahalanobis distance (AMH). The data for this research is simulated by using the Monte Carlo simulation technique with 500 repetitions for each situation by S-plus 2000 package. The results of this research are as follows:

According to the comparison of AMH from three referred methods, it is found that when the average probability of success in population are equal to 0.1 and 0.3, ML method would give the lowest AMH, while PC and WE are the second and the third lowest, respectively, for all groups and values of n . In case the average probability of success in population are equal to 0.5 and 0.8, PC method would give the lowest AMH, while WE and ML are the second and the third lowest, respectively, for all groups and values of n . It is found that the closer the average probability of success between population and sample are, the more efficiency that parameter estimation from WE and PC methods will be.