บทคัดย่อ

T160338

การวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบ วิธีการประมาณค่าพารามิเตอร์ในสมการถดถอยโลจิสติค วิธีการประมาณค่าพารามิเตอร์ที่ใช้ในการศึกษาครั้งนี้ คือ วิธีการแบบริดจ์ (RE) วิธีภาวะน่าจะเป็นสูงสุด (MLE) และวิธีฟังก์ชันจำแนกประเภท (DF) โดยที่ตัวแปรตาม (Y) เป็นตัวแปรเชิงคุณภาพมี 2 ค่าคือ 0 หรือ 1 และมีตัวแปรอธิบาย(X) 1 หรือ 2 ตัวแปร การเปรียบเทียบกระทำภายใต้ขนาดตัวอย่าง $n =$ 10 , 20 , 30 และ 40 โดยมีสัดส่วนของตัวแปรตาม $(Y=1)$ ค่อนข้างสูง ($P \geq 0.75$) คือ 0.75 , 0.80 , 0.85 , 0.90 และ 0.95 และการแจกแจงของตัวแปรอธิบาย 3 แบบ คือ การแจกแจงแบบปกติ การแจกแจงแบบชี้กำลัง และการแจกแจงแบบไวบูลล์

ค่าที่ใช้ในการเปรียบเทียบคือ ค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Square Root Mean Square Error (RMSE)) และ ตัวสถิติ Deviance(DV) โดยข้อมูลที่ใช้ในการวิจัยได้จากการจำลองระบบด้วยเทคนิคมอนติคาร์โล กระทำซ้ำ 500 รอบ ในแต่ละสถานการณ์ ผลการวิจัยพบว่า กรณีตัวแปรอธิบาย 1 ตัว วิธี RE ให้ค่า RMSE และค่า DV มากกว่าวิธีอื่นในทุกการแจกแจงของตัวแปรอธิบาย และทุกเงื่อนไขที่ทำการศึกษา ในขณะที่ วิธี MLE ให้ค่า RMSE และค่า DV น้อยกว่า วิธี DF ยกเว้นเมื่อขนาดตัวอย่าง $n =$ 30 และ 40 วิธี DF ให้ค่า RMSE และค่า DV น้อยกว่า วิธี MLE กรณีตัวแปรอธิบาย 2 ตัว ที่มีการแจกแจงแบบปกติและการแจกแจงแบบชี้กำลัง หรือ การแจกแจงแบบปกติและการแจกแจงแบบไวบูลล์ หรือการแจกแจงแบบชี้กำลังและการแจกแจงแบบไวบูลล์ วิธี RE ให้ค่า RMSE และค่า DV มากกว่าวิธีอื่นในทุกการแจกแจง และทุกเงื่อนไขที่ทำการศึกษา ในขณะที่ วิธี MLE ให้ค่า RMSE และค่า DV น้อยกว่า วิธี DF

ABSTRACT

TE160338

The objective of this study is to compare the estimation of parameters in logistic regression. The methods of estimating parameters under consideration in this study are Ridge Estimation method (RE), Maximum Likelihood Estimation method (MLE), and Discriminant Function method (DF). Dependent variable (Y) is dichotomous (0 or 1) with one or two explanatory variables (X). The comparison are done under conditions of sample sizes 10, 20, 30 and 40, high proportion $(P \geq 0.75)$of dependent variable $(Y = 1)$: 0.75, 0.80, 0.85, 0.90 and 0.95 , and 3 distributions of explanatory variables : Normal distribution , Exponential distribution and Weibull distribution.

The criteria employed for the comparison are Square Root Mean Squares Error (RMSE) and Deviance (DV). The data of the experiment are generated through The Monte Carlo simulation technique with 500 repetitions. It was found that in case of one explanatory variable , the RE method always gave larger RMSE and DV than the other two methods for every given conditions , whlile the MLE method gave smaller RMSE and DV than DF method , except $n =$ 30 and 40 which the DF method gave smaller RMSE and DV than MLE. In case of two explanatory variables (Normal and Exponential , Normal and Weibull , or Exponential and Weibull) , the RE method always gave larger RMSE and DV than the other two methods for every given conditions , whlile the MLE method gave smaller RMSE and DV than DF.