

วิธีทางสถิติที่นิยมใช้ในการทำนายการเป็นสมาชิกกลุ่มที่ถูกต้องในกรณีที่ตัวแปรตอบสนองมี 2 ค่า คือวิธีวิเคราะห์จำแนกกลุ่มในเชิงเส้นตรง วิธีวิเคราะห์การถดถอยโลจิสติกและวิธีวิเคราะห์การถดถอยพหุคูณ ซึ่งข้อสมมติในการใช้วิธีวิเคราะห์จำแนกกลุ่มในเชิงเส้นตรงและวิธีวิเคราะห์การถดถอยพหุคูณจะคล้ายคลึงกัน การใช้วิธีวิเคราะห์การถดถอยพหุคูณในกรณีที่เกิดพหุสัมพันธ์ระหว่างตัวแปรที่ใช้ทำนายจะทำให้ตัวประมาณค่าที่ได้จากวิธีกำลังสองน้อยที่สุดมีความแปรปรวนมาก วิธีที่สามารถแก้ปัญหานี้ได้คือวิธีรีดจ์รีเกรสชัน ในขณะที่วิธีวิเคราะห์การถดถอยโลจิสติกไม่มีข้อสมมติเกี่ยวกับการแจกแจงปกติ ความสัมพันธ์เชิงเส้นและความแปรปรวนที่เท่ากันของตัวแปรที่ใช้ทำนาย

ในการศึกษาครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพในการทำนายการเป็นสมาชิกกลุ่มที่ถูกต้องด้วยวิธีวิเคราะห์จำแนกกลุ่มในเชิงเส้นตรง วิธีวิเคราะห์การถดถอยโลจิสติก วิธีวิเคราะห์การถดถอยพหุคูณและวิธีรีดจ์รีเกรสชัน โดยใช้ค่าเฉลี่ยของความคลาดเคลื่อนในการจำแนกกลุ่ม ค่าเฉลี่ยของ B และค่าเฉลี่ยของ C ผลการศึกษาพบว่าในทุกกรณีวิธีวิเคราะห์การถดถอยโลจิสติกและวิธีวิเคราะห์จำแนกกลุ่มในเชิงเส้นตรง ให้ผลการทำนายไม่แตกต่างกันและสามารถทำนายได้ดีกว่าวิธีอื่นๆ

คำสำคัญ : วิธีวิเคราะห์จำแนกกลุ่มในเชิงเส้นตรง วิธีวิเคราะห์การถดถอยโลจิสติก
วิธีวิเคราะห์การถดถอยพหุคูณ วิธีรีดจ์รีเกรสชัน พหุสัมพันธ์

Abstract

243307

The most widely used statistical methods for analyzing categorical outcome variables are Linear Discriminant Analysis and Logistic Regression. If a dependent variable is a binary outcome, an analyst can choose among Discriminant Analysis, Logistic and Multiple Regression. The statistical assumptions required for Discriminant Analysis are essentially the same as for Multiple Regression. In the presence of multicollinearity the ordinary least squares (OLS) estimator could become unstable due to their large variance, which leads to poor prediction. The one of the popular solution of this problem is Ridge regression. Logistic Regression makes no assumption about the distribution of the independent variables. They do not have to be normally distributed, linearly related or of equal variance within each group.

The purpose of this study was to compare the accuracy of the Classifications of Group membership of the Multiple regression Ridge regression Logistic Regression and Linear Discrimination Analysis with the mean of classification error mean of B and mean of C. The results showed that Logistic Regression and Linear discrimination analysis performed better than the others.

Keywords: Linear Discriminant Analysis , Logistic Regression , Multiple Regression , Ridge regression , multicollinearity