

จริยา แสงสุวรรณ 2551: การศึกษาเปรียบเทียบวิธีการประมาณค่าสูญหาย  
ในการวิเคราะห์การถดถอยพหุคูณ ปริญญาวิทยาศาสตรมหาบัณฑิต (สถิติ) สาขาสถิติ  
ภาควิชาสถิติ อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก: รองศาสตราจารย์ประสิทธิ์ พัยคมพงษ์,  
M.S. 84 หน้า

งานวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบวิธีการประมาณค่าสูญหายของตัวแปรตามใน  
การวิเคราะห์การถดถอยพหุคูณ โดยทำการประมาณค่าสูญหายของตัวแปรตาม 4 วิธี คือ วิธีสูญหาย  
วิธีค่าเฉลี่ย วิธีสมการถดถอย และวิธีการใส่ค่าหลายค่าแทนข้อมูลที่สูญหายแต่ละค่า (วิธีเอ็มไอ)  
เกณฑ์ที่ใช้ในการเปรียบเทียบคือ ค่าประมาณของรากที่สองของค่าเฉลี่ยของความคลาดเคลื่อน  
กำลังสอง (RMSE) การเปรียบเทียบกระทำภายใต้สถานการณ์ของขนาดตัวอย่างเท่ากับ 50 70 100  
และ 200 ค่าเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนเท่ากับ 1 5 และ 15 เปอร์เซนต์การสูญหายของ  
ตัวแปรตามเท่ากับ 5% 10% 20% และ 30% ตามลำดับ และตัวแปรอิสระมีการแจกแจงแบบปกติ  
หลายตัวแปร ซึ่งระดับความสัมพันธ์ระหว่างตัวแปรอิสระมี 3 ระดับ คือ ระดับต่ำ (0.20)  
ระดับปานกลาง (0.50) และระดับสูง (0.70) ในการวิจัยครั้งนี้ได้ทำการจำลองด้วยวิธีมอนติคาร์โล  
ซึ่งกระทำซ้ำ 5,000 ครั้ง ในแต่ละสถานการณ์

ผลการวิจัยสรุปได้ดังนี้ เมื่อเปอร์เซนต์การสูญหายเพิ่มขึ้น วิธีสมการถดถอยและวิธีเอ็มไอ  
ให้ค่าประมาณของ RMSE ลดลง และวิธีการประมาณค่าสูญหายทั้ง 4 วิธี ให้ค่าประมาณของ RMSE  
แตกต่างกัน วิธีสมการถดถอยและวิธีเอ็มไอ ให้ค่าประมาณของ RMSE ใกล้เคียงกัน แต่เนื่องจาก  
วิธีสมการถดถอยเป็นวิธีที่ง่ายและไม่ซับซ้อน ดังนั้น วิธีสมการถดถอย จึงเป็นวิธีที่เหมาะสมสำหรับ  
การประมาณค่าสูญหายของตัวแปรตามในการวิเคราะห์การถดถอยพหุคูณ



ลายมือชื่อนิสิต



ลายมือชื่ออาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

17 / ๗ / ๒๕๖๑

Chariya Saengsuwan 2008: A Comparative Study of Missing Data Estimation Methods in Multiple Regression Analysis. Master of Science (Statistics), Major Field: Statistics, Department of Statistics. Thesis Advisor: Associate Professor Prasit Payakkapong, M.S. 84 pages.

The purpose of this research was to compare missing data estimation methods of the dependent variables in multiple regression analysis. Four estimation methods of missing data of the dependent variables were loss method, mean method, regression method and multiple imputation method (mi). The criterion of comparison was the estimated squares root of mean squares error (RMSE). The comparisons were done under the conditions of sample sizes 50, 70, 100 and 200; standard deviations of error 1, 5 and 15; missing percentage 5%, 10%, 20% and 30% and distribution of independent variables were multivariate normal distribution with the levels of correlations among independent variables classified into 3 levels : low levels (0.20), middle levels (0.50) and high levels (0.70). The study used the monte carlo simulation method with repeated 5,000 times under each situation.

The results of the study are concluded as follows. When missing percentage are increased the estimated values of RMSE of regression method and mi method are decreased and four estimation methods of missing data give different increasing values of RMSE. The estimated values of RMSE of regression method and mi method give similar results but regression method is easier and less complicate, so that regression method is suitable method for estimation the missing data of dependent variables in multiple regression analysis.

Chariya Saengsuwan  
Student's signature

Prasit Payakkapong  
Thesis Advisor's signature

17 / March / 2008