170183

รุ่งฤทัย ไทยสม : การประมาณค่าพารามิเตอร์ของตัวแบบการถดถอยเชิงเส้นเชิงเดียวด้วยวิธีเบส์และวิธีประมาณความควรจะเป็นสูงสุด (PARAMETERS ESTIMATION IN SIMPLE LINEAR REGRESSION MODEL BY BAYESIAN METHOD AND MAXIMUM LIKELIHOOD METHOD ) อ.ที่ปรึกษา : รศ. รอ. มานพ วราภักดิ์ , 105 หน้า .ISBN 974-53-1445-5

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของการประมาณค่าพารามิเตอร์ถดถอยของตัวแบบเชิงเส้นเชิงเดียว กรณีที่ให้ตัวแปรตามมีการแจกแจงแบบปกติและกรณีที่ตัวแปรตามมีการแจกแจงแบบล็อกนอร์มัลและแปลงเป็นการแจกแจงแบบปกติ โดยจะเปรียบเทียบวิธีการประมาณพารามิเตอร์การถดถอย 2 วิธี ได้แก่วิธีความควรจะเป็นสูงสุด (Maximum Likelihood Method (MLE)) และวิธีเบส์ (Bayesian Method (BAYES)) เมื่อกำหนดให้การแจกแจงก่อนของพารามิเตอร์การถดถอยเป็นแบบปกติสองตัวแปร เกณฑ์ที่ใช้ในการเปรียบเทียบประสิทธิภาพคือค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเฉลี่ย (Average Mean Squares Error (AMSE)) สถานการณ์ที่ศึกษาคือตัวแปรอิสระ $X_i$ จำลองมาจากการแจกแจงปกติที่มีค่าเฉลี่ยเท่ากับ 50 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 10 การแจกแจงความคลาดเคลื่อนสุ่ม $\varepsilon_i$ เป็นแบบปกติที่มีค่าเฉลี่ย $E(\varepsilon_i)$ เท่ากับ 0 ส่วนเบี่ยงเบนมาตรฐาน $SD(\varepsilon_i)$ เท่ากับ 1.0 3.0 5.0 7.0 และ 9.0 ค่า $\underline{\beta} = (1.0,1.0)'$ สำหรับการแจกแจงก่อนปกติสองตัวแปรของพารามิเตอร์การถดถอย $\underline{\beta}$ มีค่าเฉลี่ย $\underline{\mu} = \begin{pmatrix}0.5\\1.0\end{pmatrix}, \begin{pmatrix}1.0\\2.0\end{pmatrix}, \begin{pmatrix}2.0\\3.0\end{pmatrix}$, ความแปรปรวน $\sigma_1^2$ และ $\sigma_2^2$ มีค่าสอดคล้องกับสัมประสิทธิ์ความแปรผัน $(C.V.)$ของการแจกแจงในระดับต่ำ กลาง สูง ซึ่งในที่นี้กำหนดเป็นค่า 0.6 1.3 และ 1.8 ตามลำดับ และค่าสหสัมพันธ์ $(\rho)$ระหว่างพารามิเตอร์มีค่า -0.3 -0.1 0.5 0.7 และ 0.9 และขนาดตัวอย่าง ($n$) ที่ศึกษาเท่ากับ 10 20 30 50 70 และ 90 จำลองสถานการณ์การทดลองด้วยเทคนิคมอนติคาร์โลทำซ้ำ 500 ครั้ง ในแต่ละสถานการณ์ของการทดลอง

ผลการวิจัยสามารถสรุป ได้ดังนี้

1. กรณีที่ตัวแปรตามมีการแจกแจงแบบล็อกนอร์มัลและแปลงเป็นการแจกแจงแบบปกติ ให้ผลสรุปเหมือนกับกรณีที่ตัวแปรตามมีการแจกแจงแบบปกติ ซึ่งอธิบายตามกรณีของสหสัมพันธ์ระหว่าง $\beta_0$ และ $\beta_1$ ได้ดังนี้

1.1 เมื่อ $\beta_0$ และ $\beta_1$ มีสหสัมพันธ์ในระดับต่ำ ($-0.3 \le \rho \le 0.3$) เมื่อ $0 < SD(\varepsilon_i) \le 2.1$ ทุกค่า $n$ และ $2.1 < SD(\varepsilon_i) < 8.2$, $n \ge 46$ จะได้ว่า MLE ให้ประสิทธิภาพในการประมาณมากที่สุด แต่ถ้า$2.1 < SD(\varepsilon_i) < 8.2$, $0 < n < 45$ และ $SD(\varepsilon_i) \ge 8.2$ ทุกค่า $n$ จะได้ว่า BAYES ให้ประสิทธิภาพในการประมาณมากที่สุด

1.2 เมื่อ $\beta_0$ และ $\beta_1$ มีสหสัมพันธ์ในระดับปานกลาง ($-0.7 \le \rho < -0.3$ หรือ $0.3 < \rho \le 0.7$) เมื่อ $0 < SD(\varepsilon_i) \le 3.3$ ทุกค่า $n$ และ $SD(\varepsilon_i) > 3.3$, $n \ge 27$ จะได้ว่า MLE ให้ประสิทธิภาพในการประมาณมากที่สุด แต่ถ้า $SD(\varepsilon_i) > 3.3$, $0 < n < 26$ จะได้ว่า BAYES ให้ประสิทธิภาพในการประมาณมากที่สุด

1.3 เมื่อ $\beta_0$ และ $\beta_1$ มีสหสัมพันธ์ในระดับสูง( $\rho < -0.7$ หรือ $\rho > 0.7$ ) ที่ค่า $0 < SD(\varepsilon_i) \le 8.5$ ทุกค่า $n$ และที่ค่า $SD(\varepsilon_i) \ge 8.5$, $n \ge 28$ จะได้ว่า MLE ให้ประสิทธิภาพในการประมาณมากที่สุด แต่ถ้า $SD(\varepsilon_i) \ge 8.5$, $0 < n < 27$ จะได้ว่า BAYES ให้ประสิทธิภาพในการประมาณมากที่สุด

2. เมื่อ $SD(\varepsilon_i), \sigma_1^2, \sigma_2^2$ หรือ $\rho$ เพิ่มขึ้นค่า AMSE มีแนวโน้มเพิ่มขึ้น แต่เมื่อ $n$ เพิ่มขึ้นค่า AMSE มีแนวโน้มลดลง

ภาควิชา สถิติ
สาขาวิชา สถิติ
ปีการศึกษา 2547

ลายมือชื่อนิสิต.....รุ่งฤทัย ไทยสม.....
ลายมือชื่ออาจารย์ที่ปรึกษา.....

170183

## 4482374726 : MAJOR STATISTICS

KEY WORD : LINEAR REGRESSION / BAYES METHOD / MAXIMUM LIKELIHOOD METHOD

ROOGRUTHAI THAISOM : PARAMETERS ESTIMATION IN SIMPLE LINEAR REGRESSION MODEL BY BAYESIAN METHOD AND MAXIMUM LIKELIHOOD METHOD.

THESIS ADVISOR : ASSOC. PROF. CAPT. MANOP VARAPHAKDI , M.S. 105 pp.

ISBN 974-53-1445-5

The objective of this research is to compare the efficiency of regression – coefficient estimation in simple linear regression model under the normal and lognormal distributions of the dependent variable. The estimation methods are Maximum Likelihood Method (MLE) and Bayesian Method (Bayes). The measurement for the efficiency of the methods is Average Mean Square Error (AMSE). This research specified the parameter $\underline{\beta} = (1.0,1.0)'$. The observations of independent variable ($X_i$) are generated from the normal distribution with mean 50 and standard deviation 10. Random errors ($\varepsilon_i$) are independent and identically disdributed normal with mean zero and standard deviation 1.0, 3.0, 5.0, 7.0, and 9.0. The joint prior distribution of the parameters is bivariate normal distribution with mean vectors $\underline{\mu} = \begin{pmatrix}0.5\\1.0\end{pmatrix}, \begin{pmatrix}1.0\\2.0\end{pmatrix}, \begin{pmatrix}2.0\\3.0\end{pmatrix}$ and variances $\sigma_1^2$ and $\sigma_2^2$ which have the values according to the cofficient of variation that are low level, medium level and high level : 0.6, 1.3, and 1.8, respectively, and the level of correlation among parameters $(\rho)$ are –0.3, -0.1, 0.5, 0.7, and 0.9. The sample sizes ($n$) are 10, 20, 30, 50, 70, and 90. The AMSE of the estimates are computed through the Monte Carlo Simulation method. The simulation is repeated 500 times in each situation.

The results of this research are as follows,

1. For the correlation among parameter $\beta_0$ and $\beta_1$ are $-0.3 \le \rho \le 0.3$ at $0 < SD(\varepsilon_i) \le 2.1$ for all $n$, and $2.1 < SD(\varepsilon_i) < 8.2$, $n \ge 46$, MLE is the most efficiency. But for $2.1 < SD(\varepsilon_i) < 8.2$, $0 < n < 45$, and $SD(\varepsilon_i) \ge 8.2$ for all $n$, BAYES is the most efficiency.

2. For the correlation among parameter $\beta_0$ and $\beta_1$ are $-0.7 \le \rho < -0.3$ or $0.3 < \rho \le 0.7$ at $0 < SD(\varepsilon_i) \le 3.3$ for all $n$, and at $SD(\varepsilon_i) > 3.3$, $n \ge 27$, MLE is the most efficiency. But for $SD(\varepsilon_i) > 3.3$, $0 < n < 26$, BAYES is the most efficiency.

2. For the correlation among parameter $\beta_0$ and $\beta_1$ are $\rho > 0.7$ at $0 < SD(\varepsilon_i) \le 8.5$ for all $n$, and at $SD(\varepsilon_i) \ge 8.5$, $n \ge 28$, MLE is the most efficiency. But at $SD(\varepsilon_i) \ge 8.5$, $0 < n < 27$, BAYES is the most efficiency.

4. The value of AMSE increases while the value of $SD(\varepsilon_i)$, $\sigma_1^2$, $\sigma_2^2$ or $\rho$ increases. But the value of AMSE decreases while the sample size $n$ increases.

Department Statistics
Field of study Statistics
Academic 2004

Student 's signature.....
Advisor 's signature.....