

# A Comparison of Bayesian and Maximum Likelihood Estimations in a SUR Tobit Regression Model

Seung-Chun Lee<sup>a,1</sup>, Byongsu Choi<sup>b</sup>

<sup>a</sup>Department of Applied Statistics, Hanshin University

<sup>b</sup>Department of Multimedia Engineering, Hansung University

(Received August 13, 2014; Revised October 7, 2014; Accepted October 7, 2014)

---

## Abstract

Both Bayesian and maximum likelihood methods are efficient for the estimation of regression coefficients of various Tobit regression models (see. *e.g.* Chib, 1992; Greene, 1990; Lee and Choi, 2013); however, some researchers recognized that the maximum likelihood method tends to underestimate the disturbance variance, which has implications for the estimation of marginal effects and the asymptotic standard error of estimates. The underestimation of the maximum likelihood estimate in a seemingly unrelated Tobit regression model is examined. A Bayesian method based on an objective noninformative prior is shown to provide proper estimates of the disturbance variance as well as other regression parameters

Keywords: Seemingly unrelated Tobit regression model, maximum likelihood estimate, EM algorithm, Bayes estimation, Gibbs sampling.

---

## 1. 서론

중도절단된 관측값은 생존분석 등에서 흔하게 접할 수 있으며, 반응변수가 중도절단된 관측값을 포함하고 있는 회귀모형을 중도절단 회귀모형(censored regression model)이라고 한다. 특히, 반응변수가 0에서 절단된 경우를 흔히 토빗회귀모형(Tobit regression model)이라고 하지만 단방향으로 절단된 경우, 중도절단 회귀모형과 토빗회귀모형을 사실한 동일하기 때문에 여기에서 토빗회귀모형은 단방향 중도절단 회귀모형을 의미한다.

일반적으로 회귀모형의 모수는 최소제곱법에 의해 추정되고 있으나 토빗회귀모형에서 최소제곱추정량은 편의추정량(biased estimator)이라는 것이 잘 알려져 있다. 특히, 중도절단된 데이터의 비중이 클 경우, 편이의 정도가 커지기 때문에 설명변수의 영향을 과소, 또는 과대 추정하게 된다. 이러한 이유로 Amemiya (1984), Greene (1990) 등은 토빗모형에서 모수 추정방법을 연구하였다. 결론적으로 빈도학파의 추정방법은 최대가능도추정법(maximum likelihood estimation)으로 귀결된다. 한편, 베이즈안 추정에서는 Bruno (2004), Cowles 등 (1996), Hamilton (1999), Lee와 Choi (2013), Lee와 Choi

---

This work was supported by Hanshin University research grant.

This research was financially supported by Hansung University.

<sup>1</sup>Corresponding author: Department of Applied Statistics, Hanshin University, 441 Yangsan-Dong, Osan, Gyeonggi-do, 447-791, Korea. E-mail: [seung@hs.ac.kr](mailto:seung@hs.ac.kr)

(2014) 등, 대부분의 연구자들이 깁스샘플링(Gibbs sampling)에 의한 추정방법을 제시하고 있다. 특히, Chib (1992)은 모수의 사후분포 기대값을 몬테카를로 적분, 라플라스 근사법 (Tierney와 Kadane, 1986) 및 깁스샘플링(Gibbs sampling)으로 구하는 방법에 대한 모의실험을 하여 깁스샘플링을 추천한 바 있다. 이 모의실험에서 Chib (1992)은 표본크기가 큰 경우는 베이저안 추정과 최대가능도 추정은 큰 차이를 보이지 않지만, 표본크기가 작은 경우에는 베이저안 추정값이 최대가능도 추정값보다 실제값에 가까운 것으로 결론지었고, 베이저안 추정값의 표준오차가 최대가능도 추정의 표준오차보다 모든 표본 크기에서 대체로 큰 값을 갖는다고 하였다.

Chib (1992)의 모의실험에서 베이저안 추정값과 최대가능도 추정값이 큰 차이를 보이지 않음에도 불구하고 표준오차의 크기에서는 차이가 나는 이유에 Lee와 Choi (2013)는 최대가능도 추정방법이 표준오차를 과소 추정하기 때문이라고 설명하고 있다. 즉, 최대가능도 추정에서 정규근사에 의한 Wald 신뢰구간은 명목 신뢰수준보다 작은 포함확률(coverage probability)을 갖는 것으로 나타났으며, 이는 최대가능도 추정의 표준오차가 과소 추정되었기 때문이라고 하였다. 또, 이와 같이 표준오차가 과소추정된 원인의 하나는 오차항 분산의 과소추정에 있다고 하였다.

최대가능도 추정에서 오차항 분산의 과소 추정문제는 Greene (2004a,b)에서도 언급되었다. 그는 고정효과를 갖는 패널 회귀모형(fixed effect panel regression model)에서 절편(intercept)의 수가 과다하게 많기 때문에 발생하는 문제, 즉 incidental parameters problem (Lancaster, 2000) 때문에 최대가능도 추정에서 오차항 분산이 과소 추정된다고 하였다. 그러나, 변량효과를 갖는 패널 회귀모형(random effect panel regression model)에서도 분산요소(variance component)의 과소추정 문제가 Lee와 Choi (2014)에 의해 발견되었기 때문에 오차항 분산의 과소 추정이 incidental parameters problem 때문이라기 보다는 최대가능도 추정의 일반적인 특성일 가능성이 높다고 하겠다. 본 연구의 동기는 보기에 무관한 회귀모형에서 반응변수가 중도 절단된(seemingly unrelated Tobit regression model; SUR 토빗 회귀모형)에서 최대가능도 추정방법이 표준오차 및 오차항 분산을 과소 추정할 가능성과 베이저안 추정방법이 이러한 과소 추정 문제를 해결할 수 있는지에 대해 알아 보는데 있다.

Chib (1992)은 토빗모형의 베이저안 추정에서 깁스샘플링에 의한 추정이 매우 효율적이었다고 기술하고 있다. 그러나 베이저안 추정의 효율은 일반적으로 사후 기대값을 구하는 방법보다는 사전분포(prior distribution)에 의존하는 바가 크다. 즉, 베이저안 방법론에서 사전분포는 분석자의 사전정보를 반영하게 되는데 이와 같이 주관적인 분석은 빈도학파적 입장에서의 효율성을 담보하기 어렵다. 그러나 무정보사전분포(noninformative prior)에 의한 베이저안 추론에서는 분석결과가 사전정보 보다는 데이터 자체에 의존하는 바가 크므로 빈도학파적 효율성을 갖는 경우가 많다. 그러므로 여기에서는 무정보 사전분포에 의한 베이저안 추정과 최대가능도 추정과 효율성을 비교하기로 한다.

## 2. SUR 토빗회귀모형

두 잠재변수(latent variable)  $Z_1$ 과  $Z_2$ 에 대한 SUR 회귀모형

$$\begin{bmatrix} Z_{1t} \\ Z_{2t} \end{bmatrix} = \begin{bmatrix} x'_{1t} & 0 \\ 0 & x'_{2t} \end{bmatrix} \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} + \begin{bmatrix} U_{1t} \\ U_{2t} \end{bmatrix}, \quad t = 1, \dots, n, \quad (2.1)$$

$$Z_t = x'_t B + U$$

에서 관측값  $y_{it}$ 가 다음과 같이 주어진다고 하고,

$$y_{it} = \begin{cases} Z_{it}, & \text{if } Z_{it} > 0, \\ 0, & \text{if } Z_{it} \leq 0, \end{cases} \quad (2.2)$$

오차항  $U_t = (U_{1t}, U_{2t})'$ 은 독립적으로 이변량 정규분포  $N(\mathbf{0}, \Sigma)$ 에 따른다고 가정하자. 단,  $\Sigma$ 는

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix}$$

와 같다. 즉, 본 연구에서 고려된 SUR 토빗모형의 로그가능도 함수(log-likelihood function)는  $\sigma^{ij}$ 를  $\Sigma^{-1}$ 의  $(i, j)$  요소라고 할 때

$$\begin{aligned} \ell(\boldsymbol{\theta}; \mathbf{Z}) &= \sum_{t=1}^n \ell(\boldsymbol{\theta}; Z_t) \\ &= \sum_{t=1}^n \left\{ c + \frac{1}{2} \left[ \ln |\Sigma^{-1}| - \sum_{i=1}^2 \sigma^{ii} (Z_{it} - x'_{it} B_i)^2 - 2\sigma^{ij} (Z_{1t} - x'_{1t} B_1) (Z_{2t} - x'_{2t} B_2) \right] \right\} \end{aligned} \quad (2.3)$$

와 같다. 단,  $\boldsymbol{\theta} = (B, \Sigma^{-1})$ 이다.

### 2.1. 베이지안추정

$\pi(\boldsymbol{\theta})$ 를 모수  $\boldsymbol{\theta}$ 에 대한 사전분포라고 하면, 베이지안 방법론에서  $\boldsymbol{\theta}$ 는 사후기대값

$$E(\boldsymbol{\theta}|\mathbf{y}) = \frac{\int \boldsymbol{\theta} L(\boldsymbol{\theta}; \mathbf{y}) \pi(\boldsymbol{\theta}) d\boldsymbol{\theta}}{\int L(\boldsymbol{\theta}; \mathbf{y}) \pi(\boldsymbol{\theta}) d\boldsymbol{\theta}} \quad (2.4)$$

으로 추정된다. 여기서  $L(\boldsymbol{\theta}; \mathbf{y})$ 은 가능도함수(likelihood function)를 나타내는데, 사전분포가 부적절 분포(improper distribution)인 경우도 식 (2.4)의 분모항이 유한이면 사후기대값이 존재하는 것으로 알려져 있다.

일반적으로 베이지안 추론에서는 식 (2.4)의 값을 해석적(analytic) 방법으로 구하지 못하는 경우가 많기 때문에 이를 해결하기 위한 여러 가지 수치적 방법론이 연구되어 왔다. Chib (1992)은 왼쪽에서 절단된(left-censored) 토빗모형에서 사후기대값을 구하기 위해 몬테칼로 적분, 라플라스 근사 및 깃스샘플링 방법을 적용해 보고 모의실험을 통해 깃스샘플링 방법에 의한 추정이 상대적으로 효율성이 높다고 하였다. 깃스샘플링 방법은 사후 기대값을 구하는 과정에서 사후분포에 대한 추론이 가능하기 때문에 구간추정 또는 가설검정 등의 분석에도 용이하다는 장점도 있다. 이러한 여러 가지 이유로 토빗모형의 베이지안 추정에 관한 대부분의 연구와 R 패키지들이 깃스샘플링 방법을 채택하고 있다.

베이지안 추론의 특징은 사후기대값을 구하는 방법보다는 사전분포에 영향을 받는 것으로 알려져 있다. 전통적인 베이지안 방법론에서는 사전분포는 분석자의 주관적인 사전정보를 반영하기 때문에 빈도학파적 효율성을 담보하기 어려울 수 있다. 그러나 무정보 사전분포(noninformative prior)에 의한 베이지안 추론은 분석결과가 사전정보 보다는 데이터 자체에 의존하는 특성을 갖으므로 빈도학파에서 요구하는 효율성 및 객관성 갖는 경우가 많다. 이러한 이유로 Lee와 Choi (2013)와 Lee와 Choi (2014)에서와 같이 최대가능도 추정의 표준오차 과소 추정문제를 해결할 수 있는 대안이 될 수 있을 것으로 예상할 수 있다.

선형모형에서 베이지안 추론에 대한 자세한 내용은 이미 잘 알려져 있기 때문에 세부적인 내용의 기술은 피하기로 한다. 비록 적용된 사전분포는 다르지만 SUR 토빗모형의 베이지안 추정의 세부적인 내용은 Huang (2001)와 Wichitaksorn과 Choy (2011) 등을 참조할 수 있다.

(2.1)의  $Z_t|B, \Sigma \sim N(x_t B, \Sigma)$ 에 대한 베이지안 모형의 확립을 위해  $B$ 와  $\Sigma$ 에 대한 사전분포를 정의하여야 한다.  $B$ 의 사전분포는 Huang (2001)와 같이 공액사전분포인 다변량정규분포 또는 무정보 사

전분포인 균등분포가 많이 이용되고 있다. 다변량정규분포의 경우 분산값을 크게하여 무정보 사전분포에 가깝게 설정할 수 있다. 균등분포는 부적절 사전분포(improper prior distribution)이지만 사후분포는 진분포(proper distribution)가 된다는 것이 잘 알려져 있으며, 다변량정규분포와는 달리 분산값을 설정할 필요가 없기 때문에 선호된다. 한편, 분산공분산행렬  $\Sigma$ 의 무정보 사전분포는 부적절 사전분포인 Jeffreys가 잘 알려져 있으나, 이 경우 사후분포가 부적절 분포(improper distribution)가 되는 것 (Natarajan과 Kass, 2000)으로 알려져 있다.  $\Sigma$ 에 대한 사전분포로는 공액분포인 역위샤트 분포(inverse Wishart distribution)가 널리 이용되고 있으며, 여기에서도 Daniels와 Kass (1999)에서 추천한 바와 같이 자유도 2와 척도모수(scale parameter)  $\Psi$ 인 역위샤트 분포를 적용하기로 한다. 즉, 사전분포는

$$p(B, \Sigma) \propto |\Sigma|^{-\frac{5}{2}} \exp\left(\frac{1}{2} \text{tr}(\Psi \Sigma^{-1})\right)$$

과 같다. Daniels와 Kass (1999)는 척도모수  $\Psi$ 에 대해 최대가능도 추정값을 제시하고 있으나, 여기에서는 비교를 위해 절단되지 않은 반응변수들의 최소제곱 추정값을 사용하였다.

김스표본을 위한 완전조건부 분포(full-conditional distribution)은 다음과 같이 설정될 수 있으며,

$$\begin{aligned} p(B, \Sigma | Z) &\propto p(Z | B, \Sigma) p(B, \Sigma) \\ &\propto |\Sigma|^{-\frac{n+2}{2}} \exp\left[-\frac{1}{2} \left( \sum_{t=1}^n (Z_t - X'_t B)' \Sigma^{-1} (Z_t - X'_t B) + \text{tr}(\Psi \Sigma^{-1}) \right)\right] \end{aligned}$$

모수들의 완전조건부 분포가 다음과 같이 유도된다.

$$\begin{aligned} B | \Sigma, Z &\sim N\left(\left[\sum_{t=1}^n X'_t \Sigma^{-1} X_t\right]^{-1} \left[\sum_{t=1}^n X'_t \Sigma^{-1} Z_t\right], \left[\sum_{t=1}^n X'_t \Sigma^{-1} X_t\right]^{-1}\right) \\ \Sigma | B, Z &\sim \text{Inv-Wishart}_{n+2}\left(\Psi + \sum_{t=1}^n (Z_t - X'_t B) (Z_t - X'_t B)'\right) \end{aligned}$$

마지막으로 확대 데이터(data-augment, Tanner and Wong, 1987) 단계를 위한 잠재변수의 완전조건부 분포는 반응변수들의 절단 여부에 따라 다음과 같이 4 가지 경우로 구분하여 구할 수 있다.

1.  $Z_t | y_t, B, \Sigma = y_t$ 로 퇴화(degenerate) if  $t \in S_1 = \{t | y_{1t} > 0, y_{2t} > 0\}$ .
2.  $Z_{1t} | y_t, B, \Sigma \sim \text{TN}_{(-\infty, 0)}(x'_{1t} B_1 + \sigma_{12} \sigma_{22}^{-1} (y_{2t} - x'_{2t} B_2), \sigma_{11} - \sigma_{12}^2 / \sigma_{22})$ 이고,  $Z_{2t} | y_t, B, \Sigma = y_{2t}$ 로 퇴화 if  $t \in S_2 = \{t | y_{1t} = 0, y_{2t} > 0\}$ .
3.  $Z_{1t} | y_t, B, \Sigma = y_{1t}$ 로 퇴화하고,  $Z_{2t} | y_t, B, \Sigma \sim \text{TN}_{(-\infty, 0)}(x'_{2t} B_2 + \sigma_{12} \sigma_{11}^{-1} (y_{1t} - x'_{1t} B_1), \sigma_{22} - \sigma_{12}^2 / \sigma_{11})$  if  $t \in S_3 = \{t | y_{1t} > 0, y_{2t} = 0\}$ .
4.  $Z_t | y_t, B, \Sigma \sim \text{TN}_{(-\infty, 0)}(x'_t B, \Sigma)$  if  $t \in S_4 = \{t | y_{1t} = 0, y_{2t} = 0\}$ .

여기서  $\text{TN}_{(a,b)}(\mu, \sigma)$ 는 범위  $(a, b)$  밖에서 절단되고, 평균(또는 평균벡터)  $\mu$ , 분산(또는 분산공분산행렬)  $\sigma$ 인 절단 정규분포(truncated normal distribution)을 나타낸다.

초기치  $B^{(0)}, \Sigma^{(0)}$ 가 주어지면 베이지안 추정을 위한 김스표본의  $i$ -번째 반복은 다음과 같이 진행된다.

단계 1.  $p(Z_t | B^{(i-1)}, \Sigma^{(i-1)}, y_t)$ 에서  $Z_t^{(i)}, t = 1, \dots, n$  추출.

단계 2.  $p(B | \Sigma^{(i-1)}, Z^{(i)})$ 에서  $B^{(i)}$  추출.

단계 3.  $(\Sigma | B^{(i)}, Z^{(i)})$ 에서  $\Sigma^{(i)}$  추출.

2.2. 최대가능도추정

토빗회귀모형에서 최대가능도 추정값은 EM 알고리즘을 이용하여 쉽게 구할 수 있다. EM 알고리즘에 대해서는 이미 잘 알려져 있을 뿐 아니라, SUR 토빗모형에 대한 EM 알고리즘 적용 방법이 Huang 등 (1987)에 잘 설명되어 있으므로 여기에서는 Huang 등 (1987)에서 언급되지 않은 표준오차의 계산에 대해 그들의 기호를 이용하여 설명하기로 한다.

최대가능도 추정량의 근사분산은 정보행렬(information matrix)의 역행렬에서 구할 수 있다. EM 알고리즘에 의해 최대가능도 추정값을 구할 경우 Louis (1982)에 의해 제안된 방법에 따라 관측 정보행렬(observed information matrix)을 구할 수 있다. 관측 정보는 일반적으로 기대 정보(expected information) 보다 정보의 측정치로서 적절 (Efron과 Hinkley, 1978)한 것으로 알려져 있기 때문에 관측 정보행렬 계산을 위한 Louis의 (3.2') 식

$$I_Y(\hat{\theta}) = \sum_{t=1}^n E_{\hat{\theta}} \{B_t(Z_t; \hat{\theta}) | y_t\} - \sum_{t=1}^n E_{\hat{\theta}} \{S_t(Z_t; \hat{\theta}) S_t^T(Z_t; \hat{\theta}) | y_t\} - 2 \sum_{t < t'} E_{\hat{\theta}} \{S_t(Z_t; \hat{\theta}) | y_t\} E_{\hat{\theta}}^T \{S_{t'}(Z_{t'}; \hat{\theta}) | y_{t'}\} \tag{2.5}$$

은 최대가능도 추정량의 표준오차 계산에 유용하다.

식 (2.5)의  $S_t(Z_t; \theta)$ 는  $B_t(Z_t; \theta)$ 는 각각 식 (2.3)의  $\ell(\theta; Z_t)$ 를  $\theta$ 에 대해 1차 미분한 스코어 함수(score function)와 2차 미분에  $-1$ 을 곱한 것으로  $x'_{1t}B_1 = \mu_{1t}, x'_{2t}B_2 = \mu_{2t}$ 라고 쓸 때

$$S_t(Z_t; \theta) = \begin{bmatrix} \sigma^{11}x_{1t}(Z_{1t} - \mu_{1t}) + \sigma^{22}x_{1t}(Z_{2t} - \mu_{2t}) \\ \sigma^{12}x'_{2t}(Z_{2t} - \mu_{2t}) + \sigma^{12}x'_{2t}(Z_{1t} - \mu_{1t}) \\ \frac{\sigma_{11} - (Z_{1t} - \mu_{1t})^2}{2} \\ \frac{\sigma_{22} - (Z_{2t} - \mu_{2t})^2}{2} \\ \sigma_{12} - (Z_{1t} - \mu_{1t})(Z_{2t} - \mu_{2t}) \end{bmatrix},$$

$$B_t(Z_t; \theta) = \begin{bmatrix} \sigma^{11}x_{1t}x'_{1t} & \sigma^{12}x_{1t}x'_{2t} & -x_{1t}(Z_{1t} - \mu_{1t}) & 0 & -x_{1t}(Z_{2t} - \mu_{2t}) \\ \sigma^{12}x_{2t}x'_{1t} & \sigma^{22}x_{2t}x'_{2t} & 0 & -x_{2t}(Z_{2t} - \mu_{2t}) & -x_{2t}(Z_{1t} - \mu_{1t}) \\ -(Z_{1t} - \mu_{1t})x'_{1t} & 0 & \frac{\sigma_{11}^2}{2} & \frac{\sigma_{12}^2}{2} & \sigma_{11}\sigma_{12} \\ 0 & -(Z_{2t} - \mu_{2t})x'_{2t} & \frac{\sigma_{12}^2}{2} & \frac{\sigma_{22}^2}{2} & \sigma_{22}\sigma_{12} \\ -(Z_{2t} - \mu_{2t})x'_{1t} & -(Z_{1t} - \mu_{1t})x'_{2t} & \sigma_{11}\sigma_{12} & \sigma_{22}\sigma_{12} & |\Sigma| + 2\sigma_{12}^2 \end{bmatrix}$$

와 같이 나타낼 수 있다.

식 (2.5)의 오른쪽 첫 번째와 세 번째 항을 구하기 위해서는  $E_{\hat{\theta}} \{Z_{it} - \hat{\mu}_{it} | y_t\}, E_{\hat{\theta}} \{(Z_{it} - \hat{\mu}_{it})^2 | y_t\}, i = 1, 2, E_{\hat{\theta}} \{(Z_{1t} - \hat{\mu}_{1t})(Z_{2t} - \hat{\mu}_{2t}) | y_t\}$ 의 계산이 요구되는 바, 이 값들은 Huang 등 (1987)에서 기술된  $M(1), M(2), M(1, 0), M(2, 0)$  및  $M(1, 1)$ 을 이용하여 구할 수 있으나,  $M(2, 0)$ 에서  $\Phi(h_{1t})$ 와  $\Phi(h_{2t})$ 는 각각  $\Phi(A_{1t})$ 와  $\Phi(A_{2t})$ 의 오타이다. EM 알고리즘의 적용을 위해서는  $E_{\hat{\theta}} \{Z_{it}Z_{jt} | y_t\}$ 와 같이 잠재변수들만의 기대값을 구하는 것이 편리하지만, 정보행렬의 계산에서는 위에서도 같이 잠재변수에서 평균을 차감한 변수들의 기대값이 유용하다. 그러나 한 형태의 기대값을 구하면 다른 유형의 기대값은 쉽게 얻을 수 있다. 한편, 두 번째 항에서는  $E_{\hat{\theta}} \{(Z_{it} - \hat{\mu}_{it})^3 | y_t\}, E_{\hat{\theta}} \{(Z_{it} - \hat{\mu}_{it})^4 | y_t\}, i = 1, 2$ 와  $E_{\hat{\theta}} \{(Z_{it} - \hat{\mu}_{it})(Z_{jt} - \hat{\mu}_{jt})^2 | y_t\}, E_{\hat{\theta}} \{(Z_{it} - \hat{\mu}_{it})(Z_{jt} - \hat{\mu}_{jt})^3 | y_t\}, E_{\hat{\theta}} \{(Z_{it} - \hat{\mu}_{it})^2(Z_{jt} - \hat{\mu}_{jt})^2 | y_t\},$

$i, j = 1, 2$ , 단  $i \neq j$ 의 기대값이 필요하며, 이 값들은  $t$ -번째 반응변수들의 절단 여부에 따라 다음과 같이 4 가지 경우로 구분하여 구하게 된다.

1.  $t \in S_1$  일 경우, 임의의  $r$ 과  $s$ 에 대해 다음이 성립된다.

$$E_{\hat{\theta}} \{ (Z_{1t} - \hat{\mu}_{1t})^r (Z_{2t} - \hat{\mu}_{2t})^s | y_t \} = (y_{1t} - \hat{\mu}_{1t})^r (y_{2t} - \hat{\mu}_{2t})^s .$$

2.  $t \in S_2 = \{t | y_{1t} = 0, y_{2t} > 0\}$ 이면

$$E_{\hat{\theta}} \{ (Z_{1t} - \hat{\mu}_{1t})^r (Z_{2t} - \hat{\mu}_{2t})^s | y_t \} = E_{\hat{\theta}} \{ (Z_{1t} - \hat{\mu}_{1t})^r | y_t \} (y_{2t} - \hat{\mu}_{2t})^s$$

이 된다. 또,  $T_{1t} = (Z_{1t} - \mu_{1t})/\sigma_1, T_{2t} = (Z_{2t} - \mu_{2t})/\sigma_2$  (단,  $Z_{2t} = y_{2t}$ )라고 할 때  $T_{1t}|y_t$ 는 평균  $\rho T_{2t}$ , 분산  $(1 - \rho^2)$ 인 절단 정규분포  $TN_{(-\infty, h_{1t})}(\rho T_{2t}, 1 - \rho^2)$ 을 따르게 된다. 여기서  $h_{1t} = -\mu_{1t}/\sigma_1$ 이다. 그러므로  $M(r) = E(T_{1t}^r | y_t)$ 라고 할 때,  $E_{\hat{\theta}} \{ (Z_{1t} - \hat{\mu}_{1t})^r | y_t \} = \hat{\sigma}_1^r M(r)$ 과 같이 구할 수 있다.

$TN_{(-\infty, a)}(\mu, \sigma^2)$ 의 적률생성함수(moment generating function)  $M_Y(t)$ 는

$$M_Y(t) = \frac{1}{\Phi\left(\frac{a-\mu}{\sigma}\right)} \exp\left(\mu t + \frac{1}{2}\sigma^2 t^2\right) \Phi\left(\frac{a-\mu}{\sigma} - \sigma t\right)$$

이므로 이를 이용하면  $M(3)$ 과  $M(4)$ 는 각각

$$\begin{aligned} M(3) &= 3M(1)M(2) - 2M(1)^3 - \sqrt{1-\rho^2}D_2, \\ M(4) &= M(1)M(3) + 3M(2)(M(2) - M(1)^2) - \sqrt{1-\rho^2}(3M(1)D_2 + D_3) \end{aligned}$$

와 같이 구할 수 있다. 여기서  $\Phi(\cdot)$ 와  $\phi(\cdot)$ 은 표준정규분포의 분포함수와 밀도함수를 나타내고,  $D_2$ 와  $D_3$ 는  $D(t) = \phi(\alpha_{2t} - \sigma t)/\Phi(\alpha_{2t} - \sigma t)$ 를  $t$ 에 대해 각각 2차 및 3차 미분하여 얻은 것으로

$$\begin{aligned} D_2 &= D''(0) = (1 - \rho^2) D_0 ((\alpha_{2t} + D_0)(\alpha_{2t} + 2D_0) - 1), \\ D_3 &= D'''(0) = \sqrt{1 - \rho^2} \{ D_2(\alpha_{2t} + 2D_0) + 2(1 - \rho^2) D_0(\alpha_{2t} + D_0) [D_0(\alpha_{2t} + D_0) - 1] \} \end{aligned}$$

와 같이 정의된다. 이때,  $\alpha_{2t} = (h_{1t} - \rho T_{2t})/\sqrt{1 - \rho^2}, D_0 = D(0) = \phi(\alpha_{2t})/\Phi(\alpha_{2t})$ 이다.

3.  $t \in S_3 = \{t | y_{1t} > 0, y_{2t} = 0\}$ 일 경우, 첨자를 바꾸면 2의 결과를 이용할 수 있다.

4.  $t \in S_4 = \{t | y_{1t} = 0, y_{2t} = 0\}$ 이면  $(T_{1t}, T_{2t})$ 는 평균과 분산이 각각 0과 1이고 공분산이  $\rho$ 인 표준 절단 이변량정규분포(standard truncated bivariate normal distribution)을 따르게 된다. 이 때  $(T_{1t}, T_{2t})$ 의 하한과 상한은 각각  $(-\infty, -\infty)$ 와  $(h_{1t}, h_{2t})$ (단,  $h_{2t} = -\mu_{2t}/\sigma_2$ )이다. 이 경우  $M(r, s) = E\{T_{1t}^r T_{2t}^s | T_{1t} > h_{1t}, T_{2t} > h_{2t}\}$ 는  $r + s \leq 2$ 일 때 Huang 등 (1987)과 같이 해석적(analytic) 방법으로 구할 수 있으나,  $r + s > 2$ 일 때는 수치적(numerical) 방법에 의존할 수 밖에 없다. 즉, 임의표본에 의한 몬테칼로 적분(Monte Carlo integration)을 통해  $M(r, s)$ 를 추정하거나 수치적분을 통해 근사값을 구하여야 한다.  $r + s \leq 4$ 에서는 수치적분을 이용하여 매우 안정적으로  $M(r, s)$ 을 구할 수 있었으며, 특히 **R**의 cubature 패키지 (Johnson과 Narasimhan, 2014)를 이용하여 오차한도  $1.e - 5$ 이내에서  $M(r, s)$ 를 구할 수 있었다. 이렇게  $M(r, s)$  값이 구하면 정보행렬을 구하기 위해 필요한 값들을

$$E_{\hat{\theta}} \{ (Z_{it} - \hat{\mu}_{it})^r (Z_{it} - \hat{\mu}_{it})^s | y_t \} = \hat{\sigma}_1^r \hat{\sigma}_2^s M(r, s)$$

의 식에 의해 얻을 수 있다.

**Table 3.1.** Estimates and standard error of SUR tobit model for Testscore data.

	Bayesian				Maximum likelihood			
	V21		V23		V21		V23	
	Estimate	STE	Estimate	STE	Estimate	STE	Estimate	STE
Intercept	7.9832	1.7575	5.2783	1.4769	8.1132	1.6850	5.3184	1.4284
V1	0.1235	0.0841	0.0171	0.0745	0.1242	0.0811	0.0184	0.0719
V2	0.2498	0.0936	0.2628	0.0805	0.2454	0.0904	0.2604	0.0784
V3	0.2027	0.2006	0.3185	0.1650	0.1894	0.1906	0.3138	0.1592
Variance of Error	13.375		10.815		12.321		10.237	
Covariance of Errors	3.5719				3.3581			

Huang (1999)은 식 (2.5)의 기대값 자체를 몬테칼로 적분에 의존하여 구하는 방법을 제시하고 있다. 이 방법은 정보행렬을 매우 간단히 구할 수 있다는 장점이 있으나 정확한 계산이 가능함에도 불구하고 수치적 방법에 의해 근사값을 구할 이유는 없다고 판단된다.

### 3. 베이지안 추정과 최대가능도 추정의 비교

#### 3.1. Testscore 자료 분석

Testscore는 Scientific Software International, LISREL에서 절단 다변량 회귀모형의 예제로 소개된 데이터로써 스웨덴어의 메타음운 문자(meta-phonological character)의 연구를 위해 90명 학생들의 읽기, 쓰기 등 11개 항목의 점수를 수집한 것이다. 이 데이터는 Jöreskog (2004)에서 다변량 절단회귀의 예제로 소개되었으나, 2개의 반응변수(V21, V23)과 3개의 독립변수(V1, V2, V3)에 대한 설명이 결여되어 있기 때문에 분석결과에 대한 의미를 찾기는 어려운 상황이다. 여기에서는 단지 최대가능도 추정값과 베이지안 추정값을 비교하는데 의미를 갖기로 한다.

반응변수 V21과 V23은 각각 17과 16의 오른쪽에서 41개(45.56%)와 21개(23.33%)의 관찰값이 절단되어 있어 비교적 높은 절단율을 가지고 있다. Jöreskog (2004)에서는 변수의 가장 작은 값, 또는 가장 큰 값이 2개 이상 관찰이 되면 절단된 것으로 간주하였고, V21은 가장 작은 3이 3개 관찰되어 3의 왼쪽에서도 절단된 것으로 판단하고 있다. 그러나, 왼쪽 절단 비율(3.33%)은 비교적 작기 때문에 왼쪽에서의 절단은 고려하지 않기로 한다. 이와같이 Testscore 데이터의 반응변수들은 오른쪽에서 절단되었기 때문에 반응변수 V21과 V23에 -1을 곱하고, 다시 각각 17과 16을 더하여 SUR 토빗모형으로 변환하였다. 이렇게 변환된 데이터에서 추정된 회귀계수들에게 -1을 곱하고, 절편에 각각 17과 16을 더하면 추정오차의 변함이 없이 원래 데이터의 회귀계수 추정값을 구할 수 있다. 이러한 과정을 거쳐 얻어진 베이지안과 최대가능도 추정값 및 추정오차는 Table 3.1과 같다.

Testscore 데이터에서 SUR 토빗모형의 회귀계수들은 베이지안과 최대가능도 방법에 의해 거의 같은 값으로 추정되었다. 이러한 유사성은 Figure 3.1에서 확인할 수 있다. Figure 3.1은 베이지안 추정값과 최대가능도 추정값 차이의 상대적 크기를 나타내기 위해 모수들의 사후분포와 같이 추정값을 표시한 그림으로, 상단에서 반응변수 V21에 대한 V1, V2, V3의 회귀계수 추정값을 상대적으로 비교할 수 있다. Figure 3.1에 그려지지는 않았지만 V23에 대한 추정값 비교에서도 V21과 거의 차이가 없다는 것을 Table 3.1에서 확인할 수 있다.

Chib (1992), Lee와 Choi (2013), Lee와 Choi (2014) 등에서 언급된 결과와 같이 SUR 토빗모형에서도 베이지안과 최대가능도 추정이 유사하기 때문에 이는 토빗모형의 일반적인 특징인 것으로 판단된다. 그러나 앞에서 언급된 바와 같이 본 연구는 추정값의 비교보다는 추정오차에 대한 비교에 주안점이 있

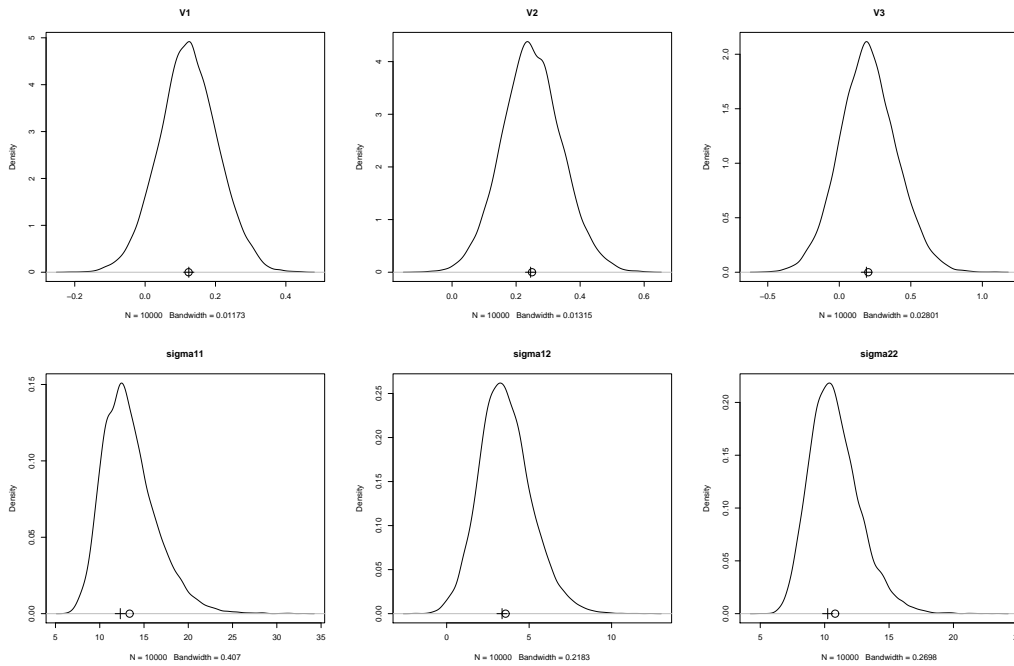


Figure 3.1. Estimates(Bayesian(“○”) and MLE(“+”)) and Posterior distribution of some parameters

며, Table 3.1에서 주목하여야 할 부분은 표준오차에 대한 비교이다. 즉, 최대가능도 추정의 표준오차가 베이저안 추정의 표준오차보다 일률적으로 작다는 것에 유의할 필요가 있다.

두 추정값의 유사성에도 불구하고 표준오차의 크기에서 차이가 나는 것은 베이저안 추정의 표준오차가 과대 추정되었거나 또는 최대가능도 추정의 표준오차가 과소추정되었음을 시사한다. 또, 이 차이는 오차항 분산의 과대 또는 과소 추정이 하나의 원인이 되는 것으로 해석이 된다. 즉, Figure 3.1의 하단에는 오차항의 분산을 비교하기 위해 그려진 것인데, 회귀계수의 추정과는 달리 분산의 추정에서는 유의미한 차이가 있는 것으로 나타나고 있다. 오차항 분산 추정값의 차이가 표준오차의 차이에 대한 하나의 원인이 되는 것으로 추측할 수 있기 때문에 표준오차에 대한 적절한 평가와 함께 오차항 분산의 추정에 대한 효율성을 비교하여야 하는 바, 다음 절에서는 모의실험을 통해 이를 검토하기로 한다.

3.2. 표준오차 및 오차항 분산 추정의 비교를 위한 모의실험

표준오차의 적절성 평가 및 오차항 분산 추정의 효율성을 비교하기 위해

$$\begin{aligned}
 y_1 &= \max\{\beta_{10} + \beta_{11}x_{11} + \beta_{12}x_{12} + \beta_{13}x_{13} + \epsilon_1, 0\}, \\
 y_2 &= \max\{\beta_{20} + \beta_{21}x_{21} + \beta_{22}x_{22} + \beta_{23}x_{23} + \epsilon_2, 0\}, \\
 \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \end{pmatrix} &\sim N\left(0, \begin{bmatrix} \sigma_{11} & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_{22} \end{bmatrix}\right)
 \end{aligned}$$

과 같은 SUR 토빗모형에서 표본크기  $n = 100$ 인 데이터를 발생시키고 베이저안 및 최대가능도 추정을 실시하였다. 이러한 과정을 2000회 반복하여 추정값 및 표준오차의 평균을 구하였다. 베이저안과 최



**Table 3.2.** Averages of estimate, coverage and width of confidence interval with relatively small sample size.

$\theta$		$\beta_{10}$	$\beta_{11}$	$\beta_{12}$	$\beta_{13}$	$\sigma_{11}$	$\beta_{20}$	$\beta_{21}$	$\beta_{22}$	$\beta_{23}$	$\sigma_{22}$	$\sigma_{12}$	
parameter		-3	1	-1.5	0.1	1	8	4	3.5	-0.5	2		
0.1	Average	Bayes	-3.014	1.004	-1.513	0.100	1.004	7.981	4.031	3.529	-0.503	2.000	0.143
	Estimate	MLE	-2.989	1.000	-1.506	0.099	0.957	7.987	4.009	3.512	-0.501	1.835	0.137
	Average	Bayes	0.522	0.185	0.213	0.010		0.954	0.209	0.429	0.026		
	SE	MLE	0.512	0.182	0.209	0.010		0.924	0.203	0.415	0.026		
	Average	Bayes	0.428	0.151	0.169	0.008	0.132	0.794	0.175	0.351	0.022	0.359	0.176
	AbsError	MLE	0.425	0.151	0.168	0.008	0.133	0.789	0.174	0.350	0.022	0.364	0.176
	Coverage	Bayes	0.948	0.940	0.949	0.942		0.942	0.944	0.935	0.942		
		MLE	0.941	0.937	0.946	0.936		0.936	0.938	0.926	0.933		
	Average	Bayes	-3.017	1.003	-1.514	0.100	1.006	7.972	4.030	3.531	-0.503	2.002	0.426
	Estimate	MLE	-2.991	0.998	-1.507	0.099	0.960	7.978	4.009	3.515	-0.500	1.846	0.420
	Average	Bayes	0.513	0.182	0.209	0.010		0.922	0.202	0.414	0.026		
	SE	MLE	0.502	0.178	0.205	0.010		0.892	0.196	0.400	0.025		
Average	Bayes	0.419	0.148	0.167	0.008	0.133	0.772	0.171	0.339	0.021	0.360	0.179	
AbsError	MLE	0.416	0.147	0.166	0.008	0.134	0.767	0.169	0.338	0.021	0.365	0.179	
Coverage	Bayes	0.951	0.939	0.947	0.945		0.939	0.941	0.937	0.938			
	MLE	0.943	0.936	0.941	0.939		0.935	0.934	0.926	0.928			
-0.1	Average	Bayes	-3.014	1.005	-1.512	0.100	1.005	7.992	4.031	3.526	-0.503	2.000	-0.138
	Estimate	MLE	-2.988	1.000	-1.505	0.099	0.958	7.998	4.009	3.508	-0.501	1.835	-0.144
	Average	Bayes	0.522	0.185	0.213	0.010	0.166	0.954	0.209	0.429	0.026	0.453	0.224
	SE	MLE	0.512	0.182	0.209	0.010	0.188	0.924	0.203	0.415	0.026	0.129	0.134
	Average	Bayes	0.427	0.152	0.169	0.008	0.131	0.792	0.175	0.349	0.022	0.358	0.176
	AbsError	MLE	0.424	0.151	0.168	0.008	0.131	0.787	0.173	0.347	0.022	0.362	0.177
	Coverage	Bayes	0.946	0.945	0.950	0.939		0.939	0.940	0.937	0.935		
		MLE	0.941	0.940	0.945	0.931		0.931	0.936	0.930	0.928		
	Average	Bayes	-3.018	1.007	-1.511	0.100	1.006	8.000	4.030	3.522	-0.503	2.002	-0.418
	Estimate	MLE	-2.992	1.002	-1.504	0.099	0.960	8.007	4.009	3.505	-0.501	1.846	-0.425
	Average	Bayes	0.514	0.182	0.209	0.010		0.925	0.203	0.415	0.026		
	SE	MLE	0.503	0.178	0.205	0.010		0.893	0.197	0.401	0.025		
Average	Bayes	0.417	0.149	0.166	0.008	0.129	0.768	0.170	0.336	0.021	0.357	0.177	
AbsError	MLE	0.414	0.149	0.166	0.008	0.128	0.763	0.168	0.334	0.021	0.362	0.178	
Coverage	Bayes	0.946	0.944	0.946	0.944		0.940	0.942	0.937	0.934			
	MLE	0.942	0.941	0.944	0.937		0.936	0.930	0.926	0.927			

$\rho = 0.1 \rightarrow \sigma_{12} = 0.1414, \rho = 0.3 \rightarrow \sigma_{12} = 0.4243, \rho = -0.1 \rightarrow \sigma_{12} = -0.1414, \rho = -0.3 \rightarrow \sigma_{12} = -0.4243$

대가능도 추정량은 근사적으로 정규분포에 수렴하기 때문에 Wald 신뢰구간  $\hat{\theta} \pm z_{\alpha/2}SE(\hat{\theta})$ 의 포함확률은 근사적으로  $1 - \alpha$ 에 수렴하여야 한다. 만약 Wald 신뢰구간의 포함확률이 명목 신뢰수준(nominal confidence level)에 미치지 못하면 표준오차가 과소추정된 것으로 볼 수 있고, 이와 반대로 명목 신뢰수준보다 크다면 표준오차는 과대추정되었을 가능성이 높기 때문에 회귀계수의 경우, 표준오차 크기의 적절성을 평가하기 위해 신뢰구간의 포함확률 추정값을 구하였으며, 오차항 분산인 경우는 평균제곱오차(mean squared error)를 계산하였다.

모의실험에서 독립변수들은 각각  $x_{11} = 2U_1, x_{12} = 1(U_2 > 0.5), x_{13} = DU_1(25, 60), x_{21} = 5U_3, x_{22} = 1(U_4 > 0.5), x_{23} = DU_2(25, 60)$ 에 의해 독립적으로 생성하였다. 여기서  $U_i$ 와  $DU_i$ 는 연속형과 이산형 균등분포를 나타낸다. 또, 모수값은  $(\beta_{10}, \beta_{11}, \beta_{12}, \beta_{13}) = (-3, 1, -1.5, 0.1), (\beta_{20}, \beta_{21}, \beta_{22}, \beta_{23}) =$

(8, 4, 3.5, -0.5),  $\sigma_{11} = 1, \sigma_{22} = 2$ 와 같이 설정하였고  $\rho = -0.3, -0.1, 0.1, 0.3$ 에서 실험을 실시하였다. 이때,  $\sigma_{12}$ 는 -0.4243, -1.414, 0.1414, 0.4243의 값을 갖게 된다. 이러한 설정은 절단율의 조정을 위해  $\beta_{10}$ 와  $\beta_{20}$  값을 변경하였을 뿐, 다른 값들은 Wichitaksorn과 Choy (2011)와 동일하다.

$y_1$ 와  $y_2$  모두 대략 18% ~ 22%의 절단율을 갖는다. Wichitaksorn과 Choy (2011)는 표본크기  $n = 500$ 에서 62% ~ 65%의 절단율을 갖는다고 하는데, 베이지안과 최대가능도 추정의 비교는 소표본에서 만 의미가 있으므로 본 연구에서는 표본크기  $n = 100$ 을 정하였으며, 절단율은 계산시간을 고려하여 조절하였다. 모의실험 결과는 Table 3.2에 수록하였다.

모의실험에서 회귀계수의 최대가능도 추정값은 평균적으로 베이지안 추정값 보다 실제 모수에 가까운 경우가 많았으나 평균 절대오차(average absolute error) 값이 거의 같기 때문에 회귀계수 추정에서는 큰 차이가 없다고 할 수 있다. 이 결과는 Chib (1992)의 결론과도 유사하다. 그러나 최대가능도 추정의 표준오차가 베이지안 추정보다 항상 같거나 작기 때문에 최대가능도 추정이 효율적인 것으로 오해할 수 있다. Lee와 Choi (2014)에서는 이에 대해 최대가능도 추정의 표준오차를 이용한 신뢰구간이 명목 신뢰수준보다 현저히 낮기 때문에 최대가능도 추정의 표준오차가 과소 추정되었을 것으로 판단하고 있다. 본 연구의 모의실험에서도 베이지안 추정의 포함확률이 최대가능도 추정보다 명목 신뢰수준에 근사하기 때문에 최대가능도 추정의 표준오차는 과소 추정된 것으로 볼 수 있으며, 그 원인의 하나가 오차항 분산의 과소 추정에 있을 것으로 판단된다. 즉, 최대가능도 추정은 회귀계수의 추정과는 달리 오차항 분산을 현저히 과소 추정하고 있는 것을 확인할 수 있다.

#### 4. 결론

회귀모형에서 회귀계수 추정의 표준오차를 구하기 위해서는 오차항 분산의 추정값이 필요하다. 즉, 오차항 분산의 추정은 회귀계수에 대한 가설검정 및 구간추정에 영향을 미친 뿐 아니라, 독립변수의 주변효과(marginal effect; Greene, 2012, p.889) 추정에도 영향을 미치기 때문에 회귀분석에서 매우 중요한 부분이라고 할 수 있다. 모의실험에서 최대가능도 추정방법은 회귀계수를 효율적으로 추정하고 있으나 오차항 분산을 과소추정하는 경향이 있기 때문에 모수추정 이후의 분석에서는 효율성이 떨어지는 것을 확인할 수 있었다. 이와 비교하여 베이지안 추정방법은 회귀계수는 물론 오차항 분산의 추정에서도 실제값에 가깝게 추정하고 있으므로 토빗회귀모형의 추정에서 최대가능도 추정보다 선호된다고 하겠다. 모의실험이 비록 제한된 범위내에서 수행되었으나 이와 같은 결론을 내리기에 큰 무리가 없을 것이다.

#### References

- Amemiya, T. (1984). Tobit models: A survey, *Journal of Econometrics*, **24**, 3-61
- Bruno, G. (2004). Limited dependent panel models: A comparative analysis of classical and Bayesian inference among econometrics packages, *Computing in Economics and Finance, Society for Computational Economics*, [http://editorialexpress.com/cgi-bin/conference/download.cgi?db\\_name=SCE2004&paper\\_id=41](http://editorialexpress.com/cgi-bin/conference/download.cgi?db_name=SCE2004&paper_id=41).
- Chib, S. (1992). Bayesian inference in the Tobit censored regression model, *Journal of Econometrics*, **51**, 77-99.
- Cowles, M. K., Carlin, B. P. and Connett, J. E. (1996). Bayesian Tobit modeling of longitudinal ordinal clinical trial compliance data with nonignorable missingness, *Journal of American Statistical Association*, **91**, 86-98.
- Daniels, M. D. and Kass, R. E. (1999). Nonconjugate Bayesian estimation of covariance matrices and its use in hierarchical models, *Journal of American Statistical Association*, **94**, 1254-1263.
- Efron, B. and Hinkley, D. V. (1978). The observed versus expected information, *Biometrika*, **65**, 163-168.

- Greene, W. H. (1990). *Econometric Analysis*, Macmillan, New York.
- Greene, W. (2004a). The behavior of the maximum likelihood estimator of limited dependent variable models in the presence of fixed effects, *Econometrics Journal*, **7**, 98–119.
- Greene, W. (2004b). Fixed effects and bias due to the incidental parameters problem in the Tobit model, *Econometrics Reviews*, **23**, 125–147.
- Greene, W. (2012). *Econometric Analysis*. 7<sup>th</sup> edition, Pearson.
- Hamilton, B. H. (1999). HMO selection and medicare costs: Bayesian MCMC estimation of a robust panel data tobit model with survival, *Health Economics and Econometrics*, **8**, 403–414.
- Huang, C. J., Sloan, F. and Adamache, K. W. (1987). Estimation of seemingly unrelated Tobit regression via EM algorithm, *Journal of Business & Economic Statistics*, **5**, 425–430.
- Huang, H. C. (1999). Estimation of the SUR Tobit model via the MCECM algorithm, *Economics Letters*, **64**, 25–30.
- Huang, H. C. (2001). Bayesian analysis of the SUR Tobit model, *Applied Economics Letters*, **8**, 617–622.
- Jöreskog (2004). Multivariate censored regression, available at [www.ssicentral.com/lisrel/column12.htm](http://www.ssicentral.com/lisrel/column12.htm).
- Johnson, S. G. and Narasimhan, B. (2014). *Package Cubature*, <http://ab-initio.mit.edu/wiki/index.php/Cubature>.
- Lancaster, T. (2000). The incidental parameter problem since 1948, *Journal of Econometrics*, **95**, 391–413.
- Lee, S.C. and Choi, B. (2013). Bayesian interval estimation of Tobit regression model, *The Korean Journal of Applied Statistics*, **26**, 737–746.
- Lee, S.C. and Choi, B. (2014). Bayesian inference for censored panel regression model, *Communications for Statistical Applications and Methods*, **21**, 192–200.
- Louis, T. A. (1982). Finding the observed information matrix when using the EM algorithm, *Journal of the Royal Statistical Society: Series B*, **44**, 226–233.
- Natarajan, R. and Kass, R. R. (2000). Bayesian methods for generalized linear mixed models, *Journal of the American Statistical Association*, **95**, 227–237.
- Tanner, M. A. and Wong, W. H. (1987). The calculation of posterior distributions by data augmentation (with discussion), *Journal of the American Statistical Association*, **82**, 528–550.
- Tierney, L. and Kadane, J. B. (1986). Accurate approximations for posterior moments and marginal densities, *Journal of American Statistical Association*, **81**, 82–86.
- Wichitakorn, N. and Choy, S. T. B. (2011). Modeling dependence of seemingly unrelated Tobit model through copula: A Bayesian analysis, *Thailand Econometrics Society*, **3**, 6–19.

# SUR 토빗회귀모형에서 베이지안 추정과 최대가능도 추정의 비교

이승천<sup>a,1</sup> · 최병수<sup>b</sup>

<sup>a</sup>한신대학교 응용통계학과, <sup>b</sup>한성대학교 멀티미디어학과

(2014년 8월 13일 접수, 2014년 10월 7일 수정, 2014년 10월 7일 채택)

---

## 요약

Greene (2004a,b), Lee와 Choi (2014) 등의 연구에서 토빗 회귀모형의 최대가능도 추정은 표준오차를 과소추정한다는 것이 알려졌고, 그 원인의 하나는 오차항 분산의 과소 추정에 있다고 한다. 오차항 분산의 과소 추정은 회귀계수에 대한 가설 검정 및 구간추정에 영향을 미칠 뿐 아니라 독립변수들의 주변효과를 추정하는데에도 영향을 미치게 되므로 토빗 회귀모형에 대한 적절한 분석이 수행되려면 최대가능도 추정의 과소 추정 문제를 해결하여야 한다. 일반적으로 무정보 사전분포에 의한 베이지안 추론 방법은 빈도학파들이 요구하는 효율성을 갖는 경우가 많다. 본 연구에서도 무정보 사전분포에 의한 베이지안 추론을 적용하여, 베이지안 방법론이 SUR 토빗 회귀모형에서 최대가능도 추정의 과소 추정 문제를 해결할 수 있는 하나의 대안이 될 수 있다는 것을 보였다.

주요용어: SUR 토빗 회귀모형, 최대가능도 추정, EM 알고리즘, 베이지안 추정, 김스샘플링.

---

---

이 논문은 한신대학교 학술연구비 지원에 의하여 연구되었음.

이 논문은 한성대학교 학술연구비 지원에 의하여 연구되었음.

<sup>1</sup>교신저자: (447-791) 경기도 오산시 양산동 411, 한신대학교 응용통계학과. E-mail: seung@kss.or.kr