

다중 베이지요인에 의한 회귀모형 오차항의 자기상관 검정

한성실¹⁾ 김혜중²⁾

요약

본 논문은 회귀분석에서 오차항의 1차 자기상관 존재 여부 및 그 값을 검정하는 방법을 베이지안 접근법으로 제안하였다. 이 방법은 모수공간의 다중분할로 인해 얻어진 여러 가설들에 대한 다중결정문제를 다중 베이지요인에 관한 이론과 일반화 Savage-Dickey 밀도비를 이용한 사후확률 추정법을 합성하여 개발되었다. 이 방법은 기존의 검정법들에서 가능한 검정 뿐 아니라 이들이 해결할 수 없는 자기상관에 대한 다중결정문제에도 사용이 가능한데 그 효용성이 있다. 모의실험을 통하여 제안된 검정법의 유효성을 평가하였다.

1. 서론

일반적으로 회귀모형의 오차항들은 평균이 0이고, 등분산이며, 서로 독립이라는 가정 아래 분석이 이루어진다. 그러나, 많은 회귀모형의 응용에서 분석대상이 되는 자료들의 오차항이 상관관계를 가지는 경우가 있으며, 이러한 경우 오차항들이 자기상관(autocorrelation)을 가진다고 한다. 실제로 회귀모형의 오차항들 사이에 자기상관이 존재할 때 일반적인 최소자승법(least square method)을 사용하면 추정량의 유효성에 심각한 문제가 발생되므로 회귀모형을 유효하게 추정하기 위해서는 자기상관 존재 여부에 대한 검정이 필수적이다.

특히, 오차항들 사이에 존재하는 1차 자기상관을 검정하기 위한 방법으로 표본이론에서는 Durbin과 Watson(1951)에 의해 제안된 D 통계량과 Marr와 Quesenberry(1991)의 NU 통계량이 널리 사용되고 있다. 이러한 표본이론에 의한 자기상관 검정법들은 귀무가설 $H_0: \rho = 0$ 이 참이라는 가정 아래에서 도출된 검정통계량을 사용하고 있기 때문에 일반적인 귀무가설 $H_0: \rho = \rho_0$ ($-1 \leq \rho \leq 1$)의 검정에는 이들을 사용할 수 없는 문제점을 갖고 있다. 한편, 베이지안 이론의 관점에서 Kim과 Han(1998)은 베이지요인을 이용하여 자기상관 검정법을 제시하였고, 이 검정법이 귀무가설 $H_0: \rho = 0$ 뿐 아니라 일반적인 귀무가설 $H_0: \rho = \rho_0$ 를 모두 유효하게 검정할 수 있는 방법이 됨을 보였다. 그러나, 미지 모수 ρ 에 대한 다중결정문제는 그 연구가 이루어지지 않은 상태이다.

본 논문에서 고려할 베이지안 다중결정문제는 미지모수 ρ 에 대한 검정에서 귀무가설이 $H_0: \rho = \rho_0$ 와 같이 한 점에 대한 가설일 경우 대립가설에 해당하는 모수공간의 분할을 통해 $H_1: \rho \in R_1$, $H_2: \rho \in R_2$ 단, $R_1 = \{-1 \leq \rho < \rho_0\}$ 과 $R_2 = \{\rho_0 < \rho \leq 1\}$, 와 같이 검정의 목적에 따라 여러 개의 대립가설을 동시에 설정하여 이들을 귀무가설과 비교할 수 있는 다중결정기법을 뜻한다. 이 기법은 0-1 손실함수하에서 얻은 베이지규칙을 사용하며,

1) (100-715) 서울시 중구 필동 3가 26, 동국대학교 통계학과, 강사

2) (100-715) 서울시 중구 필동 3가 26, 동국대학교 통계학과, 교수

회귀모형에서 제외된 독립변수들이 오차항에 미치는 영향을 평가하는데 중요한 역할을 한다. 예를 들어, 검정 결과 $H_1: \rho > 0.5$ 이 사실이고, 회귀분석에서 결정계수 값이 작으면, 이는 모형에 포함 가능한 독립변수들 중에서 높은 양의 자기상관을 가진 독립변수를 모형에 포함시키는 것이 바람직한 변수선택임을 의미한다. 또한, 베이지안 검정의 관점에서 볼 때 ρ_0 에 대해 대칭적 사전확률분포만이 사용된 기존의 베이즈 검정법과는 달리 다중결정기법은 R_1 과 R_2 에 대한 사전확률분포를 분석자의 사전정보에 따라 ρ_0 에 대해 비대칭 또는 대칭과 같이 탄력적으로 설정할 수 있는 이론적인 근거를 제공하여 ρ 에 대한 사전정보를 보다 더 현실성 있게 반영하는 검정법이 될 것이다.

이를 위하여 본 논문은 Bertolino, Piccinato와 Racugno(1995)에 의해 제안된 다중 베이즈요인(multiple Bayes factor)의 수정 형태와 이를 일반화 Savage-Dickey 밀도비로 추정할 수 있는 방법을 개발하여 자기상관에 대해 별개의 검정을 가능하게 하는 베이지안 다중 결정 문제(multiple decision problem)의 한 기법을 제시하고자 한다.

2. 다중 베이즈요인을 이용한 가설검정법

한 개의 모수에 대한 검정을 고려한 Bertolino, Piccinato와 Racugno(1995)의 경우와 달리 따름모수 $\psi \in \Psi$ 를 포함하는 모수벡터 $\omega = (\theta, \psi)'$ 에 대한 모수공간 $\Omega = \Theta \times \Psi$ 에서의 베이즈 추론에서 한 점에 대한 귀무가설 $H_0: \theta = \theta_0$ (θ_0 는 주어진 값)의 검정을 고려하자. 또한, 실제로 검정하고자 하는 모수 θ 를 효과적으로 베이지안 검정하기 위해 대립가설을 분할(partition)하였다고 하자. 이 때 각 모수공간의 분할집합(partition set)에 대해 모수 θ 의 사전확률분포(prior probability distribution)가 θ_0 에 대해 비대칭적이면 다음의 방법으로 이 점을 고려하여야 한다.

모수공간 Θ 가 $(k+1)$ 개의 집합 $\{\Theta_0, \Theta_1, \Theta_2, \dots, \Theta_k; k > 1\}$ 으로 분할되면, 각 분할집합의 사전확률이

$$\pi_i = P(\omega \in \Theta_i \times \Psi), \quad i = 0, 1, \dots, k, \quad \sum_{i=0}^k \pi_i = 1, \quad (2.1)$$

일 때 이들의 사후확률은

$$P_i = P(\omega \in \Theta_i \times \Psi | x) = \frac{\Delta_i}{\sum_{j=0}^k \Delta_j}, \quad i = 0, 1, 2, \dots, k, \quad (2.2)$$

가 된다. 여기서, $\Theta_0 = \{\theta: \theta = \theta_0\}$ 이며, $\Delta_0 = \pi_0 \int_{\Psi} p_0(\psi) \ell(\theta_0, \psi) d\psi$, 그리고 $\Delta_j = \pi_j \int_{\Psi} \int_{\Theta_j} p_j(\theta, \psi) \ell(\theta, \psi) d\theta d\psi$, $j = 1, 2, \dots, k$, 이다.

따라서, 대립가설 아래에서 정의되는 모수공간의 각 분할집합 $\Theta_j \times \Psi$, $j = 1, 2, \dots, k$,에 대한 귀무가설의 모수공간 $\Theta_0 \times \Psi$ 의 부분 베이즈요인(partial Bayes factor)은

$$B_j = \frac{\int_{\Psi} p_0(\psi) \ell(\theta_0, \psi) d\psi}{\int_{\Psi} \int_{\Theta_j} p_j(\theta, \psi) \ell(\theta, \psi) d\theta d\psi}, \quad j = 1, 2, \dots, k, \quad (2.3)$$

로서 정의된다.

한편, 식 (2.3)에서 부분 베이지요인 $B_j, j = 1, 2, \dots, k$, 는 사전확률 $\pi_0, \pi_1, \dots, \pi_k$ 와 무관하며, 모수공간 $\Omega = \Theta \times \Psi$ 에 대한 사후확률은

$$P(\omega \in \Theta \times \Psi | x) = \sum_{i=0}^k P_i = 1 \quad (2.4)$$

와 같이 표현된다.

식 (2.2)와 식 (2.4)를 이용하여 귀무가설의 사후확률을 부분 베이지요인과 사전확률로서 표현하면

$$P_0 = P(H_0 | x) = \left(1 + \frac{1}{\pi_0} \sum_{j=1}^k \frac{\pi_j}{B_j}\right)^{-1} > \prod_{j=1}^k \left(1 + \frac{\pi_j}{\pi_0 B_j}\right)^{-1} \quad (2.5)$$

이며, 대립가설 아래에서의 모수공간의 각 분할집합 $\Theta_1 \times \Psi, \Theta_2 \times \Psi, \dots, \Theta_k \times \Psi$ 에 의해 정의되는 가설을 각각 H_1, H_2, \dots, H_k 로 나타내면, 각 가설에 대한 사후확률은

$$\begin{aligned} P_j &= P(H_j | x) \quad (2.6) \\ &= \left\{1 + \frac{B_j}{\pi_j} \left(\pi_0 + \sum_{\substack{s=1 \\ s \neq j}}^k \frac{\pi_s}{B_s}\right)\right\}^{-1} > \left(1 + \frac{\pi_0}{\pi_j} B_j\right)^{-1} \prod_{\substack{s=1 \\ s \neq j}}^k \left(1 + \frac{\pi_s}{\pi_j} \frac{B_j}{B_s}\right)^{-1}, \quad j = 1, 2, \dots, k, \end{aligned}$$

이다.

이렇게 계산된 각 가설의 사후확률 $P_\ell (\ell = 0, \dots, k)$ 을 다중결정문제에 사용할 때 가설 H_ℓ 를 채택할 때 발생하는 위험(risk)은

$$R(H_\ell | x) = \sum_{\ell=0}^k L(H_\ell, H_\ell) P(H_\ell | x)$$

이다. 여기서, $L(H_\ell, H_\ell)$ 는 다중결정문제에서 발생하는 손실을 나타낸다. 만약, 손실함수를 베이시안 검정문제에서 일반적으로 사용되는 0-1 손실함수

$$L(H_\ell, H_\ell) = 1 - \delta(H_\ell, H_\ell),$$

를 사용하면, $\max_\ell P(H_\ell | x)$ 에 해당하는 가설을 채택하는 것이 베이지규칙(Bayes rule)을 따르는 다중결정방법임을 알 수 있다. 여기서, $\delta(H_\ell, H_\ell)$ 의 값은 $H_\ell = H_\ell$ 일 경우 1이고, $H_\ell \neq H_\ell$ 일 경우 0이다.

따라서, 식 (2.6)에서 구해진 사후확률 값을 이용하면, 귀무가설이 기각되는 경우 각 분할집합에 의해서 정의되는 k 개 대립가설 $H_j: \theta \in \Theta_j, j = 1, 2, \dots, k$, 에 대한 다중결정문제를 수행할 수 있게 된다. 위와 같은 검정은 식 (2.3)에서 정의된 k 개 부분 베이지요인을 쉽게 계산할 수 있음을 전제로 하며, 이들의 벡터 형태인

$$MB = (B_1, B_2, \dots, B_k)'$$

을 다중 베이지요인이라 한다. 한편, 식 (2.5)와 식 (2.6)의 부등호 양변을 비교하면 두 식 모두 좌변이 우변의 값보다 크다. 이는 분할된 모수공간하에서 정의된 가설들을 동시에 사용

하여 구한 사후확률로 가설을 검정하는 것이 개별적으로 계산된 사후확률에 의한 검정보다 더 검정의 정도가 높음을 나타낸다.

이와 같이 다중 베이지요인을 이용한 검정법은 식 (2.5)와 식 (2.6)을 통해 종래의 검정법과는 달리 대립가설 아래에서 정의되는 여러 모수공간들의 특성에 따라 여러 개의 가설을 설정하여 이들을 동시에 검정할 수 있어 기존의 방법에 비해 검정의 정도를 높일 수 있는 방법이다. 또한, 이 방법은 대립가설 아래에서 정의되는 모수공간의 각 분할집합에 따라 검정하고자 하는 모수의 사전확률분포를 각각 다르게 설정할 수 있는 특징을 가진다. 이러한 점에서 볼 때, 다중 베이지요인을 이용한 검정법은 베이지요인에 의한 기존의 검정방법보다 더 유효하고 현실적인 검정법이라 할 수 있다.

3. 중회귀모형에서 자기상관모수의 검정

3.1. 다중 베이지요인

오차항이 다음과 같은 1차 자기회귀과정을 가지는 중회귀모형을 고려하자.

$$\tilde{y} = \alpha + \tilde{X}\beta + u, \quad u = \rho u_{-1} + \varepsilon. \quad (3.1)$$

모형 (3.1)에서 $\tilde{y} = (\tilde{y}_1, \tilde{y}_2, \dots, \tilde{y}_T)'$ 와 $\tilde{X} = (\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_p)$ 는 각각 $(T \times 1)$ 종속변수 벡터와 $(T \times p)$ 독립변수 행렬을 나타낸다. 여기서, $\tilde{y}_t, t = 1, 2, \dots, T$, 는 t 번째 종속변수를 나타내며, $\tilde{x}_k = (\tilde{x}_{1k}, \tilde{x}_{2k}, \dots, \tilde{x}_{Tk})'$ 는 k 번째 독립변수 벡터를 나타낸다. 단, $k = 1, 2, \dots, p$, 이다. 그리고, α 는 절편을 나타내는 $(T \times 1)$ 벡터이며, $\beta = (\beta_1, \beta_2, \dots, \beta_p)'$ 는 회귀계수를 나타내는 $(p \times 1)$ 벡터이다. 또한, u 와 u_{-1} 은 1차 자기상관을 가지는 오차항 벡터로서

$$u = (u_1, u_2, \dots, u_T)', \quad u_{-1} = (u_0, u_1, \dots, u_{T-1})'$$

와 같이 정의되고, ρ 는 인접하는 두 시점에서의 오차항 u_t 와 u_{t-1} 사이의 상관계수를 나타내는 자기상관모수이며, $\varepsilon = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_T)'$ 는 서로 독립이고, 각각의 $\varepsilon_t, t = 1, 2, \dots, T$, 가 평균이 0, 분산이 $1/\tau$ 인 정규분포를 따르는 $(T \times 1)$ 확률변수 벡터이다. 식 (3.1)에서 $\rho = 0$ 이면 일반적인 중회귀모형과 같다.

본 연구에서는 식 (3.1)의 모형에서 정의된 1차 자기상관을 가지는 중회귀모형에 대해서 절편을 고려하지 않은 대체모형인

$$y = X\beta + u, \quad u = \rho u_{-1} + \varepsilon \quad (3.2)$$

를 고려하였다. 여기서, $y = (y_1, y_2, \dots, y_T)'$ 이며, 각각의 y_t 는 t 번째 종속변수의 관측치로서 $y_t = \tilde{y}_t - \bar{y}$ 이고, $\bar{y} = (1/T) \sum_{t=1}^T \tilde{y}_t$ 이다. 또한, $X = (x_1, x_2, \dots, x_p)$ 는 $(T \times p)$ 독립변수 행렬로 $x_k = (x_{1k}, x_{2k}, \dots, x_{Tk})'$ 는 k 번째 독립변수의 관측치 벡터이고, $x_{tk} = \tilde{x}_{tk} - \bar{x}_k, t = 1, 2, \dots, T; k = 1, 2, \dots, p$, 이며, \bar{x}_k 는 k 번째 독립변수 벡터의 평균으로 $\bar{x}_k = (1/T) \sum_{t=1}^T \tilde{x}_{tk}$ 이다.

식 (3.2)의 모형에서 $\mathbf{u} = \mathbf{y} - \mathbf{X}\beta$ 이고, $\mathbf{u}_{-1} = \mathbf{y}_{-1} - \mathbf{X}_{-1}\beta$ 이므로 이를 이용하면 식 (3.2)는

$$\mathbf{y} = \rho\mathbf{y}_{-1} + (\mathbf{X} - \rho\mathbf{X}_{-1})\beta + \varepsilon \quad (3.3)$$

와 같이 나타낼 수 있다. 여기서, \mathbf{y}_{-1} 과 \mathbf{X}_{-1} 은 각각

$$\mathbf{y}_{-1} = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_{T-1} \end{pmatrix}, \quad \mathbf{X}_{-1} = \begin{pmatrix} x_{01} & x_{02} & \cdots & x_{0p} \\ x_{11} & x_{12} & \cdots & x_{1p} \\ \vdots & \vdots & \cdots & \vdots \\ x_{(T-1)1} & x_{(T-1)2} & \cdots & x_{(T-1)p} \end{pmatrix}$$

을 나타낸다.

식 (3.3)의 모형에서 초기값 y_0 가 알려져 있다고 가정하면, 완전한 관측값의 집합인 $\{x_{tk}, y_t; t = 1, 2, \dots, T, k = 1, 2, \dots, p\}$ 는 $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_T$ 에 의해서 결정되므로 \mathbf{y} 의 결합확률밀도함수(joint probability density function)는

$$f(\mathbf{y}|\rho, \tau, \beta) \propto \tau^{\frac{T}{2}} \cdot \exp\left\{-\frac{\tau}{2}[\mathbf{y} - \rho\mathbf{y}_{-1} - (\mathbf{X} - \rho\mathbf{X}_{-1})\beta]'[\mathbf{y} - \rho\mathbf{y}_{-1} - (\mathbf{X} - \rho\mathbf{X}_{-1})\beta]\right\}, \\ -\infty < \beta_k < \infty (k = 1, 2, \dots, p), \tau > 0, |\rho| \leq 1, \quad (3.4)$$

이 된다(Wonacott와 Wonacott, 1979).

위에서 정의된 식 (3.2)의 모형에서 자기상관모수를 검정하기 위해서 세운 다중가설

$$H_0 : \rho = \rho_0, \quad H_1 : \rho > \rho_0, \quad H_2 : \rho < \rho_0, \quad (3.5)$$

을 고려해 보자. 여기서, 자기상관모수 ρ 의 모수공간은 $R = \{\rho; -1 \leq \rho \leq 1\}$ 로 제한하였고, 설정된 가설에 따라 자기상관모수 ρ 의 모수공간 R 을 $R_0 = \{\rho = \rho_0\}$, $R_1 = \{-1 \leq \rho < \rho_0\}$ 과 $R_2 = \{\rho_0 < \rho \leq 1\}$ 로 분할한다.

이와 같이 대립가설하에서 설정된 모수공간의 두 분할집합 R_1 과 R_2 에 대해서 모수 ρ 가 서로 다른 사전확률분포를 가질 때 다중 베이지요인에 의한 추론이 필요하다.

식 (3.4)에서 정의된 미지모수 $\theta = (\rho, \beta, \tau)$ 에 대한 모수공간 Θ 의 세 분할집합 $\Theta_0 = R_0 \times \Omega^p \times \Xi$ 와 $\Theta_1 = R_1 \times \Omega^p \times \Xi$, $\Theta_2 = R_2 \times \Omega^p \times \Xi$ 에서의 사전확률을 각각 $\pi_i, i = 0, 1, 2$; $\sum_{i=0}^2 \pi_i = 1$, 라 하고, 사후확률을 각각 $P_i, i = 0, 1, 2$; $\sum_{i=0}^2 P_i = 1$, 라 하면

$$P_i = P(\theta \in R_i \times \Omega^p \times \Xi | \mathbf{y}) = \frac{\Delta_i}{\sum_{i=0}^2 \Delta_i}, \quad i = 0, 1, 2, \quad (3.6)$$

이 된다. 여기서, $\Delta_0 = \pi_0 \int_{\tau} \int_{\beta_p} \cdots \int_{\beta_1} p_0(\tau, \beta) \ell(\rho_0, \tau, \beta) d\beta_1 \cdots d\beta_p d\tau$, $\Delta_j = \pi_j \int_{R_j} \int_{\tau} \int_{\beta_p} \cdots \int_{\beta_1} p_R^{(j)}(\rho, \tau, \beta) \ell(\rho, \tau, \beta) d\beta_1 \cdots d\beta_p d\tau d\rho, j = 1, 2$, $\ell(\rho, \tau, \beta)$ 는 ρ, β 와 τ 의 결합우도함수, $p_0(\tau, \beta)$ 는 R_0 에서의 결합사전밀도함수(joint prior density function), 그리고 $p_R^{(j)}(\rho, \tau, \beta), j = 1, 2$, 는 각각 R_1 과 R_2 에서의 결합사전밀도함수를 나타낸다.

다중 베이지요인에 의한 검정법은 검정하고자 하는 미지모수 ρ 의 사전확률분포 형태가 대립가설 아래에서 정의되는 모수공간의 각 분할집합에 따라 다르게 나타나는 경우에 유용하게 사용되는 방법이므로 대립가설의 각 분할집합에 대해서 서로 다른 조건부 사전확률분포(conditional prior probability distribution)를 고려해 보자.

귀무가설 아래에서의 (τ, β) 의 사전밀도함수를 $p_0(\tau, \beta)$ 라 하면, 대립가설하에서 정의되는 모수공간의 두 분할집합 $R_1 = \{-1 \leq \rho < \rho_0\}$ 과 $R_2 = \{\rho_0 < \rho \leq 1\}$ 에 대한 결합사전밀도함수는

$$p_R^{(j)}(\rho, \tau, \beta) = p_R^{(j)}(\rho)p(\tau, \beta|\rho)I_{R_j \times \Omega^p \times \Xi}(\rho, \tau, \beta), \quad j = 1, 2, \quad (3.7)$$

와 같이 나타낼 수 있으며, ρ 와 (τ, β) 가 사전독립(prior independence)이라고 가정하면

$$p_R^{(j)}(\rho, \tau, \beta) = p_R^{(j)}(\rho)p(\tau, \beta)I_{R_j \times \Omega^p \times \Xi}(\rho, \tau, \beta), \quad j = 1, 2, \quad (3.8)$$

이 된다.

이제 (τ, β) 의 사전확률분포는 모호한 사전분포(vague prior distribution)로 가정하고, ρ 의 사전확률분포는 모수공간 $R = \{\rho; -1 \leq \rho \leq 1\}$ 에서 대립가설의 두 분할에 따라 다음과 같이 서로 다르게 정의되는 절단된 정규분포(truncated normal distribution)를 따른다고 가정하자(Zellner 1971, p90 참조).

$$p(\tau, \beta) \propto \tau^{1/2}, \quad (3.9)$$

$$p_R^{(j)}(\rho) \propto \exp\left\{-\frac{(\rho - \rho_0)^2}{2r_j}\right\}I_{R_j}(\rho), \quad j = 1, 2, \quad (3.10)$$

여기서, $r_j, j = 1, 2$, 는 대립가설의 두 분할집합에 대해서 다르게 정의되는 사전확률분포에서의 분산을 나타내는 초모수(hyper-parameter)로서 항상 양의 값을 가지며, R_1 과 R_2 에 대해서 자기상관모수 ρ 는 각각 평균이 ρ_0 이고, 분산이 $r_j, j = 1, 2$, 인 정규분포가 $-1 \leq \rho \leq 1$ 에서 절단된 정규분포를 따른다. 이 때 r_j 값은 절단된 정규분포의 분산식(Johnson et. al. 1994, p158)과 ρ 에 대한 사전정보로부터 구해진다.

따라서, 식 (3.4)에서 정의된 모수들의 결합우도함수와 식 (3.9) 및 식 (3.10)에서 정의된 사전확률분포를 이용하여 (ρ, τ, β) 에 대한 결합사후밀도함수(joint posterior density function)를 구하면

$$\begin{aligned} & p_R^{(j)}(\rho, \tau, \beta|\mathbf{y}) \\ & \propto \tau^{\frac{T-3}{2}-1} \exp\left\{-\frac{\tau}{2}[\mathbf{y} - \rho\mathbf{y}_{-1} - (\mathbf{X} - \rho\mathbf{X}_{-1})\beta]'\left[\mathbf{y} - \rho\mathbf{y}_{-1} - (\mathbf{X} - \rho\mathbf{X}_{-1})\beta\right] \right. \\ & \quad \left. - \frac{(\rho - \rho_0)^2}{2r_j}\right\} I_{R_j}(\rho), \quad \text{for } \rho \in R_j, j = 1, 2, \quad (3.11) \end{aligned}$$

와 같고, 식 (3.11)을 이용하여 중회귀모형에서 자기상관모수를 검정하기 위한 부분 베이지요인을

$$B_j = \frac{\int_{\beta_p} \cdots \int_{\beta_1} \int_{\tau} \ell(\rho_0, \beta, \tau) p_0(\beta, \tau) d\tau d\beta_1 \cdots d\beta_p}{\int_{R_j} \int_{\beta_p} \cdots \int_{\beta_1} \int_{\tau} \ell(\rho, \beta, \tau) p_R^{(j)}(\rho, \beta, \tau) d\tau d\beta_1 \cdots d\beta_p d\rho}, \quad j = 1, 2, \quad (3.12)$$

와 같이 얻을 수 있다. 그러나, 식 (3.12)에서 정의된 부분 베이즈요인을 실제로 계산하는 문제는 복잡하다. 따라서, 부분 베이즈요인 B_1 과 B_2 를 Verdinelli와 Wasserman(1995)이 제안한 일반화 Savage-Dickey 밀도비(generalized Savage-Dickey density ratio)의 형태로 변환시키고, Gibbs 추출법에 의하여 자기상관모수의 존재 여부를 검정하기 위한 부분 베이즈요인을 근사적으로 추정하고자 한다.

3.2. 부분 베이즈요인 추정

식 (3.12)에서 $\int_{R_j} \int_{\beta_p} \cdots \int_{\beta_1} \int_{\tau} \ell(\rho, \beta, \tau) p_R^{(j)}(\rho, \beta, \tau) d\tau d\beta_1 \cdots d\beta_p d\rho$ 를 m_j 로 나타내고 분자와 분모에 $p_R^{(j)}(\rho_0, \beta, \tau | \mathbf{y})$ 를 곱하면

$$B_j = p_R^{(j)}(\rho_0 | \mathbf{y}) \int_{\beta_p} \cdots \int_{\beta_1} \int_{\tau} \frac{\ell(\rho_0, \beta, \tau) p_0(\beta, \tau) p(\beta, \tau | \rho_0, \mathbf{y})}{p_R^{(j)}(\rho_0, \beta, \tau | \mathbf{y}) m_j} d\tau d\beta_1 \cdots d\beta_p, \quad j = 1, 2 \tag{3.13}$$

와 같이 쓸 수 있다. 여기서, $p_R^{(j)}(\rho | \mathbf{y})$, $j = 1, 2$, 는 모수공간 $R = \{\rho; -1 \leq \rho \leq 1\}$ 에서의 대립가설의 두 분할집합에 대한 ρ 의 주변사후밀도함수(marginal posterior density function)이다. 이것은 식 (3.11)을 τ 와 $\beta_1, \beta_2, \dots, \beta_p$ 에 대해서 차례로 적분함으로써 구할 수 있고, 아식의 ρ 값에 ρ_0 를 대입시켜 $p_R^{(j)}(\rho_0 | \mathbf{y})$ 를 구한다, 또한, $\rho = \rho_0$ 인 경우 ρ 와 β, τ 의 절단된 결합사후밀도함수는 $p_R^{(j)}(\rho_0, \beta, \tau | \mathbf{y}) = \ell(\rho_0, \beta, \tau) p_R^{(j)}(\rho_0, \beta, \tau) / m_j$ 로 나타낼 수 있으므로, 이것을 식 (3.13)에 대입하여 정리하면 부분 베이즈요인은

$$B_j = \frac{p_R^{(j)}(\rho_0 | \mathbf{y})}{p_R^{(j)}(\rho_0)} E \left[\frac{p_0(\beta, \tau)}{p(\beta, \tau | \rho_0)} \right], \quad j = 1, 2, \tag{3.14}$$

와 같이 나타낼 수 있으며, 식 (3.14)에서 $E[\cdot]$ 는 조건부 사후확률밀도함수 $p(\beta, \tau | \rho_0, \mathbf{y})$ 에 대한 기대값을 나타낸다(Kim, 1999 참조).

한편, 식 (3.8)에 의해 $p(\beta, \tau | \rho_0) = p_0(\beta, \tau)$ 이 성립되므로 식 (3.14)에서 정의된 부분 베이즈요인 B_j 는

$$B_j = \frac{p_R^{(j)}(\rho_0 | \mathbf{y})}{p_R^{(j)}(\rho_0)}, \quad j = 1, 2, \tag{3.15}$$

와 같이 간단하게 나타낼 수 있다. 여기서, $p_R^{(j)}(\rho)$, $j = 1, 2$, 는 각 대립가설의 두 분할집합에 대한 사전확률밀도함수로서

$$p_R^{(1)}(\rho) = \frac{(2\pi r_1)^{-1/2} \exp\{- (\rho - \rho_0)^2 / 2r_1\}}{\int_{-1}^{\rho_0} (2\pi r_1)^{-1/2} \exp\{- (\rho - \rho_0)^2 / 2r_1\} d\rho}, \tag{3.16}$$

$$p_R^{(2)}(\rho) = \frac{(2\pi r_2)^{-1/2} \exp\{- (\rho - \rho_0)^2 / 2r_2\}}{\int_{\rho_0}^1 (2\pi r_2)^{-1/2} \exp\{- (\rho - \rho_0)^2 / 2r_2\} d\rho} \tag{3.17}$$

이 된다. 식 (3.16)과 식 (3.17)에서 분모를 적분하여 간단히 나타내고 $\rho = \rho_0$ 를 대입하여 정리하면

$$p_R^{(1)}(\rho_0) = \frac{1}{\sqrt{2\pi r_1}} \left[\Phi(0) - \Phi\left(\frac{-1 - \rho_0}{\sqrt{r_1}}\right) \right]^{-1},$$

$$p_R^{(2)}(\rho_0) = \frac{1}{\sqrt{2\pi r_2}} \left[\Phi\left(\frac{1 - \rho_0}{\sqrt{r_2}}\right) - \Phi(0) \right]^{-1}$$

을 얻을 수 있다. 단, $\Phi(\cdot)$ 는 표준정규분포의 누적확률분포함수를 나타낸다.

따라서, 위 식을 이용하면 식 (3.15)에서 정의된 부분 베이지요인은 각각

$$B_1 = \sqrt{2\pi r_1} \left[\Phi(0) - \Phi\left(\frac{-1 - \rho_0}{\sqrt{r_1}}\right) \right] \cdot p_R^{(1)}(\rho_0 | \mathbf{y}), \quad (3.18)$$

$$B_2 = \sqrt{2\pi r_2} \left[\Phi\left(\frac{1 - \rho_0}{\sqrt{r_2}}\right) - \Phi(0) \right] \cdot p_R^{(2)}(\rho_0 | \mathbf{y}) \quad (3.19)$$

로 나타낼 수 있다.

이와 같이 얻어진 부분 베이지요인에는 비폐쇄형(unclosed form)인 $p_R^{(j)}(\rho | \mathbf{y})$, $j = 1, 2$,가 존재하여 이들을 직접 계산하기에는 어려움이 있다. 그래서, Gibbs 추출법을 이용하여 이들을 근사적으로 추정하고자 한다.

오차항이 1차 자기회귀과정을 가지는 중회귀모형 아래에서 ρ 와 β , τ 사이의 조건부 사후확률분포를 유도하면 Gibbs 추출법에 필요한 모든 조건부 사후확률분포들은

$$\tau | \beta, \rho, \mathbf{y} \sim G(\alpha^{**}, \beta^{**}),$$

$$\beta | \tau, \rho, \mathbf{y} \sim N(B_1^{-1}(\rho)B_2(\rho), \tau^{-1}B_1^{-1}(\rho)),$$

이며, $\rho \in R_j$ 에 대해

$$\rho | \beta, \tau, \mathbf{y} \sim N_{R_j} \left(\frac{r_j B_2(\beta) + \rho_0 \tau^{-1}}{r_j B_1(\beta) + \tau^{-1}}, \frac{r_j \tau^{-1}}{r_j B_1(\beta) + \tau^{-1}} \right), \quad j = 1, 2,$$

로써 영역이 R_j 인 절단정규분포를 따른다. 또한, τ 와 ρ 가 주어진 경우 β 에 대한 조건부 사후확률분포는 평균벡터가 $B_1^{-1}(\rho)B_2(\rho)$ 이고, 분산-공분산 행렬(variance-covariance matrix)이 $\tau B_1^{-1}(\rho)$ 인 다변량 정규분포(multivariate normal distribution)이 된다.

여기서, $\alpha^{**} = \frac{T+3}{2}$, $\beta^{**} = 2 / [(\mathbf{y} - \rho \mathbf{y}_{-1} - (\mathbf{X} - \rho \mathbf{X}_{-1})\beta)]' [(\mathbf{y} - \rho \mathbf{y}_{-1} - (\mathbf{X} - \rho \mathbf{X}_{-1})\beta)]$,

$$B_1(\rho) = (\mathbf{X} - \rho \mathbf{X}_{-1})' (\mathbf{X} - \rho \mathbf{X}_{-1}), \quad B_2(\rho) = (\mathbf{X} - \rho \mathbf{X}_{-1})' (\mathbf{y} - \rho \mathbf{y}_{-1}),$$

$$B_1(\beta) = (\mathbf{y}_{-1} - \mathbf{X}_{-1}\beta)' (\mathbf{y}_{-1} - \mathbf{X}_{-1}\beta), \quad B_2(\beta) = (\mathbf{y}_{-1} - \mathbf{X}_{-1}\beta)' (\mathbf{y} - \mathbf{X}\beta),$$

$$R_1 = \{-1 \leq \rho < \rho_0\} \text{과 } R_2 = \{\rho_0 < \rho \leq 1\} \text{이다.}$$

위에서 유도한 조건부 사후확률분포를 이용하여 Gibbs 추출법으로 τ 와 β 및 ρ 의 사후표본을 추출할 수 있다. 특히, 절단정규분포를 따르는 ρ 의 사후표본은 Devroy(1986)가 제안한 one-for-one 추출방법을 사용하였다.

이와 같은 방법으로 추출된 사후표본을 이용하면 $p_R^{(j)}(\rho_0 | \mathbf{y})$ 는

$$\hat{p}_R^{(j)}(\rho_0 | \mathbf{y}) = \frac{1}{N} \sum_{k=1}^N p_R^{(j)}(\rho_0 | \beta^{(km)}, \tau^{(km)}, \mathbf{y}), \quad j = 1, 2, \quad (3.20)$$

으로 추정된다. 단, $\beta^{(km)} = (\beta_1^{(km)}, \beta_2^{(km)}, \dots, \beta_p^{(km)})$

따라서, 식 (3.18)과 식 (3.19)에서 정의된 부분 베이지요인의 추정은

$$\hat{B}_1 = \sqrt{2\pi r_1} \left\{ \Phi(0) - \Phi\left(\frac{-1 - \rho_0}{\sqrt{r_1}}\right) \right\} \left\{ \frac{1}{N} \sum_{k=1}^N p_R^{(1)}(\rho_0 | \beta^{(km)}, \tau^{(km)}, \mathbf{y}) \right\}, \quad (3.21)$$

$$\hat{B}_2 = \sqrt{2\pi r_2} \left\{ \Phi\left(\frac{1 - \rho_0}{\sqrt{r_2}}\right) - \Phi(0) \right\} \left\{ \frac{1}{N} \sum_{k=1}^N p_R^{(2)}(\rho_0 | \beta^{(km)}, \tau^{(km)}, \mathbf{y}) \right\} \quad (3.22)$$

으로 이루어진다. 식 (3.21)과 식 (3.22)에 의해 추정된 부분 베이지요인 값과 사전확률 값이 주어지면 귀무가설 및 대립가설의 분할집합에 대한 사후확률을 구할 수 있다.

본 연구에서는 ρ 의 모수공간 $R = \{\rho; -1 \leq \rho \leq 1\}$ 을 R_0, R_1 및 R_2 로 분할하였으므로 세 분할집합에 대한 사전확률이 (A) $\pi_0 = 1/2, \pi_1 = \pi_2 = 1/4$ 인 경우와 (B) $\pi_0 = \pi_1 = \pi_2 = 1/3$ 인 경우에 대해 각각 사후확률 P_0, P_1 및 P_2 를 계산하면 다음과 같다.

(A) $\pi_0 = \frac{1}{2}, \pi_1 = \pi_2 = \frac{1}{4}$ 인 경우 :

$$P_0 = P(H_0 | \mathbf{y}) = \left(1 + \frac{1}{2} \left(\frac{1}{\hat{B}_1} + \frac{1}{\hat{B}_2} \right) \right)^{-1}, \quad (3.23)$$

$$P_1 = P(H_1 | \mathbf{y}) = \left(1 + 2\hat{B}_1 + \frac{\hat{B}_1}{\hat{B}_2} \right)^{-1}, \quad (3.24)$$

$$P_2 = P(H_2 | \mathbf{y}) = \left(1 + 2\hat{B}_2 + \frac{\hat{B}_2}{\hat{B}_1} \right)^{-1}. \quad (3.25)$$

(B) $\pi_0 = \pi_1 = \pi_2 = \frac{1}{3}$ 인 경우 :

$$P_0 = P(H_0 | \mathbf{y}) = \left(1 + \frac{1}{\hat{B}_1} + \frac{1}{\hat{B}_2} \right)^{-1}, \quad (3.26)$$

$$P_1 = P(H_1 | \mathbf{y}) = \left(1 + \hat{B}_1 + \frac{\hat{B}_1}{\hat{B}_2} \right)^{-1}, \quad (3.27)$$

$$P_2 = P(H_2 | \mathbf{y}) = \left(1 + \hat{B}_2 + \frac{\hat{B}_2}{\hat{B}_1} \right)^{-1}. \quad (3.28)$$

이와 같이 구해진 각 가설에 대한 사후확률 값을 비교하여 가장 큰 사후확률 값을 가지는 가설을 채택함으로써 식 (3.5)에서 설정된 다중 가설을 베이지규칙하에서 동시에 검정할 수 있다(2장 참조). 한편, 식 (3.10)의 ρ 의 사전확률밀도함수를 ρ_0 에 대해 대칭이 되도록 가정하면(즉, $r_1 = r_2$ 인 경우), $H_0: \rho = \rho_0$ 대 $H_A: \rho \neq \rho_0$ 에 대한 일반적인 베이지요인 B 는

$$B = \frac{1}{\frac{1}{B_1} + \frac{1}{B_2}}$$

의 관계에 있음을 식 (3.12)로부터 쉽게 유도할 수 있다.

4. 모의실험

이 장에서는 제안된 Gibbs 알고리즘에 의해 추정되는 부분 베이지요인이 자기상관모수를 유효하게 검정하는지를 평가하기 위해서 모의실험하였다. 모의실험은 다음의 중회귀모형하에서 이루어졌다.

$$y_i = 2x_{1i} + 3x_{2i} + 5x_{3i} + u_i, \quad u_i = \rho u_{i-1} + \epsilon_i,$$

단, $u_0 = 0$, $\epsilon_i \stackrel{iid}{\sim} N(0, 1)$, $x_{ji} \stackrel{iid}{\sim} U(0, 1)$, $j = 1, 2, 3$.

위 모형하에서 주어진 ρ 의 값($\rho = 0, \pm 0.3, \pm 0.5, \pm 0.7, \pm 0.9$)에 따라 y_i 와 x_{ji} , $j = 1, 2, 3$, 값들을 각각 20개씩 발생시켜 얻은 자료 (자료의 발생에는 초기값 $u_0 = 0$ 의 영향을 제거시키기 위해서 처음 발생된 100개의 자료는 제외시키고, 101번째에서 120번째에 발생된 자료를 사용하였음)를 가지고 부분 베이지요인 값을 추정하여 $H_0 : \rho = \rho_0$, $H_1 : -1 \leq \rho < \rho_0$, $H_2 : \rho_0 < \rho \leq 1$, 를 검정하는 절차를 200번씩 반복 실험하였다. 이 결과를 주어진 ρ_0 값에 따라 표 4.1($\rho_0 = 0$ 인 경우), 표 4.2($\rho_0 = 0.5$ 인 경우), 표 4.3($\rho_0 = -0.5$ 인 경우)에 구분하여 나타내었다.

먼저 Best, Cowles와 Vines(1996)에 의해 개발된 S-PLUS함수인 CODA 함수를 사용하여 제안된 알고리즘에 의해 생성된 Gibbs 사후표본의 수렴성을 검토하였다. 이러한 Gibbs 사후표본의 수렴성을 검증하기 위한 여러 측도들이 많은 학자들에 의해 제안되었으며, 본 연구에서는 Gelman과 Rubin(1992)에 의해 제안된 방법을 사용하였고, 추정된 모수들 사이의 체인별 자기상관을 검토하였다. 그 결과 알고리즘을 1000번 반복하면 Gibbs 사후표본이 수렴하는 것으로 나타났으며, 자기상관계수도 시차 5 이후에는 0이 됨을 보였다. 이를 토대로 1001번째 사후표본부터 Gibbs 추출열을 5로 하여 크기가 1000개인 Gibbs 표본을 구하였다.

이와 같은 절차에 의해 얻어진 Gibbs 표본을 이용하여 3장에서 정의된 각 부분 베이지요인을 추정하고, 추정된 부분 베이지요인과 모수공간의 각 분할집합에 대한 두 가지 사전 확률의 가정, (A) $\pi_0 = 1/2$, $\pi_1 = \pi_2 = 1/4$ 와 (B) $\pi_0 = \pi_1 = \pi_2 = 1/3$ 에 대해서 가설 H_0 , H_1 및 H_2 에 대한 사후확률을 구하여 제안된 검정법의 유효성을 검증하였다.

표 4.1은 중회귀모형 아래에서 ρ 의 사전확률분포에서의 분산이 $r_1 > r_2$ 인 경우, ρ 의 변화에 따른 부분 베이지요인과 사후확률을 나타내고 있다. 또한, 일반적인 가설 $H_0 : \rho = 0.5$ 와 $H_0 : \rho = -0.5$ 에 대해서도 같은 방법에 의해 모의실험을 실시하고, 이에 대한 결과를 표 4.2와 표 4.3에 제시하였다. 여기서, $r_1 = 2(|\rho| + 0.5)$, $r_2 = (|\rho| + 0.5)$ 이고, ρ 는 주어진 자기상관모수의 참값을 나타낸다. 표에 제시된 값들은 200번 모의실험한 결과 얻어진 부분 베이지요인과 사후확률의 평균을, 괄호속의 값은 표준편차를 나타낸다.

제시된 표에 의하면 모든 경우에 있어서 제안된 다중 베이지요인에 의한 검정 기준인 가설들의 사후확률이 자기상관모수 ρ 의 참값이 포함된 가설에서 다른 가설들보다 높은 확률을 나타내고 있음을 알 수 있다. 이러한 현상은 (A)와 (B) 두 경우 뿐 아니라 귀무가설에 설정된 ρ_0 값에 관계없이 동일하여 이 장에서 행한 모의실험은 제시된 검정법의 유효성과 로버스트성을 함께 보여주고 있다.

표 4.1: 귀무가설 $H_0 : \rho = 0$ 에 대한 부분 베이지요인과 사후확률

ρ 의 참값	부 분 베이지요인		사 후 확 률					
	B_1	B_2	(A) $\pi_0 = \frac{1}{2}, \pi_1 = \pi_2 = \frac{1}{4}$			(B) $\pi_0 = \pi_1 = \pi_2 = \frac{1}{3}$		
			$P(H_0 \mathbf{y})$	$P(H_1 \mathbf{y})$	$P(H_2 \mathbf{y})$	$P(H_0 \mathbf{y})$	$P(H_1 \mathbf{y})$	$P(H_2 \mathbf{y})$
-0.9	0.00011 (0.0007)	13.6559 (1.3504)	0.00017 (0.0010)	0.99981 (0.0011)	0.00001 (0.0001)	0.00009 (0.0005)	0.99988 (0.0007)	0.00002 (0.0001)
-0.7	0.00664 (0.0144)	11.4264 (1.4828)	0.00942 (0.0179)	0.98967 (0.0194)	0.00090 (0.0015)	0.00523 (0.0103)	0.99373 (0.0121)	0.00102 (0.0018)
-0.5	0.10957 (0.1249)	8.78461 (1.4929)	0.11512 (0.1115)	0.87298 (0.1219)	0.01188 (0.0138)	0.07015 (0.0717)	0.91508 (0.0856)	0.01476 (0.0171)
-0.3	0.62571 (0.4446)	6.17015 (1.1803)	0.39166 (0.1728)	0.56218 (0.1921)	0.04614 (0.0243)	0.26907 (0.1346)	0.66643 (0.1660)	0.06445 (0.0364)
0.0	2.77333 (0.9880)	2.65595 (0.8561)	0.64820 (0.0913)	0.18001 (0.0902)	0.17177 (0.0707)	0.49059 (0.0859)	0.25787 (0.1038)	0.25152 (0.0895)
0.3	6.26586 (1.5114)	0.74449 (0.4369)	0.43493 (0.1707)	0.05831 (0.0541)	0.50674 (0.1907)	0.30245 (0.1333)	0.08030 (0.0615)	0.61724 (0.1666)
0.5	8.91438 (1.5427)	0.19224 (0.1709)	0.18537 (0.1361)	0.01767 (0.0139)	0.79695 (0.1483)	0.11601 (0.0907)	0.02270 (0.0188)	0.86127 (0.1075)
0.7	11.2845 (1.5340)	0.01720 (0.0294)	0.02393 (0.0360)	0.00235 (0.0035)	0.97370 (0.0391)	0.01345 (0.0212)	0.00272 (0.0043)	0.98383 (0.0251)
0.9	13.7715 (1.5272)	0.00020 (0.0004)	0.00035 (0.0007)	0.00004 (0.0001)	0.99960 (0.0008)	0.00018 (0.0003)	0.00004 (0.0001)	0.99976 (0.0004)

표 4.2: 귀무가설 $H_0 : \rho = 0.5$ 에 대한 부분 베이지요인과 사후확률

ρ 의 참값	부 분 베이지요인		사 후 확 률					
	B_1	B_2	(A) $\pi_0 = \frac{1}{2}, \pi_1 = \pi_2 = \frac{1}{4}$			(B) $\pi_0 = \pi_1 = \pi_2 = \frac{1}{3}$		
			$P(H_0 \mathbf{y})$	$P(H_1 \mathbf{y})$	$P(H_2 \mathbf{y})$	$P(H_0 \mathbf{y})$	$P(H_1 \mathbf{y})$	$P(H_2 \mathbf{y})$
-0.9	0.00001 (0.0001)	6.43434 (0.4924)	0.00001 (0.0001)	0.99999 (0.0001)	0.00001 (0.0001)	0.00001 (0.0001)	0.99999 (0.0001)	0.00001 (0.0001)
-0.7	0.00003 (0.0001)	6.25083 (0.4921)	0.00005 (0.0001)	0.99994 (0.0001)	0.00001 (0.0001)	0.00003 (0.0001)	0.99996 (0.0001)	0.00001 (0.0001)
-0.5	0.00126 (0.0017)	5.93339 (0.6274)	0.00199 (0.0025)	0.99767 (0.0029)	0.00032 (0.0004)	0.00107 (0.0013)	0.99857 (0.0018)	0.00035 (0.0005)
-0.3	0.02253 (0.0301)	5.56601 (0.6595)	0.02983 (0.0317)	0.96592 (0.0363)	0.00424 (0.0046)	0.01686 (0.0189)	0.97825 (0.0245)	0.00488 (0.0057)
0.0	0.27966 (0.2534)	4.66094 (0.7361)	0.24122 (0.1119)	0.72433 (0.1320)	0.03443 (0.0221)	0.15223 (0.0804)	0.80333 (0.1103)	0.04442 (0.0321)
0.3	1.94589 (1.0695)	3.17274 (0.7591)	0.27200 (0.0893)	0.61212 (0.1198)	0.11586 (0.0436)	0.37035 (0.0854)	0.45593 (0.1387)	0.17370 (0.0678)
0.5	5.17727 (2.3062)	1.86108 (0.7501)	0.66361 (0.0820)	0.10184 (0.0508)	0.23453 (0.1119)	0.50501 (0.0809)	0.15201 (0.0735)	0.34296 (0.1326)
0.7	11.7154 (4.0252)	0.64582 (0.4647)	0.43801 (0.2063)	0.02998 (0.0201)	0.53201 (0.2242)	0.30657 (0.1637)	0.04299 (0.0308)	0.65043 (0.1921)
0.9	20.5449 (5.1254)	0.07811 (0.0935)	0.10091 (0.1045)	0.00463 (0.0047)	0.89446 (0.1092)	0.05945 (0.0650)	0.00562 (0.0060)	0.93491 (0.0709)

표 4.3: 귀무가설 $H_0 : \rho = -0.5$ 에 대한 부분 베이즈요인과 사후확률

ρ 의 참값	부 분		사 후 확 률					
	베이즈요인		(A) $\pi_0 = \frac{1}{2}, \pi_1 = \pi_2 = \frac{1}{4}$			(B) $\pi_0 = \pi_1 = \pi_2 = \frac{1}{3}$		
	B_1	B_2	$P(H_0 y)$	$P(H_1 y)$	$P(H_2 y)$	$P(H_0 y)$	$P(H_1 y)$	$P(H_2 y)$
-0.9	0.03109 (0.0655)	28.5330 (6.2851)	0.04138 (0.0787)	0.95694 (0.0818)	0.00166 (0.0032)	0.02415 (0.0477)	0.97384 (0.0517)	0.00200 (0.0040)
-0.7	0.49845 (0.4566)	15.4376 (5.3065)	0.36588 (0.2146)	0.61372 (0.2295)	0.02039 (0.0159)	0.25090 (0.1662)	0.72016 (0.1898)	0.02892 (0.0248)
-0.5	1.60063 (0.7107)	7.06594 (3.2416)	0.64743 (0.1223)	0.27846 (0.1480)	0.07410 (0.0379)	0.49183 (0.1108)	0.39586 (0.1555)	0.11229 (0.0858)
-0.3	2.79436 (0.7447)	2.85579 (1.7311)	0.35234 (0.0657)	0.14367 (0.0627)	0.50398 (0.0937)	0.29430 (0.0656)	0.21661 (0.0879)	0.48908 (0.1170)
0.0	4.02853 (0.6586)	0.43178 (0.3630)	0.29217 (0.1257)	0.04780 (0.0285)	0.66001 (0.1497)	0.19174 (0.0922)	0.06382 (0.0411)	0.74442 (0.1285)
0.3	4.97484 (0.6988)	0.03312 (0.0377)	0.03757 (0.0348)	0.00604 (0.0057)	0.95638 (0.0403)	0.02177 (0.0209)	0.00714 (0.0071)	0.97108 (0.0278)
0.5	5.51142 (0.7207)	0.00309 (0.0071)	0.00380 (0.0069)	0.00072 (0.0013)	0.99547 (0.0081)	0.00211 (0.0040)	0.00082 (0.0015)	0.99705 (0.0055)
0.7	5.86924 (0.5451)	0.00004 (0.0001)	0.00009 (0.0001)	0.00001 (0.0001)	0.99989 (0.0001)	0.00004 (0.0001)	0.00001 (0.0001)	0.99993 (0.0001)
0.9	5.94196 (0.6039)	0.00001 (0.0001)	0.00001 (0.0001)	0.00001 (0.0001)	0.99999 (0.0001)	0.00001 (0.0001)	0.00001 (0.0001)	0.99999 (0.0001)

표 4.4: MCMC 모의실험에 의한 $H_0 : \rho = 0$ 가 참일 때의 검정결과

	가 설	H_0	H_1	H_2
채택 비율	(A) $\pi_0 = \frac{1}{2}, \pi_1 = \pi_2 = \frac{1}{4}$	194 (0.970)	4 (0.020)	2 (0.010)
	(B) $\pi_0 = \pi_1 = \pi_2 = \frac{1}{3}$	173 (0.865)	11 (0.055)	16 (0.080)

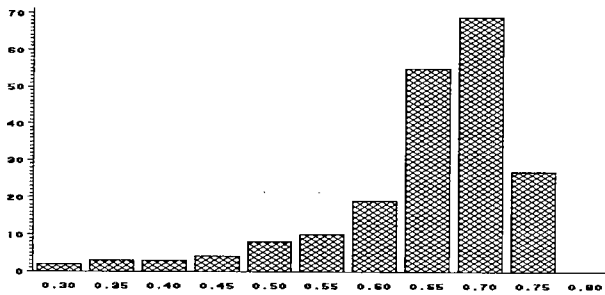


그림 4.1: $\pi_0 = \frac{1}{2}, \pi_1 = \pi_2 = \frac{1}{4}$ 인 경우 사후확률 $P(H_0|y)$ 의 히스토그램

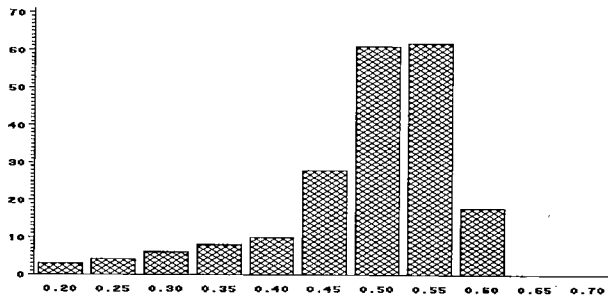


그림 4.2: $\pi_0 = \pi_1 = \pi_2 = \frac{1}{3}$ 인 경우 사후확률 $P(H_0|\mathbf{y})$ 의 히스토그램

앞에 제시된 표 4.4는 그림 4.1, 그림 4.2와 함께 $H_0 : \rho = 0$ 가 참일 때 모의실험에 의해 얻은 베이지 검정결과를 나타낸다. 이를 보면 200번의 모의실험 중 $H_0 : \rho = 0$ 가 참일 때 H_0 의 채택횟수가 가장 높게 나타나고 있어 제안된 검정법의 유효성을 보여주고 있다. 여기서, ()의 값은 채택비율을 나타낸다. 또한, 두 가지 사전확률의 가정 (A)와 (B)에 대해 각각 사후확률 $P(H_0|\mathbf{y})$ 의 히스토그램(그림 4.1, 그림 4.2)으로 미루어 볼 때 사전확률의 가정에 관계없이 세 가설 H_0, H_1, H_2 중 참가설인 H_0 의 사후확률이 현저히 높음을 보이고 있다.

5. 결론

본 논문은 회귀모형의 오차항들 사이에 존재하는 자기상관모수의 존재 여부 (또는 자기상관모수의 값)를 검정하는 방법으로 Bertolino, Piccinato와 Racugno(1995)에 의해 제안된 다중 베이지요인을 적용하는 방법을 제안하였다. 이를 위해서 따름모수를 동반한 대립가설에서 정의된 관심 모수공간의 분할에 따라 사전분포의 형태가 다르게 나타나는 경우 정보적 사전분포(informative prior distribution)의 가정 아래에서 각 부분 베이지요인을 유도하였으며, 이를 자기상관모수 검정에 적용하는 방법을 제시하였다. 또한, 부분 베이지요인의 계산에서 발생하는 복잡한 수치적 적분 문제를 해결하기 위해 이를 일반화 Savage-Dickey 밀도비의 형태로 변환시킨 다음 Gibbs 추출법을 이용하여 근사적으로 추정하는 방법을 고안하였다. 또한, 추정된 부분 베이지요인 값과 사전확률을 이용하여 모수공간의 각 분할집합에 따라 설정된 다중 가설들의 사후확률을 계산 및 이를 통한 다중결정법을 제시하였다.

이 방법은 표본이론에서 유도된 기존의 검정법으로는 불가능한 일반적인 귀무가설 $H_0 : \rho = \rho_0$ 의 검정을 가능하게 할 뿐 아니라, 이것이 베이지 규칙에 의한 다중결정기법이라는 점에서 그 특징을 찾아볼 수 있다. 따라서, 본 논문에서 제안된 다중 베이지요인을 이용한 자기상관모수 검정법은 베이지 규칙으로 모수공간의 각 분할에 대해 다양하게 설정된 대립가설들의 진위를 한번에 검정할 수 있으며, 기존의 표본이론에 의한 방법으로 해결할 수 없었던 문제점들을 해결할 수 있는 방법임을 알 수 있다.

본 논문에서는 대립가설의 모수공간을 2개의 공간으로 분할하여 검정법을 유도하였으나, 본 논문의 내용을 일반화시키면 k 개($k > 2$)로 분할된 공간하에서 설정되는 가설들의 다중결정문제로 쉽게 해결이 가능하며, 이에 대한 연구 및 응용분야의 개발은 앞으로의 연구과제로 남겼다.

참고문헌

- [1] Bertolino, F., Piccinato, L. and Racugno, W.(1995), Multiple Bayes Factors for Testing Hypotheses, *Journal of the American Statistical Association*, vol. 90, 213 – 219.
- [2] Best, N. G., Cowles, M. K and Vines, S. K.(1996), *CODA : Convergence Diagnosis and Output Analysis Software for Gibbs Sampling Output*, Version 0.40, University of Nebraska Medical Center, USA.
- [3] Devroy, L.(1986), *Non-uniform Random Variate Generation*, Springer – Verlag, New York.
- [4] Durbin, J. and Watson, G. S.(1951), Testing for Serial Correlation in Least Square Regression, *Biometrika*, vol. 38, 159 – 177.
- [5] Gelman, A. and Rubin, D. B.(1992), Inference from Iterative Simulation Using Multiple Sequences, *Statistical Science*, vol. 4, 457 – 511.
- [6] Johnson, N. L., Kotz, S. and Balakrishnan, N.(1994), *Continuous Univariate Distributions*, vol. 1, John Wiley & Sons, Inc., New york.
- [7] Kim, H. J.(1999), On Computing Multiple Bayes Factor Using Modified Savage-Dickey Density Ratio, *Annals of the Institute of Statistical Mathematics*, to appear.
- [8] Kim, H. J. and Han, S. S.(1998), A Bayesian Test for the First-order Autocorrelations in Regression Analysis, *The Korean Journal of Applied Statistics*, vol. 11, 97 – 111.
- [9] Marr, R. L. and Quesenberry, C. P.(1991), A NU Test for Serial Correlation of Residuals from One or More Regression Regimes, *Technometrics*, vol. 33, 441 – 457.
- [10] Verdinelli, I. and Wasserman, L.(1995), Computing Bayes Factors Using a Generalization of the Savage-Dickey Density Ratio, *Journal of the American Statistical Association*, vol. 90, 614–618.
- [11] Wonnacott, R. J. and Wonnacott, T. H.(1979), *Econometrics*, John Wiley & Sons, New York.
- [12] Zellner, A.(1971), *An Introduction to Bayesian Inference in Econometrics*, John Wiley & Sons, New York.

[1999년 3월 접수, 1999년 8월 최종수정]

On Testing the First-order Autocorrelation of the Error Term in a Regression Model via Multiple Bayes Factor

Sung-Sil, Han¹⁾ Hea-Jung, Kim²⁾

ABSTRACT

This article suggests a Bayesian method for testing the presence(or a given value) of serial correlation of the error term in a multiple linear regression model. Development of the method is based upon a synthesis of multiple decision method using multiple Bayes factor defined by partitions of parameters space and posterior probability estimation method utilizing a generalized Savage-Dickey density ratio. This yields a way of dealing with Bayesian multiple decision problem about the serial correlation that has not been available yet. Performance of the method is examined via a Monte Carlo simulation study.

1) Lecturer, Department of Statistics, Dongguk University, 100-715, Seoul, Korea

2) Professor, Department of Statistics, Dongguk University, 100-715, Seoul, Korea