응용통계연구 제 16권 1호, 2003년, pp. 101-115

# 왜도 타원형 분포를 이용한 준모수적 계층적 선택 모형 *

정 윤 식 [1) 장 정 훈 [2)

요 약

본 논문에서는 Chen, Dey와 Shao(1999), Branco와 Dey(2001)가 제안한 왜도가 있는 두터운 꼬리를 가지는 오차 분포와 디리슈레 과정 사전분포를 이용한 베이지안 메타분석(meta-analysis)을 하고자 한다. 베이지안 메타분석을 위하여 가중함수를 고려한 계층적 선택 모형을 이용한다. 이때의 오차항은 왜도가 있는 비정규 분포로 가정한다. 이를 위하여 우선 왜도 타원형 분포의 일반적인 족을 소개한다. 이 분포족중 왜도 정규분포와 왜도 $t$ 분포를 오차항 분포로 이용한 베이지안 계층적 선택 모형을 고려하며, 이때 발생하는 복잡한 베이지안 계산은 MCMC 방법으로 해결한다. 마지막으로, 실제 자료(Johnson, 1993)인 두 가지의 충치예방약의 효과에 대한 차이를 비교하기 위해 얻어진 12개의 연구 자료를 이용하여 본 연구에서 제시된 베이지안 방법을 이용하여 메타분석을 한다.

주요용어: 가중함수, 계층적 선택 모형, 깁스 샘플러, 디리슈레 과정 사전분포, 마코프 체인 몬테 칼로 방법, 밀도 생성함수, 베이지안 메타분석, 왜도 타원형 분포

## 1. 서론

메타분석은 서로 독립적으로 연구되어진 결과들을 종합하여 전체적인 하나의 결과를 도출하기 위한 통계적 방법으로 특히 의학분야에서 어떤 약의 효과라든지 새로운 연구에 대한 계획 등에 널리 쓰이고 있다. 이러한 메타분석에는 두 가지의 본질적인 문제점을 지니고 있는데 하나는 메타분석에 포함되는 연구들은 각기 다른 시점, 장소에서 서로 독립적으로 연구된 결과로써 각각의 연구효과가 이질적이다. 따라서, Morris와 Normand(1992)는 임의효과 모형(random effect model)이나 계층적 모형 (hierarchical model)을 이용하여 설명하였다. 또 다른 문제점은 출판편의가 존재하는 것이다. 즉, 유의한 결과를 보여주는 연구는 그렇지 못한 연구에 비해 출판될 확률이 높다는 것으로 자료를 수집할 때 통계적으로 유의한 결과를 보여주는 연구가 메타분석에 포함될 가능성이 커지는 것을 의미하므로 이 때 편의가 발생한다. 이와 같이 일정한 수준을 넘어선 것들만을 얻게 되므로 이를 출판편의라 한다. 그러므로 출판편의가 존재할 때 이러한 편의를 보정하기 위하여 Larose와 Dey(1996)는 일반적으로 가중함수를 사용하였다. Silliman(1997)은 계층적 모형에 가중함

* 본 연구는 한국과학재단 목적기초연구(2000-1-10400-005-2)지원으로 수행되었음
1) (609-735) 부산시 금정구 장전동, 부산대학교 자연과학대학 통계학과 교수
E-mail: yschung@pusan.ac.kr
2) (609-735) 부산시 금정구 장전동, 부산대학교 자연과학대학 통계학과 박사과정
E-mail: statjjh@hanmail.net

수를 포함시킨 계층적 선택 모형(hierarchical selection models)을 소개했다. 이 모형은 각각의 연구효과의 이질성과 출판편의를 동시에 설명하고 각 미관측된 연구효과가 정규분포를 따르는 임의효과로 간주함으로서 이 모형은 모수적 임의효과 모형과 유사하다. Chung, Dey와 Jang(2002)는 준 모수적 계층적 선택모형의 준모수적 베이지안 방법을 제시하였다. 이러한 연구들에서 오차항들은 모두 정규분포를 가정하였다. 그러나, 최근들어 많은 통계학자들이 비 대칭과 두터운 꼬리를 갖는 오차항에 대한 중요성을 강조하면서, 비 정규오차를 갖는 선형모형에 대한 이론적, 실용적인 관심이 매우 증가하고 있다. 최초로, Zellner(1976)는 다변량 스튜던트 $t$ 분포를 이용한 선형 회귀모형을 베이즈 방법과 고전적 방법으로 연구, 비교하였다. Azzalini와 Dalla-Valle(1996)은 정규분포에 모양모수(shape parameter)를 추가한 족을 확장시킨 왜도 정규분포의 다변량 형태에 대한 일반적인 이론을 제시했다. 최근에 Branco와 Dey(2001)는 다변량 정규분포, 스튜던트 $t$분포, 멱 지수분포와 피어슨 타입 II분포를 포함하는 다변량 왜도 타원형 분포의 일반적인 족을 제안하면서 왜도 모수를 정의하였으며, Sahu, Dey와 Branco(2001)는 왜도 타원형 오차를 갖는 회귀모형을 연구하여 회귀계수에 대한 베이지안 방법을 제안하였다.

이에 우리는 이렇게 제시된 비 대칭 오차항을 갖는 모형을 이용하여 메타분석에 사용되는 모형을 구축하려고 한다. 즉, Chung, Dey와 Jang(2002)에 의해 제시된 메타분석의 두 가지 문제점을 설명해 줄 수 있는 준 모수적 계층적 선택 모형과 Sahu, Dey와 Branco(2001)등이 제안한 왜도 타원형 분포를 결합하여 왜도 분포를 갖는 오차항을 가정한 베이지안 해석을 소개한다. 예로써, 우리는 두 가지 충치 예방약의 효과에 대한 차이를 비교하기 위해 12개의 연구 자료에 대한 메타분석을 한다. 이때 오차항이 왜도 타원형 분포를 갖는 베이지안 계층적 선택 모형을 가정하고, 더욱이 미 관측연구효과는 비모수적 사전분포인 디리슈레 과정 사전분포를 이용한 베이지안 해석을 한다.

본 논문은 다음과 같이 구성되어 있다. 2장에서는 다변량 왜도 타원형 분포를 소개한다. 특별한 형태인 왜도 정규분포와 왜도 스튜던트 $t$ 분포가 예제로서 설명될 것이다. 3장에서는 오차항이 왜도 타원형 분포를 갖는 계층적 선택 모형에서 미 관측연구효과는 비모수적 사전분포인 디리슈레 과정 사전분포를 가정한 모형식을 설명한다. 또한 복잡한 베이지안 계산은 MCMC방법을 이용하여 왜도 정규분포와 왜도 $t$분포하에서의 자세한 계산적 문제를 다룬다. 마지막으로 4장에서는 Johnson(1993)의 실제 자료를 예제로 소개하고 이 자료를 제시된 방법으로 적용할 것이다.

## 2. 대칭오차의 준 모수적 계층적 선택 모형

1장에서 언급한 메타분석의 중요한 문제점을 해결하기 위하여 Silliman(1997)은 모수적 접근인 계층적 선택 모형을 제시하였으며, Chung, Dey 와 Jang(2002)은 계층적 선택 모형 아래서 임의효과에 비모수적 접근인 준 모수적 계층적 선택모형으로 확장하였다.

Morris 와 Normand(1992)이 제시한 베이지안 계층적 모형은 다음과 같다. $i=1,\dots,n$에 대하여

$$Y_i|\alpha_i, \sigma_i \sim N(\alpha_i, \sigma_i^2),$$
$$\alpha_i|\mu, \sigma_\alpha^2 \sim N(\mu, \sigma_\alpha^2), \quad \mu, \sigma_i^2, \sigma_\alpha^2 \sim \pi(\mu, \sigma_i^2, \sigma_\alpha^2) \tag{2.1}$$

여기서 $\pi(\cdot)$은 미지의 모수들에 대한 결합 사전분포를 나타내며, $Y_i$는 $i$번째 관측된 확률 표본, $\alpha_i$는 모수 $\mu$와 $\sigma_\alpha^2$ 를 갖는 확률표본에 대응하는 임의효과이다.

메타분석에서 모형 (2.1)은 그 자체가 통계적으로 유의한 쪽으로 편의된 출판연구의 하나이기 때문에 자주 쓰이는 모형은 아니다. 다시 말하면, 유의한 결과를 보여주는 연구는 그렇지 못한 연구에 비해 출판될 확률이 더 높다는 것을 말하는데 이것이 앞서 언급한 출판편의이다. 따라서 메타분석에 이러한 문제점을 해결하고자 Silliman(1997)은 계층적 모형 (2.1)에 가중함수를 포함시킨 계층적 선택모형을 제시하였다. 이 모형은 각각의 연구효과의 이질성과 출판편의를 동시에 설명할 수 있는 모형으로 다음과 같이 정의되어진다.

$$Y_i|\alpha_i, \sigma_i \sim f^w(y_i|\alpha_i, \sigma_i),$$
$$\alpha_i|\mu, \sigma_\alpha^2 \sim N(\mu, \sigma_\alpha^2), \quad \mu, \sigma_i^2, \sigma_\alpha^2 \sim \pi(\mu, \sigma_i^2, \sigma_\alpha^2) \tag{2.2}$$

여기서 가중분포(weighted distribution)는

$$f^w(y_i|\alpha_i, \sigma_i) = \frac{w(y_i)\ f(y_i|\alpha_i, \sigma_i)}{C_w} \tag{2.3}$$

와 같이 정의되어지고 $w(\cdot)$는 가중함수이고, 정규상수는

$$C_w = \int w(x)\ f(x|\alpha_i, \sigma_i) dx \tag{2.4}$$

이다. 물론 비 가중 분포 $f(y_i|\alpha_i, \sigma_i) = N(y_i; \alpha_i, \sigma_i)$는 (2.1)에서 정의된 것과 같다.

모형 (2.2)는 모형의 이질성과 출판 편의를 동시에 설명할 수 있지만 미관측된 연구효과에 가정된 분포가 잘못되었을 경우에는 결과를 잘못 도출할 수 있는 문제점을 가지고 있다. 이런 관점에서 Chung, Dey 와 Jang(2002)은 연구효과에 비모수적 접근인 디리슈레 과정 사전분포(Ferguson, 1973)를 이용한 준 모수적 계층적 선택모형을 다음과 같이 소개하였다.

$$Y_i|\alpha_i, \sigma_i \sim f^w(y_i|\alpha_i, \sigma_i),$$
$$\alpha_i|G \sim G, \quad G|M, \mu, \sigma_\alpha^2 \in DP(M \cdot G_0(\mu, \sigma_\alpha^2)),$$
$$\mu, \sigma_\alpha^2, \sigma_i^2 \sim \pi(\mu, \sigma_\alpha^2, \sigma_i^2), \quad M \sim p(M) \tag{2.5}$$

여기서 $G$는 미지 분포함수이고 $DP$는 평균이 $G_0$이고 정도(precision)가 $M$인 디리슈레 과정 사전분포(DPP)를 나타낸다. $G_0$는 기저 사전분포 (baseline prior)라 한다. 또한 $\mu$와 $\sigma_\alpha^2$는 이 기저 사전분포의 평균과 분산이 된다. DPP를 이용했을 때 나타나는 계산적 어려움은 MCMC방법을 통해서 해결될 수 있으며, Chung, Dey 와 Jang(2002) 에 잘 설명되어 있다. 그러나 관측된 확률표본의 오차항이 비 정규분포이면 모형 (2.5)는 역시 적당하지 않다. 따라서 우리는 모형 (2.5)를 확장한 오차분포가 대칭인 것을 포함하는 3장에서 논의될 왜도 오차 모형을 제시한다.

## 3. 왜도 오차분포에서의 준 모수적 계층적 선택 모형

모형 (2.1)에서, 오차항 $\epsilon_i$에 대하여 Chen, Dey와 Shao(1999)가 제시한 방법으로 왜도 타원형 분포를 정의한다. 오차항이 일변량이므로 일변량 확률변수로 취급한다.

### 3.1. 일변량 왜도 타원형 분포

이 절에서는 Chen, Dey와 Shao(1999)가 제시한 방법을 사용하여 왜도 타원형 분포의 족을 제시한다. 왜도 확률변수는 대칭인 확률변수와 양의 확률변수의 합으로부터 얻어진다. 즉,

$$\epsilon = U + \delta Z. \tag{3.1}$$

여기서 $U$는 대칭이고 단봉인 분포를 가지고 $Z$는 양의 치우친 분포(positive skewed distribution)를 갖는다. 만약 $\delta = 0$이면 오차항은 원래의 대칭인 분포가 되며 따라서 이를 왜도 모수로서 쉽게 해석할 수 있다. Chen, Dey와 Shao(1999)는 $\delta > 0(\delta < 0)$이라는 것은 오른쪽(왼쪽)으로 치우쳐있다는 것을 보였다. Branco와 Dey(2001)에 의하여 $U \sim El(\mu, \sigma; g_2)$와 $Z \sim El(0, 1; g_1)$의 형태인 간단한 타원형 족으로 수정한다. 여기서 $El(\mu, \sigma; g)$는 위치모수가 $\mu$, 척도모수가 $\sigma$ 그리고 밀도 생성함수(density generator)가 $g$인 타원형 분포를 의미한다. 즉, $Y \sim El(\mu, \sigma; g)$라 함은 $Y$의 밀도함수의 형태가

$$f_{el}(y|\mu, \sigma; g^{(1)}) = \sigma^{-1/2} g^{(1)} \left[ \frac{(y-\mu)^2}{\sigma} \right] \tag{3.2}$$

이고

$$g^{(1)}(u) = \frac{\Gamma(1/2)}{\sqrt{\pi}} \frac{g(u)}{\int_0^\infty u^{-1/2} g(u) du}, \quad u \geq 0 \tag{3.3}$$

를 만족하는데 여기서 $g(u)$는 $\int_0^\infty u^{-1/2} g(u) du$의 값이 존재하도록 해주는 비 증가 함수이다. 이런 $g$를 밀도생성함수라고 한다. 또한 본 논문에서는 타원형 분포의 밀도함수 (3.2)가 항상 존재한다고 가정한다.

다음의 변수변환을 고려해보자.

$$\epsilon = U + \delta Z. \tag{3.4}$$

Azzalini와 Dalla-Valle(1996)의 결과로부터 왜도 타원형 족(skew elliptical class)은 다음의 조건부 분포

$$[\epsilon | Z > 0] \tag{3.5}$$

로부터 발전되어질 수 있고 이 분포를 $SE(\mu, \sigma, \delta; g)$로 표기한다.

**정 리 3.1.**(Sahu, Dey and Branco; 2001) 위의 가정 하에서 (3.5)에 정의된 $\epsilon | Z > 0$의 확률밀도함수는

$$\begin{aligned} f(\epsilon|\mu, \sigma, \delta; g^{(1)}) = \quad & 2 f_{el}(\epsilon | \mu, \sigma + \delta^2; g^{(1)}) \times \\ & F_{el} \left( \frac{\delta y_*}{(\sigma + \delta^2)\sqrt{1 - \delta^2 (\sigma + \delta^2)^{-1}}} | 0, I; g^{(1)}_{q(y_*)} \right) \end{aligned} \tag{3.6}$$

이고 여기서 $f_{el}(\epsilon|\mu,\sigma;g)$는 위치모수가 $\mu$, 척도모수가 $\sigma$ 그리고 밀도 생성함수가 $g$인 타원형 밀도함수를 나타내고 $F_{el}$는 그에 대응되는 타원형 분포함수를 나타내며

$$g_a^{(1)}(u) = \frac{\Gamma(\frac{1}{2})}{\pi^{\frac{1}{2}}} \frac{g(a+u;2)}{\int_0^\infty r^{\frac{1}{2}-1} g(a+u;2)dr}, \quad a > 0 \tag{3.7}$$

와

$$q(y_*) = y_*^t(1+\delta^2)^{-1}y_* \text{ and } y_* = y - \mu \tag{3.8}$$

이다.

**예 제 3.1.** 왜도 정규 분포

만약 $g(u) = e^{-\frac{u}{2}}$이면 $g^{(1)}(u) = \frac{1}{\sqrt{2\pi}}e^{-\frac{u}{2}}$가 됨을 쉽게 보일 수 있고 $g^{(1)}_{q(y_*)}$는 $q(y_*)$의 영향을 받지 않는다. 따라서 왜도 정규분포의 확률밀도함수는

$$f(\epsilon|\mu,\sigma,\delta) = 2\phi\left[\frac{\epsilon-\mu}{\sqrt{\sigma+\delta^2}}\right] \times \Phi\left[\frac{\delta(\epsilon-\mu)}{\sqrt{\sigma(\sigma+\delta^2)}}\right] \tag{3.9}$$

이고 여기서 $\phi$와 $\Phi$는 각각 표준정규분포의 밀도함수와 누적분포함수를 나타낸다. (3.9)의 밀도함수는 $U \sim N(\mu,\sigma)$와 $Z \sim N^+(0,1)$일 때 $\epsilon = \delta Z + U$의 분포와 정확히 일치한다. 여기서 $N^+(0,1)$는 0에서 절단된 절단 정규분포이다. 그러므로 $\epsilon$에 대한 왜도 정규분포는 $U \sim N(\mu,\sigma)$와 $Z \sim N^+(0,1)$일 때 $\epsilon = \delta Z + U$의 분포라고 정의할 수 있고 이 결과는 Azzalini와 Dalla-Valle(1996)에서도 볼 수 있다.

**예 제 3.2.** 왜도 $t$ 분포

$g(u;\nu) = \left[1+\frac{u}{\nu}\right]^{-\frac{\nu+2}{2}}$라고 하면

$$g_a^{(1)}(u;\nu) = \Gamma(1)[\pi(\nu+1)]^{-1/2}\left(\frac{\nu+1}{\nu+a}\right)\left(1+\frac{u}{\nu+1}\frac{\nu+1}{\nu+a}\right)^{-(\nu+2)/2} \tag{3.10}$$

이고 왜도 $t$분포의 밀도함수는

$$f(\epsilon|\mu,\sigma,\nu) = 2t_\nu(y|\mu,1+\delta^2) \times T_{\nu+1}\left[\left(\frac{\nu+q(y_*)}{\nu+p}\right)^{-\frac{1}{2}} \frac{\delta(\epsilon-\mu)}{\sqrt{\sigma(\sigma+\delta^2)}}\right] \tag{3.11}$$

이고 여기서 $t_\nu$와 $T_\nu$는 각각 자유도 $\nu$인 스튜던트 $t$분포의 밀도함수와 누적분포함수를 나타낸다. 왜도 정규분포와 마찬가지로 이 분포는 $U \sim t_\nu(\mu,\sigma)$와 $Z \sim t^+(0,1)$일 때 $\epsilon = U + \delta Z$의 분포와 일치하고 여기서 $t_\nu(\mu,\sigma)$는 위치모수가 $\mu$, 척도모수가 $\sigma$ 그리고 자유도 $\nu$인 $t$분포를 나타내며 $t^+(0,1)$는 0에서 절단된 절단 $t$분포를 나타낸다.

위와 같은 방법으로 우리는 왜도가 있는 준 모수적 계층적 선택모형을 모형화 할 수 있다. (2.1)의 베이지안 계층적 모형에서 오차항, $\epsilon_i$,에 왜도 정규분포(Azzalini와 Dalla-Valle, 1996)를 포함하는 왜도 타원형 분포 (3.6)을 가정한다. 따라서 우리의 모형식은 다음과 같다.

$$Y_i = \alpha_i + \epsilon_i^*, \quad \epsilon_i^* \sim SE(0,\sigma_i^2;\delta), \tag{3.12}$$

여기서 $SE(0,\sigma_i^2;\delta)$는 (3.5)에서 정의된 것과 같다. Chen, Dey와 Shao(1999)에 의하여 다음과 같이 표현할 수 있다.

$$\epsilon_i^* = u_i + \delta z_i, \tag{3.13}$$

여기서 $z_i$는 양의 확률변수이고 $u_i$는 대칭이고 단봉인 분포를 가진다.

본 논문에서 $z_i$에는 0에서 절단된 절단 정규분포(양의 정규분포)와 절단 $t$분포(양의 $t$분포)가 이용될 것이고 $u_i$에는 위치모수가 0인 정규분포와 스튜던트 $t$분포가 이용될 것이다. 전자의 경우가 예제 3.1에 있는 왜도 정규분포이고 후자의 경우가 예제 3.2에 있는 왜도 $t$분포라 한다.

### 3.2. 왜도 정규분포를 이용한 준 모수적 계층적 선택모형

$z_i$와 $u_i$가 각각 양의 정규분포와 정규분포를 따른다고 가정하자. 또한 관측된 확률표본 $Y_i$는 (2.3)의 가중분포와 임의효과 $\alpha_i$는 디리슈레 과정 사전분포를 가정한다. 구조는 다음과 같다.

$$\begin{aligned}
Y_i &= \alpha_i + \delta(z_i - E(z_i)) + u_i,\\
\alpha_i|G &\sim G, \quad z_i \sim N^+(0,1), \quad u_i \sim N(0,\sigma_i^2)\\
G|M,\mu,\sigma_\alpha^2 &\in DP(M\cdot N(\mu,\sigma_\alpha^2))\\
\mu \sim N(a,b),\ &\sigma_\alpha^2 \sim IG(c_1,d_1),\ \sigma_i^2 \sim IG(c_2,d_2)\\
\delta &\sim N(m,\tau), \quad M \sim Gam(e_1,e_2),
\end{aligned} \tag{3.14}$$

여기서 $Gam(g,h)$는 모양모수 $g$와 척도모수 $h$를 갖는 감마분포, $IG(p,q)$는 모양모수 $p$와 척도모수 $q$를 갖는 역 감마 분포를 나타내고 $E(z_i)$는 $E(Y_i)$가 일반화 선형 모형의 구조처럼 $\alpha_i$의 예측값이 되도록, 즉 $E(Y_i)=\alpha_i$, 해주는 $z_i$의 기대값으로 수정항(correction factor)이라 불린다. 확률표본이 가중분포로부터 관측되므로 $i=1,\ldots,n$에 대하여

$$y_i|\alpha_i,\sigma_i^2,\delta,z_i \sim (C(\alpha_i,\sigma_i,\delta,z_i))^{-1}w(y_i)N(\alpha_i+\delta(z_i-E(z_i)),\sigma_i^2) \tag{3.15}$$

여기서 $C(\alpha_i,\sigma_i^2,\delta,z_i)$는 (2.4)에서 정의된 정규화 상수로 $(\alpha_i,\sigma_i^2,\delta,z_i)$의 함수이다.

이제 MCMC방법을 사용하기 위한 완전 조건부 분포를 구해보자. 이제부터 각 괄호(bracket)는 밀도함수(density)를 나타내기로 한다. 즉, 예를 들어 $[X,Y]$, $[X|Y]$와 $[X]$는 각각 $X$와 $Y$의 결합밀도함수, $Y$가 주어졌을 때의 $X$의 조건부 밀도함수와 $X$의 주변 밀도함수를 나타낸다. 기저 사전분포 하에서

$$[\mu|\alpha,\mathbf{y},\mathbf{z},\delta,\sigma_\alpha,\sigma] = N\left(\frac{b\sum\alpha_i+\sigma_\alpha^2 a}{bn+\sigma_\alpha^2}, \frac{b\sigma_\alpha^2}{bn+\sigma_\alpha^2}\right), \tag{3.16}$$

$$[\delta|\alpha,\mathbf{y},\mu,\mathbf{z},\sigma_\alpha,\sigma] \propto \prod_{i=1}^{n}(C(\alpha_i,\sigma_i,\delta,z_i))^{-1}N(\beta_1(\delta),\gamma_1(\delta)) \tag{3.17}$$

여기서 $\beta_1(\delta) = \frac{\sum \frac{(z_i - E(z_i)(y_i - \alpha_i)}{\sigma_i^2} + \frac{m}{\tau}}{\sum \frac{(z_i - E(z_i))^2}{\sigma_i^2} + \frac{1}{\tau}}$이고 $\gamma_1(\delta) = \left(\sum \frac{(z_i - E(z_i))^2}{\sigma_i^2} + \frac{1}{\tau}\right)^{-1}$가 되고,

$$[z_i|\alpha, \mathbf{y}, \mu, z_j(j \neq i), \delta, \sigma_\alpha, \sigma] \quad \propto \quad (C(\alpha_i, \sigma_i, \delta, z_i))^{-1} N^+(\beta_1(z_i), \gamma_1(z_i)) \tag{3.18}$$

여기서 $\beta_1(z_i) = \frac{\delta(y_i - \alpha_i + \delta E(z_i))}{\sigma_i^2 + \delta^2}$와 $\gamma_1(z_i) = \frac{\sigma_i^2}{(\sigma_i^2 + \delta^2)}$가 되며,

$$[\sigma_\alpha^2|\alpha, \mathbf{y}, \mu, \mathbf{z}, \delta, \sigma] = IG\left(\frac{n}{2} + c_1, \frac{1}{2}\sum(\alpha_i - \mu)^2 + d_1\right), \tag{3.19}$$

$$[\sigma_i^2|\alpha, \mathbf{y}, \mu, \mathbf{z}, \delta, \sigma_\alpha, \sigma_j(j \neq i)] \quad \propto \quad (C(\alpha_i, \sigma_i, \delta, z_i))^{-1} IG(\xi_1(\sigma_i), \rho_1(\sigma_i)) \tag{3.20}$$

여기서 $\xi_1(\sigma_i) = \frac{1}{2} + c_2$이고 $\rho_1(\sigma_i) = \frac{1}{2}(y_i - \alpha_i - \delta(z_i - E(z_i))^2 + d_2$ 이다.

$\alpha_i$와 $M$의 완전 조건부 분포는 West, Müller와 Escobar(1994)에 의해서 다음과 같이 나타낼 수 있다. 먼저 $\alpha_i$는

$$[\alpha_i|\mathbf{y}, \sigma, \mathbf{z}, \mu, \delta, \sigma_\alpha^2, \alpha_j(j \neq i), M] \propto q_0 g_p(\alpha_i) + \sum_{j \neq i} q_j I(\alpha_i|\alpha_j) \tag{3.21}$$

여기서 $q_0$와 $q_j$는 $q_0 \propto \int f^w(y_i|\alpha_i, \sigma_i, \delta, z_i) dG_0(\alpha_i)$와 $q_j \propto f^w(y_i|\alpha_j, \sigma_j, \delta, z_j)$이며 $q_0 + \sum_{j \neq i} q_j = 1$를 만족한다. $g_p(\alpha_i) \propto f^w(y_i|\alpha_i, \sigma_i, \delta, z_i) dG_0$이고 $I(t|u)$는 $u$에서 확률 1을 갖는 퇴화분포(degenerate distribution)를 나타낸다. 여기서 $g_p(\alpha_i)$는 기저 사후분포(baseline posterior)라 불린다. 또한, 디리슈레 과정 사전분포의 척도 모수 $M$은

$$[M|\eta, k] = \pi_\eta Gam(e_1 + k, B(\eta)) + (1 - \pi_\eta) Gam(e_1 + k - 1, B(\eta)) \tag{3.22}$$

와 같이 혼합 감마 분포가 되고 여기서 $\eta$는 평균이 $\frac{M+1}{M+n+1}$, 인 베타분포로부터 나온 잠재변수이며 $B(\eta) = e_2 - \log(\eta)$이고 가중치 $\pi_\eta$는

$$\frac{\pi_\eta}{1 - \pi_\eta} = \frac{e_1 + k - 1}{n(e_2 - \log(\eta))}. \tag{3.23}$$

와 같이 정의된다. 자세한 계산 과정은 Chung, Dey와 Jang(2002)에 있다.

### 3.3. 왜도 $t$분포를 이용한 준 모수적 계층적 선택 모형

이제 오차항에 다소 두터운 꼬리를 갖는 분포를 고려하여 보자. 스튜던트 $t$분포가 널리 쓰이므로 $z_i$와 $u_i$에 각각 양의 $t$분포와 스튜던트 $t$분포를 가정한다. 그러면 모형은 다음과

같다.

$$\begin{aligned} Y_i &= \alpha_i + \delta(z_i - E(z_i)) + u_i, \\ \alpha_i|G \sim G, &\quad z_i \sim t^+(0,1;\nu_2), \quad u_i \sim t(0,\sigma_i^2;\nu_1) \\ G|M,\mu,\sigma_\alpha^2 &\in DP(M\cdot N(\mu,\sigma_\alpha^2)) \\ \mu \sim N(a,b), &\ \sigma_\alpha^2 \sim IG(c_1,d_1),\ \sigma_i^2 \sim IG(c_2,d_2) \\ \delta \sim N(m,\tau), &\ \ M \sim Gam(e_1,e_2), \end{aligned} \tag{3.24}$$

여기서 $\nu_1$과 $\nu_2$는 각각 $u_i$와 $z_i$의 자유도를 나타낸다. 앞절과 마찬가지로 $i=1,\ldots,n$에 대하여

$$y_i|\alpha_i,\sigma_i^2,\delta,z_i \sim (C_t(\alpha_i,\sigma_i,\delta,z_i))^{-1}w(y_i)t(\alpha_i+\delta(z_i-E(z_i)),\sigma_i^2;\nu_1), \tag{3.25}$$

가 되고 $(C_t(\alpha_i,\sigma_i,\delta,z_i))^{-1}$는 $t$분포에 대한 정규화 상수이다.

(3.24)의 $t$분포를 바로 사용하게 되면 계산적인 어려움이 생기므로 $t$분포는 정규분포의 척도혼합형태로 나타낼 수 있다는 사실을 이용한다. 또한 양의 $t$분포도 같은 방법으로 사용할 수 있다. 즉, $i=1,\ldots,n$에 대하여

$$\begin{aligned} y_i|\alpha_i,\sigma_i,\delta,z_i,w_i &\sim (C_t(\alpha_i,\sigma_i,\delta,z_i))^{-1}N(\alpha_i+\delta(z_i-E(z_i)),w_i\sigma_i^2), \\ w_i &\sim IG\left(\frac{\nu_1}{2},\frac{\nu_1}{2}\right), \end{aligned} \tag{3.26}$$

와

$$z_i|\lambda_i \sim N^+(0,\lambda_i), \quad \lambda_i \sim IG\left(\frac{\nu_2}{2},\frac{\nu_2}{2}\right) \tag{3.27}$$

로 놓으면 $y_i|\alpha_i,\sigma_i,\delta,z_i$의 주변분포는 (3.25)과 일치하고 $z_i$의 주변분포는 (3.24)와 일치한다. 왜도 정규분포의 경우와 마찬가지로 완전 조건부 분포는 다음과 같고 $\mu,\sigma_\alpha^2$와 $M$의 완전 조건부 분포는 3.2절과 동일하다.

$$[\delta|\alpha,\mathbf{y},\mu,\mathbf{z},\sigma_\alpha,\sigma,\mathbf{w},\lambda] \propto \prod_{i=1}^n (C_t(\alpha_i,\sigma_i,\delta,z_i))^{-1}N(\beta_2(\delta),\gamma_2(\delta)) \tag{3.28}$$

여기서 $\beta_2(\delta)=\frac{\sum\frac{(y_i-\alpha_i)(z_i-E(z_i))}{w_i\sigma_i^2}+\frac{m}{\tau}}{\sum\frac{(z_i-E(z_i))^2}{w_i\sigma_i^2}+\frac{1}{\tau}}$와 $\gamma_2(\delta)=\left(\sum\frac{(z_i-E(z_i))^2}{w_i\sigma_i^2}+\frac{1}{\tau}\right)^{-1}$이고,

$$[z_i|\alpha,\mathbf{y},\mu,z_j(j\neq i),\delta,\sigma_\alpha,\sigma,\mathbf{w},\lambda] \propto (C_t(\alpha_i,\sigma_i,\delta,z_i))^{-1}N^+(\beta_3(z_i),\gamma_3(z_i)) \tag{3.29}$$

여기서 $\beta_3(z_i)=\frac{\lambda_i\delta(y_i-\alpha_i-\delta E(z_i))}{w_i\sigma_i^2+\lambda_i\delta^2}$와 $\gamma_3(z_i)=\left(\frac{\lambda_i w_i\sigma_i^2}{\lambda_i\delta^2+w_i\sigma_i^2}\right)$이다.

$$[\sigma_i^2|\alpha,\mathbf{y},\mu,\mathbf{z},\delta,\sigma_\alpha,\sigma_j(j\neq i),\mathbf{w},\lambda] \propto (C_t(\alpha_i,\sigma_i,\delta,z_i))^{-1}IG(\xi_2(\sigma_i),\rho_2(\sigma_i)) \tag{3.30}$$

여기서 $\xi_2(\sigma_i) = \frac{1}{2} + c_2$, $\rho_2(\sigma_i) = \frac{1}{2w_i}(y_i - \alpha_i - \delta(z_i - E(z_i)))^2 + d_2$이다. 또한,

$$[w_i|\alpha, \mathbf{y}, \mu, \mathbf{z}, \delta, \sigma_\alpha, w_j(j \neq i), \sigma, \lambda] = IG(\xi_2(w_i), \rho_2(w_i)) \tag{3.31}$$

여기서 $\xi_2(w_i) = \frac{\nu_1+1}{2}$와 $\rho_2(w_i) = \frac{1}{2}\left(\frac{(y_i-\alpha_i-\delta(z_i-E(z_i)))^2}{\sigma_i^2} + \nu_1\right)$ 이다.

$$[\lambda_i|\alpha, \mathbf{y}, \mu, \mathbf{z}, \delta, \sigma_\alpha, \lambda_j(j \neq i), \sigma, \mathbf{w}] = IG\left(\frac{\nu_2+1}{2}, \frac{1}{2}(z_i^2 + \nu_2)\right) \tag{3.32}$$

이 된다. 또한 $\alpha_i$에 대한 기저사후 분포는 3.2절에서와 비슷하게

$$\begin{aligned} g_p(\alpha_i) \quad &\propto \quad (C_t(\alpha_i, \sigma_i, \delta, z_i))^{-1} \\ & \qquad N\left(\frac{\sigma_\alpha^2(y_i - \delta(z_i - E(z_i)) + \mu w_i \sigma_i^2}{\sigma_\alpha^2 + w_i \sigma_i^2}, \frac{w_i \sigma_\alpha^2 \sigma_i^2}{\sigma_\alpha^2 + w_i \sigma_i^2}\right). \end{aligned} \tag{3.33}$$

이다.

## 4. 예제

Johnson(1993)은 두 가지 충치 예방약- NaF와 SMFP- 의 효과를 비교하기 위해 이미 연구되어진 12개의 연구를 종합했다. 각 연구에서 $y_i$는 두 가지 약의 평균 차이를 나타낸다. 즉, 이것은 SMFP를 사용한 환자의 충치범위에서 NaF를 사용한 환자의 충치범위를 뺀 값이 되므로 $y_i$의 값이 양이라면 NaF가 SMFP보다 더 효과적이라는 것을 나타낸다. 또한 각 연구의 표준오차의 추정치 $\hat{\sigma}_i$와 표본 크기 $N$이 표 4.1에 주어져 있다. 이런 연구로부터 전체적인 효과 $\mu$를 추정하는 것이 우리의 관심이다. Johnson(1993)은 가중평균으로 $\mu$의 값을 추정하였고 그 추정치는 .32, 95%신뢰구간은 (.13,.52)로서 NaF가 더 효과적이라는 가설을 뒷받침한다.

Johnson(1993)의 결과가 비록 NaF의 효과가 더 좋다는 것을 나타내지만 그의 메타분석에는 몇 가지 문제점이 있을 수 있다. 특히 그는 전 세계의 문헌에서 찾아낸 12개의 연구를 고려했지만 실제로 찾아낸 연구는 13개였다. 13번째 연구에서는 분산의 추정치를 구할 수가 없어서 메타분석에서 제외시켰고 또한 이 12개의 연구 중에서 5번째 연구를 제외하면 나머지 11개의 연구효과가 양의 값을 나타내고 있다. 이러한 연구로 Johnson의 연구에는 출판편의가 존재한다고 생각하는 것이 타당하여 Silliman(1997)은 이 자료에 대하여 계층적 선택모형을 적용하였고 Chung, Dey와 Jang(2002)은 오차항이 정규본포를 따르는 준모수적 계층적 선택모형을 적용시켜 분석하였다.

**표 4.1.** The Original Johnson Data

| | Studies | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| $y_i$ | .86 | .33 | .47 | .50 | -.28 | .04 | .80 | .19 | .49 | .49 | .01 | .67 |
| $\hat{\sigma}_i$ | .57 | .56 | .35 | .25 | .54 | .28 | .78 | .13 | .28 | .24 | .08 | .17 |
| $N$ | 247 | 326 | 277 | 363 | 343 | 1490 | 418 | 2273 | 1352 | 2762 | 2222 | 2126 |

이제 표 4.1에 주어져 있는 자료에 3장에서 제시한 왜도가 있는 준 모수적 계층적 선택 모형을 적용한다. $N(\mu, \sigma_\alpha^2)$와 $t(\mu, \sigma_\alpha^2; 5)$의 두 분포가 $\alpha_i$의 기저 사전분포로 사용될 것이다. 또한 $w(y) = |y|$와 $w(y) = |y|^2$가 가중함수로 사용될 것이며, 이 때 $w(y) = |y|$는 유의한 결과를 보여주는 연구가 좀 더 관측될 가능성이 많다는 것을 의미하고 $w(y) = |y|^2$는 출판편의가 매우 심하다는 것을 나타낸다. 이 두 가중함수가 지정되었을 때의 정규화 상수는 해석적으로 계산이 가능하며 그 형태는 다음과 같다. 먼저 왜도 정규분포하에서 $w(y) = |y|$일 때,

$$\begin{aligned} C(\alpha_i, \sigma_i, \delta, z_i) &= (\alpha_i + \delta(z_i - E(z_i)))\left(1 - 2\Phi\left(-\frac{\alpha_i + \delta(z_i - E(z_i))}{\sigma_i}\right)\right) \\ &\quad + \sqrt{\frac{2}{\pi}}\sigma_i \exp\left(-\frac{(\alpha_i + \delta(z_i - E(z_i)))^2}{2\sigma_i^2}\right) \end{aligned} \tag{4.1}$$

이고 여기서 $\Phi(\cdot)$는 표준정규분포의 누적분포함수를 나타낸다. $w(y) = |y|^2$의 경우에는 정규화 상수는 바로 $E(Y^2)$과 일치하므로 (4.1)보다 훨씬 간단하다. 즉,

$$C(\alpha_i, \sigma_i, \delta, z_i) = (\alpha_i + \delta(z_i - E(z_i))^2 + \sigma_i^2. \tag{4.2}$$

이다. 왜도 $t$의 경우를 살펴보자. 마찬가지로 $w(y) = |y|$가 지정되었다면

$$\begin{aligned} C_t(\alpha_i, \sigma_i, \delta, z_i) &= (\alpha_i + \delta(z_i - E(z_i)))\left(1 - 2T_{\nu_1}\left(-\frac{\alpha_i + \delta(z_i - E(z_i))}{\sigma_i}\right)\right) \\ &\quad + c(\nu_1)\sigma_i \frac{2}{\nu_1 - 1}\left(\nu_1 + \frac{(\alpha_i + \delta(z_i - E(z_i)))^2}{\sigma_i^2}\right)^{-\frac{\nu_1 - 1}{2}}, \end{aligned} \tag{4.3}$$

이고 여기서 $T_{\nu_1}(\cdot)$ 는 자유도 $\nu_1$인 스튜던트 $t$분포의 누적분포함수이고 상수 $c(\nu_1)$은

$$c(\nu_1) = \frac{\Gamma\left(\frac{\nu_1+1}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)\sqrt{\pi}}\nu_1^{\nu_1/2} \tag{4.4}$$

이다. 또한 $w(y) = |y|^2$일 때는

$$C_t(\alpha_i, \sigma_i, \delta, z_i) = (\alpha_i + \delta(z_i - E(z_i))^2 + \frac{\nu_1}{\nu_1 - 2}\sigma_i^2. \tag{4.5}$$

의 형태로 역시 $E(Y^2)$과 일치한다.

이제 미지의 모수에 대한 사전분포를 정의하여 보자. 먼저 $\mu \sim N(0, 0.04)$ 와 $\sigma_\alpha^2 \sim IG(2.0016, 24.96)$로 주어지는데 이는 DerSimonian과 Laird(1986)에 의해 제시된 비 반복 추정치 (noniterative estimate)에 기인한다. 또한 주어진 자료를 바탕으로 하여 $\sigma_i^2 \sim IG(1.2, 5)$와 왜도모수 $\delta$의 사전분포는 Branco와 Dey(2001)의 경우와 같이 $\delta \sim N(0, 10^2)$로 가정하며 디리슈레 과정 사전분포의 정도모수인 $M$의 사전분포로서 세 개의 감마분포를 정의하는데 각각의 평균이 1,5와 30이 되도록 가정한다. 여기서 $E(M)$의 선택은 기저 사전분포에 대한 믿음의 정도가 작고, 보통이고 크다는 것을 의미한다. 기호의 편이를 위해 다음을 정의하자.

$$M_1 = Gam(2, 0.5), \quad M_2 = Gam(5, 1), \quad M_3 = Gam(15, 2).$$

예를 들면 $M \sim M_1$이라는 것은 $M$의 사전분포가 $Gam(2, 0.5)$라는 것을 나타낸다. 우리는 $M$의 사전분포와 기저분포가 변함에 따라서 관심있는 모수의 변화를 알아볼 것이다. 결과는 표 4.2, 4.3과 4.4에 주어져있다. 메트로폴리스 알고리즘이 10,000번 포함된 깁스 샘플러는 5,000번 수행되었고 최초 3,000번은 번-인 주기(burn-in period)로 버렸다. 알고리즘의 수렴성은 통계 패키지인 S-plus에서 쉽게 사용할 수 있는 CODA의 Geweke(1992)방법으로 검사되었다.

**표 4.2.** $\mu$, $\sigma^2_\alpha$, $M$와 $\delta$의 추정치

| 가중함수 | 모형 | DPP | $\hat{\mu}$ | $\hat{\sigma}^2_\alpha$ | $\hat{M}$ | $\hat{\delta}$ |
|---|---|---|---|---|---|---|
| $w(y) = \lvert y \rvert$ | $SN$ | $M \sim M_1$ | -0.004(0.465) | 57E | 11.7 | -0.472 |
| | | | | | | (-2.820,1.202) |
| | | $M \sim M_2$ | -0.002(0.456) | 57E | 20.3 | -0.860 |
| | | | | | | (-4.358,2.179) |
| | | $M \sim M_3$ | 0.003 (0.497) | 57E | 55.1 | -0.685 |
| | | | | | | (-3.937,2.179) |
| | $ST$ | $M \sim M_1$ | 0.001 (0.490) | 57E | 7.5 | -0.640 |
| | | | | | | (-0.916,-0.378) |
| | | $M \sim M_2$ | 0.002 (0.498) | 57E | 12.9 | -0.641 |
| | | | | | | (-0.923,-0.374) |
| | | $M \sim M_3$ | 0.001 (0.495) | 57E | 40.4 | -0.638 |
| | | | | | | (-0.905,-0.360) |
| $w(y) = \lvert y \rvert^2$ | $SN$ | $M \sim M_1$ | 0.003 (0.526) | 57E | 11.6 | -0.636 |
| | | | | | | (-3.573,2.009) |
| | | $M \sim M_2$ | 0.001 (0.512) | 57E | 19.5 | -0.721 |
| | | | | | | (-3.872,1.859) |
| | | $M \sim M_3$ | 0.011 (0.565) | 57E | 54.8 | -0.981 |
| | | | | | | (-5.056,1.712) |
| | $ST$ | $M \sim M_1$ | -0.001(0.489) | 57E | 7.7 | -0.666 |
| | | | | | | (-0.950,-0.406) |
| | | $M \sim M_2$ | 0.001 (0.494) | 57E | 13.3 | -0.665 |
| | | | | | | (-0.941,-0.408) |
| | | $M \sim M_3$ | 0.001 (0.503) | 57E | 40.9 | -0.661 |
| | | | | | | (-0.932,-0.403) |

표 4.2에 우리가 제시한 왜도 오차모형의 결과가 주어져 있다. 모형에서 $SN$과 $ST$는 본 논문에서 제시된 것처럼 오차항에 대한 왜도 정규분포와 자유도 5인 왜도 $t$분포를 나타낸다. 세 번째 열에 있는 $\mu$의 추정치 옆의 괄호의 값은 $Pr(\mu > 0|data)$을 나타내고 네 번째 열은 $\sigma_\alpha^2$의 추정값을 나타내며 $E$는 $10^{-4}$을 나타낸다. 즉, 57E는 0.0057이라는 뜻이 되는 것이고, $\delta$의 추정치 밑의 괄호의 두 값은 95% 신뢰구간(credible interval)을 나타낸다. 만약, 이 신뢰구간이 0을 포함하고 있으면 주어진 자료에는 왜도가 없다는 것이고 그렇지 않으면 주어진 자료에 왜도가 존재한다는 것을 뜻한다.

표 4.2에 제시된 결과들을 분석하여 보자. 먼저, 전체적인 연구효과를 나타내는 왜도 오차모형 하에서의 $\mu$의 사후 추정치는 가중함수와 기저 사전분포의 믿음의 정도에 관계없이 거의 0을 나타내고 있으며 0보다 클 사후확률도 0.5근방을 나타내고 있으므로 NaF의 우수성은 입증되지 않는다고 할 수 있다. 다음으로 연구간 분산을 나타내는 $\sigma_\alpha^2$은 오차항의 분포에 매우 로버스트함을 알 수 있고 디리슈레 과정 사전분포의 정도모수인 $M$의 사후 추정치 역시 오차항의 분포에는 관계없이 $M$의 사전분포의 선택에만 어느정도 민감성을 보이고 있다. 이제 왜도모수인 $\delta$에 대하여 추론해 보자. $\delta$의 추정치는 왜도 정규분포와 왜도 $t$ 분포에 따라 매우 다른 결과를 말해주고 있다. 왜도 정규분포가 가정되어 있을 때의 $\delta$의 값이 왜도 $t$ 분포가 가정되어 있을 때보다 변동이 상당히 크다는 것을 알 수 있고 또한 왜도 정규분포 하에서는 주어진 자료에 왜도가 존재하지 않는다는 것을 알 수 있으며 표 4.2에 의해서 기저 사전분포에 대한 믿음의 정도가 커질수록 95% 신뢰구간의 길이는 더 커지는 추세를 보이는 것을 알 수 있다. 그러나 왜도 $t$ 분포가 가정된다면 왜도모수 $\delta$는 가중함수와 $M$의 사전분포의 선택에 매우 로버스트하며 또한 95% 신뢰구간이 0을 포함하고 있지 않다는 것으로 주어진 자료에는 왜도가 존재한다는 것을 암시하고 있다. 이는 왜도 정규분포에 비해서 자유도 5인 왜도 $t$ 분포가 가지고 있는 두꺼운 꼬리확률이 $\delta$의 변동을 어느정도 설명하고 있는 것으로 보인다. 표 4.3은 왜도 정규분포와 왜도 $t$ 분포가 가정되었을 때의 평균 군집(cluster)의 수를 나타내는 데 여기서 군집이라고 하는 것은 서로 다른 값을 갖는 $\alpha_i$들의 값의 개수이고 이에 대한 자세한 설명은 Chung과 Jang(2001)에서 볼 수 있다. 표 4.3에서 왜도 정규분포하에서는 기저 사전분포의 믿음의 정도가 커질수록 군집의 개수가 많아지지만 왜도 $t$ 분포하에서는 이런 믿음의 정도에 관계없이 약 3개의 군집만 형성하고 있는 것을 알 수 있다.

**표 4.3.** $E(M)$과 $w(y)$의 변화에 따른 평균 군집 수

| | $w(y) = \|y\|$ | | $w(y) = \|y\|^2$ | |
|---|---|---|---|---|
| | $SN$ | $ST$ | $SN$ | $ST$ |
| $M \sim M_1$ | 7.324 | 3.270 | 7.033 | 3.557 |
| $M \sim M_2$ | 9.017 | 3.188 | 8.404 | 3.483 |
| $M \sim M_3$ | 10.687 | 3.235 | 10.336 | 3.417 |

# 참고문헌


[1] Azzalini, A. and Dalla-Valle, A. (1996) The Multivariate Skew-Normal Distribution, *Biometrika*, **83**, 715-726.

[2] Branco, M.D. and Dey, D. K. (2001) A General Class of Multivariate Skew-elliptical Distributions, *Journal of Multivariate Analysis*, **79**, 99-113.

[3] Chen, M-H., Dey, D. K. and Shao, Q-M.(1999) A new skewed link model for dichotomous quantal response data. *Journal of the American Statistical Association*, **94**, 1172-1186.

[4] Chung, Y., Dey, D. K. and Jang, J. (2002) Semiparametric hierarchical selection model for Bayesian meta analysis, *Journal of Statistical Computation and Simulation*, (in press).

[5] Chung, Y. and Jang, J. (2001) A Bayesian Method to Semiparametric Hierarchical Selection Models, *The Korean Journal of Applied Statistics*, **14**, 161-175.

[6] DerSimonian, R. and Laird, N. (1986) Meta-Analysis in Clinical Trials, *Controlled Clinical Trials*, **7**, 177-188.

[7] Ferguson, T. S. (1973) A Bayesian Analysis of Some Nonparametric Problems, *The Annals of Statistics*, **1**, 209-230.

[8] Geweke, J. (1992) Evaluating the Accuracy of Sampling-Based Approaches to Calculating Posterior Moments, in *Bayesian Statistics 4*, (J. M. Bernado, J. O. Berger, A. P. Dawid, and A. F. M. Smith eds.), Oxford, UK; Oxford University Press, pp. 169-193.

[9] Johnson, M. F. (1993) Comparative Efficacy of Naf and SMFP Dentifrices in Caries Prevention : A Meta-Analysis Overview, *Journal of the European Organization for Caries Research (ORCA)*, **27**, 328-336.

[10] Larose, D. and Dey, D. K. (1996) Weighted Distributions Viewed in the Context of Model Selection : a Bayesian Perspective, *Test*, **5**, 227-246

[11] Morris, C. N. and Normand, S. L. (1992) Hierarchical Models for Combining Information and for Meta-Analysis, in *Bayesian Statistics 4*, (J. M. Bernado, J. O. Berger, A. P. Dawid, and A. F. M. Smith eds.), Oxford, UK; Oxford University Press, pp. 321-344.

[12] Sahu, S. K., Dey, D. K. and Branco, M. D. (2001) A New Class of Multivariate Skew Distribution with Applications to Bayesian Regression Models, Technical Report 0030, University of Southampton.

[13] Silliman, N. P. (1997) Hierarchical Selection Models with Applications in Meta-Analysis, *Journal of the American Statistical Association*, **92**, 926-936.

[14] West, M., Müller, P. and Escobar, M. D. (1994) Hierarchical Priors and Mixture Models, with Applications in Regression and Density Estimation, *Aspects of Uncertainty: A Tribute to D. V. Lindley* ed. A. F. M. Smith, and P. R. Freeman, London: John Wiley and Sons, pp. 363-386.

[15] Zeller, A.(1976) Bayesian and non-bayesian analysis of regression model with multivariate Student-t term. *Journal of the American Statistical Association*, **71**, 400-405.

[ 2002년 7월 접수, 2002년 11월 채택 ]

# Semiparametric Bayesian Hierarchical Selection Models with Skewed Elliptical Distribution

Younshik Chung [1] Junghoon Jang [2]

ABSTRACT

Lately there has been much theoretical and applied interest in linear models with non-normal heavy tailed error distributions. Starting Zellner(1976)'s study, many authors have explored the consequences of non-normality and heavy-tailed error distributions. We consider hierarchical models including selection models under a skewed heavy-tailed error distribution proposed originally by Chen, Dey and Shao(1999) and Branco and Dey(2001) with Dirichlet process prior(Ferguson, 1973) in order to use a meta-analysis. A general calss of skewed elliptical distribution is reviewed and developed. Also, we consider the detail computational scheme under skew normal and skew $t$ distribution using MCMC method. Finally, we introduce one example from Johnson(1993)'s real data and apply our proposed methodology.

*Keywords:* Bayesian meta analysis;Density generator function; Dirichlet process prior; Gibbs sampler; MCMC method; Skew elliptical distribution; Weighted function

1) Professor, Department of Statistics, Pusan University, 609-735, Pusan, Korea
E-mail: yschung@pusan.ac.kr
2) Graduate, Department of Statistics, Pusan University, 609-735, Pusan, Korea
E-mail: statjjh@hanmail.net