

# 한우의 도체중, 배장근단면적 및 근내지방도의 유전모수 추정방법

이 특 환  
한경대학교

## Methods for Genetic Parameter Estimations of Carcass Weight, Longissimus Muscle Area and Marbling Score in Korean Cattle

D. H. Lee  
Hankyong National University

### ABSTRACT

This study is to investigate the amount of biased estimates for heritability and genetic correlation according to data structure on marbling scores in Korean cattle. Breeding population with 5 generations were simulated by way of selection for carcass weight, Longissimus muscle area and latent values of marbling scores and random mating. Latent variables of marbling scores were categorized into five by the thresholds of 0, 1, 2, and 3 SD(DS1) or seven by the thresholds of -2, -1, 0, 1, 2, and 3 SD(DS2). Variance components and genetic parameters(Heritabilities and Genetic correlations) were estimated by restricted maximum likelihood on multivariate linear mixed animal models and by Gibbs sampling algorithms on multivariate threshold mixed animal models in DS1 and DS2. Simulation was performed for 10 replicates and averages and empirical standard deviation were calculated. Using REML, heritability of marbling score were under-estimated as 0.315 and 0.462 on DS1 and DS2, respectively, with comparison of the parameter(0.500). Otherwise, using Gibbs sampling in the multivariate threshold animal models, these estimates did not significantly differ to the parameter. Residual correlations of marbling score to other traits were reduced with comparing the parameters when using REML algorithm with assuming linear and normal distribution. This would be due to loss of information and therefore, reduced variation on marbling score. As concluding, genetic variation of marbling would be well defined if liability concepts were adopted on marbling score and implemented threshold mixed model on genetic parameter estimation in Korean cattle.

(Key words : Categorical data, Threshold model, Variance component, Modeling, Hanwoo)

### I 서 론

1990년대 이후 우리나라 한우의 개량은 육량보다는 육질중심의 개량에 크게 관심을 기울여 왔으며 따라서 한우 종축선발을 위한 대표적인 개량형질로써 도축시 측정 가능한 육질관련 형질에 크게 비중을 두었고 이러한 형질을 기준으로 종축선발을 실시하여 왔다. 실제로 한우 후대검정을 통한 종모우 선발에 기준이 되는 형질들은 도체중, 배장근단면적 및 근내지방도 등이었다(축산기술연구소, 2003). 한우 후대검정방법에 있

어서 1996년 가을 이전 출생 검정송아지들에 대하여는 거세를 실시하지 않고 후대검정을 실시한 반면에 이후 후대 검정송아지들에 대하여 거세를 실시하여 후대검정을 실시하여 왔다. 한편 거세우와 비거세우 간에 근내지방의 침착정도는 상당한 차이가 있는 것으로 알려져 있으며 비거세우에 있어서 근내지방의 침착정도를 점수화하여 표현한 근내지방도의 점수분포를 살펴보면 1점이 45%, 2점이 24%으로 자료분포에 있어서 정규성에 크게 위배된 분포특성을 보이고 있는 것으로 조사되었다(이 등, 2001). 반면에 23차 한우 후대

Corresponding author : D. H. Lee, Dept. of Animal Life and Resources, Hankyong National University, Seokjong-dong 67, Ansong-si, Kyonggi-do, 456-749, Korea, Tel : 031-670-5091, Fax : 031-676-5091, E-mail : dhlee@hnu.hankyong.ac.kr

검정 이후부터 실시한 거세 검정우들의 근내지방도 자료를 살펴보면 정규성에 가까이 접근하는 분포특성을 보이고 있었다(윤 등, 2002). 주요 개량의 대상이 되는 이러한 형질들에 대한 유전분석은 현재까지 각 자료의 정규분포 특성을 갖는다는 가정하에 개체들의 유전효과를 포함하는 단형질 선형모형을 설정하고 유전분석을 실시하여 왔다(축산기술연구소, 2003). 하지만 정규성에 크게 위배되는 범주형 자료의 유전분석을 선형모형의 가정하에 실시하는 것은 분석의 오류를 범할 수 있다고 하겠다. 이론적으로 이산형 자료특성을 갖는 형질에 있어서 유전능력 평가는 범주형 모형하에서 실현가능하며(Gianola and Foulley, 1983) 이러한 모형에서의 분석방법들에 대한 연구가 다수 수행되었다(Varona 등, 1999; Ramirez-Valverde 등, 2001). 일반적으로 범주형 모형은 순서화된 범주형자료(Ordered categorical data)에서 정규 분포하는 잠재변수가 존재한다(Wright, 1934)는 가정하에 잠재변수를 생성하고 이를 이용한 혼합모형식을 설정하여 유전분석을 실시하는 방법인데 일반적으로 범주형자료에 있어서 유전변이는 연속변량이며 정규분포한다는 가정하에 선형모형을 설정하여 변이를 추정할 경우와 비교할 때, 범주형 모형에 의한 유전변이가 1~2배 크게 추정된다고 보고하였다(Matós 등, 1997, Val Tassel 등, 1999; Phocas and Lae, 2003). 하지만 자료의 분포특성에 따른 편의추정의 크기에 대한 연구는 아직 미진한 상태이며 그 원인에 대한 명확한 해석이 미흡한 실정이다. 특히 현재 한우 종모우 선발을 위하여 고려되고 있는 형질들에 대한 종축평가에 있어서 연속변량으로 간주되는 도체중 및 배장근단면적과 이산형이며 비정규분포의 특성을 갖는 근내지방도에 대한 유전변이 추정에 있어서 다수의 형질을 동시에 평가할 수 있는 방법에 대한 연구가 선행되어야 한다. 따라서 본 연구는 한우개량에 있어서 육질에 크게 영향하며 순서화된 범주형 자료로 조사되는 근내지방도 및 연속형 변량으로 간주되는 도체중 및 배장근단면적에 있어서 비거세우에서 조사된 근내지방도의 경우와 거세우에서 조사된 근내지방도의 경우에 다형질 선형모형과 범주형 모형에서 유전변이의 차이를 알아보고 범주형 자료의 특성

에 따른 유전변이의 크기를 제시함으로써 실제 한우 종모우의 유전능력 평가를 실시하기 위한 최적 분석모형 및 분석방법을 알아보고 이를 한우 종모우 평가방법에 활용하기 위한 참고자료로 제시하고자 본 모의실험을 통한 연구를 수행하였다.

## II 분석방법

### 1. 모의시험 자료생성

본 연구를 위해 수행한 모의실험은 현실 타당성에 근접한 육종계획을 수립하고 이에 대한 검정모형을 설정하여 Monte Carlo 모의실험을 실시하였다. 본 모의실험에 육질개량을 위하여 고려된 형질들은 실제 한우개량을 위한 선발형질로써 도축하여 조사되는 도체중(CW), 배장근단면적(LMA), 및 근내지방도(MS) 등이었다.

본 모의실험에 있어서 각 세대별 종축선발은 3가지 고려된 형질들에 대한 다변량 혼합개체모형식으로부터 추정된 각 형질별 육종가(최적선형불편예측치)에 상대적 경제 가중치를 동일하게 부여하여 종축을 선발하였다. 종축선발을 위한 모의실험은 숫소 100두 및 암소 5,000두를 기초축으로 하여 매 세대 수컷 50두와 암컷 2,000두를 선발하여 5세대 동안 진행되었고 출생년도는 세대와 중첩되도록 설계하였으며 개체별 혈연관계를 고려하였다. 세대간격은 수컷의 경우 보증종모우로 선발된 후에 정액을 활용한다는 가정하에 6세부터 10세까지 이용하였으며 암컷의 경우는 2세부터 6세까지 이용하였다. 본 모형에서 자료생성을 위하여 고려한 고정효과는 암컷과 수컷으로 구분한 성별 효과, 5세대에 걸쳐 세대와 중첩되고 평균 세대당 10수준으로 하여 총 50수준을 임의 배정한 검정년도 - 계절 효과를 고려하였다. 기초 집단에 대한 도체중, 배장근단면적 및 근내지방도의 잠재변수에 대한 평균은 각각 300kg, 100cm<sup>2</sup> 및 50으로 가정하였고 각각의 형질들에 대한 분산 - 공분산 성분은 Table 1에 제시된 값을 모수로 이용하였다. 각각의 고정효과는 각 형질별 평균을 중심으로  $\pm 2$  SD 이내의 균일분포함수에서 임의 생성하였다. 육종가에 대한 자료생성은

Falconer와 Mackay(1996)가 제시한 infinitesimal genetic model에 기초하여 다음과 같이 생성하였다.

Table 1. Parameters of genetic and residual (co)variances for simulation

	CW	LMA	MS <sub>L</sub>
CW	250(750) <sup>(1)</sup>	25	15
LMA	100	20(50)	2
MS <sub>L</sub>	5	7	10(10)

Values in blanket and below diagonal are residual (co)variances and values upper diagonal are genetic (co)variances; CW: carcass weight, LMA: Longissimus Muscle Area, MS<sub>L</sub>: Liability scale of marbling score

$$a_0 = 0.5a_s + 0.5a_d + m \dots\dots\dots (1)$$

여기서  $a_0$ 는 자손의 해당 형질에 대한 육종가이고  $a_s$ 는 부의 육종가  $a_d$ 는 모의 육종가 및  $m$ 은 Mendelian sampling contribution이다.

부모를 모르는 기초집단의 경우 3가지 도축형질에 대한 개체의 육종가는 평균이 "0"이고 분산이  $G$ 인 다변량 정규분포함수( $a \sim MVN(0, G)$ )에서 생성하였으며 부모를 알 경우에는 상기의 식 (1)에서 Mendelian sampling에 의한 기여도  $m \sim MVN(0, \{0.5 - 0.25[F_s + F_d]\}G)$ 을 고려하여 생성하였다. 여기서  $F_s(F_d)$ 는 아버지(어미)의 근교계수이며  $G$ 는 제시한 해당 형질들에 대한 선발을 가하기 전의 초기 유전 (공)분산성분이다. 근교계수 추정방법은 Meuwissen과 Luo (1992)가 제시한 방법을 recursive algorithm으로 변형하여 계산하였다. 개체별 해당형질들에 대한 육종가 생성은 다형질 정규분포함수에서 Cholesky 분해방법으로 다수의 형질을 동시에 임의 생성하였으며 임의오차에 대한 생성은  $e \sim MVN(0, I \otimes R)$ 의 가정하에 다형질 정규분포함수에서 동시 생성하였다. 또한 각 형질별 반복기록이 없는 것을 가정하였다. 본 모형을 식으로 표현하면 다음과 같다.

$$y_{im} = \mu_i + Sex_{ij} + YS_{ik} + a_{im} + e_{im}$$

여기서  $y_i$ 는  $i$ 형질에 대한  $m$ 개체의 관측치,  $\mu_i$ 는  $i$ 형질의 전체평균,  $Sex_{ij}$ 는  $i$ 형질의  $j$ 성에 대한 고정효과(수준 = 2),  $YS_{ik}$ 는  $i$ 형질의  $k$ 년도 - 계절에 대한 고정효과(수준 50),  $a_{im}$ 은  $i$

형질의  $m$ 개체의 임의 상가적 유전효과 및  $e_m$ 은  $i$ 형질의  $m$ 개체에 대한 임의 환경효과이다.

상기의 모형으로부터 세대별 모든 개체에 대하여 모수 및 생성된 모수를 모형에 적합시켜 생성된 표현형 자료를 이용하여 표현형 순위 및 육종가를 추정하였다. 종축의 육종가 추정방법은 상기의 모형에 적합하여 매 세대 preconditioned conjugate gradient 방법을 이용한 다변량 BLUP 추정(Tsuruta 등, 2001)방법으로 추정하였으며 각 형질별 육종가를 표준화한 값에 동일가중치를 부여하여 선발지수식을 계산하였고 지수에 의한 상위 종축을 선발하였으며 선발된 종축은 임의교배를 실시하였다. 최종 생성된 자료를 이용하여 범주형 자료를 생성하기 위하여 표준화한 근내지방도 잠재변수에 대하여 0, 1, 2, 및 3 SD의 값을 임계값으로 하여 1~5점의 근내지방도(순서화된 범주형자료)를 생성하였고(DS1) 또한 동일한 근내지방도 잠재변수에 대하여 -2, -1, 0, 1, 2 및 3 SD를 임계값으로 한 1~7점의 근내지방도 자료(DS2)를 생성하였다. 본 모의실험은 10반복을 실시하여 모수와 추정치와의 차이를 계산하고 통계적 t-검정을 통하여 추정의 정확도를 알아보았다.

2. 분산성분 추정

(1) 선형 혼합개체모형에서 EM-REML

근내지방도에 대하여 5단계로 점수화하여 표현한 범주형 자료를 갖는 모의시험자료(DS1)와 7단계로 점수화한 모의시험자료(DS2)에 대하여 범주형 자료를 연속형 변량으로 가정하고 CW와 LMA를 포함하는 다변량 혼합 개체모형을 설정하여 분산 - 공분 성분을 다음과 같은 모형을 설정하여 추정하였다.

$$\begin{bmatrix} y_1' \\ y_2' \\ y_3' \end{bmatrix} = \begin{bmatrix} X_1 & 0 & 0 \\ 0 & X_2 & 0 \\ 0 & 0 & X_3 \end{bmatrix} \begin{bmatrix} \beta_1' \\ \beta_2' \\ \beta_3' \end{bmatrix} + \begin{bmatrix} Z_1 & 0 & 0 \\ 0 & Z_2 & 0 \\ 0 & 0 & Z_3 \end{bmatrix} \begin{bmatrix} a_1' \\ a_2' \\ a_3' \end{bmatrix} + \begin{bmatrix} e_1' \\ e_2' \\ e_3' \end{bmatrix} \dots\dots\dots (2)$$

여기서  $y_1, y_2$  및  $y_3$ 는 각각 CW, LMA 및 MS에 대한 관측치이고  $\beta_1, \beta_2$  및  $\beta_3$ 는 각각의 형

질들에 대한 고정효과이며  $a_1, a_2$  및  $a_3$ 는 각각의 형질들에 대한 동물개체의 유전능력을 나타내는 임의효과이다. 임의효과에 대한 평균은 "0"으로 가정하고 이들의 분산 - 공분산 성분은 다음과 같이 가정하였다.

3가지 형질들의 상가적 유전효과를 포함하는 다변량 혼합 개체모형에서 EM-REML 방법으로 임의변량에 대한 (공)분산 성분을 추정하였는데 추정의 수렴속도를 증가시키기 위하여 Aitken acceleration 방법(Little and Rubin, 1987)으로 수렴속도를 빠르게 구현하여 분석을 실시하였다.

(2) 범주형 혼합개체모형에서 Gibbs sampling 방법

다변량 범주형 혼합개체모형을 설정하고 Geman과 Geman(1984)이 제시한 Gibbs sampling algorithm과 Sorensen 등(1995)이 제시한 단형질 범주형 자료의 분석방법을 응용하였고 다수의 형질로 확장하는 이론에 대한 연구는 이(2002)의 연구결과에 기초하여 실시하였다. 다변량 통계분석에서 연속형 종속변수는 도체중량, 배장근단면적이고 순서화된 범주형 자료로써 근내지방도를 범주형 모형을 설정하여 Bayesian inference의 일종인 Gibbs sampling 방법으로 분산성분을 추정하였다. 분석모형에서 범주형 관측치는 아래와 같이 주어진 범주구간내에 관측되지 않는 잠재변수( $u$ )가 존재한다는 가정하에 잠재변수를 생성하였다. 즉, 범주형 자료의  $i$ 번째 관측치에서  $y_i = \lambda$ , if  $t_{\lambda-1} < u_i \leq t_\lambda$ 이며  $\lambda$ 은 범주형 자료 1 ~ C의 구간내에 관측된 관측치이며 각각의 임계값에 대한 벡터  $t = (t_1, t_2, \dots, t_{C-1})$ '는  $-\infty = t_0 < t_1 < \dots < t_{C-1} < t_C = \infty$ 의 범위내에 존재한다. 여기서부터 생성된 근내지방도에 대한 잠재변수는 연속형 변량으로써 연속형 변량인 CW 및 LMA를 고려하여 다형질 혼합개체모

형을 설정하고 유전분석을 실시하였다. 여기서 분석모형은 모의실험을 위한 자료생성시 설정된 모형과 동일하게 분석모형을 설정하였다. 즉, 3가지 종속변수에 대한 벡터를  $U$ 로 정의한

$$U = \begin{pmatrix} u_{CW} \\ u_{LMA} \\ u_{MS} \end{pmatrix} \sim N \left( \begin{pmatrix} W\theta_{CW} \\ W\theta_{LMA} \\ W\theta_{MS} \end{pmatrix}, R \otimes I_n \right)$$

여기서  $R$ 은  $3 \times 3$ 크기의 잔차분산 - 공분산성분 행렬이고  $\theta$ 는 각 형질별 위치모수이며  $W$ 는 이에 대응되는 계수행렬이다. 근내지방도 관측치에 대한 잠재변수는 다음과 같은 분포특성을 갖는다. 즉 근내지방도가  $i$ 번째 형질이고  $j$ 번째 관측치가  $\lambda$ 이라면 이에 대한 잠재변수는 다음의 확률변수 특성을 갖는 것으로 가정하였다.

$$f(u_{ij} | y_{ij} = \lambda, u_{-ij}, \theta, R, t) = \frac{\phi\left(\frac{u_{ij} - \xi_{ij}}{\sigma_{ij}}\right)}{\Phi\left(\frac{t_{i\lambda} - \xi_{ij}}{\sigma_{ij}}\right) - \Phi\left(\frac{t_{i\lambda-1} - \xi_{ij}}{\sigma_{ij}}\right)} I(u_{ij} \in [t_{i\lambda-1}, t_{i\lambda}]) \dots \dots \dots (3)$$

따라서 주어진 임계값 기타형질에 대한 관측치, 지역모수, 잔차분산 및 관측치  $\lambda$ 에서 잠재변수의 대한 조건확률분포는 다음과 같이 표현할 수 있다.

$$u_{ij} | y_{ij} = \lambda, u_{-ij}, \theta, r, t \sim TN_{t_{i\lambda-1}, t_{i\lambda}}(\xi_{ij}, \sigma_{ij}^2) \dots (4)$$

여기서  $u_{-ij}$ 는  $u_{ij}$ 를 제외한 기타 잠재변수 및 연속변량을 갖는 형질들에 대한 관측치를 의미하고  $\phi(\cdot)$ 와  $\Phi(\cdot)$ 는 정규분포에서 확률밀도함수(probability density function)와 누적밀도함수(cumulated density function)를 의미하며,  $TN$ 는 절단형 정규분포함수를 의미한다. 또한 절단형 정규분포함수에서 평균  $\xi_{ij} = w'_{ij}\theta_i + R_{i(-i)}$ ,  $R_{(-i)(-i)}^{-1} e_{-ij}$ 와 분산  $\sigma_{ij}^2 = r_{ii} - R_{i(-i)} R_{(-i)(-i)}^{-1} R_{(-i)i}$ 으로 계산된다. 여기서 잔차분산  $R = \begin{bmatrix} R_{ii} & R_{i(-i)} \\ R_{(-i)i} & R_{(-i)(-i)} \end{bmatrix}$ 이다.

임의로 생성된 잠재변수에 대한 변이의 크기와 중심치를 결정해야 하는 문제(Identifiability)가 발생한다. 이를 위하여 Sorensen 등(1995)은 임계값의 하나를 "0"으로 제한하고 잔차분산을 "1"로 제한하는 방법을 사용할 수 있다고 하였는데 본 연구에서는 2개의 인접한 임계값을 "0"

과 “1”로 제한하여 잠재변수를 생성하였다 (VanTassel 등, 1998).

Gibbs sampling에 의한 임계값의 생성은 Sorensen 등(1995)이 제시한 방법으로 다음과 같이 생성하였다.

$$t_{ik}^n | \theta, R, G, t_{-ik}, U, y^0 \sim UN(\max(\max_j(u_{ij}^* | y_{ij} = \lambda), t_{ik-1}^n), \min(\min_j(u_{ij}^* | y_{ij} = \lambda + 1), t_{ik+1}^n))) \dots\dots\dots (5)$$

여기서 UN은 균일분포함수를 의미한다. 위치모수(Location parameter)와 이산모수(dispersion parameter)에 대한 사후분포특성을 규명하기 위해서는 이들 모수에 대한 사전분포 특성을 정의해야 한다. 본 연구에서는 각 모수에 대한 사전분포를 다음과 같이 가정하였다.

(a)  $f(\beta) \propto \text{cons} \dots\dots\dots (6)$

(b)  $f(a | G) \sim N(0, G \otimes A) \dots\dots\dots (7)$

여기서  $G$ 는 3가지 형질들에 대한  $3 \times 3$  유전 분산 - 공분산 행렬이고  $A$ 는 개체들간의 혈연계수 행렬이다. 유전분산 - 공분산 성분과 환경 분산 - 공분산 성분에 대한 사전정보를 모르기 때문에 Jeffrey의 noninformative priors(Gelman et al., 1995)를 적용하여 아래와 같이 가정하였다.

(c)  $f(R) \propto |R|^{-k/2} \dots\dots\dots (8)$

(d)  $f(G) \propto |G|^{-k} \dots\dots\dots (9)$

각 모수에 대한 사전정보에 따라 사후조건확률분포는 지역모수의 경우 Van Tassel 등(1998)이 제시한 방법에 따라 다변량 정규분포 함수에서 Gibbs sampler를 적용하였으며 분산-공분산 성분은 inverted Wishart 분포함수에서 생성하였다. 각 모수에 대한 Gibbs sampler를 적용하기 위한 모수의 초기치로써 임계값의 초기치는 0부터 1씩 증가하여 부여하였으며, 지역모수의 초기치는 각 요인들의 수준별 단순평균을 적용하였고, 분산 - 공분산 성분은 모의실험자료를 생성하기 위하여 사용된 모수를 적용하였다.

### III 결과 및 고찰

1998년부터 2002년까지 한우 검정우의 거세

를 통한 후대검정에서 조사된 근내지방도의 분포는 Fig. 1에 제시된 바와 같이 정규분포에서 벗어나고 있었다(윤 등, 2002). 하지만 한우 비거세우를 이용한 후대검정에서 조사된 근내지방도의 분포(이 등, 2001)와 비교해 볼 때 비거세우에서 보다 정규성에 상당히 근접하고 있음을 알 수 있었다. 따라서 본 연구에 이용된 모의실험 자료 역시 범주형 자료인 근내지방도의 생성을 비거세우의 자료분포 및 거세우에서의 자료분포와 유사하게 생성하였다. 모의실험을 통하여 생성된 근내지방도에 대한 자료의 분포는 5가지 범주로 조사된 경우(DS1)에 상당히 오른쪽으로 치우친 분포(왜도 = 0.94)인 반면에 DS2의 경우에는 정규분포에 상당히 접근하는 분포특성을 갖도록 구성하였다(Table 2).

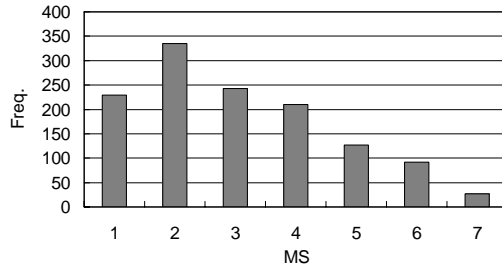


Fig. 1. Frequency of marbling scores(MS) of Hanwoo steers on progeny tests from 1998 to 2002.

Table 2. Information of distributions for linear and categorical traits on pooled simulated data(replicates=10)

	Mean	SD	CV	Skewness	Kurtosis*
CW	310.83	32.96	10.61	-0.00	-0.02
LMA	102.15	8.69	8.50	-0.01	-0.02
MS <sub>L</sub>	51.83	4.90	9.45	-0.00	-0.03
MS <sub>5</sub>	1.68	0.80	47.33	-0.94	-0.21
MS <sub>7</sub>	3.47	1.07	31.00	-0.01	-0.26

CW: carcass weight, LMA: Longissimus Muscle Area, MS<sub>L</sub>: Liability scale of marbling score, MS<sub>5</sub>: 5 grade scale of marbling score, MS<sub>7</sub>: 7 grade scale of marbling score. Kurtosis\*=Kurtosis-3

모의실험을 통하여 얻어진 자료에 대하여 REML방법과 Gibbs sampling 방법으로 추정된 (공)분산 성분과 이를 이용하여 계산된 유전력 및 유전상관 추정결과를 모수와 비교하는 방법

으로 분석방법에 다른 추정의 정확도를 비교하였다. 선형모형에서 REML 방법으로 추정된 분산 및 공분산 성분은 Aitken 가속알고리즘을 적용함으로써 수렴속도를 높일 수 있었다. Gibbs sampling에 의한 각 모수의 사후분포에 대한 안정화는 1,000회 Gibbs sampling 이내에 빠르게 안정화되었다. 특히 근내지방도의 잠재변수 생성을 위하여 2개의 임계값을 제한함으로써 모수를 하나 줄임으로써 기타 임계값에 대한 Gibbs Samples이 빠르게 안정화되어 수렴하는 것을 알 수 있었다. 또한 이들 사후분포에 대한 평균으로 표현한 추정치는 모수와 차이가 없는 것을 알 수 있었다(Table 3). 여기서 잠재변수의 Identifiability를 위한 제한을 가할 때 잔차분산을 “1”로 제한하는 방법 또는

잔차분산의 크기를 변환하는 방법(Lee, 2002)과 비교해 보면, 이산형 자료가 하나일 경우에는 2개의 인접한 임계값을 제한하는 방법은 프로그램이 용이하고 기타 임계값에 대한 사후분포의 수렴이 빨랐는데 이는 모수가 하나 감소하였기 때문인 것으로 사료된다.

Table 4에는 5구간으로 조사된 근내지방도를 포함하는 자료(DS1)에서 REML 방법과 Gibbs 방법으로 추정된 모수의 결과를 비교하여 제시하였다. Table 3에 제시된 바와 같이 일반 선형혼합모형에서 REML 방법으로 분석한 유전력의 추정치는 연속변량이며 선형인 도체중 및 배장 근단면적에서 모수와 t-검정 결과, 유의적인 차이가 없는 것으로 분석되었으며 또한 이들 형질들 간의 유전상관 및 잔차상관에 있어서도 모수와 유의적인 차이가 없는 것으로 추정되었다. 반면에 근내지방도의 유전력 추정치는 모수와 비교하여 유의적으로 낮게 추정되었다( $P < 0.01$ ). 이는 타 연구결과와 일치하는 결과로써 과소 추정되는 정도는 자료의 특성에 따라 차이가 있는 것으로 사료된다. 근내지방도의 연속변량과의 유전상관 추정치는 모수와 유의적인 차이가 없는 것으로 추정되었으나 잔차상관의 경우에는 모수보다 다소 낮게 추정되었다. 이는 형질들 간의 잔차상관이 적어지고 있는 것으로 사료되지만 이에 대한 연구가 필요할 것으로 생각한다.

Table 3. Means and empirical standard deviations of posterior means of thresholds by Gibbs sampler in simulated data(replicates=10)

Threshold Param.	DS with 5 Cat.	DS with 7 Cat			
		Mean	SD	Mean	SD
3	2.000	2.015	0.033	1.957	0.020
4	3.000	3.059	0.078	3.026	0.034
5	4.000	N/A		4.091	0.057
6	5.000	N/A		5.188	0.107

Table 4. Means, empirical standard deviations of Estimates of generic parameters for carcass weight, longissimus muscle area, marbling score with 5 categories by two different models and algorithms(replicates=10)

Param.	REML with LM				GIBBS with TM				
	Est.	SD	Diff.	P value	Est.	SD	Diff.	P value	
<b>Heritability</b>									
CW	0.250	0.254	0.017	0.004	0.472	0.257	0.017	0.007	0.226
LMA	0.286	0.288	0.013	0.002	0.667	0.285	0.013	-0.001	0.900
MS	0.500	0.315	0.011	-0.185	0.000	0.492	0.017	-0.008	0.186
<b>Genetic correlation</b>									
CW-LMA	0.354	0.360	0.032	0.007	0.522	0.366	0.032	0.012	0.245
CW-MS	0.300	0.297	0.044	-0.003	0.808	0.322	0.046	0.022	0.158
LMA-MS	0.141	0.134	0.043	-0.007	0.605	0.158	0.045	0.017	0.263
<b>Residual correlation</b>									
CW-LMA	0.516	0.512	0.013	-0.004	0.349	0.511	0.012	-0.006	0.180
CW-MS	0.058	0.050	0.014	-0.008	0.115	0.048	0.018	-0.010	0.111
LMA-MS	0.313	0.241	0.015	-0.072	0.000	0.303	0.020	-0.010	0.146

Table 5. Means, empirical standard deviations of estimates of generic parameters for carcass weight, longissimus muscle area, marbling score with 7 categories by two different models and algorithms(replicates=10)

Param.	REML with LM				GIBBS with TM				
	Est.	SD	Diff.	P value	Est.	SD	Diff.	P value	
<b>Heritability</b>									
CW	0.250	0.256	0.017	0.006	0.272	0.258	0.017	0.008	0.190
LMA	0.286	0.284	0.013	-0.001	0.765	0.286	0.013	0.000	0.968
MS	0.500	0.462	0.015	-0.038	0.000	0.500	0.016	0.000	0.936
<b>Genetic correlation</b>									
CW-LMA	0.354	0.365	0.031	0.012	0.265	0.366	0.031	0.013	0.227
CW-MS	0.300	0.309	0.039	0.009	0.508	0.307	0.040	0.007	0.580
LMA-MS	0.141	0.151	0.039	0.009	0.473	0.151	0.040	0.009	0.471
<b>Residual correlation</b>									
CW-LMA	0.516	0.511	0.012	-0.006	0.168	0.511	0.012	-0.006	0.166
CW-MS	0.058	0.047	0.016	-0.011	0.056	0.052	0.017	-0.006	0.284
LMA-MS	0.313	0.283	0.015	-0.030	0.000	0.308	0.017	-0.005	0.350

반면에 7개의 범주로 구성된 근내지방도를 포함하는 자료(DS2)에서 유전력 추정치를 살펴보면 (Table 5) 근내지방도를 선형으로 간주하고 REML 방법으로 추정된 근내지방도의 유전력은 모수 0.500보다 0.038이 적은 0.462으로 추정되었는데 이는 5개의 범주형 자료의 근내지방도를 포함하는 DS1에서의 REML 방법에 의한 유전력 추정치 0.315 보다는 다소 높고 모수에 근접한 추정치이었다. REML 방법에 의한 유전상관 및 잔차상과 추정치는 DS1에서와 유사한 경향을 보였지만 편의된 정도의 크기는 DS1 보다는 다소 낮은 것을 알 수 있었다. 반면에 Gibbs Sampling 방법으로 각 모수의 사후 분포함수에 대한 평균치로 예측한 추정치는 유전력, 유전상관 및 잔차상관 모두에서 모수와 유의적인 차이가 없는 것을 알 수 있었다. 따라서 범주형자료에 대한 유전분석을 실시할 때, 이들 자료를 선형으로 간주하고 유전분석하는 것은 실제 모수보다 과소 추정될 수 있으며 그 크기는 범주형 자료의 크기 및 분포특성에 따라 다르다는 것을 알 수 있었다. 반면에 이들 자료를 이산형으로 간주하고 잠재변수가 존재한다는 가정하에 임계형 모형을 설정하고 유전분석을 실시하면 모수에 근접한 추정치를 얻을 수 있음을 알 수 있었다. 따라서 이러한 자료에 대한 유전분석을 실시할 때, 임계형 모형이 타당할 것으로 사료되었다.

#### IV 요약

한우 중모우 선발을 위한 유전능력 평가에서 고려되는 형질들 중 이산형 형태로 조사되는 근내지방도의 유전변이가 추정방법에 따라 어느 정도 차이가 있는지 알아보기 위한 모의실험을 실시하였다. 모의실험 자료는 연속변량으로 간주되는 도체중 및 배장근단면적과 근내지방도의 잠재변수를 다변량 정규분포함수에서 생성하였고 근내지방도의 잠재변수를 이용하여 특정 임계값을 중심으로 순서화된 근내지방도 점수로 변환하였다. 근내지방도의 점수 부여방법으로써 비거세우에서 조사된 근내지방도의 점수 1~5점 사이에 정규분포에서 크게 어긋나는 분포특성을 갖도록 자료(DS1)를 생성하였고 또한 한우 거세우에서 현재 조사되고 있는 점수 1~7점 사이에 정규분포에 좀더 접근한 분포특성을 갖는 모의 자료(DS2)를 생성하였다. 분석방법간에 유전변이 추정의 정확도를 알아보기 위하여 1) 생성된 이들 자료를 선형으로 간주하고 다형질 혼합 선형 개체모형에서 REML 분석방법으로 유전변이를 추정하였고 2) 특정 임계치를 중심으로 잠재변수가 존재한다는 가정하에 다형질 임계 개체 혼합모형을 설정하여 Gibbs sampling 방법으로 유전변이를 추정하였다. 여기서 추정된 유전변이(유전력, 유전상관 및 잔차상관)에 대하여 모수와의 차이를

검정함으로써 편의되는 정도를 알아보았다. 모의 실험은 각 자료에 대하여 10회 실시하였다.

분석결과, 근내지방도의 유전력 추정치는 DS1에서는 다형질 임계개체혼합모형을 설정하여 Gibbs sampling 방법으로 모수에 대한 사후분포의 평균으로 계산한 결과 참값과 유의적인 차이가 없는 것으로 분석되었다. 반면에 근내지방도를 선형으로 간주하고 다형질 선형 개체혼합모형에 의한 유전력 추정치는 모수보다 매우 낮은 유전력을 보였다(0.500 vs 0.315). 유전상관 추정치는 선형모형에서의 REML방법 또는 임계모형에서의 Gibbs sampling 방법에서 모두 모수와 유의적인 차이가 없는 것으로 분석되었으나 근내지방도의 잔차상관에 있어서 REML 방법으로 분석하였을 경우에 모수보다 낮게 추정되었다. 반면에 범주형 모형에서는 모수와 추정치 간에 유의적인 차이가 없는 것으로 분석되었다. 또한 7개의 범주형으로 조사된 자료(DS2)에서 이들 추정치는 DS1에서와 동일한 경향을 보였는데 그 편의 정도는 다소 적어지는 경향을 보였다. 따라서 이산형으로 조사되는 근내지방도에 대한 유전변이를 추정하기 위해서는 범주형 임계모형이 선형모형 보다 다소 정확한 추정을 할 수 있을 것으로 판단되었다.

## V 사 사

본 연구는 한국학술진흥재단 2003년 신진교수연구지원사업(KRF-2003-003-F00029)의 연구비 지원에 의하여 수행되었음.

## VI 인용 문헌

- Falconer, D. S. and Mackay, T. F. C. 1996. Introduction to Quantitative Genetics. Edinburgh, Longman. 4<sup>th</sup> Ed.
- Gelman, A., Carlin, J. B., Stern, H. S. and Rubin, D. B. 1995. Bayesian data analysis. Chapman & Hall.
- Geman, S. and Geman, D. 1984. Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. IEEE Trans. Pattern Anal. Mach. Intel. 6:721-741.
- Gianola, D. and Fouley, J. L. 1983. Sire evaluation for ordered categorical data with a threshold model. Genet. Sel. Evol. 15:201-244.
- Lee, D. H. 2002. Investigation of biases for variance components on multiple traits with varying number of categories in threshold models using Bayesian inferences. Asian-Aust. J. Anim. Sci. 15:925-1070.
- Little, R. J. A. and Robin, D. B. 2002. *Statistical analysis with missing data*. John Wiley & Sons. Inc.
- Matos, C. A. P., Thomas, D. L., Gianola, D., Tempelman, R. J. and Young, L. D. 1997. Genetic analysis of discrete reproductive traits in sheep using linear and nonlinear models: I. Estimation of genetic parameters. J. Anim. Sci. 75:76-87.
- Meuwissen, T. H. E. and Luo, Z. 1992. Computing inbreeding coefficients in large populations. Genet. Sel. Evol. 24:305-313.
- Phocas, F. and Laloe, D. 2003. Evaluation models and genetic parameters for calving difficulty in beef cattle. J. Anim. Sci. 81:933-938.
- Ramirez-Valverde, R., Misztal, I. and Bertrand, J. K. 2001. Comparison of threshold vs linear and animal vs sire models for predicting direct and maternal genetic effects on calving difficulty in beef cattle. J. Anim. Sci. 79:333-338.
- Sorensen, D. A., Andersen, S., Gianola, D. and Korsgaard, I. 1995. Bayesian inference in threshold models using Gibbs sampling. Genet. Sel. Evol. 27: 229-249.
- Tsuruta, S., Misztal, I. and Strandén, I. 2001. Use of preconditioned conjugate gradient algorithm as a generic solver for mixed-model equations in animal breeding applications. J. Animal Sci. 1166-1172.
- VanTassell, C. P., VanVleck, L. D. and Gregory, K. E. 1998. Bayesian analysis of twinning and ovulation rates using a multiple-trait threshold model and Gibbs sampling. J. Anim. Sci. 76: 2048-2061.
- Varona, L., Misztal, I. and Bertrand, J. K. 1999. Threshold-linear versus linear-linear analysis of birth weight and calving ease using an animal model: I. Variance component estimation. J. Anim. Sci. 77:1944-2002.
- 윤호백, 김시동, 나승환, 장은미, 이학교, 전광주, 이득환. 2002. 거세한우의 도체형질에 대한 유전 모수 추정. 한국동물자원과학회지 44(4):383-390.
- 이득환. 2002. 다형질 Threshold 개체모형에서 missing 기록을 포함한 이산형 자료에 대한 Bayesian 분석. 한국 동물 자원 과학 회지 44(2):151-164.
- 이득환, Misztal, I. and Bertrand, J. K. 2001. Bayesian analysis of carcass traits using multivariate threshold animal models and Gibbs sampling with missing records in Korean cattle. 한국동물자원과학회지 43(1):9-22.
- 축산기술연구소. 2003. 2003년 한우유전능력평가 보고서. 농촌진흥청 축산기술연구소. (접수일자 : 2004. 2. 19. / 채택일자 : 2004. 7. 1.)