Turk J Vet Anim Sci
25 (2001) 855-862

Hayvan Islahında Bayesian Yöntemi Kullanılarak Çoklu Karakter Analizi: Gibbs Örneklemesi Yaklaşımı

Mehmet Ziya FIRAT
Akdeniz Üniversitesi, Ziraat Fakültesi, Zootekni Bölümü, Antalya - TÜRKİYE

Geliş Tarihi: 27.01.2000

Özet: Birçok ıslah denemesi, herbir bireyden ölçülen ekonomik olarak önemli birden fazla karaktere dayanmaktadır. Elde edilen verilerin istatistiksel analizinde, aynı bireylerden kaydedilen karakterler çoğu kez teker teker dikkate alınırlar. Bununla birlikte, genellikle sadece belirli bir karakterin kalıtsallığının şekli ile değil fakat aynı zamanda onun diğer karakterlerle ilişkileri ile de ilgilenilmektedir. Çoklu karakter analizleri karakterler arasındaki genetik ve fenotipik korelasyonlar hakkında yorumlamalar yapmayı gerektirmektedir, fakat böylesi parametrelerin nokta tahminleri, yüzlerce hayvandan elde edilen ıslah verileri kullanılsa dahi zayıf olabilir. Bayesian yöntemleri parametre uzayı içersinde olmayan varyans matrislerini dahil etmemektedirler ve sadece nokta tahminleri verme yerine olabilirlik fonksiyonu ve prior dağılışındaki parametreler hakkındaki tüm bilgileri kullanırlar. Bu çalışmada, baba-bir üvey kardeş ailelerini temsil eden dengeli tek yönlü çok karakterli bir boğa modeli, iki farklı prior tanımlamalı bir Gibbs örneklemesi yaklaşımı kullanılarak incelenmiştir. İki varyans matrisi için invers-Wishart dağılışları ve ortalama vektörü için üniform dağılış prior olarak kullanılmıştır. Gibbs örneklemesinin sonuçları varyans analiz metodundan elde edilen parametre tahminleri ile karşılaştırılmıştır. Gibbs örneklemesi algoritmasını kullanarak yapılan Bayesian analizinin herbir bilinmeyen parametrelerin tam marjinal posterior dağılışının bir tahminini verdiği ve aynı zamanda geleneksel yöntemlerin aksine parametre uzayı içersinde nokta tahminleri verdiği gösterilmiştir.

Anahtar Sözcükler: Bayesian yorumlama, çok değişkenli boğa modeli, Gibbs örneklemesi, genetik ve fenotipik varyans matrisleri

Multiple-trait Analysis Using Bayesian Methods in Animal Breeding: A Gibbs Sampling Approach

Abstract: Most animal breeding experiments are based on more than one economically important trait measured in each individual. In the statistical analysis of the resulting data, traits recorded in the same individuals are often considered one at a time. Usually we are interested, however, not only in the mode of inheritance of a particular trait but also in its relationships with other traits. Multivariate analyses are required to make inferences about genetic and phenotypic correlations between traits, but point estimates of such parameters can be poor even when breeding data on hundreds of animals are used. Bayesian methods exclude variance matrices which are not within the parameter space, and make use of all the information on parameters in the likelihood function and prior distribution, rather than just providing point estimates. In this study, a balanced one-way multiple-trait sire model representing half-sib families is investigated using a Gibbs sampling approach with two different prior specifications. Inverse Wishart distributions for the two variance matrices and a uniform distribution for the mean vector are used as priors. The results of Gibbs sampling are compared with estimates of the parameters obtained from the analysis of variance method. It is shown that a Bayesian analysis using a Gibbs sampling algorithm provides an estimate of the complete marginal posterior distribution of each unknown parameter and also gives point estimates which are within the parameter space, in contrast to conventional procedures.

Key Words: Bayesian inference, multivariate sire model, Gibbs sampling, genetic and phenotypic variance matrices

Giriş

Hayvan ıslahının amacı, ıslah edilen damızlık adayı hayvanların seçimi ile çiftlik hayvanlarının genetik ilerlemesini sağlamaktır. Seleksiyon, herbir bireyden ölçülen birden fazla sayıdaki ekonomik değeri yüksek özelliğe göre yapılır. Sadece bir karakter ekonomik öneme sahip olsa dahi, seleksiyonun etkinliğini geliştirmek için bununla ilgili karakterler de ölçülebilir. Bundan dolayı, böylesi gözlemlerin analizi için çok karakterli modele gereksinim vardır.

Genetik değerlendirmelerde çoklu karakter analiz yönteminin başlıca iki avantajı vardır. Bunlardan ilki, çoklu karakter yöntemlerinin, tekli metodlara göre bireyleri değerlendirmek için daha fazla bilgi kullanmasıdır. Genetik ve çevresel korelasyonların mutlak değerleri yüksek olabileceğinden, bütün karakterlerin çoklu analizini yapmakla, tahmindeki doğruluk ve sonuç olarak seleksiyondaki ilerleme artar. Diğeri ise, bazı karakterler sınırlı sayıdaki bireyler üzerinde ölçülmüşlerdir ve bu karakterlerin populasyondaki daha fazla sayıdaki bireyler üzerinde ölçülen diğer karakterlerle birlikte analiz edilerek tahminde daha fazla doğruluk kazanılabilmesidir (1).

Bununla birlikte, çoklu karakter analizinin en büyük dezavantajı, çözülecek eşitlik sayısında artma olmasından dolayı ilave hesaplamalara ihtiyaç göstermesidir. Diğer bir dezavantajı, analiz edilecek karakter sayısı arttıkça negatif tanımlı varyans matrisi tahminlerinin olasılıklarında artma olmasıdır (1, 2). Varyans kovaryans unsurlarının tahmini, kantitatif genetikçiler tarafından iki veya daha fazla sayıdaki karakterler arasındaki genetik ve çevresel ilişkilerin ölçümünde ve ıslah yöntemlerini formule etmede kullanılırlar. Varyans analiz (ANOVA) yöntemleri genetik ve fenotipik parametreleri tahmin etmede yaygın olarak kullanılmışlardır. ANOVA yöntemine göre yapılan çoklu karakter analizinde, kareler ve çarpımlar toplamı beklenen değerlerine eşitlemek suretiyle varyans kovaryans matrislerinin tahminlerini elde etmek mümkündür. Fakat kareler ve çarpımlar ortalaması arasındaki farka dayalı metodlar, özellikle çok sayıdaki karakterden ölçüm yapıldığı zaman negatif tanımlı genetik varyans matrisleri tahmini elde edilmesine neden olurlar (2).

Son yıllarda, hayvan ıslahında varyans unsurları hakkında yorumlamalar yapmak için Bayesian yöntemleri geliştirilmiştir (3, 4). Posterior dağılışların tahminleri için kullanılan Gibbs örneklemesi yöntemi Wang ve ark. (5, 6) ve Fırat ve ark. (7) tarafından tek değişkenli karışık bir doğrusal modeldeki varyans unsurları tahmini için geliştirilmiştir. Foulley ve ark. (8) t binom karakteri için çok değişkenli deneysel Bayesian yaklaşımı kullanmışlardır. Sabit ve şansa bağlı parametreler için sabit prior varsaymışlardır. Bununla birlikte, hayvan ıslahı uygulamalarında karakterlerin çoğu sürekli dağılış göstermektedir ve böyle karakterlerin çoklu-karakter analizleri yakın yıllarda Varona ve ark. (9), Jensen ve ark. (10) ve Jensen (11) tarafından Bayesian yöntemleri kullanılarak yapılmıştır.

Bu makalenin amacı, baba-bir üvey kardeş ailelerini temsil eden dengeli tek yönlü çok karakterli bir boğa modeli için iki farklı prior tanımlamalı bir Gibbs örneklemesi yaklaşımı kullanmaktır. Foulley ve ark. (8)'nın aksine, varyans unsurları ve bunların fonksiyonları hakkında Bayesian marjinal yorumlamalar yapmada şansa bağlı parametreler için beli başlı dağılışlardan priorlar kullanılmıştır. Gibbs örneklemesinin sonuçları varyans analiz metodundan elde edilen parametre tahminleri ile karşılaştırılmıştır.

Metot

Model ve Varsayımlar

Eşit büyüklükteki, n, baba-bir üvey kardeş gruplarından (s) ibaret dengeli tek yönlü bir sınıflama olduğu ve herbir bireyden t karakter ölçüldüğü varsayılsın. y_{ij} (i=1,...,s; j=1,...,n) i'inci boğa grubundaki j'inci hayvandan gözlenen t karaktere ait değerden oluşan vektör olsun. Bu durumda çok değişkenli doğrusal şansa bağlı model şöyle verilebilir

$$y_{ij}=\mu+s_i+e_{ij} \qquad (i = 1,...,s; j = 1,...,n), \qquad [1]$$

burada μ, ortalamaların t vektörü, s_i boğa etkilerinin t vektörü ve e_{ij} t hata vektörüdür. Boğa etkileri ve hataların vektörlerinin birbirlerinden bağımsız ve normal dağılış gösterdikleri varsayılır; $s_i \sim N_t(0, \Sigma_s)$ ve $e_{ij} \sim N_t(0, \Sigma_e)$ olup Σ_e pozitif tanımlıdır. Böylece y_{ij} vektörleri müşterek çoklu normal dağılış gösterirler ve ortalamaları μ ve ikinci momentleri şöyledir

$Var(y_{ij}) = \Sigma_p = \Sigma_s + \Sigma_e$, $Cov(y_{ij}, y_{ij'}) = \Sigma_s$ $(j \neq j')$, ve $Cov(y_{ij}, y_{i'j'}) = 0$ $(i \neq i')$,

burada Σ_p fenotipik varyans matrisidir.

Tek yönlü çoklu karakter varyans analizinden, fenotipik, genetik, boğa ve hata varyans kovaryans matrislerinin tahminleri sırasıyla şöyle elde edilir:

$$\widehat{\Sigma}_p = \left[M_b + (n-1)M_w\right] / n, \ \widehat{\Sigma}_g = 4(M_b - M_w) / n,$$

$$\widehat{\Sigma}_s = (M_b - M_w) / n, \ \widehat{\Sigma}_e = M_w.$$

M_b ve M_w gruplar arası ve gruplar içi kareler ve çarpımlar ortalamalarının matrisleridir

$$M_b = n \sum_{i=1}^{s} (\bar{y}_{i.}-\bar{y}_{..}) (\bar{y}_{i.}-\bar{y}_{..})'/(s-1) \text{ ve}$$

$$M_W = \sum_{i=1}^{s} \sum_{j=1}^{n} (y_{ij}-\bar{y}_{i.}) (y_{ij}-\bar{y}_{i.})'/(s(n-1))$$

burada $\bar{y}_{i.}$ i'inci gruba ait ortalama vektörü, ve $\bar{y}_{..}$ genel ortalama vektörüdür.

Eşitlik [1]'de çoklu karakter baba bir üvey kardeş modeli Σ_s ve Σ_e matrisleri üzerine bazı kısıtlamalar getirir. Hayvan ıslahında, c'y karakterlerinin herhangi bir doğrusal kombinasyonu kalıtım derecesini, $4c'\Sigma_s c/c'(\Sigma_s + \Sigma_e)c'$, verir. Kalıtım derecesi ebeveynlerdeki varyasyonun döllere aktarılabilen oranıdır. Böyle bütün kombinasyonlar için kalıtım derecesinin 1'den büyük olmaması gerekmektedir. Bu tür kısıtlama $\Sigma_e - 3\Sigma_s$'nin pozitif tanımlı olması anlamına gelmektedir. Bu kısıtlamayı dikkate almayan tahmin metodu, bazı karakterler veya bunların doğrusal kombinasyonlarına ait kalıtım derecesinin negatif veya 1'den büyük tahminlerini verebilir.

Bayesian Formülasyonu

Bayesian analizi modeldeki bilinmeyen parametrelere, μ, $\{s_i\}$, Σ_s ve Σ_e, prior dağılış atamayı gerektirmektedir. μ'nün üniform dağıldığı varsayılırsa olasılık fonksiyonu şöyledir

$$f(\mu) = \text{sabite}. \qquad [2]$$

Σ_s verildiğinde $\{s_i\}$'nin prior dağılışı çok değişkenli normal dağılış $N_t(0, \Sigma_s)$'dır, böylece

$$f(\{s_i\} \mid \Sigma_s) \propto |\Sigma_s|^{-\frac{1}{2}s} \exp\left[-\frac{1}{2}\mathrm{tr}\left(\sum_{i=1}^{s} s_i s'_i \Sigma_s^{-1}\right)\right] \qquad [3]$$

burada $|\Sigma_s|$ Σ_s'nin determinantını temsil etmektedir. Σ_s ve Σ_e'nin aşağıdaki müşterek prior dağılışına sahip oldukları varsayılsın

$$f(\Sigma_s, \Sigma_e | v_s, S_s, v_e, S_e) \propto |\Sigma_s|^{-\frac{1}{2}(v_s+t+1)} |\Sigma_e|^{-\frac{1}{2}(v_e+t+1)}$$
$$\times \exp\left[-\frac{1}{2}\mathrm{tr}\left(v_s \Sigma_s^{-1} S_s + v_e \Sigma_e^{-1} S_e\right)\right] \qquad 0<\Sigma_s<1/3\ \Sigma_e. \qquad [4]$$

$\Sigma_e - 3\Sigma_s$'nin pozitif tanımlı olma koşulu olmaksızın, [4] nolu eşitlik Σ_s ve Σ_e'nin invers-Wishart dağılışlarına, $W^{-1}(v_s, v_s S_s)$ ve $W^{-1}(v_e, v_e S_e)$, sahip oldukları anlamına gelmektedir. Eşitlik [4]'deki S_s^{-1} ve S_e^{-1}, Σ_s^{-1} ve Σ_e^{-1}'nin prior beklenen değerleri olarak ve v_s ve v_e serbestlik dereceleri olarak yorumlanabilirler. Model [1], μ, $\{s_i\}$, Σ_s ve Σ_e verildiğinde veriler $\{y_{ij}\}$'nin aşağıdaki olabilirlik fonksiyonuna sahip olduğunu belirtmektedir

$$f(\{y_{ij}\} | \mu, \{s_i\}, \Sigma_s, \Sigma_e) \propto |\Sigma_e|^{-\frac{1}{2}sn}$$
$$\exp\left[-\frac{1}{2}\sum_{i=1}^{s}\sum_{j=1}^{n}(y_{ij}-\mu-s_i)'\Sigma_e^{-1}(y_{ij}-\mu-s_i)\right]. \qquad [5]$$

μ, $\{s_i\}$, Σ_s ve Σ_e için [2], [3] ve [4]'de verilen prior yogunluk fonksiyonları ve [5]'deki olabilirlik fonksiyonunu kullanarak, $\{y_{ij}\}$ verildiğinde parametrelerin müşterek yogunluk fonksiyonu aşağıdaki gibi elde edilebilir

$$f(\mu, \{s_i\}, \Sigma_s, \Sigma_e | \{y_{ij}\})$$
$$\propto |\Sigma_s|^{-\frac{1}{2}(s+v_s+t+1)} |\Sigma_e|^{-\frac{1}{2}(sn+v_e+t+1)}$$
$$\times \exp\left\{-\frac{1}{2}\mathrm{tr}\left[\Sigma_s^{-1}\left(v_s S_s + \sum_{i=1}^{s} s_i s_i'\right)\right]\right\}, \qquad [6]$$
$$\times \exp\left\{-\frac{1}{2}\mathrm{tr}\left[\Sigma_e^{-1}\left(S_w + v_e S_e + n\sum_{i=1}^{s}(\bar{y}_{i.}-\mu-s_i)(\bar{y}_{i.}-\mu-s_i)'\right)\right]\right\}$$

burada S_w gruplar arası kareler ve çarpımlar toplamı matrisini temsil etmektedir.

Parametrelerin hepsi, μ, $\{s_i\}$, Σ_s ve Σ_e, θ ile gösterilsin ve $\pi(\theta)$ ilgi duyulan bir fonksiyon olsun. $\pi(\theta)$'nın posterior beklenen değerinin bulunmak istendiği varsayılsın

$$E[\pi(\theta | \{y_{ij}\})] = \int_\Omega \pi(\theta) f(\theta | \{y_{ij}\}) d\theta \qquad [7]$$

burada $f(\theta|\{y_{ij}\})$ [6]'da verilen müşterek posterior fonksiyonu ve Ω θ'nın tanım aralığıdır. Bununla ilgili olarak en az iki problem mevcuttur. Bunlardan ilki, [7]'nin analitik yöntemlerle elde edilmesinin mümkün olmamasıdır. İkincisi, standart Monte Carlo yaklaşımlarının böylesine çok katlı integral problemine çözüm olabilmesine ragmen bunu yapmanın kolay olmamasıdır. Çünkü marjinal posterior yogunluk fonksiyonu bilinmeyen formda olabilir ve böylece bu fonksiyondan örnek çekmek zordur. Bu sorun verilerin artırılması yöntemleri uygulanarak hafifletilebilir. Fakat, modeldeki diğer bütün parametreler verildiğnde μ, $\{s_i\}$, Σ_s ve Σ_e'nin herbirinin şartlı posterior dağılışlarını kullanarak müşterek posterior dağılıştan örnekler çekmeyi sağlayan Gibbs örneklemesi yaklaşımı bu problemin üstesinden gelmek için uygulanabilir.

Model [1]'deki model için Gibbs örneklemesi yapabilmek için, μ, $\{s_i\}$, Σ_s ve Σ_e'nin tam şartlı posterior dağılışlarına, yani diğer parametreler verildiğinde bunların herbirinin şartlı dağılışına, ihtiyaç vardır. Şartlı posterior dağılışlar, hem posterior dağılışın yapısını ve hem de etkili hesaplamanın esasını anlamayı sağlamaktadır. Tam şartlı posterior dağılışlar şunlardır:

$$\left[\mu \mid \{s_i\},\Sigma_s,\Sigma_e,\{y_{ij}\}\right]=N_t\left(\bar{y}_{..}-\bar{s}_{.},(sn)^{-1}\Sigma_e\right) \quad [8]$$

$$\left[\{s_i\} \mid \mu,\Sigma_s,\Sigma_e,\{y_{ij}\}\right]=N_t\left(n\Sigma_s(n\Sigma_s+\Sigma_e)^{-1}(\bar{y}_{i.}-\mu),\Sigma_s(n\Sigma_s+\Sigma_e)^{-1}\Sigma_e\right) \quad [9]$$

$$\left[\Sigma_s \mid \mu,\{s_i\},\Sigma_e,\{y_{ij}\}\right]=W_t^{-1}\left(s+v_s,\sum_{i=1}^{s} s_i s_i'+v_s S_s\right) \quad [10]$$

$$\left[\Sigma_e \mid \mu,\{s_i\},\Sigma_s,\{y_{ij}\}\right]=W_t^{-1}\left(sn+v_e,\sum_{i=1}^{s}\sum_{j=1}^{n} (y_{ij}-\mu-s_i)'(y_{ij}-\mu-s_i)+v_e S_e\right) \quad [11]$$

Böylece, Gibbs örneklemesi, μ, ve $\{s_i\}$ için [8] ve [9]'daki çok değişkenli normal dağılışlardan ve Σ_s ve Σ_e için [10] ve [11]'deki invers Wishart dağılışlarından örnekleme yapmaktadır. [8]-[11]'deki dağılışlar Wang ve ark. (5, 6) ve Gelfand ve ark. (12) tarafından tek değişkenli tek-yönlü sınıflandırma için verilen dağılışları çok değişkenliler için genelleştirmektedirler.

Gibbs örneklemesi

Geman ve Geman (13) tarafından sunulan Gibbs örneklemesi, karmaşık stokastik modeller hakkında yorumlamalar yapmak için oldukça güçlü ve giderek artan bir biçimde önem kazanan bir yöntemdir. Gelfand ve ark. (12) ve Gelfand ve Smith (14) bunun Bayesian yorumlamaya olan ilişkisini göstermişlerdir. Gibbs örneklemesi sıra ile bütün şartlı dağılışlardan örnekleme yaparak modeldeki parametrelerin hepsinin müşterek posterior yogunluk dağılışına örnek değerler üretir. Basit mantıksal dayanağı ve uygulama kolaylağından dolayı bu yöntem oldukça yaygın hale gelmiştir. Yukarıdaki model için algoritmanın tamamı (Σ_e -3Σ_s için verilen kısıtlama dahil) aşağıdaki gibidir:

i) Her parametreye keyfi bir başlangıç değeri $\theta_0=(\mu_0, \{s_{i0}\}, \Sigma_{s0}, \Sigma_{e0})$ seçilir;

ii) $[\mu|\{s_{i0}\}, \Sigma_{s0}, \Sigma_{e0}, \{y_{ij}\}]$'dan bir değer μ_1 çekilir ve μ güncelleştirilir;

iii) $[s_i \mid \mu_1, \Sigma_{s0}, \Sigma_{e0}, \{y_{ij}\}]$'dan bir değer $\{s_{i1}\}$ çekilir ve $\{s_i\}$ güncelleştirilir;

iv) $[\Sigma_s \mid \mu_1, \{s_{i1}\}, \Sigma_{e0}, \{y_{ij}\}]$'dan bir değer Σ_{S1} çekilir ve Σ_s güncelleştirilir;

v) $[\Sigma_e \mid \mu_1, \{s_{i1}\}, \Sigma_{S1}, \{y_{ij}\}]$'dan bir değer Σ_{e1} çekilir ve Σ_e güncelleştirilir;

vi) Σ_e -3Σ_s'nin pozitif tanımlı olup olmadığı kontrol edilir; eğer değilse, pozitif tanımlı olana kadar iv) ve v)'i tekrar edilir.

Bu altı adım Gibbs örneklemesinin tek bir döngüsünü oluşturmaktadır.

vii) Adım ii)-vi) güncelleşmiş değerler kullanarak m defa tekrarlanır ve bir dizi $\{\theta_l\}=(\mu_l, \{s_{il}\}, \Sigma_{sl}, \Sigma_{el})$, l=1,...,m değerleri elde edilir.

m arttıkça, $\{\theta_l\}$ $f(\mu, \{s_i\}, \Sigma_s, \Sigma_e \mid \{y_{ij}\})$ müşterek posterior dağılışından bir tesadüf örneğine yaklaşır. Gibbs örneklemesi teorisine dayanarak, bir dizi $\{\theta_l\}$, (l=1,...,m), tesadüf örnekleri çekilebilir ve posterior ortalama [7]'nin sayısal yaklaşımı aşağıdaki gibi verilir:

$$E\left[\pi\left(\theta \mid \{y_{ij}\}\right)\right]=\frac{1}{m}\sum_{l=1}^{m}\pi(\theta_l). \quad [12]$$

Gibbs örneklemesinin bu şekildeki uygulanması örnek değerlerinin tamamını kaydeder. Bununla birlikte, algoritmayı uygulamanın en az iki değişik yolu vardır: farklı başlangıç değerlerine sahip çoklu kısa zincirler veya her k'ıncı iterasyonu kaydeden uzunca tek bir zincir. Dizinin populasyon parametresine yaklaşmasını sağlamak için bazı başlangıç iterasyonları atılabilir. Gelfand ve ark. (12) ve Gelfand ve Smith (14) çoklu kısa zincir algoritmasını kullanmışlardır. Simüle edilmiş verilerle olan deneyimler ve Raftery ve Lewis (15)'nin teorik varsayımları, mevcut durum için posterior beklenen değerlere yakın başlangıç değerleri kullanılması halinde böylesi karmaşalıkların gereksiz olduğu inancını pekiştirmiştir. Bu yöntem tek değişkenli karışık bir doğrusal model için Wang ve ark. (6) tarafından da kullanılmıştır.

Sonuçlar ve Tartışma

Çok sayıdaki karakter için herbir boğaya eşit sayıda döl olacak şekilde gözlemler elde etmek üzere model [1]'e dayalı olarak farklı miktardaki istatistiksel bilgiyi temsil eden farklı aile büyüklükleri, kalıtım dereceleri ve karakter sayısına sahip deneme desenleri Monte Carlo yöntemiyle simüle edilmiştir. Boğa başına düşen döl sayısı, n, 8 veya 20 iken boğa sayısı 25 veya 80 olarak alınmıştır. Simülasyonda kullanılacak Σ_s ve Σ_e'nin değerlerinin seçiminde, eğer Σ_e pozitif tanımlı ise bu durumda Σ_e=PP' ve Σ_s=PΛP' olacak biçimde singüler olmayan bir t x t P matrisi bulmanın daima mümkün olacağına dikkat edilmelidir, burada Λ bir köşegen matrisdir ve P genotipik ve fenotipik olarak aralarında korelasyon bulunmayan karakterlerin kanonik transformasyonunu temsil eder. Böylece matrisler $\Sigma_e+\Sigma_s = I_p$ koşulunu sağlayacak duruma indirgenir, Σ_s burada köşegen matristir. Bu durumda, Σ_s'nin köşegen elemanları kalıtım derecelerinden oluşan bir vektör h^2 ile tanımlanır; herbir köşegen elemanı karşılığı olan kalıtım derecesinin dörtde biridir. Burada σ_s^2 ve σ_e^2 boğa ve hata varyanslarının, Σ_s ve Σ_e, köşegen elemanlarına karşılık gelmektedir ve $h^2 = 4\ \sigma_s^2 / (\sigma_s^2 + \sigma_e^2)$'dir.

Simülasyonlarda 500 defa tekrarlanan veri setleri kullanılmış ve parametrelerin ANOVA tahminleri ve posterior beklenen değerlerine ait istatistikler her deneme deseni için bu veri setlerinin ortalaması alınarak bulunmuştur. Sayısal sonuçlar t=2, 4 ve 6 karakter için elde edilmiş fakat bu bölümdeki tablolar sadece t=4 karakter için verilmiştir. Farklı kalıtım dereceleri ve aile büyüklüklerine sahip dört karakter (t=4) için 500 örnek üzerinden elde edilen σ_s^2, σ_e^2 ve h^2 parametrelerinin ANOVA tahminlerinin ortalama ve standart sapmaları Tablo 1'de verilmiştir. Bu tablodaki sonuçlar parametrelerin tanım aralığı dışındaki değerleri de içermektedir, yani negatif kalıtım derecesi olduğu durumlar. Tablodan da görülebileceği gibi, aile büyüklüğü arttıkça parametre tahminleri gerçek parametre değerlerine yaklaşmaktadırlar.

Bu çalışmada, Raftery ve Lewis (15)'nin yöntemi kullanılarak 'yakınsama' tayininin incelenmesi gerçekleştirilmiştir ve atılması gereken başlangıç adımlarının sayısının ihmal edilebilir olduğu bulunmuştur. Böylece 'yakınsama' sadece 1,000 adımla başarılmıştır. 500 veri setinin herbiri için 1,000 adımlık Gibbs örneklemesi kullanılmıştır ve ilgi duyulan parametreler hakkındaki yorumlamalar tüm adımlar esas alınarak yapılmıştır. Bayesian analizleri iki prior tanımlaması kullanılarak yapılmıştır. Bunlar aşağıda verilmiştir:

i) Prior1 ile gösterilecek olan ilk prior tanımlamasında, prior parametreleri S_s ve S_e gerçek parametre değerlerinin, Σ_s ve Σ_e, aynısı olarak seçilmişlerdir.

ii) Prior2 ile gösterilen ikinci tanımlamada, prior parametreleri birim matrisine oransal olarak seçilmişlerdir, yani $S_s=(1-a)I_t$ ve $S_e=aI_t$ burada a hata varyans matrisinin, Σ_e, köşegen elemanlarının ortalamasıdır. Örneğin, köşegen elemanları Σ_s=köş(0.05 0.15) ve Σ_e=köş(0.95 0.85) olduğunda $S_s=0.10I_2$ ve $S_e=0.90I_2$'dir.

Marginal posterior beklenen değerlerin prior tanımlamalardaki farklılıklardan nasıl etkilendiklerini göstermek amacıyla iki farklı prior tanımlaması kullanılmıştır. Her iki prior tanımlamasında da, serbestlik dereceleri karakter sayısına eşitlenmiştir, $v_s=v_e=t$; bunlar prior dağılışların belirli bir dağılıştan olması için en küçük değerlerdir. Farklı kalıtım dereceleri ve aile büyüklüklerine sahip iki prior tanımlaması için parametrelerin posterior beklenen değerlerine ait özellikler Tablo 2'de verilmiştir.

Tablo 2'nin sonuçları, küçük aile büyüklüğü ve düşük kalıtım dereceli deneme desenleri için Prior1 tanımlamalı Bayesian yönteminin varyans unsurları ve kalıtım derecelerini gereğinden daha fazla tahmin ettiğini belirtmektedir. Bunun dışında, parametrelerin posterior beklenen değerlerinin neredeyse tamamen gerçek parametre değerlerinin aynısı olduğu görülmektedir. Posterior ortalamaları esas alan tahminlerin örnekleme varyansı sistematik olarak ANOVA'nınkinden daha küçüktür.

Tablo 2'nin Prior2 için verilen sonuçları biraz faklı yorumlama vermektedir. Bayesian yönteminin büyük olan gerçek kalıtım derecesine karşılık gelen varyans unsuru ve kalıtım derecelerini daha az tahmin ettiği ve aile büyüklüğü küçük olduğu zaman küçük kalıtım derecesine karşılık gelen aynı parametreleri daha fazla tahmin ettiği görülmektedir. Bunun sebebi, marjinal posterior beklenen değerlerin küçük aileler için prior dağılışın etkisini göstermesi ve verilerin parametreler hakkında, özellikle Σ_s ve bunun fonksiyonları hakkında, çok az bilgi sağlamasıdır.

Bu makalede, Gibbs örneklemesinin, hayvan ıslahı uygulamalarında ortaya çıkan dengeli çok karakterli bir boğa modelindeki bütün parametrelerin Bayesian analizinde başarılı bir biçimde kullanılabileceği ortaya konulmuştur. Gibbs örneklemesi parametrelerin integralini mümkün kılmakta ve ilgi duyulan parametrelerin marjinal posterior dağılışlarının Monte Carlo tahminini vermektedir. Gibbs örneklemesi kullanılarak yapılan bir Bayesian analizi her parametrenin tam marjinal posterior dağılışının bir tahminini verir ve aynı zamanda ANOVA gibi geleneksel yöntemlerin aksine, parametre uzayı içersinde olan nokta tahminleri verir.

Marjinal posterior beklenen değerlerin küçük örnek büyüklüğüne sahip deneme desenleri için farklı prior tanımlamalarından nasıl etkilendiği de gösterilmiştir. Yeterince büyük ailelerde, prior tanımlamalardaki farklılık herbir parametre hakkında esas itibariyle aynı marjinal posterior yorumlamayı verir. Buradan, marjinal posterior yogunluk fonksiyonunun prior tanımlamalardaki değişikliklere oldukça güçlü olduğu sonucu çıkarılabilir.

Tablo 1. Farklı Kalıtım Dereceleri ve Aile Büyüklüklerine Sahip Dört Karaktere Ait ANOVA Tahminleri.

h^{2*}	σ^2_s		σ^2_e		h^2	
	Ort.[a]	SS[b]	Ort.	SS	Ort.	SS
			s=25 n=8			
.1	0.026	0.047	0.975	0.106	0.102	0.182
.1	0.027	0.045	0.976	0.107	0.107	0.178
.2	0.051	0.052	0.944	0.102	0.201	0.199
.2	0.049	0.049	0.946	0.107	0.197	0.195
.1	0.027	0.043	0.972	0.111	0.106	0.171
.3	0.072	0.059	0.926	0.096	0.282	0.219
.4	0.103	0.066	0.912	0.096	0.398	0.245
.6	0.153	0.075	0.851	0.089	0.598	0.267
			s=25 n=20			
.1	0.023	0.021	0.977	0.063	0.093	0.082
.1	0.026	0.022	0.976	0.060	0.103	0.086
.2	0.049	0.029	0.950	0.062	0.198	0.111
.2	0.050	0.028	0.954	0.059	0.197	0.105
.1	0.025	0.021	0.982	0.060	0.100	0.081
.3	0.074	0.036	0.927	0.060	0.294	0.133
.4	0.100	0.042	0.899	0.059	0.394	0.152
.6	0.153	0.058	0.850	0.053	0.601	0.196
			s=80 n=8			
.1	0.026	0.024	0.972	0.057	0.102	0.096
.1	0.026	0.024	0.970	0.058	0.104	0.095
.2	0.051	0.028	0.946	0.054	0.203	0.109
.2	0.049	0.028	0.949	0.053	0.194	0.110
.1	0.026	0.024	0.974	0.060	0.102	0.097
.3	0.075	0.031	0.924	0.056	0.300	0.117
.4	0.101	0.036	0.902	0.051	0.399	0.132
.6	0.148	0.042	0.848	0.051	0.592	0.153
			s=80 n=20			
.1	0.025	0.012	0.937	0.037	0.099	0.049
.1	0.025	0.012	0.974	0.034	0.099	0.047
.2	0.050	0.016	0.952	0.034	0.199	0.063
.2	0.051	0.016	0.949	0.033	0.203	0.064
.1	0.026	0.012	0.971	0.038	0.102	0.047
.3	0.074	0.019	0.925	0.033	0.297	0.074
.4	0.099	0.023	0.899	0.033	0.395	0.085
.6	0.151	0.029	0.850	0.031	0.599	0.099

[a]Ortalama [b]Standart sapma

* İlk sütun simülasyonda kullanılan değerleri temsil etmektedir.

Tablo 2. Farklı Kalıtım Dereceleri ve Aile Büyüklüklerine Sahip Dört Karaktere Ait İki Farklı Prior Tanımlamalı Gibbs Örneklemesinin Posterior Beklenen Değerleri.

h^{2*}	Prior1						Prior2					
	σ^2_s		σ^2_e		h^2		σ^2_s		σ^2_e		h^2	
	Ort.[a]	SS[b]	Ort.	SS	Ort.	SS	Ort.	SS	Ort.	SS	Ort.	SS
						s=25 n=8						
.1	0.035	0.014	0.980	0.102	0.140	0.052	0.044	0.013	0.982	0.096	0.174	0.053
.1	0.036	0.013	0.981	0.103	0.144	0.053	0.046	0.016	0.992	0.102	0.178	0.060
.2	0.063	0.022	0.950	0.100	0.249	0.084	0.052	0.020	0.963	0.095	0.207	0.078
.2	0.063	0.020	0.951	0.101	0.250	0.083	0.054	0.020	0.962	0.100	0.215	0.078
.1	0.035	0.009	0.979	0.105	0.122	0.040	0.073	0.018	0.969	0.101	0.281	0.068
.3	0.080	0.024	0.937	0.093	0.313	0.087	0.096	0.029	0.919	0.090	0.376	0.107
.4	0.106	0.028	0.926	0.092	0.411	0.106	0.108	0.037	0.912	0.094	0.421	0.132
.6	0.148	0.032	0.873	0.084	0.578	0.109	0.142	0.055	0.861	0.092	0.556	0.186
						s=25 n=20						
.1	0.029	0.010	0.979	0.063	0.117	0.038	0.036	0.010	0.978	0.064	0.143	0.039
.1	0.031	0.011	0.978	0.060	0.124	0.043	0.038	0.011	0.978	0.058	0.148	0.042
.2	0.056	0.018	0.953	0.061	0.229	0.069	0.051	0.018	0.954	0.061	0.201	0.071
.2	0.057	0.018	0.957	0.059	0.222	0.065	0.051	0.019	0.958	0.059	0.202	0.070
.1	0.029	0.010	0.985	0.059	0.116	0.037	0.056	0.012	0.971	0.065	0.218	0.045
.3	0.080	0.024	0.931	0.060	0.315	0.087	0.087	0.026	0.927	0.059	0.340	0.093
.4	0.105	0.029	0.904	0.059	0.413	0.104	0.107	0.034	0.906	0.056	0.416	0.119
.6	0.153	0.035	0.857	0.053	0.602	0.118	0.149	0.047	0.859	0.054	0.580	0.159
						s=80 n=8						
.1	0.030	0.010	0.974	0.055	0.118	0.040	0.036	0.011	0.970	0.055	0.142	0.041
.1	0.030	0.010	0.971	0.057	0.119	0.039	0.036	0.010	0.968	0.057	0.144	0.040
.2	0.055	0.016	0.949	0.053	0.218	0.064	0.049	0.017	0.952	0.053	0.197	0.065
.2	0.054	0.017	0.951	0.053	0.213	0.065	0.049	0.017	0.954	0.053	0.193	0.066
.1	0.029	0.010	0.976	0.057	0.114	0.041	0.052	0.011	0.962	0.054	0.206	0.043
.3	0.077	0.022	0.929	0.055	0.304	0.083	0.081	0.022	0.925	0.052	0.320	0.082
.4	0.102	0.027	0.907	0.051	0.401	0.099	0.098	0.028	0.905	0.051	0.389	0.101
.6	0.148	0.032	0.854	0.050	0.589	0.116	0.144	0.038	0.854	0.051	0.571	0.133
						s=80 n=20						
.1	0.026	0.008	0.974	0.037	0.103	0.031	0.029	0.008	0.973	0.036	0.117	0.035
.1	0.026	0.008	0.976	0.033	0.104	0.030	0.030	0.008	0.974	0.033	0.118	0.030
.2	0.050	0.013	0.912	0.035	0.201	0.053	0.048	0.014	0.955	0.035	0.190	0.053
.2	0.052	0.014	0.950	0.033	0.205	0.052	0.049	0.014	0.951	0.034	0.196	0.054
.1	0.026	0.008	0.973	0.038	0.104	0.030	0.039	0.007	0.970	0.036	0.155	0.029
.3	0.076	0.017	0.927	0.033	0.300	0.064	0.077	0.017	0.929	0.034	0.306	0.067
.4	0.101	0.021	0.901	0.033	0.399	0.077	0.101	0.023	0.900	0.032	0.402	0.083

[a]Ortalama [b]Standart sapma

* İlk sütun simülasyonda kullanılan değerleri temsil etmektedir.

Kaynaklar

1. Fırat, M.Z.: Hayvan ıslahında kullanılan çoklu karakter analizinin getirdiği problemler ve çözüm yolları. Turk. J. Vet. Anim. Sci., 1997; 21, 227-231.
2. Hill, W.G. and Thompson, R.: Probabilities of non-positive definite between-group or genetic covariance matrices. Biometrics, 1978; 34, 429-439.
3. Gianola, D. and Fernando, R.L.: Bayesian methods in animal breeding theory. J. Anim. Sci., 1986; 63, 217-244.
4. Gianola, D. and Foulley, J.L.: Variance estimation from integrated likelihoods (VEIL). Genet. Sel. Evol., 1990; 18, 485-498.
5. Wang, C.S., Rutledge, J.J. and Gianola, D.: Marginal inferences about variance components in a mixed linear model using Gibbs sampling. Genet. Sel. Evol., 1993; 25, 41-62.
6. Wang, C.S., Rutledge, J.J. and Gianola, D.: Bayesian analysis of mixed linear models via Gibbs sampling with an application to litter size in Iberian pigs. Genet. Sel. Evol., 1994; 26, 91-115.
7. Firat, M.Z., Theobald, C.M. and Thompson, R.: Univariate analysis of test day milk yields of British Holstein Friesian heifers using Gibbs sampling. Acta Agric. Scand., Sect. A, Anim. Sci. 1977; 47, 213-220.
8. Foulley, J.L., Im, S., Gianola, D. and Hoschele, I.: Empirical Bayes estimation of parameters for n polygenic binary traits. Genet. Sel. Evol., 1987; 19, 197-224.
9. Varona, L., Moreno, C., Garcia-Cortes, L. and Altarriba, J.: Estimacion multicaracter de componentes de varianza y covarianza en vacuno lechero mediante muestreo de Gibbs. Revista Portugueza de Zootecnia, 1994; 1, 185-195.
10. Jensen, J., Wang, C.S., Sorensen, D.A. and Gianola, D.: Bayesian inference on variance and covariance components for traits influenced by maternal and direct genetic effects using the Gibbs sampler. Acta Agric. Scand., 1994; 44, 193-201.
11. Jensen, J.: Bayesian analysis of bivariate mixed models with one continuous and one binary trait using the Gibbs sampler. Proceedings of the fifth world congress on genetics applied to livestock production, 1994; 18, 333-336.
12. Gelfand, A.E., Hills, S.E., Racine-Poon, A. and Smith, A.F.M.: Illustration of Bayesian inference in normal data models using Gibbs sampling. J. Amer. Statist. Assoc., 1990; 85, 972-985.
13. Geman, S. and Geman, D.: Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1984; 6, 721-741.
14. Gelfand, A.E. and Smith, A.F.M.: Sampling-based approaches to calculating marginal densities. J. Amer. Statist. Assoc., 1990; 85, 398-409.
15. Raftery, A.E. and Lewis, S.M.: How many iterations in the Gibbs sampler? In Bayesian Statistics 4, Bernardo, J.M., Berger, J.O., David, A.P. and Smith, A.F.M. (eds). Oxford: Clarendon Press, 1992; 763-773.