

مقایسه رهیافت بیزی مرسوم و رهیافت بیزی تقریبی در برآورد مؤلفه‌های واریانس مدل‌های حیوانی

فاطمه حسینی^{۱*}، امید کریمی^۱ و نیلوفر جواهری^۲

۱ و ۲. استادیاران و دانشجوی سابق کارشناسی ارشد، گروه آمار، دانشکده ریاضی، آمار و علوم کامپیوتر، دانشگاه سمنان، ایران

(تاریخ دریافت: ۱۳۹۴/۸/۲۶ - تاریخ پذیرش: ۱۳۹۵/۴/۶)

چکیده

مدل‌های حیوانی برای مدل‌بندی مشاهده‌های مربوط به عملکرد حیوان با همبستگی ژنتیکی استفاده می‌شوند. این مدل‌ها متعلق به کلاس مدل‌های مختلط خطی تعمیم‌یافته هستند و همبستگی ژنتیکی موجود بین داده‌ها توسط تأثیر تصادفی با ارزش اصلاحی به مدل اضافه می‌شود. از جمله هدف‌های این مدل‌ها، برآورد مؤلفه‌های واریانس است. در این پژوهش رهیافت بیزی تقریبی برای برآورد مؤلفه‌های واریانس مدل حیوانی ارائه و با رهیافت بیزی مرسوم مقایسه شد. برای این منظور مجموعه داده شبیه‌سازی شده با ۱۰۸۴ رکورد مربوط به حیوان فرضی استفاده شد. مشاهده‌های وزن حیوان هنگام تولد است. این مجموعه داده شامل کد حیوان، کد مادر، کد پدر، جنسیت و سال تولد است. جنسیت به‌عنوان اثر ثابت و اثرهای مادری، حیوان و سال تولد به‌صورت عامل‌های تصادفی در نظر گرفته شدند. چهار مدل با رهیافت بیزی مرسوم برازش و با یک معیار مدل‌گزینی مدل مناسب انتخاب شد. رهیافت بیزی تقریبی روی مدل مناسب پیاده‌سازی شد. زمان محاسبات با رایانه به ویژگی‌های (Intel Core i7, 4GB, 2.7 GHz) برای رهیافت بیزی مرسوم حدود ۱۲۰ ثانیه و با رهیافت بیزی تقریبی کمتر از ۱۰ ثانیه به طول انجامید. برای بررسی نکویی برازش معیار ریشه میانگین توان دوم خطای نسبی محاسبه شد که به ترتیب ۰/۱۵۶۸ و ۰/۱۴۹۹ برای رهیافت بیزی مرسوم و بیزی تقریبی به دست آمد. برای بررسی تفاوت نداشتن معنی‌دار بین برازش وزن حیوان توسط دو رهیافت، آزمون تی (T) استفاده شد که میزان احتمال آزمون ۰/۹۸ به دست آمد و فرضیه صفر برابر بودن میانگین برازش وزن حیوان از دو رهیافت برابری، پذیرفته شد.

واژه‌های کلیدی: رهیافت بیزی تقریبی، ژنتیک افزایشی، مدل حیوانی، وراثت.

Comparison of conventional bayesian and approximate bayesian approaches in estimation of variance components using animal models

Fateme Hosseini^{1*}, Omid Karimi¹ and Niloofar Javaheri²

1, 2. Assistant Professors and Former M.Sc. Student, Department of Statistics, Faculty of Mathematics, Statistics and Computer Science, Semnan University, Semnan, Iran

(Received: Nov. 17, 2015 - Accepted: Jun. 26, 2016)

ABSTRACT

Animal models are used to model the observations of animal performance that are genetically dependent. These models are considered as generalized linear mixed models and the genetic correlation structure of data is considered through random effects of breeding values. One goal of the mentioned models is to estimate variance components. In this research, an approximate Bayesian approach presented to estimate variance components in animal model and compared with the conventional Bayesian approach. A generated data set for hypothetical animal population with 1084 records was used. The observations are the animal's birth weight and the data includes dam ID, sire ID, sex and birth year. The effect of gender was considered as fixed effect and the effects of dam, animal and year of birth were used as random effect. Four different models were fitted by the conventional Bayesian approach and the appropriate model was selected by deviance information criteria. The approximate Bayesian approach was applied on it. Time consuming with a PC with configuration (Intel Core i7, 4GB, 2.7 GHz) was about 120 second for the conventional Bayesian approach and little than 10 second for the approximate Bayesian approach. Goodness of fit was computed by relative root mean squared error of prediction that was respectively 0.1568 and 0.1499 for conventional Bayesian and the approximate Bayesian approaches. T-test was used to illustrate lack of significant different to fit weight of animals between two approaches. The null hypothesis was accepted with p-value 0.98 that it shows mean of fitted animal weights for two approaches are equal.

Keywords: animal model, approximate bayesian approach, breeding value, heritability.

مقدمه

مدل‌های خطی تعمیم‌یافته^۱ (GLM) شامل دامنه گسترده‌ای از مدل‌های آماری هستند که می‌توان با به‌کارگیری این مدل‌ها افزون بر مشاهده‌های پیوسته، مشاهده‌های گسسته و رسته‌ای را مدل‌بندی کرد. در این مدل‌ها با فرض استقلال مشاهده‌های بین میانگین مشاهده‌ها و متغیرهای کمکی با استفاده از یک تابع پیوند^۲ معلوم، ارتباط برقرار می‌شود. پیرامون مدل‌های خطی تعمیم‌یافته McCullagh & Nelder (1989) به‌طور مفصل بررسی کردند. در حالتی که بین مشاهده‌های همبستگی وجود دارد مدل‌های مختلط خطی تعمیم‌یافته^۳ (GLMM) که تعمیمی از مدل خطی تعمیم‌یافته است، به کار گرفته می‌شود. همبستگی موردنظر با اضافه کردن متغیرهای پنهان^۴ یا همان اثرگذاری تصادفی به مدل در نظر گرفته می‌شود. در برخی پژوهش‌ها هدف بررسی متغیرهایی است که بر صفات تولیدی، تولیدمثلی، سلامتی و تیپ دام و طیور تأثیر می‌گذارند. صفاتی مانند وزن دام هنگام تولد، تولید تخم‌مرغ در هفته‌های مختلف تخم‌گذاری در ماکیان، میزان شیر و ... از جمله بررسی‌های موردی در علوم دامی هستند. برای تحلیل این نوع متغیرها از مدل‌های حیوانی^۵ استفاده می‌شود. مدل‌های حیوانی متعلق به کلاس مدل‌های مختلط خطی تعمیم‌یافته هستند که برای شناسایی بخش ژنتیکی صفات در اصلاح‌نژاد حیوانات و زیست‌شناسی تکاملی استفاده می‌شوند. در اغلب بررسی‌های اصلاح نژاد حیوانات، صفت مربوط به عملکرد حیوان می‌تواند پیوسته یا گسسته باشند، این متغیرها مستقل نیستند و وابستگی‌های ژنتیکی و محیطی دارند. این وابستگی‌ها، با اضافه کردن متغیرهای پنهان به صورت اثرگذاری تصادفی به مدل در نظر گرفته می‌شوند. بنابراین هدف از مدل‌های حیوانی بررسی تأثیر برخی اثرگذاری‌های ثابت و تصادفی روی صفات مربوط به اصلاح نژاد حیوانات است. روش‌های استنباط آماری

مختلف برای بررسی این مدل‌ها پیشنهاد شده است. به‌عنوان مثال تابعیت تصادفی در یک مدل خطی برای تجزیه رکوردهای روز آزمون به‌منظور ارزیابی ژنتیکی گاوهای شیری به کار گرفته شده است (Schaeffer, 1994). به‌طور معمول در بیشتر کشورها برای ارزیابی ژنتیکی گاوهای شیری از رکوردهای ماهیانه مربوط به گاوهای شیری یا همان رکوردهای روز آزمون استفاده می‌شود و از مدل آماری تابعیت تصادفی حیوانی استفاده می‌شود (Cobuci, 2011). برای برآورد فراسنجه‌های ژنتیکی گاوهای هلشتاین در ایران اغلب رکوردهای روز آزمون و تابعیت تصادفی حیوانی به کار گرفته شده است، (Mohammadi *et al.*, 2013; Mohammadi *et al.*, 2014; Razm kabir *et al.*, 2014; Moghbeli, 2014; Bohlouli, 2012). برای برآورد مؤلفه‌های واریانس مربوط به تأثیر تصادفی Simm (1998) در بررسی خود پیرامون افزایش شیر و گوشت گاو، از الگوریتم بیشینه درست‌نمایی محدود شده^۶ (REML) استفاده کرد. Seraj *et al.* (2011) با مدل حیوانی اثرگذاری عامل‌های مادری را بر وزن یک‌روزگی جوجه‌های مازندران بررسی و برآورد فراسنجه‌ها را با روش بیشینه درست‌نمایی محدود شده به دست آورد. در مدل‌های مختلط خطی تعمیم‌یافته به دلیل وجود همبستگی بین مشاهده‌ها و در برخی موارد کیفی بودن مشاهده‌ها و نرمال نبودن، تابع درست‌نمایی شکل مشخصی ندارد و برای به دست آوردن برآوردهای بیشینه درست‌نمایی از روش‌های مرسوم نمی‌توان استفاده کرد و باید از الگوریتم‌های تکرارشونده مانند الگوریتم‌های مونت کارلویی استفاده شود، که گاهی بسیار زمان‌بر است، (McCulloch, 1997). برای حل این مشکل در مدل‌های مختلط خطی تعمیم‌یافته، استفاده از رهیافت بیزی پیشنهاد شده است (Fong, 2010). مدل‌های حیوانی متعلق به کلاس مدل‌های مختلط خطی تعمیم‌یافته هستند و در این مدل‌ها ممکن است صفت موردبررسی کیفی باشد و به دلیل وجود متغیرهای پنهان ژنتیکی و محیطی حیوانی تابع درست‌نمایی شکل مشخصی ندارد، می‌توان

1. Generalized Linear Models
2. Link Function
3. Generalized Linear Mixed Models
4. Latent Variables
5. Animal models

6. Restricted maximum likelihood

نرم‌افزاری INLA قابل اجرا در محیط نرم‌افزار R ارائه شده است. هدف اصلی این پژوهش، نحوهٔ استفاده از دو رهیافت بیزی تقریبی و بیزی معمولی در مدل‌های حیوانی و مقایسهٔ دو رهیافت بیزی تقریبی و بیزی مرسوم از نظر زمان محاسبات و دقت است. برای این منظور در آغاز نظریهٔ رهیافت بیزی تقریبی و بیزی مرسوم در مدل‌های حیوانی ارائه و آنگاه این دو روش روی یک مجموعه دادهٔ شبیه‌سازی شدهٔ مربوط به صفت وزن یک حیوان فرضی موجود در نرم‌افزار R پیاده‌سازی می‌شود. چهار مدل متفاوت برای این مجموعهٔ داده در نظر گرفته می‌شود و با معیار انتخاب مدل میزان اطلاع انحرافی (DIC)^۵ مدل مناسب انتخاب می‌شود. برآورد مؤلفه‌های واریانس و وراثت‌پذیری از هر دو روش محاسبه و زمان محاسبات و نتایج دو رهیافت بیزی مقایسه می‌شوند.

مدل حیوانی

مدل حیوانی، نوعی مدل مختلط خطی تعمیم‌یافته است، به طوری که اگر همبستگی موجود در مدل‌های مختلط خطی تعمیم‌یافته از نوع محیطی و ژنتیکی باشد، این نوع از مدل‌ها، مدل‌های حیوانی نامیده می‌شوند. در این مدل‌ها مشاهده‌های صفت موردبررسی با n_d رکورد به صورت بردار $\mathbf{y} = (y_1, \dots, y_{n_d})^T$ نشان داده می‌شود، به طوری که توزیع \mathbf{y}_i متعلق به خانوادهٔ نمایی است. میانگین به صورت $\mu_i = E(y_i)$ است و توسط یک تابع پیوند معلوم $g(\cdot)$ با اثرگذاری ثابت مربوط به متغیرهای کمکی و غیر کمکی، تأثیر تصادفی مربوط به اثرگذاری محیطی فردی و مادری و مقادیر ارزش اصلاحی که بیانگر تأثیر ژنتیکی افزایشی هستند، به صورت رابطه (۱) مرتبط می‌شود:

$$\eta_i = \mathbf{z}_i^T \boldsymbol{\beta} + u_i + \varepsilon_i \quad (1)$$

$u_i \sim N(0, \sigma_u^2)$, $\varepsilon_i \sim N(0, \sigma_\varepsilon^2)$ که در آن $g(\mu_i) = \eta_i$ پیش‌بینی‌کنندهٔ خطی، $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_{n_j})^T$ تأثیر ثابت، u_i ارزش اصلاحی فرد i ام، ε_i باقی‌ماندهٔ تصادفی و \mathbf{z}_i بردار با

از رهیافت بیزی استفاده کرد (Tempelman & Gianola, 1994; Jamrozika, 2001; Bolker *et al.*, 2009).

در رهیافت بیزی با توجه به اطلاعات گذشته پیرامون فراسنجه‌های مدل، برای فراسنجه‌های مدل توزیع‌های پیشین در نظر گرفته می‌شود و سپس توزیع پسین متناسب با حاصل ضرب تابع درستنمایی در توزیع توأم پیشین در نظر گرفته می‌شود و در نهایت برآورد بیزی فراسنجه‌ها از میانگین توزیع پسین محاسبه می‌شود. در صورتی که توزیع پسین شکل بسته‌ای نداشته باشد، برای استفاده از رهیافت بیزی باید از الگوریتم‌های مونت‌کارلوی زنجیر مارکوفی^۱ (MCMC) مانند الگوریتم متروپولیس هستینگس و نمونه‌گیری گیبز استفاده کرد، این الگوریتم‌ها به دلیل نیاز به تولید نمونه‌های مونت‌کارلویی با شمار تکرار مناسب برای رسیدن به ماندگاری، به طور معمول در مدل‌های پیچیده نیاز به صرف زمان طولانی برای انجام محاسبات دارند. بسته‌های نرم‌افزاری زیادی برای تحلیل این مدل‌ها با رهیافت بیزی معرفی شده است که از مهم‌ترین آن‌ها می‌توان به MCMCglmm (Hadfield, 2010) اشاره کرد. برای حل مشکل زمان محاسبات رهیافت بیزی مرسوم با الگوریتم‌های MCMC، رهیافت بیزی تقریبی^۲ در مدل‌های حیوانی معرفی شده است، (Holand *et al.*, 2013). در رهیافت بیزی تقریبی برای حل انتگرال‌ها از تقریب لاپلاس استفاده می‌شود و به همین دلیل این روش بیزی رهیافت بیزی تقریبی با تقریب لاپلاس آشیانی جمع‌بسته^۳ (INLA) نامیده می‌شود، (Rue & Martino, 2007). رهیافت بیزی تقریبی INLA برای تحلیل مدل‌های سلسله‌مراتبی فضایی^۴ ارائه شده است، (Rue *et al.*, 2009). در مدل‌های آمیختهٔ خطی تعمیم‌یافته فضایی این رهیافت بیزی تقریبی بررسی شده و با رهیافت بیزی مرسوم مقایسه شده است، (Eidsvik *et al.*, 2009; Hosseini, 2016; Hosseini *et al.*, 2011).

برای استفاده از رهیافت بیزی تقریبی بسته

1. Markov Chain Monte Carlo
2. Approximate Bayesian approach
3. Integrated Nested Laplace Approximation
4. Hierarchical Spatial Models

5. Deviance information criterion

هستند، لذا برای نشان دادن تغییرپذیری‌ها، از تفاوت در نسبت‌های واریانس‌های یادشده استفاده می‌شود، که تغییر و تفاوت‌های مؤلفه‌های واریانس را در تقسیم‌بندی‌های مختلف نشان می‌دهد و لذا تفاوت بین صفات مختلف (یک صفت در چند محیط) مشخص می‌شود. به‌عنوان مثال در مدل حیوانی گاوسی وراثت‌پذیری مستقیم (\mathbf{h}_u^2) از رابطه $\mathbf{h}_u^2 = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2}$ به دست می‌آید، که متناسب با واریانس فنوتیپی ($\sigma_u^2 + \sigma_e^2$) است. در بسیاری از بررسی‌های پیرامون مدل‌های حیوانی اگر بتوان تأثیر اثرگذاری مؤثر دیگر مانند تأثیر محیطی دائمی مادری و اثرگذاری ژنتیکی مادری را روی صفت موردبررسی و مؤلفه‌های واریانس شناسایی و به مدل اضافه کرد، نتایج دقت بیشتری دارند (Jasouri et al., 2014). اگر مدل حیوانی به‌صورت رابطه (۲) بیان شود:

$$\eta_i = \mathbf{z}_i^T \boldsymbol{\beta} + u_i + pe_i + \varepsilon_i, i = 1, \dots, n_d \quad (2)$$

بیان شود، که در آن pe_i تأثیر تصادفی محیطی دائمی مادری است، به‌طوری‌که:

$$n_i \sim N(0, \sigma_u^2), \varepsilon_i \sim N(0, \sigma_e^2), pe_i \sim N(0, \sigma_{pe}^2)$$

هستند، آنگاه فراسنجه‌های مدل (۲) عبارت‌اند از $\boldsymbol{\theta} = (\boldsymbol{\beta}^T, \sigma_u^2, \sigma_{pe}^2, \sigma_e^2)^T$ و ماتریس واریانس-کوواریانس این مدل به شکل

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_u^2 \mathbf{A} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \sigma_{pe}^2 \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \sigma_e^2 \mathbf{I} \end{pmatrix}$$

مؤلفه واریانس ژنتیک افزایشی مستقیم، \mathbf{A} ماتریس خویشاوندی، \mathbf{I} ماتریس واحد، σ_{pe}^2 مؤلفه واریانس تأثیر محیطی دائمی مادری و σ_e^2 مؤلفه واریانس خطای تصادفی هستند. در صورتی‌که مشاهده‌های صفت موردبررسی در مدل حیوانی (۲) کمی و به‌صورت نرمال باشد، وراثت‌پذیری مستقیم از رابطه

$$\mathbf{h}_u^2 = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2 + \sigma_{pe}^2}$$

بعد به تحلیل بیزی مدل (۲) پرداخته می‌شود.

ابعاد $(n_j + 1) \times$ متغیرهای کمکی مربوط به اثرگذاری ثابت هستند که عنصر اول آن برای در نظر گرفتن عرض از مبدأ در مدل منظور می‌شود. بنابراین تأثیر تصادفی ژنتیکی که با اثرگذاری ژنتیکی افزایشی شناخته می‌شوند با ارزش اصلاحی گاوسی (u_i) و تأثیر تصادفی محیطی و دیگر اثرگذاری‌های تصادفی با جمله باقی‌مانده تصادفی گاوسی (ε_i) وارد مدل شده است. فراسنجه‌های مدل (۱) عبارتند از $\boldsymbol{\theta} = (\boldsymbol{\beta}^T, \sigma_u^2, \sigma_e^2)^T$ که در آن σ_e^2 مؤلفه واریانس خطای تصادفی و σ_u^2 مؤلفه واریانس ژنتیک افزایشی مستقیم هستند.

تابع توزیع توأم شرطی $(\mathbf{y}, \mathbf{u} | \boldsymbol{\theta})$ به‌صورت $\pi(\mathbf{y}, \mathbf{u} | \boldsymbol{\theta}) = \pi(\mathbf{y} | \mathbf{u}, \boldsymbol{\beta}, \sigma_e^2) \pi(\mathbf{u} | \sigma_u^2)$ است. $\pi(\mathbf{y} | \mathbf{u}, \boldsymbol{\beta}, \sigma_e^2)$ تابع توزیع شرطی بردار \mathbf{y} است که متعلق به خانواده‌ی نمایی در نظر گرفته می‌شود و $\pi(\mathbf{u} | \sigma_u^2)$ تابع توزیع پیشین بردار ارزش اصلاحی است که نرمال $N(0, \sigma_u^2 \mathbf{A})$ فرض می‌شود که در آن \mathbf{A} ماتریس خویشاوندی^۱ است که به‌طور معمول بر پایه‌ی شجره‌نامه یعنی کد پدر، کد مادر و کد حیوان ساخته می‌شود، بردار ارزش اصلاحی به‌صورت $\mathbf{u} = \{u_1, \dots, u_{n_p}\}$ فرض می‌شود، به‌طوری‌که در آن n_p شمار رکورد شجره و $n_d \leq n_p$ است. بنابراین تابع درستنمایی به شکل زیر به دست می‌آید:

$$L(\boldsymbol{\theta}; \mathbf{y}) = \int \dots \int \prod_{i=1}^{n_d} \pi(y_i | u_i, \sigma_e^2, \boldsymbol{\beta}) \pi(\mathbf{u} | \sigma_u^2) d\mathbf{u}$$

که شکل مشخصی ندارد. برای استفاده از روش بیشینه درستنمایی نیاز به حل عددی انتگرال‌های چندبعدی است. یک حالت ساده از مدل‌های حیوانی، مدل‌های حیوانی با مشاهده‌های کمی برای صفت موردبررسی و به‌صورت نرمال است. در این صورت تابع پیوند یک تابع همانی $\eta_i = \mu_i$ فرض می‌شود. به‌طور معمول در برخی بررسی‌های مربوط به مدل‌های حیوانی یک صفت چند بار برای افراد اندازه‌گیری می‌شود. به‌عنوان مثال در ارزیابی ژنتیکی گاوهای شیری از مدل‌های روز آزمون استفاده می‌شود. چون واریانس‌های ژنتیک افزایشی و محیطی متفاوت

خاصی از توزیع‌های شناخته‌شده ندارند و برای تولید نمونه از الگوریتم متروپولیس هستیگس استفاده می‌شود.

البته اگر توزیع مشاهده‌های صفت موردبررسی نرمال باشد توزیع شرطی کامل بردار ضریب‌های رگرسیونی با در نظر گرفتن توزیع پیشین نرمال، شکل مشخص نرمال خواهد داشت. پس از به دست آوردن برآورد فراسنجه‌ها، برای پیش‌بینی بیزی رکورد حیوان جدید در آغاز باید متغیر پنهان ارزش اصلاحی برای آن فرد را از توزیع پیش‌بینی به صورت

$$\pi(u_0 | \mathbf{y}) = \int \dots \int \pi(u_0 | \mathbf{u}, \sigma_u^2) \pi(\mathbf{u}, \boldsymbol{\beta}, \sigma_u^2, \sigma_e^2 | \mathbf{y}) d\mathbf{u} d\boldsymbol{\beta} d\sigma_u^2 d\sigma_e^2$$

محاسبه کرد و آنگاه از توزیع پیش‌بینی برای $\pi(y_0 | \mathbf{y}) = \int_{u_0} \pi(y_0 | u_0) \pi(u_0 | \mathbf{y}) du_0$ پیش‌بینی میزان رکورد برای حیوان جدید استفاده کرد. در بخش بعد رهیافت بیزی تقریبی برای تحلیل مدل‌های حیوانی ارائه می‌شود که نسبت به رهیافت بیزی مرسوم زمان محاسبات کوتاه‌تری دارد.

تحلیل بیزی تقریبی مدل‌های حیوانی

در بردار تصادفی \mathbf{X} فرض کنید توزیع شرطی کامل هر مؤلفه از \mathbf{X} به صورت توزیع شرطی آن مؤلفه به شرط دیگر مؤلفه‌ها شود، یعنی $\pi(x_i | \mathbf{X}_{-i}), i = 1, \dots, n$ که در آن \mathbf{X}_{-i} بردار \mathbf{X} به جز مؤلفه x_i است. اکنون اگر توزیع شرطی کامل x_i تنها به شمار کمی از مؤلفه‌های \mathbf{X} بستگی داشته باشد، \mathbf{X} خاصیت مارکوفی دارد. اگر این مجموعه از مؤلفه‌ها با δ_i مشخص شود که تشکیل‌دهندهٔ مجموعه‌ای از همسایگی‌های مؤلفه i ام است، آنگاه $\pi(x_i | \mathbf{X}_{-i}) = \pi(x_i | \mathbf{X}_{\delta_i})$ می‌شود جمله‌های x_i و $\mathbf{X}_{-\{i, \delta_i\}}$ که در آن $\mathbf{X}_{-\{i, \delta_i\}}$ بردار \mathbf{X} به جز مؤلفه x_i و همسایگی‌هایش است، مستقل هستند. هنگامی از میدان تصادفی مارکوفی استفاده می‌شود به جای استفاده از ماتریس کوواریانس از عکس ماتریس کوواریانس استفاده می‌شود و ماتریس

مدل حیوانی بیزی

به منظور ارائهٔ یک تحلیل بیزی برای مدل حیوانی (۲)، نیاز است تا توزیع‌های پیشین برای فراسنجه‌های مدل ضریب‌های تأثیر ثابت ($\boldsymbol{\beta}$) توزیع پیشین به طور معمول به صورت توزیع یکنواخت و یا توزیع نرمال با واریانس بزرگ معلوم فرض می‌شود. متغیر تصادفی پنهان ارزش اصلاحی گاوسی به شکل $\mathbf{u} | \mathbf{A}, \sigma_u^2 \sim N(0, \sigma_u^2 \mathbf{A})$ فرض می‌شود، که در آن \mathbf{A} ماتریس خویشاوندی و مؤلفهٔ واریانس ژنتیک افزایشی است، (Sorensen, 2002). توزیع پیشین برای مؤلفه‌ها واریانس $(\sigma_u^2, \sigma_e^2, \sigma_{pe}^2)$ به طور معمول توزیع گامای معکوس^۱ فرض می‌شود. با فرض استقلال پیشین‌ها، توزیع پیشین توأم به صورت:

$$\pi(\boldsymbol{\theta}) = \pi(\boldsymbol{\beta}, \sigma_e^2, \sigma_u^2, \sigma_{pe}^2) = \pi(\boldsymbol{\beta}) \cdot \pi(\sigma_e^2) \cdot \pi(\sigma_u^2) \cdot \pi(\sigma_{pe}^2),$$

به دست می‌آید. لذا توزیع پسین متناسب با:

$$\pi(\mathbf{u}, \boldsymbol{\beta}, \sigma_u^2, \sigma_e^2, \sigma_{pe}^2 | \mathbf{y}) = \pi(\mathbf{y} | \mathbf{u}, \boldsymbol{\beta}, \sigma_e^2) \cdot \pi(\mathbf{u} | \sigma_u^2) \cdot \pi(\boldsymbol{\beta}) \cdot \pi(\sigma_e^2) \cdot \pi(\sigma_u^2) \cdot \pi(\sigma_{pe}^2),$$

خواهد شد، که شکل پیچیده‌ای دارد. بنابراین از روش‌های MCMC مانند گیبز استفاده می‌شود. برای به کارگیری الگوریتم گیبز توزیع شرطی کامل فراسنجه‌ها موردنیاز است. توزیع شرطی کامل برای فراسنجه‌ها به صورت:

$$\begin{aligned} \pi(\boldsymbol{\beta} | \mathbf{y}, \mathbf{u}, \sigma_e^2, \sigma_u^2, \sigma_{pe}^2) &\propto \pi(\mathbf{y} | \mathbf{u}, \boldsymbol{\beta}, \sigma_e^2, \sigma_{pe}^2) \cdot \pi(\boldsymbol{\beta}) \\ \pi(\sigma_u^2 | \mathbf{y}, \mathbf{u}, \boldsymbol{\beta}, \sigma_e^2, \sigma_{pe}^2) &\propto \pi(\mathbf{u} | \sigma_u^2) \cdot \pi(\sigma_u^2) \\ \pi(\sigma_e^2 | \mathbf{y}, \mathbf{u}, \boldsymbol{\beta}, \sigma_u^2, \sigma_{pe}^2) &\propto \pi(\mathbf{y} | \mathbf{u}, \boldsymbol{\beta}, \sigma_e^2, \sigma_{pe}^2) \cdot \pi(\sigma_e^2) \\ \pi(\sigma_{pe}^2 | \mathbf{y}, \mathbf{u}, \boldsymbol{\beta}, \sigma_u^2, \sigma_e^2) &\propto \pi(\mathbf{y} | \mathbf{u}, \boldsymbol{\beta}, \sigma_e^2, \sigma_{pe}^2) \cdot \pi(\sigma_{pe}^2) \end{aligned}$$

تشکیل می‌شوند. توزیع شرطی کامل برای هر مؤلفهٔ بردار \mathbf{u} به دلیل وابستگی مؤلفه‌ها به یکدیگر به صورت:

$$\begin{aligned} \pi(u_i | \mathbf{u}_{-i}, \mathbf{y}, \boldsymbol{\beta}, \sigma_u^2, \sigma_e^2, \sigma_{pe}^2) &\propto \\ \pi(\mathbf{y} | \mathbf{u}, \boldsymbol{\beta}, \sigma_e^2, \sigma_{pe}^2) \pi(u_i | \mathbf{u}_{-i}, \sigma_u^2), i = 1, \dots, n_d, \end{aligned}$$

که در آن $\mathbf{u}_{-i} = \{u_1, \dots, u_{i-1}, u_{i+1}, \dots, u_{n_d}\}$ است. توزیع‌های شرطی کامل تشکیل شده شکل

$$\begin{aligned}\pi(x_i | \mathbf{y}) &= \iiint \pi(x_i | \sigma_u^2, \sigma_e^2, \sigma_{pe}^2, \mathbf{y}) \pi(\sigma_u^2, \sigma_e^2, \sigma_{pe}^2 | \mathbf{y}) d\sigma_u^2 d\sigma_e^2 d\sigma_{pe}^2 \\ \pi(\sigma_u^2 | \mathbf{y}) &= \iint \pi(\sigma_u^2, \sigma_e^2, \sigma_{pe}^2 | \mathbf{y}) d\sigma_e^2 d\sigma_{pe}^2 \\ \pi(\sigma_e^2 | \mathbf{y}) &= \iint \pi(\sigma_u^2, \sigma_e^2, \sigma_{pe}^2 | \mathbf{y}) d\sigma_u^2 d\sigma_{pe}^2 \\ \pi(\sigma_{pe}^2 | \mathbf{y}) &= \iint \pi(\sigma_u^2, \sigma_e^2, \sigma_{pe}^2 | \mathbf{y}) d\sigma_u^2 d\sigma_e^2\end{aligned}$$

که انتگرال‌های بالا نیاز به محاسبه $\pi(x_i | \sigma_u^2, \sigma_e^2, \sigma_{pe}^2, \mathbf{y})$ و $\pi(\sigma_u^2, \sigma_e^2, \sigma_{pe}^2 | \mathbf{y})$ است. برای تحلیل بیزی این نوع مدل با ساختار پیچیده رهیافت بیزی تقریبی را می‌توان استفاده کرد.

با به کار بردن تقریب لاپلاس برای توزیع‌های $\pi(\sigma_u^2, \sigma_e^2, \sigma_{pe}^2 | \mathbf{y})$ و $\pi(x_i | \sigma_u^2, \sigma_e^2, \sigma_{pe}^2, \mathbf{y})$ توزیع‌های تقریبی به صورت $\tilde{\pi}(x_i | \sigma_u^2, \sigma_e^2, \sigma_{pe}^2, \mathbf{y})$ و $\tilde{\pi}(\sigma_u^2, \sigma_e^2, \sigma_{pe}^2 | \mathbf{y})$ به دست می‌آید، به طوری که

$\tilde{\pi}(x_i | \sigma_u^2, \sigma_e^2, \sigma_{pe}^2, \mathbf{y})$ توزیع تقریبی نرمال دارد (Rue et al., 2009). برای تقریب

$\tilde{\pi}(\sigma_u^2, \sigma_e^2, \sigma_{pe}^2 | \mathbf{y})$ در آغاز با در نظر گرفتن مد تابع $\theta = (\sigma_u^2, \sigma_e^2, \sigma_{pe}^2)$ با روش

بهینه‌سازی به صورت θ^* تعیین می‌شود. ماتریس هسین $\mathbf{H} = \frac{d^2 \log \tilde{\pi}(\theta | \mathbf{y})}{d\theta^i d\theta^j}$ ^۳ محاسبه و عکس آن

به صورت $\mathbf{H}^{-1} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}'$ تجزیه می‌شود، که در آن \mathbf{V} ماتریس بردارهای ویژه ^۴ است و $\mathbf{\Lambda}$ ماتریس قطری

مقادیر ویژه ^۵ \mathbf{H} است، به طوری که روی قطر اصلی آن مقادیر ویژه و دیگر عناصر ماتریس صفر هستند. مبدأ

مختصات را به مد θ^* انتقال داده، فرمول مختصات در مبدأ θ^* با استفاده از مقادیر استاندارد t (برای مثال

در حالت دو بعدی $t = (0, 0)^T$)، به صورت $\theta(t) = \theta^* + \mathbf{V}\mathbf{\Lambda}^{-1}t$ تعریف می‌شود. با آغاز از مبدأ

مختصات جدید روی هر یک از محورها نقطه‌ای به فاصله مقادیر صحیح δ_i به گونه‌ای اختیار می‌شوند که شرط زیر برقرار باشد:

$$\log \tilde{\pi}(\theta(0) | \mathbf{y}) - \log \tilde{\pi}(\theta(t) | \mathbf{y}) < \delta_i$$

سپس به طور همسان نقاط درون صفحات نیز تعیین می‌شوند. با به کار بردن این الگوریتم نمونه‌هایی از توزیع $\pi(\sigma_u^2, \sigma_e^2, \sigma_{pe}^2 | \mathbf{y})$ تولید می‌شوند.

اکنون با استفاده از نمونه‌های تولیدشده

دقت ^۱ نامیده می‌شود و با علامت Q بیان می‌شود. بیشتر درایه‌های ماتریس دقت صفر است و درایه‌های غیر صفر این ماتریس توسط ساختار همسایگی مشخص می‌شوند و در راستای راحتی و کوتاه کردن محاسبات سودمند است، (Rue & Held, 2005). در واقع میدان تصادفی مارکوفی با یک ماتریس دقت که ساختار پراکنده دارد، تعریف می‌شود و استفاده از روش‌های عددی محاسباتی مؤثر را امکان‌پذیر می‌سازد.

در مدل‌های حیوانی شجره ساختار مارکوفی دارد به این صورت که یک کد حیوان به شمار محدود در ستون کد پدر و کد مادر ممکن است تکرار شود به همین دلیل برخی از عناصر ماتریس خویشاوندی صفر می‌شود به ویژه هنگامی از عکس ماتریس استفاده می‌شود شمار بیشتری از درایه‌ها صفر خواهد شد. بنابراین بردار متغیر ارزش اصلاحی \mathbf{u} توسط میدان تصادفی مارکوفی گاوسی قابل‌بیان و وارون ماتریس خویشاوندی \mathbf{A} یعنی \mathbf{A}^{-1} یک ماتریس تنک ^۲ است، یعنی بیشتر عناصر آن صفر است (Steinsland & Jensen, 2010). اکنون ارزش اصلاحی و ضریب‌های رگرسیونی در یک بردار، به صورت $\mathbf{x} = (\boldsymbol{\beta}^T, \mathbf{u}^T)^T$ در نظر گرفته می‌شود. با توجه به نرمال بودن توزیع ضریب‌های رگرسیونی و ارزش اصلاحی بردار متغیرهای پنهان \mathbf{x} با یک میدان تصادفی مارکوفی گاوسی $\mathbf{x} | \sigma_u^2 : N(\mathbf{0}, \mathbf{Q}^{-1})$ مدل‌بندی می‌شود، که در آن ماتریس دقت است.

با توجه به استقلال شرطی \mathbf{y} می‌توان نوشت:

$$\pi(\mathbf{y} | \mathbf{x}, \sigma_e^2, \sigma_{pe}^2) = \prod_{i=1}^{n_d} \pi(y_i | \mathbf{x}, \sigma_e^2, \sigma_{pe}^2)$$

و بنابراین توزیع پسین می‌شود:

$$\begin{aligned}\pi(\mathbf{x}, \sigma_u^2, \sigma_e^2, \sigma_{pe}^2 | \mathbf{y}) &\propto \pi(\sigma_u^2, \sigma_e^2, \sigma_{pe}^2) \\ &\pi(\mathbf{x} | \sigma_u^2) \prod_{i=1}^{n_d} \pi(y_i | \mathbf{x}, \sigma_e^2, \sigma_{pe}^2)\end{aligned}$$

این توزیع پسین شکل بسته‌ای ندارد و از روش‌های تکراری برای به دست آوردن برآورد فراسنجه‌ها استفاده می‌شود که بسیار زمان‌بر هستند. توزیع‌های پسین حاشیه‌ای به صورت زیر هستند:

3. Hessian Matrix
4. Eigenvectors
5. Eigenvalues

1. Positive Definite Precision Matrix
2. Sparse matrix

کمترین DIC را دارد بهترین مدل است و ادامهٔ بررسی‌ها روی مدل انتخاب‌شده انجام می‌شود. برآورد بیزی فراسنجه‌ها (میانگین تابع چگالی حاشیه‌ای پسین فراسنجه)، انحراف معیار برآوردها و فاصلهٔ قابل باور HPD^۲ برای فراسنجه‌های مدل با هردو رهیافت بیزی مرسوم و بیزی تقریبی ارائه می‌شود.

نتایج و بحث

آماره‌های توصیفی مربوط به صفت وزن حیوان و ساختار کلی شجره در جدول ۱ و نمودار جعبه‌ای و هیستوگرام مربوط به مشاهده‌های وزن حیوان در شکل ۱ ارائه شده‌اند. با توجه به کمی بودن صفت وزن و شکل ۱ توزیع نرمال برای این صفت در نظر گرفته می‌شود.

از بستهٔ نرم‌افزاری MCMCglmm استفاده شد و با اجرای الگوریتم MCMC با ۵۰۰۰۰ تکرار از رهیافت بیزی مرسوم هر چهار مدل برازش و معیار انتخاب مدل DIC برای چهار مدل به ترتیب ۳۹۲۹/۸۴۱، ۳۷۰۲/۲۹۵، ۳۶۵۹/۶۹ و ۳۳۲۸/۸۲۹ به دست آمد. با توجه به مقادیر DIC مدل چهارم به‌عنوان مدل مناسب انتخاب شد. نتایج برآورد بیزی فراسنجه‌ها، انحراف معیار برآوردها و مجموعهٔ قابل باور HPD برای مدل چهارم در جدول ۲ آورده شده است.

برای مدل چهارم تابع چگالی حاشیه‌ای پسین و نمودار اثر نمونه‌های مونت کارلویی مؤلفه‌های واریانس مربوط به تأثیر تصادفی ارزش اصلاحی، باقی‌مانده، مادر و سال تولد به ترتیب در شکل‌های ۲ الی ۵ رسم شده است. میانگین تابع چگالی حاشیه‌ای پسین برآورد بیزی فراسنجه است و نمودارهای اثر نمونه‌های تولیدشده از الگوریتم MCMC را نشان می‌دهد.

$\theta^{(1)}, \dots, \theta^{(R)}$ توزیع پسین $\pi(x_i | \mathbf{y})$ به صورت زیر تقریب زده می‌شود (Rue et al., 2009):

$$\pi(x_i | \mathbf{y}) \approx \sum_{r=1}^R \tilde{\pi}(x_i | \theta^{(r)}, \mathbf{y}) \tilde{\pi}(\theta^{(r)} | \mathbf{y}) \Delta_r$$

مواد و روش‌ها

در این بخش به مقایسهٔ دو رهیافت بیزی تقریبی و مرسوم پرداخته می‌شود. برای این منظور از یک مجموعه داده شبیه‌سازی‌شدهٔ موجود در بستهٔ نرم‌افزاری MCMCglmm استفاده می‌شود. این مجموعه داده مربوط به وزن تولد برای نوعی حیوان فرضی است که به شمار ۱۰۸۴ رکورد شبیه‌سازی شده است و افزون بر وزن تولد، جنسیت حیوان و سال تولد مشخص است (Wilson et al., 2010). در فایل شجره شمار حیوان رکورد ۱۳۰۹، شمار مادر ۱۰۸۴ و شمار پدر ۴۸۷ دارند. چهار مدل حیوانی به شرح زیر برای تحلیل این مجموعهٔ داده در نظر گرفته می‌شود.

$$M_1: y_i = \beta_0 + u_i + \varepsilon_i, \quad i = 1, \dots, 1084,$$

$$M_2: y_i = \beta_0 + (\beta_{sex})_i + u_i + \varepsilon_i,$$

$$M_3: y_i = \beta_0 + (\beta_{sex})_i + (year)_i + u_i + \varepsilon_i,$$

$$M_4: y_i = \beta_0 + (\beta_{sex})_i + (year)_i + (mother)_i + u_i + \varepsilon_i,$$

در مدل حیوانی اول β_0 عرض از مبدأ یا میانگین جمعیت بدون در نظر گرفتن هیچ اثر ثابت و تصادفی، u_i اثر ارزش اصلاحی فرد نام، ε_i اثر باقی‌مانده و y_i وزن حیوان نام است. در مدل دوم، به مدل اول اثر ثابت "جنسیت" اضافه می‌شود. در مدل سوم، اثر "سال تولد" به‌صورت اثر تصادفی به مدل دوم اضافه می‌شود و در مدل چهارم اثر تصادفی ژنتیکی "مادر" به مدل سوم افزوده شده است. پس از برازش چهار مدل با رهیافت بیزی مرسوم، مدل مناسب با معیار انتخاب مدل میزان اطلاع انحرافی (DIC)^۱ برگزیده می‌شود. مدلی که

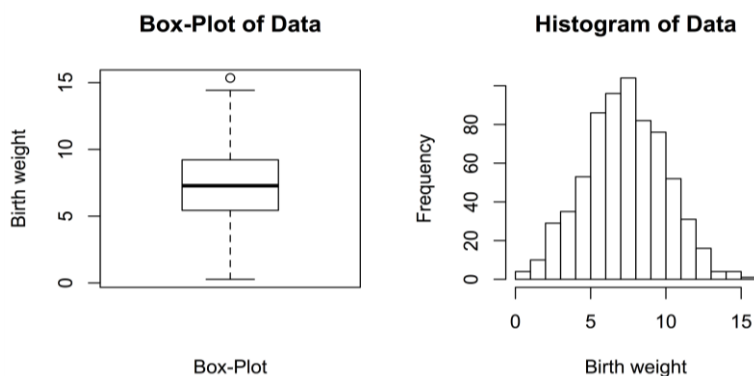
جدول ۱. آماره‌های توصیفی برای مشاهده‌های وزن حیوان هنگام تولد و ساختار کلی شجره

Table 1. Descriptive statistics for the animal birth weight observations and general structure of the pedigree

Data		Pedigree	
No. Records	1084	No. animals in pedigree	1309
Mean	7.368	No. sires	487
SD	2.689	No. dams	1084
Median	7.355	No. of animals with both parents unknown	225
Minimum	0.270	No. of animals with both parents known	487
Maximum	15.350	No. of animals with at least one parent known	1084
First quartile	5.455	No. sires no-repeated record	159
Third quartile	9.260	No. dams no-repeated record	430

1. Deviance information criterion

2. Highest posterior density credible interval



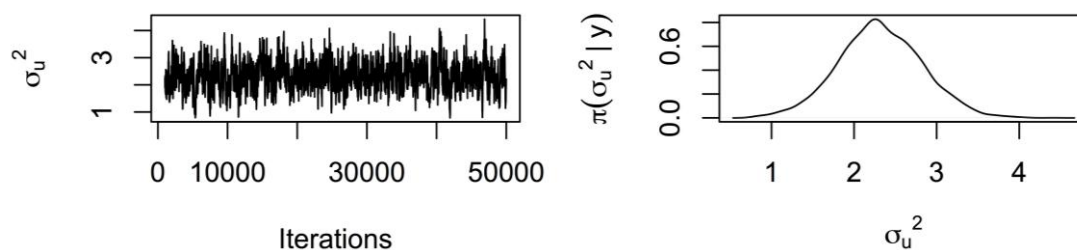
شکل ۱. (چپ) نمودار جعبه‌ای و (راست) نمودار ستونی برای وزن حیوان هنگام تولد
Figure 1. (Left) Box Plot and (right) histogram for the animal's birth weight

جدول ۲. نتایج رهیافت بیزی مرسوم و بیزی تقریبی برآورد بیزی فراسنجه (چارک اول و چارک سوم)، انحراف معیار و فاصله ۹۵ درصد قابل باور بیزی (HPD)

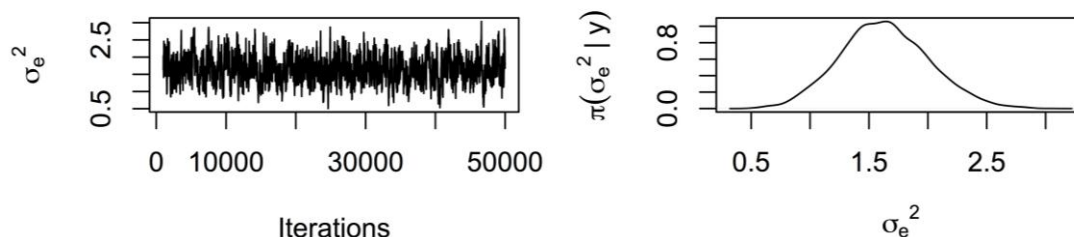
Table 2. Results of approximate Bayesian and conventional Bayesian approaches: Bayesian Estimate of parameter (first and third quartiles), standard deviation and 95% Bayesian credibility interval (HPD)

Par.	Conventional Bayesian			Approximate Bayesian		
	Bayesian estimate (first and third quartiles)	SD	95% HPD	Bayesian estimate (first and third quartiles)	SD	95% HPD
σ_u^2	2.33 (1.93,3.12)	0.50	(1.32,3.26)	2.32 (1.73,3.06)	0.339	(1.69,3.00)
σ_e^2	1.63 (1.41,2.27)	0.37	(0.96,2.41)	1.89 (1.32,2.66)	0.342	(1.27,2.58)
σ_{By}^2	0.98 (0.77,1.51)	0.30	(0.47,1.54)	0.93 (0.49,1.64)	0.296	(0.43,1.53)
σ_m^2	1.14 (0.98, 1.54)	0.23	(0.70,1.61)	1.29 (0.79,1.96)	0.299	(0.74,1.89)

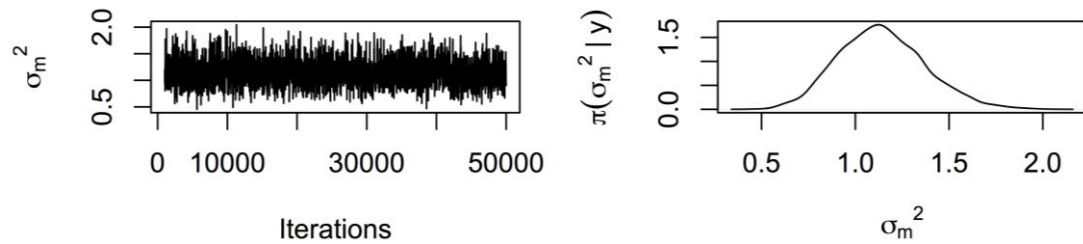
در جدول σ_u^2 ، σ_e^2 ، σ_{By}^2 و σ_m^2 به ترتیب مؤلفه‌های واریانس مربوط به تأثیر تصادفی ارزش اصلاحی حیوان، باقی‌مانده، سال تولد و مادری هستند.



شکل ۲. (چپ) نمودار اثر و (راست) توزیع حاشیه‌ای پسین مؤلفه واریانس مربوط به اثر تصادفی ارزش اصلاحی
Figure 2. (Left) the trace plot and (right) the marginal posterior distribution the variance component related to the additive genetic random effect

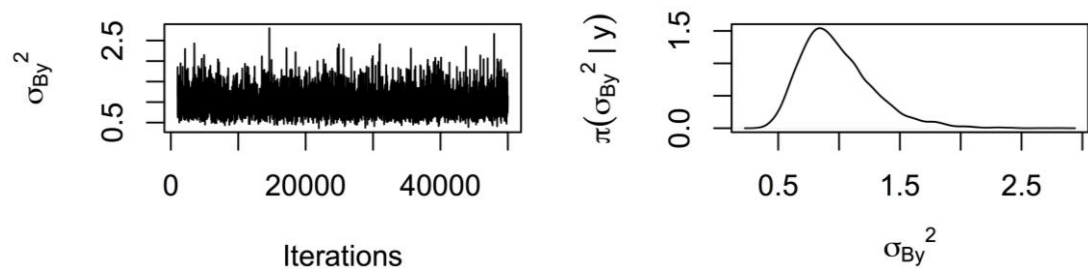


شکل ۳. (چپ) نمودار اثر و (راست) توزیع حاشیه‌ای پسین مؤلفه واریانس مربوط به اثر تصادفی باقی‌مانده
Figure 3. (Left) the trace plot and (right) the marginal posterior distribution the variance component related to the residual random effect



شکل ۴. (چپ) نمودار اثر و (راست) توزیع حاشیه‌ای پسین مؤلفه‌وارینانس مربوط به اثر تصادفی مادر

Figure 4. (Left) the trace plot and (right) the marginal posterior distribution the variance component related to the maternal random effect



شکل ۵. (چپ) نمودار اثر و (راست) توزیع حاشیه‌ای پسین مؤلفه‌وارینانس مربوط به اثر تصادفی سال تولد

Figure 5. (Left) the trace plot and (right) the marginal posterior distribution the variance component related to the year of birth random effect

برای مدل چهارم، وراثت‌پذیری مستقیم از رابطه زیر محاسبه شد:

$$h_u^2 = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2 + \sigma_m^2 + \sigma_{By}^2}$$

تأثیر مؤلفه‌های واریانس اثر تصادفی مادری، باقی‌مانده و سال تولد به‌صورت زیر بررسی شدند:

$$e^2 = \frac{\sigma_e^2}{\sigma_u^2 + \sigma_e^2 + \sigma_m^2 + \sigma_{By}^2}$$

$$m^2 = \frac{\sigma_m^2}{\sigma_u^2 + \sigma_e^2 + \sigma_m^2 + \sigma_{By}^2}$$

$$h_{By}^2 = \frac{\sigma_{By}^2}{\sigma_u^2 + \sigma_e^2 + \sigma_m^2 + \sigma_{By}^2}$$

نتایج برآورد وراثت‌پذیری و نسبت واریانس دیگر مؤلفه‌ها، انحراف معیار و چارک‌های برآورد بر پایه رهیافت بیزی مرسوم در جدول ۳ ارائه شده‌است. ضریب همبستگی کندال بین اثرگذاری ژنتیکی افزایشی مستقیم و اثرگذاری ژنتیکی مادری برای مدل چهارم محاسبه و میزان آن برابر ۰/۰۳۱ به دست آمد و میزان احتمال آزمون فرضیه صفر بودن ضریب همبستگی ۰/۳۸۳۱ به دست آمد و لذا صفر بودن ضریب همبستگی رد نشد.

نتیجه نشان می‌دهد که حدود ۳۸ درصد از تغییرپذیری وزن حیوان توسط متغیر ارزش اصلاحی بیان می‌شود درحالی‌که این میزان برای سال تولد حیوان ۱۶ درصد، برای اثر مادری حدود ۱۹ درصد و برای خطای تصادفی ۲۷ درصد است که بیانگر این مطلب است که ارزش اصلاحی دام بیشترین تأثیر و سال تولد کمترین تأثیر را روی وزن حیوان نسبت به دیگر اثرگذاری‌ها دارند. فاصله اطمینان ۹۵ درصد برای وراثت‌پذیری مستقیم برابر (۰/۵۰۵ و ۰/۳۲۹) است که نشان می‌دهد با اطمینان ۹۵ درصد وراثت‌پذیری جمعیت در این فاصله قرار می‌گیرد به‌طوری‌که دست‌کم حدود ۳۳ درصد و بیشینه حدود ۵۰ درصد از وراثت‌پذیری توسط ارزش اصلاحی دام در سطح معنی‌داری ۵ درصد قابل‌بیان است. فاصله قابل باور HPD در سطح ۹۵ درصد برای وراثت‌پذیری محاسبه شد که به‌صورت (۰/۵۳۴ و ۰/۲۲۳) به دست آمد و باور ۹۵ درصد وراثت‌پذیری جمعیت در این فاصله قرار می‌گیرد. نتایج برآورد وراثت‌پذیری و نسبت واریانس دیگر مؤلفه‌ها، انحراف معیار و چارک‌های برآورد بر پایه رهیافت بیزی تقریبی در جدول ۳ ارائه شده‌است. با رهیافت بیزی تقریبی به دست آمده است که نشان می‌دهد

نتایج برآورد وراثت‌پذیری و نسبت واریانس دیگر مؤلفه‌ها، انحراف معیار و چارک‌های برآورد بر پایه رهیافت بیزی مرسوم در جدول ۳ ارائه شده‌است. ضریب همبستگی کندال بین اثرگذاری ژنتیکی افزایشی مستقیم و اثرگذاری ژنتیکی مادری برای مدل چهارم محاسبه و میزان آن برابر ۰/۰۳۱ به دست آمد و میزان احتمال آزمون فرضیه صفر بودن ضریب همبستگی ۰/۳۸۳۱ به دست آمد و لذا صفر بودن ضریب همبستگی رد نشد.

$$h_u^2 = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2 + \sigma_m^2 + \sigma_{By}^2}$$

تأثیر مؤلفه‌های واریانس اثر تصادفی مادری، باقی‌مانده و سال تولد به‌صورت زیر بررسی شدند:

$$e^2 = \frac{\sigma_e^2}{\sigma_u^2 + \sigma_e^2 + \sigma_m^2 + \sigma_{By}^2}$$

$$m^2 = \frac{\sigma_m^2}{\sigma_u^2 + \sigma_e^2 + \sigma_m^2 + \sigma_{By}^2}$$

$$h_{By}^2 = \frac{\sigma_{By}^2}{\sigma_u^2 + \sigma_e^2 + \sigma_m^2 + \sigma_{By}^2}$$

نتایج برآورد وراثت‌پذیری و نسبت واریانس دیگر مؤلفه‌ها، انحراف معیار و چارک‌های برآورد بر پایه رهیافت بیزی مرسوم در جدول ۳ ارائه شده‌است. ضریب همبستگی کندال بین اثرگذاری ژنتیکی افزایشی مستقیم و اثرگذاری ژنتیکی مادری برای مدل چهارم محاسبه و میزان آن برابر ۰/۰۳۱ به دست آمد و میزان احتمال آزمون فرضیه صفر بودن ضریب همبستگی ۰/۳۸۳۱ به دست آمد و لذا صفر بودن ضریب همبستگی رد نشد.

$$MSEP = n^{-1} \sum (y - \hat{y})^2$$

$$RMSEP = \sqrt{n^{-1} \sum (y - \hat{y})^2}$$

$$RRMSEP = \sqrt{n^{-1} \sum \{(y - \hat{y})\hat{y}^{-1}\}^2}$$

برای بررسی نبود تفاوت معنی‌دار بین مقادیر برازش‌شده توسط دو رهیافت از آزمون تی (T) استفاده شد، که آماره آزمون و میزان احتمال به ترتیب ۰/۲۵۳ و ۰/۹۷۹۸ به دست آمدند که فرضیه صفر که نبود تفاوت معنی‌دار بین برازش وزن حیوان با رهیافت بیزی مرسوم و رهیافت بیزی تقریبی پذیرفته می‌شود. زمان محاسبات رهیافت بیزی مرسوم با به کار بردن بسته MCMCgImm برای ۵۰۰۰۰ تکرار الگوریتم MCMC با یک رایانه (Intel Core i7, 4GB, 2.7 GHz) حدود ۱۲۰ ثانیه و زمان محاسبات با رهیافت بیز تقریبی با به کار بردن بسته INLA کمتر از ده ثانیه به طول انجامید. در مدل‌های حیوانی حجم مشاهده‌های بزرگ است و کاهش زمان محاسبات یکی از هدف‌های مهم در این مدل‌ها است. گاهی صفت موردبررسی غیرنرمال و گسسته است که باعث پیچیدگی مدل و افزایش زمان محاسبات الگوریتم‌های MCMC می‌شود.

حدود ۳۶ درصد از تغییرپذیری وزن حیوان توسط متغیر ارزش اصلاحی بیان می‌شود درحالی‌که این میزان برای سال تولد حیوان ۱۴ درصد، برای اثر مادری حدود ۲۰ درصد و برای خطای تصادفی ۲۹ درصد است. فاصله اطمینان ۹۵ درصد برای وراثت‌پذیری مستقیم برابر (۰/۳۹۴ و ۰/۳۵۱) است که نشان می‌دهد با اطمینان ۹۵ درصد وراثت‌پذیری جمعیت در این فاصله قرار می‌گیرد به طوری که دست‌کم حدود ۳۵ درصد و بیشینه حدود ۴۰ درصد از وراثت‌پذیری توسط ارزش اصلاحی دام در سطح معنی‌داری ۵ درصد قابل‌بیان است. همان‌طور که ملاحظه می‌شود بین نتایج به دست آمده از رهیافت تقریبی و رهیافت بیزی مرسوم تفاوت محسوسی وجود ندارد.

برای مقایسه بیشتر بین دو رهیافت به کار گرفته شده و بررسی تفاوت معنی‌دار بین این دو رهیافت، مقادیر برازش‌شده وزن حیوان محاسبه شد و آنگاه معیارهای میانگین توان دوم خطای پیش‌بینی^۱ (MSEP)، ریشه میانگین توان دوم خطای پیش‌بینی^۲ (RMSEP) و ریشه میانگین توان دوم خطای نسبی پیش‌بینی^۳ (RRMSEP) از رابطه‌های زیر محاسبه شدند. نتایج در جدول ۴ آورده شده است.

جدول ۳. برآورد وراثت‌پذیری مستقیم، مادری، باقی‌مانده و نسبت مؤلفه واریانس سال تولد به واریانس فنوتیپی، انحراف معیار برآوردها و چارک‌های برآورد با دو رهیافت: بیزی مرسوم و بیزی تقریبی

Table 3. Estimate of the direct, maternal, residual heritability and year of birth variance component to phenotypic variance, standard deviation and quartiles of estimate with two approaches: the approximate Bayesian and the conventional Bayesian

Par.	Conventional Bayesian			Approximate Bayesian		
	Bayesian estimate	SD	(First and third quartiles)	Bayesian estimate	SD	(First and third quartiles)
h_u^2	0.381	0.078	(0.329,0.505)	0.364	0.018	(0.351,0.394)
e^2	0.271	0.067	(0.224,0.385)	0.294	0.005	(0.291,0.304)
m^2	0.188	0.036	(0.163,0.252)	0.199	0.007	(0.196,0.209)
h_{By}^2	0.160	0.041	(0.131,0.234)	0.142	0.016	(0.130,0.170)

جدول ۴. معیارهای MSEP، RMSEP و RRMSEP برای دو رهیافت بیزی مرسوم و بیزی تقریبی

Table 4. The MSEP, RMSEP and RRMSEP criterions for two approaches the conventional and the approximate Bayesian

Criterion	Conventional Bayesian approach	Approximate Bayesian approach
MSEP	0.775	0.696
RMSEP	0.880	0.834
RRMSEP	0.157	0.150

1. Mean Squared Error Prediction
2. Root Mean Squared Error Prediction
3. Relative Root Mean Squared Error Prediction

نتیجه‌گیری کلی

نظر می‌رسد. در مدل‌های آمیخته خطی تعمیم‌یافته فضایی نشان داده شده است که استفاده از یک کلاس توزیعی بزرگ‌تر و انعطاف‌پذیرتر از توزیع نرمال مانند چوله نرمال بسته برای اثر مکانی پنهان باعث افزایش دقت در نتایج می‌شود (Hosseini *et al.*, 2011). در مدل‌های حیوانی چون ارزش اصلاحی پنهان است و توزیع نرمال برای آن فرض می‌شود، ممکن است استفاده از خانواده توزیع چوله نرمال بسته که شامل خانواده نرمال نیز است، برای مدل‌بندی ارزش اصلاحی منجر به نتایج با دقت بالاتری شود که نیاز به بررسی دارد.

در مجموعه داده‌های موردبررسی با معیار مدل‌گزینی DIC مدل شامل اثرگذاری مادری مدل مناسب‌تر معرفی شد و نتایج دو رهیافت بیزی تفاوت معنی‌داری نداشت، اما زمان محاسبات روش بیزی تقریبی کمتر از روش بیزی مرسوم بود. حجم نمونه داده‌های موردبررسی در این مقاله کوچک است. در اغلب مسائل حیوانی با مجموعه داده‌ها با حجم بالا سروکار داریم. استفاده از رهیافت بیزی تقریبی برای مدل‌های حیوانی که به‌طورمعمول حجم مشاهده‌های بزرگ است سودمند به

REFERENCES

1. Bohlouli, M. & Alijani, S. (2012). Genotype by environment interaction for milk production traits in Iranian Holstein dairy cattle using random regression model. *Livestock Research for Rural Development*, 24, 1-7.
2. Bolker, B. M., Brooks, C. J., Clark, S. W., Geange, J. R. & Poulsen *et al.* (2009). Generalized linear mixed models: a practical guide for ecology and evolution, *Trends in ecology & evolution*, 24, 127-135.
3. Cobuci JA, Costa CN, Neto JB & Freitas AF, (2011). Genetic parameters for milk production by using random regression models with different alternatives of fixed regression modeling. *Revist Brasileira de Zootecnia*, 40, 557-567.
4. Eidsvik, J., Martino, S. & Rue, H. (2009). Approximate Bayesian Inference in Spatial Generalized Linear Mixed Models, *Scandinavian Journal of Statistics*, 36, 1-22.
5. Fong, Y., Rue, H., & Wakefield, J. (2010). Bayesian inference for generalized linear mixed models. *Bio-statistics*, 11, 397-412
6. Hadfield, J. D. (2010). MCMC Methods for Multi-Response Generalized Linear Mixed Models: The MCMCglmm R Package, *Journal of Statistical Software*, 33(2), 1-22.
7. Holand, A. M., Steinsland, I., Martino, S. & Jensen, H. (2013). Animal Models and Integrated Nested Laplace Approximations, *G3- Genes Genom Genet*, 3, 1241-1251.
8. Hosseini, F., Eidsvik, J. & Mohammadzadeh, M. (2011). Approximate Bayesian Inference in Spatial GLMM with Skew Normal Latent Variables, *Computational Statistics and Data Analysis*, 55, 1791-1806.
9. Hosseini, F. (2016). A new algorithm for estimating the parameters of the spatial generalized linear mixed models, *Environmental and Ecological Statistic*, Online publication. doi: 10.1007/s10651-015-0335-6.
10. Jamrozika, J., Gianolab, D. & Schaeffera, L. R. (2001). Bayesian estimation of genetic parameters for test day records in dairy cattle using linear hierarchical models, *Livestock Production Science*, 71, 223-240
11. Jasouri, M., Alijani, S., Talebi, R. & Hasanzadeh, A. (2014). Influence of maternal effects on estimation of genetic parameters of growth traits in Ghezel sheep using bayesian via Gibbs sampling technique, *Animal Science Researchs*, 34(1), 47-55.
12. McCulloch, C. (1997). Maximum likelihood algorithm for generalized linear mixed models, *journal of the American Statistical Association*, 92, 162-170.
13. McCullagh, P. & Nelder, J. A. (1989). *Generalized Linear Models*, London, Chapman and Hall.
14. Moghbeli Damane, M., Ayatollahi Mehrgardi, A., Asadi Foz, M. & Molaei Moghbeli, S. (2014). Estimation of genetic parameters for production traits and somatic cell score in Iranian Holstein dairy cattle using random regression model, *Journal of Livestock Science and Technologies*, 2(2), 43-52
15. Mohammadi, Alijani S, Rafet, S.A., Taghizadeh, A. & Buhloli, M. (2013). Comparison of fitting performance of polynomial functions in random regression model for test day milk yield in of Iranian Holstein dairy cattle, *Research on Animal Production*, 3(6), 46-63.
16. Mohammadi, Alijani S., Rafet, S.A. & Taghizadeh, A. (2014). Comparison of fitting performance of random regression animal and sire models for yield traits of Iranian Holstein dairy cattle, *Animal Science researches*, 23(4), 159-178.
17. Nelder, J. A. & Wedderburn, R. W. M. (1972). Generalized Linear Mixed Models, *Journal of the Royal Statistical Associate*, Ser. A, 135, 370-384.

18. Razmkabir, M., Moradi-Shahre Babak, M., Pakdel, A. & Neati Javaremi, A. (2012). Estimation of genetic parameters for test day records of milk yield in Holstein dairy cattle of Iran, *Iranian Journal of Animal Science*, 42, 171-178.
19. Rue, H. & Held, L. (2005). Gaussian Markov Random Fields: Theory and Applications, volume 104 of Monographs on Statistics and Applied Probability. Chapman & Hall, London.
20. Rue, H. & Martino, S. (2007). Approximate Bayesian Inference for Hierarchical Gaussian Markov Random fields models. *Journal of Statistical Planning and Inference*, 137, 3177-3199.
21. Rue, H., Martino, S. & Chopin, N. (2009). Approximate Bayesian Inference for Latent Gaussian Models using integrated nested Laplace approximations. *Journal of the Royal Statistical Society, Series B* 71, 319-392.
22. Schaeffer, LR. & Dekkers, JCM, (1994). Random regressions in animal models for test-day production in dairy cattle. Proc. *5th World Congress of Genetics Applied to Livestock Production, Guelph, Ontario, Canada XVIII*, 443-446.
23. Seraj, A., Vaez Torshizi, R. & Pakdel, A. (2011). Influence of fitting different animal models on genetic parameters of Day-old Chick Weight, *Iranian Journal of Animal Science*, 41, 363-371.
24. Simm, G. (1998). *Genetic Improvement of Cattle and Sheep*, Ipswich, U.K, Farming Press.
25. Sorensen, D. & Gianola, D. (2002), *Likelihood Bayesian and MCMC Methods in Genetics*, Springer.
26. Steinsland, I. & Jensen, H. (2010). Utilizing Gaussian Markov Random Field Properties of Bayesian Animal Models, *Biometrics*, 66(3), 763-771.
27. Tempelman, R. J. & Gianola, D. (1994). Assessment of a Poisson animal model for embryo yields in a simulated multiple ovulation-embryo transfer schemes. *Genetics selection evolution*, 26, 263-290.
28. Wilson, A. J., Reale, D., Clements, M. N., Morrissey, M. M., Postma, E., Walling, C. A., Kruuk, L. E. B. & Nussey, D. H. (2010). An ecologist's guide to the animal model. *Journal of Animal Ecology*, 79, 13-26.