

Note

On the Metropolis-Hastings Acceptance Probability to Add or Drop a Quantitative Trait Locus in Markov Chain Monte Carlo-Based Bayesian Analyses

Jean-Luc Jannink^{*,1} and Rohan L. Fernando[†]

^{*}Department of Agronomy and [†]Department of Animal Science, Iowa State University, Ames, Iowa 50010-1010

Manuscript received June 13, 2003
Accepted for publication October 15, 2003

ABSTRACT

The Metropolis-Hastings algorithm used in analyses that estimate the number of QTL segregating in a mapping population requires the calculation of an acceptance probability to add or drop a QTL from the model. Expressions for this acceptance probability need to recognize that sets of QTL are unordered such that the number of equivalent sets increases with the factorial of the QTL number. Here, we show how accounting for this fact affects the acceptance probability and review expressions found in the literature.

In a Bayesian quantitative trait loci (QTL) analysis it is possible to consider the number of QTL to be an unobserved parameter subject to estimation (SILLANPÄÄ and ARJAS 1998). Markov chain Monte Carlo (MCMC) implementations of such an analysis can use the Metropolis-Hastings algorithm to move between models with differing numbers of QTL and hence differing dimensionalities (SORENSEN and GIANOLA 2002, Chap. 11). The Metropolis-Hastings algorithm requires calculation of the probability to accept a move from the current model to the model proposed for the next iteration of the Markov chain. We denote this acceptance probability α below. Several expressions for the calculation of α have been presented in the literature (HEATH 1997; UIMARI and HOESCHELE 1997; SILLANPÄÄ and ARJAS 1998; STEPHENS and FISCH 1998; LEE and THOMAS 2000). Here, we show that a careful derivation of α , and in particular of the prior for the vector of QTL parameters, indicates that some of these expressions for α contain an error or at least a risk being misinterpreted.

The QTL model and derivation of α : Linear models in previous reports all differ in details that are not pertinent to the present derivation. In a very general model, the vector of phenotypes \mathbf{y} can be given as

$$\mathbf{y} = \boldsymbol{\mu} + \mathbf{X}\boldsymbol{\beta} + \mathbf{e}, \tag{1}$$

where the dimension of the incidence matrix \mathbf{X} depends on the number of QTL and the number of possible

genotypes at the QTL and the content of \mathbf{X} depends on the QTL genotypes of each individual, the vector $\boldsymbol{\beta}$ contains the phenotypic effects of all genotypes of all QTL, and $\mathbf{e} \sim N(0, \mathbf{I}\sigma^2)$ is a vector of residuals. We present here the acceptance probability for a move that adds a QTL; the acceptance probability for a move to drop a QTL follows easily. Denote $\boldsymbol{\theta}_q = (\tau_{1q}, \tau_{2q}, \dots, \tau_{rq})$ the parameters for QTL q , where each QTL is associated with r different types of parameters. Models exist in the literature with some subset of the following QTL parameters: QTL genotypes, QTL allelic effects, QTL variance, QTL allelic frequencies, and QTL position. Denote $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \dots, \boldsymbol{\theta}_{N_{\text{qtl}}})$, the vector of all QTL parameters excluding N_{qtl} , which is the number of QTL. A new QTL is proposed with parameters $\boldsymbol{\theta}_{N_{\text{qtl}}+1} = (\tau_{1N_{\text{qtl}}+1}, \tau_{2N_{\text{qtl}}+1}, \dots, \tau_{rN_{\text{qtl}}+1})$, so that the complete parameter vector for the candidate model with $N_{\text{qtl}} + 1$ QTL is $\boldsymbol{\theta}^+ = (\boldsymbol{\theta}, \boldsymbol{\theta}_{N_{\text{qtl}}+1})$. The acceptance probability for a Markov chain with reversible jumps takes the form

$$\alpha = \min \left\{ 1, \frac{p_d}{p_a} \times \frac{\pi(N_{\text{qtl}} + 1, \boldsymbol{\mu}, \sigma^2, \boldsymbol{\theta}^+)}{\pi(N_{\text{qtl}}, \boldsymbol{\mu}, \sigma^2, \boldsymbol{\theta})} \times \frac{q(\mathbf{u}^+ | N_{\text{qtl}} + 1, \boldsymbol{\mu}, \sigma^2, \boldsymbol{\theta}^+)}{q(\mathbf{u} | N_{\text{qtl}}, \boldsymbol{\mu}, \sigma^2, \boldsymbol{\theta})} \times J \right\} \tag{2}$$

(GREEN 1995), where p_d and p_a are, respectively, the probabilities of proposing to drop and add a QTL to the model; $\pi(N_{\text{qtl}}, \boldsymbol{\mu}, \sigma^2, \boldsymbol{\theta})$ is the posterior density of the model $(N_{\text{qtl}}, \boldsymbol{\mu}, \sigma^2, \boldsymbol{\theta})$; $q(\mathbf{u} | N_{\text{qtl}}, \boldsymbol{\mu}, \sigma^2, \boldsymbol{\theta})$ is the probability of proposing a stochastic vector \mathbf{u} conditional on the current model such that $\boldsymbol{\theta}^+ = (\boldsymbol{\theta}, \boldsymbol{\theta}_{N_{\text{qtl}}+1}) = g_{N_{\text{qtl}}+1}(\boldsymbol{\theta}, \mathbf{u})$, and $g_{N_{\text{qtl}}+1}(\boldsymbol{\theta}, \mathbf{u})$ is a function whose Jacobian is J (SORENSEN and GIANOLA 2002, Chap. 11); finally, $q(\mathbf{u}^+ | N_{\text{qtl}} + 1, \boldsymbol{\mu}, \sigma^2, \boldsymbol{\theta}^+)$ is the probability of the proposal to drop

¹Corresponding author: Department of Agronomy, Iowa State University, 1208 Agronomy Hall, Ames, IA 50010-1010.
E-mail: jjannink@iastate.edu

from $\boldsymbol{\theta}^+$ back down to $\boldsymbol{\theta}$. If $g_{N_{\text{qtl}}+1}$ is chosen as the identity transformation such that $\boldsymbol{\theta}_{N_{\text{qtl}}+1} = \mathbf{u}$, then $J = 1$ and this term can be left out (SILLANPÄÄ and ARJAS 1998). Usually, p_d and p_a are taken to be equal so that their ratio can also be left out.

The posterior density $\pi(N_{\text{qtl}}, \boldsymbol{\mu}, \sigma^2, \boldsymbol{\theta})$ given observable and prior information is

$$\pi(N_{\text{qtl}}, \boldsymbol{\mu}, \sigma^2, \boldsymbol{\theta}) \propto p(\mathbf{y}|N_{\text{qtl}}, \boldsymbol{\mu}, \sigma^2, \boldsymbol{\theta})p(N_{\text{qtl}})p(\boldsymbol{\mu})p(\sigma^2)p(\boldsymbol{\theta}|N_{\text{qtl}}), \quad (3)$$

where $p(\mathbf{y}|N_{\text{qtl}}, \boldsymbol{\mu}, \sigma^2, \boldsymbol{\theta})$ represents the likelihood assuming $\mathbf{e} \sim N(0, \mathbf{I}\sigma^2)$ and $p(\ast)$ is the prior distribution for parameter \ast . To add a QTL, $\boldsymbol{\theta}_{N_{\text{qtl}}+1}$ is chosen from its prior, while to drop a QTL, one QTL among the $(N_{\text{qtl}} + 1)$ is chosen at random and its parameters are removed from the model, such that

$$\alpha = \min\left\{1, \frac{p(\mathbf{y}|N_{\text{qtl}} + 1, \boldsymbol{\mu}, \sigma^2, \boldsymbol{\theta}^+)p(N_{\text{qtl}} + 1)p(\boldsymbol{\mu})p(\sigma^2)p(\boldsymbol{\theta}^+|N_{\text{qtl}} + 1)}{p(\mathbf{y}|N_{\text{qtl}}, \boldsymbol{\mu}, \sigma^2, \boldsymbol{\theta})p(N_{\text{qtl}})p(\boldsymbol{\mu})p(\sigma^2)p(\boldsymbol{\theta}|N_{\text{qtl}})} \times \frac{(N_{\text{qtl}} + 1)^{-1}}{p(\boldsymbol{\theta}_{N_{\text{qtl}}+1})}\right\}. \quad (4)$$

In some previous expressions for α , authors have equated $p(\boldsymbol{\theta}^+|N_{\text{qtl}} + 1)$ to $p(\boldsymbol{\theta}|N_{\text{qtl}})p(\boldsymbol{\theta}_{N_{\text{qtl}}+1})$ such that these terms cancel out of Equation 4. Here we show that if $\boldsymbol{\theta}$ is unordered, this equality does not hold. The vector $\boldsymbol{\theta}$ has N_{qtl} subvectors, where each subvector contains the parameters for a locus. Consider a model with QTL defined by the specific subvectors $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \dots, \boldsymbol{\theta}_{N_{\text{qtl}}}$. There are $N_{\text{qtl}}!$ ways of assigning these $\boldsymbol{\theta}_i$ to the subvectors of $\boldsymbol{\theta}$. That is, for example, $\{\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \dots, \boldsymbol{\theta}_{N_{\text{qtl}}}\}$ represents the same model as $\{\boldsymbol{\theta}_2, \boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_{N_{\text{qtl}}}\}$. Since there are $N_{\text{qtl}}!$ permutations that result in the same model, the prior for $\boldsymbol{\theta}$ is $p(\boldsymbol{\theta}|N_{\text{qtl}}) = N_{\text{qtl}}! \prod_{q=1}^{N_{\text{qtl}}} p(\boldsymbol{\theta}_q)$. Consider more formally the case of two QTL, A and B, with parameters $\boldsymbol{\theta}_A$ and $\boldsymbol{\theta}_B$. The prior probability that $\boldsymbol{\theta}_A = \boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_B = \boldsymbol{\theta}_2$ is $p(\boldsymbol{\theta}_1)p(\boldsymbol{\theta}_2)$ and $\int_{\Theta} p(\boldsymbol{\theta}_1)p(\boldsymbol{\theta}_2)d\theta_1d\theta_2 = 1$, where Θ is the domain of $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$. Note that, for $\boldsymbol{\theta}_A \neq \boldsymbol{\theta}_B$, both $(\boldsymbol{\theta}_1 = \boldsymbol{\theta}_A)(\boldsymbol{\theta}_2 = \boldsymbol{\theta}_B)$ and $(\boldsymbol{\theta}_1 = \boldsymbol{\theta}_B)(\boldsymbol{\theta}_2 = \boldsymbol{\theta}_A)$ contribute to the integral. However, we consider these two assignments to represent the same model such that $p(\boldsymbol{\theta} = \{\boldsymbol{\theta}_A, \boldsymbol{\theta}_B\}) = 2p(\boldsymbol{\theta}_1)p(\boldsymbol{\theta}_2)$. This reasoning can be extended to higher numbers of QTL, leading us back to $p(\boldsymbol{\theta}|N_{\text{qtl}}) = N_{\text{qtl}}! \prod_{q=1}^{N_{\text{qtl}}} p(\boldsymbol{\theta}_q)$. Consequently, the prior and proposal terms for QTL parameters simplify to

$$\frac{p(\boldsymbol{\theta}^+|N_{\text{qtl}} + 1)}{p(\boldsymbol{\theta}|N_{\text{qtl}})p(\boldsymbol{\theta}_{N_{\text{qtl}}+1})} = \frac{(N_{\text{qtl}} + 1)! \prod_{q=1}^{N_{\text{qtl}}+1} p(\boldsymbol{\theta}_q)}{N_{\text{qtl}}! (\prod_{q=1}^{N_{\text{qtl}}} p(\boldsymbol{\theta}_q)) \times p(\boldsymbol{\theta}_{N_{\text{qtl}}+1})} = N_{\text{qtl}} + 1. \quad (5)$$

After all simplifications are performed, we obtain

$$\alpha = \min\left\{1, \frac{p(\mathbf{y}|N_{\text{qtl}} + 1, \boldsymbol{\theta}^+)p(N_{\text{qtl}} + 1)}{p(\mathbf{y}|N_{\text{qtl}}, \boldsymbol{\theta})p(N_{\text{qtl}})}\right\}. \quad (6)$$

If the prior for N_{qtl} is taken to be Poisson with mean λ , this gives

$$\alpha = \min\left\{1, \frac{p(\mathbf{y}|N_{\text{qtl}} + 1, \boldsymbol{\theta}^+)}{p(\mathbf{y}|N_{\text{qtl}}, \boldsymbol{\theta})} \times \frac{\lambda}{(N_{\text{qtl}} + 1)}\right\}. \quad (7)$$

This expression is identical to that given in SILLANPÄÄ and ARJAS (1998) under step 1.2 in their Appendix A, except that they square the $N_{\text{qtl}} + 1$ term in the denominator. The following reasoning reinforces our belief that Equation 7 is correct: consider running the sampler in the absence of phenotypic data. In that case, $p(\mathbf{y}|N_{\text{qtl}}, \boldsymbol{\mu}, \sigma^2, \boldsymbol{\theta}) = 1$ for all models, and the posterior should be equal to the prior. Running a Metropolis-Hastings sampler using (7) will indeed give samples from a Poisson distribution with mean λ , while the acceptance probabilities given in SILLANPÄÄ and ARJAS (1998) would not.

Other derivations of α in the literature: In their derivation for α , UIMARI and HOESCHELE (1997) discuss a reversible-jump step moving between one and two QTL. In their model, they force the second QTL to be to the right of the first QTL. Thus, their model does not allow for the two permutations discussed above. Consequently, the results given in this note are irrelevant to their derivation. HEATH (1997) does not discuss the number of permutations that results in the same vector of QTL parameters. The equation he provides for α (his Equation B2) appears to decompose the stationary distribution π into the product of the prior with the likelihood as we do in Equation 4. A $(N_{\text{qtl}} + 1)$ term nevertheless appears in the denominator of Equation B2 (he denotes N_{qtl} by k) and we therefore believe Equation B2 to be incorrect. STEPHENS and FISCH (1998) also appear to make the mistake that is the main point of this note. In addition, they decompose the QTL position into a marker interval and a recombination frequency within the interval. They account for the prior of the recombination frequency but do not account for the prior of the specific interval in which the QTL is placed. LEE and THOMAS (2000) present what is, to the best of our ability to discern, a correct acceptance probability with respect to the main point of this note. Their derivation of α , however, is unclear. They document the need for the $(N_{\text{qtl}} + 1)^{-1}$ term because of the random choice of which QTL to drop (their Equation A4). They also recognize the $\lambda \times (N_{\text{qtl}} + 1)^{-1}$ term as the ratio of the Poisson priors. Since they do not discuss the permutations of the vector of QTL parameters, however, the attentive reader will be puzzled by their final expression for α , Equation A5. A different variant of the problem presented in this note arises in BINK (2002). That study presents research on a finite polygenic model in which unlinked loci are modeled, that is, the QTL have no position. Our derivation above is general to whether or not the QTL have positions and therefore is also relevant to this case. Consequently, the term in the denominator of the acceptance probability given (BINK 2002, p. 248)

should not be squared. Finally, SORESENSEN and GIANOLA (2002, pp. 690–699) order θ by labeling each of its subvectors according to when the subvector was sampled. The Metropolis-Hastings proposal they discuss is different from that discussed above in that it allows only the last QTL sampled to be dropped. Thus, they have generalized UIMARI and HOESCHELE's (1997) ordered vector of two QTL to an ordered vector of an arbitrary number of QTL, and this note is also irrelevant to this derivation. It is heartening to see that, despite this difference in ordered *vs.* unordered QTL vector, SORESENSEN and GIANOLA (2002) arrive at the same expression for α as we do in Equation 7. Indeed, considering the reasoning given above of running the sampler in the absence of phenotypic data, it is necessary that they should arrive at the same expression.

The fact that this error has remained in the literature for over 5 years underscores the view that while Bayesian analysis using Markov chain Monte Carlo is incredibly flexible and therefore powerful, the devil is in the details. Furthermore, incorrect analyses can give results that seem quite reasonable.

We acknowledge discussions with Dr. Mikko Sillanpää.

LITERATURE CITED

- BINK, M. C. A. M., 2002 On flexible finite polygenic models for multiple-trait evaluation. *Genet. Res.* **80**: 245–256.
- GREEN, P. J., 1995 Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika* **82**: 711–732.
- HEATH, S. C., 1997 Markov chain Monte Carlo segregation and linkage analysis for oligogenic models. *Am. J. Hum. Genet.* **61**: 748–760.
- LEE, J. K., and D. C. THOMAS, 2000 Performance of Markov chain-Monte Carlo approaches for mapping genes in oligogenic models with an unknown number of loci. *Am. J. Hum. Genet.* **67**: 1232–1250.
- SILLANPÄÄ, M. J., and E. ARJAS, 1998 Bayesian mapping of multiple quantitative trait loci from incomplete inbred line cross data. *Genetics* **148**: 1373–1388.
- SORENSEN, D., and D. GIANOLA, 2002 *Likelihood, Bayesian and MCMC Methods in Genetics*. Springer Verlag, Berlin.
- STEPHENS, D. A., and R. D. FISCH, 1998 Bayesian analysis of quantitative trait locus data using reversible jump Markov chain Monte Carlo. *Biometrics* **54**: 1334–1347.
- UIMARI, P., and I. HOESCHELE, 1997 Mapping-linked quantitative trait loci using Bayesian analysis and Markov chain Monte Carlo algorithms. *Genetics* **146**: 735–743.

Communicating editor: R. W. DOERGE

