

# Genomic Selection: Bayesian Methods

Rohan L. Fernando

Iowa State University

June 6, 2009

# Part I

## Bayesian Inference: Theory

# More parameters than observations

- ▶ Model:

$$y_i = \mu + \sum_j X_{ij}\alpha_j + \mathbf{e}_i$$

$X_{ij}$  is genotype of  $i$  at marker locus  $j$

- ▶ Problem: more markers than animals with genotypes and phenotypes
- ▶ Expect many of the  $\alpha_j$  to be zero or close to zero
- ▶ Combine this prior information with phenotypic data to estimate  $\alpha_j$  (Meuwissen et al., 2001; Genetics 157:1819-1829)
- ▶ Bayesian inference provides a formal way to combine prior information with data

# Meaning of probability

- ▶ In Bayesian inference, probabilities are used to quantify your beliefs or knowledge about possible values of parameters
  - ▶ What is the probability that  $h^2 > 0.5$ ?
  - ▶ What is the probability that milk yield is controlled by more than 100 loci?
- ▶ In the frequency approach, probability is a limiting frequency

# Essentials of Bayesian approach

- ▶ **Prior** probabilities quantify beliefs about parameters before the data are analyzed
- ▶ Parameters are related to the data through the model or “**likelihood**”, which is the conditional probability density for the data given the parameters
- ▶ The prior and the likelihood are combined using Bayes theorem to obtain **posterior** probabilities, which are conditional probabilities for the parameters given the data
- ▶ Inferences about parameters are based on the posterior

# Bayes Theorem

- ▶ Let  $f(\theta)$  denote the prior probability density for  $\theta$
- ▶ Let  $f(\mathbf{y}|\theta)$  denote the likelihood
- ▶ Then, the posterior probability of  $\theta$  is:

$$f(\theta|\mathbf{y}) = \frac{f(\mathbf{y}|\theta)f(\theta)}{f(\mathbf{y})}$$
$$\propto f(\mathbf{y}|\theta)f(\theta)$$

# Computing posteriors

- ▶ Often no closed form for  $f(\boldsymbol{\theta}|\mathbf{y})$
- ▶ Further, even if computing  $f(\boldsymbol{\theta}|\mathbf{y})$  is feasible, obtaining  $f(\theta_i|\mathbf{y})$  would require integrating over many dimensions
- ▶ Thus, in many situations, inferences are made using the empirical posterior constructed by drawing samples from  $f(\boldsymbol{\theta}|\mathbf{y})$
- ▶ Gibbs sampler is widely used for drawing samples from posteriors

# Gibbs sampler

- ▶ Want to draw samples from  $f(x_1, x_2, \dots, x_n)$
- ▶ Even though it may be possible to compute  $f(x_1, x_2, \dots, x_n)$ , it is difficult to draw samples directly from  $f(x_1, x_2, \dots, x_n)$
- ▶ Gibbs:
  - ▶ Get valid a starting point  $\mathbf{x}^0$
  - ▶ Draw sample  $\mathbf{x}^t$  as:

$$\begin{array}{ll} x_1^t & \text{from } f(x_1 | x_2^{t-1}, x_3^{t-1}, \dots, x_n^{t-1}) \\ x_2^t & \text{from } f(x_2 | x_1^t, x_3^{t-1}, \dots, x_n^{t-1}) \\ x_3^t & \text{from } f(x_3 | x_1^t, x_2^t, \dots, x_n^{t-1}) \\ & \vdots \\ & \vdots \\ x_n^t & \text{from } f(x_n | x_1^t, x_2^t, \dots, x_{n-1}^t) \end{array}$$

- ▶ The sequence  $\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^n$  is a Markov chain with stationary distribution  $f(x_1, x_2, \dots, x_n)$

# Inference from Markov chain

Can show that samples obtained from the **Markov chain** can be used to draw inferences from  $f(x_1, x_2, \dots, x_n)$  provided the chain is:

- ▶ **Irreducible**: can move from any state  $i$  to any other state  $j$
- ▶ **Positive recurrent**: return time to any state has finite expectation
- ▶ *Markov Chains*, J. R. Norris (1997)

# Metropolis-Hastings sampler

- ▶ Sometimes may not be able to draw samples directly from  $f(x_i | \mathbf{x}_{i-})$
- ▶ Convergence of the Gibbs sampler may be too slow
- ▶ Metropolis-Hastings (MH) for sampling from  $f(x)$ :
  - ▶ a candidate sample,  $y$ , is drawn from a proposal distribution  $q(y | x^{t-1})$

- ▶
$$x^t = \begin{cases} y & \text{with probability } \alpha \\ x^{t-1} & \text{with probability } 1 - \alpha \end{cases}$$

- ▶
$$\alpha = \min\left(1, \frac{f(y)q(x^{t-1} | y)}{f(x^{t-1})q(y | x^{t-1})}\right)$$

- ▶ The sequence of samples from MH is a Markov chain with stationary distribution  $f(x)$

# Proposal distributions

Two main types:

- ▶ Approximations of the target density:  $f(x)$ 
  - ▶ Not easy to find approximation that is easy to sample from
  - ▶ High acceptance rate is good!
- ▶ Random walk type: stay close to the previous sample
  - ▶ Generally easy to construct proposal
  - ▶ High acceptance rate may indicate that candidate is too close to previous sample
  - ▶ Intermediate acceptance rate is good

## Part II

# Bayesian Inference: Application to Whole Genome Analyses

# Model

Model:

$$y_i = \mu + \sum_j X_{ij} \alpha_j + \mathbf{e}_i$$

Priors:

- ▶  $\mu \propto \text{constant}$  (not proper, but posterior is proper)
- ▶  $\mathbf{e}_i \sim (\text{iid})\text{N}(0, \sigma_e^2)$ ;  $\sigma_e^2 \sim \nu_e \mathbf{S}_e^2 \chi_{\nu_e}^{-2}$
- ▶ Consider several different priors for  $\alpha_j$

# Normal

- ▶ Prior:  $(\alpha_j | \sigma_\alpha^2) \sim (\text{iid}) \text{N}(0, \sigma_\alpha^2)$ ;  $\sigma_\alpha^2$  is known
- ▶ What is  $\sigma_\alpha^2$ ?
- ▶ Assume the QTL genotypes are a subset of those available for the analysis
  - ▶ Then, the genotypic value of  $i$  can be written as:

$$g_i = \mu + \mathbf{x}'_i \boldsymbol{\alpha}$$

- ▶ Note that  $\boldsymbol{\alpha}$  is common to all  $i$
  - ▶ Thus, the variance of  $g_i$  comes from  $\mathbf{x}'_i$  being random
- ▶ So,  $\sigma_\alpha^2$  is not the genetic variance at a locus
- ▶ If locus  $j$  is randomly sampled from all the loci available for analysis:
  - ▶ Then,  $\alpha_j$  will be a random variable
  - ▶  $\sigma_\alpha^2 = \text{Var}(\alpha_j)$

## Relationship of $\sigma_\alpha^2$ to genetic variance

Assume loci with effect on trait are in linkage equilibrium. Then, the additive genetic variance is

$$V_A = \sum_j^k 2p_j q_j \alpha_j^2,$$

where  $p_j = 1 - q_j$  is gene frequency at SNP locus  $j$ .  
Letting  $U_j = 2p_j q_j$  and  $V_j = \alpha_j^2$ ,

$$V_A = \sum_j^k U_j V_j$$

For a randomly sampled locus, covariance between  $U_j$  and  $V_j$  is

$$C_{UV} = \frac{\sum_j U_j V_j}{k} - \left(\frac{\sum_j U_j}{k}\right)\left(\frac{\sum_j V_j}{k}\right)$$

## Relationship of $\sigma_\alpha^2$ to genetic variance

Rearranging the previous expression for  $C_{UV}$  gives

$$\sum_j U_j V_j = kC_{UV} + \left(\sum_j U_j\right)\left(\frac{\sum_j V_j}{k}\right)$$

So,

$$V_A = kC_{UV} + \left(\sum_j 2p_j q_j\right)\left(\frac{\sum_j \alpha_j^2}{k}\right)$$

Letting  $\sigma_\alpha^2 = \frac{\sum_j \alpha_j^2}{k}$  gives

$$V_A = kC_{UV} + \left(\sum_j 2p_j q_j\right)\sigma_\alpha^2$$

and,

$$\sigma_\alpha^2 = \frac{V_A - kC_{UV}}{\sum_j 2p_j q_j}$$

# Blocked Gibbs sampler

- ▶ Let  $\theta' = [\mu, \alpha']$
- ▶ Can show that  $(\theta | \mathbf{y}, \sigma_e^2) \sim N(\hat{\theta}, \mathbf{C}^{-1} \sigma_e^2)$

▶

$$\hat{\theta} = \mathbf{C}^{-1} \mathbf{W}' \mathbf{y}; \quad \mathbf{W} = [\mathbf{1}, \mathbf{X}]$$

▶

$$\mathbf{C} = \begin{bmatrix} \mathbf{1}'\mathbf{1} & \mathbf{1}'\mathbf{X} \\ \mathbf{X}'\mathbf{1} & \mathbf{X}'\mathbf{X} + I \frac{\sigma_e^2}{\sigma_\alpha^2} \end{bmatrix}$$

- ▶ Blocked Gibbs sampler
  - ▶ García-Cortés and Sorensen (1996, GSE 28:121-126)
  - ▶ *Likelihood, Bayesian and MCMC Methods* ... (LBMMQG, Sorensen and Gianola, 2002)

# Full conditionals for single-site Gibbs

▶  $(\mu | \mathbf{y}, \boldsymbol{\alpha}, \sigma_e^2) \sim \text{N}\left(\frac{\mathbf{1}'(\mathbf{y} - \mathbf{X}\boldsymbol{\alpha})}{n}, \frac{\sigma_e^2}{n}\right)$

▶  $(\alpha_j | \mathbf{y}, \mu, \boldsymbol{\alpha}_{j-}, \sigma_e^2) \sim \text{N}\left(\hat{\alpha}_j, \frac{\sigma_e^2}{c_j}\right)$

▶

$$\hat{\alpha}_j = \frac{\mathbf{x}_j' \mathbf{w}}{c_j}$$

▶

$$\mathbf{w} = \mathbf{y} - \mathbf{1}\mu - \sum_{j' \neq j} \mathbf{x}_{j'} \alpha_{j'}$$

▶

$$c_j = (\mathbf{x}_j' \mathbf{x}_j + \frac{\sigma_e^2}{\sigma_\alpha^2})$$

▶  $(\sigma_e^2 | \mathbf{y}, \mu, \boldsymbol{\alpha}) \sim [(\mathbf{y} - \mathbf{W}\boldsymbol{\theta})'(\mathbf{y} - \mathbf{W}\boldsymbol{\theta}) + \nu_e \mathbf{S}_e^2] \chi_{(\nu_e + n)}^{-2}$

## Derive: full conditional for $\alpha_j$

From Bayes' Theorem,

$$f(\alpha_j | \mathbf{y}, \mu, \boldsymbol{\alpha}_{j_-}, \sigma_e^2) = \frac{f(\alpha_j, \mathbf{y}, \mu, \boldsymbol{\alpha}_{j_-}, \sigma_e^2)}{f(\mathbf{y}, \mu, \boldsymbol{\alpha}_{j_-}, \sigma_e^2)}$$

$$\propto f(\mathbf{y} | \alpha_j, \mu, \boldsymbol{\alpha}_{j_-}, \sigma_e^2) f(\alpha_j) f(\mu, \boldsymbol{\alpha}_{j_-}, \sigma_e^2)$$

$$\propto (\sigma_e^2)^{-n/2} \exp\left\{-\frac{(\mathbf{w} - \mathbf{x}_j \alpha_j)' (\mathbf{w} - \mathbf{x}_j \alpha_j)}{2\sigma_e^2}\right\} (\sigma_\alpha^2)^{-1/2} \exp\left\{-\frac{\alpha_j^2}{2\sigma_\alpha^2}\right\}$$

where

$$\mathbf{w} = \mathbf{y} - \mathbf{1}\mu - \sum_{j \neq j'} \mathbf{x}_{j'} \alpha_{j'}$$

## Derive: full conditional for $\alpha_j$

The exponential terms in the joint density can be written as:

$$-\frac{1}{2\sigma_e^2} \left\{ \mathbf{w}'\mathbf{w} - 2\mathbf{x}'_j\mathbf{w}\alpha_j + [\mathbf{x}'_j\mathbf{x}_j + \frac{\sigma_e^2}{\sigma_\alpha^2}]\alpha_j^2 \right\}$$

Completing the square in this expression with respect to  $\alpha_j$  gives

$$-\frac{1}{2\sigma_e^2} \left\{ c_j(\alpha_j - \hat{\alpha}_j)^2 + \mathbf{w}'\mathbf{w} - c_j\hat{\alpha}_j^2 \right\}$$

where

$$\hat{\alpha}_j = \frac{\mathbf{x}_j\mathbf{w}}{c_j}$$

So,

$$f(\alpha_j | \mathbf{y}, \mu, \boldsymbol{\alpha}_{j-}, \sigma_e^2) \propto \exp \left\{ -\frac{(\alpha_j - \hat{\alpha}_j)^2}{2\frac{\sigma_e^2}{c_j}} \right\}$$

## Full conditional for $\sigma_e^2$

From Bayes' theorem,

$$\begin{aligned}f(\sigma_e^2 | \mathbf{y}, \mu, \boldsymbol{\alpha}) &= \frac{f(\sigma_e^2, \mathbf{y}, \mu, \boldsymbol{\alpha})}{f(\mathbf{y}, \mu, \boldsymbol{\alpha})} \\ &\propto f(\mathbf{y} | \sigma_e^2, \mu, \boldsymbol{\alpha}) f(\sigma_e^2) f(\mu, \boldsymbol{\alpha})\end{aligned}$$

where

$$f(\mathbf{y} | \sigma_e^2, \mu, \boldsymbol{\alpha}) \propto (\sigma_e^2)^{-n/2} \exp\left\{-\frac{(\mathbf{w} - \mathbf{x}_j \alpha_j)'(\mathbf{w} - \mathbf{x}_j \alpha_j)}{2\sigma_e^2}\right\}$$

and

$$f(\sigma_e^2) = \frac{(S_e^2 \nu_e / 2)^{\nu_e / 2}}{\Gamma(\nu_e / 2)} (\sigma_e^2)^{-(2 + \nu_e) / 2} \exp\left(-\frac{\nu_e S_e^2}{2\sigma_e^2}\right)$$

## Full conditional for $\sigma_e^2$

So,

$$f(\sigma_e^2 | \mathbf{y}, \mu, \boldsymbol{\alpha}) \propto (\sigma_e^2)^{-(2+n+\nu_e)/2} \exp\left(-\frac{SSE + \nu_e S_e^2}{2\sigma_e^2}\right)$$

where

$$SSE = (\mathbf{w} - \mathbf{x}_j \alpha_j)' (\mathbf{w} - \mathbf{x}_j \alpha_j)$$

So,

$$f(\sigma_e^2 | \mathbf{y}, \mu, \boldsymbol{\alpha}) \sim \tilde{\nu}_e \tilde{S}_e^2 \chi_{\tilde{\nu}_e}^{-2}$$

where

$$\tilde{\nu}_e = n + \nu_e; \quad \tilde{S}_e^2 = \frac{SSE + \nu_e S_e^2}{\tilde{\nu}_e}$$

## Alternative view of Normal prior

Consider fixed linear model:

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{X}\alpha + \mathbf{e}$$

This can be also written as

$$\mathbf{y} = [\mathbf{1} \quad \mathbf{X}] \begin{bmatrix} \mu \\ \alpha \end{bmatrix} + \mathbf{e}$$

Suppose we observe for each locus:

$$y_j^* = \alpha_j + \epsilon_j$$

# Least Squares with Additional Data

Fixed linear model with the additional data:

$$\begin{bmatrix} \mathbf{y} \\ \mathbf{y}^* \end{bmatrix} = \begin{bmatrix} \mathbf{1} & \mathbf{X} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mu \\ \alpha \end{bmatrix} + \begin{bmatrix} \mathbf{e} \\ \epsilon \end{bmatrix}$$

OLS Equations:

$$\begin{bmatrix} \mathbf{1}' & \mathbf{0}' \\ \mathbf{X}' & \mathbf{I}' \end{bmatrix} \begin{bmatrix} I_n \frac{1}{\sigma_e^2} & \mathbf{0} \\ \mathbf{0} & I_k \frac{1}{\sigma_\epsilon^2} \end{bmatrix} \begin{bmatrix} \mathbf{1} & \mathbf{X} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \hat{\mu} \\ \hat{\alpha} \end{bmatrix} = \begin{bmatrix} \mathbf{1}' & \mathbf{0}' \\ \mathbf{X}' & \mathbf{I}' \end{bmatrix} \begin{bmatrix} I_n \frac{1}{\sigma_e^2} & \mathbf{0} \\ \mathbf{0} & I_k \frac{1}{\sigma_\epsilon^2} \end{bmatrix} \begin{bmatrix} \mathbf{y} \\ \mathbf{y}^* \end{bmatrix}$$

$$\begin{bmatrix} \mathbf{1}'\mathbf{1} & \mathbf{1}'\mathbf{X} \\ \mathbf{X}'\mathbf{1} & \mathbf{X}'\mathbf{X} + I \frac{\sigma_e^2}{\sigma_\epsilon^2} \end{bmatrix} \begin{bmatrix} \hat{\mu} \\ \hat{\alpha} \end{bmatrix} = \begin{bmatrix} \mathbf{1}'\mathbf{y} \\ \mathbf{X}'\mathbf{y} + \mathbf{y}^* \frac{\sigma_e^2}{\sigma_\epsilon^2} \end{bmatrix}$$

# Univariate- $t$

Prior:

$$(\alpha_j | \sigma_j^2) \sim \mathbf{N}(\mathbf{0}, \sigma_j^2)$$

$$\sigma_j^2 \sim \nu_\alpha \mathbf{S}_{\nu_\alpha}^2 \chi_{\nu_\alpha}^{-2}$$

Can show that the unconditional distribution for  $\alpha_j$  is

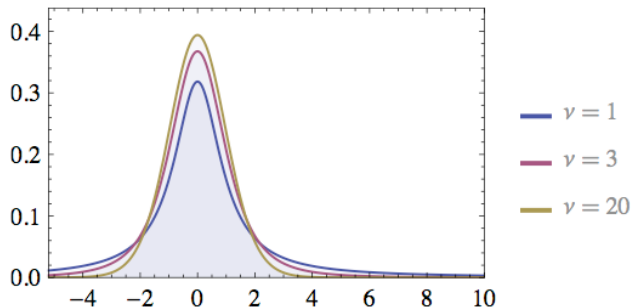
$$\alpha_j \sim (\text{iid}) t(\mathbf{0}, \mathbf{S}_{\nu_\alpha}^2, \nu_\alpha)$$

(Sorensen and Gianola, 2002, LBMMQG pages 28,60)

This is Bayes-A (Meuwissen et al., 2001; Genetics 157:1819-1829)

# Univariate- $t$

Plots of PDF for typical parameters:



---

Generated by Wolfram|Alpha ([www.wolframalpha.com](http://www.wolframalpha.com))

## Full conditional for single-site Gibbs

Full conditionals are the same as in the "Normal" model for  $\mu$ ,  $\alpha_j$ , and  $\sigma_e^2$ . Let

$$\boldsymbol{\xi} = [\sigma_1^2, \sigma_2^2, \dots, \sigma_k^2]$$

Full conditional conditional for  $\sigma_j^2$ :

$$\begin{aligned} f(\sigma_j^2 | \mathbf{y}, \mu, \boldsymbol{\alpha}, \boldsymbol{\xi}_{j-}, \sigma_e^2) &\propto f(\mathbf{y}, \mu, \boldsymbol{\alpha}, \boldsymbol{\xi}, \sigma_e^2) \\ &\propto f(\mathbf{y} | \mu, \boldsymbol{\alpha}, \boldsymbol{\xi}, \sigma_e^2) f(\alpha_j | \sigma_j^2) f(\sigma_j^2) f(\mu, \boldsymbol{\alpha}_{j-}, \boldsymbol{\xi}_{j-}, \sigma_e^2) \\ &\propto (\sigma_j^2)^{-1/2} \exp\left\{-\frac{\alpha_j^2}{2\sigma_j^2}\right\} (\sigma_j^2)^{-(2+\nu_\alpha)/2} \exp\left\{\frac{\nu_\alpha \mathbf{S}_\alpha^2}{2\sigma_j^2}\right\} \\ &\propto (\sigma_j^2)^{-(2+\nu_\alpha+1)/2} \exp\left\{\frac{\alpha_j^2 + \nu_\alpha \mathbf{S}_\alpha^2}{2\sigma_j^2}\right\} \end{aligned}$$

## Full conditional for $\sigma_\alpha^2$

So,

$$(\sigma_\alpha^2 | \mathbf{y}, \mu, \boldsymbol{\alpha}, \boldsymbol{\xi}_-, \sigma_\epsilon^2) \sim \tilde{\nu}_\alpha \tilde{\mathbf{S}}_\alpha^2 \chi_{\tilde{\nu}_\alpha}^{-2}$$

where

$$\tilde{\nu}_\alpha = \nu_\alpha + 1$$

and

$$\tilde{\mathbf{S}}_\alpha^2 = \frac{\alpha_j^2 + \nu_\alpha \mathbf{S}_\alpha^2}{\tilde{\nu}_\alpha}$$

# Multivariate- $t$

Prior:

$$(\alpha_j | \sigma_\alpha^2) \sim (\text{iid}) \mathbf{N}(\mathbf{0}, \sigma_\alpha^2)$$

$$\sigma_\alpha^2 \sim \nu_\alpha \mathbf{S}_{\nu_\alpha}^2 \chi_{\nu_\alpha}^{-2}$$

Can show that the unconditional distribution for  $\alpha$  is

$$\alpha \sim \text{multivariate-}t(\mathbf{0}, \mathbf{I} \mathbf{S}_{\nu_\alpha}^2, \nu_\alpha)$$

(Sorensen and Gianola, 2002, LBMMQG page 60)

We will see later that this is Bayes-C with  $\pi = 0$ .

## Full conditional for $\sigma_\alpha^2$

We will see later that

$$(\sigma_\alpha^2 | \mathbf{y}, \mu, \boldsymbol{\alpha}, \sigma_e^2) \sim \tilde{\nu}_\alpha \tilde{\mathbf{S}}_\alpha^2 \chi_{\nu_\alpha}^{-2}$$

where

$$\tilde{\nu}_\alpha = \nu_\alpha + k$$

and

$$\tilde{\mathbf{S}}_\alpha^2 = \frac{\boldsymbol{\alpha}'\boldsymbol{\alpha} + \nu_\alpha \mathbf{S}_\alpha^2}{\tilde{\nu}_\alpha}$$

## Spike and univariate- $t$

Prior:

$$(\alpha_j | \pi, \sigma_j^2) \begin{cases} \sim \text{N}(0, \sigma_j^2) & \text{probability } (1 - \pi), \\ = 0 & \text{probability } \pi \end{cases}$$

and

$$(\sigma_j^2 | \nu_\alpha, \mathbf{S}_\alpha^2) \sim \nu_\alpha \mathbf{S}_\alpha^2 \chi_{\nu_\alpha}^{-2}$$

Thus,

$$(\alpha_j | \pi) \text{(iid)} \begin{cases} \sim \text{univariate-}t(0, \mathbf{S}_\alpha^2, \nu_\alpha) & \text{probability } (1 - \pi), \\ = 0 & \text{probability } \pi \end{cases}$$

This is Bayes-B (Meuwissen et al., 2001; Genetics 157:1819-1829)

## Notation for sampling from mixture

The indicator variable  $\delta_j$  is defined as

$$\delta_j = 1 \Rightarrow (\alpha_j | \sigma_j^2) \sim \mathbf{N}(\mathbf{0}, \sigma_j^2)$$

and

$$\delta_j = 0 \Rightarrow (\alpha_j | \sigma_j^2) = \mathbf{0}$$

## Sampling strategy in MHG (2001)

- ▶ Sampling  $\sigma_e^2$  and  $\mu$  are as under the Normal prior.
- ▶ MHG proposed to use a Metropolis-Hastings sampler to draw samples for  $\sigma_j^2$  and  $\alpha_j$  jointly from their full-conditional distribution.
- ▶ First,  $\sigma_j^2$  is sampled from

$$f(\sigma_j^2 | \mathbf{y}, \mu, \alpha_{j-}, \xi_-, \sigma_e^2)$$

- ▶ Then,  $\alpha_j$  is sampled from its full-conditional, which is identical to that under the Normal prior

## Sampling $\sigma_j^2$

The prior for  $\sigma_j^2$  is used as the proposal. In this case, the MH acceptance probability becomes

$$\alpha = \min\left(1, \frac{f(\mathbf{y}|\sigma_{can}^2, \boldsymbol{\theta}_{j\_})}{f(\mathbf{y}|\sigma_j^2, \boldsymbol{\theta}_{j\_})}\right)$$

where  $\sigma_{can}^2$  is used to denote the candidate value for  $\sigma_j^2$ , and  $\boldsymbol{\theta}_{j\_}$  all the other parameters. It can be shown that,  $\alpha_j$  depends on  $\mathbf{y}$  only through  $r_j = \mathbf{x}'_j \mathbf{w}$  (look here). Thus

$$f(\mathbf{y}|\sigma_j^2, \boldsymbol{\theta}_{j\_}) \propto f(r_j|\sigma_j^2, \boldsymbol{\theta}_{j\_})$$

## "Likelihood" for $\sigma_j^2$

Recall that

$$\mathbf{w} = \mathbf{y} - \mathbf{1}\mu - \sum_{j' \neq j} \mathbf{x}_{j'} \alpha_{j'} = \mathbf{x}_j \alpha_j + \mathbf{e}$$

Then,

$$E(\mathbf{w} | \sigma_j^2, \boldsymbol{\theta}_{j-}) = \mathbf{0}$$

When  $\delta = 1$ :

$$\text{Var}(\mathbf{w} | \delta_j = 1, \sigma_j^2, \boldsymbol{\theta}_{j-}) = \mathbf{x}_j \mathbf{x}_j' \sigma_j^2 + \mathbf{I} \sigma_e^2$$

and  $\delta = 0$ :

$$\text{Var}(\mathbf{w} | \delta_j = 0, \sigma_j^2, \boldsymbol{\theta}_{j-}) = \mathbf{I} \sigma_e^2$$

## "Likelihood" for $\sigma_j^2$

So,

$$E(r_j | \sigma_j^2, \boldsymbol{\theta}_{j-}) = 0$$

and

$$\text{Var}(r_j | \delta_j = 1, \sigma_j^2, \boldsymbol{\theta}_{j-}) = (\mathbf{x}'_j \mathbf{x}_j)^2 \sigma_j^2 + \mathbf{x}'_j \mathbf{x}_j \sigma_e^2 = v_1$$

$$\text{Var}(r_j | \delta_j = 0, \sigma_j^2, \boldsymbol{\theta}_{j-}) = \mathbf{x}'_j \mathbf{x}_j \sigma_e^2 = v_0$$

So,

$$f(r_j | \delta_j, \sigma_j^2, \boldsymbol{\theta}_{j-}) \propto (v_\delta)^{-1/2} \exp\left\{-\frac{r_j^2}{2v_\delta}\right\}$$

## MH acceptance probability when prior is used as proposal

Suppose we want to sample  $\theta$  from  $f(\theta|\mathbf{y})$  using the MH with its prior as proposal. Then, the MH acceptance probability becomes:

$$\alpha = \min\left(1, \frac{f(\theta_{can}|\mathbf{y})f(\theta^{t-1})}{f(\theta^{t-1}|\mathbf{y})f(\theta_{can})}\right)$$

where  $f(\theta)$  is the prior for  $\theta$ . Using Bayes' theorem, the target density can be written as:

$$f(\theta|\mathbf{y}) = f(\mathbf{y}|\theta)f(\theta)$$

Then, the acceptance probability becomes

$$\alpha = \min\left(1, \frac{f(\mathbf{y}|\theta_{can})f(\theta_{can})f(\theta^{t-1})}{f(\mathbf{y}|\theta^{t-1})f(\theta^{t-1})f(\theta_{can})}\right)$$

## Alternative algorithm for spike and univariate-t

Rather than use the prior as the proposal for sampling  $\sigma_j^2$ , we

- ▶ sample  $\delta_j = 1$  with probability 0.5
- ▶ when  $\delta = 1$ , sample  $\sigma_j^2$  from a scaled inverse chi-squared distribution with
  - ▶ scale parameter =  $\sigma_j^{2(t-1)}/2$  and 4 degrees of freedom when  $\sigma_j^{2(t-1)} > 0$ , and
  - ▶ scale parameter =  $S_\alpha^2$  and 4 degrees of freedom when  $\sigma_j^{2(t-1)} = 0$

# Multivariate- $t$ mixture

Prior:

$$(\alpha_j | \pi, \sigma_\alpha^2) \begin{cases} \sim \mathbf{N}(\mathbf{0}, \sigma_\alpha^2) & \text{probability } (1 - \pi), \\ = 0 & \text{probability } \pi \end{cases}$$

and

$$(\sigma_\alpha^2 | \nu_\alpha, \mathbf{S}_\alpha^2) \sim \nu_\alpha \mathbf{S}_\alpha^2 \chi_{\nu_\alpha}^{-2}$$

Further,

$$\pi \sim \text{Uniform}(0, 1)$$

- ▶ The  $\alpha_j$  variables with their corresponding  $\delta_j = 1$  will follow a multivariate- $t$  distribution.
- ▶ This is what we have called Bayes-C $\pi$

## Full conditionals for single-site Gibbs

Full-conditional distributions for  $\mu$ ,  $\alpha$ , and  $\sigma_e^2$  are as with the Normal prior.

Full-conditional for  $\delta_j$ :

$$\Pr(\delta_j | \mathbf{y}, \mu, \alpha_{-j}, \boldsymbol{\delta}_{-j}, \sigma_\alpha^2, \sigma_e^2, \pi) = \Pr(\delta_j | r_j, \boldsymbol{\theta}_{j-})$$

$$\begin{aligned} \Pr(\delta_j | r_j, \boldsymbol{\theta}_{j-}) &= \frac{f(\delta_j, r_j | \boldsymbol{\theta}_{j-})}{f(r_j | \boldsymbol{\theta}_{j-})} \\ &= \frac{f(r_j | \delta_j, \boldsymbol{\theta}_{j-}) \Pr(\delta_j | \pi)}{f(r_j | \delta_j = 0, \boldsymbol{\theta}_{j-}) \pi + f(r_j | \delta_j = 1, \boldsymbol{\theta}_{j-}) (1 - \pi)} \end{aligned}$$

## Full conditional for $\sigma_\alpha^2$

This can be written as

$$f(\sigma_\alpha^2 | \mathbf{y}, \mu, \boldsymbol{\alpha}, \boldsymbol{\delta}, \sigma_e^2) \propto f(\mathbf{y} | \sigma_\alpha^2, \mu, \boldsymbol{\alpha}, \boldsymbol{\delta}, \sigma_e^2) f(\sigma_\alpha^2, \mu, \boldsymbol{\alpha}, \boldsymbol{\delta}, \sigma_e^2)$$

But, can see that

$$f(\mathbf{y} | \sigma_\alpha^2, \mu, \boldsymbol{\alpha}, \boldsymbol{\delta}, \sigma_e^2) \propto f(\mathbf{y} | \mu, \boldsymbol{\alpha}, \boldsymbol{\delta}, \sigma_e^2)$$

So,

$$f(\sigma_\alpha^2 | \mathbf{y}, \mu, \boldsymbol{\alpha}, \boldsymbol{\delta}, \sigma_e^2) \propto f(\sigma_\alpha^2, \mu, \boldsymbol{\alpha}, \boldsymbol{\delta}, \sigma_e^2)$$

Note that  $\sigma_\alpha^2$  appears only in  $f(\boldsymbol{\alpha} | \sigma_\alpha^2)$  and  $f(\sigma_\alpha^2)$ :

$$f(\boldsymbol{\alpha} | \sigma_\alpha^2) \propto (\sigma_\alpha^2)^{-k/2} \exp\left\{-\frac{\boldsymbol{\alpha}'\boldsymbol{\alpha}}{2\sigma_\alpha^2}\right\}$$

and

$$f(\sigma_\alpha^2) \propto (\sigma_\alpha^2)^{-(\nu_\alpha+2)/2} \exp\left\{\frac{\nu_\alpha S_\alpha^2}{2\sigma_\alpha^2}\right\}$$

## Full conditional for $\sigma_\alpha^2$

Combining these two densities gives:

$$f(\sigma_\alpha^2 | \mathbf{y}, \mu, \boldsymbol{\alpha}, \boldsymbol{\delta}, \sigma_e^2) \propto (\sigma_\alpha^2)^{-(k+\nu_\alpha+2)/2} \exp\left\{-\frac{\boldsymbol{\alpha}'\boldsymbol{\alpha} + \nu_\alpha \mathbf{S}_\alpha^2}{2\sigma_\alpha^2}\right\}$$

So,

$$(\sigma_\alpha^2 | \mathbf{y}, \mu, \boldsymbol{\alpha}, \boldsymbol{\delta}, \sigma_e^2) \sim \tilde{\nu}_\alpha \tilde{\mathbf{S}}_\alpha^2 \chi_{\tilde{\nu}_\alpha}^{-2}$$

where

$$\tilde{\nu}_\alpha = k + \nu_\alpha$$

and

$$\tilde{\mathbf{S}}_\alpha^2 = \frac{\boldsymbol{\alpha}'\boldsymbol{\alpha} + \nu_\alpha \mathbf{S}_\alpha^2}{\tilde{\nu}_\alpha}$$

## Hyper parameter: $S_\alpha^2$

If  $\sigma^2$  is distributed as a scaled, inverse chi-square random variable with scale parameter  $S^2$  and degrees of freedom  $\nu$

$$E(\sigma^2) = \frac{\nu S^2}{\nu - 2}$$

Recall that under some assumptions

$$\sigma_\alpha^2 = \frac{V_a}{\sum_j 2p_j q_j}$$

So, we take

$$S_\alpha^2 = \frac{(\nu_\alpha - 2) V_a}{\nu_\alpha k(1 - \pi) 2p\bar{q}}$$

## Full conditional for $\pi$

Using Bayes' theorem,

$$f(\pi|\boldsymbol{\delta}, \mu, \boldsymbol{\alpha}, \sigma_{\alpha}^2, \sigma_{\epsilon}^2, \mathbf{y}) \propto f(\mathbf{y}|\pi, \boldsymbol{\delta}, \mu, \boldsymbol{\alpha}, \sigma_{\alpha}^2, \sigma_{\epsilon}^2)f(\pi, \boldsymbol{\delta}, \mu, \boldsymbol{\alpha}, \sigma_{\alpha}^2, \sigma_{\epsilon}^2)$$

But,

- ▶ Conditional on  $\boldsymbol{\delta}$  the likelihood is free of  $\pi$
- ▶ Further,  $\pi$  only appears in probability of the vector of bernoulli variables:  $\boldsymbol{\delta}$

Thus,

$$f(\pi|\boldsymbol{\delta}, \mu, \boldsymbol{\alpha}, \sigma_{\alpha}^2, \sigma_{\epsilon}^2, \mathbf{y}) = \pi^{(k-m)}(1 - \pi)^m$$

where  $m = \boldsymbol{\delta}'\boldsymbol{\delta}$ , and  $k$  is the number of markers. Thus,  $\pi$  is sampled from a beta distribution with  $a = k - m + 1$  and  $b = m + 1$ .

# Simulation I

- ▶ 2000 unlinked loci in LE
- ▶ 10 of these are QTL:  $\pi = 0.995$
- ▶  $h^2 = 0.5$
- ▶ Locus effects estimated from 250 individuals

## Results for Bayes-B

Correlations between true and predicted additive genotypic values estimated from 32 replications

$\pi$	$S^2$	Correlation
0.995	0.2	0.91 (0.009)
0.8	0.2	0.86 (0.009)
0.0	0.2	0.80 (0.013)
0.995	2.0	0.90 (0.007)
0.8	2.0	0.77 (0.009)
0.0	2.0	0.35 (0.022)

## Simulation II

- ▶ 2000 unlinked loci with  $Q$  loci having effect on trait
- ▶  $N$  is the size of training data set
- ▶ Heritability = 0.5
- ▶ Validation in an independent data set with 1000 individuals
- ▶ Bayes-B and Bayes-C $\pi$  with  $\pi = 0.5$

# Results

Results from 15 replications

$N$	$Q$	$\pi$	$\hat{\pi}$	$\text{Corr}(g, \hat{g})$	
				Bayes-C $\pi$	Bayes-B
2000	10	0.995	0.994	0.995	0.937
2000	200	0.90	0.899	0.866	0.834
2000	1900	0.05	0.202	0.613	0.571
4000	1900	0.05	0.096	0.763	0.722

## Simulation II

- ▶ Genotypes: 50k SNPs from 1086 Purebred Angus animals, ISU
- ▶ Phenotypes:
  - ▶ QTL simulated from 50 randomly sampled SNPs
  - ▶ substitution effect sampled from  $N(0, \sigma_\alpha^2)$
  - ▶  $\sigma_\alpha^2 = \frac{\sigma_g^2}{502\bar{p}\bar{q}}$
  - ▶  $h^2 = 0.25$
- ▶ QTL were included in the marker panel
- ▶ Marker effects were estimated for 50k SNPs

# Validation

- ▶ Genotypes: 50k SNPs from 984 crossbred animals, CMP
- ▶ Additive genetic merit ( $g_i$ ) computed from the 50 QTL
- ▶ Additive genetic merit predicted ( $\hat{g}_i$ ) using estimated effects for 50k SNP panel

# Results

Correlations between  $g_i$  and  $\hat{g}_i$  estimated from 3 replications

$\pi$	Correlation	
	Bayes-B	Bayes-C
0.999	0.86	0.86
0.25	0.70	0.26

BayesC $\pi$ :

- ▶  $\hat{\pi} = 0.999$
- ▶ Correlation = 0.86