Bayesian Methods in Genome Association Studies

Rohan L. Fernando

Iowa State University

February, 2010
Outline of Part I

Fundamentals

Bayesian Inference
  Theory
  Computing Posteriors
Outline of Part II

Bayesian Regression Models
- Normal
- Student-$t$
- Mixture Models

Simulations
Part I

Bayesian Inference: Theory
Bayes Theorem

The conditional probability of $X$ given $Y$ is

$$\Pr(X \mid Y) = \frac{\Pr(X, Y)}{\Pr(Y)} = \frac{\Pr(Y \mid X) \Pr(X)}{\Pr(Y)}$$

where $\Pr(X, Y)$ is the joint probability of $X$ and $Y$, $\Pr(X)$ is the probability of $X$, and $\Pr(Y)$ is the probability of $Y$. 
Conditional Probability by Example

Joint distribution of smoking and lung cancer in a hypothetical population of 1,000,000:

<table>
<thead>
<tr>
<th>Smoking</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>42,500</td>
<td>7,500</td>
</tr>
<tr>
<td>No</td>
<td>207,500</td>
<td>742,500</td>
</tr>
<tr>
<td>Total</td>
<td>250,000</td>
<td>750,000</td>
</tr>
</tbody>
</table>

Question: What is the relative frequency of lung cancer among smokers?

Answer: \( \frac{42,500}{250,000} = 0.17 \)
Conditional Probability by Example

Joint distribution of smoking and lung cancer in a hypothetical population of 1,000,000:

<table>
<thead>
<tr>
<th>Smoking</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>42,500</td>
<td>7,500</td>
</tr>
<tr>
<td>No</td>
<td>207,500</td>
<td>742,500</td>
</tr>
<tr>
<td></td>
<td>250,000</td>
<td>750,000</td>
</tr>
</tbody>
</table>

Question: What is the relative frequency of lung cancer among smokers?

Answer: \[ \frac{42,500}{250,000} = 0.17 \]
Conditional Probability by Example

Joint distribution of smoking and lung cancer in a hypothetical population of 1,000,000:

<table>
<thead>
<tr>
<th>Lung Cancer</th>
<th>Smoking</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
<td>42,500</td>
<td>7,500</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>207,500</td>
<td>742,500</td>
</tr>
</tbody>
</table>

Question: What is the relative frequency of lung cancer among smokers?

Answer: \( \frac{42,500}{250,000} = 0.17 \)
Conditional Probability by Example

- As explained below, this relative frequency is also the conditional probability of lung cancer given smoking.
  - The frequentist definition of probability of an event is the limiting value of its relative frequency in a large number of trials.
  - Suppose we sample with replacement individuals from the 250,000 smokers and compute the relative frequency of lung cancer incidence.
  - It can be shown that as the sample size goes to infinity, this relative frequency will approach \( \frac{42,500}{250,000} = 0.17 \).

- This conditional probability is usually written as \( \frac{42,500}{1,000,000} / \frac{250,000}{1,000,000} = 0.17 \).

- The ratio in the numerator is joint probability of smoking and lung cancer, and the ratio in the denominator is the marginal probability of smoking.
As explained below, this relative frequency is also the conditional probability of lung cancer given smoking.

The frequentist definition of probability of an event is the limiting value of its relative frequency in a large number of trials.

Suppose we sample with replacement individuals from the 250,000 smokers and compute the relative frequency of lung cancer incidence.

It can be shown that as the sample size goes to infinity, this relative frequency will approach \( \frac{42,500}{250,000} = 0.17 \).

This conditional probability is usually written as \( \frac{42,500/1,000,000}{250,000/1,000,000} = 0.17 \).

The ratio in the numerator is joint probability of smoking and lung cancer, and the ratio in the denominator is the marginal probability of smoking.
Conditional Probability by Example

- As explained below, this relative frequency is also the conditional probability of lung cancer given smoking.
  - The frequentist definition of probability of an event is the limiting value of its relative frequency in a large number of trials.
  - Suppose we sample with replacement individuals from the 250,000 smokers and compute the relative frequency of lung cancer incidence.
    - It can be shown that as the sample size goes to infinity, this relative frequency will approach $\frac{42,500}{250,000} = 0.17$.
  - This conditional probability is usually written as $\frac{42,500}{1,000,000} = 0.17$.
  - The ratio in the numerator is joint probability of smoking and lung cancer, and the ratio in the denominator is the marginal probability of smoking.
As explained below, this relative frequency is also the conditional probability of lung cancer given smoking.

The frequentist definition of probability of an event is the limiting value of its relative frequency in a large number of trials.

Suppose we sample with replacement individuals from the 250,000 smokers and compute the relative frequency of lung cancer incidence.

It can be shown that as the sample size goes to infinity, this relative frequency will approach \( \frac{42,500}{250,000} = 0.17 \).

This conditional probability is usually written as \( \frac{42,500}{1,000,000} = 0.17 \).

The ratio in the numerator is joint probability of smoking and lung cancer, and the ratio in the denominator is the marginal probability of smoking.
Conditional Probability by Example

As explained below, this relative frequency is also the conditional probability of lung cancer given smoking.

- The frequentist definition of probability of an event is the limiting value of its relative frequency in a large number of trials.
- Suppose we sample with replacement individuals from the 250,000 smokers and compute the relative frequency of lung cancer incidence.
- It can be shown that as the sample size goes to infinity, this relative frequency will approach \( \frac{42,500}{250,000} = 0.17 \).

This conditional probability is usually written as 
\[ \frac{42,500}{250,000} / \frac{1,000,000}{1,000,000} = 0.17. \]

The ratio in the numerator is joint probability of smoking and lung cancer, and the ratio in the denominator is the marginal probability of smoking.
As explained below, this relative frequency is also the conditional probability of lung cancer given smoking.

The frequentist definition of probability of an event is the limiting value of its relative frequency in a large number of trials.

Suppose we sample with replacement individuals from the 250,000 smokers and compute the relative frequency of lung cancer incidence.

It can be shown that as the sample size goes to infinity, this relative frequency will approach \( \frac{42,500}{250,000} = 0.17 \).

This conditional probability is usually written as \( \frac{42,500}{1,000,000} \div \frac{250,000}{1,000,000} = 0.17 \).

The ratio in the numerator is joint probability of smoking and lung cancer, and the ratio in the denominator is the marginal probability of smoking.
Meaning of Probability in Bayesian Inference

- In the frequency approach, probability is a limiting frequency.
- 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?
Meaning of Probability in Bayesian Inference

- In the frequency approach, probability is a limiting frequency.
- 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?
Meaning of Probability in Bayesian Inference

- In the frequency approach, probability is a limiting frequency.
- 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?
Meaning of Probability in Bayesian Inference

- In the frequency approach, probability is a limiting frequency.
- 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?
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.
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.
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 probabilities.
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 in Bayesian Inference

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

\[
f(\theta|y) = \frac{f(y|\theta)f(\theta)}{f(y)} = \propto f(y|\theta)f(\theta)
\]
Let $f(\theta)$ denote the prior probability density for $\theta$.

Let $f(y|\theta)$ denote the likelihood.

Then, the posterior probability of $\theta$ is:

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

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

$\propto f(y|\theta)f(\theta)$
Computing posteriors

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

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

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

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

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

\[
\begin{align*}
  x_1^t & \text{ from } f(x_1 | x_2^{t-1}, x_3^{t-1}, \ldots, x_n^{t-1}) \\
  x_2^t & \text{ from } f(x_2 | x_1^t, x_3^{t-1}, \ldots, x_n^{t-1}) \\
  x_3^t & \text{ from } f(x_3 | x_1^t, x_2^t, \ldots, x_n^{t-1}) \\
  & \vdots \\
  x_n^t & \text{ from } f(x_n | x_1^t, x_2^t, \ldots, x_{n-1}^t)
\end{align*}
\]

- The sequence \( x^1, x^2, \ldots, x^n \) is a Markov chain with stationary distribution \( f(x_1, x_2, \ldots, x_n) \)
Gibbs sampler

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

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

- The sequence $x^1, x^2, \ldots, x^n$ is a Markov chain with stationary distribution $f(x_1, x_2, \ldots, x_n)$
Gibbs sampler

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

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

- The sequence $x^1, x^2, \ldots, x^n$ is a Markov chain with stationary distribution $f(x_1, x_2, \ldots, x_n)$
Gibbs sampler

▶ Want to draw samples from $f(x_1, x_2, \ldots, x_n)$

▶ Even though it may be possible to compute $f(x_1, x_2, \ldots, x_n)$, it is difficult to draw samples directly from $f(x_1, x_2, \ldots, x_n)$

▶ Gibbs:
  ▶ Get valid a starting point $x^0$
  ▶ Draw sample $x^t$ as:

\[
\begin{align*}
  x_1^t & \text{ from } f(x_1 | x_2^{t-1}, x_3^{t-1}, \ldots, x_n^{t-1}) \\
  x_2^t & \text{ from } f(x_2 | x_1^t, x_3^{t-1}, \ldots, x_n^{t-1}) \\
  x_3^t & \text{ from } f(x_3 | x_1^t, x_2^t, \ldots, x_n^{t-1}) \\
  \vdots & \text{ } \vdots \\
  x_n^t & \text{ from } f(x_n | x_1^t, x_2^t, \ldots, x_{n-1}^t)
\end{align*}
\]

▶ The sequence $x^1, x^2, \ldots, x^n$ is a Markov chain with stationary distribution $f(x_1, x_2, \ldots, x_n)$
Gibbs sampler

- Want to draw samples from $f(x_1, x_2, \ldots, x_n)$
- Even though it may be possible to compute $f(x_1, x_2, \ldots, x_n)$, it is difficult to draw samples directly from $f(x_1, x_2, \ldots, x_n)$

Gibbs:
- Get valid a starting point $x^0$
- Draw sample $x^t$ as:

$$
\begin{align*}
  x^t_1 & \text{ from } f(x_1 | x^{t-1}_2, x^{t-1}_3, \ldots, x^{t-1}_n) \\
  x^t_2 & \text{ from } f(x_2 | x^t_1, x^{t-1}_3, \ldots, x^{t-1}_n) \\
  x^t_3 & \text{ from } f(x_3 | x^t_1, x^t_2, \ldots, x^{t-1}_n) \\
  \vdots & \\
  x^t_n & \text{ from } f(x_n | x^t_1, x^t_2, \ldots, x^t_{n-1})
\end{align*}
$$

- The sequence $x^1, x^2, \ldots, x^n$ is a Markov chain with stationary distribution $f(x_1, x_2, \ldots, x_n)$
Gibbs sampler

- Want to draw samples from $f(x_1, x_2, \ldots, x_n)$
- Even though it may be possible to compute $f(x_1, x_2, \ldots, x_n)$, it is difficult to draw samples directly from $f(x_1, x_2, \ldots, x_n)$
- Gibbs:
  - Get valid a starting point $x^0$
  - Draw sample $x^t$ as:
    \[
    \begin{align*}
    x_1^t & \text{ from } f(x_1 | x_2^{t-1}, x_3^{t-1}, \ldots, x_n^{t-1}) \\
    x_2^t & \text{ from } f(x_2 | x_1^t, x_3^{t-1}, \ldots, x_n^{t-1}) \\
    x_3^t & \text{ from } f(x_3 | x_1^t, x_2^t, \ldots, x_n^{t-1}) \\
    \vdots & \quad \vdots \\
    x_n^t & \text{ from } f(x_n | x_1^t, x_2^t, \ldots, x_{n-1}^t)
    \end{align*}
    \]
- The sequence $x^1, x^2, \ldots, x^n$ is a Markov chain with stationary distribution $f(x_1, x_2, \ldots, x_n)$
Gibbs sampler

- Want to draw samples from $f(x_1, x_2, \ldots, x_n)$
- Even though it may be possible to compute $f(x_1, x_2, \ldots, x_n)$, it is difficult to draw samples directly from $f(x_1, x_2, \ldots, x_n)$

Gibbs:

- Get valid a starting point $x^0$
- Draw sample $x^t$ as:

$$
\begin{align*}
  x_1^t & \text{ from } f(x_1 | x_2^{t-1}, x_3^{t-1}, \ldots, x_n^{t-1}) \\
  x_2^t & \text{ from } f(x_2 | x_1^t, x_3^{t-1}, \ldots, x_n^{t-1}) \\
  x_3^t & \text{ from } f(x_3 | x_1^t, x_2^t, \ldots, x_n^{t-1}) \\
  \vdots & \quad \vdots \\
  x_n^t & \text{ from } f(x_n | x_1^t, x_2^t, \ldots, x_{n-1}^t)
\end{align*}
$$

- The sequence $x^1, x^2, \ldots, x^n$ is a Markov chain with stationary distribution $f(x_1, x_2, \ldots, x_n)$
Gibbs sampler

- Want to draw samples from $f(x_1, x_2, \ldots, x_n)$
- Even though it may be possible to compute $f(x_1, x_2, \ldots, x_n)$, it is difficult to draw samples directly from $f(x_1, x_2, \ldots, x_n)$
- Gibbs:
  - Get valid a starting point $x^0$
  - Draw sample $x^t$ as:
    
    \[
    \begin{align*}
    x_1^t & \quad \text{from} \quad f(x_1 | x_2^{t-1}, x_3^{t-1}, \ldots, x_n^{t-1}) \\
    x_2^t & \quad \text{from} \quad f(x_2 | x_1^t, x_3^{t-1}, \ldots, x_n^{t-1}) \\
    x_3^t & \quad \text{from} \quad f(x_3 | x_1^t, x_2^t, \ldots, x_n^{t-1}) \\
    \vdots & \\
    x_n^t & \quad \text{from} \quad f(x_n | x_1^t, x_2^t, \ldots, x_{n-1}^t)
    \end{align*}
    \]

- The sequence $x^1, x^2, \ldots, x^n$ is a Markov chain with stationary distribution $f(x_1, x_2, \ldots, x_n)$
Gibbs sampler

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

$$
x_1^t \text{ from } f(x_1 | x_2^{t-1}, x_3^{t-1}, \ldots, x_n^{t-1})
$$

$$
x_2^t \text{ from } f(x_2 | x_1^t, x_3^{t-1}, \ldots, x_n^{t-1})
$$

$$
x_3^t \text{ from } f(x_3 | x_1^t, x_2^t, \ldots, x_n^{t-1})
$$

$$
\vdots
$$

$$
x_n^t \text{ from } f(x_n | x_1^t, x_2^t, \ldots, x_{n-1}^t)
$$

- The sequence $x^1, x^2, \ldots, x^n$ is a Markov chain with stationary distribution $f(x_1, x_2, \ldots, x_n)$
Gibbs sampler

- Want to draw samples from \( f(x_1, x_2, \ldots, x_n) \)
- Even though it may be possible to compute \( f(x_1, x_2, \ldots, x_n) \), it is difficult to draw samples directly from it.
- Gibbs:
  - Get a valid starting point \( x^0 \)
  - Draw sample \( x^t \) as:
    \[
    \begin{align*}
    x^t_1 & \quad \text{from} \quad f(x_1 | x^t_2, x^t_3, \ldots, x^t_n) \\
    x^t_2 & \quad \text{from} \quad f(x_2 | x^t_1, x^t_3, \ldots, x^t_n) \\
    x^t_3 & \quad \text{from} \quad f(x_3 | x^t_1, x^t_2, \ldots, x^t_n) \\
    & \quad \vdots \\
    x^t_n & \quad \text{from} \quad f(x_n | x^t_1, x^t_2, \ldots, x^t_{n-1})
    \end{align*}
    \]
- The sequence \( x^1, x^2, \ldots, x^n \) is a Markov chain with stationary distribution \( f(x_1, x_2, \ldots, x_n) \)
Can show that samples obtained from the Markov chain can be used to draw inferences from $f(x_1, x_2, \ldots, x_n)$ 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

Inference from Markov chain

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

Inference from Markov chain

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

Inference from Markov chain

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

Example

Let $f(x)$ be a bivariate normal density with means

$$\mu' = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

and covariance matrix

$$V = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 2.0 \end{bmatrix}$$

Suppose we do not know how to draw samples from $f(x)$, but know how to draw samples from $f(x_i|x_j)$, which is univariate normal with mean:

$$\mu_{i,j} = \mu_i + \frac{V_{ij}}{V_{jj}} (x_j - \mu_j)$$

and variance

$$V_{i,j} = V_{ii} - \frac{V_{ij}^2}{V_{jj}}$$
Gibbs sampler

- Gibbs:
  - Start with $x^0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$
  - Draw sample $x^t$ as:
    
    \[
    \begin{align*}
    x_1^t & \quad \text{from} \quad f(x_1 | x_2^{t-1}) \\
    x_2^t & \quad \text{from} \quad f(x_2 | x_1^t)
    \end{align*}
    \]

- Use the sequence $x^1, x^2, \ldots, x^n$ to compute any property of $f(x)$, for example

\[
\Pr(x_1 > \mu_1 \text{ and } x_2 > \mu_2)
\]
Gibbs sampler

- Gibbs:
  - Start with \( x^0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \)
  - Draw sample \( x^t \) as:
    - \( x_1^t \) from \( f(x_1 | x_{2}^{t-1}) \)
    - \( x_2^t \) from \( f(x_2 | x_1^t) \)

- Use the sequence \( x^1, x^2, \ldots, x^n \) to compute any property of \( f(x) \), for example

\[
\Pr(x_1 > \mu_1 \text{ and } x_2 > \mu_2)
\]
Gibbs sampler

- Gibbs:
  - Start with $\mathbf{x}^0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$
  - Draw sample $\mathbf{x}^t$ as:
    
    $x_1^t$ from $f(x_1|x_2^{t-1})$
    $x_2^t$ from $f(x_2|x_1^t)$

- Use the sequence $\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^n$ to compute any property of $f(\mathbf{x})$, for example

  $\Pr(x_1 > \mu_1 \text{ and } x_2 > \mu_2)$
Gibbs sampler

- Gibbs:
  - Start with $\mathbf{x}^0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$
  - Draw sample $\mathbf{x}^t$ as:
    - $x^t_1$ from $f(x_1 | x^t_{2-1})$
    - $x^t_2$ from $f(x_2 | x^t_1)$

- Use the sequence $\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^n$ to compute any property of $f(x)$, for example

$$\Pr(x_1 > \mu_1 \text{ and } x_2 > \mu_2)$$
Gibbs sampler

- **Gibbs:**
  - Start with \( \mathbf{x}^0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \)
  - Draw sample \( \mathbf{x}^t \) as:
    
    \[
    \begin{align*}
    x_{1}^{t} & \quad \text{from} \quad f(x_1 | x_{2}^{t-1}) \\
    x_{2}^{t} & \quad \text{from} \quad f(x_2 | x_{1}^{t})
    \end{align*}
    \]

- Use the sequence \( \mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^n \) to compute any property of \( f(\mathbf{x}) \), for example
  
  \[ \Pr(x_1 > \mu_1 \text{ and } x_2 > \mu_2) \]
MCMC Estimates of \( \Pr(x_1 > \mu_1 \text{ and } x_2 > \mu_2) \)
Metropolis-Hastings sampler

- Sometimes may not be able to draw samples directly from $f(x_i|x_{i-1})$
- 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 $q(y|x^{t-1})$
  - $x^t = \begin{cases} y & \text{with probability } \alpha \\ x^{t-1} & \text{with probability } 1 - \alpha \end{cases}$
  - $\alpha = \min(1, \frac{f(y)q(x^{t-1}|y)}{f(x^{t-1})q(y|x^{t-1})})$
- The samples from MH is a Markov chain with stationary distribution $f(x)$
Metropolis-Hastings sampler

- Sometimes may not be able to draw samples directly from $f(x_i | x_{i-1})$
- 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 $q(y | x^{t-1})$
  - $x^t = \begin{cases} y & \text{with probability } \alpha \\ x^{t-1} & \text{with probability } 1 - \alpha \end{cases}$
  - $\alpha = \min(1, \frac{f(y)q(x^{t-1} | y)}{f(x^{t-1})q(y | x^{t-1})})$
- The samples from MH is a Markov chain with stationary distribution $f(x)$
Metropolis-Hastings sampler

- Sometimes may not be able to draw samples directly from $f(x_i|x_{i-1})$
- 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 $q(y|x^{t-1})$
  - $x^t = \begin{cases} y & \text{with probability } \alpha \\ x^{t-1} & \text{with probability } 1 - \alpha \end{cases}$
  - $\alpha = \min(1, \frac{f(y)q(x^{t-1}|y)}{f(x^{t-1})q(y|x^{t-1})})$
- The samples from MH is a Markov chain with stationary distribution $f(x)$
Metropolis-Hastings sampler

- Sometimes may not be able to draw samples directly from $f(x_i|x_{i-1})$
- 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 $q(y|x^{t-1})$
  
  $x^t = \begin{cases} 
  y & \text{with probability } \alpha \\
  x^{t-1} & \text{with probability } 1 - \alpha 
  \end{cases}$
  
  $\alpha = \min(1, \frac{f(y)q(x^{t-1}|y)}{f(x^{t-1})q(y|x^{t-1})})$

- The samples from MH is a Markov chain with stationary distribution $f(x)$
Metropolis-Hastings sampler

- Sometimes may not be able to draw samples directly from $f(x_i | x_{i-1})$
- 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 $q(y | x^{t-1})$
  - $x^t = \begin{cases} y & \text{with probability } \alpha \\ x^{t-1} & \text{with probability } 1 - \alpha \end{cases}$

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

- The samples from MH is a Markov chain with stationary distribution $f(x)$
Metropolis-Hastings sampler

- Sometimes may not be able to draw samples directly from $f(x_i|x_{i-1})$
- 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 $q(y|x_{t-1})$
  - $x^t = \begin{cases} y & \text{with probability } \alpha \\ x^{t-1} & \text{with probability } 1 - \alpha \end{cases}$
  - $\alpha = \min(1, \frac{f(y)q(x^{t-1}|y)}{f(x^{t-1})q(y|x^{t-1})})$

- The samples from MH is a Markov chain with stationary distribution $f(x)$
Metropolis-Hastings sampler

- Sometimes may not be able to draw samples directly from \( f(x_i|x_{i-1}) \)
- 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 \( q(y|x^{t-1}) \)
  - \[
  x^t = \begin{cases} 
  y & \text{with probability } \alpha \\
  x^{t-1} & \text{with probability } 1 - \alpha 
  \end{cases}
  \]
  - \[
  \alpha = \min(1, \frac{f(y)q(x^{t-1}|y)}{f(x^{t-1})q(y|x^{t-1})})
  \]
- The 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
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
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
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
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
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
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
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
MH Sampler to Estimate $\Pr(x_1 > \mu_1 \text{ and } x_2 > \mu_2)$

MH Sampler:

- Start with $x^0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$
- Draw sample $x^t$ as:
  
  $y_1 = x_{1}^{t-1} + u_1$
  $y_2 = x_{2}^{t-1} + u_2$

  where $u_i$ is Uniform($-\nu_{ii}^{1/2}, \nu_{ii}^{1/2}$).

- Compute
  
  $\alpha = \min(1, \frac{f(y)}{f(x^{t-1})})$

  and

  $x^t = \begin{cases} 
  y & \text{with probability } \alpha \\
  x^{t-1} & \text{with probability } 1 - \alpha 
  \end{cases}$
MH Sampler to Estimate $\Pr(x_1 > \mu_1 \text{ and } x_2 > \mu_2)$

MH Sampler:

- Start with $x^0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$
- Draw sample $x^t$ as:
  
  $y_1 = x_1^{t-1} + u_1$
  $y_2 = x_2^{t-1} + u_2$

where $u_i$ is Uniform($-\nu_i^{1/2}, \nu_i^{1/2}$).

- Compute
  
  $\alpha = \min(1, \frac{f(y)}{f(x^{t-1})})$

and

  $x^t = \begin{cases} 
  y & \text{with probability } \alpha \\
  x^{t-1} & \text{with probability } 1 - \alpha 
  \end{cases}$
MH Sampler to Estimate $\Pr(x_1 > \mu_1 \text{ and } x_2 > \mu_2)$

MH Sampler:

- Start with $x^0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$
- Draw sample $x^t$ as:

  $$
  y_1 = x_1^{t-1} + u_1 \\
  y_2 = x_2^{t-1} + u_2
  $$

  where $u_i$ is Uniform($-v_{ii}^{1/2}, v_{ii}^{1/2}$).
- Compute

  $$
  \alpha = \min(1, \frac{f(y)}{f(x^{t-1})})
  $$

  and

  $$
  x^t = \begin{cases} 
  y & \text{with probability } \alpha \\
  x^{t-1} & \text{with probability } 1 - \alpha
  \end{cases}
  $$
MCMC Estimates of $\Pr(x_1 > \mu_1 \text{ and } x_2 > \mu_2)$
Distribution of $y_1$ Sampled Using MH
Part II

Bayesian Inference: Application to Whole Genome Analyses
Model

Model:

\[ y_i = \mu + \sum_j X_{ij} \alpha_j + e_i \]

Priors:

- \( \mu \propto \text{constant} \) (not proper, but posterior is proper)
- \( (e_i|\sigma_e^2) \sim (\text{iid})N(0, \sigma_e^2); \quad \sigma_e^2 \sim \nu_e S_e^2 \chi_{v_e}^{-2} \)
- Consider several different priors for \( \alpha_j \)
Model

Model:

\[ y_i = \mu + \sum_j X_{ij} \alpha_j + e_i \]

Priors:

- \( \mu \propto \text{constant} \) (not proper, but posterior is proper)
- \( (e_i|\sigma^2_e) \sim \text{(iid)}N(0, \sigma^2_e); \sigma^2_e \sim \nu_e S^2_e \chi^{-2}_{\nu_e} \)
- Consider several different priors for \( \alpha_j \)
Model

Model:

\[ y_i = \mu + \sum_j X_{ij} \alpha_j + e_i \]

Priors:

- \( \mu \propto \) constant (not proper, but posterior is proper)
- \( (e_i | \sigma^2_e) \sim (\text{iid})N(0, \sigma^2_e); \sigma^2_e \sim \nu_e S^2_e \chi^2_{\nu_e} \)
- Consider several different priors for \( \alpha_j \)
Model:

\[ y_i = \mu + \sum_j X_{ij} \alpha_j + e_i \]

Priors:

- \( \mu \propto \) constant (not proper, but posterior is proper)
- \((e_i | \sigma^2_e) \sim (\text{iid})N(0, \sigma^2_e)\); \( \sigma^2_e \sim \nu_e S_e^2 \chi^{-2}_{\nu_e} \)
- Consider several different priors for \( \alpha_j \)
Prior: \((\alpha_j | \sigma^2_\alpha) \sim \text{(iid)N}(0, \sigma^2_\alpha); \ \sigma^2_\alpha \text{ is known}\)

What is \(\sigma^2_\alpha\)?

Assume the QTL genotypes are a subset of those available for the analysis

Then, the genotypic value of \(i\) can be written as:

\[
g_i = \mu + x_i'\alpha
\]

Note that \(\alpha\) is common to all \(i\)
Thus, the variance of \(g_i\) comes from \(x_i'\) being random

So, \(\sigma^2_\alpha\) 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
\[
\sigma^2_\alpha = \text{Var}(\alpha_j)
\]
Normal

- Prior: \((\alpha_j | \sigma^2_\alpha) \sim (\text{iid}) \text{N}(0, \sigma^2_\alpha)\); \(\sigma^2_\alpha\) is known
- What is \(\sigma^2_\alpha\)?
  - Assume the QTL genotypes are a subset of those available for the analysis
    - Then, the genotypic value of \(i\) can be written as:
      \[
g_i = \mu + \mathbf{x}_i' \alpha
\]
    - Note that \(\alpha\) is common to all \(i\)
    - Thus, the variance of \(g_i\) comes from \(\mathbf{x}_i'\) being random
  - So, \(\sigma^2_\alpha\) 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
    - \(\sigma^2_\alpha = \text{Var}(\alpha_j)\)
Prior: \((\alpha_j | \sigma^2_\alpha) \sim (\text{iid}) N(0, \sigma^2_\alpha)\); \(\sigma^2_\alpha\) is known

What is \(\sigma^2_\alpha\)?

Assume the QTL genotypes are a subset of those available for the analysis

Then, the genotypic value of \(i\) can be written as:

\[
g_i = \mu + x'_i \alpha
\]

Note that \(\alpha\) is common to all \(i\)

Thus, the variance of \(g_i\) comes from \(x'_i\) being random

So, \(\sigma^2_\alpha\) 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

\[
\sigma^2_\alpha = \text{Var}(\alpha_j)
\]
Prior: \((\alpha_j | \sigma^2_\alpha) \sim \text{(iid)}N(0, \sigma^2_\alpha)\); \(\sigma^2_\alpha\) is known

What is \(\sigma^2_\alpha\)?

Assume the QTL genotypes are a subset of those available for the analysis

Then, the genotypic value of \(i\) can be written as:

\[
g_i = \mu + x'_i \alpha
\]

Note that \(\alpha\) is common to all \(i\)

Thus, the variance of \(g_i\) comes from \(x'_i\) being random

So, \(\sigma^2_\alpha\) 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

\(\sigma^2_\alpha = \text{Var}(\alpha_j)\)
Normal

- Prior: \((\alpha_j | \sigma^2_\alpha) \sim \text{(iid)N}(0, \sigma^2_\alpha)\); \(\sigma^2_\alpha\) is known
- What is \(\sigma^2_\alpha\)?
- Assume the QTL genotypes are a subset of those available for the analysis
  - Then, the genotypic value of \(i\) can be written as:
    \[
    g_i = \mu + x'_i \alpha
    \]
  - Note that \(\alpha\) is common to all \(i\)
    - Thus, the variance of \(g_i\) comes from \(x'_i\) being random
  - So, \(\sigma^2_\alpha\) 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
    - \(\sigma^2_\alpha = \text{Var}(\alpha_j)\)
Prior: \((\alpha_j|\sigma^2_\alpha) \sim (\text{iid})N(0, \sigma^2_\alpha)\); \(\sigma^2_\alpha\) is known

What is \(\sigma^2_\alpha\)?

Assume the QTL genotypes are a subset of those available for the analysis

Then, the genotypic value of \(i\) can be written as:

\[ g_i = \mu + x'_i \alpha \]

Note that \(\alpha\) is common to all \(i\)

Thus, the variance of \(g_i\) comes from \(x'_i\) being random

So, \(\sigma^2_\alpha\) 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

\(\sigma^2_\alpha = \text{Var}(\alpha_j)\)
Prior: \((\alpha_j | \sigma^2_\alpha) \sim (\text{iid})N(0, \sigma^2_\alpha)\); \(\sigma^2_\alpha\) is known

What is \(\sigma^2_\alpha\)?

Assume the QTL genotypes are a subset of those available for the analysis

Then, the genotypic value of \(i\) can be written as:

\[
g_i = \mu + x'_i \alpha
\]

Note that \(\alpha\) is common to all \(i\)

Thus, the variance of \(g_i\) comes from \(x'_i\) being random

So, \(\sigma^2_\alpha\) 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

\(\sigma^2_\alpha = \text{Var}(\alpha_j)\)
Normal

- Prior: \((\alpha_j | \sigma^2_\alpha) \sim (\text{iid})\text{N}(0, \sigma^2_\alpha); \sigma^2_\alpha \text{ is known}\)
- What is \(\sigma^2_\alpha\)?
- Assume the QTL genotypes are a subset of those available for the analysis
  - Then, the genotypic value of \(i\) can be written as:
    \[
g_i = \mu + x'_i \alpha
    \]
    - Note that \(\alpha\) is common to all \(i\)
    - Thus, the variance of \(g_i\) comes from \(x'_i\) being random
- So, \(\sigma^2_\alpha\) 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
  - \(\sigma^2_\alpha = \text{Var}(\alpha_j)\)
Prior: \((\alpha_j | \sigma^2_\alpha) \sim (iid)N(0, \sigma^2_\alpha)\); \(\sigma^2_\alpha\) is known

What is \(\sigma^2_\alpha\)?

Assume the QTL genotypes are a subset of those available for the analysis

- Then, the genotypic value of \(i\) can be written as:
  \[ g_i = \mu + \mathbf{x}_i' \alpha \]

- Note that \(\alpha\) is common to all \(i\)
- Thus, the variance of \(g_i\) comes from \(\mathbf{x}_i'\) being random

So, \(\sigma^2_\alpha\) 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

\[ \sigma^2_\alpha = \text{Var}(\alpha_j) \]
Normal

- Prior: \((\alpha_j | \sigma^2_\alpha) \sim (\text{iid})N(0, \sigma^2_\alpha)\); \(\sigma^2_\alpha\) is known
- What is \(\sigma^2_\alpha\)?
- Assume the QTL genotypes are a subset of those available for the analysis
  - Then, the genotypic value of \(i\) can be written as:
    \[ g_i = \mu + x'_i \alpha \]
  - Note that \(\alpha\) is common to all \(i\)
  - Thus, the variance of \(g_i\) comes from \(x'_i\) being random
- So, \(\sigma^2_\alpha\) 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
  - \(\sigma^2_\alpha = \text{Var}(\alpha_j)\)
Relationship of $\sigma^2_\alpha$ to genetic variance

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

$$V_A = \sum_{j}^{k} 2p_jq_j\alpha_j^2,$$

where $p_j = 1 - q_j$ is gene frequency at SNP locus $j$.

Letting $U_j = 2p_jq_j$ 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 $\sigma^2_{\alpha}$ to genetic variance

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

$$V_A = \sum_{j=1}^{k} 2p_j q_j \alpha_j^2,$$

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

$$V_A = \sum_{j=1}^{k} U_j V_j$$

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 $\sigma^2_{\alpha}$ to genetic variance

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

\[ V_A = \sum_{j} 2p_j q_j \alpha_j^2, \]

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

\[ V_A = \sum_{j} U_j V_j \]

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

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

So,

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

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

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

and,

$$\sigma_\alpha^2 = \frac{V_A - kC_{UV}}{\sum_j 2p_j q_j}$$
Relationship of $\sigma_{\alpha}^2$ to genetic variance

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

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

So,

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

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

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

and,

$$\sigma_{\alpha}^2 = \frac{V_A - kC_{UV}}{\sum_j 2p_j q_j}$$
Relationship of $\sigma^2_{\alpha}$ 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,

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

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

$$
V_A = kC_{UV} + \left(\sum_j 2p_jq_j\right)\sigma^2_{\alpha}
$$

and,

$$
\sigma^2_{\alpha} = \frac{V_A - kC_{UV}}{\sum_j 2p_jq_j}
$$
Blocked Gibbs sampler

Let $\theta' = [\mu, \alpha']$

Can show that $(\theta | y, \sigma^2_e) \sim N(\hat{\theta}, C^{-1}\sigma^2_e)$

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

$$C = \begin{bmatrix}
1'1 & 1'X \\
X'1 & X'X + I\frac{\sigma^2_e}{\sigma^2_\alpha}
\end{bmatrix}$$

Blocked Gibbs sampler

- *Likelihood, Bayesian and MCMC Methods* · · · (LBMMMQG, Sorensen and Gianola, 2002)
Let $\theta' = [\mu, \alpha']$

Can show that $(\theta | y, \sigma^2_e) \sim N(\hat{\theta}, C^{-1}\sigma^2_e)$

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

\[
C = \begin{bmatrix}
1'1 & 1'X \\
X'1 & X'X + I\frac{\sigma^2_e}{\sigma^2_\alpha}
\end{bmatrix}
\]

Blocked Gibbs sampler


Likelihood, Bayesian and MCMC Methods ... (LBMMQG, Sorensen and Gianola, 2002)
Let $\theta' = [\mu, \alpha']$

Can show that $(\theta | y, \sigma_e^2) \sim N(\hat{\theta}, C^{-1}\sigma_e^2)$

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

\[ C = \begin{bmatrix} 1' 1 & 1' X \\ X' 1 & X'X + I \sigma_e^2 / \sigma_\alpha^2 \end{bmatrix} \]

Blocked Gibbs sampler


Likelihood, Bayesian and MCMC Methods · · · (LBMMQG, Sorensen and Gianola, 2002)
Blocked Gibbs sampler

Let $\theta' = [\mu, \alpha']$

Can show that $(\theta | y, \sigma_\epsilon^2) \sim N(\hat{\theta}, C^{-1}\sigma_\epsilon^2)$

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

\[
C = \begin{bmatrix}
1'1 & 1'X \\
X'1 & X'X + I \frac{\sigma_\epsilon^2}{\sigma_\alpha^2}
\end{bmatrix}
\]

Blocked Gibbs sampler

- *Likelihood, Bayesian and MCMC Methods* · · · (LBMMQG, Sorensen and Gianola, 2002)
Let $\theta' = [\mu, \alpha']$

Can show that $(\theta | y, \sigma^2_e) \sim N(\hat{\theta}, C^{-1} \sigma^2_e)$

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

\[
C = \begin{bmatrix}
1'1 & 1'X \\
X'1 & X'X + I \frac{\sigma^2_e}{\sigma^2_\alpha}
\end{bmatrix}
\]

Blocked Gibbs sampler

- *Likelihood, Bayesian and MCMC Methods* · · · (LBMMQG, Sorensen and Gianola, 2002)
Full conditionals for single-site Gibbs

\[ (\mu | y, \alpha, \sigma^2_e) \sim N\left( \frac{1'(y - X\alpha)}{n}, \frac{\sigma^2_e}{n} \right) \]

\[ (\alpha_j | y, \mu, \alpha_j, \sigma^2_e) \sim N(\hat{\alpha}_j, \frac{\sigma^2_e}{c_j}) \]

\[ \hat{\alpha}_j = \frac{x_j'w}{c_j} \]

\[ w = y - 1\mu - \sum_{j' \neq j} x_{j'}\alpha_{j'} \]

\[ c_j = (x_j'x_j + \frac{\sigma^2_e}{\sigma^2_\alpha}) \]

\[ (\sigma^2_e | y, \mu, \alpha) \sim [((y - W\theta)'(y - W\theta) + \nu_eS^2_e] \chi^2_{(\nu_e+n)} \]
Full conditionals for single-site Gibbs

\begin{itemize}
  \item \((\mu | y, \alpha, \sigma^2_e) \sim \mathcal{N}(\frac{1'}{n}(y - X\alpha), \frac{\sigma^2_e}{n})\)
  \item \((\alpha_j | y, \mu, \alpha_j, \sigma^2_e) \sim \mathcal{N}(\hat{\alpha}_j, \frac{\sigma^2_e}{c_j})\)
  \item \(\hat{\alpha}_j = \frac{x_j'w}{c_j}\)
  \item \(w = y - 1\mu - \sum_{j' \neq j} x_{j'}\alpha_{j'}\)
  \item \(c_j = (x_j'x_j + \frac{\sigma^2_e}{\sigma^2_\alpha})\)
  \item \((\sigma^2_e | y, \mu, \alpha) \sim [(y - W\theta)'(y - W\theta) + \nu eS^2_e] \chi_{(\nu e + n)}^{-2}\)
\end{itemize}
Full conditionals for single-site Gibbs

\[ (\mu|\mathbf{y}, \alpha, \sigma^2_\epsilon) \sim N\left(\frac{1'}{n}(\mathbf{y} - \mathbf{X} \alpha), \frac{\sigma^2_\epsilon}{n}\right) \]

\[ (\alpha_j|\mathbf{y}, \mu, \alpha_j, \sigma^2_\epsilon) \sim N(\hat{\alpha}_j, \frac{\sigma^2_\epsilon}{c_j}) \]

\[ \hat{\alpha}_j = \frac{x_j' \mathbf{w}}{c_j} \]

\[ \mathbf{w} = \mathbf{y} - 1 \mu - \sum_{j' \neq j} x_{j'} \alpha_{j'} \]

\[ c_j = (x_j' x_j + \frac{\sigma^2_\epsilon}{\sigma^2_\alpha}) \]

\[ (\sigma^2_\epsilon|\mathbf{y}, \mu, \alpha) \sim \left[ (\mathbf{y} - \mathbf{W} \theta)'(\mathbf{y} - \mathbf{W} \theta) + \nu S^2_\epsilon\right] \chi^{-2}_{(\nu + n)} \]
Full conditionals for single-site Gibbs

\[ (\mu | y, \alpha, \sigma^2_e) \sim N\left(\frac{1'(y - X\alpha)}{n}, \frac{\sigma^2_e}{n}\right) \]

\[ (\alpha_j | y, \mu, \alpha_{j-}, \sigma^2_e) \sim N(\hat{\alpha}_j, \frac{\sigma^2_e}{c_j}) \]

\[ \hat{\alpha}_j = \frac{x_j'w}{c_j} \]

\[ w = y - 1\mu - \sum_{j' \neq j} x_{j'}\alpha_{j'} \]

\[ c_j = (x_j'x_j + \frac{\sigma^2_e}{\sigma^2_\alpha}) \]

\[ (\sigma^2_e | y, \mu, \alpha) \sim \left[(y - W\theta)'(y - W\theta) + \nu_e S^2_e\right] \chi^{-2}_{(\nu_e + n)} \]
Full conditionals for single-site Gibbs

\[
\begin{align*}
\mu | y, \alpha, \sigma_e^2 & \sim N \left( \frac{1'(y-X\alpha)}{n}, \frac{\sigma_e^2}{n} \right) \\
\alpha_j | y, \mu, \alpha_j, \sigma_e^2 & \sim N (\hat{\alpha}_j, \frac{\sigma_e^2}{c_j}) \\
\hat{\alpha}_j & = \frac{x'_j w}{c_j} \\
w & = y - 1\mu - \sum_{j' \neq j} x'_j \alpha_{j'} \\
c_j & = (x'_j x_j + \frac{\sigma_e^2}{\sigma_\alpha^2}) \\
\sigma_e^2 | y, \mu, \alpha & \sim \left[ (y - W\theta)'(y - W\theta) + \nu e S_e^2 \right] \chi_{(\nu_e + n)}^{-2}
\end{align*}
\]
Full conditionals for single-site Gibbs

1. \((\mu | \mathbf{y}, \alpha, \sigma^2_e) \sim N(\frac{1'}n (\mathbf{y} - \mathbf{X}\alpha), \frac{\sigma^2_e}{n})\)

2. \((\alpha_j | \mathbf{y}, \mu, \alpha_{j-}, \sigma^2_e) \sim N(\hat{\alpha}_j, \frac{\sigma^2_e}{c_j})\)

\[\hat{\alpha}_j = \frac{x_j' \mathbf{w}}{c_j}\]

3. \(\mathbf{w} = \mathbf{y} - 1\mu - \sum_{j' \neq j} x_{j'} \alpha_{j'}\)

4. \(c_j = (x_j'x_j + \frac{\sigma^2_e}{\sigma^2_\alpha})\)

5. \((\sigma^2_e | \mathbf{y}, \mu, \alpha) \sim [(\mathbf{y} - \mathbf{W}\theta)'(\mathbf{y} - \mathbf{W}\theta) + \nu_e S^2_e] \chi_{(\nu_e + n)}^{-2}\)
Derive: full conditional for $\alpha_j$

From Bayes’ Theorem,

$$f(\alpha_j|\mathbf{y}, \mu, \alpha_-, \sigma^2_e) = \frac{f(\alpha_j, \mathbf{y}, \mu, \alpha_-, \sigma^2_e)}{f(\mathbf{y}, \mu, \alpha_-, \sigma^2_e)}$$

$$\propto f(\mathbf{y}|\alpha_j, \mu, \alpha_-, \sigma^2_e) f(\alpha_j) f(\mu, \alpha_-, \sigma^2_e)$$

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

where

$$\mathbf{w} = \mathbf{y} - 1\mu - \sum_{j \neq j'} \mathbf{x}_{j'}\alpha_{j'}$$
Derive: full conditional for $\alpha_j$

From Bayes’ Theorem,

$$f(\alpha_j | y, \mu, \alpha_{-j}, \sigma^2_e) = \frac{f(\alpha_j, y, \mu, \alpha_{-j}, \sigma^2_e)}{f(y, \mu, \alpha_{-j}, \sigma^2_e)}$$

$$\propto f(y | \alpha_j, \mu, \alpha_{-j}, \sigma^2_e) f(\alpha_j) f(\mu, \alpha_{-j}, \sigma^2_e)$$

$$\propto (\sigma^2_e)^{-n/2} \exp\left\{ -\frac{(w - x_j \alpha_j)'(w - x_j \alpha_j)}{2\sigma^2_e} \right\} (\sigma^2_\alpha)^{-1/2} \exp\left\{ -\frac{\alpha_j^2}{2\sigma^2_\alpha} \right\}$$

where

$$w = y - 1\mu - \sum_{j \neq j'} x_j \alpha_{j'}$$
Derive: full conditional for $\alpha_j$

From Bayes’ Theorem,

$$f(\alpha_j|y, \mu, \alpha_{j -}, \sigma_e^2) = \frac{f(\alpha_j, y, \mu, \alpha_{j -}, \sigma_e^2)}{f(y, \mu, \alpha_{j -}, \sigma_e^2)}$$

$$\propto f(y|\alpha_j, \mu, \alpha_{j -}, \sigma_e^2)f(\alpha_j)f(\mu, \alpha_{j -}, \sigma_e^2)$$

$$\propto (\sigma_e^2)^{-n/2} \exp\left\{-\frac{(w - x_j\alpha_j)'(w - x_j\alpha_j)}{2\sigma_e^2}\right\}(\sigma^2_\alpha)^{-1/2} \exp\left\{-\frac{\alpha_j^2}{2\sigma^2_\alpha}\right\}$$

where

$$w = y - 1\mu - \sum_{j \neq j'} x_{j'}\alpha_{j'}$$
Derive: full conditional for $\alpha_j$

From Bayes’ Theorem,

$$f(\alpha_j | y, \mu, \alpha_j_, \sigma^2_\theta) = \frac{f(\alpha_j, y, \mu, \alpha_j_, \sigma^2_\theta)}{f(y, \mu, \alpha_j_, \sigma^2_\theta)}$$

$$\propto f(y | \alpha_j, \mu, \alpha_j_, \sigma^2_\theta)f(\alpha_j)f(\mu, \alpha_j_, \sigma^2_\theta)$$

$$\propto (\sigma^2_\theta)^{-n/2} \exp\{-\frac{(w - x_j\alpha_j)'(w - x_j\alpha_j)}{2\sigma^2_\theta}\}(\sigma^2_\alpha)^{-1/2} \exp\{-\frac{\alpha_j^2}{2\sigma^2_\alpha}\}$$

where

$$w = y - 1\mu - \sum_{j\neq j'} x_j'\alpha_j'$$
Derive: full conditional for $\alpha_j$

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

$$-rac{1}{2\sigma_e^2}\{w'w - 2x'_jw\alpha_j + [x'_jx_j + \frac{\sigma_e^2}{\sigma_\alpha^2}]\alpha_j^2\}$$

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

$$-rac{1}{2\sigma_e^2}\{c_j(\alpha_j - \hat{\alpha}_j)^2 + w'w - c_j\hat{\alpha}_j^2\}$$

where

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

So,

$$f(\alpha_j|y, \mu, \alpha_j, \sigma_e^2) \propto \exp\{-\frac{(\alpha_j - \hat{\alpha}_j)^2}{2\frac{\sigma_e^2}{c_j}}\}$$
Derive: full conditional for $\alpha_j$

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

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

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

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

where

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

So,

$$f(\alpha_j|\mathbf{y}, \mu, \alpha_{j-}, \sigma_e^2) \propto \exp\left\{ -\frac{(\alpha_j - \hat{\alpha}_j)^2}{2\frac{\sigma_e^2}{c_j}} \right\}$$
Derive: full conditional for $\alpha_j$

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

$$-\frac{1}{2\sigma^2_e} \{w'w - 2x_j'w\alpha_j + [x_j'x_j + \frac{\sigma^2_e}{\sigma^2_\alpha}]\alpha^2_j\}$$

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

$$-\frac{1}{2\sigma^2_e} \{c_j(\alpha_j - \hat{\alpha}_j)^2 + w'w - c_j\hat{\alpha}_j^2\}$$

where

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

So,

$$f(\alpha_j|y, \mu, \alpha_j, \sigma^2_e) \propto \exp\left\{-\frac{(\alpha_j - \hat{\alpha}_j)^2}{2\frac{\sigma^2_e}{c_j}}\right\}$$
Derive: full conditional for $\alpha_j$

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

$$-\frac{1}{2\sigma^2_e}\left\{w'w - 2x_j'w\alpha_j + [x_j'x_j + \frac{\sigma^2_e}{\sigma^2_\alpha}]\alpha^2_j\right\}$$

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

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

where

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

So,

$$f(\alpha_j|y, \mu, \alpha_j_, \sigma^2_e) \propto \exp\left\{-\frac{(\alpha_j - \hat{\alpha}_j)^2}{2\frac{\sigma^2_e}{c_j}}\right\}$$
Full conditional for $\sigma^2_e$

From Bayes’ theorem,

$$f(\sigma^2_e | y, \mu, \alpha) = \frac{f(\sigma^2_e, y, \mu, \alpha)}{f(y, \mu, \alpha)}$$

$$\propto f(y | \sigma^2_e, \mu, \alpha) f(\sigma^2_e) f(\mu, \alpha)$$

where

$$f(y | \sigma^2_e, \mu, \alpha) \propto (\sigma^2_e)^{-n/2} \exp\left\{ - \frac{(w - x_j \alpha_j)'(w - x_j \alpha_j)}{2\sigma^2_e} \right\}$$

and

$$f(\sigma^2_e) = \frac{(S^2_e \nu_e/2)^{\nu_e/2}}{\Gamma(\nu/2)} (\sigma^2_e)^{-(2+\nu_e)/2} \exp\left( - \frac{\nu_e S^2_e}{2\sigma^2_e} \right)$$
Full conditional for $\sigma^2_e$

From Bayes’ theorem,

$$f(\sigma^2_e|y, \mu, \alpha) = \frac{f(\sigma^2_e, y, \mu, \alpha)}{f(y, \mu, \alpha)}$$

$$\propto f(y|\sigma^2_e, \mu, \alpha) f(\sigma^2_e) f(\mu, \alpha)$$

where

$$f(y|\sigma^2_e, \mu, \alpha) \propto (\sigma^2_e)^{-n/2} \exp\left\{-\frac{(w - x_j\alpha_j)'(w - x_j\alpha_j)}{2\sigma^2_e}\right\}$$

and

$$f(\sigma^2_e) = \frac{(S_e^2\nu_e/2)^{\nu_e/2}}{\Gamma(\nu/2)} (\sigma^2_e)^{-(2+\nu_e)/2} \exp\left(-\frac{\nu_eS_e^2}{2\sigma^2_e}\right)$$
Full conditional for $\sigma_\epsilon^2$

From Bayes’ theorem,

$$f(\sigma_\epsilon^2|y, \mu, \alpha) = \frac{f(\sigma_\epsilon^2, y, \mu, \alpha)}{f(y, \mu, \alpha)}$$

$$\propto f(y|\sigma_\epsilon^2, \mu, \alpha)f(\sigma_\epsilon^2)f(\mu, \alpha)$$

where

$$f(y|\sigma_\epsilon^2, \mu, \alpha) \propto (\sigma_\epsilon^2)^{-n/2} \exp\left\{ -\frac{(w - x_j\alpha_j)'(w - x_j\alpha_j)}{2\sigma_\epsilon^2} \right\}$$

and

$$f(\sigma_\epsilon^2) = \frac{(S_\epsilon^2\nu_e/2)^{\nu_e/2}}{\Gamma(\nu/2)}(\sigma_\epsilon^2)^{-(2+\nu_e)/2} \exp\left( -\frac{\nu_eS_\epsilon^2}{2\sigma_\epsilon^2} \right)$$
Full conditional for $\sigma_e^2$

From Bayes’ theorem,

$$f(\sigma_e^2 | y, \mu, \alpha) = \frac{f(\sigma_e^2, y, \mu, \alpha)}{f(y, \mu, \alpha)}$$

$$\propto f(y | \sigma_e^2, \mu, \alpha)f(\sigma_e^2)f(\mu, \alpha)$$

where

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

and

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

So,

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

where

$$SSE = (w - x_j \alpha_j)'(w - x_j \alpha_j)$$

So,

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

where

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

So,

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

where

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

So,

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

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:

\[ y = 1\mu + X\alpha + e \]

This can be also written as

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

Suppose we observe for each locus:

\[ y_j^* = \alpha_j + \epsilon_j \]
Consider fixed linear model:

\[ y = 1\mu + X\alpha + e \]

This can be also written as

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

Suppose we observe for each locus:

\[ y^*_j = \alpha_j + \epsilon_j \]
Alternative view of Normal prior

Consider fixed linear model:

\[ y = 1_\mu + X_\alpha + e \]

This can be also written as

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

Suppose we observe for each locus:

\[ y_j^* = \alpha j + \epsilon_j \]
Least Squares with Additional Data

Fixed linear model with the additional data:

\[
\begin{bmatrix}
  y \\
  y^*
\end{bmatrix} =
\begin{bmatrix}
  1 & X \\
  0 & I
\end{bmatrix}
\begin{bmatrix}
  \mu \\
  \alpha
\end{bmatrix} +
\begin{bmatrix}
  e
\end{bmatrix}
\]

OLS Equations:

\[
\begin{bmatrix}
  1' & 0' \\
  X' & I'
\end{bmatrix}
\begin{bmatrix}
  \frac{1}{\sigma^2_e} & 0 \\
  0 & \frac{1}{\sigma^2_e}
\end{bmatrix}
\begin{bmatrix}
  1 & X \\
  0 & I
\end{bmatrix}
\begin{bmatrix}
  \hat{\mu} \\
  \hat{\alpha}
\end{bmatrix} =
\begin{bmatrix}
  1' & 0' \\
  X' & I'
\end{bmatrix}
\begin{bmatrix}
  \frac{1}{\sigma^2_e} & 0 \\
  0 & \frac{1}{\sigma^2_e}
\end{bmatrix}
\begin{bmatrix}
  y \\
  y^*
\end{bmatrix}
\]

\[
\begin{bmatrix}
  1'1 \\
  X'1 \\
  X'X + I \frac{\sigma^2_e}{\sigma^2_e}
\end{bmatrix}
\begin{bmatrix}
  \hat{\mu} \\
  \hat{\alpha}
\end{bmatrix} =
\begin{bmatrix}
  1'y \\
  X'y + y^* \frac{\sigma^2_e}{\sigma^2_e}
\end{bmatrix}
\]
Least Squares with Additional Data

Fixed linear model with the additional data:

\[
\begin{bmatrix}
  y \\
  y^*
\end{bmatrix} =
\begin{bmatrix}
  1 & X \\
  0 & I
\end{bmatrix}
\begin{bmatrix}
  \mu \\
  \alpha
\end{bmatrix} +
\begin{bmatrix}
  e
\end{bmatrix}
\]

OLS Equations:

\[
\begin{bmatrix}
  1' \\
  X'
\end{bmatrix}
\begin{bmatrix}
  I_n \frac{1}{\sigma_e^2} & 0 \\
  0 & I_k \frac{1}{\sigma_e^2}
\end{bmatrix}
\begin{bmatrix}
  1' \\
  X'
\end{bmatrix}
\begin{bmatrix}
  \hat{\mu} \\
  \hat{\alpha}
\end{bmatrix} =
\begin{bmatrix}
  1' \\
  X'
\end{bmatrix}
\begin{bmatrix}
  I_n \frac{1}{\sigma_e^2} & 0 \\
  0 & I_k \frac{1}{\sigma_e^2}
\end{bmatrix}
\begin{bmatrix}
  y \\
  y^*
\end{bmatrix}
\]

\[
\begin{bmatrix}
  1'1 & 1'X \\
  X'1 & X'X + I \frac{\sigma_e^2}{\sigma_e^2}
\end{bmatrix}
\begin{bmatrix}
  \hat{\mu} \\
  \hat{\alpha}
\end{bmatrix} =
\begin{bmatrix}
  1'y \\
  X'y + y^* \frac{\sigma_e^2}{\sigma_e^2}
\end{bmatrix}
\]
Least Squares with Additional Data

Fixed linear model with the additional data:

\[
\begin{bmatrix}
  y \\
y^\star
\end{bmatrix}
= \begin{bmatrix}
  1 & X \\
  0 & I
\end{bmatrix}
\begin{bmatrix}
  \mu \\
  \alpha
\end{bmatrix}
+ \begin{bmatrix}
  e
\end{bmatrix}
\]

OLS Equations:

\[
\begin{bmatrix}
  1' & 0' \\
  X' & I'
\end{bmatrix}
\begin{bmatrix}
  I_n \frac{1}{\sigma_e^2} & 0 \\
  0 & I_k \frac{1}{\sigma_e^2}
\end{bmatrix}
\begin{bmatrix}
  1 & X \\
  0 & I
\end{bmatrix}
\hat{\alpha}
= \begin{bmatrix}
  1' & 0' \\
  X' & I'
\end{bmatrix}
\begin{bmatrix}
  I_n \frac{1}{\sigma_e^2} & 0 \\
  0 & I_k \frac{1}{\sigma_e^2}
\end{bmatrix}
\begin{bmatrix}
  y \\
y^\star
\end{bmatrix}
\]

\[
\begin{bmatrix}
  1'1 \\
  X'1 \\
  X' & X'X + I \frac{\sigma_e^2}{\sigma_c^2}
\end{bmatrix}
\begin{bmatrix}
  \hat{\mu} \\
  \hat{\alpha}
\end{bmatrix}
= \begin{bmatrix}
  1'y \\
  X'y + y^\star \frac{\sigma_e^2}{\sigma_c^2}
\end{bmatrix}
\]
Univariate-$t$

Prior:

\[ (\alpha_j|\sigma_j^2) \sim N(0, \sigma_j^2) \]
\[ \sigma_j^2 \sim \nu_{\alpha} S_{\nu_{\alpha}}^2 \chi_{\nu_{\alpha}}^{-2} \]

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

\[ \alpha_j \sim (\text{iid}) t(0, S_{\nu_{\alpha}}^2, \nu_{\alpha}) \]

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

This is Bayes-A (Meuwissen et al., 2001; Genetics 157:1819-1829)
Univariate-\textit{t}

Prior:

\[(\alpha_j|\sigma^2_j) \sim \text{N}(0, \sigma^2_j)\]

\[\sigma^2_j \sim \nu_{\alpha} S_{\nu_{\alpha}}^2 \chi_{\nu_{\alpha}}^{-2}\]

Can show that the unconditional distribution for \(\alpha_j\) is

\[\alpha_j \sim (\text{iid})t(0, S_{\nu_{\alpha}}^2, \nu_{\alpha})\]

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

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

Prior:

\[
(\alpha_j | \sigma_j^2) \sim N(0, \sigma_j^2)
\]
\[
\sigma_j^2 \sim \nu_\alpha S_{\nu_\alpha}^2 \chi_{\nu_\alpha}^{-2}
\]

Can show that the unconditional distribution for \(\alpha_j\) is

\[
\alpha_j \sim (iid) t(0, S_{\nu_\alpha}^2, \nu_\alpha)
\]

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

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

Plots of PDF for typical parameters:

\[
\begin{align*}
\text{\( v = 1 \)} & \quad \text{\( v = 3 \)} & \quad \text{\( v = 20 \)} \\
0.4 & \quad 0.3 & \quad 0.2 \\
0.1 & \quad 0.0 \\
-4 & \quad -2 & \quad 0 & \quad 2 & \quad 4 & \quad 6 & \quad 8 & \quad 10
\end{align*}
\]

Generated by Wolfram|Alpha (www.wolframalpha.com)
Full conditional for single-site Gibbs

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

$$\xi = [\sigma^2_1, \sigma^2_2, \ldots, \sigma^2_k]$$

Full conditional conditional for $\sigma^2_j$:

$$f(\sigma^2_j | y, \mu, \alpha, \xi, \sigma^2_e) \propto f(y, \mu, \alpha, \xi, \sigma^2_e)$$

$$\propto f(y | \mu, \alpha, \xi, \sigma^2_e) f(\alpha_j | \sigma^2_j) f(\sigma^2_j) f(\mu, \alpha_j _, \xi_j _, \sigma^2_e)$$

$$\propto (\sigma^2_j)^{-1/2} \exp\left\{-\frac{\alpha^2_j}{2\sigma^2_j}\right\} (\sigma^2_j)^{-(2+\nu_\alpha)/2} \exp\left\{\frac{\nu_\alpha S^2_\alpha}{2\sigma^2_j}\right\}$$

$$\propto (\sigma^2_j)^{-(2+\nu_\alpha+1)/2} \exp\left\{\frac{\alpha^2_j + \nu_\alpha S^2_\alpha}{2\sigma^2_j}\right\}$$
Full conditional for single-site Gibbs

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

$$\xi = [\sigma^2_1, \sigma^2_2, \ldots, \sigma^2_k]$$

Full conditional conditional for $\sigma^2_j$:

$$f(\sigma^2_j | y, \mu, \alpha, \xi_{-j}, \sigma^2_e) \propto f(y, \mu, \alpha, \xi, \sigma^2_e)$$

$$\propto f(y | \mu, \alpha, \xi, \sigma^2_e) f(\alpha_j | \sigma^2_j) f(\sigma^2_j) f(\mu, \alpha_{-j}, \xi_{-j}, \sigma^2_e)$$

$$\propto (\sigma^2_j)^{-1/2} \exp\left\{-\frac{\alpha^2_j}{2\sigma^2_j}\right\} (\sigma^2_j)^{-(2+\nu_\alpha)/2} \exp\left\{\frac{\nu_\alpha S^2_\alpha}{2\sigma^2_j}\right\}$$

$$\propto (\sigma^2_j)^{-(2+\nu_\alpha+1)/2} \exp\left\{\frac{\alpha^2_j + \nu_\alpha S^2_\alpha}{2\sigma^2_j}\right\}$$
Full conditional for single-site Gibbs

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

$$\xi = [\sigma^2_1, \sigma^2_2, \ldots, \sigma^2_k]$$

Full conditional conditional for $\sigma^2_j$:

$$f(\sigma^2_j | y, \mu, \alpha, \xi_{-j}, \sigma^2_e) \propto f(y, \mu, \alpha, \xi, \sigma^2_e)$$

$$\propto f(y | \mu, \alpha, \xi, \sigma^2_e)f(\alpha_j | \sigma^2_j)f(\sigma^2_j)f(\mu, \alpha_{-j}, \xi_{-j} \sigma^2_e)$$

$$\propto (\sigma^2_j)^{-1/2} \exp\left\{-\frac{\alpha^2_j}{2\sigma^2_j}\right\}(\sigma^2_j)^{-2+\nu\alpha/2} \exp\left\{\frac{\nu\alpha S^2_\alpha}{2\sigma^2_j}\right\}$$

$$\propto (\sigma^2_j)^{-2+\nu\alpha+1/2} \exp\left\{\frac{\alpha^2_j + \nu\alpha S^2_\alpha}{2\sigma^2_j}\right\}$$
Full conditional for single-site Gibbs

Full conditionals are the same as in the "Normal" model for \( \mu, \alpha_j, \) and \( \sigma^2_e \). Let

\[ \xi = [\sigma^2_1, \sigma^2_2, \ldots, \sigma^2_k] \]

Full conditional conditional for \( \sigma^2_j \):

\[
f(\sigma^2_j | y, \mu, \alpha, \xi, \sigma^2_e) \propto f(y, \mu, \alpha, \xi, \sigma^2_e) \propto f(y | \mu, \alpha, \xi, \sigma^2_e) f(\alpha_j | \sigma^2_j) f(\sigma^2_j) f(\mu, \alpha_j, \xi, \sigma^2_e) \]

\[
\propto (\sigma^2_j)^{-1/2} \exp\{-\frac{\alpha^2_j}{2\sigma^2_j}\}(\sigma^2_j)^{-\frac{2+\nu\alpha}{2}} \exp\{\frac{\nu\alpha S^2_{\alpha}}{2\sigma^2_j}\} \]

\[
\propto (\sigma^2_j)^{-\frac{(2+\nu\alpha+1)}{2}} \exp\{\frac{\alpha^2_j + \nu\alpha S^2_{\alpha}}{2\sigma^2_j}\} \]
Full conditional for single-site Gibbs

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

$$\xi = [\sigma^2_1, \sigma^2_2, \ldots, \sigma^2_k]$$

Full conditional conditional for $\sigma^2_j$:

$$f(\sigma^2_j | y, \mu, \alpha, \xi_{-j}, \sigma^2_e) \propto f(y, \mu, \alpha, \xi, \sigma^2_e)$$

$$\propto f(y | \mu, \alpha, \xi, \sigma^2_e) f(\alpha_j | \sigma^2_j) f(\sigma^2_j) f(\mu, \alpha_{-j}, \xi_{-j} \sigma^2_e)$$

$$\propto (\sigma^2_j)^{-1/2} \exp\left\{-\frac{\alpha_j^2}{2\sigma^2_j}\right\} (\sigma^2_j)^{-(2+\nu_\alpha)/2} \exp\left\{\frac{\nu_\alpha S^2_\alpha}{2\sigma^2_j}\right\}$$

$$\propto (\sigma^2_j)^{-(2+\nu_\alpha+1)/2} \exp\left\{\frac{\alpha_j^2 + \nu_\alpha S^2_\alpha}{2\sigma^2_j}\right\}$$
Full conditional for single-site Gibbs

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

$$\xi = [\sigma^2_1, \sigma^2_2, \ldots, \sigma^2_k]$$

Full conditional conditional for $\sigma^2_j$:

$$f(\sigma^2_j | y, \mu, \alpha, \xi, \sigma^2_\epsilon) \propto f(y, \mu, \alpha, \xi, \sigma^2_\epsilon)$$

$$\propto f(y | \mu, \alpha, \xi, \sigma^2_\epsilon) f(\alpha_j | \sigma^2_j) f(\sigma^2_j) f(\mu, \alpha_j, \xi_j \sigma^2_\epsilon)$$

$$\propto (\sigma^2_j)^{-1/2} \exp\left\{-\frac{\alpha_j^2}{2\sigma^2_j}\right\} (\sigma^2_j)^{-(2+\nu_\alpha)/2} \exp\left\{\frac{\nu_\alpha S^2_\alpha}{2\sigma^2_j}\right\}$$

$$\propto (\sigma^2_j)^{-2+\nu_\alpha+1/2} \exp\left\{\frac{\alpha_j^2 + \nu_\alpha S^2_\alpha}{2\sigma^2_j}\right\}$$
Full conditional for $\sigma_j^2$

So,

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

where

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

and

$$\tilde{S}_\alpha^2 = \frac{\alpha_j^2 + \nu_\alpha S_\alpha^2}{\tilde{\nu}_\alpha}$$
Multivariate-\(t\)

Prior:

\[
(\alpha_j | \sigma^2_\alpha) \sim \text{(iid)N}(0, \sigma^2_\alpha)
\]

\[
\sigma^2_\alpha \sim \nu_\alpha S^2_{\nu_\alpha} \chi^{-2}_{\nu_\alpha}
\]

Can show that the unconditional distribution for \(\alpha\) is

\[
\alpha \sim \text{multivariate-}t(0, IS^2_{\nu_\alpha}, \nu_\alpha)
\]

(Sorensen and Gianola, 2002, LBMMQG page 60)

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

Prior:

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

$$\sigma^2_\alpha \sim \nu_\alpha S^2_{\nu_\alpha} \chi^{-2}_{\nu_\alpha}$$

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

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

(Sorensen and Gianola, 2002, LBMMQG page 60)

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

Prior:

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

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

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

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

(Sorensen and Gianola, 2002, LBMMQG page 60)

We will see later that this is Bayes-C with $\pi = 0$. 
Full conditional for $\sigma^{2}_{\alpha}$

We will see later that 

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

where

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

and

$$\tilde{S}^{2}_{\alpha} = \frac{\alpha' \alpha + \nu_{\alpha} S^{2}_{\alpha}}{\tilde{\nu}_{\alpha}}$$
Spike and univariate-t

Prior:

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

and

\[
(\sigma_j^2 | \nu_\alpha, S_\alpha^2) \sim \nu_\alpha S_\alpha^2 \chi_{\nu_\alpha}^{-2}
\]

Thus,

\[
(\alpha_j | \pi) (\text{iid}) \begin{cases} \sim \text{univariate-t}(0, S_\alpha^2, \nu_\alpha) & \text{probability } (1 - \pi), \\ = 0 & \text{probability } \pi \end{cases}
\]

This is Bayes-B (Meuwissen et al., 2001; Genetics 157:1819-1829)
Spike and univariate-$t$

Prior:

\[
(\alpha_j | \pi, \sigma^2_j) \begin{cases} 
\sim N(0, \sigma^2_j) & \text{probability } (1 - \pi), \\
= 0 & \text{probability } \pi
\end{cases}
\]

and

\[
(\sigma^2_j | \nu_\alpha, S^2_\alpha) \sim \nu_\alpha S^2_\alpha \chi^{-2}_{\nu_\alpha}
\]

Thus,

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

This is Bayes-B (Meuwissen et al., 2001; Genetics 157:1819-1829)
Spike and univariate-\(t\)

Prior:

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

and

\[
(\sigma_j^2 | \nu_\alpha, S_\alpha^2) \sim \nu_\alpha S_\alpha^2 \chi_{\nu_\alpha}^{-2}
\]

Thus,

\[
(\alpha_j | \pi) (\text{iid}) \begin{cases} 
\sim \text{univariate-}t(0, S_\alpha^2, \nu_\alpha) & \text{probability } (1 - \pi), \\
= 0 & \text{probability } \pi
\end{cases}
\]

This is Bayes-B (Meuwissen et al., 2001; Genetics 157:1819-1829)
The indicator variable $\delta_j$ is defined as

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

and

$$\delta_j = 0 \Rightarrow (\alpha_j | \sigma_j^2) = 0$$
Sampling strategy in MHG (2001)

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

$$f(\sigma^2_j|y, \mu, \alpha_j, \xi, \sigma^2_\theta)$$

using MH with prior as proposal.
- Then, $\alpha_j$ is sampled from its full-conditional, which is identical to that under the Normal prior.
Sampling strategy in MHG (2001)

- Sampling $\sigma^2$ and $\mu$ are as under the Normal prior.
- MHG proposed to use a Metropolis-Hastings sampler to draw samples for $\sigma^2_j$ and $\alpha_j$ jointly from their full-conditional distribution.

  - First, $\sigma^2_j$ is sampled from

    $$f(\sigma^2_j | y, \mu, \alpha_j, \xi, \sigma^2_\theta)$$

    using MH with prior as proposal.

  - Then, $\alpha_j$ is sampled from its full-conditional, which is identical to that under the Normal prior.
Sampling strategy in MHG (2001)

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

$$f(\sigma^2_j | y, \mu, \alpha_j, \xi_-, \sigma^2_\epsilon)$$

using MH with prior as proposal.
- Then, $\alpha_j$ is sampled from its full-conditional, which is identical to that under the Normal prior.
Sampling strategy in MHG (2001)

- Sampling $\sigma^2_\varepsilon$ and $\mu$ are as under the Normal prior.
- MHG proposed to use a Metropolis-Hastings sampler to draw samples for $\sigma^2_j$ and $\alpha_j$ jointly from their full-conditional distribution.
- First, $\sigma^2_j$ is sampled from
  \[
  f(\sigma^2_j | y, \mu, \alpha_j, \xi, \sigma^2_\varepsilon)
  \]
  using MH with prior as proposal.
- Then, $\alpha_j$ is sampled from its full-conditional, which is identical to that under the Normal prior.
Suppose we want to sample $\theta$ from $f(\theta|y)$ using the MH with its prior as proposal. Then, the MH acceptance probability becomes:

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

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

$$f(\theta|y) \propto f(y|\theta)f(\theta)$$

Then, the acceptance probability becomes

$$\alpha = \min(1, \frac{f(y|\theta_{can})f(\theta_{can})f(\theta^{t-1})}{f(y|\theta^{t-1})f(\theta^{t-1})f(\theta_{can})})$$
MH acceptance probability when prior is used as proposal

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

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

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

$$f(\theta|y) \propto f(y|\theta)f(\theta)$$

Then, the acceptance probability becomes

$$\alpha =\min(1, \frac{f(y|\theta_{can})f(\theta_{can})f(\theta^{t-1})}{f(y|\theta^{t-1})f(\theta^{t-1})f(\theta_{can})})$$
MH acceptance probability when prior is used as proposal

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

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

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

$$f(\theta|y) \propto f(y|\theta)f(\theta)$$

Then, the acceptance probability becomes

$$\alpha = \min(1, \frac{f(y|\theta_{\text{can}})f(\theta_{\text{can}})f(\theta^{t-1})}{f(y|\theta^{t-1})f(\theta^{t-1})f(\theta_{\text{can}})})$$
MH acceptance probability when prior is used as proposal

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

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

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

$$f(\theta|y) \propto f(y|\theta)f(\theta)$$

Then, the acceptance probability becomes

$$\alpha = \min(1, \frac{f(y|\theta_{\text{can}})f(\theta_{\text{can}})f(\theta^{t-1})}{f(y|\theta^{t-1})f(\theta^{t-1})f(\theta_{\text{can}})})$$
Sampling $\sigma_j^2$

Thus when the prior for $\sigma_j^2$ is used as the proposal, the MH acceptance probability becomes

$$\alpha = \min(1, \frac{f(y|\sigma_{can}^2, \theta_{j\_})}{f(y|\sigma_j^2, \theta_{j\_})})$$

where $\sigma_{can}^2$ is used to denote the candidate value for $\sigma_j^2$, and $\theta_{j\_}$ all the other parameters. It can be shown that, $\alpha_j$ depends on $y$ only through $r_j = x'_j w$ (page 30). Thus

$$f(y|\sigma_j^2, \theta_{j\_}) \propto f(r_j|\sigma_j^2, \theta_{j\_})$$
Sampling $\sigma_j^2$

Thus when the prior for $\sigma_j^2$ is used as the proposal, the MH acceptance probability becomes

$$\alpha = \min(1, \frac{f(y|\sigma_{can}^2, \theta_\perp)}{f(y|\sigma_j^2, \theta_\perp)})$$

where $\sigma_{can}^2$ is used to denote the candidate value for $\sigma_j^2$, and $\theta_\perp$ all the other parameters. It can be shown that, $\alpha_j$ depends on $y$ only through $r_j = x_j'w$ (page 30). Thus

$$f(y|\sigma_j^2, \theta_\perp) \propto f(r_j|\sigma_j^2, \theta_\perp)$$
"Likelihood" for $\sigma_j^2$

Recall that

$$w = y - 1\mu - \sum_{j' \neq j} x_{j'} \alpha_{j'} = x_j \alpha_j + e$$

Then,

$$E(w|\sigma_j^2, \theta_j_\cdot) = 0$$

When $\delta = 1$:

$$\text{Var}(w|\delta_j = 1, \sigma_j^2, \theta_j_\cdot) = x_j x_j' \sigma_j^2 + I\sigma_e^2$$

and $\delta = 0$:

$$\text{Var}(w|\delta_j = 0, \sigma_j^2, \theta_j_\cdot) = I\sigma_e^2$$
"Likelihood" for $\sigma_j^2$

Recall that

$$w = y - 1\mu - \sum_{j' \neq j} x_{j'}\alpha_{j'} = x_j\alpha_j + e$$

Then,

$$E(w|\sigma_j^2, \theta_j) = 0$$

When $\delta = 1$:

$$\text{Var}(w|\delta_j = 1, \sigma_j^2, \theta_j) = x_jx_j'\sigma_j^2 + I\sigma_e^2$$

and $\delta = 0$:

$$\text{Var}(w|\delta_j = 0, \sigma_j^2, \theta_j) = I\sigma_e^2$$
"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,

$$\mathbb{E}(\mathbf{w}|\sigma_j^2, \theta_j) = 0$$

When $\delta = 1$:

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

and $\delta = 0$:

$$\text{Var}(\mathbf{w}|\delta_j = 0, \sigma_j^2, \theta_j) = \mathbf{I}\sigma_e^2$$
"Likelihood" for $\sigma_j^2$

Recall that

$$w = y - \mathbf{1}\mu - \sum_{j' \neq j} x_{j'} \alpha_{j'} = x_j \alpha_j + e$$

Then,

$$\mathbb{E}(w | \sigma_j^2, \theta_{j_-}) = 0$$

When $\delta = 1$:

$$\text{Var}(w | \delta_j = 1, \sigma_j^2, \theta_{j_-}) = x_j x_j' \sigma_j^2 + I \sigma_e^2$$

and $\delta = 0$:

$$\text{Var}(w | \delta_j = 0, \sigma_j^2, \theta_{j_-}) = I \sigma_e^2$$
"Likelihood" for $\sigma_j^2$

So,

$$E(r_j|\sigma_j^2, \theta_j) = 0$$

and

$$\text{Var}(r_j|\delta_j = 1, \sigma_j^2, \theta_j) = (x'_j x_j)^2 \sigma_j^2 + x'_j x_j \sigma_e^2 = v_1$$

$$\text{Var}(r_j|\delta_j = 0, \sigma_j^2, \theta_j) = x'_j x_j \sigma_e^2 = v_0$$

So,

$$f(r_j|\delta_j, \sigma_j^2, \theta_j) \propto (v_\delta)^{-1/2} \exp\left\{-\frac{r_j^2}{2v_\delta}\right\}$$
"Likelihood" for $\sigma^2_j$

So,

$$E(r_j | \sigma^2_j, \theta_j_) = 0$$

and

$$\text{Var}(r_j | \delta_j = 1, \sigma^2_j, \theta_j_) = (x'_j x_j)^2 \sigma^2_j + x'_j x_j \sigma^2_e = \nu_1$$

$$\text{Var}(r_j | \delta_j = 0, \sigma^2_j, \theta_j_) = x'_j x_j \sigma^2_e = \nu_0$$

So,

$$f(r_j | \delta_j, \sigma^2_j, \theta_j_) \propto (\nu_\delta)^{-1/2} \exp\left\{-\frac{r_j^2}{2\nu_\delta}\right\}$$
Alternative View of Prior in BayesB

- How much information is being added by the prior?
- BayesB is identical to ML with additional data!
- Can “see” how much additional data in BayesB prior.
Alternative View of Prior in BayesB

- How much information is being added by the prior?
- BayesB is identical to ML with additional data!
- Can “see” how much additional data in BayesB prior.
Alternative View of Prior in BayesB

- How much information is being added by the prior?
- BayesB is identical to ML with additional data!
- Can “see” how much additional data in BayesB prior.
Suppose at locus $j$, $\delta_j = 1$, and we observe additional data:

$$u_j \sim N(0, I_q \sigma_j^2)$$

Assume that only unknown is $\sigma_j^2$

So, adjust phenotypes as:

$$w = y - 1\mu - \sum_{j' \neq j} x_{j'} \alpha_{j'}$$

Likelihood:

$$L(\sigma_j^2; w, u_j) = L(\sigma_j^2; \hat{\alpha}_j, u_j)$$
Likelihood with Additional Data

\[ L(\sigma_j^2; \hat{\alpha}_j, u_j) \propto f_1(\hat{\alpha}_j|\sigma_j^2) \times f_2(u_j|\sigma_j^2) \]

\[ f_2(u_j|\sigma_j^2) \propto (\sigma_j^2)^{-q/2} \exp\left[\frac{-u_j'u_j}{2\sigma_j^2}\right] \]

\[ \propto (\sigma_j^2)^{-[\nu/2+1]} \exp\left[\frac{-\nu S^2}{2\sigma_j^2}\right] \]

\[ \nu = q - 2, \quad S^2 = \frac{u_j'u_j}{\nu} \]
Likelihood with Additional Data

\[ L(\sigma_j^2; \hat{\alpha}_j, u_j) \propto f_1(\hat{\alpha}_j|\sigma_j^2) \times f_2(u_j|\sigma_j^2) \]

\[ f_2(u_j|\sigma_j^2) \propto (\sigma_j^2)^{-q/2} \exp\left[\frac{-u'_j u_j}{2\sigma_j^2}\right] \]

\[ \propto (\sigma_j^2)^{-[\nu/2+1]} \exp\left[\frac{-\nu S^2}{2\sigma_j^2}\right] \]

\[ \nu = q - 2, \ S^2 = \frac{u'_j u_j}{\nu} \]
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 $= \frac{\sigma_j^{2(t-1)}}{2}$ and 4 degrees of freedom when $\delta_j^{(t-1)} = 1$, and
  - scale parameter $= S_{\alpha}^2$ and 4 degrees of freedom when $\delta_j^{(t-1)} = 0$
Alternative algorithm for spike and univariate-t

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

- sample $\delta_j = 1$ with probability 0.5
- when $\delta = 1$, sample $\sigma^2_j$ from a scaled inverse chi-squared distribution with
  - scale parameter $= \sigma^2_j (t-1)/2$ and 4 degrees of freedom when $\delta_j^{(t-1)} = 1$, and
  - scale parameter $= S^2_\alpha$ and 4 degrees of freedom when $\delta_j^{(t-1)} = 0$
Alternative algorithm for spike and univariate-t

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

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

Prior:

\[
(\alpha_j | \pi, \sigma^2_\alpha) \begin{cases} 
\sim \text{N}(0, \sigma^2_\alpha) & \text{probability } (1 - \pi), \\
= 0 & \text{probability } \pi
\end{cases}
\]

and

\[
(\sigma^2_\alpha | \nu_\alpha, S^2_\alpha) \sim \nu_\alpha S^2_\alpha \chi_{\nu_\alpha}^{-2}
\]

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\(\pi\)
Multivariate-\( t \) mixture

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

and
\[
(\sigma_\alpha^2 | \nu_\alpha, S_\alpha^2) \sim \nu_\alpha S_\alpha^2 \chi_{\nu_\alpha}^{-2}
\]

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\( \pi \)
Multivariate-\( t \) mixture

Prior:
\[
\begin{align*}
(\alpha_j | \pi, \sigma^2_\alpha) & \sim \begin{cases} 
    N(0, \sigma^2_\alpha) & \text{probability } (1 - \pi), \\
    = 0 & \text{probability } \pi
\end{cases}
\end{align*}
\]

and
\[
(\sigma^2_\alpha | \nu_\alpha, S^2_\alpha) \sim \nu_\alpha S^2_\alpha \chi^{-2}_{\nu_\alpha}
\]

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\( \pi \)
Multivariate-\(t\) mixture

Prior:

\[
(\alpha_j | \pi, \sigma^2_\alpha) \begin{cases} \sim N(0, \sigma^2_\alpha) & \text{probability } (1 - \pi), \\ = 0 & \text{probability } \pi \end{cases}
\]

and

\[
(\sigma^2_\alpha | \nu_\alpha, S^2_\alpha) \sim \nu_\alpha S^2_\alpha \chi_{\nu_\alpha}^{-2}
\]

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 $\sigma^2_e$ are as with the Normal prior.

Full-conditional for $\delta_j$:

$$
\Pr(\delta_j|\mathbf{y}, \mu, \alpha_{-j}, \delta_{-j}, \sigma^2_{\alpha}, \sigma^2_e, \pi) = \Pr(\delta_j|r_j, \theta_{j_-})
$$

$$
\Pr(\delta_j|r_j, \theta_{j_-}) = \frac{f(\delta_j, r_j|\theta_{j_-})}{f(r_j|\theta_{j_-})}
$$

$$
= \frac{f(r_j|\delta_j, \theta_{j_-})\Pr(\delta_j|\pi)}{f(r_j|\delta_j = 0, \theta_{j_-})\pi + f(r_j|\delta_j = 1, \theta_{j_-})(1 - \pi)}
$$
Full conditionals for single-site Gibbs

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

Full-conditional for $\delta_j$:

$$\Pr(\delta_j|y, \mu, \alpha_{-j}, \delta_{-j}, \sigma^2_\alpha, \sigma_e^2, \pi) = \Pr(\delta_j|r_j, \theta_{j_-})$$

$$\Pr(\delta_j|r_j, \theta_{j_-}) = \frac{f(\delta_j, r_j|\theta_{j_-})}{f(r_j|\theta_{j_-})}$$

$$= \frac{f(r_j|\delta_j, \theta_{j_-}) \Pr(\delta_j|\pi)}{f(r_j|\delta_j = 0, \theta_{j_-})\pi + f(r_j|\delta_j = 1, \theta_{j_-})(1 - \pi)}$$
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 $\delta_j$:

$$
\Pr(\delta_j|y, \mu, \alpha_{-j}, \delta_{-j}, \sigma^2_\alpha, \sigma^2_e, \pi) = 
\Pr(\delta_j|r_j, \theta_{j_\_})
$$

$$
\Pr(\delta_j|r_j, \theta_{j_\_}) = \frac{f(\delta_j, r_j|\theta_{j_\_})}{f(r_j|\theta_{j_\_})}
$$

$$
= \frac{f(r_j|\delta_j, \theta_{j_\_}) \Pr(\delta_j|\pi)}{f(r_j|\delta_j = 0, \theta_{j_\_})\pi + f(r_j|\delta_j = 1, \theta_{j_\_})(1 - \pi)}
$$
Full conditionals for single-site Gibbs

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

Full-conditional for $\delta_j$:

\[
\Pr(\delta_j|y, \mu, \alpha_{-j}, \delta_{-j}, \sigma^2_\alpha, \sigma^2_e, \pi) = \Pr(\delta_j|r_j, \theta_{j_{-}})
\]

\[
\Pr(\delta_j|r_j, \theta_{j_{-}}) = \frac{f(\delta_j, r_j|\theta_{j_{-}})}{f(r_j|\theta_{j_{-}})}
\]

\[
= \frac{f(r_j|\delta_j, \theta_{j_{-}})\Pr(\delta_j|\pi)}{f(r_j|\delta_j = 0, \theta_{j_{-}})\pi + f(r_j|\delta_j = 1, \theta_{j_{-}})(1 - \pi)}
\]
Full conditional for $\sigma^2_\alpha$

This can be written as

$$f(\sigma^2_\alpha | y, \mu, \alpha, \delta, \sigma^2_e) \propto f(y | \sigma^2_\alpha, \mu, \alpha, \delta, \sigma^2_e) f(\sigma^2_\alpha, \mu, \alpha, \delta, \sigma^2_e)$$

But, can see that

$$f(y | \sigma^2_\alpha, \mu, \alpha, \delta, \sigma^2_e) \propto f(y | \mu, \alpha, \delta, \sigma^2_e)$$

So,

$$f(\sigma^2_\alpha | y, \mu, \alpha, \delta, \sigma^2_e) \propto f(\sigma^2_\alpha, \mu, \alpha, \delta, \sigma^2_e)$$

Note that $\sigma^2_\alpha$ appears only in $f(\alpha | \sigma^2_\alpha)$ and $f(\sigma^2_\alpha)$:

$$f(\alpha | \sigma^2_\alpha) \propto (\sigma^2_\alpha)^{-k/2} \exp\{ -\frac{\alpha' \alpha}{2\sigma^2_\alpha} \}$$

and

$$f(\sigma^2_\alpha) \propto (\sigma^2_\alpha)^{-(\nu_\alpha+2)/2} \exp\{ \frac{\nu_\alpha S^2_\alpha}{2\sigma^2_\alpha} \}$$
Full conditional for $\sigma^2_\alpha$

This can be written as

$$f(\sigma^2_\alpha | y, \mu, \alpha, \delta, \sigma^2_\epsilon) \propto f(y | \sigma^2_\alpha, \mu, \alpha, \delta, \sigma^2_\epsilon) f(\sigma^2_\alpha, \mu, \alpha, \delta, \sigma^2_\epsilon)$$

But, can see that

$$f(y | \sigma^2_\alpha, \mu, \alpha, \delta, \sigma^2_\epsilon) \propto f(y | \mu, \alpha, \delta, \sigma^2_\epsilon)$$

So,

$$f(\sigma^2_\alpha | y, \mu, \alpha, \delta, \sigma^2_\epsilon) \propto f(\sigma^2_\alpha, \mu, \alpha, \delta, \sigma^2_\epsilon)$$

Note that $\sigma^2_\alpha$ appears only in $f(\alpha | \sigma^2_\alpha)$ and $f(\sigma^2_\alpha)$:

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

and

$$f(\sigma^2_\alpha) \propto (\sigma^2_\alpha)^{-(\nu_\alpha+2)/2} \exp\left\{ \frac{\nu_\alpha S^2_\alpha}{2\sigma^2_\alpha} \right\}$$
Full conditional for $\sigma^2_\alpha$

This can be written as

$$f(\sigma^2_\alpha | y, \mu, \alpha, \delta, \sigma^2_\theta) \propto f(y | \sigma^2_\alpha, \mu, \alpha, \delta, \sigma^2_\theta) f(\sigma^2_\alpha, \mu, \alpha, \delta, \sigma^2_\theta)$$

But, can see that

$$f(y | \sigma^2_\alpha, \mu, \alpha, \delta, \sigma^2_\theta) \propto f(y | \mu, \alpha, \delta, \sigma^2_\theta)$$

So,

$$f(\sigma^2_\alpha | y, \mu, \alpha, \delta, \sigma^2_\theta) \propto f(\sigma^2_\alpha, \mu, \alpha, \delta, \sigma^2_\theta)$$

Note that $\sigma^2_\alpha$ appears only in $f(\alpha | \sigma^2_\alpha)$ and $f(\sigma^2_\alpha)$:

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

and

$$f(\sigma^2_\alpha) \propto (\sigma^2_\alpha)^{-(\nu_\alpha+2)/2} \exp\left\{ \frac{\nu_\alpha S^2_\alpha}{2 \sigma^2_\alpha} \right\}$$
Full conditional for $\sigma^2_\alpha$

This can be written as

$$f(\sigma^2_\alpha | y, \mu, \alpha, \delta, \sigma^2_e) \propto f(y | \sigma^2_\alpha, \mu, \alpha, \delta, \sigma^2_e) f(\sigma^2_\alpha, \mu, \alpha, \delta, \sigma^2_e)$$

But, can see that

$$f(y | \sigma^2_\alpha, \mu, \alpha, \delta, \sigma^2_e) \propto f(y | \mu, \alpha, \delta, \sigma^2_e)$$

So,

$$f(\sigma^2_\alpha | y, \mu, \alpha, \delta, \sigma^2_e) \propto f(\sigma^2_\alpha, \mu, \alpha, \delta, \sigma^2_e)$$

Note that $\sigma^2_\alpha$ appears only in $f(\alpha | \sigma^2_\alpha)$ and $f(\sigma^2_\alpha)$:

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

and

$$f(\sigma^2_\alpha) \propto (\sigma^2_\alpha)^{-(\nu_\alpha+2)/2} \exp\left\{ \frac{\nu_\alpha S^2_\alpha}{2\sigma^2_\alpha} \right\}$$
Full conditional for $\sigma^2_\alpha$

This can be written as

$$f(\sigma^2_\alpha | y, \mu, \alpha, \delta, \sigma^2_\epsilon) \propto f(y | \sigma^2_\alpha, \mu, \alpha, \delta, \sigma^2_\epsilon) f(\sigma^2_\alpha, \mu, \alpha, \delta, \sigma^2_\epsilon)$$

But, can see that

$$f(y | \sigma^2_\alpha, \mu, \alpha, \delta, \sigma^2_\epsilon) \propto f(y | \mu, \alpha, \delta, \sigma^2_\epsilon)$$

So,

$$f(\sigma^2_\alpha | y, \mu, \alpha, \delta, \sigma^2_\epsilon) \propto f(\sigma^2_\alpha, \mu, \alpha, \delta, \sigma^2_\epsilon)$$

Note that $\sigma^2_\alpha$ appears only in $f(\alpha | \sigma^2_\alpha)$ and $f(\sigma^2_\alpha)$:

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

and

$$f(\sigma^2_\alpha) \propto (\sigma^2_\alpha)^{-(\nu_\alpha+2)/2} \exp\left\{\frac{\nu_\alpha S^2_\alpha}{2\sigma^2_\alpha}\right\}$$
Full conditional for $\sigma^2_\alpha$

Combining these two densities gives:

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

So,

$$(\sigma^2_\alpha | y, \mu, \alpha, \delta, \sigma^2_\epsilon) \sim \tilde{\nu}_\alpha \tilde{S}^2_\alpha \chi^{-2}_{\tilde{\nu}_\alpha}$$

where

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

and

$$\tilde{S}^2_\alpha = \frac{\alpha' \alpha + \nu_\alpha S^2_\alpha}{\tilde{\nu}_\alpha}$$
Full conditional for $\sigma^2_\alpha$

Combining these two densities gives:

$$f(\sigma^2_\alpha \mid y, \mu, \alpha, \delta, \sigma^2_\theta) \propto (\sigma^2_\alpha)^{-(k+\nu_\alpha+2)/2} \exp\left\{ \frac{\alpha'\alpha + \nu_\alpha S^2_\alpha}{2\sigma^2_\alpha} \right\}$$

So,

$$(\sigma^2_\alpha \mid y, \mu, \alpha, \delta, \sigma^2_\theta) \sim \tilde{\nu}_\alpha \tilde{S}^2_\alpha \chi_{\tilde{\nu}_\alpha}^{-2}$$

where

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

and

$$\tilde{S}^2_\alpha = \frac{\alpha'\alpha + \nu_\alpha S^2_\alpha}{\tilde{\nu}_\alpha}$$
Hyper parameter: $S^2_\alpha$

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

$$\sigma^2_\alpha = \frac{V_a}{\sum_j 2p_jq_j}$$

So, we take

$$S^2_\alpha = \frac{(\nu_\alpha - 2)V_a}{\nu_\alpha k(1 - \pi)2pq}$$
Hyper parameter: $S^2_\alpha$

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

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

So, we take

$$S^2_\alpha = \frac{(\nu_\alpha - 2)V_a}{\nu_\alpha k(1 - \pi)2pq}$$
Hyper parameter: $S^2_\alpha$

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

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

So, we take

$$S^2_\alpha = \frac{(\nu_\alpha - 2) V_a}{\nu_\alpha k (1 - \pi) 2pq}$$
Full conditional for $\pi$

Using Bayes’ theorem,

$$f(\pi | \delta, \mu, \alpha, \sigma_\alpha^2, \sigma_\theta^2, y) \propto f(y | \pi, \delta, \mu, \alpha, \sigma_\alpha^2, \sigma_\theta^2) f(\pi, \delta, \mu, \alpha, \sigma_\alpha^2, \sigma_\theta^2)$$

But,

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

Thus,

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

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

Using Bayes’ theorem,

$$f(\pi|\delta, \mu, \alpha, \sigma^2_\alpha, \sigma^2_\theta, y) \propto f(y|\pi, \delta, \mu, \alpha, \sigma^2_\alpha, \sigma^2_\theta)f(\pi, \delta, \mu, \alpha, \sigma^2_\alpha, \sigma^2_\theta)$$

But,

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

Thus,

$$f(\pi|\delta, \mu, \alpha, \sigma^2_\alpha, \sigma^2_\theta, y) = \pi^{(k-m)}(1 - \pi)^m$$

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

Full conditional for $\pi$

Using Bayes’ theorem,

$$f(\pi | \delta, \mu, \alpha, \sigma^2_\alpha, \sigma^2_\theta, y) \propto f(y | \pi, \delta, \mu, \alpha, \sigma^2_\alpha, \sigma^2_\theta) f(\pi, \delta, \mu, \alpha, \sigma^2_\alpha, \sigma^2_\theta)$$

But,

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

Thus,

$$f(\pi | \delta, \mu, \alpha, \sigma^2_\alpha, \sigma^2_\theta, y) = \pi^{(k-m)}(1 - \pi)^m$$

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

Using Bayes’ theorem,

$$f(\pi | \delta, \mu, \alpha, \sigma^2_{\alpha}, \sigma^2_{\epsilon}, y) \propto f(y | \pi, \delta, \mu, \alpha, \sigma^2_{\alpha}, \sigma^2_{\epsilon}) f(\pi, \delta, \mu, \alpha, \sigma^2_{\alpha}, \sigma^2_{\epsilon})$$

But,

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

Thus,

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

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

Using Bayes’ theorem,

$$f(\pi | \delta, \mu, \alpha, \sigma_\alpha^2, \sigma_e^2, y) \propto f(y | \pi, \delta, \mu, \alpha, \sigma_\alpha^2, \sigma_e^2) f(\pi, \delta, \mu, \alpha, \sigma_\alpha^2, \sigma_e^2)$$

But,

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

Thus,

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

where $m = \delta' \delta$, and $k$ is the number of markers. Thus, $\pi$ is sampled from a beta distribution with $a = k - m + 1$ and $b = m + 1$. 
BayesC$\pi$ with Unknown $S^2_\alpha$

- Prior for $S^2_\alpha$: Gamma(a, b)

$$f(S^2_\alpha|a, b) \propto b^a(S^2_\alpha)^{a-1} \exp\{-bS^2_\alpha\}$$

- Using Bayes theorem,

$$f(S^2_\alpha|\delta, \mu, \alpha, \sigma^2_\alpha, \sigma^2_\varepsilon, y) \propto f(y|S^2_\alpha, \sigma^2_\alpha, \ldots)f(S^2_\alpha, \sigma^2 \ldots)$$

- Given $\mu, \alpha, \sigma^2_\varepsilon$, $f(y|S^2_\alpha, \sigma^2_\alpha, \ldots)$ does not depend on $S^2_\alpha$.

- In $f(S^2_\alpha, \sigma^2 \ldots)$, $S^2_\alpha$ is only in $f(S^2_\alpha|a, b)$ and $f(\sigma^2_\alpha|S^2_\alpha, \nu_\alpha)$
BayesC with Unknown $S^2_{\alpha}$

- Prior for $S^2_{\alpha}$: Gamma(a, b)
  
  $$f(S^2_{\alpha}|a, b) \propto b^a(S^2_{\alpha})^{a-1} \exp\{-bS^2_{\alpha}\}$$

- Using Bayes theorem,
  
  $$f(S^2_{\alpha}|\delta, \mu, \alpha, \sigma^2_{\alpha}, \sigma^2_e, y) \propto f(y|S^2_{\alpha}, \sigma^2_{\alpha}, \ldots)f(S^2_{\alpha}, \sigma^2 \ldots)$$

  - Given $\mu, \alpha$, and $\sigma^2_e$, $f(y|S^2_{\alpha}, \sigma^2_{\alpha}, \ldots)$ does not depend on $S^2_{\alpha}$.
  - In $f(S^2_{\alpha}, \sigma^2 \ldots)$, $S^2_{\alpha}$ is only in $f(S^2_{\alpha}|a, b)$ and $f(\sigma^2_{\alpha}|S^2_{\alpha}, \nu_{\alpha})$.
BayesC with Unknown $S^2_{\alpha}$

- Prior for $S^2_{\alpha}$: Gamma(a,b)

\[
f(S^2_{\alpha}| a, b) \propto b^a (S^2_{\alpha})^{a-1} \exp\{-bS^2_{\alpha}\}
\]

- Using Bayes theorem,

\[
f(S^2_{\alpha}| \delta, \mu, \alpha, \sigma^2_{\alpha}, \sigma^2_e, y) \propto f(y| S^2_{\alpha}, \sigma^2_{\alpha}, \ldots) f(S^2_{\alpha}, \sigma^2 \ldots)
\]

- Given $\mu, \alpha, \text{ and } \sigma^2_e$, $f(y| S^2_{\alpha}, \sigma^2_{\alpha}, \ldots)$ does not depend on $S^2_{\alpha}$.

- In $f(S^2_{\alpha}, \sigma^2 \ldots)$, $S^2_{\alpha}$ is only in $f(S^2_{\alpha}| a, b)$ and $f(\sigma^2_{\alpha}| S^2_{\alpha}, \nu_{\alpha})$.
BayesC$\pi$ with Unknown $S^2_{\alpha}$

- Prior for $S^2_{\alpha}$: Gamma(a,b)

\[
f(S^2_{\alpha}|a, b) \propto b^a(S^2_{\alpha})^{a-1} \exp\{-bS^2_{\alpha}\}
\]

- Using Bayes theorem,

\[
f(S^2_{\alpha}|\delta, \mu, \alpha, \sigma^2_{\alpha}, \sigma^2_e, y) \propto f(y|S^2_{\alpha}, \sigma^2_{\alpha}, \ldots)f(S^2_{\alpha}, \sigma^2 \ldots)
\]

- Given $\mu, \alpha$, and $\sigma^2_e$, $f(y|S^2_{\alpha}, \sigma^2_{\alpha}, \ldots)$ does not depend on $S^2_{\alpha}$.

- In $f(S^2_{\alpha}, \sigma^2 \ldots)$, $S^2_{\alpha}$ is only in $f(S^2_{\alpha}|a, b)$ and $f(\sigma^2_{\alpha}|S^2_{\alpha}, \nu_{\alpha})$
BayesC\(\pi\) with Unknown \(S^2_\alpha\)

- Prior for \(S^2_\alpha\): Gamma\((a,b)\)
  \[
f(S^2_\alpha|a, b) \propto b^a(S^2_\alpha)^{a-1} \exp\{-bS^2_\alpha\}
\]

- Prior for \(\sigma^2_\alpha\):
  \[
f(\sigma^2_\alpha) \propto (\sigma^2_\alpha)^{-(\nu+2)/2} \exp\{\frac{\nu S^2_\alpha}{2\sigma^2_\alpha}\}
\]

- Combining these gives:
  \[
f(S^2_\alpha|\sigma^2_\alpha, y, \ldots) \propto S^2_\alpha^{(a-1+\nu/2)} \exp\{-S^2_\alpha\left(\frac{\nu}{2\sigma^2_\alpha} + b\right)\}
\]
BayesC$_\pi$ with Unknown $S^2_{\alpha}$

- Prior for $S^2_{\alpha}$: Gamma(a,b)

\[
f(S^2_{\alpha} \mid a, b) \propto b^a (S^2_{\alpha})^{a-1} \exp\{-bS^2_{\alpha}\}
\]

- Prior for $\sigma^2_{\alpha}$:

\[
f(\sigma^2_{\alpha}) \propto (\sigma^2_{\alpha})^{-(\nu_{\alpha}+2)/2} \exp\left\{\frac{\nu_{\alpha}S^2_{\alpha}}{2\sigma^2_{\alpha}}\right\}
\]

- Combining these gives:

\[
f(S^2_{\alpha} \mid \sigma^2_{\alpha}, y, \ldots) \propto S^2_{\alpha}^{(a-1+\nu/2)} \exp\{-S^2_{\alpha}(\frac{\nu_{\alpha}}{2\sigma^2_{\alpha}} + b)\}
\]
BayesC$^\pi$ with Unknown $S^2_{\alpha}$

- Prior for $S^2_{\alpha}$: Gamma(a,b)
  \[ f(S^2_{\alpha}|a, b) \propto b^a(S^2_{\alpha})^{a-1} \exp\{-bS^2_{\alpha}\} \]

- Prior for $\sigma^2_{\alpha}$:
  \[ f(\sigma^2_{\alpha}) \propto (\sigma^2_{\alpha})^{-(\nu_{\alpha}+2)/2} \exp\{\frac{\nu_{\alpha}S^2_{\alpha}}{2\sigma^2_{\alpha}}\} \]

- Combining these gives:
  \[ f(S^2_{\alpha}|\sigma^2_{\alpha}, y, \ldots) \propto S^2_{\alpha}^{(a-1+\nu/2)} \exp\{-S^2_{\alpha}(\frac{\nu_{\alpha}}{2\sigma^2_{\alpha}} + b)\} \]
BayesC$\pi$ with Unknown $S^2_\alpha$

So, $f(S^2_\alpha|a, b)$ is Gamma($a^*, b^*$), where

$$a^* = a + \nu_\alpha / 2$$

and

$$b^* = b + \frac{\nu_\alpha}{2\sigma^2_\alpha}$$
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
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
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
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

<table>
<thead>
<tr>
<th>$\pi$</th>
<th>$S^2$</th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.995</td>
<td>0.2</td>
<td>0.91 (0.009)</td>
</tr>
<tr>
<td>0.8</td>
<td>0.2</td>
<td>0.86 (0.009)</td>
</tr>
<tr>
<td>0.0</td>
<td>0.2</td>
<td>0.80 (0.013)</td>
</tr>
<tr>
<td>0.995</td>
<td>2.0</td>
<td>0.90 (0.007)</td>
</tr>
<tr>
<td>0.8</td>
<td>2.0</td>
<td>0.77 (0.009)</td>
</tr>
<tr>
<td>0.0</td>
<td>2.0</td>
<td>0.35 (0.022)</td>
</tr>
</tbody>
</table>
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 $\pi = 0.5$
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 $\pi = 0.5$
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 $\pi = 0.5$
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 $\pi = 0.5$
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**

<table>
<thead>
<tr>
<th>$N$</th>
<th>$Q$</th>
<th>$\pi$</th>
<th>$\hat{\pi}$</th>
<th>$\text{Corr}(g, \hat{g})$</th>
<th>$\text{Bayes-C}_\pi$</th>
<th>$\text{Bayes-B}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000</td>
<td>10</td>
<td>0.995</td>
<td>0.994</td>
<td>0.995</td>
<td>0.995</td>
<td>0.937</td>
</tr>
<tr>
<td>2000</td>
<td>200</td>
<td>0.90</td>
<td>0.899</td>
<td>0.866</td>
<td>0.866</td>
<td>0.834</td>
</tr>
<tr>
<td>2000</td>
<td>1900</td>
<td>0.05</td>
<td>0.202</td>
<td>0.613</td>
<td>0.613</td>
<td>0.571</td>
</tr>
<tr>
<td>4000</td>
<td>1900</td>
<td>0.05</td>
<td>0.096</td>
<td>0.763</td>
<td>0.763</td>
<td>0.722</td>
</tr>
</tbody>
</table>
Simulation III

- **Genotypes:** 50k SNPs from 1086 Purebred Angus animals, ISU
- **Phenotypes:**
  - QTL simulated from 50 randomly sampled SNPs
  - Substitution effect sampled from $N(0, \sigma^2_\alpha)$
  - $\sigma^2_\alpha = \frac{\sigma^2_g}{502 \rho q}$
  - $h^2 = 0.25$
- QTL were included in the marker panel
- Marker effects were estimated for 50k SNPs
Simulation III

- Genotypes: 50k SNPs from 1086 Purebred Angus animals, ISU

- Phenotypes:
  - QTL simulated from 50 randomly sampled SNPs
  - Substitution effect sampled from $\mathcal{N}(0, \sigma^2_\alpha)$
  - $\sigma^2_\alpha = \frac{\sigma^2_g}{502pq}$
  - $h^2 = 0.25$

- QTL were included in the marker panel
- Marker effects were estimated for 50k SNPs
Simulation III

- Genotypes: 50k SNPs from 1086 Purebred Angus animals, ISU
- Phenotypes:
  - QTL simulated from 50 randomly sampled SNPs
  - Substitution effect sampled from $N(0, \sigma^2_\alpha)$
    - $\sigma^2_\alpha = \frac{\sigma^2_g}{50pq}$
    - $h^2 = 0.25$
- QTL were included in the marker panel
- Marker effects were estimated for 50k SNPs
Simulation III

- **Genotypes:** 50k SNPs from 1086 Purebred Angus animals, ISU

- **Phenotypes:**
  - QTL simulated from 50 randomly sampled SNPs
  - Substitution effect sampled from $N(0, \sigma^2_\alpha)$
    $$\sigma^2_\alpha = \frac{\sigma^2_g}{502\rho \overline{pq}}$$
  - $h^2 = 0.25$

- QTL were included in the marker panel
- Marker effects were estimated for 50k SNPs
Simulation III

- Genotypes: 50k SNPs from 1086 Purebred Angus animals, ISU

- Phenotypes:
  - QTL simulated from 50 randomly sampled SNPs
  - substitution effect sampled from N(0, $\sigma_\alpha^2$)
  - $\sigma_\alpha^2 = \frac{\sigma_g^2}{502pq}$
  - $h^2 = 0.25$

- QTL were included in the marker panel
- Marker effects were estimated for 50k SNPs
Simulation III

- **Genotypes:** 50k SNPs from 1086 Purebred Angus animals, ISU
- **Phenotypes:**
  - QTL simulated from 50 randomly sampled SNPs
  - Substitution effect sampled from $N(0, \sigma^2_{\alpha})$
  - $\sigma^2_{\alpha} = \frac{\sigma^2_g}{502pq}$
  - $h^2 = 0.25$
- QTL were included in the marker panel
- Marker effects were estimated for 50k SNPs
Simulation III

- Genotypes: 50k SNPs from 1086 Purebred Angus animals, ISU

- Phenotypes:
  - QTL simulated from 50 randomly sampled SNPs
  - substitution effect sampled from $N(0, \sigma^2_\alpha)$
  - $\sigma^2_\alpha = \frac{\sigma^2_g}{502\rho q}$
  - $h^2 = 0.25$

- QTL were included in the marker panel

- Marker effects were estimated for 50k SNPs
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
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
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
Correlations between $g_i$ and $\hat{g}_i$ estimated from 3 replications

<table>
<thead>
<tr>
<th>$\pi$</th>
<th>Bayes-B</th>
<th>Bayes-C</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.999</td>
<td>0.86</td>
<td>0.86</td>
</tr>
<tr>
<td>0.25</td>
<td>0.70</td>
<td>0.26</td>
</tr>
</tbody>
</table>

BayesC$\pi$:

- $\hat{\pi} = 0.999$
- Correlation = 0.86
## Results

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

<table>
<thead>
<tr>
<th>( \pi )</th>
<th>Bayes-B</th>
<th>Bayes-C</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.999</td>
<td>0.86</td>
<td>0.86</td>
</tr>
<tr>
<td>0.25</td>
<td>0.70</td>
<td>0.26</td>
</tr>
</tbody>
</table>

**BayesC\(\pi\):**

- \( \hat{\pi} = 0.999 \)
- Correlation = 0.86
Correlations between $g_i$ and $\hat{g}_i$ estimated from 3 replications

<table>
<thead>
<tr>
<th>$\pi$</th>
<th>Bayes-B</th>
<th>Bayes-C</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.999</td>
<td>0.86</td>
<td>0.86</td>
</tr>
<tr>
<td>0.25</td>
<td>0.70</td>
<td>0.26</td>
</tr>
</tbody>
</table>

**BayesC$\pi$:**
- $\hat{\pi} = 0.999$
- Correlation $= 0.86$
Results

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

<table>
<thead>
<tr>
<th>( \pi )</th>
<th>Correlation</th>
<th>Bayes-B</th>
<th>Bayes-C</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.999</td>
<td>0.86</td>
<td>0.86</td>
<td></td>
</tr>
<tr>
<td>0.25</td>
<td>0.70</td>
<td>0.26</td>
<td></td>
</tr>
</tbody>
</table>

BayesC\( \pi \):

- \( \hat{\pi} = 0.999 \)
- Correlation = 0.86