Introduction to Bayesian Linear Regression, Posterior Inference, and MCMC

Peter Sørensen

Introduction (Applications and motivation in genomics and healthcare)

Bayesian Linear Regression Models provide a flexible statistical framework for modeling complex biological and healthcare data. They support key applications such as:

- **Genome-wide association studies (GWAS)** and **fine-mapping** of causal variants
- **Polygenic risk scoring (PRS)** for predicting complex traits and disease risk
- **Gene and pathway enrichment analyses** to test biological hypotheses
- **Integrative multi-omics modeling** across the genome, transcriptome, epigenome, and proteome
- Applications to **registry-based healthcare data**, enabling population-level **risk prediction** and **disease modeling**

Introduction (Definition and advantages of BLR)

- **Bayesian Linear Regression (BLR)** extends classical regression by incorporating **prior information** and producing **posterior distributions** over model parameters.
- **Advantages:**
  - Handles **high-dimensional** and **small-sample** problems.
  - Provides **full uncertainty quantification**.
  - Enables **regularization** and integration of **prior biological knowledge**.

Overview (Structure of the session)

- **Classical Linear Regression**
  - Model, inference, and limitations
- **Bayesian Linear Regression**
  - Motivation, priors, and posteriors
  - Conditional posteriors and inference
- **Computation and Applications**
  - MCMC and Gibbs sampling
  - Diagnostics and R implementation

Classical Linear Regression

The standard linear regression model, which assumes that the observed outcomes can be expressed as a linear combination of predictors plus random noise:

$$y = X\beta + e, \quad e \sim \mathcal{N}(0, \sigma^2 I_n)$$

- y: $n \times 1$ vector of observed outcomes
- X: $n \times p$ design matrix of predictors
- β: $p \times 1$ vector of unknown coefficients
- e: Gaussian noise with mean 0 and variance σ^2

Estimation

Given the linear model, we can estimate the regression coefficients and residual variance using the method of ordinary least squares (OLS), which minimizes the sum of squared residuals:

Regression effects:

$$\hat{\beta} = (X^\top X)^{-1} X^\top y$$

Residual variance:

$$\hat{\sigma}^2 = \frac{1}{n-p} \sum_i (y_i - x_i^\top \hat{\beta})^2$$

Inference via standard errors and t-tests, confidence intervals, and prediction intervals.

Limitations

While ordinary least squares estimation is simple and widely used, it has several important limitations that motivate the use of regularized or Bayesian approaches:

- No explicit control over **effect size distribution**
- Sensitive when **collinearity** is high
- **Not identifiable** when $p > n$
- Uncertainty largely **asymptotic** unless normality assumptions hold

Why Bayesian Linear Regression?

The Bayesian framework extends linear regression by incorporating prior beliefs about the model parameters and updating them with observed data through Bayes' theorem:

- Combines **likelihood** and **prior** to form the **posterior**.
- Priors express beliefs about **effect sizes**:
  - Normal $\rightarrow$ many small effects
  - Spike-and-slab $\rightarrow$ sparse effects
- Acts as a **regularizer**:
  - Shrinks small/noisy effects toward 0
  - Preserves large, important effects
- **Stable when** $p > n$ due to prior information.
- Provides **full posterior distributions** for β and σ^2.

Bayesian Linear Regression with Gaussian Priors

Bayesian linear regression starts with the same model structure as classical linear regression.

$$y = X\beta + e, \quad e \sim \mathcal{N}(0, \sigma^2 I_n)$$

- y: $n \times 1$ vector of observed outcomes
- X: $n \times p$ design matrix of predictors
- β: $p \times 1$ vector of unknown coefficients
- e: Gaussian noise with mean 0 and variance σ^2

Likelihood in Bayesian Linear Regression

Because the residuals are assumed to be Gaussian,

$$e \sim \mathcal{N}(0, \sigma^2 I_n)$$

it follows that the **response vector** y follows a multivariate normal distribution:

$$y = X\beta + e \quad \Rightarrow \quad y \sim \mathcal{N}(X\beta, \sigma^2 I_n)$$

This defines the **likelihood**, i.e. the probability of the observed data given the model parameters β and σ^2:

$$p(y \mid X, \beta, \sigma^2) = \mathcal{N}(y \mid X\beta, \sigma^2 I_n)$$

Introducing Priors

In Bayesian linear regression, we specify **prior distributions** to express our beliefs about the model parameters before observing the data.

A common **conjugate prior** for the regression coefficients is

$$\beta \mid \sigma_b^2 \sim \mathcal{N}(0, \sigma_b^2 I_p)$$

This expresses the belief that most effect sizes are small and centered around zero — consistent with the **polygenic assumption** often used in genetics.

Using **conjugate priors** ensures that the **posterior distributions** remain in the same family as the priors (e.g., scaled inverse-chi-squared for variance parameters), enabling **closed-form Gibbs sampling updates**.

Role of the Prior Variance σ_b^2

The parameter σ_b^2 acts as a **shrinkage** (or **regularization**) parameter:

- Small $\sigma_b^2 \rightarrow$ stronger shrinkage toward zero
- Large $\sigma_b^2 \rightarrow$ weaker shrinkage, allowing larger effects

It determines the **strength of regularization** and is often treated as an **unknown hyperparameter** to be estimated from the data.

Priors on Variance Components

We also place priors on the variance components to complete the hierarchical model:

$$\sigma_b^2 \mid S_b, v_b \sim S_b \, \chi^{-2}(v_b), \quad \sigma^2 \mid S, v \sim S \, \chi^{-2}(v)$$

- S_b and v_b are hyperparameters that control the prior distribution for the **effect size variance** σ_b^2.
- S and v are hyperparameters for the **residual variance** σ^2.

These scaled inverse-chi-squared priors ensure conjugacy, enabling **closed-form updates** for the variance parameters in Gibbs sampling.

Typical choices:

- Small degrees of freedom (e.g., $v_b = v = 4$) give weakly informative, heavy-tailed priors.
- Scale parameters S_b and S are often set based on expected variance magnitudes (e.g., empirical estimates).

Posterior Distribution

In Bayesian linear regression, we combine the **likelihood** and **prior distributions** using **Bayes' rule** to obtain the **joint posterior**:

$$p(\beta, \sigma_b^2, \sigma^2 \mid y) \propto p(y \mid \beta, \sigma^2)\, p(\beta \mid \sigma_b^2)\, p(\sigma_b^2)\, p(\sigma^2)$$

The posterior distribution represents all **updated knowledge** about the unknown parameters after observing the data.
It serves as the foundation for computing **posterior means**, **credible intervals**, and **predictions**.

In practice, the posterior is often too complex to evaluate directly, so we use **sampling-based methods** such as Gibbs sampling to approximate it.

Conjugacy and Gibbs Sampling

With **conjugate priors**, each parameter's **full conditional distribution** has a closed-form solution.
This makes **Gibbs sampling** a natural and efficient inference method.

- Parameters are updated one at a time, each from its conditional posterior.
- The resulting Markov chain explores the **joint posterior** of $(\beta, \sigma_b^2, \sigma^2)$.

Gibbs sampling thus provides an easy way to approximate the full posterior in Bayesian linear regression.

Full Conditional for β

Given σ^2, σ_b^2, and the data y, the regression coefficients have a **multivariate normal** conditional posterior:

$$\beta \mid \sigma^2, \sigma_b^2, y \sim \mathcal{N}(\mu_\beta, \Sigma_\beta)$$

where

$$\Sigma_\beta = \left(\frac{X^\top X}{\sigma^2} + \frac{I}{\sigma_b^2} \right)^{-1}, \quad \mu_\beta = \Sigma_\beta \frac{X^\top y}{\sigma^2}$$

This distribution represents our **updated belief** about β after observing the data, while holding σ_b^2 and σ^2 fixed.

Comparison to Classical OLS

In classical regression, the OLS estimator is

$$\hat{\beta}_{\text{OLS}} = (X^\top X)^{-1} X^\top y$$

The estimate of β is **independent of** σ^2, since σ^2 only scales the likelihood, not its maximum.

In Bayesian regression, σ^2 appears explicitly in the posterior:

$$\Sigma_\beta = \left(\frac{X^\top X}{\sigma^2} + \frac{I}{\sigma_b^2} \right)^{-1}, \quad \mu_\beta = \Sigma_\beta \frac{X^\top y}{\sigma^2}$$

The term $\frac{I}{\sigma_b^2}$ introduces **shrinkage**, regularizing estimates and stabilizing inference especially when $p > n$ or predictors are highly correlated. Thus, the Bayesian posterior mean is a **regularized, uncertainty-aware generalization** of OLS.

Full Conditional for β_j

Instead of sampling β jointly, we can update each coefficient β_j **one at a time**, holding all others fixed efficient.

Let X_j be the jth column of X and define the **partial residual**:

$$r_j = y - X_{-j}\beta_{-j}$$

Then the conditional posterior for β_j is univariate normal:

$$\beta_j \mid D \sim \mathcal{N}\left(\frac{X_j^\top r_j}{X_j^\top X_j + \sigma^2/\sigma_b^2},\ \frac{\sigma^2}{X_j^\top X_j + \sigma^2/\sigma_b^2}\right)$$

This corresponds to a **regularized least-squares update**. Residual updates **avoid matrix inversion**, scale to high dimensions, and extend naturally to **sparse (spike-and-slab)** models.

Full Conditional for σ_b^2

The conditional distribution of the **prior variance** σ_b^2, given β and the hyperparameters, is a **scaled inverse-chi-squared**:

$$\sigma_b^2 \mid \beta \sim \tilde{S}_b \, \chi^{-2}(\tilde{v}_b)$$

where

$$\tilde{v}_b = v_b + p, \quad \tilde{S}_b = \frac{\beta^\top \beta + v_b S_b}{\tilde{v}_b}$$

At each Gibbs iteration, σ_b^2 is sampled directly given β. This update reflects our revised belief about the **variability of effect sizes** after observing the current posterior draw of β.

Full Conditional for σ^2

The conditional distribution of the **residual variance** σ^2, given β and the data, is also **scaled inverse-chi-squared**:

$$\sigma^2 \mid \beta, y \sim \tilde{S}\,\chi^{-2}(\tilde{v})$$

where

$$\tilde{v} = v + n, \quad \tilde{S} = \frac{(y - X\beta)^\top (y - X\beta) + vS}{\tilde{v}}$$

At each Gibbs iteration, σ^2 is sampled directly given β.
This captures our updated belief about the **residual variability** after accounting for the current linear predictor $X\beta$.

Gibbs Sampling: Motivation

Bayesian inference often involves **complex posteriors** that lack closed-form solutions. To approximate these, we use **Markov Chain Monte Carlo (MCMC)** methods.

MCMC builds a **Markov chain** whose stationary distribution is the target posterior. Once the chain has **converged**, its samples can be used to estimate:

- Posterior means, variances, and credible intervals
- Predictive distributions
- Other functions of interest

Among MCMC algorithms, the **Gibbs sampler** is especially useful when all **full conditional distributions** are available in **closed form**.

Gibbs Sampling: The Algorithm

For Bayesian linear regression with conjugate priors, the joint posterior is:

$$p(\beta, \sigma_b^2, \sigma^2 \mid y) \propto p(y \mid \beta, \sigma^2)\, p(\beta \mid \sigma_b^2)\, p(\sigma_b^2)\, p(\sigma^2)$$

We iteratively draw from the following **full conditionals**:

1. Sample $\beta \mid \sigma_b^2, \sigma^2, y$
2. Sample $\sigma_b^2 \mid \beta$
3. Sample $\sigma^2 \mid \beta, y$

Each step updates one parameter given the latest values of the others. Repeating this sequence yields samples from the **joint posterior** $p(\beta, \sigma_b^2, \sigma^2 \mid y)$.

Because each conditional is **standard** (Normal or scaled inverse-χ^2), Gibbs sampling is both **efficient** and **easy to implement**.

Posterior Summaries

After running the Gibbs sampler, we obtain a sequence of posterior draws $\{\theta^{(t)}\}_{t=1}^{T}$ for parameters such as β_j, σ^2, or σ_b^2, where T denotes the total number of MCMC iterations (after burn-in).
We can summarize the posterior distribution using:

- **Posterior mean**

$$\mathbb{E}[\theta \mid y] \approx \frac{1}{T}\sum_{t=1}^{T}\theta^{(t)}$$

- **Posterior median**: the median of $\{\theta^{(t)}\}$
- **95% credible interval**: the interval between the 2.5th and 97.5th percentiles of $\{\theta^{(t)}\}$

These summaries describe the most probable values of θ and their associated uncertainty after combining data with prior beliefs.

Estimating Uncertainty

Bayesian inference provides **full posterior distributions**, not just point estimates. Uncertainty is quantified directly from the posterior samples:

- **Posterior standard deviation**

$$\mathrm{SD}(\theta \mid y) \approx \sqrt{\frac{1}{T-1}\sum_{t=1}^{T}(\theta^{(t)} - \bar{\theta})^2}$$

The **width** of the credible interval reflects this uncertainty. Parameters with broader posteriors are estimated with less precision, and the degree of uncertainty depends on both the data and the prior.

Posterior Prediction

Given a new observation x_{new}, we can predict using posterior draws:

1. Compute predicted means for each sample:

$$\hat{y}_{\text{new}}^{(t)} = x_{\text{new}}^{\top} \beta^{(t)}$$

2. Add residual uncertainty:

$$y_{\text{new}}^{(t)} \sim \mathcal{N}\left(x_{\text{new}}^{\top} \beta^{(t)},\ \sigma^{2(t)}\right)$$

The resulting samples $\{y_{\text{new}}^{(t)}\}$ form a **posterior predictive distribution**, from which we can derive **predictive intervals** and evaluate **predictive accuracy**.

Model Checking and Hypothesis Testing

Posterior samples enable rich **model diagnostics** and **hypothesis testing**:

- **Posterior probability of an event**

$$\Pr(\beta_j \neq 0 \mid y) \approx \frac{1}{T} \sum_{t=1}^{T} \mathbf{1}\left(\beta_j^{(t)} \neq 0\right)$$

- **Posterior predictive checks**
  Simulate new datasets using posterior draws and compare them to the observed data to assess model fit.
- **Model comparison**
  Bayes factors and marginal likelihoods can be approximated to formally test or compare competing models.

These tools extend Bayesian inference beyond estimation to **model validation**, **uncertainty quantification**, and **decision-making**.

Convergence Diagnostics

Before interpreting MCMC results, we must check that the Gibbs sampler has **converged** to the target posterior distribution.

Convergence diagnostics assess whether the Markov chain has reached its **stationary distribution** and is producing valid samples.

Two basic strategies are:

- **Burn-in** – Discard early iterations (e.g., first 1000) to remove dependence on starting values.
- **Thinning** – Keep every k-th sample to reduce autocorrelation.

These steps improve sample quality and ensure reliable posterior summaries.

Trace Plots

A simple yet powerful diagnostic is the **trace plot**,
showing sampled parameter values $\theta^{(t)}$ over iterations t.

- A **converged chain** fluctuates around a stable mean — no trend or drift.
- Multiple chains from different starting points should **overlap** and **mix well**.

Trace plots help detect: - Lack of stationarity (upward/downward trends) - Poor mixing or multimodality - Burn-in issues

Visual inspection is often the **first step** in assessing convergence.

Autocorrelation

Samples from a Gibbs sampler are **correlated**, especially for tightly coupled parameters.
The **autocorrelation function (ACF)** quantifies dependence across lags k:

$$\hat{\rho}_k = \frac{\sum_{t=1}^{T-k}(\theta^{(t)} - \bar{\theta})(\theta^{(t+k)} - \bar{\theta})}{\sum_{t=1}^{T}(\theta^{(t)} - \bar{\theta})^2}$$

- High $\hat{\rho}_k \rightarrow$ slow mixing and fewer effective samples
- Low $\hat{\rho}_k \rightarrow$ better mixing and faster convergence

Reducing autocorrelation may require **more iterations**, **reparameterization**, or **thinning** the chain.

Effective Sample Size (ESS)

Autocorrelation reduces the number of *independent* samples obtained.

The **effective sample size (ESS)** adjusts for this:

$$\text{ESS}(\theta) = \frac{T}{1 + 2\sum_{k=1}^{K} \hat{\rho}_k}$$

- Small ESS $\rightarrow$ chain is highly correlated, less informative
- Rule of thumb: ESS > 100 per parameter for stable inference

ESS provides a quantitative measure of **sampling efficiency** and helps determine whether more iterations are needed.

Gelman–Rubin Diagnostic ($\hat{R}$)

When running multiple chains, the **Gelman–Rubin statistic** compares **between-chain** and **within-chain** variability.

For m chains with T iterations each:

$$W = \frac{1}{m}\sum_{i=1}^{m} s_i^2, \quad B = \frac{T}{m-1}\sum_{i=1}^{m}(\bar{\theta}_i - \bar{\theta})^2$$

The potential scale reduction factor:

$$\hat{R} = \sqrt{\frac{\hat{V}}{W}}, \quad \hat{V} = \frac{T-1}{T}W + \frac{1}{T}B$$

Values of $\hat{R}$ close to 1 indicate convergence, whereas $\hat{R} > 1.1$ suggests that the chains have not yet converged.

Geweke Diagnostic

The **Geweke test** checks whether early and late portions of a single chain have the same mean, indicating **stationarity**.

$$Z = \frac{\bar{\theta}_A - \bar{\theta}_B}{\sqrt{\text{Var}(\bar{\theta}_A) + \text{Var}(\bar{\theta}_B)}}$$

Typically:

- Segment A = first 10% of the chain
- Segment B = last 50% of the chain

Under convergence, $Z \sim \mathcal{N}(0, 1)$.

- $|Z| \leq 2 \rightarrow$ chain likely stationary
- $|Z| > 2 \rightarrow$ potential non-convergence

These diagnostics ensure that posterior summaries reflect the **true target distribution**.

Motivation for the Spike-and-Slab Prior

In standard Bayesian linear regression,

$$\beta_j \sim \mathcal{N}(0, \sigma_b^2)$$

This **Gaussian (shrinkage) prior** assumes that all predictors have small effects but does **not allow exact zeros**, limiting its ability to perform variable selection.

The **spike-and-slab prior** addresses this by mixing two components:

- A **spike** at zero $\rightarrow$ represents excluded predictors
- A **slab** (wide normal) $\rightarrow$ represents active predictors

This formulation produces **sparse**, interpretable models that automatically select the most relevant variables.

Spike-and-Slab Bayesian Linear Regression

As in classical Bayesian linear regression, the outcome is modeled as

$$y = Xb + e, \quad e \sim \mathcal{N}(0, \sigma^2 I_n)$$

where y is the $n \times 1$ response vector, X is the $n \times p$ design matrix of predictors, b is the $p \times 1$ vector of regression coefficients, and σ^2 is the residual variance.

This defines the **likelihood**:

$$y \mid b, \sigma^2 \sim \mathcal{N}(Xb, \sigma^2 I_n)$$

The goal is to estimate b and determine which predictors truly contribute to explaining variation in y.

The Spike-and-Slab Mixture Prior

Each regression coefficient b_i is drawn from a **two-component mixture prior**:

$$p(b_i \mid \sigma_b^2, \pi) = \pi \, \mathcal{N}(0, \sigma_b^2) + (1 - \pi) \, \delta_0$$

where:

- π is the **prior probability** that b_i is nonzero (active predictor)
- δ_0 is a **point mass at zero** (excluded predictor)

Thus, with probability π a predictor belongs to the **slab** (included), and with probability $1 - \pi$ it belongs to the **spike** (excluded).

This prior induces **sparsity**, allowing the model to automatically select relevant predictors while shrinking others exactly to zero.

Hierarchical (Indicator) Representation

The spike-and-slab prior can be expressed hierarchically by introducing a **binary inclusion indicator** δ_i:

$$b_i = \alpha_i \, \delta_i$$

where

$$\alpha_i \mid \sigma_b^2 \sim \mathcal{N}(0, \sigma_b^2), \quad \delta_i \mid \pi \sim \text{Bernoulli}(\pi)$$

- α_i: effect size when the predictor is **included**
- δ_i: binary variable indicating **inclusion (1)** or **exclusion (0)**

This representation separates **effect size** (α_i) from **inclusion** (δ_i), making inference straightforward via Gibbs sampling. Marginalizing over δ_i recovers the **spike-and-slab mixture prior** defined earlier.

Prior for the Inclusion Probability π

The overall **sparsity level** of the model is controlled by π, which represents the prior probability that a predictor is included.
We assign π a **Beta prior**:

$$\pi \sim \text{Beta}(\alpha, \beta)$$

- Small α and large $\beta \rightarrow$ favor **sparser models**
- $\alpha = \beta = 1 \rightarrow$ **uniform prior** (no preference)
- Larger $\alpha \rightarrow$ favor **denser models**

This prior allows the **data to inform the degree of sparsity** through posterior updating of π.

Priors for Variance Components

Variance parameters are typically assigned **scaled inverse-chi-squared** priors:

$$\sigma_b^2 \sim S_b \chi^{-2}(v_b), \quad \sigma^2 \sim S\chi^{-2}(v)$$

These priors are **conjugate**, yielding **closed-form conditional updates** for both variance components.
The hyperparameters (S_b, v_b) and (S, v) encode prior beliefs about the variability of **effect sizes** and **residual noise**, respectively.

In the spike-and-slab model, the **sum of squares** for updating σ_b^2 is computed only over the **included effects**, i.e.,

$$\sum_{i:\delta_i=1} \alpha_i^2,$$

ensuring that the variance of inactive predictors (where $\delta_i = 0$) does not influence the estimate of σ_b^2.

Joint Posterior Structure

As in Bayesian linear regression with normal priors, we combine the **likelihood** and **priors** to obtain the **joint posterior** over all model parameters:

$$p(\mu, \alpha, \delta, \pi, \sigma_b^2, \sigma^2 \mid y) \propto p(y \mid \mu, \alpha, \delta, \sigma^2)\, p(\alpha \mid \sigma_b^2)\, p(\delta \mid \pi)\, p(\pi)\, p(\sigma_b^2)\, p$$

This captures our **updated beliefs** about effect sizes, inclusion indicators, and variance components after observing the data.

The inference procedure follows the same principle as for standard BLR — we use **Gibbs sampling** to draw from each parameter's full conditional distribution.
Next, we derive these **full conditional distributions** from the joint posterior.

Gibbs Sampling: The Algorithm (Spike-and-Slab BLR)

We iteratively draw from the following **full conditionals**:

1. Sample $\alpha \mid \delta, \sigma_b^2, \sigma^2, y$
2. Sample $\delta \mid \alpha, \pi, y$
3. Sample $\pi \mid \delta$
4. Sample $\sigma_b^2 \mid \alpha, \delta$
5. Sample $\sigma^2 \mid \alpha, \delta, y$

Each step updates one parameter block given the others, and iterating the sequence yields samples from the joint posterior. Since all conditionals have **standard forms** (Normal, Bernoulli, Beta, scaled inverse-χ^2), Gibbs sampling is **straightforward and efficient**.

Posterior Inclusion Probabilities

The **posterior inclusion probability (PIP)** quantifies how likely each predictor is to be **included in the model** (i.e., truly associated with y):

$$\widehat{\Pr}(\delta_i = 1 \mid y) = \frac{1}{T} \sum_{t=1}^{T} \delta_i^{(t)}$$

- **High PIP** $\rightarrow$ predictor is likely important
- **Low PIP** $\rightarrow$ predictor is likely irrelevant

PIPs provide a direct measure of **variable relevance** and form the basis for **Bayesian feature selection**.

Advantages of Spike-and-Slab Priors

This hierarchical mixture prior offers several key benefits:

- **Sparsity** — allows exact zeros for irrelevant predictors
- **Interpretability** — binary indicators yield posterior inclusion probabilities (PIPs)
- **Adaptivity** — the inclusion probability π is inferred from the data
- **Balance** — captures both strong signals (for detection) and small effects (for prediction)

Thus, spike-and-slab models naturally combine **variable selection** with **Bayesian uncertainty quantification**.

Summary of Bayesian Linear Regression

Bayesian Linear Regression combines the **likelihood** and **prior** to form the **posterior**, enabling principled modeling, regularization, and uncertainty quantification.

- Inference is performed via **MCMC**, typically **Gibbs sampling**, producing posterior draws for **means**, **credible intervals**, and **predictions**.
- **Spike-and-slab priors** introduce **sparsity** and support **variable selection**, assigning **exact zeros** to irrelevant predictors and identifying key variables through **posterior inclusion probabilities (PIPs)**.
- **Conjugate** and **mixture** priors allow for **efficient** and **robust inference**, even when $p > n$.
- With appropriate **convergence diagnostics**, Bayesian models yield **stable and reliable inference** across diverse data settings.

Applications in Genomics

We have now covered the basic framework of **Bayesian Linear Regression (BLR)** and will illustrate how it provides a **unified approach** for analyzing genetic and genomic data.

- **Genome-Wide Association Studies (GWAS)** and **fine-mapping** of causal variants
- **Genetic prediction** and **heritability estimation**
- **Pathway** and **gene-set enrichment** analyses

These applications demonstrate how BLR connects **statistical modeling** with **biological interpretation** in quantitative genetics.