UNIVERSIDADE DE SÃO PAULO
ESCOLA SUPERIOR DE AGRICULTURA
"LUIZ DE QUEIROZ"
DEPARTAMENTO DE GENÉTICA
LGN5825 Genética e Melhoramento de Espécies Alógamas

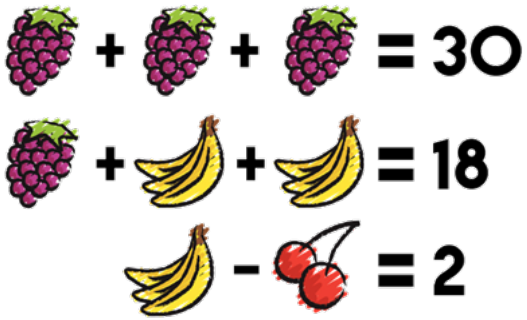
Bayesian methods of GS

Prof. Roberto Fritsche-Neto

roberto.neto@usp.br

Piracicaba, November 13th, 2019

Estimation



$$3A + 0B + 0C = 30$$
$$1A + 2B + 0C = 18$$
$$0A + 1B - 1C = 2$$

$$X*\beta = Y$$

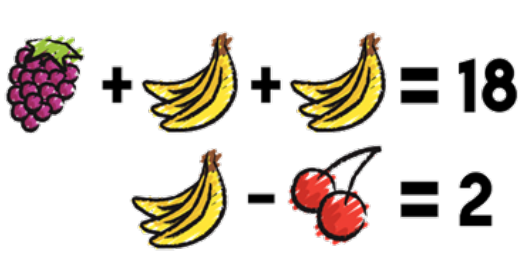
$$\begin{pmatrix} 3 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} A \\ B \\ C \end{pmatrix} = \begin{pmatrix} 30 \\ 18 \\ 2 \end{pmatrix}$$

$$\begin{pmatrix} A \\ B \\ C \end{pmatrix} = \begin{pmatrix} 3 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 1 & -1 \end{pmatrix}^{-1} \begin{pmatrix} 30 \\ 18 \\ 2 \end{pmatrix}$$

$$\beta = X^{-1}*Y$$

$$\begin{pmatrix} A \\ B \\ C \end{pmatrix} = \begin{pmatrix} 1/3 & 0 & 0 \\ -1/6 & 1/2 & 0 \\ -1/6 & 1/2 & -1 \end{pmatrix} \begin{pmatrix} 30 \\ 18 \\ 2 \end{pmatrix} = \begin{pmatrix} 10 \\ 4 \\ 2 \end{pmatrix}$$

Estimation and restrictions



Infinite soluctions

Markers > Observations

How do we solve it?

Restrictions

- Parametric
  - Penalized
    - Ridge Regression (shrinkage)
    - Elastic Net
    - LASSO (Shrinkage and selection)
  - Bayesian
    - Bayesian Ridge Regression
    - Scaled Mixtures
      - Bayesian LASSO
      - BayesA
    - Discrete Mixtures (e.g. BayesB, BayesCπ)
- Non-parametric
  - Neural Networks
  - Kernel methods (RKHS)

Bayes theory

$$P(X|Y) = \frac{P(Y|X)P(X)}{P(Y)} \begin{cases} P(X|Y): pdf conditioned\ to\ X\ given\ Y \\ P(Y|X): pdf conditioned\ to\ Y\ given\ X \\ P(X): pdf of\ X\ and P(Y): pdf of\ Y \end{cases}$$

$$\begin{cases} y: vector\ of\ observations \\ \theta: vector\ of\ parameters \end{cases}$$

$$\boldsymbol{P(\theta|y) \propto P(y|\theta)P(\theta)}$$

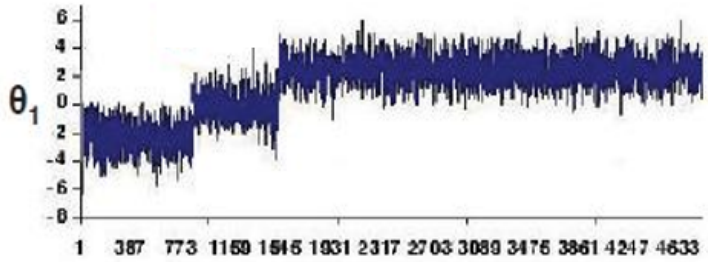
Posteriori distribution of θ

A priori distribution of θ

Genetic architecture (t, normal, exp, …

Jointed distribution of data given the parameters

Likelihood function (gaussian or ordinal)



Main difference among Bayesian methods

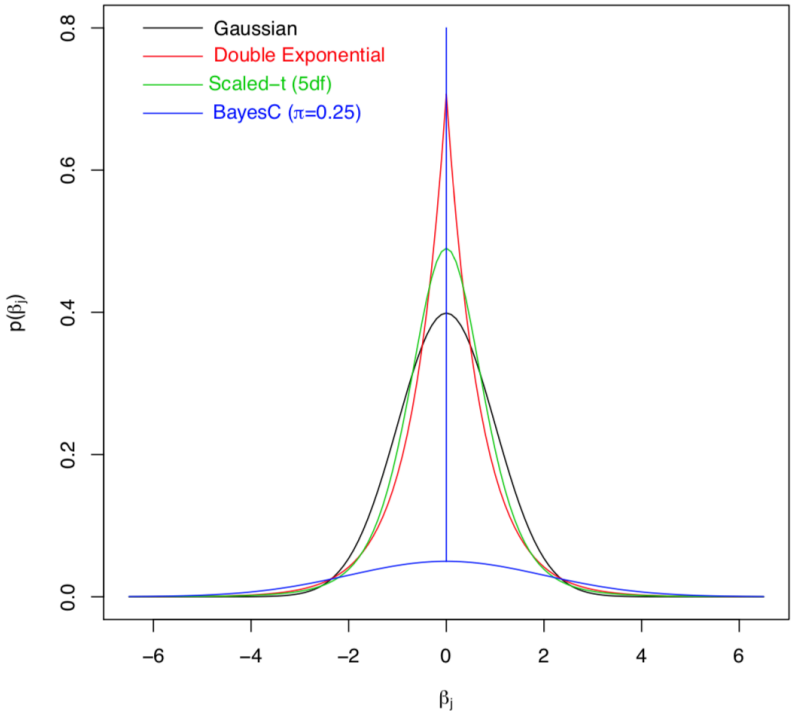
SNPs effect (a_i)

$$y = 1u + \sum_{i=1}^{m} M_i a_i + e$$

i = 1, 2,…m (SNPs)

Bayes alphabet

Model	Genetic variance (Vu)	Genetic control
Ridge Regression	Just one for all SNPs	Polygenic, small effects and explain equal amounts of variance
Bayes A	Specific for each SNP	Polygenic, small effects and explain different amounts of variance
Bayes B	Specific for each SNP	Polygenic, but some SNPs does not affect the trait ($\boldsymbol{\pi}$), where $\boldsymbol{\pi}$ is fixed
Bayes C$\boldsymbol{\pi}$	Just one for all SNPs	Polygenic, but some SNPs does not affect the trait ($\boldsymbol{\pi}$), where $\boldsymbol{\pi}$ is random
Bayesian LASSO	Specific for each SNP	Oligogenic. Many SNPs does not affect the trait. Not necessary to define $\boldsymbol{\pi}$



Bayes alphabet

Model	Conditioned fdp	Prior fdp	Vu
RR	$a_i \mid \sigma^2 \sim N(0, \sigma^2)$	$\sigma^2 \sim \chi^{-2}(S, v)$	$\sigma_u^2 = 2\sigma^2 \sum_{i=1}^{m} p_i(1-pi)$
Bayes A	$\alpha_i \mid \sigma_i^2 \sim N(0, \sigma_i^2)$	$\sigma_i^2 \sim \chi^{-2}(S, v)$	$\sigma_u^2 = 2\sum_{i=1}^{m} p_i(1-pi)\sigma_i^2$
Bayes B	$a_i \mid \boldsymbol{\pi}, \sigma_{i1}^2 \sim (1-\gamma_i)N(0, \sigma_{i0}^2 = 0) + \gamma_i N(0, \sigma_{i1}^2)$ $P(y_i=0) = \boldsymbol{\pi}$; $P(y_i=1) = 1-\boldsymbol{\pi}$	$\sigma_{i1}^2 \sim \chi^{-2}(S, v)$	$\sigma_u^2 = 2\sum_{i=1}^{m} p_i(1-pi)\sigma_i^2$
Bayes C$\boldsymbol{\pi}$	$a_i \mid \boldsymbol{\pi}, \sigma^2 \sim \boldsymbol{\pi} N(0, \sigma^2 = 0) + (1-\boldsymbol{\pi})N(0, \sigma^2)$ $\boldsymbol{\pi} \sim U(0,1)$	$\sigma^2 \sim \chi^{-2}(S^*, v)$ $S^* = \tilde{\sigma}^2 (v-2)/v$ expected priori for SNP variance (fixed)	$\sigma_u^2 = 2(1-\pi)\sigma^2 \sum_{i=1}^{m} p_i(1-pi)$
Lasso	$a_i \mid \tau_i, \sigma_e^2 \sim N(0, \tau_i^2 \sigma_e^2)$ $a_i \mid \tau_i, \sigma_e^2 \sim N(0, \sigma_i^2)$ $a \mid \lambda, \sigma_e^2 \sim \prod_i \frac{\lambda}{2\sigma_e} exp\left(\frac{-\lambda \mid a_i}{\sigma_e}\right)$	$\tau_i^2 \mid \lambda^2 \sim Exp(\lambda^2)$ $\lambda^2 \sim Gamma(\varphi_1, \varphi_2)$ $a_i \mid \lambda, \sigma_e^2 \sim Double - Exp(0, \lambda)$	$\sigma_u^2 = \sum_{i=1}^{m} \underbrace{\tau_i^2 \sigma_e^2}_{\sigma_i^2} 2p_i(1-pi)$