



MODELOS MISTOS APLICADOS AO MELHORAMENTO DE PLANTAS

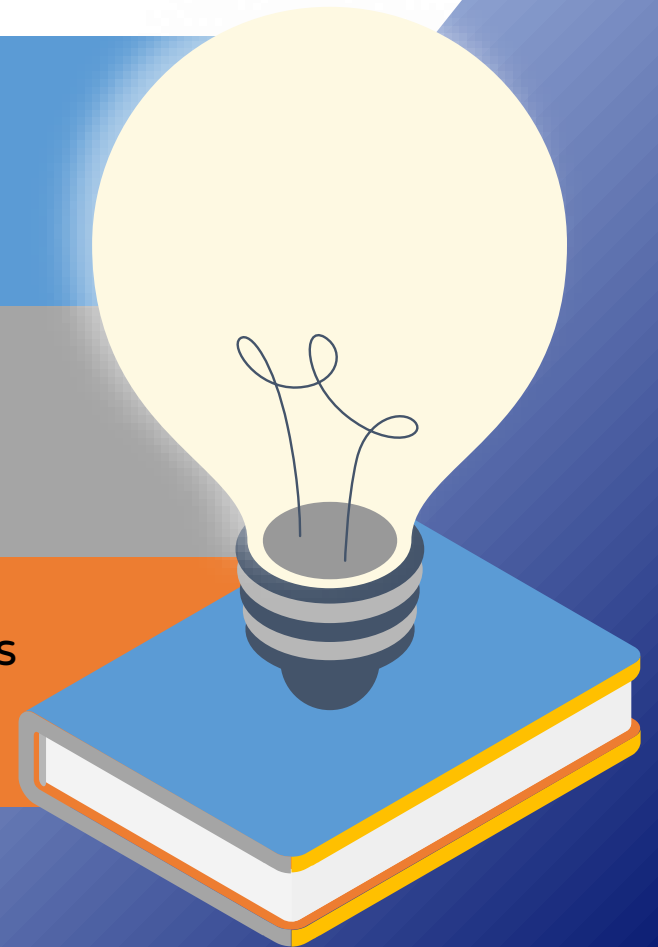
M.Sc. Renan Garcia Malikouski

M.Sc. Filipe Manoel Ferreira



OBJETIVO - NO FINAL DA AULA DE HOJE, VOCÊ SERÁ CAPAZ DE:

- 01 Entender conceitos básicos sobre modelos mistos
- 02 Entender como modelos mistos são aplicados nos melhoramentos de plantas
- 03 Aprender a analisar e interpretar algumas análises básicas utilizando o software SELEGEN – REML/BLUP



ROTEIRO: 1º DIA



01

Apresentação
dos
participantes



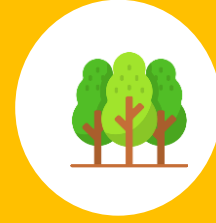
02

Introdução a
teoria de
modelos mistos



03

Apresentação
do software
Selegen
REML/BLUP



04

Modelo 105 –
Est. Geral;

ROTEIRO: 2º DIA



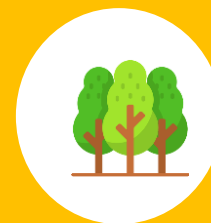
Modelo 96 –
Efeitos genéticos
e ambientais
(LRT)



Modelo 54-
Interação GxA



Modelo 55 -
Repetibilidade



Modelo 54 –
Índice de
Seleção



Modelos 103
e 104-
Diversidade
genética e
agrupamento

ROTEIRO: 3º DIA



Modelo 76 -
Delineamento
em Blocos
Aumentados



Modelos 165-
Matriz de
parentesco
completa



Modelo 1 –
PMI e PIC



Modelo 88 –
Dialelo e fatorial
interpopulacional



Modelos 106
– Ne e
otimização de
seleção

OS FACILITADORES



Renan Garcia Malikouski · 1º

Doutorando em Genética e Melhoramento na Universidade Federal de Viçosa

Formação:

- Eng. Agrônomo (IFES)
- M. Sc. Agricultura tropical (UFES)
- Doutorando Genética e Melhoramento (UFV) – Laboratório de Biometria UFV

Interesse:

- Estatística Experimental;
- Biometria no Melhoramento de plantas;
 - Diversidade Genética;
 - Interação GxA;
 - Análises de medidas repetidas;
- Seleção Genômica.

OS FACILITADORES



Filipe Manoel Ferreira 🗣️ (He/Him)

Tree Genetics and Breeding | Quantitative Genetics | Prediction-based Modeling

Formação:

- Eng. Florestal (UFV)
- M. Sc. Genética e Melhoramento (UFV)
- Doutorando Genética e Melhoramento (UFV) – Laboratório de Biometria UFV

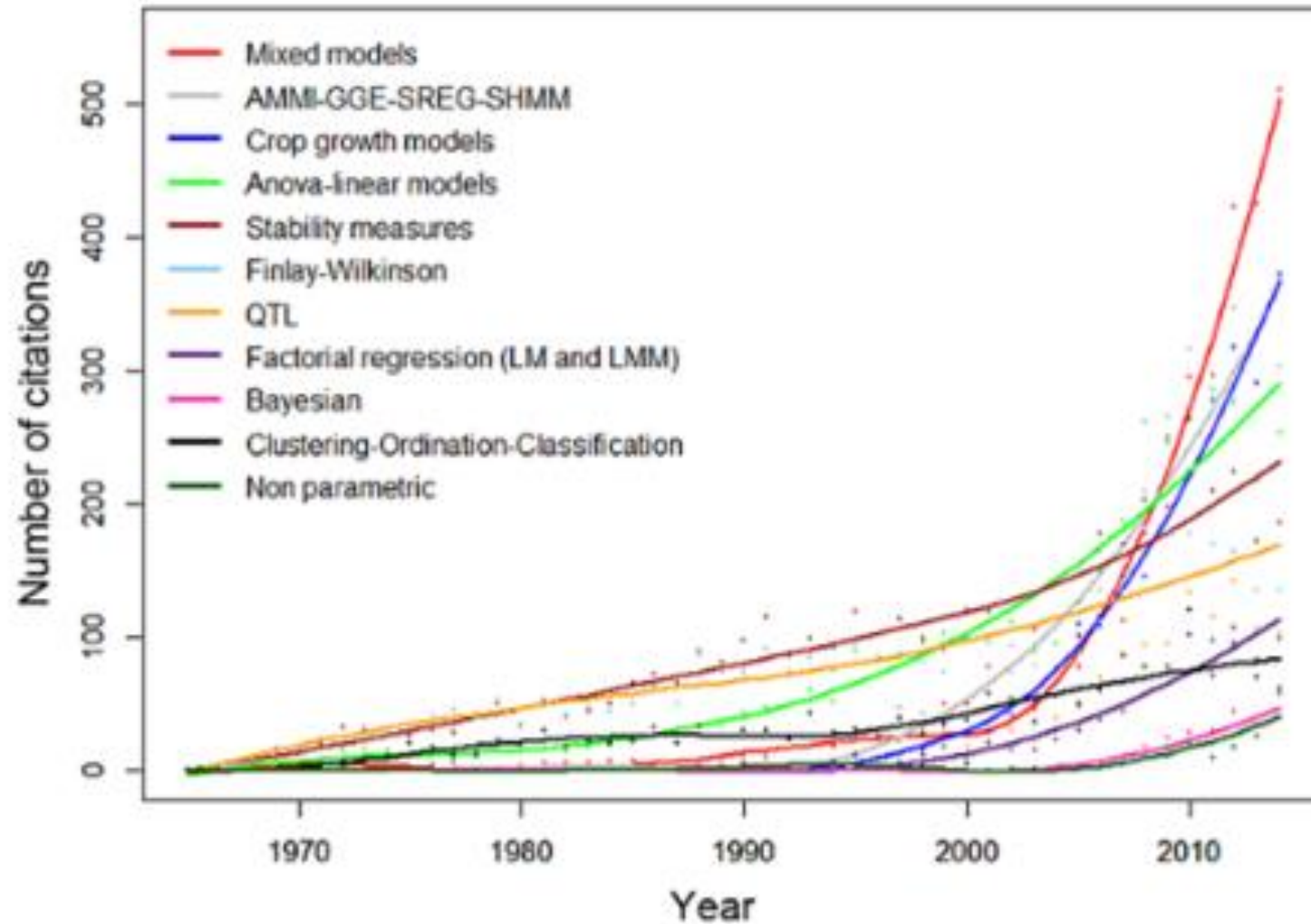
Interesses:

- Seleção genômica;
- Genética quantitativa;
- Melhoramento Florestal;
- Programação.

INTRODUÇÃO À TEORIA DE MODELOS MISTOS

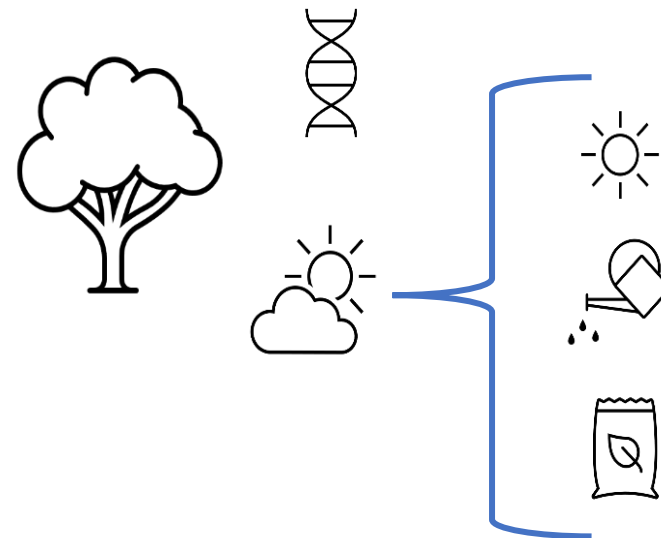
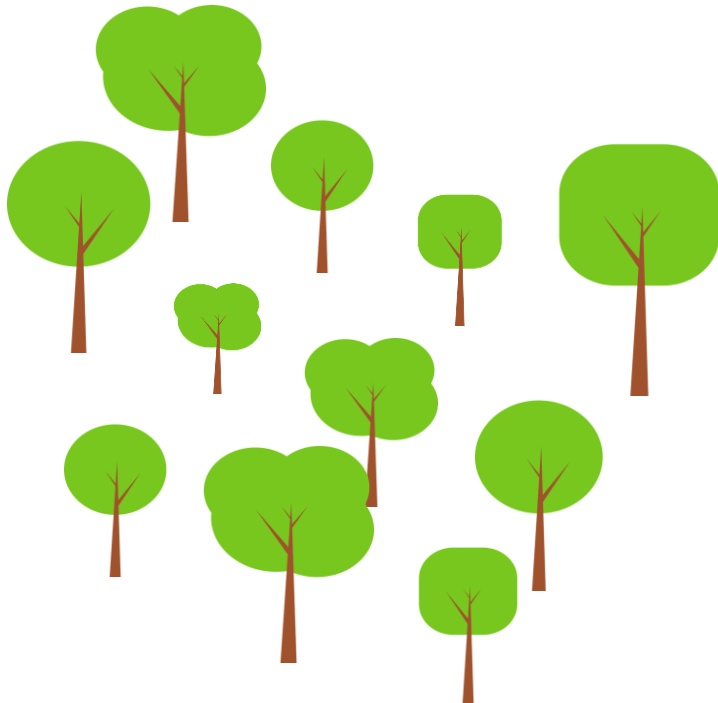


MOTIVAÇÃO

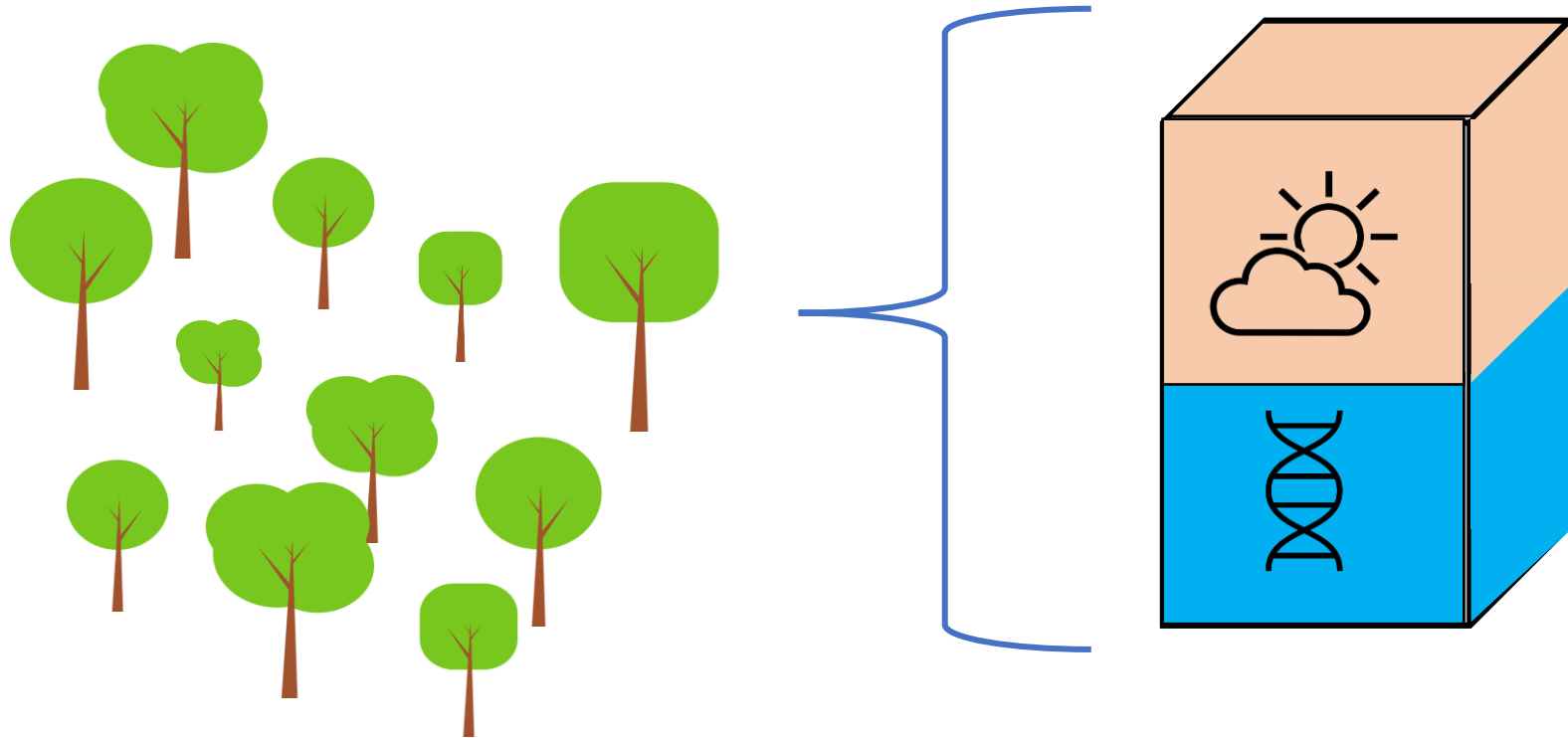


MELHORAMENTO GENÉTICO

- Deslocar a média de uma dada característica na direção desejada;
- Necessitam de conceitos estatísticos para decompor a variação observada em causas genéticas e ambientais.



DECOMPONDO A VARIÂNCIA OBSERVADA

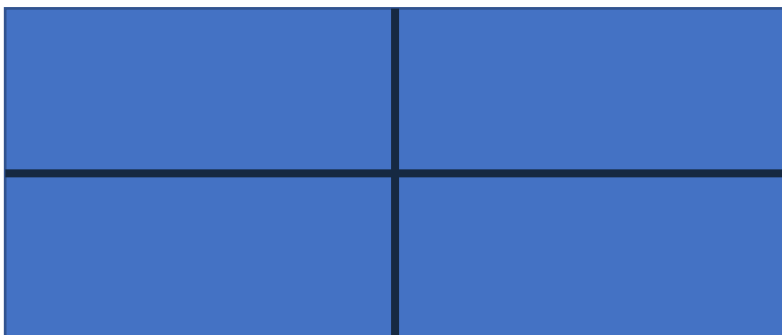


Após **estimar** os componentes de variância, consigo **prever** os valores genéticos para cada indivíduo

POR QUÊ DEVO DECOMPOR A VARIÂNCIA OBSERVADA?

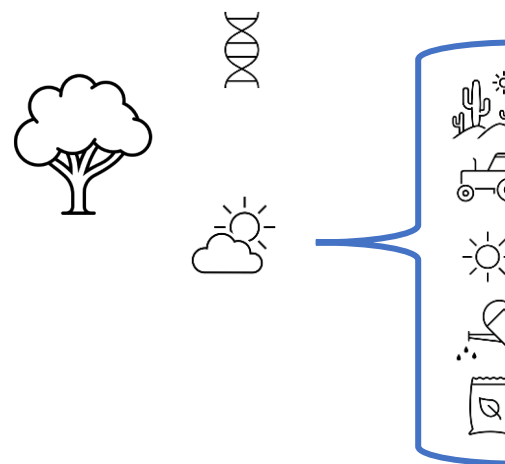
FENOTÍPICA

Seleção massal (controle ambiental é mínimo);



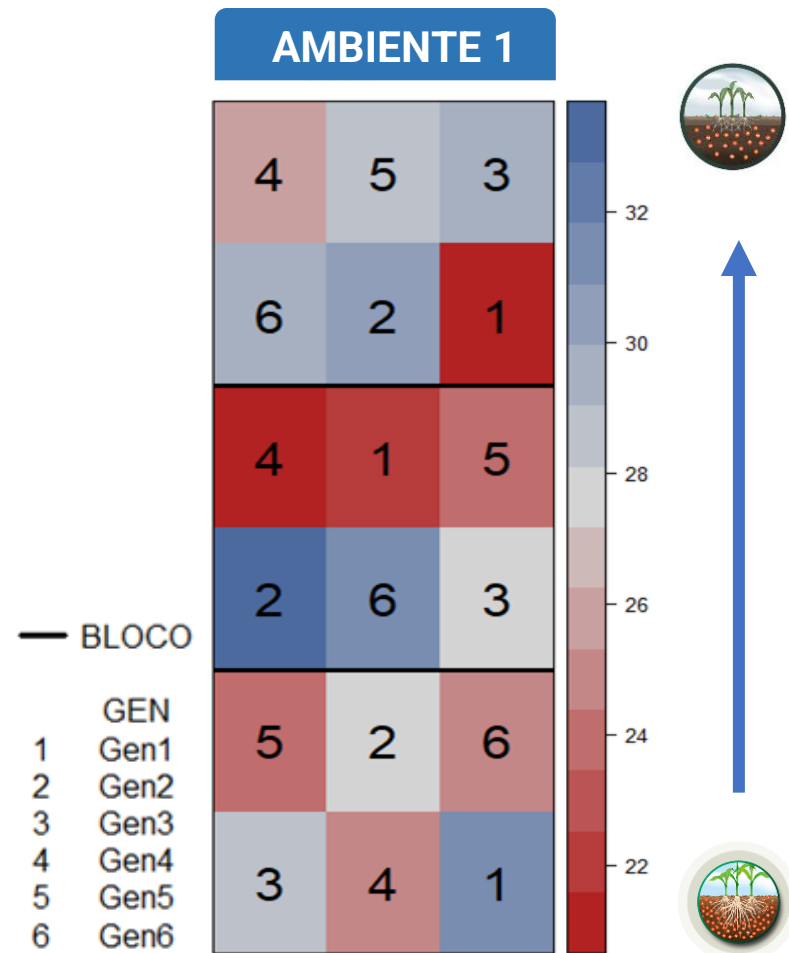
GENOTÍPICA

Utilizar **desenhos experimentais** e conhecimentos de **estatística** e **genética quantitativa** para separar o efeito genotípico dos demais efeitos (ambiental);



MODELOS ESTATÍSTICOS

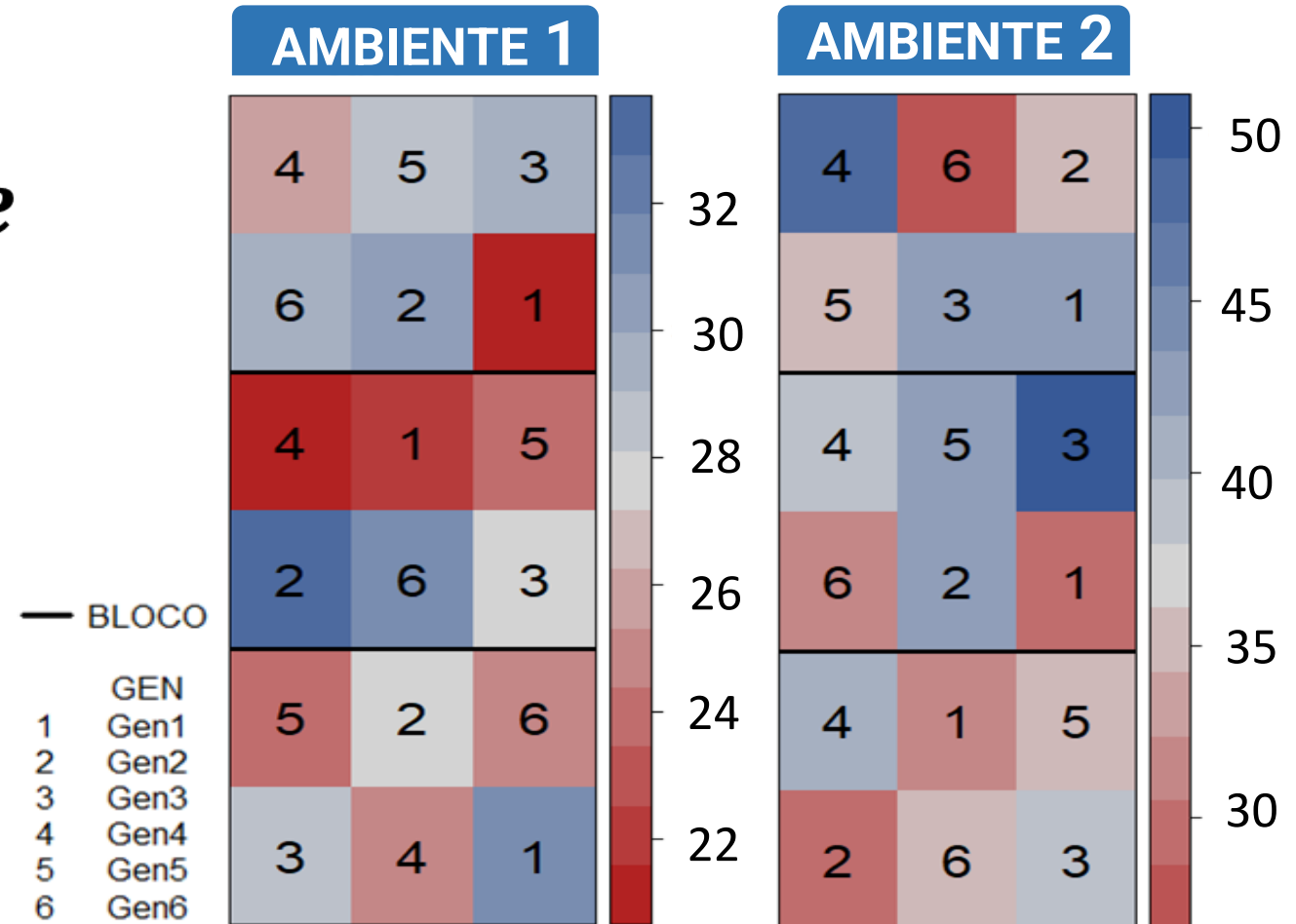
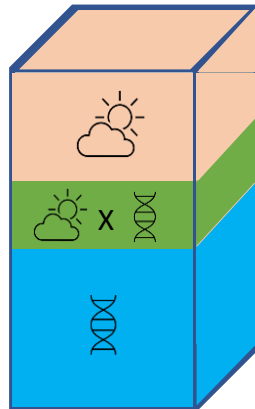
- São modelos estocásticos que são representações mais simplificadas da realidade através de equações matemáticas



$$y = X\beta + Zu + e$$

MODELOS ESTADÍSTICOS

$$y = X\beta + Zu + Wi + e$$

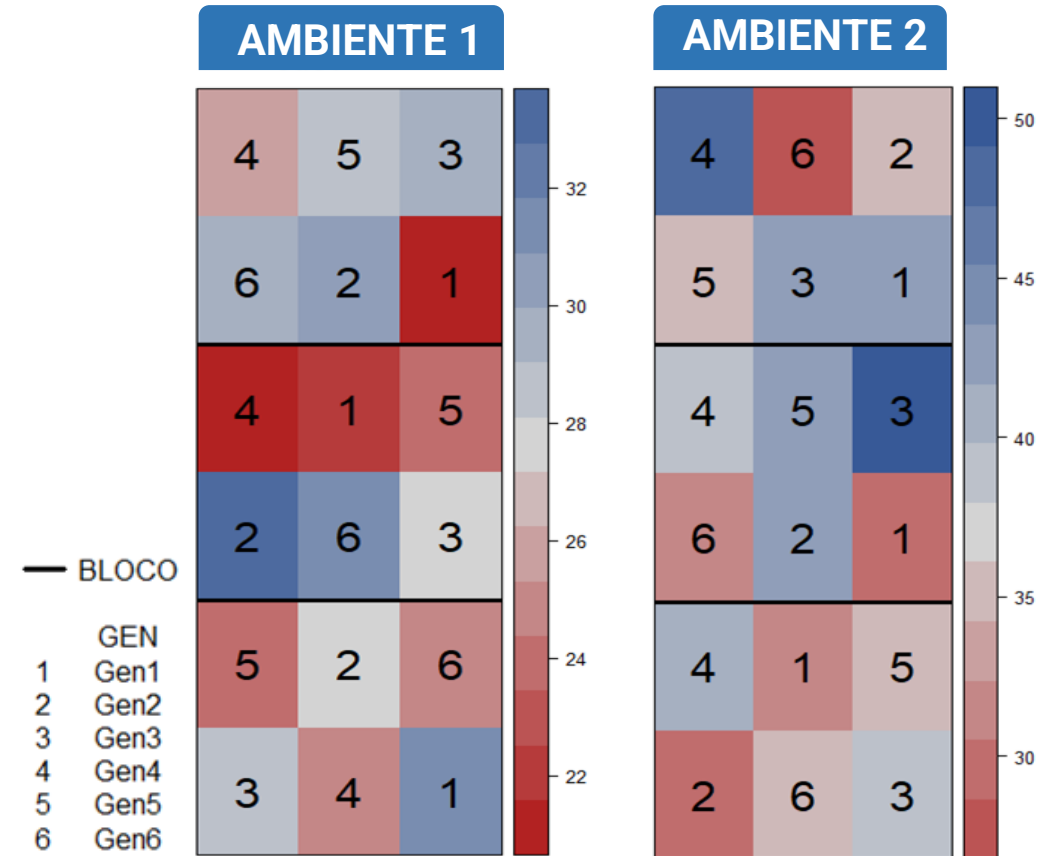


MEDIDAS REPETIDAS

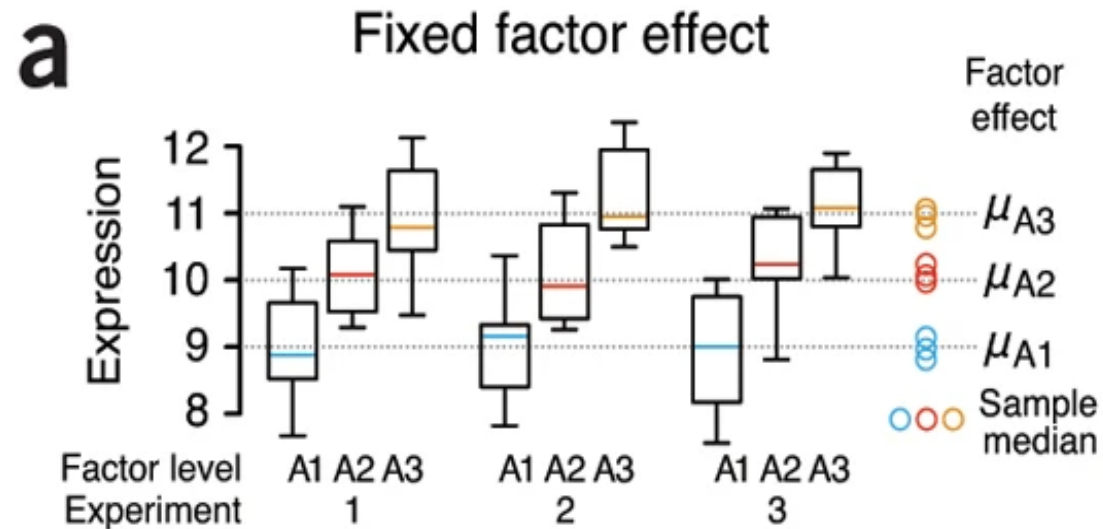
$$y = X\beta + Zu + Wi + e$$



$$y = X\beta + Zu + Wi + Tp + e$$

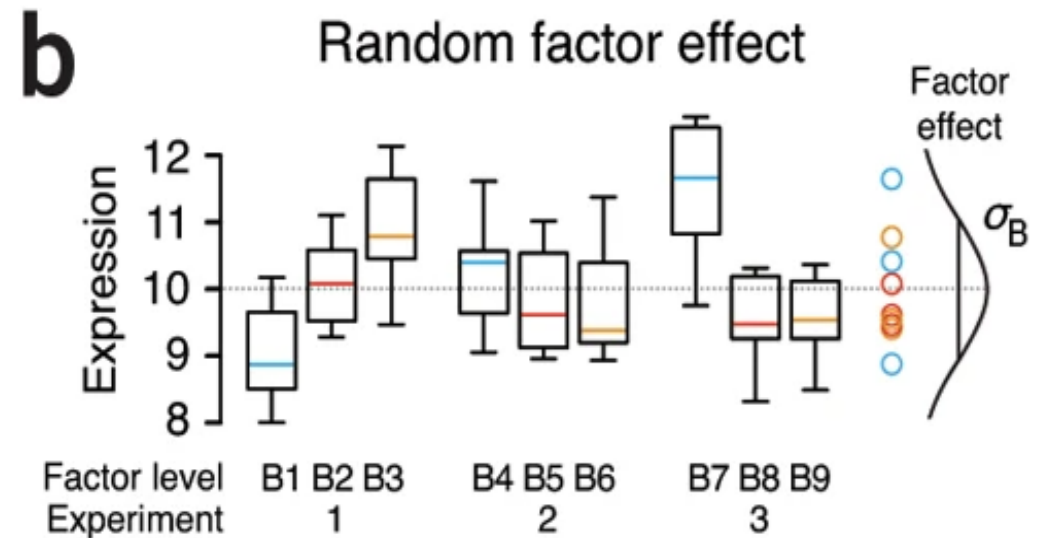


NATUREZA DO EFEITO



Níveis dos fatores não alteram (Princípios ativos de defensivos);

$$H_0 \rightarrow m1 = m2 = m3 = \dots = mt$$

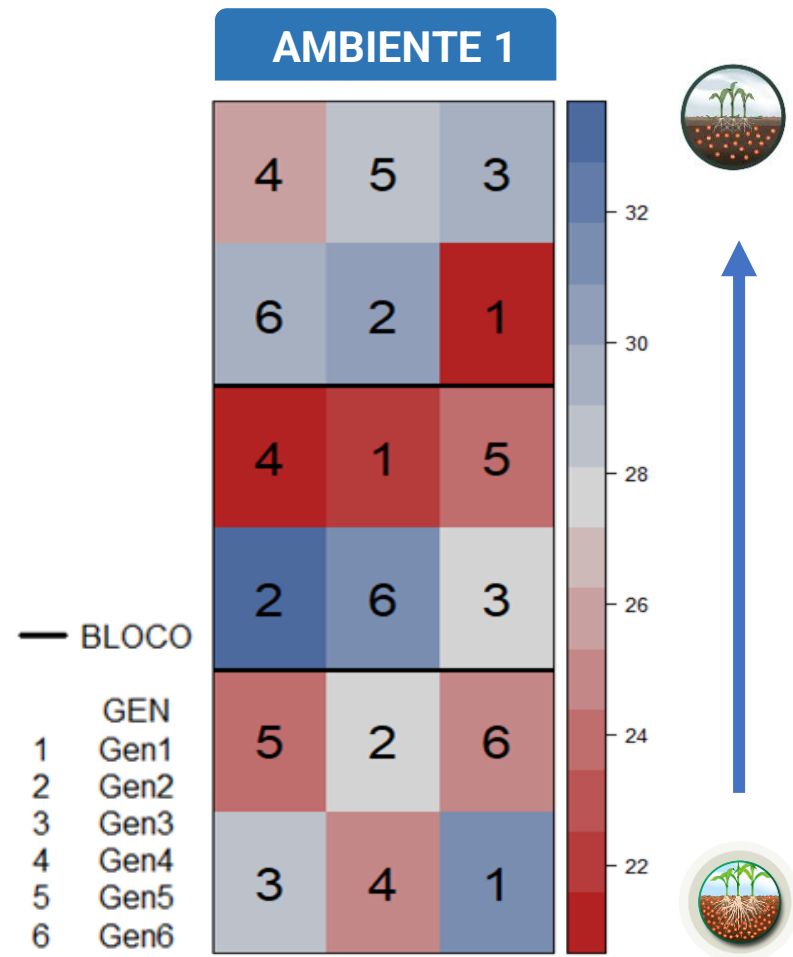


Não repete os mesmos níveis (amostragem aleatória de uma distribuição de densidade).

Posso expandir as inferência para outros níveis

$$H_0 \rightarrow \sigma_x^2 = 0$$

MODELOS ESTATÍSTICOS



$$y = X\beta + Zu + e$$

β = bloco (fixo)

u = genético (aleatório)

e = resíduo (aleatório)

$$\begin{bmatrix} u \\ e \end{bmatrix} \sim N \left[\mathbf{0}, \begin{pmatrix} \sigma_u^2 \mathbf{K} & 0 \\ 0 & \sigma_e^2 \mathbf{R} \end{pmatrix} \right]$$

EQUAÇÕES DE MODELOS MISTOS (HENDERSON)

$$y = X\beta + Zu + e$$

BLUEs

$$\begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z} \\ \mathbf{Z}'\mathbf{X} & \mathbf{Z}'\mathbf{Z} + \lambda\mathbf{K}^{-1} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{b}} \\ \hat{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{Z}'\mathbf{y} \end{bmatrix}$$



BLUPs

Shrinkage coefficient

SHRINKAGE COEFFICIENT

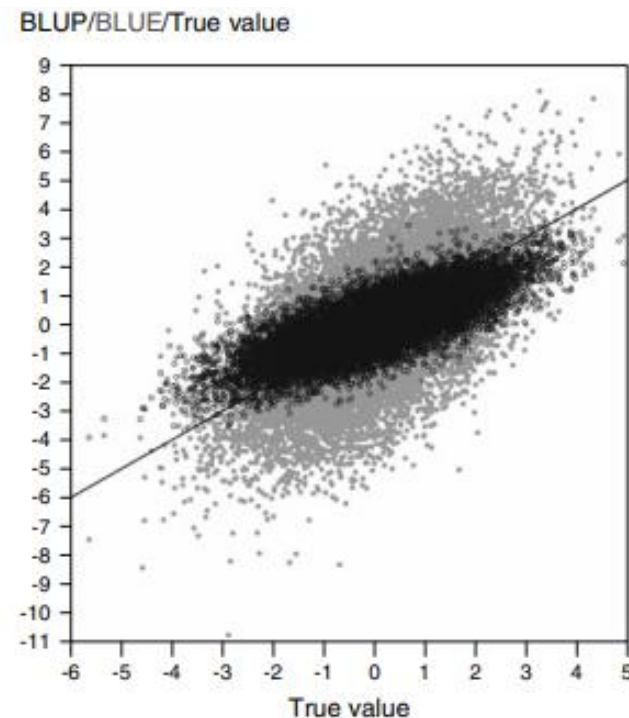
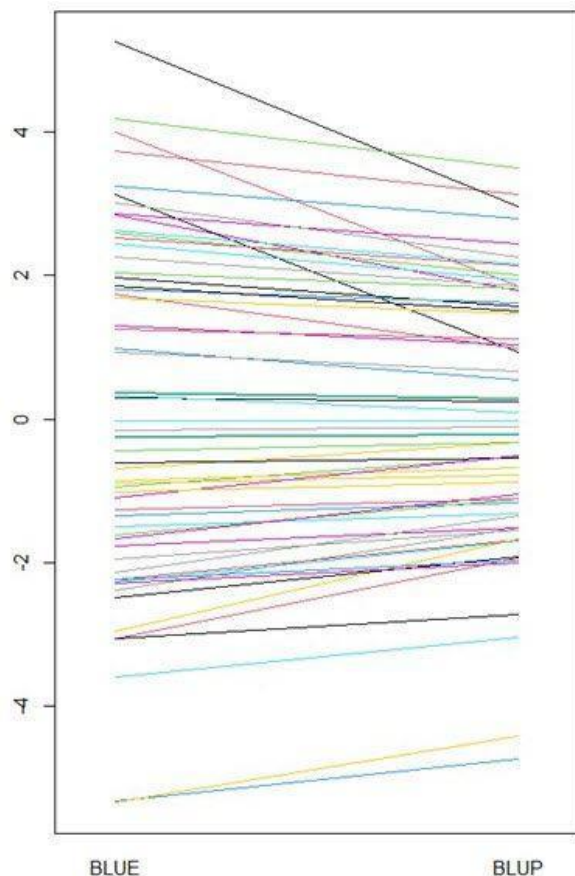


Fig. 4 Plot of simulated BLUP (black circles), BLUE (grey dots), and true genetic values (solid line) versus true values

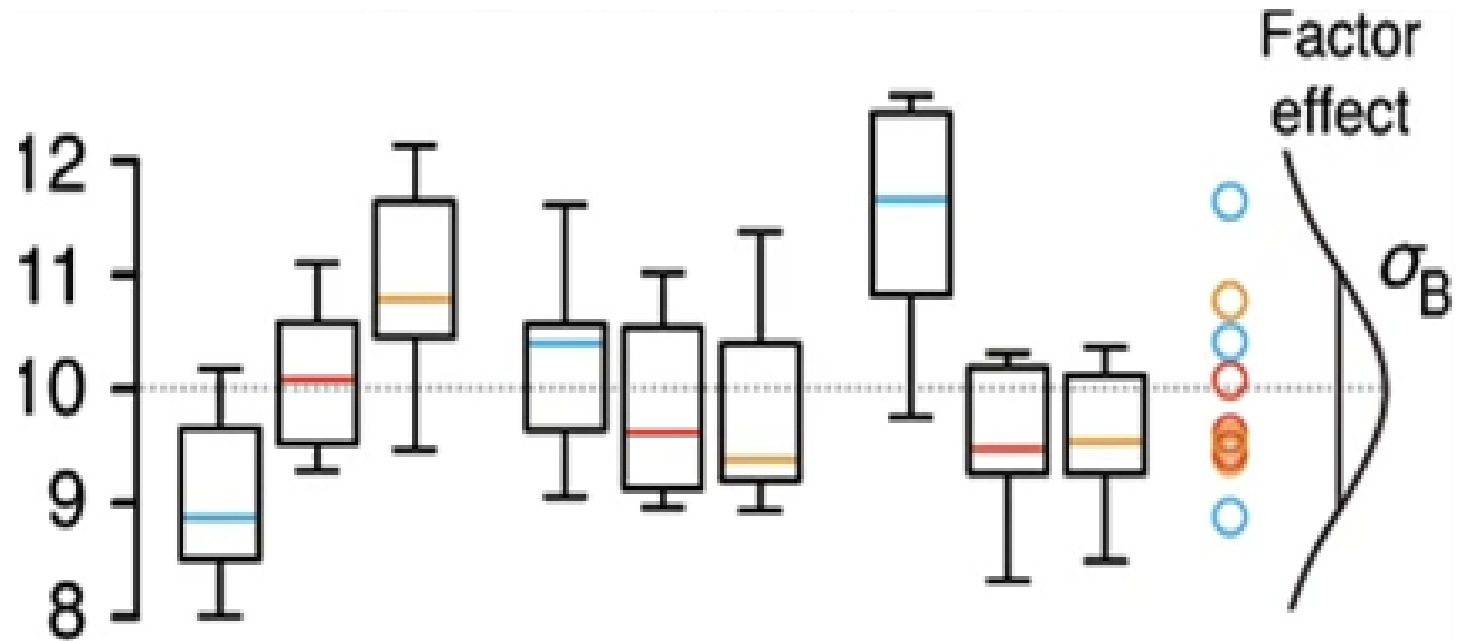
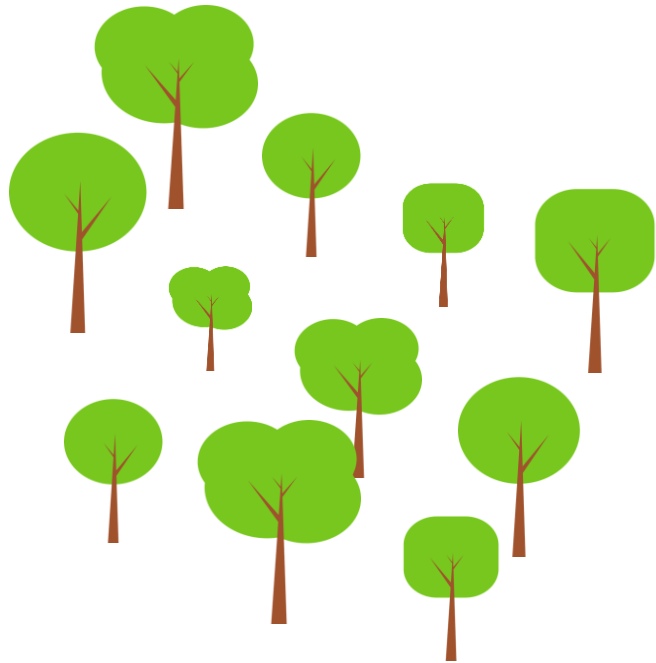
$$BLUE_i = \hat{a}_i = (\bar{y}_i - \bar{y})$$

$$BLUP_i = \hat{a}_i = \left[\frac{\sigma_a^2}{\sigma_a^2 + \sigma_e^2/n} \right] (\bar{y}_i - \bar{y})$$

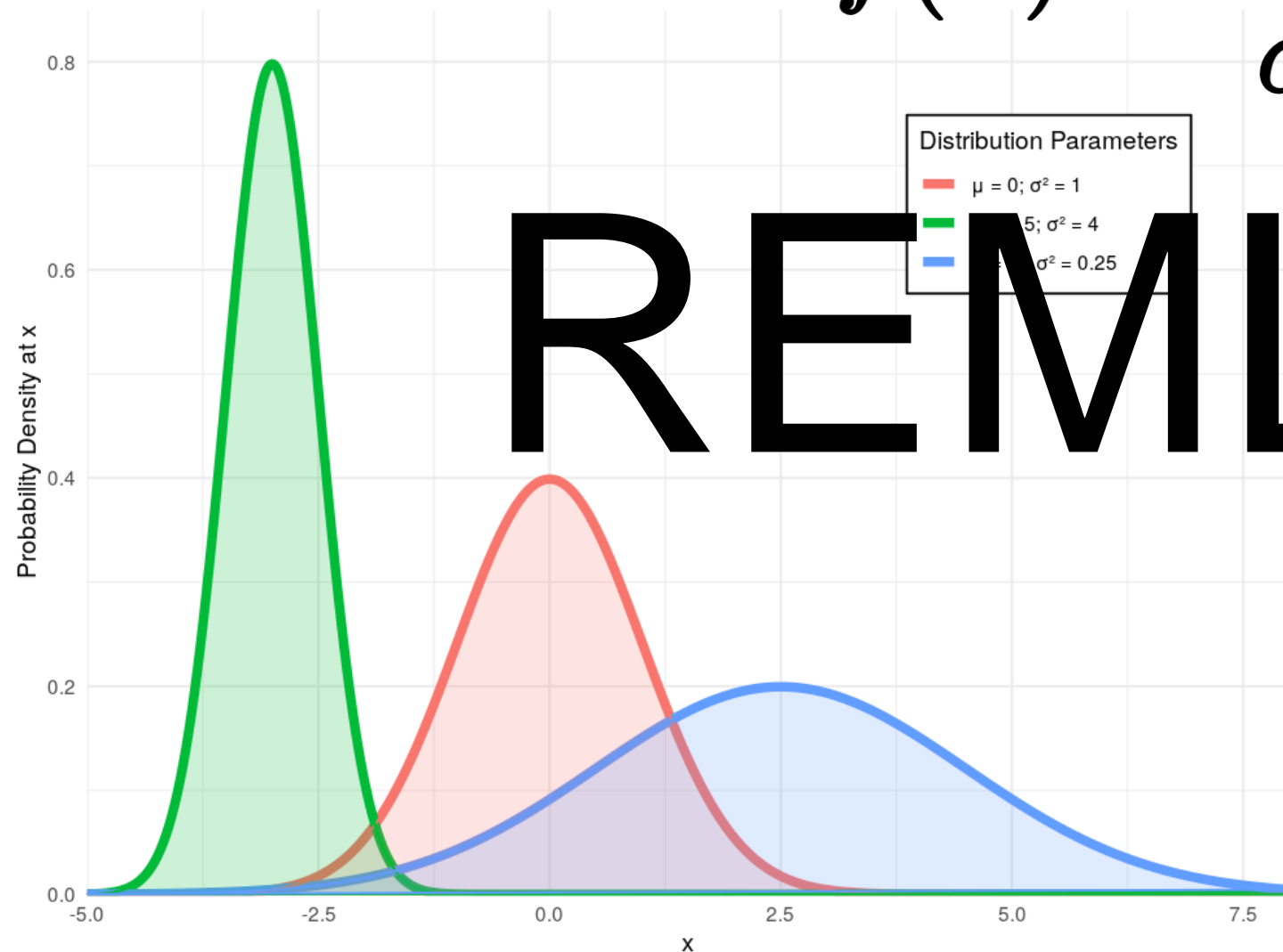
Predizer os valores genéticos

Estimar os componentes de
variância

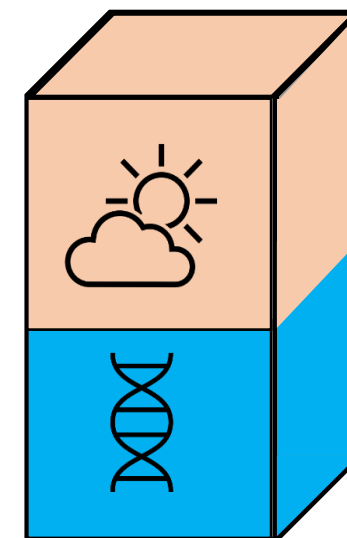
REML/BLUP



$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$



REML

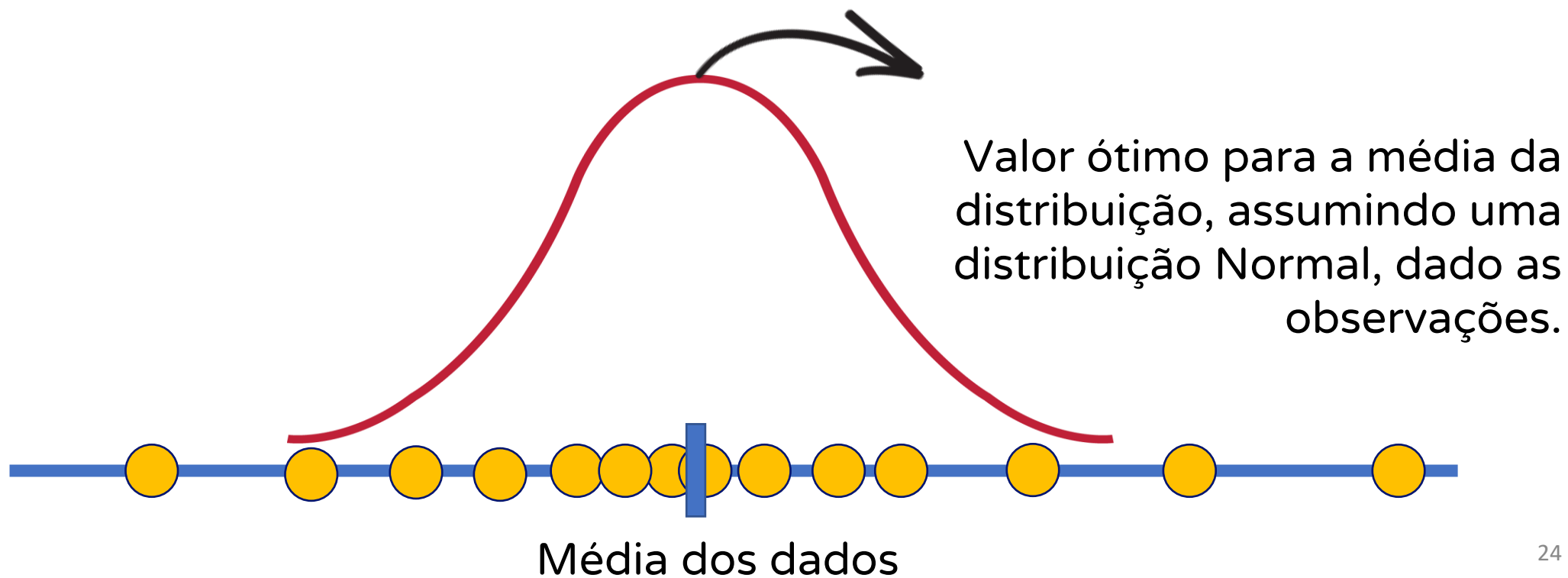


REML

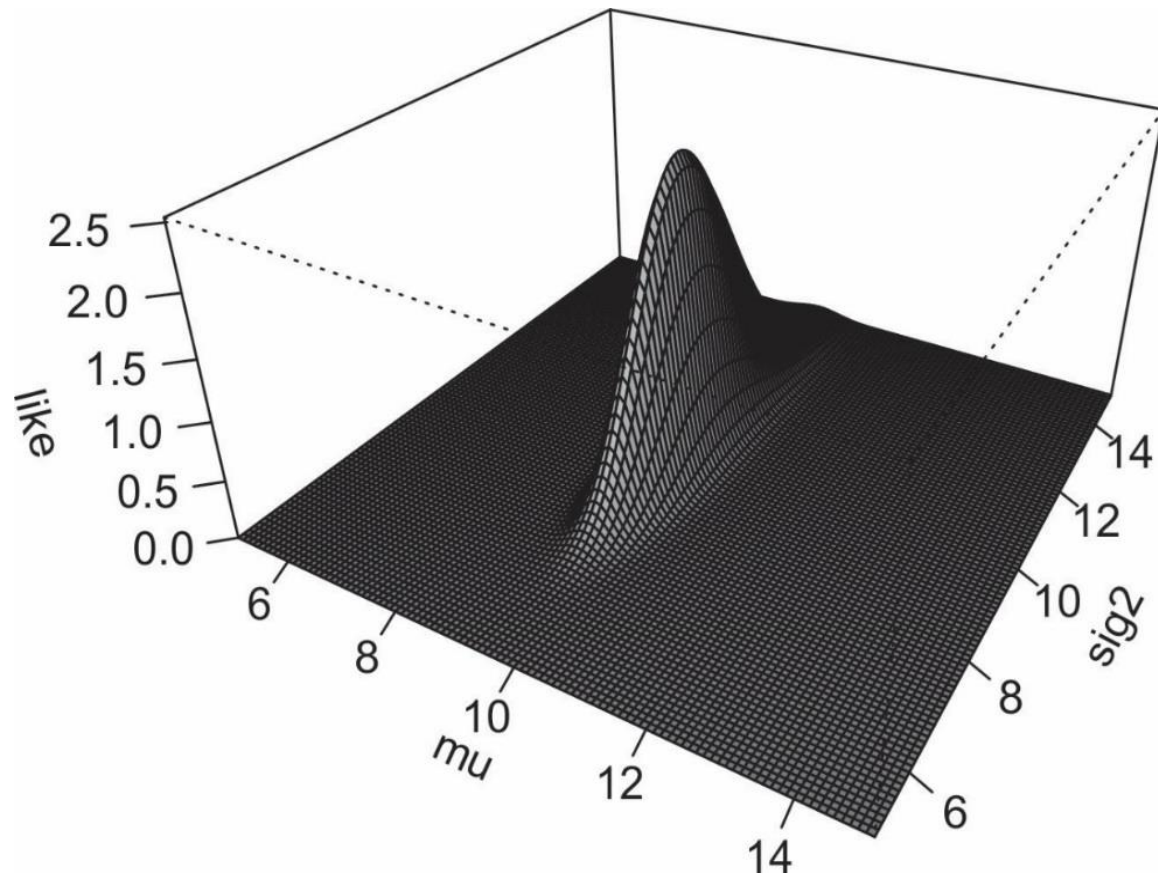
Restricted/residual Maximum Likelihood

“Encontrar uma boa maneira de ajustar uma distribuição aos dados”

Verossimilhança
de observar os
dados



$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$



BLUP

Best linear unbiased prediction

“Calcular valores genéticos errando pouco”

EQUAÇÕES DE MODELOS MISTOS

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e}$$

$$\begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z} \\ \mathbf{Z}'\mathbf{X} & \mathbf{Z}'\mathbf{Z} + \mathbf{I}^{-1}\lambda \end{bmatrix} \begin{bmatrix} \hat{\mathbf{b}} \\ \tilde{\mathbf{g}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{Z}'\mathbf{y} \end{bmatrix}$$

$$\lambda = (1 - h^2)/h^2$$

$$h^2 = \frac{\sigma_g^2}{\sigma_f^2}$$

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{W}\mathbf{p} + \mathbf{e}$$

$$\lambda = (1 - h^2)/h^2$$

$$h^2 = \frac{\sigma_g^2}{\sigma_f^2}$$

$$\begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z} & \mathbf{X}'\mathbf{W} \\ \mathbf{Z}'\mathbf{X} & \mathbf{Z}'\mathbf{Z} + \mathbf{I}^{-1}\lambda & \mathbf{Z}'\mathbf{W} \\ \mathbf{W}'\mathbf{X} & \mathbf{W}'\mathbf{Z} & \mathbf{W}'\mathbf{W} + \mathbf{I}^{-1}\lambda_2 \end{bmatrix} \begin{bmatrix} \hat{\mathbf{b}} \\ \tilde{\mathbf{g}} \\ \tilde{\mathbf{p}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{Z}'\mathbf{y} \\ \mathbf{W}'\mathbf{y} \end{bmatrix}$$

$$\lambda_2 = (1 - h^2 - C^2)/C^2$$

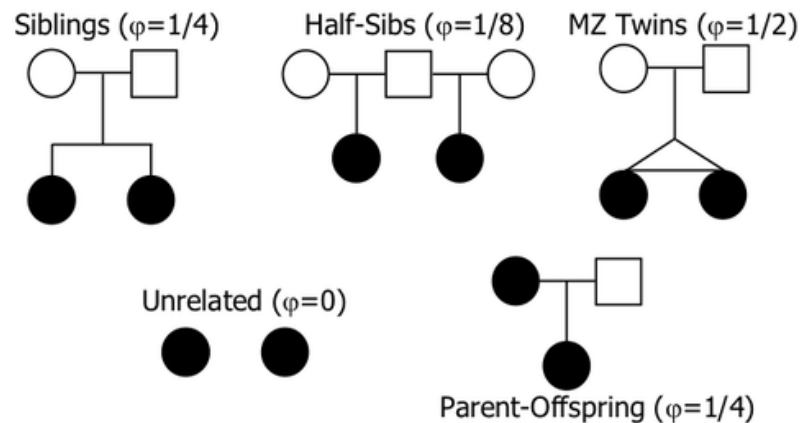
$$C^2 = \frac{\sigma_p^2}{\sigma_f^2}$$

EQUAÇÕES DE MODELOS MISTOS

$$\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + I^{-1}\lambda \end{bmatrix} \begin{bmatrix} \hat{b} \\ \tilde{g} \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \end{bmatrix}$$

Não considerando o parentesco

$$I = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$



Considerando o parentesco
Ex. Meio irmãos

$$I \rightarrow A = \begin{bmatrix} 1 & 1/4 & 1/4 \\ 1/4 & 1 & 1/4 \\ 1/4 & 1/4 & 1 \end{bmatrix}$$

EMM – MATRIZES DE INCIDÊNCIA

$$y = X\beta + Zu + e$$

Ind	Gen	Bloc	y
1	1	1	18,36
2	1	2	21,54
3	2	1	8,23
4	2	2	7,25
5	3	1	16
6	3	2	10
7	4	1	18,25
8	4	2	20
9	5	1	9,95
10	5	2	10,01

$$\begin{array}{c}
 y \\
 \left[\begin{array}{c} 18,36 \\ 21,54 \\ 8,23 \\ 7,25 \\ 16,00 \\ 10,00 \\ 18,25 \\ 20,00 \\ 9,95 \\ 10,01 \end{array} \right] \\
 (n \times 1)
 \end{array}
 =
 \begin{array}{c}
 \overbrace{X}^{\substack{\beta_1 \quad \beta_2}} \\
 \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{array} \right] \\
 (n \times b)
 \end{array}
 \begin{array}{c}
 \left[\begin{array}{c} b_1 \\ b_2 \end{array} \right] \\
 (b \times 1)
 \end{array}
 +
 \begin{array}{c}
 \overbrace{Z}^{\substack{u_1 \quad u_2 \quad u_3 \quad u_4 \quad u_5}} \\
 \left[\begin{array}{ccccc} 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{array} \right] \\
 (n \times g)
 \end{array}
 \begin{array}{c}
 \left[\begin{array}{c} g_1 \\ g_2 \\ g_3 \\ g_4 \\ g_5 \end{array} \right] \\
 (n \times 1)
 \end{array}
 +
 \begin{array}{c}
 e \\
 \left[\begin{array}{c} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \\ e_7 \\ e_8 \\ e_9 \\ e_{10} \end{array} \right] \\
 (n \times 1)
 \end{array}$$

EMM – RESOLUÇÃO

$$\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + I^{-1}\lambda \end{bmatrix} \begin{bmatrix} \hat{b} \\ \tilde{g} \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \end{bmatrix}$$

$$\begin{bmatrix} \hat{b} \\ \tilde{g} \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \end{bmatrix} \begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + I^{-1}\lambda \end{bmatrix}^{-1}$$

REML / BLUP

Componente / parâmetro	Estimativa
	26,41
	6,16
	32,58
	0,81
	0,19
	13,95

$$\begin{bmatrix} b1 \\ b2 \\ g1 \\ g2 \\ g3 \\ g4 \\ g5 \end{bmatrix} = \begin{bmatrix} 14,1580 \\ 13,7600 \\ 5,3646 \\ 4,6258 \\ -0,8587 \\ -3,5630 \\ -5,5687 \end{bmatrix}$$

Best linear unbiased estimation - BLUEs

Best linear unbiased prediction - BLUPs

EQUAÇÕES DE MODELOS MISTOS

$$\begin{bmatrix} b1 \\ b2 \\ g1 \\ g2 \\ g3 \\ g4 \\ g5 \end{bmatrix} = \begin{bmatrix} 14,1580 \\ 13,7600 \\ 5,3646 \\ 4,6258 \\ -0,8587 \\ -3,5630 \\ -5,5687 \end{bmatrix}$$

Valores genéticos preditos ($v_{\hat{g}}$)

x

Valores genéticos reais (v_g)

Medida de precisão:

Acurácia seletiva ($r_{\hat{g}g}$; 0 – 1)

$$r_{\hat{g}g} = \sqrt{1 - \frac{PEV}{\sigma_g^2}} \quad \text{ou} \quad r_{\hat{g}g} = \sqrt{h_2}$$

ACURÁCIA SELETIVA

Fatores que influenciam a acurácia seletiva:

- i. Número de repetições;
- ii. Variância residual;
- iii. Proporção varE/varG .

Classificação:

- i. Baixa (0 – 0.49);
- ii. Moderada (0.50 – 0.69);
- iii. Alta (0.70 - 0.89) - Recombinação ;
- iv. Multo alta (0.90 – 1) - Recomendação.

HERDABILIDADE

$$h_g^2 = \frac{\sigma_g^2}{\sigma_f^2}$$

- Representa uma proporção da variabilidade existente em uma população segregante que é de natureza genética;
- Estabelecer estratégias de seleção;
- Informa sobre o controle genético de cada carácter a ser melhorado;
- Quanto maior a herdabilidade maior o controle genético;
- Indica a facilidade de se praticar o melhoramento daquele carácter.

HERDABILIDADE

- Herdabilidade no sentido amplo

$$h_g^2 = \frac{\sigma_g^2}{\sigma_y^2} \quad \text{ou} \quad h_g^2 = \frac{\sigma_g^2}{\sigma_g^2 + \sigma_e^2}$$

- Herdabilidade no sentido restrito

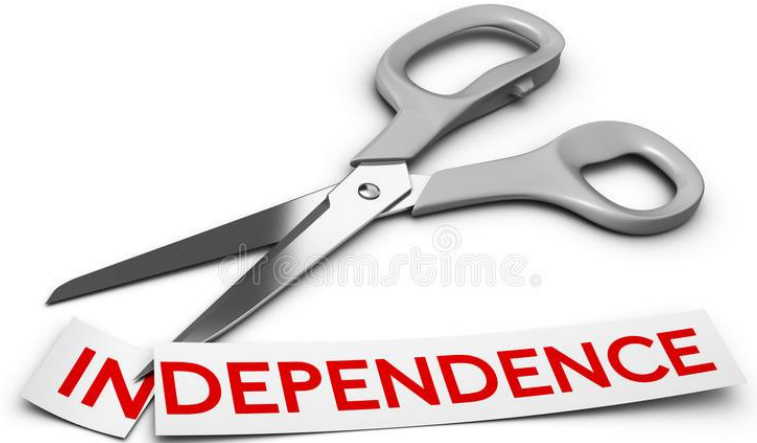
$$h_a^2 = \frac{\sigma_a^2}{\sigma_y^2} \quad \text{ou} \quad h_a^2 = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_d^2 + \sigma_e^2}$$

- Classificação das herdabilidade:
- Baixa (0 – 0.15)
- Moderada (0.15 – 0.50)
- Alta (0.50 – 0.80)
- Muito alta (0.80 – 1)

VANTAGENS DOS MM

Modelos Lineares Fixos:

- Assume independência entre níveis;
- Relação linear: $y = \alpha + \beta x$



Relaxa a pressuposição da independência entre observações

- Considera a correlação entre níveis

ANOVA X METODOLOGIA DE MODELOS MISTOS

ANOVA

Componentes de variância:

-Métodos dos momentos;

Teste de significância:

-Teste F;

Candidatos a seleção:

-Médias fenotípicas:

Met. de Modelos lineares mistos

Componentes de variância:

Restricted maximum likelihood - REML;

Teste de significância:

-Teste de LRT via Qui-quadrado

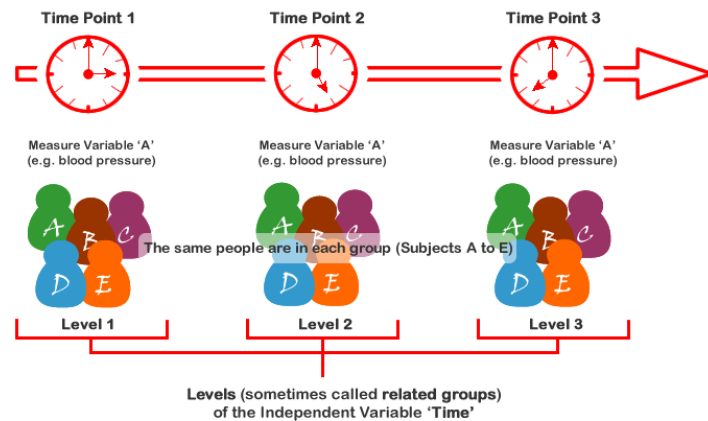
- Efeito fixo: teste de Wald

Candidatos a seleção:

-Média genética ou genotípica via Best linear unbiased prediction (BLUP);

VANTAGENS DOS MM

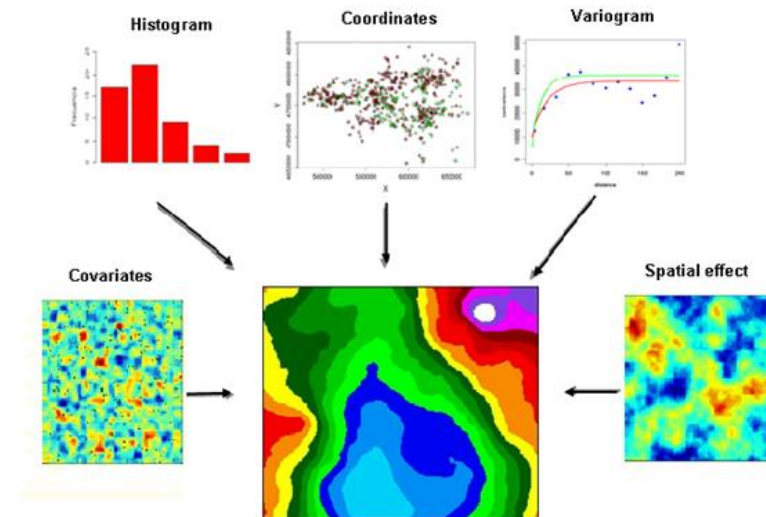
Time series (Repeated Measures)



© Lund Research Ltd 2011

<http://statistics.laerd.com>

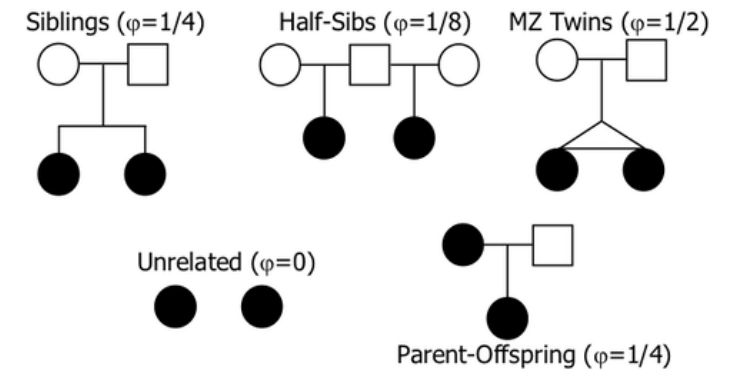
Space (Spatial analysis)



FreshBiostats

Created by Microsoft Paint and Microsoft Office Picture Manager

Family structure



VANTAGENS DOS MM

- Maximiza o ganho genético e a eficiência dos programas de melhoramento
- Não exige balanceamento dos dados
- Permite utilizar simultaneamente um grande número de informações, gerando estimativas precisas

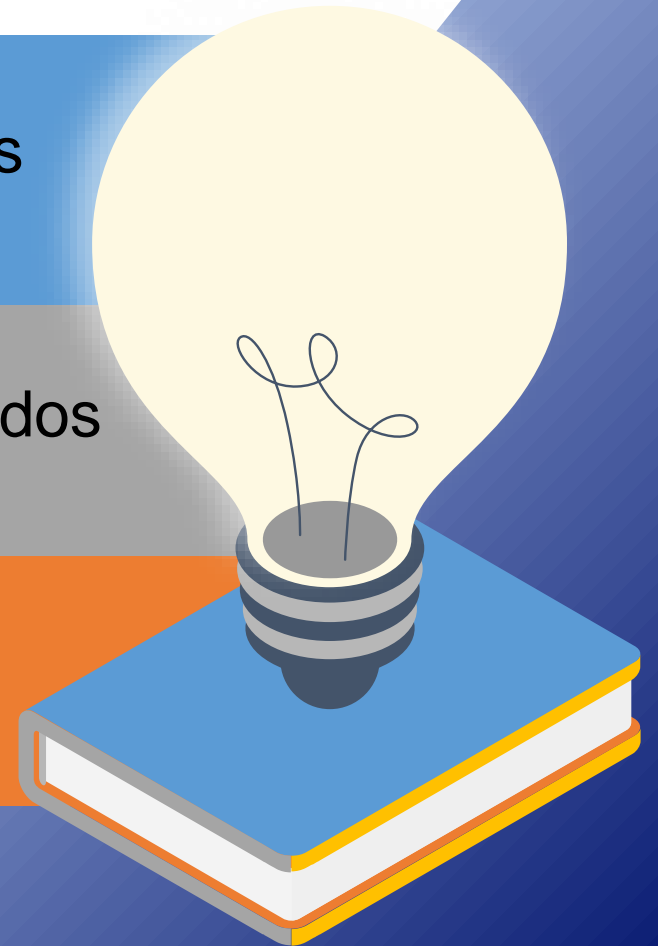
OBJETIVO - NO FINAL DA AULA DE HOJE, VOCÊ SERÁ CAPAZ DE:



Entendemos conceitos básicos sobre modelos mistos



Entendemos como modelos mistos são aplicados no melhoramentos de plantas



TO BE CONTINUED....

Aprender a analisar e interpretar análises básicas utilizando o software
SELEGEN – REML/BLUP



🎵 BREAK TIME



filipe.manoel@ufv.com



renan.malikouski@ufv.br