

GENÉTICA
E
MELHORAMENTO DE PLANTAS
2023/2024

Soluções/resoluções
dos exercícios
de Genética Quantitativa

Elsa Gonçalves, novembro 2023

PARTE I

GRUPO I

Cálculo de parâmetros genéticos em populações de clones de espécies de propagação vegetativa

1. a) Figura B.

b) Valor fenotípico médio de cada genótipo corresponde à coluna da média; estimativa da variância ambiental= 0,701; estimativa da variância fenotípica ao nível da média= 0,610.

c) Valor fenotípico médio de cada genótipo corresponde à coluna da média; estimativa da variância ambiental= 1,419; estimativa da variância fenotípica ao nível da média= 0,893.

d) Quadro 1A – investigador Y; Quadro 1B – investigador X; a atitude mais correta foi tomada pelo investigador Y. Usando a média do rendimento das plantas da unidade experimental reduzem-se os desvios ambientais (o desvio ambiental associado a uma observação individual é sempre muito maior do que o associado à média dos indivíduos que compõem a unidade experimental).

e) A alternativa do investigador B (o que propõe 5 repetições). O aumento do número de repetições aumenta a precisão de um ensaio, conduzindo, em geral, a um aumento da heritabilidade.

f) No enunciado diz-se que o ensaio foi instalado segundo um delineamento experimental em blocos casualizados completos, o que significa que o modelo de análise deverá conter os efeitos associados ao factor bloco, isto é, genericamente:

$$Y_{ij} = \mu_1 + \text{bloco}_i + \text{clone}_j + e_{ij}, \text{ com } i = 1, \dots, 4, \quad j = 1, \dots, 100.$$

Considere-se o bloco como um factor de efeitos fixos e o clone como um factor de efeitos aleatórios.

Tomando em consideração este factor, a estimativa da variância dos erros aleatórios foi de 0,678, um valor inferior ao obtido com uma análise em que não foi incluído o factor bloco (análise do quadro 1A). Esta diminuição deve-se ao facto de se conseguir um maior controlo da variação espacial no ensaio através da inclusão do factor bloco. Este modelo de análise terá como consequência o aumento da heritabilidade e, conseqüentemente, uma maior precisão na predição dos efeitos genotípicos. Este resultado revela a importância do delineamento experimental para o aumento da eficiência da seleção, assim como alerta para as consequências de uma análise não incluir todos os factores que contribuem para o valor obtido da variável resposta.

g) Sabemos que: $\hat{\sigma}_{P(\text{média})}^2 = 0,610$, $\hat{\sigma}_e^2 = QMRE = 0,677697$.

Considerando o bloco como um factor de efeitos fixos e o clone com um factor de efeitos aleatórios,

$$\hat{\sigma}_{P(\text{média})}^2 = \hat{\sigma}_g^2 + \hat{\sigma}_e^2/r, \text{ em que } r = 4 \text{ (temos 4 repetições).}$$

Assim, $\hat{\sigma}_g^2 = 0,44$.

Ou, através do quadro de análise de variância,

$$\hat{\sigma}_g^2 = \frac{QMG - QMRE}{r} = \frac{2,43834 - 0,677694}{4} = 0,44.$$

Coefficiente de variação genotípico do rendimento: $CV_g(\%) = \frac{\sqrt{\hat{\sigma}_g^2}}{MG} \times 100 = \frac{\sqrt{0,44}}{1,779849} \times 100 = 37\%$.
 Por exemplo, esta medida permite a comparação da variabilidade genética entre características.

De acordo com o modelo ajustado, a heritabilidade em sentido lato, ao nível da média dos genótipos, pode ser calculada como: $h_g^2 = \frac{\hat{\sigma}_g^2}{\hat{\sigma}_g^2 + \hat{\sigma}_e^2/r}$. Obtém-se, então $h_g^2 = 0,722$. Isto significa que 72,2% da variabilidade fenotípica ao nível da média dos genótipos é de origem genotípica.

h) Num modelo com 1 factor de efeitos aleatórios, o melhor preditor linear não enviesado empírico (EBLUP) do efeito genotípico para o genótipo i e dado por:

$$EBLUP_{gen}(i) = (\text{média do genótipo}(i) - \text{media geral}) \times \left(\frac{\hat{\sigma}_g^2}{\hat{\sigma}_g^2 + \hat{\sigma}_e^2/r} \right).$$

Sabendo que $Média_{clone0201} = 1,97$ kg/planta, $Média_{geral} = 1,779849$ kg/planta e $h_g^2 = 0,722$, então:

$EBLUP(clone\ 0201) = 0,137$ kg/planta. Devido ao seu potencial genético prevê-se que o clone 0201 produza mais 0,137 kg/planta que a média da população (média da população = 1,779849 kg/planta; *rendimento previsto devido a causas genéticas* = 1,779849 + 0,137 = 1,9168 kg/planta).

i) Trata-se de um delineamento equilibrado e de um modelo clássico (1 factor de efeitos aleatórios). Podemos aplicar a expressão clássica do ganho genético, $R = S \times h^2 = (Msel - MG) \times h^2$:

$$R = (2,912 - 1,780) \times 0,722 = 0,8173$$

$$R(\% \text{ média}) = 45,9\%$$

2. a) A função *lme* do package *nlme* do R ajusta um modelo linear misto. Primeiro é declarada a variável resposta (*rend*) e a seguir os factores de efeitos fixos (neste caso, factor *bloco*), o que corresponde ao código *rend~bloco*. Depois, são declarados os factores de efeitos aleatórios (neste caso, factor *clone*), o que corresponde ao código *random=~1|clone*.

b) As estimativas de máxima verosimilhança restrita (REML) das componentes de variância genotípica e dos erros aleatórios são obtidas com o comando *summary(macieira.lme)*, na linha dos *Random effects*: $\hat{\sigma}_g^2 = (0,2613632)^2 = 0,06831$; $\hat{\sigma}_e^2 = (0,3179316)^2 = 0,1010805$. Alternativamente, estas estimativas podem ser obtidas com o comando *VarCorr(macieira.lme)*.

c) Os melhores preditores lineares não enviesados empíricos (EBLUPs) dos efeitos genotípicos do rendimento são obtidos com o comando *ranef*. Por exemplo, o $EBLUP(clone\ C002) = -0,0047$ kg/planta.

d) De acordo com o modelo ajustado, a heritabilidade em sentido lato, ao nível da média dos genótipos, pode ser calculada como: $h_g^2 = \frac{\hat{\sigma}_g^2}{\hat{\sigma}_g^2 + \hat{\sigma}_e^2/r}$. Assim, $h_g^2 = 0,7716$. O que significa que 77,2% da variabilidade fenotípica (ao nível da média) é de origem genotípica.

e) Pede-se o valor fenotípico médio associado ao genótipo “C003”. No enunciado, temos disponível o valor da média geral ($\bar{Y}.. = 0,5221$ kg/planta, obtido com o comando *summary(macieira)*). Logo,

sabendo que neste modelo $EBLUP(\text{clone } i) = (\bar{Y}_i - \bar{Y}) \frac{\hat{\sigma}_g^2}{\hat{\sigma}_g^2 + \hat{\sigma}_e^2/r}$ e que $EBLUP(C003) = -0.2624317256$, o valor fenotípico médio do clone C003 é $\bar{Y}_{C003} = 0,181986$ kg/planta.

f) Não temos informação sobre a média dos valores fenotípicos médios dos 15 clones superiores, mas podemos calcular o ganho genético como a média dos EBLUPs dos 15 melhores clones. No enunciado os EBLUPs dos efeitos genotípicos estão ordenados, portanto, este ganho será obtido fazendo a média dos 15 melhores valores.

$R = \text{média dos 15 melhores EBLUPs} = 0,4355$ kg/planta. Em % da média, $R(\%) = 83,4\%$.

g) Ajustamento do modelo linear misto usando a função *varComp* do package *varComp* da R. O código descrito é muito semelhante ao da função *lme* do package *nlme*. Primeiro, é declarada a variável resposta (rend) e a seguir os factores de efeitos fixos (neste caso, factor bloco), o que corresponde ao código `rend~bloco`. Depois, são declarados os factores de efeitos aleatórios (neste caso, factor clone), o que corresponde ao código `random=~clone`.

h) As estimativas de máxima verosimilhança restrita (REML) das componentes de variância genotípica e dos erros aleatórios são obtidas com o comando *summary(macieiravarcomp1)*, na linha *Variance component estimates*: $\hat{\sigma}_g^2 = 0,06831$; $\hat{\sigma}_e^2 = 0,1010805$.

i) Teste à variabilidade genética do rendimento. O método de estimação das componentes de variância foi o de máxima verosimilhança restrita (deve-se escolher para estimar o parâmetro a função dos dados que torne “mais provável” a amostra observada) e o teste à componente de variância é um teste de razão de verosimilhanças restritas (ver slide das aulas teóricas).

Hipóteses: $H_0: \sigma_g^2 = 0$ vs $H_1: \sigma_g^2 > 0$; Rejeita-se H_0 , para um dado nível de significância α , se $\Lambda_{calc} > \chi_{\alpha(1)}^2$. Sabemos que $\Lambda_{calc} = 2(l_{R1} - l_{R0})$. l_{R1} é a log-verosimilhança restrita do modelo que está em H_1 (modelo completo com efeitos genotípicos) e no enunciado é dada com o comando *logLik(macieiravarcomp1)*, $l_{R1} = -5200,24$. l_{R0} é a log-verosimilhança restrita do modelo que está em H_0 (modelo sem efeitos genotípicos, um submodelo do modelo que está em H_1) e no enunciado é dada com o comando *logLik(macieiravarcomp2)*, $l_{R0} = -5281,984$. Assim, $\Lambda_{calc} = 163,486$. Para $\alpha = 0,05$, $\chi_{0,05(1)}^2 = 3,84$. Logo, rejeita-se H_0 para $\alpha = 0,05$. Conclui-se que existe variabilidade genética do rendimento.

3. Fez-se o ajustamento do modelo linear misto usando a função *asreml* do package *ASREML-R* do R. O código descrito é muito semelhante ao descrito anteriormente. Primeiro, é declarada a variável resposta (rend) e a seguir os factores de efeitos fixos (neste caso, sem factores de efeitos fixos), o que corresponde ao código `rend~1`. Depois, são declarados os factores de efeitos aleatórios, o que corresponde ao código `random=~clone+...`

a) Genericamente

Modelo 1

Valor fenotípico (rend) = média populacional + efeito genotípico + erro aleatório

$$g_i \cap \mathcal{N}(0, \sigma_g^2), \forall i; e_{ij} \cap \mathcal{N}(0, \sigma_e^2), \forall i, j; g_i \text{ e } e_{ij} \text{ são mutuamente independentes.}$$

Modelo 2

Valor fenotípico (rend) = média populacional + efeito genotípico+ efeito bloco+ erro aleatório

$g_i \cap \mathcal{N}(0, \sigma^2_g)$, $\forall i$; $b_j \cap \mathcal{N}(0, \sigma^2_b)$, $\forall j$; $e_{ij} \cap \mathcal{N}(0, \sigma^2_e)$, $\forall i, j$; g_i , b_j e e_{ij} são mutuamente independentes.

Modelo 3

Valor fenotípico (rend) = média populacional + efeito genotípico + efeito repetição resolúvel + efeito linha(rep) + efeito coluna(rep) + erro aleatório

$g_i \cap \mathcal{N}(0, \sigma^2_g)$, $\forall i$; $rep_j \cap \mathcal{N}(0, \sigma^2_r)$, $\forall j$, $linha(rep)_{kj} \cap \mathcal{N}(0, \sigma^2_{linha(r)})$, $\forall k, j$;
 $coluna(rep)_{lj} \cap \mathcal{N}(0, \sigma^2_{coluna(r)})$, $\forall l, j$; $e_{ijkl} \cap \mathcal{N}(0, \sigma^2_e)$, $\forall i, j, k, l$; g_i , rep_j , $linha(rep)_{kj}$,
 $coluna(rep)_{lj}$ e e_{ijkl} são mutuamente independentes.

Destes 3 modelos, o modelo que reflete o delineamento experimental adotado é o modelo 3 uma vez que inclui todos os fatores associados ao delineamento experimental usado. Com base na informação disponível verifica-se que a estimativa da variância dos erros aleatórios diminuiu à medida que foram introduzidos os factores relativos ao delineamento experimental e a estimativa da variância genotípica aumentou (e o respetivo erro padrão diminuiu, o que revela maior precisão). Mais uma vez, este resultado revela a importância do delineamento experimental para o aumento da eficiência da seleção, assim como alerta para as consequências de uma análise não incluir todos os factores que contribuem para o valor obtido da variável resposta (maior proporção da variabilidade não explicada pelo modelo e estimativas menos precisas da variância genética).

b) As médias das estimativas da variância dos erros de predição dos efeitos genotípicos para os 3 modelos foram, respetivamente,

$$\overline{PEV}_{modelo1} = 0,396971, \overline{PEV}_{modelo2} = 0,380954, \overline{PEV}_{modelo3} = 0,361294.$$

Uma vez que para os modelos 2 e 3 o cálculo da heritabilidade em sentido lato não obedece às condições clássicas (temos vários factores de efeitos aleatórios), calcula-se a heritabilidade em sentido lato com base na expressão

$$h_g^2 = 1 - \frac{\overline{PEV}}{\hat{\sigma}_g^2}.$$

Assim

$$\text{Modelo 1: } h_g^2 = 1 - \frac{0,396971}{0,8688656} = 0,543.$$

$$\text{Modelo 2: } h_g^2 = 1 - \frac{0,380954}{0,9565373} = 0,6017.$$

$$\text{Modelo 3: } h_g^2 = 1 - \frac{0,361294}{0,9949243} = 0,6369.$$

Estes resultados reforçam o que foi dito na alínea anterior. Com o ajustamento do modelo 3 maior proporção da variabilidade fenotípica é explicada por causas genótípicas, logo, está associada uma maior eficiência na seleção.

c)

i) Em genética quantitativa, a precisão na previsão dos efeitos genéticos é dada pelo quadrado da correlação entre os valores verdadeiros e previstos para os efeitos genéticos (REL). Assim, a precisão associada ao EBLUP do efeito genotípico do clone AR0223 é dada por:

$$REL_{AR0223} = 1 - \frac{PEV_{AR0223}}{\hat{\sigma}_g^2} = 1 - \frac{0,376519139}{0,8688656} = 0,566654$$

ii) O modelo 1 trata-se de um modelo aleatório clássico. Nestas condições, a estimativa da variância do erro de predição (PEV) do efeito genotípico para o genótipo i é dada por: $PEV_i = \frac{\hat{\sigma}_e^2 \hat{\sigma}_g^2}{r\hat{\sigma}_g^2 + \hat{\sigma}_e^2}$. Pela expressão é fácil perceber que para um delineamento equilibrado (isto é, quando todos os genótipos são avaliados no mesmo número de repetições), a estimativa da variância do erro de predição (PEV) é igual para todos os clones. Perante os resultados do enunciado, percebe-se que a PEV varia de acordo com o clone, portanto, varia com o número de repetições observadas para cada clone (na expressão acima, apenas o r muda). Quanto maior o número de repetições, menor a PEV.

De igual modo, no caso do modelo 1 ajustado, o EBLUP do efeito genotípico do genótipo i é dado por:

$$EBLUP_{gen_i} = (\text{média do genótipo } i - \text{media geral}) \times \left(\frac{\hat{\sigma}_g^2}{\hat{\sigma}_g^2 + \hat{\sigma}_e^2/r} \right).$$

Ora, verifica-se que o termo $\left(\frac{\hat{\sigma}_g^2}{\hat{\sigma}_g^2 + \hat{\sigma}_e^2/r} \right)$ varia de acordo com o r (número de repetições em que o genótipo foi avaliado). Consequentemente, num genótipo avaliado em mais repetições o termo $\frac{\hat{\sigma}_g^2}{\hat{\sigma}_g^2 + \hat{\sigma}_e^2/r}$ é maior, e num genótipo avaliado em menos repetições esse termo é menor (menor rendimento do genótipo é explicado por causas genótípicas). Assim sendo, a ordenação dos EBLUPs será diferente da ordenação dos rendimentos médios (valores fenotípicos médios).

d)

i) A seleção deverá ser feita com base na ordenação dos EBLUPs dos efeitos genotípicos, pois as ordenações segundo os valores fenotípicos médios do rendimento e os EBLUPs dos efeitos genotípicos do rendimento são diferentes. Isto acontece porque o modelo 3 ajustado tem 4 factores de efeitos aleatórios e os dados não são equilibrados. Apenas para modelos com 1 factor de efeitos aleatórios e dados equilibrados é que as ordenações coincidem.

ii) Trata-se de um modelo que não satisfaz as condições clássicas. Nestas circunstâncias, vamos calcular o ganho com base nos EBLUPs dos efeitos genotípicos. Assim, o ganho genético ao selecionar os 20 melhores clones:

$$R = \text{média dos EBLUPs dos 20 melhores clones} = 1,338 \text{ kg/planta.}$$

$$R(\% \text{ média}) = 29,4\%$$

4.

a) Teste à componente de variância genotípica através de um teste de razão de verossimilhanças restritas.

Hipóteses: $H_0: \sigma_g^2 = 0$ vs $H_1: \sigma_g^2 > 0$; Rejeita-se H_0 , para um dado nível de significância α , se $\Lambda_{calc} > \chi_{\alpha(1)}^2$. Sabemos que $\Lambda_{calc} = 273,2$. Para $\alpha = 0,05$, $\chi_{0,05(1)}^2 = 3,84$. Logo, rejeita-se H_0 para $\alpha = 0,05$. Conclui-se que existe variabilidade genética do rendimento.

b) $h_g^2 = 0,778$; $R(\% \text{ média}) = 33,3\%$.

c) $EBLUP(\text{clone } 1^{\text{o}}) = 0,8602 \text{ kg/planta.}$

5. a) Trata-se de um modelo que não satisfaz as condições clássicas, diz-se que é um modelo espacial. Isto quer dizer que a matriz de covariâncias dos erros aleatórios não é uma matriz diagonal, mas sim uma matriz que assume correlação espacial entre observações. Nestas circunstâncias, vamos calcular o ganho com base nos EBLUPs dos efeitos genotípicos. Assim, o ganho genético ao selecionar os 10 melhores clones:

$$R = \text{média dos EBLUPs dos 10 melhores clones} = 0,74106 \text{ kg/planta.}$$

b) No enunciado temos os valores dos erros padrão de predição. Logo, a estimativa da variância do erro de predição (PEV) associada ao EBLUP do efeito genotípico do clone C406 é: $PEV(EBLUPC406) = (0,242)^2 = 0,058564$.

c) Com base no ajustamento deste modelo, o cálculo da heritabilidade em sentido lato não obedece às condições clássicas. Neste calcula-se uma medida aproximada da heritabilidade generalizada recorrendo-se à expressão (ver slides das aulas teóricas), $h_g^2 = 1 - \frac{PEV}{\hat{\sigma}_g^2} = 0,7798$.

GRUPO II

Cálculo de parâmetros genéticos em populações de linhas puras de espécies de autofecundação

2. $h_{(A)}^2 = 0,26$ (26% da variabilidade fenotípica na F2 é explicada por causas genéticas aditivas, ou seja, é a correspondência entre valores fenotípicos e reprodutivos).

3. $h_{(G)}^2 = 0,75$; $h_{(A)}^2 = 0,626$.

PARTE II

GRUPO I

1.

a) No enunciado estão as estimativas das componentes de variância:

$$\hat{\sigma}_{fam}^2 = 137,3825, \hat{\sigma}_e^2 = 648,3149.$$

Assim, a variância fenotípica observada no ensaio será: $\hat{\sigma}_p^2 = \hat{\sigma}_{fam}^2 + \hat{\sigma}_e^2 = 785,6974 g^2$.

A correlação intraclassa é dada por: $t_{HS} = \frac{\hat{\sigma}_{fam}^2}{\hat{\sigma}_p^2} = 0,174854$.

$t_{HS} = 0,174854$, significa que 17,5% da variabilidade fenotípica é explicada pela semelhança entre meios-irmãos.

b) $\hat{\sigma}_a^2 = 4\hat{\sigma}_{fam}^2$, uma vez que $Cov_{HS} = \frac{1}{4}V_A$ e $Cov_{HS} = Cov(y_{ij}, y_{ij'}) = \sigma_{fam}^2$. Assim, $\hat{\sigma}_a^2 = 4 \times 137,3825 = 549,53 g^2$

A variância aditiva (variância entre valores reprodutivos) dos indivíduos na população de referência, é $549,53 g^2$. Não esquecer que o valor reprodutivo representa a parte do valor do genotípico que é passado à descendência, depois de cruzamento aleatório.

c) Dado o modelo ajustado, a heritabilidade em sentido restrito é calculada como:

$$h_{(A)}^2 = \frac{4\hat{\sigma}_{fam}^2}{\hat{\sigma}_p^2}, \text{ uma vez que } h_{(A)}^2 = \frac{V_A}{V_P}.$$

Assim, $h_{(A)}^2 = 0,699417$.

É a proporção da variância fenotípica que é explicada pela variância entre valores reprodutivos dos indivíduos (progenitores femininos, neste caso, árvores mãe) na população de referência.

Nota: Uma população de referência é uma grande população onde estão aleatoriamente dispostas as árvores mãe sujeitas a cruzamento aleatório, portanto, de acordo com as condições do equilíbrio de Hardy-Weinberg.

O ensaio de descendências é o ensaio que foi estabelecido com a descendência de cada uma das árvores mãe (família de meios irmãos). É esse ensaio de famílias de meios-irmãos que foi alvo de avaliações.

d) O EBLUP da aptidão geral combinatória (GCA) do progenitor feminino F5 está na listagem dos melhores preditores lineares não enviesados empíricos (EBLUPs) dos efeitos das famílias. Ou seja, os EBLUPs dos valores da GCA do progenitor feminino de cada família são os EBLUPs do valor reprodutivo médio da família de meios-irmãos.

Assim, o EBLUP da GCA do progenitor feminino F5 é $F5 = 12,238 g$. Significa que na família F5 (ou seja, na descendência de meios-irmãos deste progenitor) o peso seco da pinha será $12,238g$ acima da média da população.

A estimativa do quadrado da correlação entre os valores verdadeiros e previstos para os efeitos genéticos das famílias é dada por:

$$REL_i = 1 - \frac{PEV_i}{(1 + F)\hat{\sigma}_{fam}^2}$$

em que PEV é a estimativa da variância do erro de predição do efeito da família i , F é o coeficiente de endogamia dos indivíduos. Neste caso, $F=0$, pois admite-se que os progenitores femininos não têm relações de parentesco.

Assim, a precisão associada ao EBLUP da família F5 (REL) é:

$$REL_{F5} = 1 - \frac{PEV_{F5}}{\hat{\sigma}_{fam}^2} = 1 - \frac{59.43488}{137,3825} = 0,56737.$$

Ou seja, a correspondência entre os valores verdadeiros e previstos para os efeitos genéticos da família F5 é de 56,7%.

e) Pede-se a previsão do valor reprodutivo do progenitor feminino F5.

O valor médio do valor reprodutivo de uma descendência (família de meios-irmãos), ou seja, o efeito da família de meios-irmãos é metade do valor de reprodutivo do seu progenitor feminino. Esta relação significa que, em média, um progenitor passa metade do seu valor reprodutivo para a sua descendência. Então o EBLUP da GCA_{F5} representa o EBLUP de $\frac{1}{2}$ do valor reprodutivo do progenitor F5.

Assim, o EBLUP do valor reprodutivo do progenitor F5 será $2 \times 12,238g = 24,476g$. Prevê-se que, devido aos efeitos médios dos seus alelos, o peso seco da pinha do progenitor F5 seja 24,476g acima da média de todas as árvores na população de referência.

f) Como os dados não são equilibrados, vamos prever o ganho genético com base nos valores reprodutivos dos 5 melhores progenitores femininos. Esses valores reprodutivos vão ser o dobro dos EBLUPs dos efeitos das famílias, logo, para os 5 melhores progenitores femininos serão:

24.47618
10.72507
10.56784
8.681142
6.354374

A média destes 5 valores reprodutivos é 12.16092 g. Logo, em % da média da população:

$$R(\%) = \frac{12.16092}{241.1385} \times 100 = 5,04\%$$

Isto significa que o ganho genético previsto no peso seco do pinha ao selecionar os 5 melhores progenitores femininos na população de referência é da ordem dos 5%.

2.

a) Ajustou-se agora modelo individual para famílias de meios-irmãos.

Agora, a estimativa da variância genética aditiva é obtida com o ajustamento do modelo, $\hat{\sigma}_a^2 = 553.0749 g^2$. Como se verifica é um valor muito parecido ao obtido na alínea 1b).

O valor da heritabilidade em sentido restrito é:

$$h_{(A)}^2 = \frac{\hat{\sigma}_a^2}{\hat{\sigma}_p^2} = \frac{553.0749}{553.0749+237.7046} = 0,6994.$$

Um valor também idêntico ao obtido na pergunta 1c).

Confirma-se que, na prática, ambos os modelos podem ser ajustados, revelando resultados similares em termos de parâmetros de genética quantitativa.

b) Com o ajustamento do modelo individual com matriz de *Pedigree*, são obtidos os EBLUPs dos valores reprodutivos de todos os indivíduos (progenitores femininos e descendentes). Se tomarmos em conta os valores dos EBLUPs dos efeitos dos progenitores femininos, verifica-se que são similares aos valores obtidos na pergunta 1. Confirma-se mais uma vez que, em termos de seleção, ambos os modelos fornecem resultados idênticos.

GRUPO II

1.

a) $h_{(A)}^2 = 0,56$. É a proporção da variabilidade fenotípica que é explicada pela variância entre valores reprodutivos dos progenitores femininos na população de referência. Ou seja, a correspondência entre valores fenotípicos e reprodutivos é de 56%.

Nota: teoricamente, admitindo a lei do equilíbrio de Hardy-Weinberg, deduz-se relação $\hat{\sigma}_a^2 = 4\hat{\sigma}_{fam}^2$. Na prática, quando se considera $\hat{\sigma}_a^2 = 4\hat{\sigma}_{fam}^2$, muitas vezes sobrestima-se a variância genética aditiva.

b) Teste $H_0: \sigma_{fam}^2 = 0$ vs $H_1: \sigma_{fam}^2 > 0$; $\Lambda_{calc} = 36,8$. Comparando com $\chi_{0,05(1)}^2 = 3,84$, rejeita-se H_0 para $\alpha = 0,05$. Conclui-se que existe variabilidade genética entre famílias.

c) O EBLUP da aptidão combinatória geral do progenitor feminino F170 é $GCA_{F170} = 0,5720212$. Este valor representa $\frac{1}{2}$ do EBLUP valor reprodutivo do progenitor feminino F170.

Assim, o EBLUP do valor reprodutivo do progenitor feminino F170 será $2 \times 0,5720212 = 1,144$. Prevê-se que, devido aos efeitos médios dos seus alelos, a altura do progenitor F170 seja 1,144 m acima da média de todas as árvores na população de referência.

d) $h_{fam}^2 = 0,765$. Este valor significa que 76,5% da variabilidade fenotípica ao nível da média da família é devida a diferenças genéticas entre famílias. É um indicador da qualidade do ensaio.

Nota: quando se desconfia da validade da lei do equilíbrio de Hardy-Weinberg na população de referência, é preferível calcular apenas a heritabilidade ao nível das médias das famílias e calcular o ganho genético com base nesta heritabilidade. Caso contrário, poderemos estar a sobrestimar os ganhos genéticos de seleção.

$$e) h_{fam}^2 = 0,76$$

$$f) h_{fam}^2 = 0,76$$

A heritabilidade ao nível da média das famílias foi calculada com base em 3 expressões equivalentes. Como o número de árvores por parcela não é igual, as expressões correntemente mais usadas são as descritas nas alíneas e) e f).

g) Pede-se ganho genético previsto ao selecionar os 10 indivíduos (progenitores femininos) mais altos.

Neste caso temos disponíveis os EBLUPs dos efeitos das famílias e a média das 10 mais altas. Sabemos que o valor reprodutivo dos progenitores femininos é o dobro dos valores que estão no enunciado.

$$\text{Logo, } R = 2 \times 0,6608 = 1,3216. R(\%) = \frac{1,3216}{10,47845} \times 100 = 12,6\%.$$

Isto significa que o ganho genético previsto em altura ao selecionar os 10 melhores progenitores femininos na população de referência é de 12,6%.

No ensaio de descendências, ao selecionar as 10 melhores famílias em altura o ganho genético previsto é de 6,3% (o ganho genético calculado como a média dos EBLUPS dos efeitos das famílias selecionadas).

2. Decidiu-se optar por fazer a análise com base nos valores médios das famílias por parcela. O valor obtido é $h_{fam}^2 = 0,77$. Como se observa, o valor da heritabilidade ao nível da média das famílias é similar ao obtido com a análise dos valores individuais por parcela (incluindo o factor plot no modelo).

Por esta via é apenas possível calcular a heritabilidade ao nível da média das famílias (isto é, não é possível calcular a heritabilidade em sentido restrito que é uma heritabilidade a nível individual; também não se pode ajustar um modelo individual com matriz de *Pedigree*). É importante que um investigador antes da colheita de dados num ensaio esteja ciente desta realidade.

3. a) Componente de variância das famílias e o erro padrão dessa estimativa:

$$\hat{\sigma}_{fam}^2 = 0,189; SE(\hat{\sigma}_{fam}^2) = \sqrt{\widehat{var}(\hat{\sigma}_{fam}^2)} = 0,084$$

b) A variância fenotípica é dada por:

$$\hat{\sigma}_p^2 = \hat{\sigma}_{fam}^2 + \hat{\sigma}_{plot}^2 + \hat{\sigma}_e^2 = 0,1890664 + 0,1993867 + 2,5259706 = 2,914424$$

c) A variância genética aditiva

$$\text{Vimos que numa família de meios-irmãos, } \hat{\sigma}_a^2 = 4\hat{\sigma}_{fam}^2 = 4 \times 0,1890664 = 0,7562656.$$

É necessário ter muito cuidado com a estimativa da variância aditiva. Teoricamente, admitindo a lei do equilíbrio de Hardy-Weinberg, deduz-se relação $\hat{\sigma}_a^2 = 4\hat{\sigma}_{fam}^2$. Na prática, quando se considera $\hat{\sigma}_a^2 = 4\hat{\sigma}_{fam}^2$, muitas vezes sobrestima-se a variância genética aditiva.

d) A heritabilidade em sentido restrito é dado por:

$$h_{(A)}^2 = \frac{\hat{\sigma}_a^2}{\hat{\sigma}_p^2} = \frac{4\hat{\sigma}_{fam}^2}{\hat{\sigma}_{fam}^2 + \hat{\sigma}_{plot}^2 + \hat{\sigma}_e^2} = 0,26.$$

Com este modelo, isto é, com a inclusão do factor proveniência e considerando o factor família subordinado à proveniência, grande parte da variabilidade anteriormente atribuída às diferenças entre famílias foi agora alocada às diferenças entre proveniências. Como resultado, a estimativa da variância das famílias é menor, pois agora o seu significado é diferente, trata-se da variabilidade das famílias na proveniência.

GRUPO III

1.

a) Estimativa da componente de variância dos erros aleatórios, $\hat{\sigma}_e^2 = 20$ ($QMRE = 20$).

b) Pela tabela de análise de variância $QMf = \hat{\sigma}_e^2 + r\hat{\sigma}_{m(f)}^2 + rb\hat{\sigma}_f^2$ e $QMm(f) = \hat{\sigma}_e^2 + r\hat{\sigma}_{m(f)}^2$

$$\text{logo, } \hat{\sigma}_f^2 = \frac{QMf - QMm(f)}{rb} = \frac{470 - 170}{30} = 10.$$

Assim, a estimativa da variância entre progenitores femininos, $\hat{\sigma}_f^2 = 10$.

c) $QMm(f) = \hat{\sigma}_e^2 + r\hat{\sigma}_{m(f)}^2$

$$\text{Logo, } \hat{\sigma}_m^2 = \frac{QMm(f) - QMRE}{r} = \frac{170 - 20}{10} = 15.$$

Assim, a estimativa da variância entre progenitores masculinos cruzados com o mesmo progenitor feminino, $\hat{\sigma}_{m(f)}^2 = 15$.

d) Com base no modelo hierárquico considerado, a estimativa da variância genética aditiva pode ser obtida como $\hat{\sigma}_a^2 = 4\hat{\sigma}_f^2$. Assim, $\hat{\sigma}_a^2 = 40$.

e). $t_{HS} = \frac{\hat{\sigma}_f^2}{\hat{\sigma}_e^2 + \hat{\sigma}_{m(f)}^2 + \hat{\sigma}_f^2} = \frac{10}{20 + 10 + 15} = 0,222$. 22,2% da variabilidade fenotípica é explicada pela semelhança entre meios-irmãos.

f) $t_{HS} = \frac{\hat{\sigma}_f^2 + \hat{\sigma}_{m(f)}^2}{\hat{\sigma}_e^2 + \hat{\sigma}_{m(f)}^2 + \hat{\sigma}_f^2} = \frac{10 + 15}{45} = 0,555$. 55,5% da variabilidade fenotípica é explicada pela semelhança entre irmãos-completos.

g) $h_{(A)}^2 = 0,89$. O grau de correspondência entre valores fenotípicos e reprodutivo é de 89%. É um valor extremamente elevado, pouco realista, resultante de um modelo que se afasta muito da lei do equilíbrio de H-W (apenas 10 progenitores femininos são cruzados com 3 progenitores masculinos e são medidos 10 descendentes em cada família de irmãos completos).

2.

a) $V_P = 476,8 \text{ mm}^2$.

b) $Cov_{HS} = 13,7$.

c) $Cov_{FS} = 90,125$.

d) $t_{HS} = 0,0287$.

e) $t_{FS} = 0,189$.

f) $h^2_{(A)} = 0,1149$.

3. Sabemos que o valor fenotípico médio para famílias de irmãos-completos, $\bar{P}_{0,FS}$ é dado por:

$$\begin{aligned}\bar{P}_{0,FS} &= \mu + \frac{1}{2}A_A + \frac{1}{2}A_B + \bar{I}_{0,FS} \\ &= \mu + GCA_A + GCA_B + SCA_{AB}\end{aligned}$$

sendo A_A o valor reprodutivo do progenitor feminino, A_B o valor reprodutivo do progenitor masculino, GCA_A a aptidão geral combinatória do progenitor feminino (metade do seu valor reprodutivo); GCA_B a aptidão geral combinatória do progenitor masculino (metade do seu valor reprodutivo); $\bar{I}_{0,FS}$ os efeitos médios da interação da família de irmãos completos, isto é, a aptidão específica combinatória (SCA_{AB}) de uma família de irmãos completos resultante do cruzamento dos dois progenitores A e B (efeitos de interação).

Assim, $SCA_{FM} = 25 - 20 - 3 - 6 = -4m$. Esta descendência de irmãos completos cresceu menos $4m$ que o previsto pelos valores reprodutivos dos progenitores (ou pelas GCA dos progenitores). Considerando os efeitos de epistasia negligenciáveis, isto é devido aos efeitos de dominância específicos do cruzamento $A \times B$.

GRUPO IV

1. Segundo o método de Finlay e Wilkinson (1963), o genótipo A é um clone com uma boa adaptação geral, o genótipo B é um genótipo com adaptação específica a ambientes favoráveis e o genótipo C é um genótipo com adaptação específica a ambientes desfavoráveis.
2. Trata-se de um *biplot* dos efeitos principais do genótipo e dos efeitos da interação genótipo×ambiente. Na interpretação deste *biplot* devemos ter em conta que o coseno do ângulo entre os vectores de dois ambientes aproxima a correlação entre esses ambientes (quanto mais pequeno o ângulo, mais positivamente correlacionados estão esses ambientes). O comprimento do vector de cada ambiente, indica a capacidade discriminante em cada ambiente. Para ordenar os genótipos num dado ambiente, traça-se uma linha perpendicular ao ambiente na origem do *biplot*.
 - a) Existem ambientes muito correlacionados (por exemplo, E1999/E1993), mas existem outros que não estão correlacionados (por exemplo, E1996/E1997), ou que o estão negativamente (por exemplo, P1989/P1998, que significa que os melhores genótipos num ambiente são os piores no outro). Ora, os dois últimos exemplos revelam a existência de interação $G \times E$.

b) Por exemplo: clone CR9106 tem bom rendimento em E1996, E1999 e restantes ambientes desse quadrante, e em E1998; tem um rendimento à volta da média no ambiente E1997; tem um mau rendimento em P1989.

Por exemplo: clone CR 4507 tem bom rendimento em P1989 e E1997; tem um rendimento à volta da média no ambiente E1994 e tem um rendimento abaixo da média em todos os ambientes restantes.

3. Trata-se de uma interpretação análoga à da pergunta 2.

GRUPO V

1. Trata-se do ajustamento de um modelo linear misto multivariado. Neste tipo de modelo, o vector Y das observações é composto por vários subvectores, cada um referente aos dados de uma dada característica. A matriz de covariâncias dos efeitos genotípicos é definida de modo a que se estime para todas as características a respectiva componente de variância genotípica e as covariâncias genéticas entre todos os pares possíveis de características. Tendo esta informação é possível estimar a correlação genética entre características. No exercício, pede-se a correlação genética entre rendimento e grau brix que pode ser calculada como: $r_{g_{rend,brix}} = \frac{\widehat{cov}_{g(rend,brix)}}{\widehat{\sigma}_{g(rend)}\widehat{\sigma}_{g(brix)}}$. Assim, $r_{g_{rend,brix}} = -0.416$, o que significa que as duas características apresentam uma correlação genética negativa moderada.

2. A correlação genética positiva entre rendimento e grau álcool provável indica que será possível ao seleccionar para uma característica obter ganhos genéticos de seleção para ambas as características consideradas. A correlação genética negativa entre acidez e grau álcool provável significa que a seleção a favor de uma característica poderá implicar perda na outra. Observa-se uma moderada-fraca correlação genética negativa entre acidez e rendimento, talvez se consiga seleccionar com ganhos genéticos para uma, sem perdas consideráveis para a outra.