PRACTICAL

Exercise. Grape Data

A genetic experiment was established for a variety of grape vine (*Vitis vinifera*) based in a resolvable IBD based on 4 replicates each with 4 incomplete blocks with row-plots based on 4 plants. A total of 16 unrelated genotypes (i.e. clones) were evaluated. The variable of interest is yield (YD, kg/plant). The data can be found in the file /Practicals/FieldT/FIELDT.txt. Note that for the response variable the records with −9 correspond to a missing value.

a) Fit the following mode to this data:

$$y = \mu + Rep + Rep(IBlock) + Genotype + Plot(Rep) + e$$

Consider all factor terms random with the exception of Rep.

b) Check residuals plots and and calculate an estimate of broad-sense heritability (H^2 = Var(Genotype)/Vtotal). Do you think you have important genetic control for this trait?

c) Obtain the predictions for each of the genotypes. Which one has the greatest EBV?

PRACTICAL

Exercise. Grape Data
Revisit Exercise 1.

a) Test the significance of the heritability (i.e. Var(Genotype)) by using a one-sided likelihood ratio test (use $\alpha = 0.05$).

b) Refit your model but this time, incorrectly, consider the term *Plot(Rep)* a fixed effect. What differences do you observe on your analyses? (Check heritability, predictions, Wald test, etc.).

PRACTICAL

Exercise. Rubber tree Data

A trial for Rubber tree was established that evaluates total height of plants at 6 months. Individuals were planted in single tree plots and they belong to several crosses established over 20 replicates as a RCBD. The data are presented in the file `/Practicals/Rubber/RUBBER.TXT`.

a) Read the data and check in how crosses each of the parents (Male or Female) are represented in this data. Do you think this is a good pattern to estimate family effects?

b) Fit a full-sib model. Estimate additive and dominance variance for this analysis. Obtain a ranking for Male, Female and Family.

c) Evaluate what are the changes occur (in variances, log-likelihood value and ranking (predictions) of parents) when you drop the term Family. Why are these changes ocurring?

PRACTICAL

Exercise. Pine Data

A tree genetic study consisting on seeds from a total of 28 female parents were collected from mass selection and tested in a RCBD together with 3 control female parents (`/Day1/Indiv/OPENPOLs.txt`. The experiment consisted in 10 replicates with 34 plots each of size 2 x 3. The response variables of interest are total height (HT, cm) and diameter at breast height (DBH, cm). For now we will concentrate in the response HT. The objective is to rank the all individuals (parents and offspring) for future selections and seed production.

a) Construct an pedigree file (PEDIND.TXT) that has the records for each of the individuals in the data file. (Note: consider all sires/males as 0).

b) Fit a ANIMAL/INDIVIDUAL model. You can use the previous parental model as a reference but make sure to include the pedigree file this time. (Note: ignore the differences between control and test genotypes).

c) Obtain your predictions and sort your genotypes from larger to smaller (use Excel). Does the prediction of the top individual agrees with the phenotypes from all individuals from the same family?

PRACTICAL

Exercise. Willow Data

A study was established using a collection of 428 willow selections from a large genetic bank. The experimental design consisted in a completely randomized design with individuals selected from a total of 32 single crosses (i.e. 32 families). The response of interest is biomass production at 3 different harvesting times (WT1, WT2, and WT3), which occurred every 3 years. The objective of the study is to be able to rank those families and select those that have the greatest yield across these 3 measurements. For the analyses a family model (i.e. FAM) will be fitted, and the different populations (i.e. Pop) will be considered as a fixed effect to denote the different blocks (or sets). For now, we will concentrate in WT3.

a) Perform a check of the data for response WT3 to evaluate its quality and to detect any potential errors or suspicious structure. Do you observe something of concern for the analysis?

b) Fit a simple family model for this response and calculate the variance explained by the family term.

c) Modify the your previous model to consider a different error variance for each of the populations. Fit this model and evaluate it using a LRT. Which model do you prefer?

PRACTICAL

Exercise. Willow Data

We will revisit Exercise 6 using willow selections. Recall that the response of interest is biomass production at 3 different harvesting times (WT1, WT2, and WT3). This time, we will perform multivariate analysis based on the family model. The most interesting correlations are the ones between WT1 and WT3, and therefore a bivariate analysis will be performed.

a) In order to get good starting values you need to fit a simple family model for the responses WT1 and WT3 individually. Are their measurement scales very different?

b) Fit a bivariate analysis for these two traits. Consider a us structure for both Family and error term.

c) Refit your model from (b) but this time consider a corgh structure for both of your random terms.

Optional: Fit a multivariate model by incorporating all3 response variables. Hint: Use corgh for the family term and us for the error term.

PRACTICAL

Exercise. Pine MET Data

A series of 3 trials was established using seeds from half-sib families. These trials have almost all the families (females) present, and they were measured for total tree height (HT) and diameter (DBH) at age 8. There is interest about evaluating the magnitude of the GxE in this population.

a) Perform an analysis for each of the sites by using a half-sib parental model. Concentrate in HT. Calculate heritabilites and identify any potential outliers. Do you need to perform any transformation?

b) Using all data fit a Explicit GxE model and calculate the Type B genetic correlation. Is this a good result?

c) Obtain the predictions for each of the females across all sites and the predictions for each of the sites. For the top overall females do you see a large change in their predictions across sites?

Optional: Refit the previous model but this time consider an Implicit GxE model using a corgh structure.