

Description of Data

We will use the data from example 11.1 in Mrode (2013). See text file uploaded to Absalon, and use `read.table` with `header=TRUE` to read it into R..

- Genotypes are coded as 0 (homozygote, first allele), 1 (heterozygote), 2 (homozygote, second allele)
- The observations for the 10 bulls are the Daughter Yield Deviations (DYDs) for fat yield, and the Effective Daughter Contributions (EDCs) are also given.
- Genetic variance for fat yield is $35.241kg^2$
- Residual variance is $245kg^2$.
- Bulls 13-20 are assumed to be reference animals
- Bulls 21-26 are assumed to be selection candidates, i.e. we pretend that do not have records

Exercise:

- 1. Compute allele frequencies, and compute the centered matrix of genotypes \mathbf{Z} as defined in Section 11.3 (Mrode, 2015).
- 2. Set-up the mixed model equations for estimating SNP effects based on reference bulls, where you for simplicity assume that matrix \mathbf{R} is the identity matrix.
- 3. What parameter value of α should you use?
- 4. Estimate the SNP effects
- 5. Compute DGV's for selection candidates
- 6. Compare the DGV for selection candidates with the actual records
- Extra: incorporate EDC as weights, i.e. redo 2-5.