

Exercise Genetics

Generell Information

This exercise requires the data set *SNP.RData*, containing 20,000 samples and 4 columns. The column description is as follows:

Variable	Description	Coding
id	Sample ID	NA
trait	phenotype	continuous
sex	sex	1 = male; 2 = female
SNP	genotype	0 = AA; 1 = AB; 2 = BB

The variable *SNP* contains data for biallelic SNP with alleles *A* and *B*.

Task 1: Allele frequencies (AF)

Determine the allele frequencies of the given SNP (rounded to three decimal places). How is the result to be interpreted?

Task 2: Hardy-Weinberg-Equilibrium (HWE)

Determine the expected genotype counts in HWE (rounded to three decimal places). Test the given SNP for deviation of HWE and state the null hypothesis to be tested. How is the result to be interpreted?

Task 3: Phenotype associations

Determine if the SNP is associated with the phenotype and test if this effect is additive, dominant or recessive (state the null hypotheses).

Task 4: Interaction

Test for a SNP-sex interaction using the genetic model determined in task 3. Visualize the result. How is the result to be interpreted?