

Computing for Statistical Genetics

Session 4: Data Manipulation

Read in `example-pheno.csv` as in Session 1, and read in corresponding genotype data from `http://faculty.washington.edu/kenrice/sisg/example-snp.txt`.

1. Using just the phenotype data, compute the mean and standard deviation blood pressure (`sbp` and `dbp`) and BMI for the whole sample, and separately for men and women.
2. Using `merge()`, merge the genotype and phenotype data sets; use the `id` variable to match up phenotype and genotype for each person. For a few subjects, check that your merging 'worked' correctly.
3. Compute the mean and standard deviation of blood pressure and BMI, by genotype, for each SNP. (Hint: note there are two loops here; over genotypes and over SNPs)