1. [50 points] In homework 4, the Pedstats software was used to analyze the data files “BMI_BIOST551.ped” and “BMI_BIOST551.dat” for heritability estimation of BMI in European pedigrees for different classes of relative pairs. You will now estimate heritability of BMI considering all pedigree individuals together in a joint analysis using mixed linear models with the QTDT software. The two data files for the European pedigrees previously used for the homework 4 BMI heritability analysis with Pedstats (available on the course website) are also in the required format for the QTDT software.

(a) Provide variance component estimates for BMI using a mixed linear model (MLM) with the following two random effects: additive polygenic and unique (or non-shared) environmental effects. Provide an estimate of the narrow sense heritability of BMI from the variance component estimates from this mixed linear model.

(b) Now provide variance components estimates for BMI assuming the following three random effects: additive polygenic, shared household, and unique environmental effects. Provide an estimate for narrow sense heritability of BMI with the variance component estimates from this mixed linear model.

(c) Is the shared household effect for BMI significant for the mixed linear model in 1(b) above? Provide evidence to support your answer.

(d) Now compare your BMI heritability estimates from the mixed linear models with variance components in 1(a) and 1(b) above to the BMI estimates obtained in homework 4 based on different relative pair types. Discuss and provide plausible explanations for any differences and/or similarities that you find in the heritability estimates.

2. [50 points] A quantitative trait $Y$ is influenced by a single autosomal locus, family effects, and unique environmental effects in a population. The autosomal locus that influences the trait is bi-allelic with alleles $A$ and $a$, and the frequency of allele $A$ is 0.15. Assume that the population is in Hardy-Weinberg equilibrium (HWE) at the locus. For an individual $j$ from family $i$ in the population, the trait has the following model:

$$Y_{ij} = 1.5 + 0.4G_{ij} + c_i + \epsilon_{ij}$$

where $G_{ij}$ is the number of copies of the $A$ allele that individual $j$ from family $i$ has, $c_i$ is the family effect on the trait for family $i$, and $\epsilon_{ij}$ is the unique environmental effect for individual $j$ in family $i$. Assume that $G_{ij}$, $c_i$, and $\epsilon_{ij}$ are independent with $c_i \sim N(0, \sigma_c^2 = 0.02)$ (i.e., $c_i$ has a normal distribution with mean 0 and variance $\sigma_c^2 = 0.02$), and $\epsilon \sim N(0, \sigma_e^2 = 0.09)$.
(a) Calculate the expected value and variance of the phenotype $Y$ in the population.

(b) What is the broad sense heritability ($H^2$) of the phenotype in the population? What is the narrow sense heritability ($h^2$) of this phenotype in the population?

(c) Consider two individuals $j$ and $k$ from family $i$. Calculate $\text{Cov}(Y_{ij}, Y_{ik})$ (i.e., the covariance of $Y_{ij}$ and $Y_{ik}$) for the following relationship types for $j$ and $k$:
   - $j$ and $k$ are mono-zygotic (MZ) twins reared in the same household.
   - $j$ and $k$ are full-siblings reared in the same household.
   - $j$ and $k$ are half-siblings reared in the same household.
   - $j$ and $k$ are half-siblings reared in different households.
   - $j$ and $k$ are first-cousins reared in different households.
   - $j$ and $k$ are unrelated but reared in the same household (e.g., one of the individuals is adopted).

(d) Now consider two individuals $j$ and $k$ who are from two different (unrelated) families $i$ and $l$, respectively, where individual $j$ is from family $i$ and individual $k$ is from family $l$. Calculate $\text{Cov}(Y_{ij}, Y_{lk})$. 