Homework 4
Due Thursday, November 20, 2014
Show All Work

1. [30 points] In lecture 8, for an outbred pair of individuals we defined k_0, k_1, and k_2, to be the probability that the pair share 0, 1, and 2 alleles, respectively, identical-by-descent (IBD) at a locus. We also obtained the mean and variance of the actual coancestry $\theta^* = \frac{1}{2}k_2^* + \frac{1}{4}k_1^*$ at a single locus in terms of k_0, k_1, and k_2, where k_i^* is 1 if the pair actually share i alleles IBD at the locus, and 0 otherwise. Now consider the average actual coancestry of 2000 independent loci, i.e., $\bar{\theta}^* = \frac{1}{2000}\sum_{m=1}^{2000}\theta_m^*$, where θ_m^* is the actual coancestry at locus m.

   (a) Calculate the mean and variance of $\bar{\theta}^*$ for two full siblings. Also obtain a 95% confidence interval for $\bar{\theta}^*$.

   (b) Redo question (a) above for a pair of relatives who are first cousins and for a pair of relatives who are double first cousins.

2. [25 points] Body mass index (BMI) is the ratio of an individual's weight (in kilograms) over the square of their height (in meters) and has traditionally been used as an indicator of obesity. A BMI of over 30 kg/m^2, for example, is often used to define obesity in a clinical setting. The files "BMI_BIOST551.ped" and "BMI_BIOST551.dat" data are from a large BMI study in European population families and can be downloaded from the course website in the "Data Sets" section. These files are in the format required for the Pedstat software. You will analyze the data to answer the following questions:

   (a) What is the full sibling correlation for BMI? Estimate heritability from this correlation? What assumption must be made for this heritability estimate to be valid?

   (b) Can a reasonable estimate for heritability be obtained using the half-siblings with this data? Explain.

   (c) What is the parent-offspring correlation for BMI? Estimate heritability from this correlation.

   (d) Now, redo your analysis stratified by sex and compare your above estimates with the sex-specific estimates (use the "–bySex" option in Pedstat). Does there appear to be evidence of a sex effect with BMI from this data?

3. [30 points] The file "BMI_TWINS.txt" contains data from a large twin study in samples with European ancestry. This file can be downloaded from the course website in the "Data Sets" section. The file contains data for 5,436 monozygotic and dizygotic same-sex twin pairs and includes the following 6 data columns: 1. PAIR (for the twin pair ID number), 2. BMI1 (BMI measurement for the first twin), 2. BMI2 (BMI measurement for the second twin) , Sex

("M" for male and "F" for female), ZYGOSITY ("MZ" or "DZ"). The following command can be used in R to read in the data:

```
mydata=read.table("BMI_TWINS.txt",header=TRUE).
```

  (a) Estimate the MZ twin correlation for BMI. What is an estimate of the heritability from this correlation?

  (b) Estimate the DZ twin correlation for BMI. What is an estimate of the heritability from this correlation?

  (c) Use both MZ and DZ correlations to estimate BMI heritability

  (d) Use both MZ and DZ correlations to estimate shared environmental variance on BMI. Is there evidence of a shared environmental effect?

  (e) Re-do your analysis stratified by sex (i.e., analyze male twins and female twins separately) and compare your above estimates with the sex-specific estimates. Does there appear to be evidence of a sex effect with BMI in the BMI twin study?

4. [15 points] Carefully read the article entitled "Genetic and Environmental Factors in Relative Body Weight and Human Adiposity", which can be found in the "Journal Article" section of the course website. Compare the correlation and heritability results and conclusions from the article to your BMI results for both the family and twin data studies (including the sex-specific analyses) in questions 2 and 3 above. Are the results from your analysis consistent with other BMI studies discussed in the article? Explain and include correlation and heritability estimates from the article for the different relative pair types considered in your analysis to support your answer.