Biostatistics 666
Problem Set 4
Due November 2, 2017

ApoE Haplotypes, LDL Cholesterol and Heart Disease.

Genetic variants in the ApoE gene have been associated with LDL cholesterol levels and susceptibility to heart disease.

a) Suppose that genotypes for several markers in the ApoE gene were measured in a set of heart disease cases and controls. Sketch out a strategy for evaluating association between ApoE haplotypes and disease outcome.

b) How could you ensure p-values calculated using the strategy you recommend above are accurate?

c) Your colleagues estimated haplotype frequencies for the sample using Clark's algorithm for haplotype frequency estimation. Describe an advantage and a disadvantage of the algorithm.

d) Propose an alternate and improved algorithm for haplotype frequency estimation. How does the algorithm improve upon Clark's method?

e) Suppose your colleagues have now measured ApoE genotypes in a population sample including 1,000 individuals for which measurements of LDL cholesterol levels are available. Sketch out a strategy for evaluating association between ApoE haplotypes and disease outcome.