BIOSTAT551 Statistical Genetics II: Quantitative Traits **Winter 2004**
Homework 1

General Instructions: For all problems, show how you determined your answer. Be sure to define all variables that you use and justify your logic. If you have used software to obtain answers, output should be attached, but this is not a substitute for showing the steps of your analysis!

1.1 Assuming a single diallelic locus that is in Hardy-Weinberg Equilibrium, show that the dominance genetic variance is equal to $\sigma_D^2 = (2p_1p_2ak)^2$. (5pts)

1.2 The apolipoprotein E polymorphism is a well-studied polymorphism that has been related to many complex traits such as abnormalities of blood lipids, cardiovascular disease and Alzheimer disease [OMIM reference #107741 APOLIPOPROTEIN E; APOE]. The following table summarizes the average cholesterol levels (mg/dl) in 600 normocholesterolemic men.

Genotype	Number in Sample	Average cholesterol (mg/dl)
ε2ε2	3	133
ε2ε4	19	183
ε3ε3	355	192
ε3ε2	75	182
ε3ε4	143	193
ε4ε4	5	207

a) Test whether the apo E polymorphism is in Hardy-Weinberg equilibrium. Be sure to state your level of significance and conclusions. (5pts)

b) Assuming Hardy Weinberg equilibrium, compute the additive and dominance variance for the apo E effect on cholesterol. (5pts)

c) Now, compute the additive and dominance variance assuming that the population is in Hardy Weinberg Disequilibrium. Note that this will require you to find the least squares solutions based on a 3 allele system. (10pts)

1.3 Up to now, we have been discussing different parameterizations for the expected phenotype given an individual's genotype. Suppose that in addition to the effect of a single major locus,G_{ij}, there is a random environmental influence, e_{ijk}:

$Y_{ijk} = G_{ij} + e_{ijk}$ where $i,j \in \{1,2\}$, $k=1,\ldots,N_{ij}$ and $e_{ijk} \sim N(0, \sigma_e^2)$.

Here i and j correspond to an individual's genotype G_{ij} with N_{ij} individuals having this genotype.

Provide a means for testing whether there is significant evidence for the presence of dominance. (10pts)

1.4 Consider the data from the second table of handout 2:

		Locus 1	
		B-	bb
Locus 2	C-	1.44	0.77
	cc	0.94	0.77

Suppose the frequency of the bb homozygote is 0.5, the frequency of the cc homozygote is 0.3, the population is in HWE and that the two loci are unlinked. Find the following effects: α_B, α_C, δ_{BB}, δ_{CC}, $(\alpha\alpha)_{BC}$, $(\alpha\delta)_{BCC}$, $(\alpha\delta)_{BBC}$ and $(\delta\delta)_{BBCC}$. (24pts)

1.5 The following data from Mourant et al. (1976) consist of genotypes for the M and S blood groups on 537 individuals.

		Locus for Blood Group S		
		SS	Ss	ss
Locus for Blood Group M	MM	91	147	85
	MN	32	78	75
	NN	5	17	7

a) Estimate the linkage disequilibrium coefficient. (5pts)
b) Is there evidence of gametic phase disequilibrium between the two loci? (5pts)

1.6 Later in the quarter, we will study methods that utilize family data to estimate the genetic variance for a phenotype. Generally, these methods assume that there is no covariance between genotype and environment. If this assumption is violated, what will be the effect on your estimate of σ_G^2 if gene-by-environment interaction and the effect of environment have been taken into account. (10pts)