

## Biostatistics 666

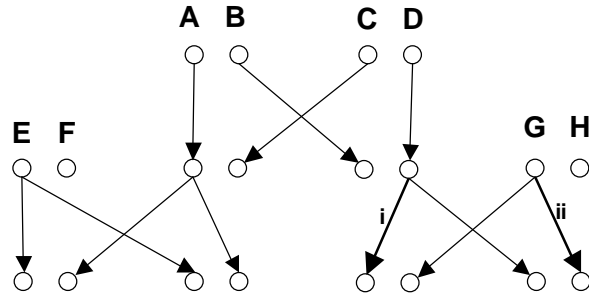
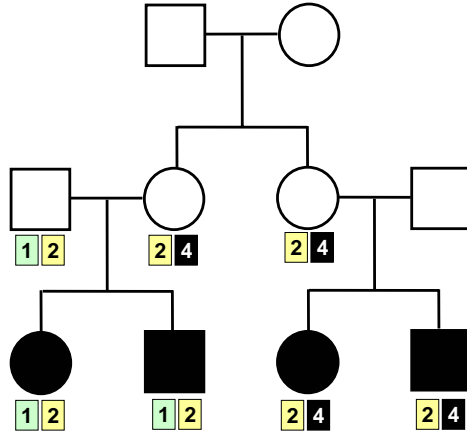
### Problem Set 7

Due April 6, 2006

### Inheritance Vector Based Pedigree Analysis

1. Consider a pair of affected siblings genotyped for a marker with 2 alleles, with frequencies  $p_1 = 0.6$  and  $p_2 = 0.4$ . The two siblings and their mother were genotyped and all three had genotype 1/2.
  - a) List all possible inheritance vectors for this pedigree. For convenience, it may be useful to apply the founder symmetry of Kruglyak and colleagues to reduce the number of inheritance vectors.
  - b) Calculate the probability of observed genotypes for each of the inheritance vectors and calculate the posterior probability of each inheritance vector conditional on the observed genotypes.
  - c) Calculate information content for this family.
  - d) Calculate the Z-score for a non-parametric linkage statistic based on the  $S_{pairs}$  scoring function. The  $S_{pairs}$  scoring function sums IBD sharing among all pairs affected individuals for each inheritance vector.
  - e) Calculate the parametric LOD score for this family, assuming a disease model where the disease allele frequency is  $p_D = 0.001$  and the disease allele frequencies are  $f_{DD} = 0.20$ ,  $f_{Dd} = 0.10$  and  $f_{dd} = 0.0$ .

2. Consider the pedigree and the inheritance graph below:



Calculate the probability of the observed genotypes conditional on the inheritance graph. In your calculations, denote the allele frequency of alleles 1, 2, 3 and 4 as  $p_1$ ,  $p_2$ ,  $p_3$  and  $p_4$  respectively. You can assume there is no genotyping error.