

The questions below are of two types: multiple choice questions and questions where an explicit answer is asked for. All the multiple choice questions have one true and three false alternatives. The other type of questions will often require a series of calculations, but we are only interested in the final answer. No extra credits will be given for partial or complete solutions.

1. Assume that n families have been genotyped at a marker locus x and that the likelihood of the observed data is 0 at $\theta = 0$, 0.1 at $\theta = 0.2$, and 0.01 at $\theta = 0.5$. What is the lod score at $\theta = 0.2$? (1p)
 - $Z(0.2) = 0.1$;
 - $Z(0.2) = 1.0$;
 - $Z(0.2) = 2.3$;
 - $Z(0.2) = 10$;
2. Consider the nuclear family in Figure 1. Assume that the affected daughter carries at least one disease allele. Which one of the following penetrance models $f = (f_0, f_1, f_2)$ is impossible given the observed disease pattern? (1p)

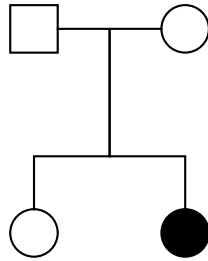


Figure 1: A nuclear family with one affected individual (filled circle).

- $f = (0, 0.8, 0.8)$;
- $f = (0, 0, 1)$;
- $f = (0, 1, 1)$;
- $f = (0.2, 0.8, 0.8)$;

3. All the individuals in Figure 2, but the affected father, have been genotyped at a marker locus x . Assuming a very rare disease allele ($p(A) = 0.0001$) and a dominant model with full penetrance and no phenocopies, is it possible to calculate a lod score at $\theta = 0$? (1p)

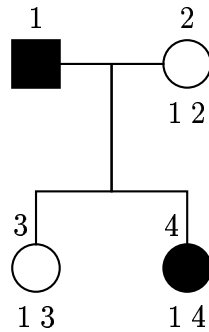


Figure 2: A partially genotyped nuclear family with two affected individuals (filled symbols).

- Yes, $Z(0) = 0$;
- Yes, $Z(0) = 0.3$;
- Yes, but $Z(0)$ will depend on the marker allele frequencies in the population;
- No;

4. Assume that we have genotyped a biallelic marker in a nuclear family with two children. The parental genotypes are (1 1) and (1 2), respectively, and the genotypes of the children are (1 1) and (1 2). The population allele frequency of the 1-allele is $p = 0.8$. What is the probability of observing the total set of genotypes: ((1 1), (1 2), (1 1), (1 2))? (1p)

Hint: Assume that the parental genotypes are independent and that the sibling genotypes are determined via Mendelian segregation. Use also the fact that genotypes of the children are independent conditional on the parental genotypes.

- $P((11), (12), (11), (12)) = 0.0419$;
- $P((11), (12), (11), (12)) = 0.0512$;
- $P((11), (12), (11), (12)) = 0.1024$;
- $P((11), (12), (11), (12)) = 0.2048$;

5. What is the interpretation of the genotypic value of a quantitative trait? (1p)

- the average trait value given the genotype;
- the average phenotypic value in the population;
- the breeding value of an individual;
- the locus-specific (broad-sense) heritability;

6. For quantitative trait influenced by a single QTL with n different alleles, how many different genotypic values are possible? (1p)

- 1;
- $n + 1$;
- $n(n + 1)/2$;
- $n(n - 1)/2 + 2$;

7. A quantitative trait is influenced by a single QTL. Only one of the following statements is true. Which one? (1p)
- The genetic correlation of two siblings sharing no alleles IBD at the QTL equals 1;
 - The genetic correlation of two siblings sharing one allele IBD at the QTL equals 1;
 - The genetic correlation of two siblings sharing two alleles IBD at the QTL equals 1;
 - The probability of two siblings sharing no alleles IBD at the QTL equals 0.5;
8. Assume that the population relative frequency of a disease allele, D , of a biallelic disease locus equals 0.001. A marker allele, M_1 , with population relative frequency 0.05 is suspected to be associated with the occurrence of the disease. Let $P(DM_1)$ denote the population relative frequency of the haplotype DM_1 . An association between D and M_1 exists if (1p)
- $P(DM_1) = 0.00005$;
 - $P(DM_1) \neq 0.00005$;
 - $P(DM_1) \leq 0.00005$;
 - $P(DM_1) \geq 0.00005$;

9. A genetic model is characterized by a disease allele frequency $p(A) = 0.01$ and by penetrances $f = (0.001, 0.001, 1)$. What is the proportion of AA-carriers among affected individuals in the population? (1p)
-

10. The probability of detecting linkage in a specific study depends on several factors like e.g. the strength of the genetic component, the sampling scheme, and the number of families included in the study. Let us assume that you are planning an ASP-study (ASP = Affected Sib Pair) and that you for some reason, e.g. based on the results of a small pilot study, suspect that the proportion of ASP:s sharing 0, 1, and 2 alleles IBD will be 10%, 45%, and 45%, respectively. Assuming perfect marker information, how many ASP-families do you need to reach an NPL-score above 4.1? (1p)
-

11. A DNA-marker M has three alleles labelled 1, 2, and 3. Fill in possible genotypes for the affected sibs in Figure 3, on the dotted lines, in such a way that it can be unambiguously decided that the number of alleles shared IBS and IBD is 1 and 0, respectively. (1p)

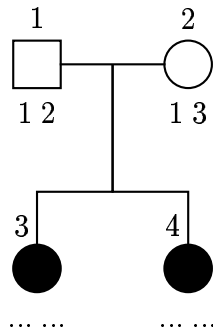


Figure 3: An affected sib pair with known parental genotypes.

12. All members of a the pedigree in Figure 4 have been genotyped at a marker locus x . Let us assume a dominant model with full penetrance and without phenocopies. Furthermore, we assume that the disease allele is so rare that the disease genotypes are "known".

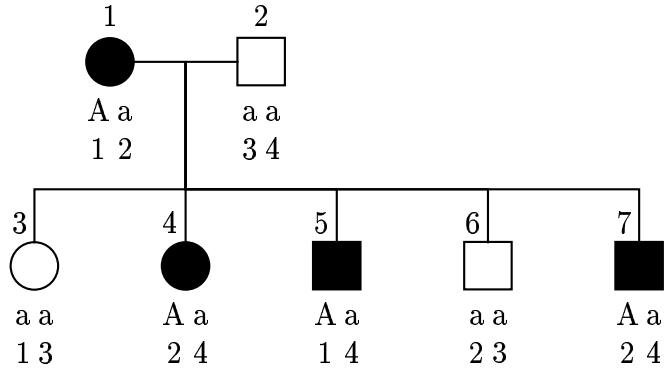


Figure 4: A nuclear family with three affected and two unaffected children.

- (a) Calculate the lod score at $\theta = 0.1$; (1p)

- (b) For which $\theta \in [0, 0.5]$ is the maximum lod score reached? (1p)

- (c) What is the maximum lod score? (1p)

13. A quantitative trait is influenced by a single QTL with 2 different alleles, A_1 and A_2 . Assume the trait value, Y , can be modelled according to $Y = X + e$, where the mean zero environmental deviation e is uncorrelated with X . Suppose that the average trait values in the population are: 1 for A_1A_1 -homozygous individuals, 3 for A_1A_2 -heterozygous individuals and 4 for A_2A_2 -homozygous individuals. Suppose further that the population relative frequency of the A_2 -allele is 10%, that the population is in Hardy-Weinberg equilibrium, and that mating is random.

(a) What is the homozygous effect, a , and the dominance coefficient, k ? (1p)

(b) Calculate the additive genetic variance? (1p)

(c) What is the (total) genetic variance of the trait values? (1p)

(d) What is the (broad-sense) heritability of the trait if the variance of the environmental deviation is assumed equal to 1? (1p)

14. For a pair of cousins:

(a) calculate the kinship coefficient, Θ ; (1p)

(b) calculate the coefficient of fraternity, Δ ; (1p)

(c) Assuming a single-locus trait with additive variance equal to 1, determine the genetic covariance. (1p)

15. In a sample of 17 family trios (each consisting of both parents and one affected child), 16 parents were found to be heterozygous at a biallelic marker locus with alleles M_1 and M_2 . In addition, 13 parents were M_1M_1 -homozygous while the remaining 5 parents were M_2M_2 -homozygous.

(a) What is the maximal possible value of the TDT test statistics for these data? (1p)

(b) What is the expected proportion of M_1 -alleles in the 17 affected children if the null hypothesis of no association between marker and disease locus is true? (1p)

(c) If precisely half of the marker alleles in the affected children are M_1 -alleles, what value of the TDT-statistic is actually observed in data? (1p)
