Statistical methods in genetic relatedness and pedigree analysis

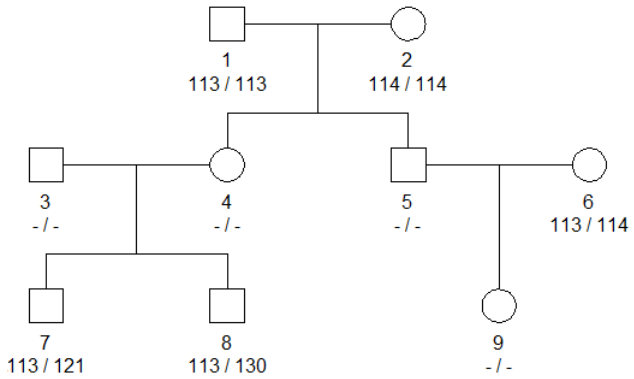
Oslo, January 2020

Magnus Dehli Vigeland and Thore Egeland

Exercise set I. Introduction to pedigrees, genetics and probabilities

Exercise I-1

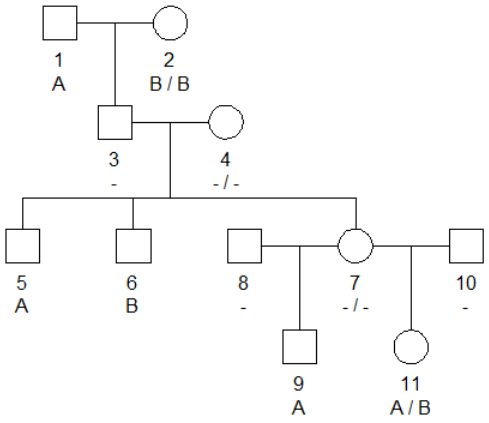
In the pedigree below some members have been typed with a single DNA marker.



a) What kind of marker is this: SNP or STR? Autosomal or X-linked? How do you know?
b) How many alleles are observed in the family. What do the allele labels (e.g. "113") mean?
c) What are the genotypes of individuals 4 and 5?
d) Can you determine the genotype of individual 3?
e) What are the possible genotypes for individual 9, and how likely is each of them?

Exercise I-2

In the pedigree below some members have been typed with a X-linked SNP marker.



a) What is the genotype of individual 3?
b) What is the genotype of individual 4?
c) What is the genotype of individual 7? Which allele did she inherit from her mother?
d) Can you determine the genotype of individual 8? What about individual 10?
e) What is the relationship between 9 and 11? Between 5 and 11? 8 and 10?

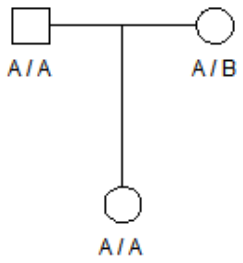
Exercise I-3

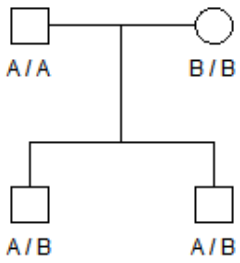
Draw pedigrees showing the following relationships.

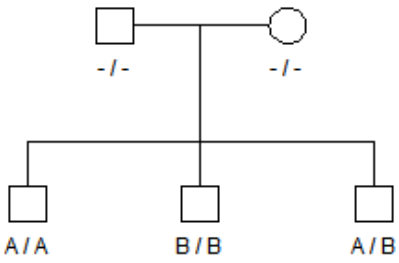
a) First cousins twice removed.
b) Half second cousins once removed.
c) Double second cousins.
d) Simultaneous half siblings and half first cousins.

Exercise I-4

Compute the likelihood of each pedigree. The genotypes are for an autosomal SNP in HWE, with allele frequencies $p_A = 0.9$ and $p_B = 0.1$.

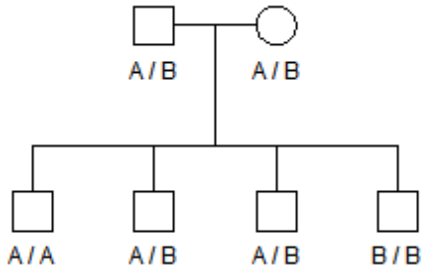
a)


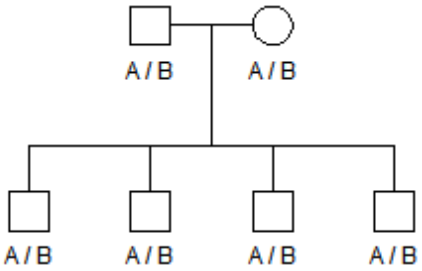
b)


c)


Exercise I-5

The two families below have been genotyped with the same SNP.





a) Which of the pedigrees has the highest likelihood? (Bonus points if you can answer without actually computing the likelihoods.)

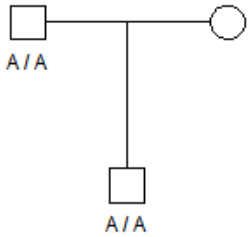
b) A student argues as follows:

The first pedigree is more likely, since that's where the children's genotypes are closest to the Hardy-Weinberg proportions $AA : AB : BB = 1 : 2 : 1$.

Explain why this argument is ridiculous.

Exercise I-6

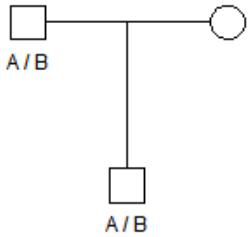
In the family below father and son have been typed with a SNP in Hardy-Weinberg equilibrium, with allele frequencies $P(A) = p$ and $P(B) = q = 1 - p$.



a) Explain that the probability of observing A/A in both father and son is p^3.
b) What is the probability that the mother also has genotype A/A?

Exercise I-7

Father and son have been genotyped with a SNP in Hardy-Weinberg equilibrium, with allele frequencies $P(A) = p$ and $P(B) = q = 1 - p$. Both of them are heterozygous.



a) Show by likelihood computations that the genotype probabilities for the mother are $P(A/A) = p^2$, $P(A/B) = 2pq$ and $P(B/B) = q^2$.
b) Why is this result somewhat surprising?

Exercise I-8

In each question below we consider a SNP marker in Hardy-Weinberg equilibrium, with alleles A and B.

a) In a sample of 100 individuals you observe 1 with genotype B/B. What is your best guess for the number of individuals with A/A and A/B?
b) In a certain population A/A occurs 4 times more often than B/B. What are the allele frequencies, and what are the genotype frequencies?
c) Is it possible to have equal frequencies for the three genotypes A/A, A/B and B/B?