Anton Amirov

1. Scientists are trying to map a rare autosomal recessive trait, in a population in which the trait allele frequency is 0.01. The trait does not impact survival or fertility of individuals. In the population, 5% of marriages are between first cousins (g = 1/16) and 20% of marriages are between second cousins (g = 1/64). The remainder are between unrelated individuals.

(a) What is the mean inbreeding coefficient in the population? What is the frequency of affected individuals?

Mean inbreeding coefficient

$$\frac{0.05}{16} + \frac{0.2}{64} = 0.00625$$

Probability to be affected is

$$0.00625 * 0.01 + (1 - 0.00625) * 0.01^2 = 0.000161875$$

(b) Scientists collect their individuals through a registry of affected individuals maintained by a Support Association (SA). Assuming each affected individual has the same probability of being registered with SA, independently of his/her relatives, what proportion of the individuals of SA's list are the offspring of first cousins, and of second cousins?
(Ans: 0.222 and 0.314)

Probability to be affected if someone is offspring of first cousins marriage is

$$\frac{0.01}{16} + \frac{15 * 0.01^2}{16} = 0.00071875$$

Probability to be affected and to be the offspring of first cousins marriage is $0.00071875 *$ 0.05 and probability to be offspring of first cousins marriage if affected is

$$\frac{0.00071875 * 0.05}{0.000161875} \approx 0.222$$

Probability to be affected if someone is offspring of first cousins marriage is

$$\frac{0.01}{64} + \frac{63 * 0.01^2}{64} = 0.0002546875$$

Probability to be to be offspring of second cousins marriage if affected is

$$\frac{0.0002546875 * 0.2}{0.000161875} \approx 0.314$$

(c) What is the probability that an affected offspring of a first cousin marriage is ibd at the trait locus? And for the offspring of a second-cousin marriage?
(Ans: 0.869 and 0.613)

For first cousins that probability will be

$$\frac{\frac{0.01}{16}}{\frac{0.01}{16} + \frac{15 * 0.01^2}{16}} = \frac{1}{1 + 15 * 0.01} \approx 0.869$$

For second cousins

$$\frac{\frac{0.01}{64}}{\frac{0.01}{64} + \frac{63 * 0.01^2}{64}} = \frac{1}{1 + 63 * 0.01} \approx 0.613$$

2. Continuing from Qu.1, through SA the scientist enroll 40 unrelated affected individuals: 20 of them are the offspring of first-cousin marriages and 20 are the offspring of a second-cousin marriages. By good luck, they have a marker so tightly linked to the trait locus that the recombination frequency r is essentially 0. At this marker there are four alleles, each having frequency 0.25.

(a) What is the probability that an affected offspring of a first-cousin marriage is homozygous at the marker locus? And for the offspring of a second-cousin marriage?
(Ans: 0.902 and 0.710).

If person is ibd at trait locus, that person is always homozygous at marker locus. Otherwise that probability is 0.25 – unconditional probability to be homozygous at marker locus. So using results from 1c, for offspring of first cousin marriage probability to be homozygous at marker locus is

$$0.869 + (1 - 0.869) * 0.25 \approx 0.902$$

and for offspring of second cousin marriage

$$0.613 + (1 - 0.613) * 0.25 \approx 0.710$$

(b) What is the probability that all 40 affected individuals are homozygous at this marker locus?

$$0.902^{20} 0.710^{20} \approx 0.0001347$$

(c) Assuming r=0, show that for the offspring of first cousins, each one homozygous at the marker contributes 0.48 to the base-10 lod score, and each one that is heterozygous contributes -0.86. For the offspring of second cousins, show that each one that is homozygous at the marker contributes 0.43, and each one that is heterozygous contributes -0.41. Hence find the overall Elod (expected lod score) from this sample.

Assuming $\rho = 1/2$ (no linkage) probability for the offspring of the first marriage to be homozygous at the marker locus is

$$\frac{1}{16} + 0.25 * \frac{15}{16} = 0.296875$$

So the contribution of homogenous offspring of first cousin marriage to the base-10 lod score is

$$\log \frac{0.902}{0.296875} \approx 0.48$$

Contribution of heterozygous offspring is

$$\log \frac{1 - 0.902}{1 - 0.296875} \approx -0.86$$

For offspring of second cousins probability to be homozygous at marker locus is

$$\frac{1}{64} + 0.25 * \frac{63}{64} = 0.26171875$$

So the contribution of homogenous offspring of second cousin marriage to the base-10 lod score is

$$\log \frac{0.710}{0.26171875} \approx 0.43$$

Contribution of heterozygous offspring is

$$\log \frac{1 - 0.710}{1 - 0.26171875} \approx -0.41$$

Elod for this sample is

$$20 * \left(0.48 * 0.902 - 0.86 * (1 - 0.902)\right) + 20 * \left(0.43 * 0.710 - 0.41 * (1 - 0.710)\right)$$
$$\approx 10.70$$

3. In this question, we assume Haldane's model of no interference. You may assume the corresponding Haldane map function relating genetic distance to recombination rate.

(a) Show that the recombination frequency between two loci is an increasing function of genetic distance, and that when the loci are tightly linked, the recombination probability is approximately the genetic distance (in Morgans).

Derivative of recombination function by distance

$$\dot\rho = \left(\frac{1 - e^{-2d}}{2}\right)' = e^{-2d} > 0$$

so the function is increasing. If d is sufficiently small (loci are tightly linked)

$$\rho(d) \approx \rho'(0)d = e^{-2*0}d = d$$

(b) (Lange, Chapter 7: number 1) Suppose loci are ordered, 1,2,3,...,L along a chromosome, and r_j is the recombination frequency between locus j and locus j+1. Show that the recombination frequency r between locus 1 and locus L satisfies the formula (according to Lange, due to Trow)
(1- 2 r) is the product of the (1- 2 r_j)

$$\prod 1 - 2r_j = \prod 1 - 2\frac{1 - e^{-2d_j}}{2} = \prod e^{-2d_j} = e^{-2\sum d_j} = e^{-2d} = 1 - 2\frac{1 - e^{-2d}}{2} = 1 - 2r$$

4. (Mather's formula, Mather (1938))
Note each chiasma in meiosis involves two of the four chromatids, and hence has probability 1/2 of being in a resulting gamete. We say there is no chromatid interference if each chiasma has probability 1/2 of being in a resulting gamete, INDEPENDENTLY of the others.
(a) If a fair coin is tossed 1,2 or 3 times, what is the probability of an odd number of heads?

In all cases probability is ½:

      For 1 toss - 1 (1) of 2 possible

      For 2 tosses – 2 (01 and 10) or 4 possible

      For 3 tosses – 4 (001, 010, 100 and 111) of 8 possible

(b) Show by induction (or elsehow if you prefer), that if a fair coin is tossed any number of N of times (N>0), the probability of an odd number of heads is 1/2.

Total number of heads is odd in 2 cases

Last toss is tail (1/2) and previous tosses resulted in odd number of heads (P_{N-1})

Last toss is heads (1/2) and previous tosses resulted in even number of heads (1- P_{N-1})

So

$$P_N = \frac{P_{N-1}}{2} + \frac{1 - P_{N-1}}{2} = \frac{1}{2}$$

(c) Deduce Mather's formula, that the recombination probability is (1/2) P(N>0), where N is the random number of chiasmata between the two loci in a meiosis.

If N=0, no recombination happens, otherwise recombination will occur if among N chiasmata there is an odd number of chiasmata involving particular gamete. Since the probability that chiasmata involves the particular gamete is ½, the probability that there is N number of those is equivalent to question about the odd number of tails in N tosses and the probability of the odd number is ½ if N>0. Therefore, recombination probability is ½ * P(N>0).

(d) Show that this means that, in a very small segment of chromosome where P(N > 1) is negligible, the recombination probability is approximately the genetic distance (in Morgans).

Distance in Morgans d is the expected amount of crossovers per section of gamete, so that amount in both gametes together is (N) is 2d. Expected amount can be expressed as

$$E(N) = 2d = \sum iP(N = i)$$

If P(N>1) are negligible, we have

$$P(N = 1) \approx 2d$$

And recombination probability

$$\frac{P(N > 0)}{2} \approx \frac{P(N = 1)}{2} \approx \frac{2d}{2} = d$$