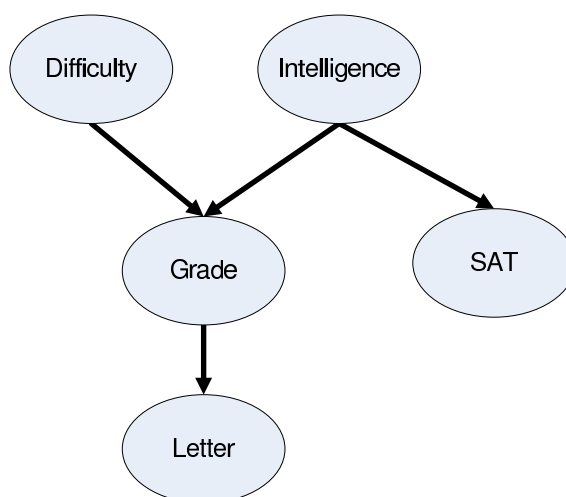


02-711: Computational Molecular Biology and Genomics Practice Midterm

February 27, 2009

Below are a set of questions that will help you prepare for the midterm. Note that the practice questions may have varied difficulties compared to the questions here. In addition, the actually midterm will have larger set of questions than the ones provided here.

1 Mathematical Background



Suppose we have the above Bayesian network where Difficult, Intelligence, Grade, SAT, and Letter has the distinct states (Easy/Hard), (Smart/Very Smart), (A/B/C), (High/Low), and (Good/Ok) respectively. Derive the MLE of the parameters for the problem.

Solution: See lecture note 2-4 and recitation slides to see how to solve it in the simple case.

2 Partial Digest Problem

You are given a partial digest (in a sorted order) of an unknown size fragment as shown below. Assume that the partial digest was impeccable where all possible combinations of fragments were obtained including the full length fragment, draw the number of restriction sites and the length between each site on the line provided below.

Fragment Length: {2, 2, 4, 5, 6, 7, 7, 9, 11, 13, 13, 14, 15, 18, 20}

Solution: 2-4-7-2-5

3 Linkage Disequilibrium

Consider a gene A with allele A_1 and A_2 at frequencies x_1 and x_2 , and a different gene B in the same population with alleles B_1 , B_2 , and B_3 at frequencies y_1 , y_2 , and y_3 . What are the expected frequencies of gametes with linkage equilibrium assuming that $x_1 = 0.3$, $y_1 = 0.2$, and $y_2 = 0.3$?

Solution: $p_{A_1B_1} = x_1y_1, p_{A_1B_2} = x_1y_2, p_{A_1B_3} = x_1(1 - y_1 - y_2), p_{A_2B_1} = (1 - x_1)y_1, p_{A_2B_2} = (1 - x_1)y_2, p_{A_2B_3} = (1 - x_1)(1 - y_1 - y_2)$

4 Hardy-Weinberg Equilibrium

The accompanying table shows the observed numbers of AA , Aa , and aa genotypes in samples of size 100 from each of four populations. Calculate the χ^2 value for goodness of fit to Hardy-Weinberg proportions. Which samples will most likely to reject the hypothesis of Hardy-Weinberg proportions?

| Population | AA | Aa | aa |
|------------|----|----|----|
| (a) | 8 | 53 | 39 |
| (b) | 9 | 61 | 30 |
| (c) | 13 | 58 | 29 |
| (d) | 18 | 35 | 47 |

Solution: (b)

5 Coalescent Model

In a population of effective size 30, how many generations are required on the average to coalesce from 4 alleles to 3? From 3 alleles to 2? From 2 alleles to 1?

Solution: $4 \rightarrow 3 : E(T_4) = \frac{2N}{6}, 3 \rightarrow 2 : E(T_3) = \frac{2N}{3}, 2 \rightarrow 1 : E(T_2) = 2N$