ETH zürich   D BSSE   Computational Biology Group

Evolutionary Dynamics

Exercises 2

Prof. Dr. Niko Beerenwinkel
Monica Dragan
Kevin Rupp
Lara Fuhrmann

6th October 2022

Exercises marked with a "💻" are programming exercises. These can be solved in a programming language of your choice. Please make sure to hand in your code along with your answers to these exercises.

Problem 1: Sequence space and Hamming distance

Consider an alphabet $\mathcal{A}$ of size $|\mathcal{A}| = B$. For a binary alphabet, one has $\mathcal{A} = \{0,1\}$ and $B = 2$, and for DNA, one has $\mathcal{A} = \{A,T,C,G\}$ and $B = 4$. We are studying sequences $S \in \mathcal{A}^L$ of length L. Assume sequences are random with a uniform distribution,

(a) How many unique binary and DNA sequences exists for $L = 28$? **(1 point)**

(b) What is the average Hamming distance between two random binary sequences? What is the expected Hamming distance for two random DNA sequences? **(1 point)**

(c) Given a binary sequence of length L, how many sequences exist at a Hamming distance three from it? How many at distance K with $K \leq L$? Repeat the calculation for DNA sequences. **(2 points)**

Problem 2: Quasispecies

Consider the quasispecies equation with two genotypes $0,1$ (i.e., binary sequences of length 1). Let the fitness of genotype 0 be $f_0 > 1$, and the fitness of genotype 1 be $f_1 = 1$. Moreover, genotypes are replicated error-free with probability q,

(a) Write down the mutation-selection matrix W and find its eigenvalues. **(2 points)**

(b) To which eigenvalue corresponds the non-trivial equilibrium point? **(1 point)**
Hint: Perron-Frobenius theorem.

(c) Examine the dynamics of the quasispecies equation and confirm the results obtained in (b). Assume that $q = 0.6$ and $f_0 = 1.5$, and initial condition $(0.65, 0.35)$. 💻 **(1 point)**

(d) What is the equilibrium point for $f_0 = f_1 = 1$? **(1 point)**

(e) Calculate the equilibrium point in the limit of low mutation rate ($q \approx 1$). **(1 point)**