ETH zürich D BSSE Computational Biology Group

Evolutionary Dynamics

Exercises 5

Prof. Dr. Niko Beerenwinkel
Dr. Katharina Jahn
Dr. Rob Noble
25th October 2018

Problem 1: Pathways of carcinogenesis

Consider three independent mutations $\{1,2,3\}$. Each mutation occurs after an exponentially distributed waiting time $T_i \sim \exp(\lambda_i)$, $i = 1,2,3$.

(a) What is the probability for the path $P = 2 \to 1 \to 3$? **(1 point)**

(b) Assume cancer arises if any two of the three genes are mutated. How many possible genotypes are there? How many pathways? Compute the expected waiting time until any two out of three genes are mutated. **(1 point)**

(c) Now consider d independent mutations. How many paths exist leading to the genotype $\{1,\ldots,d\}$ with all mutations present? If cancer already arises after any k mutations, how many different paths are there? **(1 point)**

Problem 2: Neutral Wright-Fisher process

Consider the neutral Wright-Fisher process for a system of N cells of two different types $\{A,B\}$. Let $X(t)$ denote the number of A-cells at time t. The process has the transition matrix

$$P_{i,j} = \text{Prob}[X(t) = j \mid X(t-1) = i] = \binom{N}{j} \left(\frac{i}{N}\right)^j \left(\frac{N-i}{N}\right)^{N-j},$$

that is, $X(t) \mid X(t-1) = i$ is binomially distributed with parameter $p = i/N$.

(a) Compute, by a calculation similar to the one for the neutral Moran process (Exercises 3), the conditional expectation $\mathrm{E}[X(t) \mid X(0) = i]$. **(1 point)**

Hint: You may use the results for moments of the binomial distribution.

(b) Compute the conditional variance $\mathrm{Var}[X(t) \mid X(0) = i]$. **(1 point)**

Hint: First show that

$$\mathrm{Var}[X(t) \mid X(0) = i] = V_1 + (1 - 1/N)\,\mathrm{Var}[X(t-1) \mid X(0) = i]$$

where $V_1 = \mathrm{Var}[X(1) \mid X(0) = i]$.

(c) Show that in the Wright-Fisher process, the heterozygosity H_t at time t satisfies **(1 point)**

$$\mathrm{E}[H_t \mid X_0 = i] = H_0(i)\left(1 - \frac{1}{N}\right)^t$$

and hence decreases exponentially at rate $1/N$. Compare this behaviour with the Moran model.

Note: Heterozygosity in this context is defined as the probability that two individuals chosen at random from the population are of different types.

(d) Derive an approximation for $\mathrm{Var}[X(t) \mid X(0) = i]$ for large population size N. Why does the variance increase faster for the Wright-Fisher process than for the Moran process? **(1 point)**

Problem 3: Wave approximation

Consider the wave approximation of the Wright-Fisher model for cancer progression. Here, the growth of a clone with j mutations is given by

$$\dot{x}_j = s x_j (j - \langle j \rangle).$$

For small times, the average fitness $s\langle j \rangle = s \sum_j j x_j$ can be considered constant. Use this throughout your calculations.

(a) Find the analytic solution for the initial condition $x_j(0) = 1/N$. **(1 point)**

(b) The rate at which an additional mutation occurs is given by $u d x_j(t)$. Find the time τ when the cumulative probability exceeds $1/N$. **(1 point)**

(c) Compute the waiting time until the next mutation for a mutation rate $u = 10^{-7}$/cell generation, $d = 100$ genes and a fitness advantage of $s = 2\%$ per mutation. Use that $j - \langle j \rangle \approx \sqrt{\log N}$ with $N = 10^7$ cells and assume a cell generation time of 1 day. **(1 point)**