Prof. Dr. Niko Beerenwinkel
Johannes Gawron
Norio Zimmermann
Michael Schneider

Evolutionary Dynamics

Exercises 3

10th October 2024

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: Random walk (tutorial exercise)
Consider a symmetric, time-discrete random walk in one dimension, $\{X(t) \mid X(t) \in \mathbb{Z}, t \in \mathbb{N}_0\}$. Let $a/2$ denote the probability of jumping forward or backward, respectively. Hence we have transition probabilities $P_{i,i+1} = P_{i,i-1} = a/2$, $P_{i,i} = 1 - a$ and otherwise $P_{i,j} = 0$.

Hint: Express $X(t)$ as the sum of increments $\Delta(t) \in \{-1, 0, +1\}$, i.e. $X(t) = X(0) + \Delta(1) + \cdots + \Delta(t)$ and use that the $\Delta(t)$ are iid (identically independently distributed).

(a) Argue why $\mathrm{E}[\Delta(t)] = 0$.

(b) Argue why $\mathrm{E}[X(t)] = x_0$, where $x_0 = X(0)$.

(c) Calculate the variance $\mathrm{Var}[\Delta(t)]$.

(d) Show that the variance of $X(t)$ equals $\mathrm{Var}[X(t)] = at$.

Problem 2: Neutral Moran process
Consider the neutral Moran process $\{X(t) \mid t = 0, 1, 2, \ldots\}$ with two alleles A and B, where $X(t)$ is the number of A alleles in generation t.

(a) Show that the process has a stationary mean: **(1 point)**

$$E[X(t) \mid X(0) = i] = i.$$

Hint: First calculate $E[X(t) \mid X(t-1)]$ and use the *law of total expectation*, $E_Y[Y] = E_Z[E_Y[Y \mid Z]]$ with $Y = X(t)$ and $Z = X(t-1)$.

(b) Show that the variance of $X(t)$ is given by: **(2 points)**

$$\mathrm{Var}[X(t) \mid X(0) = i] = V_1 \frac{1 - (1 - 2/N^2)^t}{2/N^2}. \tag{1}$$

Consider the following steps:

(i) Show that $V_1 := \mathrm{Var}[X(1) \mid X(0) = i] = 2i/N(1 - i/N)$.

(ii) Then use that $\forall t > 0$ $\mathrm{Var}[X(t) \mid X(t-1) = i] = \mathrm{Var}[X(1) \mid X(0) = i]$ (why?) and the *law of total variance*, $\mathrm{Var}[Y] = E_Z[\mathrm{Var}_Y[Y \mid Z]] + \mathrm{Var}_Z[E_Y[Y \mid Z]]$, to derive

$$\mathrm{Var}[X(t) \mid X(0) = i] = V_1 + (1 - 2/N^2)\,\mathrm{Var}[X(t-1) \mid X(0) = i] \tag{2}$$

(iii) The inhomogeneous recurrence equation above can be solved by bringing it into the form $x_t - a = b(x_{t-1} - a)$, from which it follows that $x_t - a = b^{t-1}(x_1 - a)$.

(c) Write a small simulation to check the results from (a) and (b). Use $N \in \{10, 100\}$ and $i = N/2$. Simulate 1000 trajectories for $t = 1, \dots, 100$, and compute empirical mean and variance in comparison to analytical mean and variance. Comment on your results shortly. 💻 **(2 points)**

Problem 3: Absorption in a birth-death process

Consider a birth-death process with state space $\{0, 1, \dots, N\}$, transition probabilities $P_{i,i+1} = \alpha_i$, $P_{i,i-1} = \beta_i > 0$, and absorbing states 0 and N.

(a) Show that the probability of ending up in state N when starting in state i is **(2 points)**

$$x_i = \frac{1 + \sum_{j=1}^{i-1} \prod_{k=1}^{j} \gamma_k}{1 + \sum_{j=1}^{N-1} \prod_{k=1}^{j} \gamma_k} \tag{3}$$

Consider the following steps:

(i) The vector $x = (x_0, \dots x_N)^T$ solves $x = Px$ where P is the transition matrix. (Why?) Set $y_i = x_i - x_{i-1}$ and $\gamma_i = \beta_i / \alpha_i$. Show that $y_{i+1} = \gamma_i y_i$.

(ii) Show that $\sum_{i=1}^{\ell} y_i = x_\ell$.

(iii) Show that $x_\ell = \left(1 + \sum_{j=1}^{\ell-1} \prod_{k=1}^{j} \gamma_k\right) x_1$.

(b) Using (3), show that for the Moran process *with selection* **(1 point)**

$$\rho = x_1 = \frac{1 - 1/r}{1 - 1/r^N},$$

where r is the relative fitness advantage. Use *l'Hôpital's rule* to calculate the limit $r \to 1$.

Problem 4: Poisson Process

Consider a Poisson process to model mutation events. We start with a population, where all individuals initially carry allele a, and new mutations result in allele b with a mutation rate u. **(2 points)**

(a) Show that the time until the first mutation appears, T_1, follows an exponential distribution $T_1 \sim \mathrm{Exp}(Nu)$.

(b) Given neutral evolution, show that the rate of evolution R, from an all-a population to an all-b population is $R = u$.

(c) Explain the assumptions and implications of this result for inferring the divergence time between species.