1. Stochastic protein dynamics. Consider a stochastic system in which particular protein in a cell has copy number n. Starting in state n, the transition rate for $n \to n+1$ is β and the transition rate for $n \to n-1$ is αn. The steady-state probability that the system is in state n is P_n.

   (a) Use detailed balance to obtain the steady-state relationship between P_n and P_{n+1}, and then derive a closed form expression for the probability distribution P_n.

   (b) Provide a closed form expression for the generating function $\tilde{P}(\phi) = \sum_{n=0}^{\infty} e^{-\phi n} P_n$; express the mean $\langle n \rangle$ and the variance $\langle n^2 \rangle - \langle n \rangle^2$ in terms of derivatives of $\tilde{P}(\phi)$; and evaluate the mean and variance in terms of model parameters.

   (c) Suppose you want to run a stochastic simulation. Starting in state n, what states can be reached? Define the random variable τ as the waiting time before the first transition. What is the mean time $\langle \tau \rangle$ for this first transition? What is the probability distribution $p(\tau)$ for the first transition?

   (d) Provide an ODE for a continuous variable $X(t)$ where $X(t)$ and $n(t)$ have the same dynamics in the limit of large n, and provide the mean and variance of $X(t)$ from the ODE starting with $X(0)$ equal to the mean.

2. Fluctuation-dissipation. Use your Gillespie simulation code for dynamics of a single protein species with copy number n and transitions

$$\begin{aligned} n &\to n+1 \\ n &\to n-1 \end{aligned}$$

   and corresponding rates β and αn. Use parameter values $\beta = 1/\text{min}$ and $\alpha = 0.1/\text{min}$. The auto-correlation function $S(t)$ is defined as $\langle n(t)n(0) \rangle - \langle n(t) \rangle \langle n(0) \rangle$. If the system is time-independent then $\langle n(t) \rangle = \langle n(0) \rangle = \langle n \rangle$ for any time t, and $S(t) = \langle n(t'+t)n(t') \rangle - \langle n(t'+t) \rangle \langle n(t') \rangle$ for any time t'.

   (a) Note that $S(0)$ is just the variance of n. Provide $S(0)$. According to the fluctuation-dissipation theorem, the shape of $S(t)$ should be identical to the shape of the relaxation of an equivalent ODE system for continuous variable $X(t)$, $S(t)/S(0) = [X(t) - \mu]/[X(0) - \mu]$, with μ equal to the steady-state value and $X(0)$ an arbitrary initial state (other than μ). Use the theorem to provide an analytic expression for $S(t)$.

   (b) Now estimate the auto-correlation function from a trajectory. Start a simulation at time 0 with n equal to the equilibrium mean, and run a trajectory for $T = 10{,}000$ min saving the state every 0.1 min. For $0 \le \tau \le 100$ min, calculate $\langle n \rangle$ and $S(t)$ from the trajectory as the discrete-time equivalents of the integrals

$$\begin{aligned} \langle n \rangle &= T^{-1} \int_0^T dt\, n(t) \\ S(t) &= (T-t)^{-1} \int_0^{T-t} dt'\, n(t) n(t'+t) - \langle n \rangle^2. \end{aligned}$$

In other words, define $\Delta t = 0.1$ min, $t_i = i\Delta t$, and $M = T/\Delta t$. Since we are including both endpoints, $i = 0$ and $i = M$, the total number of time points is $M+1$. The mean and auto-correlation are

$$\begin{aligned} \langle n \rangle &= (M+1)^{-1} \sum_{i=0}^{M} n(t_i) \\ S(j\Delta t) &= (M-j+1)^{-1} \sum_{i=0}^{M-j} n(t_i) n(t_{i+j}) - \langle n \rangle^2. \end{aligned}$$

In the discrete form, $\tau = j\Delta t$, and remember to only calculate results for $\tau \leq 100$ min, or $0 \leq j \leq 100/\Delta t$. Provide a plot comparing $S(t)$ from the simulation with the analytical result from the fluctuation-dissipation theorem.

(c) In the previous question you calculated an analytical result for P_n, the probability that the system has n copies of the protein. From the same trajectory you used for $S(t)$, estimate P_n as the fraction of time steps with particle count n, defined formally as

$$P_n = (M+1)^{-1} \sum_{i=0}^{M} \Theta[n(i\Delta t) = n],$$

where as usual Θ is 1 if its argument is True and 0 if False. Provide a table and a plot with the analytical and simulation results for P_n extending to the largest value for n observed in your simulation.

3. Stability analysis. Consider the Delta-Notch system with different exponents for proteins X and Y:

$$\begin{aligned} \dot{X}(t) &= \frac{2}{1+Y(t)^n} - X(t) \\ \dot{Y}(t) &= \frac{2}{1+X(t)^m} - Y(t). \end{aligned}$$

For convenience, we are working with units scaled so that the decay rate is 1/time, the threshold is 1 molecule. To simplify the problem, we have set the production rate to 2 molecules/time.

(a) Provide plots of the nullclines for $(n = 1, m = 2)$, $(n = 2, m = 2)$, and $(n = 4, m = 2)$.

(b) When n and m are sufficiently large, there are three fixed points: two asymmetric ($X \neq Y$) and one symmetric ($X = Y$). Provide approximate (X, Y) values for the asymmetric fixed points and provide the exact location of the symmetric fixed point.

(c) For general n and m, provide the Jacobian at the symmetric fixed point and evaluate its eigenvalues in terms of model parameters. Provide the condition on the eigenvalues that makes the symmetric fixed point unstable (which permits patterning) and describe what conditions on the Hill coefficients permit patterning.