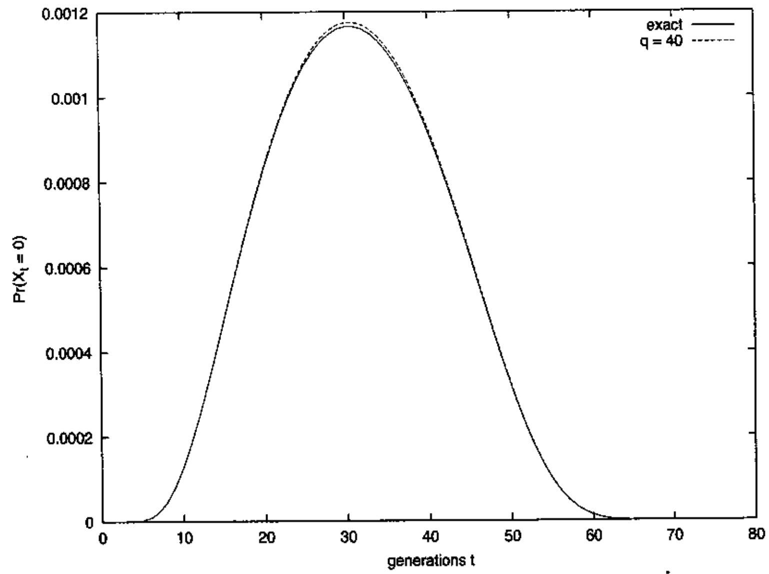

FIGURE 11.2. Extinction Probability of a Recessive Gene

Gene extinction is naturally of great interest. Figure 11.2 depicts the probability that the recessive gene is entirely absent from the population. This focuses our attention squarely on the discrete domain where we would expect the diffusion approximation to deteriorate. The solid curve of the graph shows the outcome of computing directly with the exact Wright-Fisher chain. At about generation 60, the matrix times vector multiplications implicit in the Markov chain updates start to slow the computations drastically. In this example, it took 14 minutes of computing time on a desktop PC to reach 80 generations. The hybrid algorithm with $q = 40$ intervals covering the discrete region and 500 intervals covering the continuous region takes only 11 seconds to reach generation 80. The resulting dashed curve is quite close to the solid curve in Figure 11.2, and setting $q = 50$ makes it practically identical.

11.9 Problems

1. Consider a diffusion process X_t with infinitesimal mean $\mu(t,x)$ and infinitesimal variance $\sigma^2(t,x)$. If the function $f(t)$ is strictly increasing and continuously differentiable, then argue that $Y_t = X_{f(t)}$ is a diffusion process with infinitesimal mean and variance

$$\begin{aligned} \mu_Y(t,y) &= \mu[f(t),y]f'(t) \\ \sigma_Y^2(t,y) &= \sigma^2[f(t),y]f'(t). \end{aligned}$$

Apply this result to the situation where Y_t equals y_0 at $t = 0$ and has $\mu_Y(t, y) = 0$ and $\sigma_Y^2(t, y) = \sigma^2(t)$. Show that Y_t is normally distributed with mean and variance

$$\begin{aligned} \mathrm{E}(Y_t) &= y_0 \\ \mathrm{Var}(Y_t) &= \int_0^t \sigma^2(s)\, ds. \end{aligned}$$

(Hint: Let X_t be standard Brownian motion.)

2. Show that

$$\mathrm{Cov}(Y_{t+s}, Y_t) = \frac{\sigma^2 e^{-\gamma s}(1 - e^{-2\gamma t})}{2\gamma}$$

in the Ornstein-Uhlenbeck process when s and t are nonnegative.

3. Consider a diffusion process X_t with infinitesimal mean

$$\mu(t, x) = \begin{cases} 1, & x < 0 \\ 0, & x = 0 \\ -1, & x > 0 \end{cases}$$

and infinitesimal variance 1. Find the equilibrium distribution $f(x)$ of X_t.

4. In the diffusion approximation to a branching process with immigration, we set $\mu(t, x) = (\alpha - \delta)x + \nu$ and $\sigma^2(t, x) = (\alpha + \delta)x + \nu$, where α and δ are the birth and death rates per particle and ν is the immigration rate. Demonstrate that

$$\begin{aligned} \mathrm{E}(X_t) &= x_0 e^{\beta t} + \frac{\nu}{\beta}\left[e^{\beta t} - 1\right] \\ \mathrm{Var}(X_t) &= \frac{\gamma x_0 (e^{2\beta t} - e^{\beta t})}{\beta} + \frac{\gamma \nu (e^{2\beta t} - e^{\beta t})}{\beta^2} \\ &\quad - \frac{\gamma \nu (e^{2\beta t} - 1)}{2\beta^2} + \frac{\nu (e^{2\beta t} - 1)}{2\beta} \end{aligned}$$

for $\beta = \alpha - \delta$, $\gamma = \alpha + \delta$, and $X_0 = x_0$. When $\alpha < \delta$, the process eventually reaches equilibrium. Find the limits of $\mathrm{E}(X_t)$ and $\mathrm{Var}(X_t)$.

5. In Problem 4 suppose $\nu = 0$. Verify that the process goes extinct with probability $\min\{1, e^{-2\frac{\alpha-\delta}{\alpha+\delta}x_0}\}$ by using equation (11.19) and sending c to 0 and d to ∞.

6. In Problem 4 suppose $\nu > 0$ and $\alpha < \delta$. Show that Wright's formula leads to the equilibrium distribution

$$f(x) = k\left[(\alpha + \delta)x + \nu\right]^{\frac{4\nu\delta}{(\alpha+\delta)^2} - 1} e^{\frac{2(\alpha-\delta)x}{\alpha+\delta}}$$

for some normalizing constant $k > 0$ and $x > 0$.

7. Show that formula (11.23) holds in R^2.

8. Use Stirling's formula to demonstrate that

$$\frac{\Gamma(2N\eta + \frac{1}{2})}{\sqrt{2N(1-f)}\Gamma(2N\eta)} \approx \sqrt{\frac{\eta}{1-f}}$$

when N is large in the Wright-Fisher model for a recessive disease.

9. Consider the Wright-Fisher model with no selection but with mutation from allele A_1 to allele A_2 at rate η_1 and from A_2 to A_1 at rate η_2. With constant population size N, prove that the frequency of the A_1 allele follows the beta distribution

$$f(x) = \frac{\Gamma[4N(\eta_1 + \eta_2)]}{\Gamma(4N\eta_2)\Gamma(4N\eta_1)} x^{4N\eta_2 - 1}(1-x)^{4N\eta_1 - 1}$$

at equilibrium. (Hint: Substitute $p(x) = x$ in formula (11.13) defining the infinitesimal variance $\sigma^2(t, x)$.)

10. Consider the transformed Brownian motion with infinitesimal mean α and infinitesimal variance σ^2 described in Example 11.3.2. If the process starts at $x \in [c, d]$, then prove that it reaches d before c with probability

$$u(x) = \frac{e^{-\beta x} - e^{-\beta c}}{e^{-\beta d} - e^{-\beta c}} \text{ for } \beta = \frac{2\alpha}{\sigma^2}.$$

Verify that $u(x)$ reduces to $(x-c)/(d-c)$ when $\alpha = 0$. As noted in the text, this simplification holds for any diffusion process with $\mu(x) = 0$.

11. Suppose the transformed Brownian motion with infinitesimal mean α and infinitesimal variance σ^2 described in Example 11.3.2 has $\alpha \geq 0$. If $c = -\infty$ and $d < \infty$, then demonstrate that equation (11.21) has solution

$$w(x) = e^{\gamma(d-x)} \text{ for } \gamma = \frac{\alpha - \sqrt{\alpha^2 + 2\sigma^2\theta}}{\sigma^2}.$$

Simplify $w(x)$ when $\alpha = 0$, and show by differentiation of $w(x)$ with respect to θ that the expected time $\mathrm{E}(T)$ to reach the barrier d is infinite. When $\alpha < 0$, show that

$$\Pr(T < \infty) = e^{\frac{2\alpha}{\sigma^2}(d-x)}.$$

(Hints: The variable γ is a root of a quadratic equation. Why do we discard the other root? In general, $\Pr(T < \infty) = \lim_{\theta \downarrow 0} \mathrm{E}\left(e^{-\theta T}\right)$.)