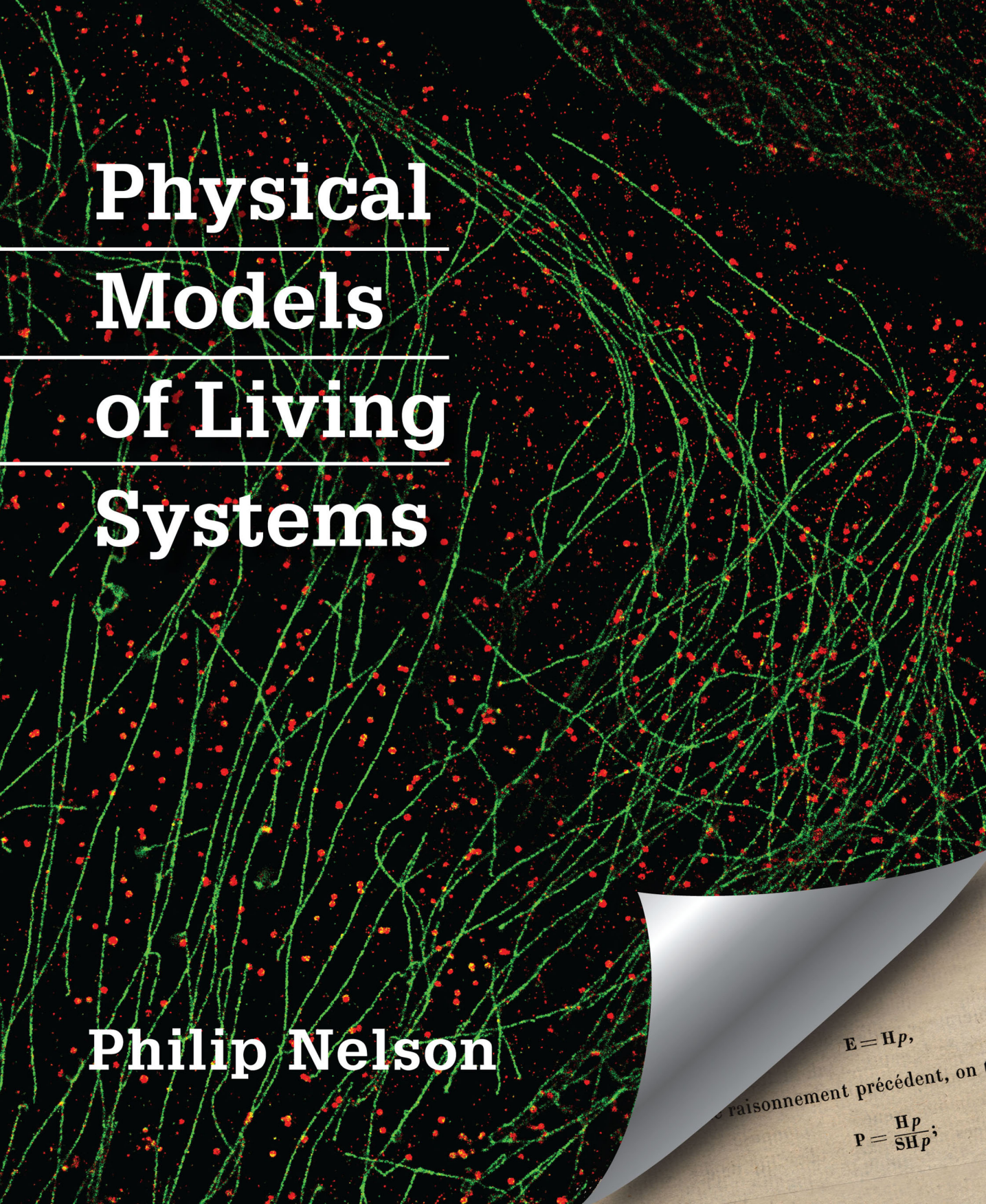
Physical Models of Living Systems

Philip Nelson



7 Poisson Processes

The objective of physics is to establish new relationships between seemingly unrelated, remote phenomena.
—Lev D. Landau

7.1 Signpost

Many key functions in living cells are performed by devices that are themselves individual molecules. These "molecular machines" generally undergo discrete steps, for example, synthesizing or breaking down some other molecules one by one. Because they are so small, they must do their jobs in spite of (or even with the help of) significant randomness from thermal motion. If we wish to understand how they work, then, we must characterize their behavior in probabilistic terms. Even with this insight, however, the challenges are daunting. Imagine an automobile engine far smaller than the wavelength of light: How can we get "under the hood" and learn about the mechanisms of such an engine?
More specifically, we will look at one aspect of this Focus Question:
Biological question: How do you detect an invisible step in a molecular motor cycle?
Physical idea: The waiting-time distributions of individual molecular motors can provide evidence for a physical model of stepping.

7.2 The Kinetics of a Single-Molecule Machine

Some molecular motors have two "feet," which "walk" along a molecular "track." The track is a long chain of protein molecules (such as **actin** or **tubulin**). The feet[1] are subunits of the motor with binding sites that recognize specific, regularly spaced sites on the track. When the energy molecule ATP is present, another binding site on the foot can bind an ATP

[1] For historical reasons, the feet are often called "heads"!

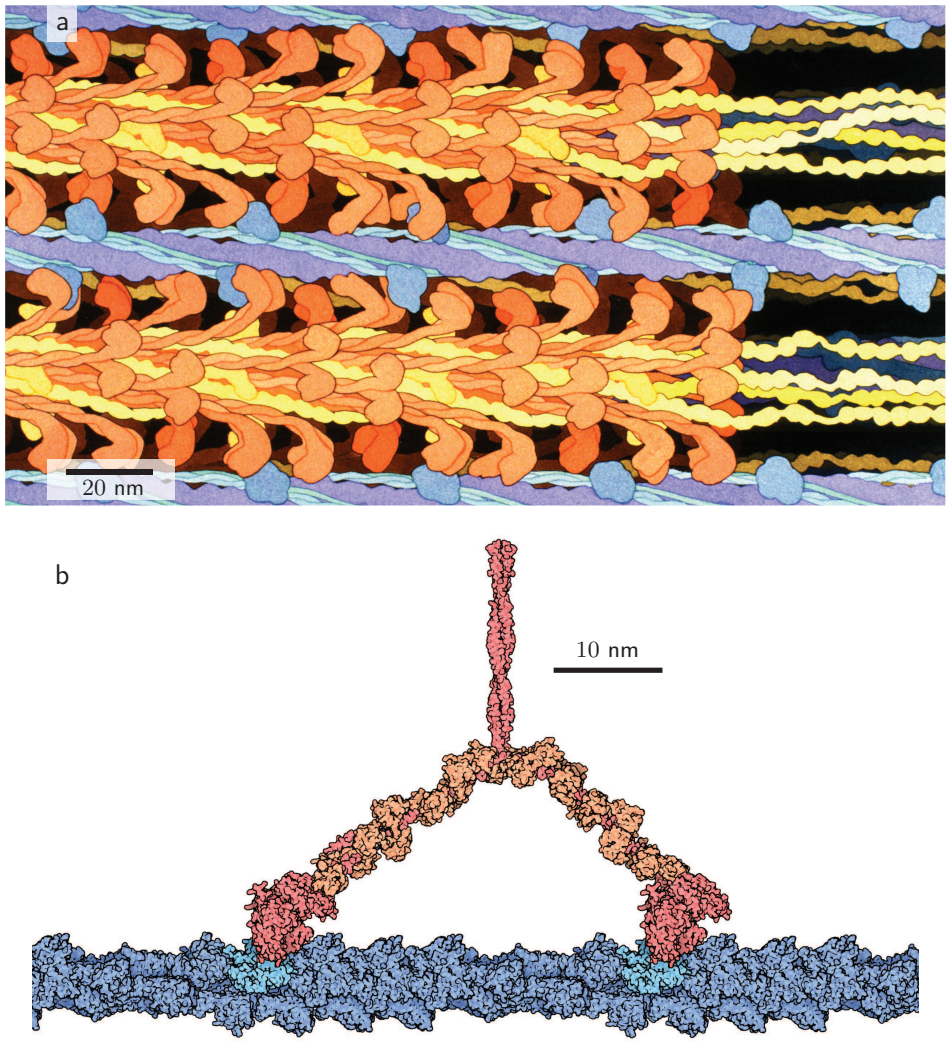


Figure 7.1 [Artist's reconstructions based on structural data.] **Molecular motors.** (a) Skeletal muscle cells contain bundles of the motor protein myosin-II (*orange*). These are interspersed with long filaments composed of the protein actin (*blue*). When activated, the myosin motors in the bundle consume ATP and step along the actin filaments, dragging the red bundle rightward relative to the blue tracks and hence causing the muscle cell to contract. (The thin snaky molecule shown in *yellow* is titin, a structural protein that keeps the actin and myosin filaments in proper alignment.) (b) The myosin-V molecule has two "legs," which join its "feet" to their common "hip," allowing it to span the 36 nm separation between two binding sites (*light blue*) on an actin filament (*blue*). [(a,b) Courtesy David S Goodsell.]

molecule. Clipping off one of the phosphate groups on the ATP yields some chemical bond energy, which is harnessed to unbind the foot from its "track" and move it in the desired direction of motion, where it can, in turn, find another binding site. In this way the motor takes a step, typically of a few nanometers but for certain motors much longer.

Figure 7.1a shows a schematic of the arrangement of many molecular motors, ganged together to exert a significant total force in our skeletal muscles. Other motors operate singly, for example, to transport small cargo from one part of a cell to another. In order to

be useful, such a motor must be able to take many steps without falling off its track; that is, it must be highly **processive**. Myosin-V, a motor in this class, is known to have a structure with two identical feet (Figure 7.1b). It is tempting to guess that myosin-V achieves its processivity (up to 50 consecutive steps in a run) by always remaining bound by one foot while the other one takes a step, just as we walk with one foot always in contact with the ground.[2]

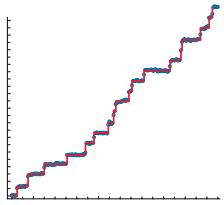


Figure 6.3c (page 134)

Chapter 6 introduced myosin-V and described how Yildiz and coauthors were able to visualize its individual steps via optical imaging. As shown in Figure 6.3c, the motor's position as a function of time looks like a staircase. The figure shows an example with rapid rises of nearly uniform height, corresponding to 74 nm steps. But the *widths* of the stairs in that figure, corresponding to the waiting times (pauses) between steps, are quite nonuniform. Every individual molecule studied showed such variation.

We may wish to measure the speed of a molecular motor, for example, to characterize how it changes if the molecule is modified. Perhaps we wish to study a motor associated with some genetic defect, or an intentionally altered form that we have engineered to test a hypothesis about the function of a particular element. But what does "speed" mean? The motor's progress consists of sudden steps, spread out between widely variable pauses. And yet, the overall trend of the trace in Figure 6.3c does seem to be a straight line of definite slope. We need to make this intuition more precise.

To make progress, imagine the situation from the motor's perspective. Each step requires that the motor bind an ATP molecule. ATPs are available, but they are greatly outnumbered by other molecules, such as water. So the motor's ATP-binding domain is bombarded by molecular collisions at a very high rate, but almost all collisions are not "productive"; that is, they don't lead to a step. Even when an ATP does arrive, it may fail to bind, and instead simply wander away.

The discussion in the previous paragraph suggests a simple physical model: We imagine that collisions occur every Δt, that each one has a tiny probability ξ to be productive, and that every collision is independent of the others. After an unproductive collision, the motor is in the same internal state as before. We also assume that after a productive collision, the internal state resets; the motor has no memory of having just taken a step. Viewed from the outside, however, its position on the track has changed. We'll call this position the system's **state variable**, because it gives all the information relevant for predicting future steps.

T_2 *Section 7.2′ (page 171) gives more details about molecular motors.*

7.3 Random Processes

Before we work out the predictions of the physical model, let's think a bit more about the nature of the problem. We can replicate our experiment, with many identical myosin-V molecules, each in a solution with the same uniform ATP concentration, temperature, and so on. The output of each trial is not a single number, however; instead, it is an entire *time series of steps* (the staircase plot). Each step advances the molecule by about the same distance; thus, to describe any particular trial, we need only state the list of times $\{t_1, t_2, \ldots, t_N\}$ when steps occurred on that trial. That is, each trial is a draw from a probability distribution whose sample space consists of increasing sequences of time values. A random system with this sort of sample space is called a **random process**.

[2] Borrowing a playground metaphor, many authors instead refer to this mechanism as "hand-over-hand stepping."

The pdf on the full sample space is a function of all the many variables t_α. In general, quite a lot of data is needed to estimate such a multidimensional distribution. But the physical model for myosin-V proposed in the preceding section gives rise to a special kind of random process that allows a greatly reduced description: Because the motor is assumed to have no memory of its past, we fully specify the process when we state the collision interval Δt and productive-step probability ξ. The rest of this chapter will investigate random processes with this Markov property.[3]

7.3.1 Geometric distribution revisited

We are considering a physical model of molecular stepping that idealizes each collision as independent of the others, and also supposes them to be simple Bernoulli trials. We (temporarily) imagine time to be a discrete variable that can be described by an integer i (labeling which "time slot"). We can think of our process as reporting a string of step/no-step results for each time slot.[4]

Let E_* denote the event that a step happened at time slot i. Then to characterize the discrete-time stepping process, we can find the probability that, given E_*, the *next* step takes place at a particular time slot $i+j$, for various positive integers j. Call this proposition "event E_j." We seek the conditional probability $\mathcal{P}(\mathsf{E}_j \mid \mathsf{E}_*)$.

More explicitly, E_* is the probability that a step occurred at slot i, *regardless of what happened on other slots.* Thus, many elementary outcomes all contribute to $\mathcal{P}(\mathsf{E}_*)$. To find the conditional probability $\mathcal{P}(\mathsf{E}_j \mid \mathsf{E}_*)$, then, we must evaluate $\mathcal{P}(\mathsf{E}_j \textbf{ and } \mathsf{E}_*)/\mathcal{P}(\mathsf{E}_*)$.[5]

- The denominator of this fraction is just ξ. Even if this seems clear, it is worthwhile to work through the logic, in order to demonstrate how our ideas fit together.

  In an interval of duration T, there are $N = T/\Delta t$ time slots. Each outcome of the random process is a string of N Bernoulli trials (step/no-step in time slot $1, \ldots, N$). E_* is the subset of all possible outcomes for which there was a step at time slot i (see Figures 7.2a–e). Its probability, $\mathcal{P}(\mathsf{E}_*)$, is the sum of the probabilities corresponding to each elementary outcome in E_*.

  Because each time slot is independent of the others, we can factor $\mathcal{P}(\mathsf{E}_*)$ into a product and use the rearrangement trick in Equation 3.14 (page 49). For each time slot prior to i, we don't care what happens, so we sum over both possible outcomes, yielding a factor of $\big(\xi + (1-\xi)\big) = 1$. Time slot i gives a factor of ξ, the probability to take a step. Each time slot following i again contributes a factor of 1. All told, the denominator we seek is

$$\mathcal{P}(\mathsf{E}_*) = \xi. \tag{7.1}$$

- Similarly in the numerator, $\mathcal{P}(\mathsf{E}_j \textbf{ and } \mathsf{E}_*)$ contains a factor of 1 for each time slot prior to i, and a factor of ξ representing the step at i. It also has $j-1$ factors of $(1-\xi)$ representing *no* step for time slots $i+1$ through $i+j-1$, another ξ for the step at time slot $i+j$, and then factors of 1 for later times:

$$\mathcal{P}(\mathsf{E}_j \textbf{ and } \mathsf{E}_*) = \xi(1-\xi)^{j-1}\xi. \tag{7.2}$$

[3] See Section 3.2.1 (page 36).
[4] This situation was introduced in Section 3.4.1.2 (page 47).
[5] See Equation 3.10 (page 45).

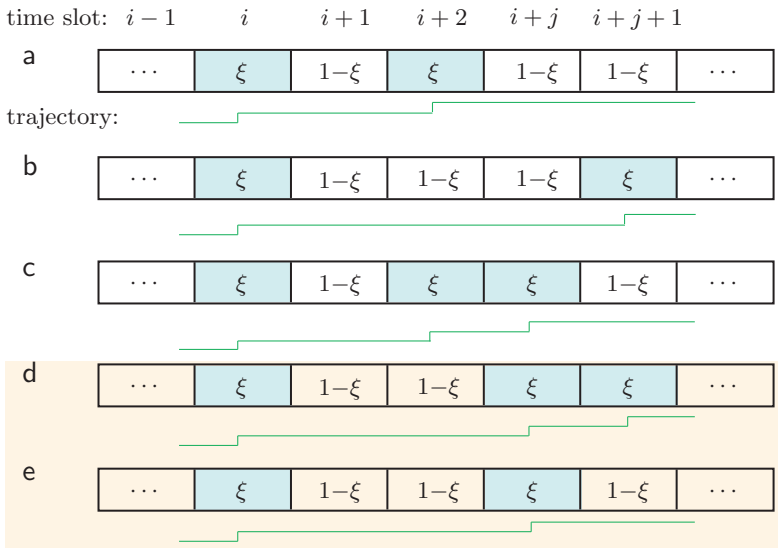


Figure 7.2 [Diagrams.] **Graphical depiction of the origin of the Geometric distribution.** (a–e) Examples of time series, and their contributions to $\mathcal{P}(\mathsf{E}_*)$, the probability that a step occurs at time slot i (Equation 7.1). *Colored boxes* represent time slots in which an event ("blip") occurred. *Green staircases* represent the corresponding motions if the blips are steps of a molecular motor. That is, they are graphs of the state variable (motor position) versus time, analogous to the real data in Figure 6.3c (page 134). (d–e) Examples of contributions to $\mathcal{P}(\mathsf{E}_j \textbf{ and } \mathsf{E}_*)$, the probability that, in addition, the *next* blip occurs at time slot $i+j$, for the case $j = 3$ (Equation 7.2). The terms shown in (d–e) differ only at position $i+j+1$, which is one of the "don't care" positions. Thus, their sum is $\cdots\xi(1-\xi)(1-\xi)\xi(\not{\xi}+1-\not{\xi})\cdots$.

The conditional probability is the quotient of these two quantities. Note that it does not depend on i, because shifting everything in time does not affect how long we must wait for the next step. In fact, $\mathcal{P}(\mathsf{E}_j \mid \mathsf{E}_*)$ is precisely the Geometric distribution (Equation 3.13):

$$\mathcal{P}(\mathsf{E}_j \mid \mathsf{E}_*) = \xi(1-\xi)^{j-1} = \mathcal{P}_{\text{geom}}(j;\xi)\,, \text{for } j = 1, 2, \ldots\,. \tag{3.13}$$

Like the Binomial, Poisson, and Gaussian distributions, this one, too, has its roots in the Bernoulli trial.

7.3.2 A Poisson process can be defined as a continuous-time limit of repeated Bernoulli trials

The Geometric distribution is useful in its own right, because many processes consist of discrete attempts that either "succeed" or "fail." For example, an animal may engage in isolated contests to establish dominance or catch prey; its survival may involve the number of attempts it must make before the next success.

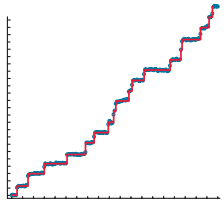


Figure 6.3c (page 134)

But often it's not appropriate to treat time as discrete. For example, as far as motor stepping is concerned, nothing interesting is happening on the time scale Δt. Indeed, the motor molecule represented by the trace in Figure 6.3c generally took a step every few seconds. This time scale is enormously slower than the molecular collision time Δt, because the vast majority of collisions are unproductive. This observation suggests that we may gain a simplification if we consider a *limit*, $\Delta t \to 0$. If such a limit makes sense, then

our formulas will have one fewer parameter (Δt will disappear). We now show that the limit does make sense, and gives rise to a one-parameter family of continuous-time random processes called Poisson processes.[6] Poisson processes arise in many contexts, so from now on we will replace the word "step" by the more generic word "blip," which could refer to a step of a molecular motor or some other sudden event.

The total number of time slots in an interval T is $T/(\Delta t)$, which approaches infinity as Δt gets smaller. If we were to hold ξ fixed, then the total number of blips expected in the interval T, that is, $\xi T/\Delta t$, would become infinite. To get a reasonable limit, then, we must imagine a series of models in which ξ is *also* taken to be small:

> *A* **Poisson process** *is a random process for which (i) the probability of a blip occurring in any small time interval Δt is $\xi = \beta \Delta t$, independent of what is happening in any other interval, and (ii) we take the continuous-time limit $\Delta t \to 0$ holding β fixed.* (7.3)

The constant β is called the **mean rate** (or simply "rate") of the Poisson process; it has dimensions $1/\mathbb{T}$. The separate values of ξ and Δt are irrelevant in the limit; all that matters is the combination β.

Your Turn 7A

Suppose that you examine a random process. Taking $\Delta t = 1\,\mu$s, you conclude that the condition in Idea 7.3 is satisfied with $\beta = 5/$s. But your friend takes $\Delta t = 2\,\mu$s. Will your friend agree that the process is Poisson? Will she agree about the value of β?

It's important to distinguish the Poisson *process* from the Poisson *distribution* discussed in Chapter 4. Each draw from the Poisson distribution is a single integer; each draw from the Poisson process is a sequence of real numbers $\{t_\alpha\}$. However, there is a connection. Sometimes we don't need all the details of arrival times given by a random process; we instead want a more manageable, reduced description.[7] Two very often used reductions of a random process involve its waiting time distribution (Section 7.3.2.1) and its count distribution (Section 7.3.2.2). For a Poisson *process*, we'll find that the second of these reduced descriptions follows a Poisson *distribution*.

7.3.2.1 Continuous waiting times are Exponentially distributed

The interval between successive blips is called the **waiting time** (or "dwell time"), t_w. We can find its distribution by taking the limit of the corresponding discrete-time result (Section 7.3.1 and Figure 7.3).

The pdf of the waiting time is the discrete distribution divided by Δt:[8]

$$\wp(t_w) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \mathcal{P}_{\text{geom}}(j; \xi). \tag{7.4}$$

In this formula, $t_w = (\Delta t)j$ and $\xi = (\Delta t)\beta$, with t_w and β held fixed as $\Delta t \to 0$. To simplify Equation 7.4, note that $1/\xi \gg 1$ because Δt approaches zero. We can exploit that

[6] In some contexts, a signal that follows a Poisson process is also called "shot noise."

[7] For example, often our experimental dataset isn't extensive enough to deduce the full description of a random process, but it does suffice to characterize one or more of its reduced descriptions.

[8] See Equation 5.1 (page 98).

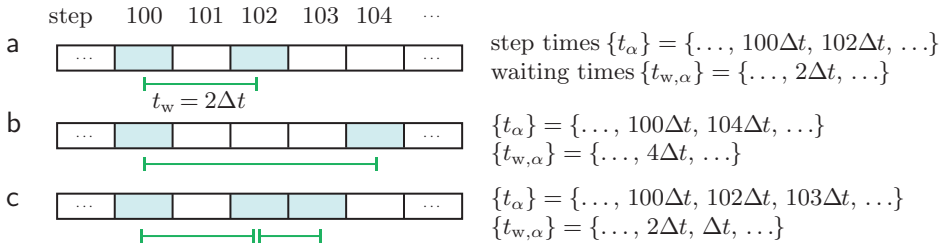


Figure 7.3 [Diagrams.] **Waiting times.** Three of the same time series as in Figure 7.2. This time we imagine the starting time slot to be number 100, and illustrate the absolute blip times t_α as well as the relative (waiting) times $t_{w,\alpha}$.

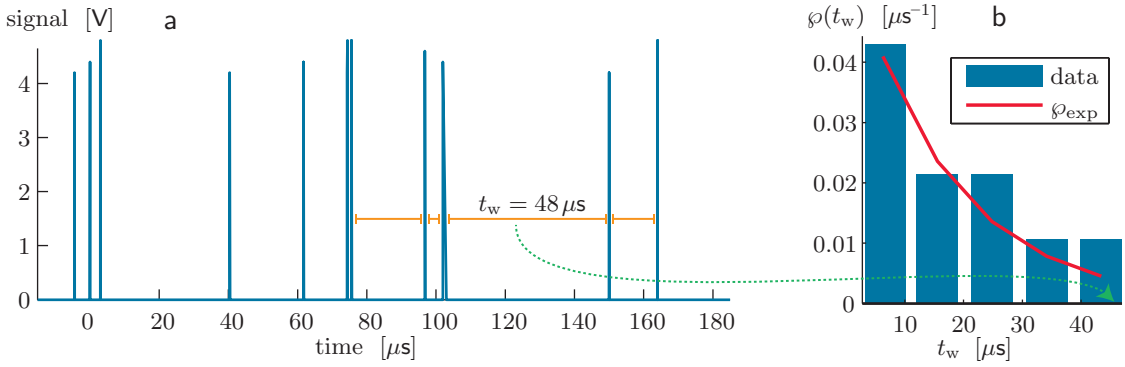


Figure 7.4 [Experimental data with fit.] **The waiting time distribution of a Poisson process (Idea 7.5).** (a) Time series of 11 blips. The *orange arrows* indicate 4 of the 10 waiting times between successive blips. The *green arrow* connects one of these to the corresponding point on the horizontal axis of a graph of $\wp(t_w)$. (b) On this graph, *bars* indicate estimates of the pdf of t_w inferred from the 10 waiting times in (a). The *curve* shows the Exponential distribution with expectation equal to the sample mean of the experimental t_w values. [Data courtesy John F Beausang (Dataset 8).]

fact by rearranging slightly:

$$\wp(t_w) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \xi (1-\xi)^{(t_w/\Delta t)-1} = \lim_{\Delta t \to 0} \frac{\xi}{\Delta t} \left((1-\xi)^{(1/\xi)}\right)^{(t_w \xi/\Delta t)} (1-\xi)^{-1}.$$

Taking each factor in turn:

- $\xi/\Delta t = \beta$.
- The middle factor involves $(1-\xi)^{(1/\xi)}$. The compound interest formula[9] says that this expression approaches e^{-1}. It is raised to the power $t_w\beta$.
- The last factor approaches 1 for small ξ.

With these simplifications, we find a family of continuous pdfs for the interstep waiting time:

> *The waiting times in a Poisson process are distributed according to the* **Exponential distribution** $\wp_{exp}(t_w; \beta) = \beta e^{-\beta t_w}$. (7.5)

Figure 7.4 illustrates this result with a very small dataset.

[9] See Equation 4.5 (page 76).

Example a. Confirm that the distribution in Idea 7.5 is properly normalized (as it must be, because the Geometric distribution has this property).
b. Work out the expectation and variance of this distribution, in terms of its parameter β. Discuss your answers in the light of dimensional analysis.

Solution a. We must compute $\int_0^\infty \mathrm{d}t_\mathrm{w}\, \beta \mathrm{e}^{-\beta t_\mathrm{w}}$, which indeed equals one.
b. The expectation of t_w is $\int_0^\infty \mathrm{d}t_\mathrm{w}\, t_\mathrm{w} \beta \mathrm{e}^{-\beta t_\mathrm{w}}$. Integrating by parts shows $\langle t_\mathrm{w} \rangle = 1/\beta$. A similar derivation[10] gives that $\langle {t_\mathrm{w}}^2 \rangle = 2\beta^{-2}$, so $\mathrm{var}\, t_\mathrm{w} = \langle {t_\mathrm{w}}^2 \rangle - (\langle t_\mathrm{w} \rangle)^2 = \beta^{-2}$. These results make sense dimensionally, because $[\beta] \sim \mathbb{T}^{-1}$.

Figure 7.4 also illustrates a situation that arises frequently: We may have a physical model that predicts that a particular system will generate events ("blips") in a Poisson process, but doesn't predict the mean rate. We do an experiment and observe the blip times $t_1, \ldots, t_N$ in an interval from time 0 to T. Next we wish to make our best estimate for the rate β of the process, for example, to compare two versions of a molecular motor that differ by a mutation. You'll find such an estimate in Problem 7.6 by maximizing a likelihood function.

7.3.2.2 Distribution of counts

A random process generates a complicated, many-dimensional random variable; for example, each draw from a Poisson process yields the entire time series $\{t_1, t_2, \ldots\}$. Section 7.3.2.1 derived a reduced form of this distribution, the ordinary (one-variable) pdf of waiting times. It's useful because it is simple and can be applied to a limited dataset to obtain the best-fit value of the one parameter β characterizing the full process.

We can get another useful reduction of the full distribution by asking, "How many blips will we observe in a fixed, finite time interval T_1?" To approach the question, we again begin with the discrete-time process, regarding the interval T_1 as a succession of $M_1 = T_1/\Delta t$ time slots. The total number of blips, ℓ, equals the sum of M_1 Bernoulli trials, each with probability $\xi = \beta \Delta t$ of success. In the continuous-time limit ($\Delta t \to 0$), the distribution of ℓ values approaches a Poisson distribution,[11] so

For a Poisson process with mean rate β, the probability of getting ℓ blips in any time interval T_1 is $\mathcal{P}_\mathrm{pois}(\ell; \beta T_1)$. (7.6)

Δt does not appear in Idea 7.6 because it cancels from the expression $\mu = M_1 \xi = \beta T_1$. Figure 7.5 illustrates this result with some experimental data.

The quantity ℓ/T_1 is different in each trial, but Idea 7.6 states that its expectation (its value averaged over many observations) is $\langle \ell/T_1 \rangle = \beta$; this fact justifies calling β the mean rate of the Poisson process.

We can also use Idea 7.6 to estimate the mean rate of a Poisson process from experimental data. Thus, if blip data have been given to us in an aggregated form, as counts in each of a series of time bins of duration T_1, then we can maximize likelihood to determine a best-fit value of $T_1\beta$, and from this deduce β.[12]

[10] See the Example on page 55.
[11] See Section 4.3.2 (page 75).
[12] See Problem 6.3.

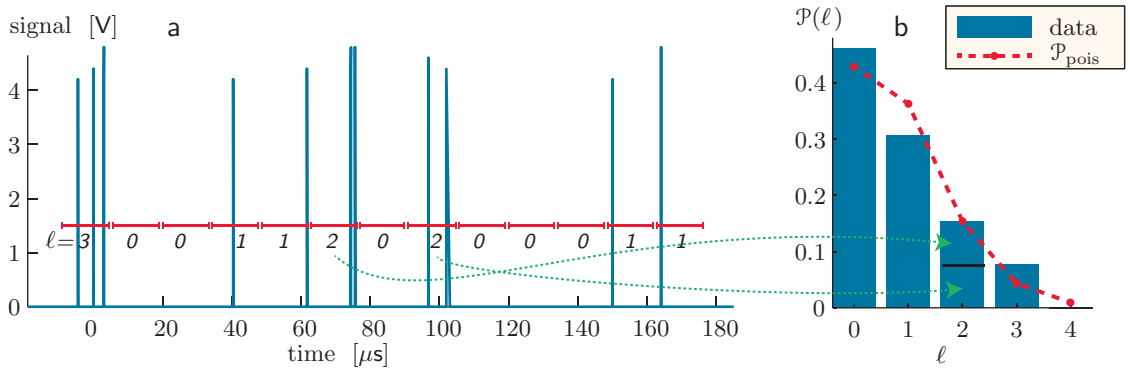


Figure 7.5 [Experimental data with fit.] **The count distribution of a Poisson process over fixed intervals (Idea 7.6).** (a) The same 11 blips shown in Figure 7.4a. The time interval has been divided into equal bins, each of duration $T_1 = 13\,\text{s}$ (*red*); the number of blips in each bin, ℓ, is given beneath its bin indicator. (b) On this graph, *bars* indicate estimates of the probability distribution of ℓ from the data in (a). *Green arrows* connect the instances of $\ell = 2$ with their contributions to the bar representing this outcome. The *red dots* show the Poisson distribution with expectation equal to the sample mean of the observed ℓ values. [Data courtesy John F Beausang (Dataset 8).]

Your Turn 7B

Alternatively, we may consider a single trial, but observe it for a long time T. Show that, in this limit, ℓ/T has expectation β and its relative standard deviation is small.

In the specific context of molecular motors, the fact you just proved explains the observation that staircase plots, like the one in Figure 6.3c, appear to have definite slope in the long run, despite the randomness in waiting times.

Figure 7.6a represents symbolically the two reduced descriptions of the Poisson process derived in this section.

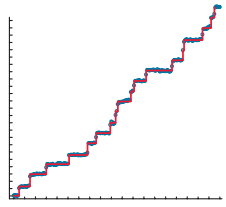


Figure 6.3c (page 134)

7.3.3 Useful Properties of Poisson processes

Two facts about Poisson processes will be useful to us later.

7.3.3.1 Thinning property

Suppose that we have a Poisson process with mean rate β. We now create another random process: For each time series drawn from the first one, we accept or reject each blip based on independent Bernoulli trials with probability ξ_{thin}, reporting only the times of the accepted blips. The **thinning property** states that the new process is also Poisson, but with mean rate reduced from β to $\xi_{\text{thin}}\beta$.

To prove this result, divide time into slots Δt so small that there is negligible probability to get two or more blips in a slot. The first process has probability $\beta\Delta t$ to generate a blip in any slot. The product rule says that in the thinned process, every time slot is again a Bernoulli trial, but with probability of a blip reduced to $(\beta\Delta t)\xi_{\text{thin}}$. Thus, the new process fulfills the condition to be a Poisson process, with mean rate $\xi_{\text{thin}}\beta$ (see Figures 7.6b1,b2).

7.3.3.2 Merging property

Suppose that we have two independent Poisson processes, generating distinct types of blips. For example, Nora may randomly throw blue balls at a wall at mean rate β_1, while Nick

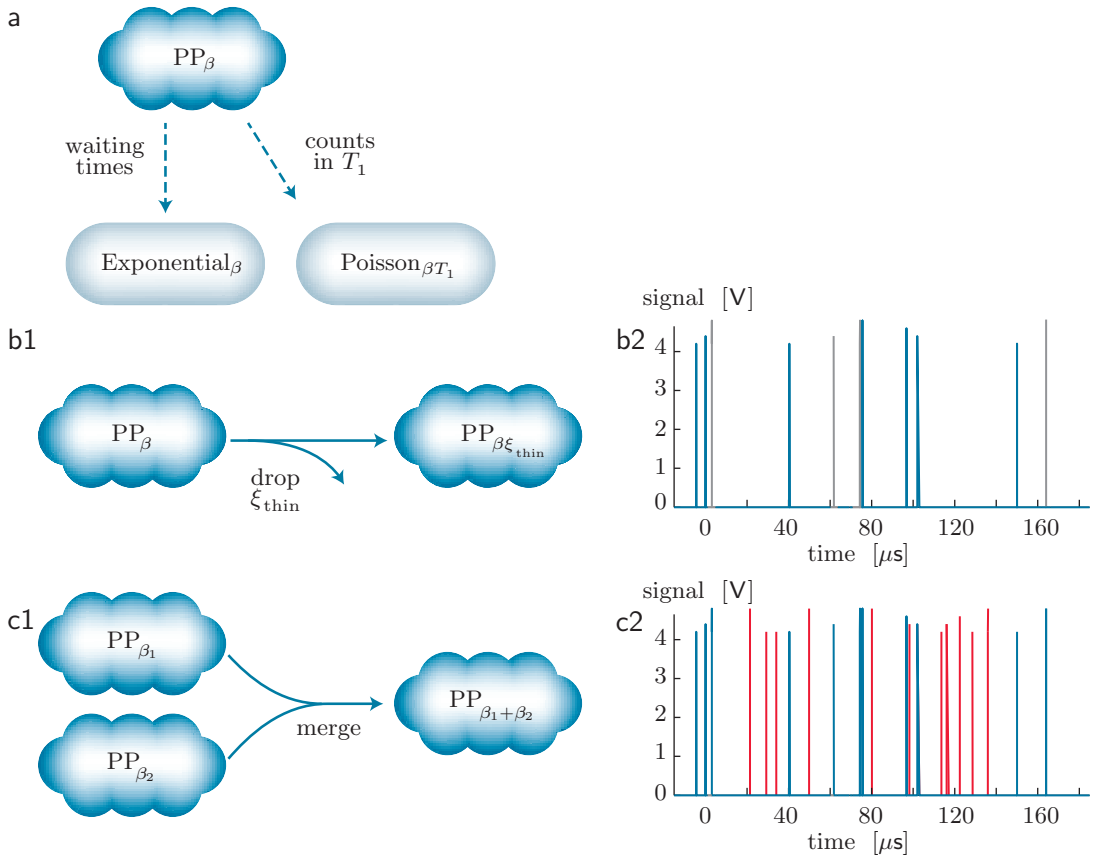


Figure 7.6 [Diagrams; experimental data.] **Some operations involving Poisson processes.** (a) A Poisson process (*upper bubble*) gives rise to two simpler reduced descriptions: Its distribution of waiting times is Exponential, whereas the distribution of blip counts in any fixed interval is Poisson (Figures 7.4, 7.5). (b1,c1) Graphical depictions of the thinning and merging properties of Poisson processes. (b2) The same data as in the two preceding figures have been thinned by randomly rejecting some blips (*gray*), with probability $\xi_{\text{thin}} = 1/2$. The remaining blips again form a Poisson process, with mean rate reduced by half. (c2) The same data have been merged with a second Poisson process with the same mean rate (*red*). The complete set of blips again forms a Poisson process, with mean rate given by the sum of the mean rates of the two contributing processes.

randomly throws red balls at the same wall at mean rate β_2. We can define a "merged process" that reports the arrival times of *either* kind of ball. The **merging property** states that the merged process is itself Poisson, with mean rate $\beta_{\text{tot}} = \beta_1 + \beta_2$. To prove it, again divide time into small slots Δt and imagine an observer who merely *hears* the balls hitting the target. Because Δt is small, the probability of getting two balls in the same time slot is negligible. Hence, the addition rule says that, in any short interval Δt, the probability of hearing a thump is $(\beta_1 \Delta t) + (\beta_2 \Delta t)$, or $\beta_{\text{tot}} \Delta t$ (see Figures 7.6c1,c2).

Example Consider three independent Poisson processes, with mean rates β_1, β_2, and β_3. Let $\beta_{\text{tot}} = \beta_1 + \beta_2 + \beta_3$.

a. After a blip of *any* type, what's the distribution of waiting times till the next event of *any* type?

b. What's the probability that any particular blip will be of type 1?

Solution a. This can be found by using the merging property and the waiting-time distribution (Idea 7.5, page 159). Alternatively, divide time into slots of duration Δt. Let's find the probability of *no* blip during a period of duration t_w (that is, $M = t_w/\Delta t$ consecutive slots). For very small Δt, the three blip outcomes become mutually exclusive, so the negation, addition, and product rules yield

$$\mathcal{P}(\text{none during } t_w) = (1 - \beta_{tot}\Delta t)^M = \left(1 - (\beta_{tot} t_w/M)\right)^M = \exp(-\beta_{tot} t_w).$$

The probability of such a period with no blip, followed by a blip of any type, is

$$\mathcal{P}(\text{none during } t_w) \times \left(\mathcal{P}(\text{type 1 during } \Delta t) + \cdots + \mathcal{P}(\text{type 3 during } \Delta t)\right) = \exp(-\beta_{tot} t_w)\beta_{tot}\Delta t.$$

Thus, the pdf of the waiting time for any type blip is $\beta_{tot}\exp(-\beta_{tot} t_w)$, as predicted by the merging property.
b. We want $\mathcal{P}(\text{blip of type 1 in } \Delta t \mid \text{blip of any type in } \Delta t) = (\beta_1 \Delta t)/(\beta_{tot}\Delta t) = \beta_1/\beta_{tot}$.

Your Turn 7C

Connect the merging property to the count distribution (Idea 7.6) and the Example on page 80.

7.3.3.3 Significance of thinning and merging properties

The two properties just proved underlie the usefulness of the Poisson process, because they ensure that, in some ways, it behaves similarly to a purely regular sequence of blips:

- Imagine a long line of professors passing a turnstile exactly once per second. You divert every third professor through a door on the left. Then the diverted stream is also regular, with a professor passing the left door once every three seconds. The thinning property states that a particular kind of *random* arrival (a Poisson process), subject to a *random* elimination (a Bernoulli trial), behaves similarly (the new process has mean rate reduced by the thinning factor).
- Imagine two long lines of professors, say of literature and chemistry, respectively, converging on a single doorway. Individuals in the first group arrive exactly once per second; those in the second group arrive once every two seconds. The stream that emerges through the door has mean rate $(1\,\text{s})^{-1} + (2\,\text{s})^{-1}$. The merging property states that a particular class of *random* processes has an even nicer property: They merge to form a new random process *of the same kind,* with mean rate again given by the sum of the two component rates.

In a more biological context,

- Section 7.2 imagined the stepping of myosin-V as a result of two sequential events: First an ATP molecule must encounter the motor's ATP-binding site, but then it must also bind and initiate stepping. It's reasonable to model the first event as a Poisson process, because most of the molecules surrounding the motor are not ATP and so cannot generate a step. It's reasonable to model the second event as a Bernoulli trial, because even when an ATP does encounter the motor, it must overcome an activation barrier to bind; thus, some fraction of the encounters will be nonproductive. The thinning property leads us to

expect that the complete stepping process will itself be Poisson, but with a mean rate lower than the ATP collision rate. We'll see in a following section that this expectation is correct.

- Suppose that two or more identical enzyme molecules exist in a cell, each continually colliding with other molecules, a few of which are substrates for a reaction that the enzymes catalyze. Each enzyme then emits product molecules in a Poisson process, just as in the motor example. The merging property leads us to expect that the *combined* production will also be a Poisson process.

7.4 More Examples

7.4.1 Enzyme turnover at low concentration

Molecular motors are examples of **mechanochemical** enzymes: They hydrolyze ATP and generate mechanical force. Most enzymes instead have purely chemical effects, for example processing substrate molecules into products.[13] At low substrate concentration, the same reasoning as in Section 7.2 implies that the successive appearances of product molecules will follow a Poisson process, with mean rate reflecting the substrate concentration, enzyme population, and binding affinity of substrate to enzyme. Chapter 8 will build on this observation.

T_2 *Section 7.5.1′a (page 171) describes some finer points about molecular turnovers.*

7.4.2 Neurotransmitter release

Nerve cells (neurons) mainly interact with each other by chemical means: One neuron releases **neurotransmitter** molecules from its "output terminal" (**axon**), which adjoins another neurons's "input terminal" (**dendrite**). Electrical activity in the first neuron triggers this release, which in turn triggers electrical activity in the second. A similar mechanism allows neurons to stimulate muscle cell contraction. When it became possible to monitor the electric potential across a muscle cell membrane,[14] researchers were surprised to find that it was "quantized": Repeated, identical stimulation of a motor neuron led to muscle cell responses with a range of peak amplitudes, and the pdf of those amplitudes consisted of a series of discrete bumps (Figure 7.7). Closer examination showed that the bumps were at integer multiples of a basic response strength. Even in the absence of any stimulus, there were occasional blips resembling those in the first bump. These observations led to the discoveries that

- Neurotransmitter molecules are packaged into bags (**vesicles**) within the nerve axon, and these vesicles all contain a roughly similar amount of transmitter. A vesicle is released either completely or not at all.
- Thus, the amount of transmitter released in response to any stimulus is roughly an integer multiple of the amount in one vesicle. Even in the absence of stimulus, an occasional vesicle can also be released "accidentally," leading to the observed spontaneous events.
- The electrical response in the muscle cell (or in another neuron's dendrite) is roughly linearly proportional to the total amount of transmitter released, and hence to the number ℓ of vesicles released.

[13] See Section 3.2.3 (page 40).

[14] The work of Katz and Miledi discussed earlier examined a much more subtle feature, the effect of discrete openings of ion channels in response to bathing a dendrite, or a muscle cell, in a fixed concentration of neurotransmitter (Section 4.3.4, page 78).

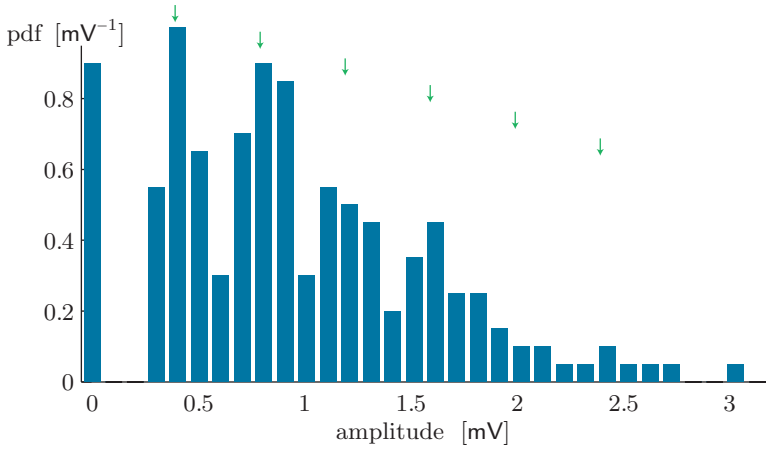


Figure 7.7 [Experimental data.] **Electrical response at a neuromuscular junction.** The bars give the estimated pdf of response amplitude, from a total of 198 stimuli. The horizontal axis gives the amplitudes (peak voltage change from rest), measured in a muscle cell in response to a set of identical stimuli applied to its motor neuron. Bumps in the distribution of amplitudes occur at 0, 1, 2, . . . , 6 times the mean amplitude of the spontaneous electrical events (*arrows*). There is some spread in each bump, mostly indicating a distribution in the number of neurotransmitter molecules packaged into each vesicle. The narrow peak near zero indicates failures to respond at all. [Data from Boyd & Martin, 1956.]

Separating the histogram in Figure 7.7 into its constituent peaks, and computing the area under each one, gave the estimated probability distribution of ℓ in response to a stimulus. This analysis showed that, at least in a certain range of stimuli, ℓ is Poisson distributed.[15] More generally,

> *If the exciting neuron is held at a constant membrane potential, then neurotransmitter vesicles are released in a Poisson process.*

7.5 Convolution and Multistage Processes

7.5.1 Myosin-V is a processive molecular motor whose stepping times display a dual character

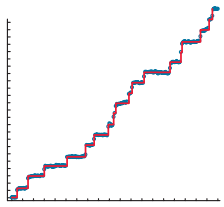


Figure 6.3c (page 134)

Figure 6.3c shows many steps of the single-molecule motor myosin-V. This motor is highly processive: Its two "feet" rarely detach simultaneously, allowing it to take many consecutive steps without ever fully unbinding from its track. The graph shows each step advancing the motor by sudden jumps of roughly 74 nm. Interestingly, however, only about one quarter of the individual myosin-V molecules studied had this character. The others *alternated* between short and long steps; the *sum* of the long and short step lengths was about 74 nm. This division at first seemed mysterious—were there two distinct kinds of myosin-V molecules? Was the foot-over-foot mechanism wrong?

Yildiz and coauthors proposed a simpler hypothesis to interpret their data:

> *All the myosin-V molecules are in fact stepping in the same way along their actin tracks. They merely differ in where the fluorescent marker, used to image the stepping, is attached to the myosin-V molecule.* (7.7)

[15] You'll explore this claim in Problem 7.14.

To see the implications of this idea, imagine attaching a light to your hip and walking in a dark room, taking 1 m steps. An observer would then see the flashlight advancing in 1 m jumps. Now, however, imagine attaching the light to your left *knee*. Each time your right foot takes a step, the left knee moves less than 1 m. Each time your left foot takes a step, however, it detaches and swings forward, moving the light by *more* than 1 m. After any two consecutive steps, the light has always moved the full 2 m, regardless of where the light was attached. This metaphor can explain the alternating stride observed in some myosin-V molecules—but is it right?

Now suppose that the light is attached to your left *ankle*. This time, the shorter steps are so short that they cannot be observed at all. All the observer sees are 2 m jumps when your left foot detaches and moves forward. The biochemical details of the fluorescent labeling used by Yildiz and coauthors allowed the fluorophore to bind in any of several locations, so they reasoned that "ankle attachment" could happen in a subpopulation of the labeled molecules. Although this logic seemed reasonable, they wanted an additional, more quantitative prediction to test it.

To find such a prediction, first recall that molecular motor stepping follows a Poisson process, with mean rate β depending on the concentration of ATP.[16] Hence, the pdf of interstep waiting times should be an Exponential distribution.[17] In fact, the subpopulation of myosin-V motors with alternating step lengths really does obey this prediction (see Figure 7.8a), as do the kinetics of many other chemical reactions. But for the other subpopulation (the motors that took 74 nm steps), the prediction fails badly (Figure 7.8b).

To understand what's going on, recall the hypothesis of Yildiz and coauthors for the nature of stepping in the 74 nm population, which is that the first, third, fifth, . . . steps are *not visible*. Therefore, what appears to be the αth interstep waiting time, $t'_{w,\alpha}$, is actually the *sum* of two consecutive waiting times:

$$t'_{w,\alpha} = t_{w,2\alpha} + t_{w,2\alpha-1}.$$

Even if the true waiting times are Exponentially distributed, we will still find that the apparent waiting times t'_w have a different distribution, namely, the convolution.[18] Thus,

$$\wp_{t'_w}(t'_w) = \int_0^{t'_w} dx\, \wp_{exp}(x;\beta) \times \wp_{exp}(t'_w - x;\beta), \tag{7.8}$$

where x is the waiting time for the first, invisible, substep.

Example a. Explain the limits on the integral in Equation 7.8.
b. Do the integral.
c. Compare your result qualitatively with the histograms in Figures 7.8a,b.
d. Discuss how your conclusion in (c) supports Idea 7.7.

Solution a. x is the waiting time for the invisible first substep. It can't be smaller than zero, nor can it exceed the specified total waiting time t'_w for the first and second substeps.

---

[16] See Section 7.2.
[17] See Idea 7.5 (page 159).
[18] See Section 4.3.5 (page 79).

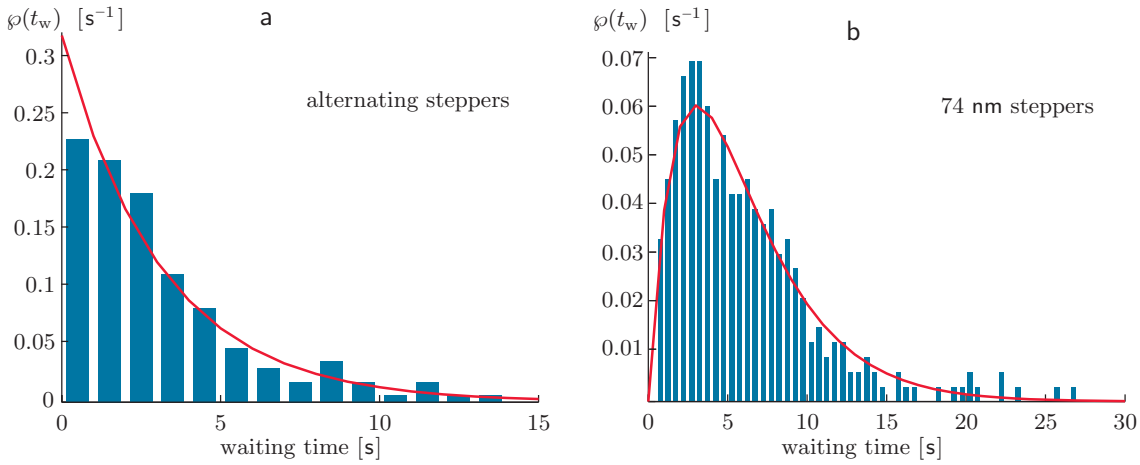


Figure 7.8 [Experimental data with fits.] **The stepping of molecular motors.** (a) Estimated pdf of the waiting times for the subpopulation of myosin-V molecules that displayed alternating step lengths, superimposed on the expected Exponential distribution (see Problem 7.8). (b) Similar graph for the other subpopulation of molecules that displayed only long steps, superimposed on the distribution derived in the Example on page 166. The shape of the curve in (b) is the signature of a random process with two alternating types of substep. Each type of substep has Exponentially distributed waiting times with the same mean rate as in (a), but only one of them is visible. [Data from Yildiz et al., 2003.]

b.

$$\beta^2 \int_0^{t'_w} dx\, \exp\bigl(-\beta x - \beta(t'_w - x)\bigr) = \beta^2 e^{-\beta t'_w} \int_0^{t'_w} dx = \beta^2 t'_w e^{-\beta t'_w}.$$

c. The function just found falls to zero at $t'_w \to 0$ and $t'_w \to \infty$. In between these extremes, it has a bump. The experimental data in Figure 7.8b have the same qualitative behavior, in contrast to those in panel (a).

d. The hypothesis under study predicted that behavior, because Figure 7.8b shows that the molecules with unimodal step length distributions are also the ones for which the hypothesis says that half the steps are invisible.

In fact, fitting the histogram in Figure 7.8a leads to a value for the mean rate β, and hence to a completely unambiguous prediction (no further free parameters) for the histogram in Figure 7.8b. That prediction was confirmed.[19] Yildiz and coauthors concluded that the correlation between which sort of step lengths a particular molecule displayed (bimodal versus single-peak histogram of step lengths) and which sort of stepping kinetics it obeyed (Exponential versus other) gave strong support for the model of myosin-V as stepping foot-over-foot.[20]

T_2 *Section 7.5.1′ (page 171) discusses more detailed descriptions of some of the processes introduced in this chapter.*

[19] See Problem 7.8.

[20] Later experiments gave more direct evidence in favor of this conclusion; see Media 11.

7.5.2 The randomness parameter can be used to reveal substeps in a kinetic scheme

The previous section discussed the probability density function $\wp(t_w) = \beta^2 t_w e^{-\beta t_w}$, which arose from a sequential process with two alternating substeps, each Poisson.

Your Turn 7D

a. Find the expectation and variance of t_w in this distribution. [*Hint:* If you recall where this distribution came from, then you can get the answers with an *extremely* short derivation.]
b. The **randomness parameter** is defined as $\langle t_w \rangle / \sqrt{\text{var } t_w}$; compute it. Compare your result with the corresponding quantity for the Exponential distribution (Idea 7.5, page 159).
c. Suggest how your answers to (b) could be used to invent a practical method for discriminating one- and two-step processes experimentally.

7.6 Computer Simulation

7.6.1 Simple Poisson process

We have seen that, for a simple process, the distribution of waiting times is Exponential.[21] This result is useful if we wish to ask a computer to simulate a Poisson process, because with it, we can avoid stepping through the vast majority of time slots in which nothing happens. We just generate a series of Exponentially distributed intervals $t_{w,1}, \ldots$, then define the time of blip α to be $t_\alpha = t_{w,1} + \cdots + t_{w,\alpha}$, the accumulated waiting time.

A computer's basic random-number function has a Uniform, not an Exponential, distribution. However, we can convert its output to get what we need, by adapting the Example on page 106. This time the transformation function is $G(t_w) = e^{-\beta t_w}$, whose inverse gives $t_w = -\beta^{-1} \ln y$.

Your Turn 7E

a. Think about how the units work in the last formula given.
b. Try this formula on a computer for various values of β, making histograms of the results.

7.6.2 Poisson processes with multiple event types

We'll need a slight extension of these ideas, called the **compound Poisson process**, when we discuss chemical reactions in Chapter 8. Suppose that we wish to simulate a process consisting of two types of blip. Each type arrives independently of the other, in Poisson processes with mean rates β_a and β_b, respectively. We could simulate each series separately and merge the lists, sorting them into a single ascending sequence of blip times accompanied by their types (a or b).

There is another approach, however, that runs faster and admits a crucial generalization that we will need later. We wish to generate a single list $\{(t_\alpha, s_\alpha)\}$, where t_α are the event times (continuous), and s_α are the corresponding event types (discrete).

[21] See Idea 7.5 (page 159).

Example a. The successive differences of t_α values reflect the waiting times for *either* type of blip to happen. Find their distribution.
b. Once something happens, we must ask *what* happened on that step. Find the discrete distribution for each s_α.

Solution a. By the merging property, the distribution is Exponential, with $\beta_{\text{tot}} = (\beta_a + \beta_b)$.[22]
b. It's a Bernoulli trial, with probability $\xi = \beta_a/(\beta_a + \beta_b)$ to yield an event of type *a*.

We already know how to get a computer to draw from each of the required distributions. Doing so gives the solution to the problem of simulating the compound Poisson process:

Your Turn 7F

Write a short computer code that uses the result just found to simulate a compound Poisson process. That is, your code should generate the list $\{(t_\alpha, s_\alpha)\}$ and represent it graphically.

THE BIG PICTURE

This chapter concludes our formal study of randomness in biology. We have moved conceptually from random systems that yield one discrete value at a time, to continuous single values, and now on to random processes, which yield a whole time series. At each stage, we found biological applications that involve both characterizing random systems and deciding among competing hypotheses. We have also seen examples, like the Luria-Delbrück experiment, where it was important to be able to simulate the various hypotheses in order to find their predictions in a precise, and hence falsifiable, form.

Chapter 8 will apply these ideas to processes involving chemical reactions.

KEY FORMULAS

- *Exponential distribution:* In the limit where $\Delta t \to 0$, holding fixed β and t_w, the Geometric distribution with probability $\xi = \beta \Delta t$ approaches the continuous form $\wp_{\text{exp}}(t_w; \beta) = \beta \exp(-\beta t_w)$. The expectation of the waiting time is $1/\beta$; its variance is $1/\beta^2$. The parameter β has dimensions $\mathbb{T}^{-1}$, as does $\wp_{\text{exp}}$.
- *Poisson process:* A random process is a random system, each of whose draws is an increasing sequence of numbers ("blip times"). The Poisson process with mean rate β is a special case, with the properties that (*i*) any infinitesimal time slot from t to $t + \Delta t$ has probability $\beta \Delta t$ of containing a blip, and (*ii*) the number in one such slot is statistically independent of the number in any other (nonoverlapping) slot.
  The waiting times in a Poisson process are Exponentially distributed, with expectation β^{-1}.
  For a Poisson process with mean rate β, the probability of getting ℓ blips in any time interval of duration T_1 is Poisson distributed, with expectation $\mu = \beta T_1$.
- *Thinning property:* When we randomly eliminate some of the blips in a Poisson process with mean rate β, by subjecting each to an independent Bernoulli trial, the remaining blips form another Poisson process with $\beta' = \xi_{\text{thin}} \beta$.

[22] See Section 7.3.3.2 (page 161).

- *Merging property:* When we combine the blips from two Poisson processes with mean rates β_1 and β_2, the resulting time series is another Poisson process with $\beta_{\text{tot}} = \beta_1 + \beta_2$.
- *Alternating-step process:* The convolution of two Exponential distributions, each with mean rate β, is not itself an Exponential; its pdf is $t_w \beta^2 e^{-\beta t_w}$.
- *Randomness parameter:* The quantity $\langle t_w \rangle / \sqrt{\text{var } t_w}$ can be estimated from experimental data. If the data form a simple Poisson process, then this quantity will be equal to one; if on the contrary the blips in the data reflect two or more obligatory substeps, each of which has Exponentially-distributed waiting times, then this quantity will be larger than one.

FURTHER READING

Semipopular:
Molecular machines: Hoffmann, 2012.

Intermediate:
Allen, 2011; Jones et al., 2009; Wilkinson, 2006.
Molecular motors: Dill & Bromberg, 2010, chapt. 29; Nelson, 2014, chapt. 10; Phillips et al., 2012, chapt. 16; Yanagida & Ishii, 2009.

Technical:
Jacobs, 2010.
Yildiz et al., 2003.
Kinetics of other enzymes and motors: Hinterdorfer & van Oijen, 2009, chapts. 6–7.

T2 **Track 2**

7.2′ More about motor stepping

Section 7.2 made some idealizations in order to arrive at a simple model. Much current research involves finding more realistic models that are complex enough to explain data, simple enough to be tractable, and physically realistic enough to be more than just data summaries.

For example, after a motor steps, the world contains one fewer ATP (and one more each of ADP and phosphate, the products of ATP hydrolysis). Our discussion implicitly assumed that so much ATP is available, and the solution is so well mixed, that depletion during the course of an experiment is negligible. In some experiments, this is ensured by constantly flowing fresh ATP-bearing solution into the chamber; in cells, homeostatic mechanisms adjust ATP production to meet demand.[23]

In other words, we assumed that both the motor *and its environment* have no memory of prior steps. Chapter 8 will develop ideas relevant for situations where this Markov property may not be assumed.

We also neglected the possibility of backward steps. In principle, a motor could bind an ADP and a phosphate from solution, step backward, and emit an ATP. Inside living cells, the concentration of ATP is high enough, and those of ADP and phosphate low enough, that such steps are rare.

T2 **Track 2**

The main text stated that enzyme turnovers follow a Poisson process. Also, Figure 3.2b suggests that the arrivals of the energy packets we call "light" follow such a random process. Although these statements are good qualitative guides, each needs some elaboration.

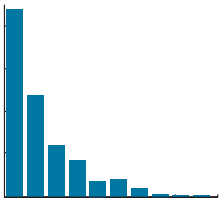


Figure 3.2b (page 38)

7.5.1′a More detailed models of enzyme turnovers

Remarkably, enzymes do display long-term "memory" effects, as seen in these examples:

- The model of myosin-V stepping discussed in the main text implicitly assumed that the motor itself has no "stopwatch" that affects its binding probability based on recent history. However, immediately after a binding event, there is a "dead time" while the step is actually carried out. During this short time, the motor cannot initiate another step. The time bins of Figure 7.8 are too long to disclose this phenomenon, but it has been seen.
- An enzyme can get into substates that can persist over many processing cycles, and that have different mean rates from other substates. The enzyme cycles through these substates, giving its apparent mean rate a long-term drift. More complex Markov models than the one in the main text are needed to account for this behavior (English et al., 2006).

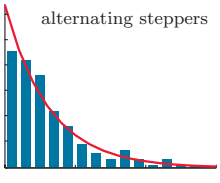


Figure 7.8a (page 167)

7.5.1′b More detailed models of photon arrivals

Actually, only laser light precisely follows a Poisson process. "Incoherent" light, for example, from the Sun, has more complicated photon statistics, with some autocorrelation.

[23] See Chapter 9.

PROBLEMS

7.1 Ventricular fibrillation

A patient with heart disease will sometimes enter "ventricular fibrillation," leading to cardiac arrest. The following table shows data on the fraction of patients failing to regain normal heart rhythm after attempts at defibrillation by electric shock, in a particular clinical trial:

number of attempts	fraction persisting in fibrillation
1	0.37
2	0.15
3	0.07
4	0.02

Assume that with 0 attempts there are no spontaneous recoveries. Also assume that the probability of recovery on each attempt is independent of any prior attempts. Suggest a formula that roughly matches these data. If your formula contains one or more parameters, estimate their values. Make a graph that compares your formula's prediction with the data above. What additional information would you need in order to assert a credible interval on your parameter value?

7.2 Basic properties of $\mathcal{P}_{\text{geom}}$

a. Continue along the lines of Your Turn 3D to find the expectation and variance of the Geometric distribution. [*Hint:* You can imitate the Example on page 77. Consider the quantity

$$\frac{\mathrm{d}}{\mathrm{d}\xi}\sum_{j=0}^{\infty}(1-\xi)^j.$$

Evaluate this quantity in two different ways, and set your expressions equal to each other. The resulting identity will be useful.]

b. Discuss how your answers to (a) behave as ξ approaches 0 and 1, and how these behaviors qualitatively conform to your expectations.

c. Review Your Turn 3D (page 48), then modify it as follows. Take the Taylor series expansion for $1/(1-z)$, multiply by $(1-z^K)$, and simplify the result (see page 19). Use your answer to find the total probability that, in the Geometric distribution, the first "success" occurs at *or before* the Kth attempt.

d. Now take the continuous-time limit of your results in (a) and compare them with the corresponding facts about the Exponential distribution (see the Example on page 160).

7.3 Radiation-induced mutation

Suppose that we maintain some single-cell organisms under conditions where they don't divide. Periodically we subject them to a dose of radiation, which sometimes induces a mutation in a particular gene. Suppose that the probability for a given individual to form a mutation after a dose is $\xi = 10^{-3}$, regardless how many doses have previously been given. Let j be the number of doses after which a particular individual develops its first mutation.

a. State the probability distribution of the random variable j.

b. What is the expectation, $\langle j \rangle$?

c. Now find the variance of j.

7.4 Winning streaks via simulation

Section 7.3.1 found a formula for the number of attempts we must make before "success" in a sequence of independent Bernoulli trials. In this problem, you'll check that result by a computer simulation. Simulation can be helpful when studying more complex random processes, for which analytic results are not available.

Computers are very fast at finding patterns in strings of symbols. You can make a long string of N random digits by successively appending the string `"1"` or `"0"` to a growing string called `flipstr`. Then you can ask the computer to search `flipstr` for occurrences of the substring `"1"`, and report a list of all the positions in the long string that match it. The differences between successive entries in this list are related to the length of runs of consecutive `"0"` entries. Then you can tabulate how often various waiting times were observed, and make a histogram.

Before carrying out this simulation, you should try to guess what your graph will look like. Nick reasoned, "Because heads is a rare outcome, once we get a tails we're likely to get a lot of them in a row, so short strings of zeros will be less probable than medium-long strings. But eventually we're bound to get a heads, so *very* long strings of zeros are also less common than medium-long strings. So the distribution should have a bump." Think about it—is that the right reasoning?

Now get your answer, as follows. Write a simple simulation of the sort described above, with $N = 1000$ "attempts" and $\xi = 0.08$. Plot the frequencies of appearance of strings of various lengths, both on regular and on semilog axes. Is this a familiar-looking probability distribution? Repeat with $N = 50\,000$.

7.5 Transformation of exponential distribution

Suppose that a pdf is known to be of Exponential form, $\wp_t(t) = \beta \exp(-\beta t)$. Let $y = \ln(t/(1\,\mathrm{s}))$ and find the corresponding function $\wp_y(y)$. Unlike the Exponential, the transformed distribution has a bump, whose location y_* tells something about the rate parameter β. Find this relation.

7.6 Likelihood analysis of a poisson process

Suppose that you measure a lot of waiting times from some random process, such as the stepping of a molecular motor. You believe that these times are draws from an Exponential distribution: $\wp(t) = Ae^{-\beta t}$, where A and β are constants. But you don't know the values of these constants. Moreover, you only had time to measure six steps, or five waiting times $t_1, \ldots, t_5$, before the experiment ended.[24]

a. A and β are not independent quantities: Express A in terms of β. State some appropriate units for A and for β.

b. Write a symbolic expression for the likelihood of any particular value of β, in terms of the measured data $t_1, \ldots, t_5$.

c. Find the maximum-likelihood estimate of the parameter β; give a short derivation of your formula.

7.7 Illustrate thinning property

a. Obtain Dataset 3, which gives blip arrival times from a sensitive light detector in dim light. Have a computer find the waiting times between events, and histogram them.

[24] Perhaps the motor detached from its track in the middle of interval #6.

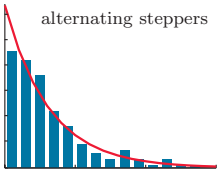


Figure 7.8a (page 167)

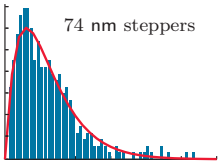


Figure 7.8b (page 167)

b. Apply an independent Bernoulli trial to each event in (a), which accepts 60% of them and rejects the rest. Again histogram the waiting times, and comment.

7.8 Hidden steps in myosin-V

If you haven't done Problem 7.6, do it before this problem. Figure 7.8 shows histograms of waiting times for the stepping of two classes of fluorescently labeled myosin-V molecules. The experimenters classified each motor molecule that they observed, according to whether it took steps of two alternating lengths or just a single length. For each class, they reported the frequencies for taking a step after various waiting times. For example, 39 motor steps were observed with t_w between 0 and 1 s.

a. Obtain Dataset 9, and use it to generate the two histograms in Figure 7.8.

b. Section 7.5.1 (page 165) proposed a physical model for this class of motors, in which the waiting times were distributed according to an Exponential distribution. Use the method in Problem 7.6 to infer from the data the value of β for the molecules that took steps of alternating lengths. [*Hint:* The model assumes that all steps are independent, so the *order* in which various waiting times were observed is immaterial. What matters is just the number of times that each t_w was observed. The data have been binned; Dataset 9 contains a list whose first entry (0.5, 39) means that the bin centered on 0.5 s contained 39 observed steps. Make the approximation that all 39 of these steps had t_w exactly equal to 0.5 s (the middle of the first bin), and so on.]

c. Graph the corresponding probability density function superimposed on the data. To make a proper comparison, rescale the pdf so that it becomes a prediction of the frequencies.

d. Section 7.5.1 also proposed that in the other class of molecules, half the steps were unobserved. Repeat (b–c) with the necessary changes.

e. Compare the values of β that you obtained in (b,d). If they are similar (or dissimilar), how do you interpret that?

f. Now consider a different hypothesis that says that each observed event is the last of a series of m sequential events, each of which is an independent, identical Poisson process. (Thus, (d) considered the special case $m = 2$.) Without doing the math, qualitatively what sort of distribution of wait times $\wp(t_w)$ would you expect for $m = 10$?

7.9 Asymmetric foot-over-foot cycle

Suppose that some enzyme reaction consists of two steps whose waiting times are independent, except that they must take place in strict alternation: $A_1 B_1 A_2 B_2 A_3 \cdots$. For example, the enzyme hexokinase alternates between cleaving a phosphate from ATP and transferring it to glucose. Or we could study a motor that walks foot-over-foot, but unlike the main text we won't assume equal rate constants for each foot.

Successive pauses are statistically independent. The pause between an A step and the next B step is distributed according to $\wp_{AB}(t) = \beta e^{-\beta t_w}$, where β is a constant with dimensions $\mathbb{T}^{-1}$. The pause between a B step and the next A step is similarly distributed, but with a different mean rate β'. Find the probability density function for the time between two successive A steps.

7.10 Staircase plot

a. Use a computer to simulate 30 draws from the Exponential distribution with mean rate $0.3\,\mathrm{s}^{-1}$. Call the results `w(1),...,w(30)`. Create a list with the cumulative sums, then duplicate them and append a 0, to get `0, w(1), w(1), w(1)+w(2),`

`w(1)+w(2),...`. Create another list `x` with entries `0, 0, step, step, 2*step, 2*step, ...`, where `step=37`, and graph the `w`'s versus `x`. Interpret your graph by identifying the entries of `w` with the interstep waiting times of length `step`.

b. Actually, your graph is not quite a realistic simulation of the observed steps of myosin-V. Adapt your code to account for the alternating step lengths observed by Yildiz and coauthors in one class of fluorescently labeled motor molecules.

c. This time adapt your code to account for the non-Exponential distribution of waiting times observed in the other class of motors. Does your graph resemble some data discussed in the main text?

7.11 Thinning via simulation

Take the list of waiting times from your computer simulation in Your Turn 7E, and modify it by deleting some blips, as follows. Walk through the list, and for each entry $t_{w,i}$ make a Bernoulli trial with some probability ξ_*. If the outcome is heads, move to the next list entry; otherwise, delete the entry $t_{w,i}$ and add its value to that of the next entry in the list. Run this modified simulation, histogram the outcomes, and so check the thinning property (Section 7.3.3.1).

7.12 Convolution via simulation

a. Use the method in Section 7.6.1 (page 168) to simulate draws from the Exponential distribution with expectation 1 s.

b. Simulate the random variable $z = x + y$, where x and y are independent random variables with the distribution used in (a). Generate a lot of draws from this distribution, and histogram them.

c. Compare your result in (b) with the distribution found in the Example on page 166.

d. Next simulate a random variable defined as the sum of *50* independent, Exponentially distributed variables. Comment on your result in the light of Problem 5.7 (page 119).

7.13 T_2 Fit count data

Radioactive tagging is important in many biological assays. A sample of radioactive substance furnishes another physical system found to produce blips in a Poisson process.

Suppose that we have a radioactive source of fixed intensity, and a detector that registers individual radiation particles emitted from the source. The average rate at which the detector emits blips depends on its distance L to the source. We measure the rate by holding the detector at a series of fixed distances $L_1, \ldots, L_N$. At each distance, we count the blips on the detector over a fixed time $\Delta T = 15$ s and record the results.

Our physical model for these data is the inverse-square law: We expect the observed number of detector blips at each fixed L to be drawn from a Poisson distribution with expectation equal to A/L^2 for some constant A.[25] We wish to test that model. Also we would like to know the constant of proportionality A, so that we can use it to deduce the rate for any value L (not just the ones that we measured). In other words, we'd like to summarize our data with an **interpolation formula**.

a. One way to proceed might be to plot the observed number of blips y versus the variable $x = (L)^{-2}$, then lay a ruler along the plot in a way that passes through $(0, 0)$ and roughly

[25] This constant reflects the intensity of the source, the duration of each measurement, and the size and efficiency of the detector.

tracks the data points. Obtain Dataset 10 and follow this procedure to estimate A as the slope of this line.

b. A better approach would make an objective fit to the data. Idea 6.8 (page 139) is not applicable to this situation—why not?

c. But the logic leading to Idea 6.8 is applicable, with a simple modification. Carry this out, plot the log-likelihood as a function of A, and choose the optimal value. Your answer from (a) gives a good starting guess for the value of A; try various values near that. Add the best-fit line according to maximum likelihood to the plot you made in (a).

d. You can estimate the integral of the likelihood function by finding its sum over the range of A values you graphed in (c) and normalizing. Use those values to estimate a 95% credible interval for the value of A.

Comment: You may still be asking, "But is the best fit *good*? Is the likelihood *big enough* to call it good?" One way to address this is to take your best-fit model, use it to generate lots of *simulated datasets* by drawing from appropriate Poisson distributions at each x_i, calculate the likelihood function for each one, and see if the typical values thus obtained are comparable to the best-fit likelihood you found by using the real data.

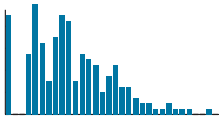


Figure 7.7 (page 165)

7.14 T_2 Quantized neurotransmitter release

The goal of this problem is to predict the data in Figure 7.7 with no fitting parameters. First obtain Dataset 11, which contains binned data on the frequencies with which various peak voltage changes were observed in a muscle cell stimulated by a motor neuron. In a separate measurement, the authors also studied spontaneous events (no stimulus), and found the sample mean of the peak voltage to be $\mu_V = 0.40\,\text{mV}$ and its estimated variance to be $\sigma^2 = 0.00825\,\text{mV}^2$.

The physical model discussed in the text states that each response is the sum of ℓ independent random variables, which are the responses caused by the release of ℓ vesicles. Each of these constituents is itself assumed to follow a Gaussian distribution, with expectation and variance given by those found for spontaneous events.

a. Find the predicted distribution of the responses for the class of events with some definite value of ℓ.

b. The model also assumes that ℓ is itself a Poisson random variable. Find its expectation μ_ℓ by computing the sample mean of all the responses in the dataset, and dividing by the mean response from a single vesicle, μ_V.

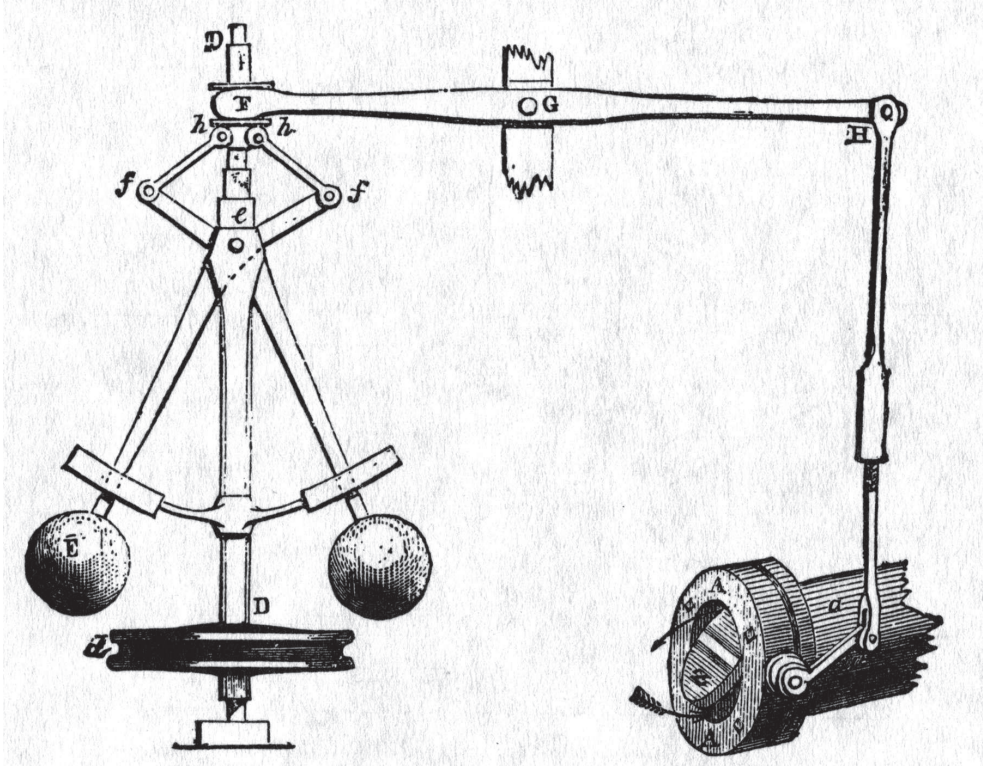
c. Take the distributions you found in (a) for $\ell > 0$, scale each by $\mathcal{P}_{\text{pois}}(\ell; \mu_\ell)$, and add them to find an overall pdf. Plot this pdf.

d. Superimpose the estimated pdf obtained from the data on your graph from (c).

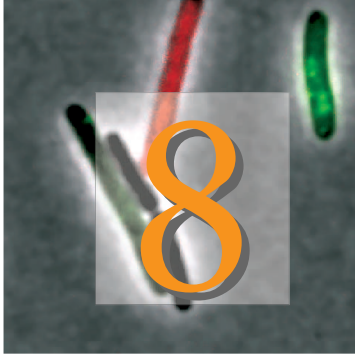
e. The experimenters also found in this experiment that, in 18 out of 198 trials, there was no response at all. Compute $\mathcal{P}_{\text{pois}}(0; \mu_\ell)$ and comment.

PART III

Control in Cells



The centrifugal governor, a mechanical feedback mechanism. [From *Discoveries and inventions of the nineteenth century*, by R Routledge, 13th edition, published 1900.]



Randomness in Cellular Processes

I think there is a world market for maybe five computers.
—Thomas Watson, Chairman of IBM, 1943

8.1 Signpost

Earlier chapters have emphasized that randomness pervades biology and physics, from subcellular actors (such as motor proteins), all the way up to populations (such as colonies of bacteria). Ultimately, this randomness has its origin in physical processes, for example, thermal motion of molecules. Although it may sound paradoxical, we have found that it is possible to characterize randomness precisely and reproducibly, sometimes with the help of physical models.

This chapter will focus on the particular arena of cellular physiology. A living cell is made of molecules, so those molecules implement all its activities. It is therefore important to understand in what ways, and to what extent, those activities are random.

The Focus Question is

Biological question: How and when will a collection of random processes yield overall dynamics that are nearly predictable?

Physical idea: Deterministic collective behavior can emerge when the copy number of each actor is large.

8.2 Random Walks and Beyond

8.2.1 Situations studied so far

8.2.1.1 Periodic stepping in random directions

One of the fundamental examples of randomness given in Chapter 3 was Brownian motion.[1] Earlier sections discussed an idealization of this kind of motion as a **random walk**: We imagine that an object periodically takes a step to the left or right, with length always equal to some constant d. The only state variable is the object's current position. This simple random process reproduces the main observed fact about Brownian motion, which is that the mean-square deviation of the displacement after many steps is proportional to the square root of the elapsed time.[2] Still, we may worry that much of the chaos of real diffusion is missing from the model. Creating a more realistic picture of random walks will also show us how to model the kinetics of chemical reactions.

8.2.1.2 Irregularly timed, unidirectional steps

We studied another kind of random process in Chapter 7: the stepping of a processive molecular motor, such as myosin-V. We allowed for randomness in the step times, instead of waiting for some fictitious clock to tick. But the step displacements themselves were predictable: The experimental data showed that they are always in the same direction, and of roughly the same length.

8.2.2 A more realistic model of Brownian motion includes both random step times and random step directions

One way to improve the Brownian motion model is to combine the two preceding ideas (random step times and random directions). As usual, we begin by simplifying the analysis to a single spatial dimension. Then one reasonable model would be to say that there is a certain fixed probability per unit time, β, of a small suspended particle being kicked to the left. Independently, there is also a fixed probability per unit time β of the particle being kicked to the right. Chapter 7 discussed one way to simulate such a process:[3] We first consider the merged Poisson process with mean rate 2β, and draw a sequence of waiting times from it. Then for each of these times, we draw from a Bernoulli trial to determine whether the step at that time was rightward or leftward. Finally, we make cumulative sums of the steps, to find the complete simulated trajectory as a function of time.

Figure 8.1 shows two examples of the outcome of such a simulation. Like the simpler random walks studied earlier, this one has the property that after enough steps are taken, a trajectory can end up arbitrarily far from its starting point. Even if several different walkers all start at the same position x_{ini} (variance is zero for the starting position), the variance of their positions after a long time, $\text{var}\big(x(t)\big)$, grows without bound as t increases. There is no limiting distribution of positions as time goes to infinity.

---

[1] See point **5** on page 36, and follow-up points **5a** and **5b**.
[2] See Problem 4.5.
[3] See Section 7.6.2 (page 168).

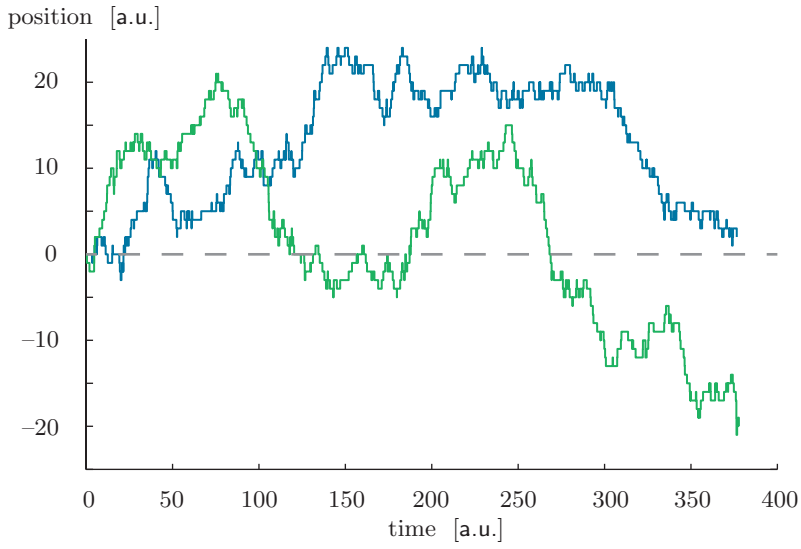


Figure 8.1 [Computer simulation.] **Computer simulation of a random walk.** The curves show two runs of the simulation. In each case, 400 steps were taken, with Exponentially distributed waiting times and mean rate of one step per unit time. Each step was of unit length, with direction specified by the outcome of a Bernoulli trial with equal probability to step up or down ($\xi = 1/2$). See Problem 8.1.

8.3 Molecular Population Dynamics as a Markov Process

Brownian motion is an example of a Markov process:[4] In order to know the probability distribution of the position at time t, all we need to know is the actual position at any *one* time t' prior to t. Any additional knowledge of the actual position at a time t'' earlier than t' gives us no additional information relevant to $\wp_{X(t)}$. If a random process has a limited amount of "state" information at any time, and the property that knowing this state at one t' completely determines the pdf of possible states at any later time t, then the process is called "Markov."

The examples discussed in Section 8.2 (stepping with irregular directions, times, or both) all have the Markov property. Chemical reactions in a well-mixed system, although also Markovian, have an additional complication:[5] They, too, occur after waiting times that reflect a Poisson process, but with a rate that depends on the concentrations of the reacting molecules.

The following discussion will introduce a number of variables, so we summarize them here:

Δt	time step, eventually taken to zero
ℓ_i	number of molecules at time $t_i = (\Delta t)i$, a random variable
ℓ_{ini}	initial value, a constant
β_s	mean rate of mRNA synthesis, a constant
$\beta_\varnothing$	mean rate of mRNA clearance (varies over time)
$k_\varnothing$	clearance rate constant
ℓ_*	steady final value

[4] See Section 3.2.1 (page 36).
[5] See point **5c** on page 40.

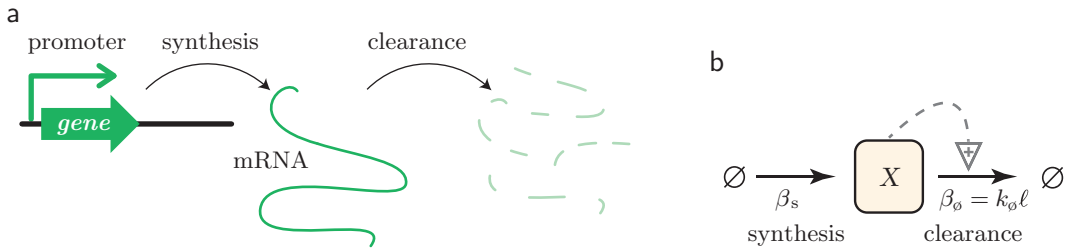


Figure 8.2 Example of a birth-death process. (a) [Schematic.] A gene (*wide arrow*) directs the synthesis of messenger RNA, which is eventually degraded (cleared) by enzymatic machinery in a cell. (b) [Network diagram.] Abstract representation as a network diagram. The *box* represents a state variable of the system, the inventory (number of copies) of some molecular species X, labeled by its name. Incoming and outgoing *black arrows* represent processes (biochemical reactions) that increase or decrease the inventory. Substrate molecules needed to synthesize X are assumed to be maintained at a fixed concentration by processes not of interest to us; they are collectively represented by the symbol $\emptyset$. Product molecules arising from the clearance of X are assumed not to affect the reaction rates; they, too, are collectively represented by a $\emptyset$. The enzymatic machinery that performs both reactions, and even the gene that directs the synthesis, are not shown at all. The two arrows are assumed to be irreversible reactions, a reasonable assumption for many cellular processes. In situations when this may not be assumed, the reverse reactions will be explicitly indicated in the network diagram by separate arrows. The *dashed arrow* is an influence line indicating that the rate of clearance depends on the level at which X is present. This particular dependence is usually tacitly assumed, however; henceforth we will not explicitly indicate it.

8.3.1 The birth-death process describes population fluctuations of a chemical species in a cell

To make these ideas concrete, imagine a very simple system called the **birth-death process** (Figure 8.2). The system involves just two chemical reactions, represented by arrows in panel (b), and one state variable, represented by the box. The state variable is the number ℓ of molecules of some species X; the processes modify this number.

The "synthesis" reaction is assumed to have fixed probability per unit time β_s to create (synthesize) new molecules of X. Such a reaction is called **zeroth order** to indicate that its mean rate is assumed to be independent of the numbers of other molecules present (it's proportional to those numbers raised to the power zero). Strictly speaking, no reaction can be independent of every molecular population. However, some cellular processes, in some situations, are effectively zeroth order, because the cell maintains roughly constant populations of the needed ingredients (substrate molecules and enzymes to process them), and the distributions of those molecules throughout the cell are unchanging in time.[6]

The other reaction, "clearance," has probability per unit time $\beta_\emptyset$ to eliminate an X molecule, for example, by converting it to something else. Unlike β_s, however, $\beta_\emptyset$ is assumed to depend on ℓ, via[7]

$$\beta_\emptyset = k_\emptyset \ell. \tag{8.1}$$

[6] T_2 We are also assuming that the population of product molecules is too small to inhibit additional production. In vitro experiments with molecular motors are another case where the zeroth-order assumption is reasonable, because substrate molecules (ATP) in the chamber are constantly replenished, and product molecules (ADP) are constantly being removed. Thus, Chapter 7 implicitly assumed that no appreciable change in the concentrations occurs over the course of the experiment.

[7] Reaction rates in a cell may also change with time due to changes in cell volume; see Section 9.4.5. Here we assume that the volume is constant.

We can think of this formula in terms of the merging property: Each of ℓ molecules has its own independent probability per unit time to be cleared, leading to a merged process for the overall population to decrease by one unit.

The constant of proportionality $k_\emptyset$ is called the **clearance rate constant**. This kind of reaction is called **first order**, because Equation 8.1 assumes that its rate is proportional to the first power of ℓ; for example, the reaction stops when the supply of X is exhausted ($\ell = 0$).

The birth-death process is reminiscent of a fundamental theme in cell biology: A gene, together with the cell's transcription machinery, implements the first arrow of Figures 8.2a,b by synthesizing messenger RNA (mRNA) molecules, a process called **transcription**. If the number of copies of the gene and the population of RNA polymerase machines are both fixed, then it seems reasonable to assume that this reaction is effectively zeroth order. The box in the figure represents the number of RNA molecules present in a cell, and the arrow on the right represents their eventual destruction. Certainly this picture is highly simplified: Cells also duplicate their genes, regulate their transcription, divide, and so on. We will add those features to the physical model step by step. For now, however, we consider only the two processes and one inventory shown in Figure 8.2. We would like to answer questions concerning both the overall development of the system, and its variability from one trial to the next.

We can make progress understanding the birth-death process by making an analogy: It is *just another kind of random walk.* Instead of wandering in ordinary space, the system wanders in its *state space,* in this case the number line of nonnegative integers ℓ. The only new feature is that, unlike in Section 8.2.2, one of the reaction rates is not a constant (see Equation 8.1). In fact, the value of that mean rate at any moment is itself a random variable, because it depends on ℓ. Despite this added level of complexity, however, the birth-death process still has the Markov property. To show this, we now find how the pdf for $\ell(t)$ depends on the system's prior history.

As usual, we begin by slicing time into slots of very short duration Δt, so that slot i begins at time $t_i = (\Delta t)i$, and by writing the population as ℓ_i instead of $\ell(t)$. During any slot i, the most probable outcome is that nothing new happens, so ℓ is unchanged: $\ell_{i+1} = \ell_i$. The next most probable outcomes are that synthesis, or clearance, takes place in the time slot. The probability of two or more reactions in Δt is negligible, for small enough Δt, and so we only need to consider the cases where ℓ changes by ± 1, or not at all. Expressing this reasoning in a formula,

$$\mathcal{P}(\ell_{i+1} \mid \ell_1, \ldots, \ell_i) = \begin{cases} (\Delta t)\beta_s & \text{if } \ell_{i+1} = \ell_i + 1; \quad \text{(synthesis)} \\ (\Delta t)k_\emptyset \ell_i & \text{if } \ell_{i+1} = \ell_i - 1; \quad \text{(clearance)} \\ 1 - (\Delta t)(\beta_s + k_\emptyset \ell_i) & \text{if } \ell_{i+1} = \ell_i; \quad \text{(no reaction)} \\ 0 & \text{otherwise.} \end{cases} \tag{8.2}$$

The right-hand side depends on ℓ_i, but not on $\ell_1, \ldots, \ell_{i-1}$, so Equation 8.2 defines a Markov process. We can summarize this formula in words:

> *The birth-death process resembles a compound Poisson process during the waiting time between any two consecutive reaction steps. After each step, however, the mean rate of the clearance reaction can change.* (8.3)

Given some starting state ℓ_* at time zero, the above characterization of the birth-death process determines the probability distribution for any question we may wish to ask about the state at a later time.

8.3.2 In the continuous, deterministic approximation, a birth-death process approaches a steady population level

Equation 8.2 looks complicated. Before we attempt to analyze the behavior arising from such a model, we should pause to get some intuition from an approximate treatment.

Some chemical reactions involve huge numbers of molecules. In this case, ℓ is a very large integer, and changing it by one unit makes a negligible relative change. In such situations, it makes sense to pretend that ℓ is actually continuous. Moreover, we have seen in several examples how large numbers imply small relative fluctuations in a discrete random quantity; so it seems likely that in these situations we may also pretend that ℓ varies deterministically. Restating Equation 8.2 with these simplifications yields the **continuous, deterministic approximation**, in which ℓ changes with time according to[8]

$$\frac{\mathrm{d}\ell}{\mathrm{d}t} = \beta_\mathrm{s} - k_\varnothing \ell. \tag{8.4}$$

Example Explain why Equation 8.4 emerges from Equation 8.2 in this limit.

Solution First compute the expectation of ℓ_{i+1} from Equation 8.2:

$$\langle \ell_{i+1} \rangle = \sum_{\ell_i} \mathcal{P}(\ell_i) \left[(\ell_i + 1)(\Delta t)\beta_\mathrm{s} + (\ell_i - 1)(\Delta t)k_\varnothing \ell_i + \ell_i\big(1 - \Delta t(\beta_\mathrm{s} + k_\varnothing \ell_i)\big) \right].$$

Now subtract $\langle \ell_i \rangle$ from both sides and divide by Δt, to find

$$\frac{\langle \ell_{i+1} \rangle - \langle \ell_i \rangle}{\Delta t} = \beta_\mathrm{s} - k_\varnothing \langle \ell_i \rangle.$$

Suppose that the original distribution has small relative standard deviation. Because ℓ is large, and the spread in its distribution only increases by less than 1 unit in a time step (Equation 8.2), the new distribution will also be sharply peaked. So we may drop the expectation symbols, recovering Equation 8.4.

To solve Equation 8.4, first notice that it has a steady state when the population ℓ equals $\ell_* = \beta_\mathrm{s}/k_\varnothing$. This makes us suspect that the equation might look simpler if we change variables from ℓ to $x = \ell - \ell_*$, and indeed it becomes $\mathrm{d}x/\mathrm{d}t = -k_\varnothing x$, whose solution is $x(t) = B\mathrm{e}^{-k_\varnothing t}$ for any constant B. Choosing B to ensure $\ell_\mathrm{ini} = 0$ (initially there are no X molecules) yields the particular solution

$$\ell(t) = (\beta_\mathrm{s}/k_\varnothing)\big(1 - \mathrm{e}^{-k_\varnothing t}\big). \tag{8.5}$$

[8] We previously met Equation 8.4 in the context of virus dynamics (Chapter 1).

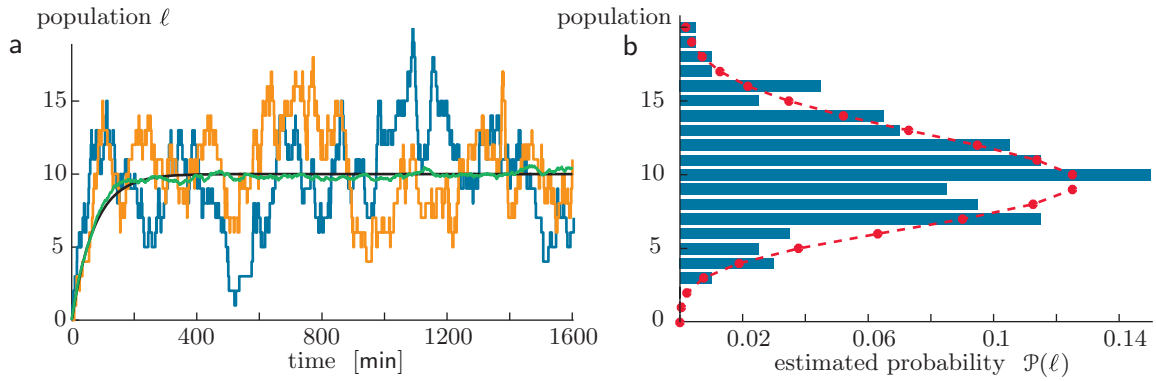


Figure 8.3 [Computer simulations.] **Behavior of a birth-death process.** (a) The *orange and blue traces* show two simulated time series (see Your Turns 8A–8B(a)). The *green trace* shows, at each time, the sample mean of the population ℓ over 200 such instances (see Problem 8.2). The *black curve* shows the corresponding solution in the continuous, deterministic approximation (Equation 8.4). (b) After the system comes to steady state, there is a broad distribution of ℓ values across instances (*bars*). The *red dots* show the Poisson distribution with $\mu = \beta_s/k_\emptyset$ for comparison (see Your Turn 8C).

Thus, initially the number of X molecules rises linearly with time, but then it levels off (**saturates**) as the clearance reaction speeds up, until a steady state is reached at $\ell(t \to \infty) = \ell_*$. The black curve in Figure 8.3a shows this solution.

8.3.3 The Gillespie algorithm

To get beyond the continuous, deterministic approximation, recall one of the lessons of the Luria-Delbrück experiment (Section 4.4, page 81): It is sometimes easier to *simulate* a random system than to derive analytic results. We can estimate whatever probabilities we wish to predict by running the simulation many times and making histograms of the quantities of interest.

Idea 8.3 suggests an approach to simulating the birth-death process, by modifying our simulation of the compound Poisson process in Section 7.6.2 (page 168). Suppose that ℓ has a known value at time zero. Then,

1. Draw the first waiting time $t_{w,1}$ from the Exponential distribution with rate $\beta_{tot} = \beta_s + k_\emptyset \ell$.
2. Next, determine which reaction happened at that time by drawing from a Bernoulli trial distribution with probability ξ, where[9]

$$\xi = \frac{(\Delta t)\beta_s}{(\Delta t)\beta_s + (\Delta t)k_\emptyset \ell} = \frac{\beta_s}{\beta_s + k_\emptyset \ell}. \tag{8.6}$$

   The probability to increase ℓ is ξ; that to decrease ℓ is $1 - \xi$. The quantities ξ and $1 - \xi$ are sometimes called the "relative propensities" of the two reactions.
3. Update ℓ by adding or subtracting 1, depending on the outcome of the Bernoulli trial.
4. Repeat.

Steps **1–4** are a simplified version of an algorithm proposed by D. Gillespie. They amount to simulating a slightly different compound Poisson process at every time step, because

[9] One way to derive Equation 8.6 is to find the conditional probability $\mathcal{P}(\ell \text{ increases} \,|\, \ell \text{ changes})$ from Equation 8.2.

both the overall rate and ξ depend on ℓ, which itself depends on the prior history of the simulated system. This dependence is quite limited, however: Knowing the state at one time determines all the probabilities for the next step (and hence all subsequent steps). That is, *the Gillespie algorithm is a method for simulating general Markov processes,* including the birth-death process and other chemical reaction networks.

The algorithm just outlined yields a set of waiting times $\{t_{w,\alpha}\}$, which we can convert to absolute times by forming cumulative sums: $t_\alpha = t_{w,1} + \cdots + t_{w,\alpha}$. It also yields a set of increments $\{\Delta\ell_\alpha\}$, each equal to ± 1, which we convert to absolute numbers in the same way: $\ell_\alpha = \ell_{ini} + \Delta\ell_1 + \cdots + \Delta\ell_\alpha$. Figure 8.3a shows a typical result, and compares it with the behavior of the continuous, deterministic approximation.

Your Turn 8A

Implement the algorithm just outlined on a computer: Write a function that accepts two input arguments `lini` and `T`, and generates two output vectors `ts` and `ls`. The argument `lini` is the initial number of molecules of X. `T` is the total time to simulate, in minutes. `ts` is the list of t_α's, and `ls` is the list of the corresponding ℓ_α's just after each of the transition times listed in `ts`. Assume that $\beta_s = 0.15/\text{min}$ and $k_\varnothing = 0.014/\text{min}$.

Your Turn 8B

a. Write a "wrapper" program that calls the function you wrote in Your Turn 8A with `lini` $= 0$ and `T` $= 1600$, then plots the resulting `ls` versus the `ts`. Run it a few times, plot the results, and comment on what you see.
b. Repeat with faster synthesis, $\beta_s = 1.5/\text{min}$, but the same clearance rate constant $k_\varnothing = 0.014/\text{min}$. Compare and contrast your result with (a).

Your answer to Your Turn 8B will include a graph similar to Figure 8.3a. It shows molecule number ℓ saturating, as expected, but still it is very different from the corresponding solution in the continuous, deterministic approximation.[10]

The Gillespie algorithm can be extended to handle cases with more than two reactions. At any time, we find the rates for all available reactions, sum them, and draw a waiting time from an appropriate Exponential distribution (step **1** on page 185). Then we find the list of all relative propensities, analogous to Equation 8.6. By definition, these numbers sum to 1, so they define a discrete probability distribution. We select which reaction occurred by drawing from this distribution;[11] then we accumulate all the changes at each time step to find the time course of ℓ.

8.3.4 The birth-death process undergoes fluctuations in its steady state

Figure 8.3 shows that the "steady" (late-time) state of the birth-death process can actually be pretty lively. No matter how long we wait, there is always a finite spread of ℓ values. In fact,

> *The steady-state population in the birth-death process is Poisson distributed, with expectation $\beta_s/k_\varnothing$.* (8.7)

[10] You'll find a connection between these two approaches in Problem 8.2.
[11] See the method in Section 4.2.5 (page 73).

Your Turn 8C

Continue Your Turns 8A–8B: Equation 8.5 suggests that the birth-death process will have come to its steady state at the end of $T = 300$ min. Histogram the distribution of final values ℓ_T across 150 trials. What further steps could you take to confirm Idea 8.7?

Despite the fluctuation, the birth-death process exhibits a bit more self-discipline than the original random walk, which never settles down to any steady state (the spread of x values grows without limit, Problem 8.1). To understand the distinction, remember that in the birth-death process there is a "hard wall" at $\ell = 0$; if the system approaches that point, it gets "repelled" by the imbalance between synthesis and clearance. Likewise, although there is no upper bound on ℓ, nevertheless if the system wanders to large ℓ values it gets "pulled back," by an imbalance in the opposite sense.

Idea 8.7 has an implication that will be important later: Because the Poisson distribution's relative standard deviation[12] is $\mu^{-1/2}$, we see that the steady-state population of a molecule will be close to its value calculated with the continuous, deterministic approximation, if that value is large. Indeed, you may have noticed a stronger result in your solution to Your Turn 8B:

The continuous, deterministic approximation becomes accurate when molecule numbers are high. (8.8)

T_2 *Section 8.3.4′ (page 195) gives an analytic derivation of Idea 8.7.*

8.4 Gene Expression

Cells create themselves by metabolizing food and making proteins, lipids, and other biomolecules. The basic synthetic mechanism is shown in Figure 8.4: DNA is **transcribed** into a messenger RNA (mRNA) molecule by an enzyme called **RNA polymerase**. Next, the resulting **transcript** is **translated** into a chain of amino acids by another enzyme complex, called the **ribosome**. The chain of amino acids then folds itself into a functioning protein (the **gene product**). The entire process is called **gene expression**. If we create a DNA sequence with two protein-coding sequences next to each other, the polymerase will generate a single mRNA containing both; translation will then create a single amino acid chain, which can fold into a combined **fusion protein**, with two domains corresponding to the two protein sequences, covalently linked into a single object.

Enzymes are themselves proteins (or complexes of protein with RNA or other cofactors). And other complex molecules, such as lipids, are in turn synthesized by enzymes. Thus, gene expression lies at the heart of all cellular processes.

T_2 *Section 8.4′ (page 197) mentions some finer points about gene expression.*

8.4.1 Exact mRNA populations can be monitored in living cells

Each step in gene expression is a biochemical reaction, and hence subject to randomness. For example, Section 8.3 suggested that it would be reasonable to model the inventory of mRNA from any particular gene via the birth-death process represented symbolically in Figure 8.2. I. Golding and coauthors tested this hypothesis in the bacterium *Escherichia coli*,

[12] See Equation 4.7 (page 78).

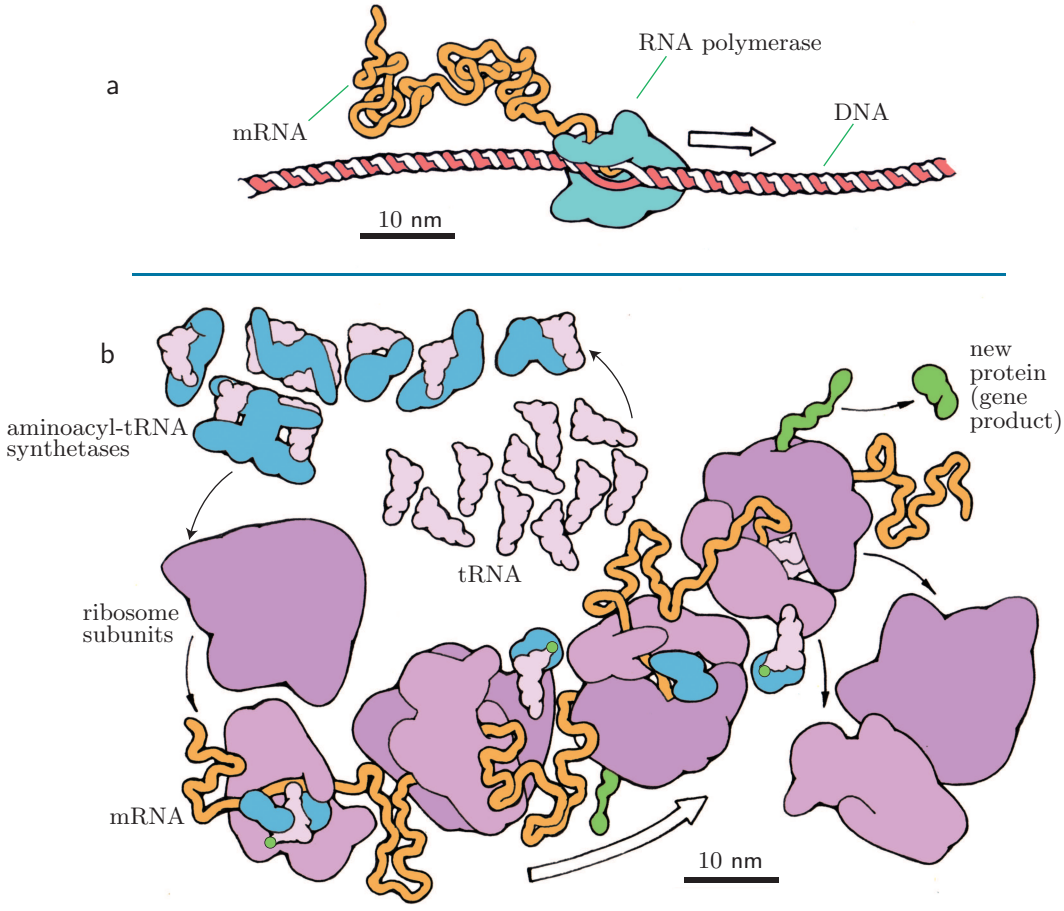


Figure 8.4 [Artist's reconstructions based on structural data.] **Transcription and translation.** (a) Transcription of DNA to messenger RNA by RNA polymerase, a processive enzyme. The polymerase reads the DNA as it walks along it, synthesizing a messenger RNA transcript as it moves. (b) The information in messenger RNA is translated into a sequence of amino acids making up a new protein by the combined action of over 50 molecular machines. In particular, aminoacyl-tRNA synthetases supply transfer RNAs, each loaded with an amino acid, to the ribosomes, which construct the new protein as they read the messenger RNA. [Courtesy David S Goodsell.]

using an approach pioneered by R. Singer. To do so, they needed a way to count the actual number of mRNA molecules in living cells, in real time.

In order to make the mRNA molecules visible, the experimenters created a cell line with an artificially designed gene. The gene coded for a gene product as usual (a red fluorescent protein), but it also had a long, noncoding part, containing 96 copies of a binding sequence. When the gene was transcribed, each copy of the binding sequence folded up to form a binding site for a protein called MS2 (Figure 8.5a). Elsewhere on the genome, the experimenters inserted another gene, for a fusion protein: One domain was a green fluorescent protein (GFP); the other coded for MS2. Thus, shortly after each transcript was produced, it began to glow brightly, having bound dozens of GFP molecules (Figure 8.5b). For each cell studied, the experimenters computed the total fluorescence intensities of all the green

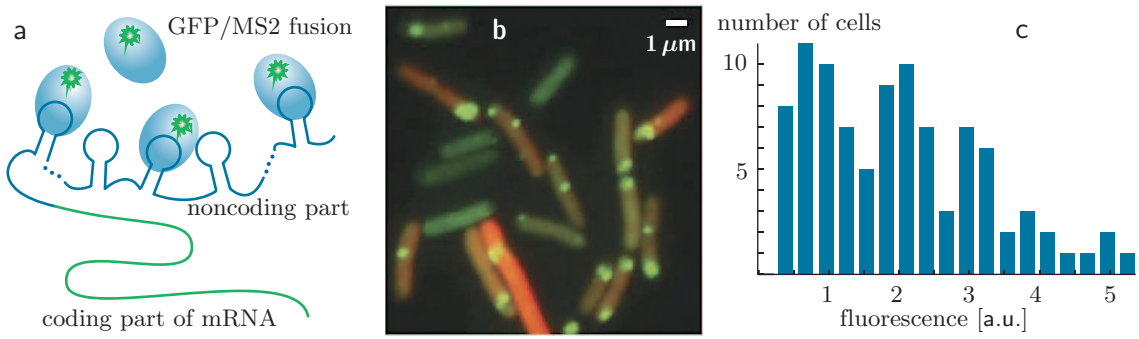


Figure 8.5 Quantification of mRNA levels in individual cells. (a) [Sketch.] Cartoon showing a messenger RNA molecule. The mRNA was designed to fold, creating multiple binding sites for a fusion protein that includes a green fluorescent protein (GFP) domain. (b) [Fluorescence micrograph.] Several individual, living bacteria, visualized via their fluorescence. Each *bright green spot* shows the location of one or more mRNA molecules labeled by GFP. The *red* color indicates red fluorescent protein (RFP), arising from translation of the coding part of the mRNA. (c) [Experimental data.] For each cell, the green fluorescence signal was quantified by finding the total photon arrival rate coming from the green spots only (minus the whole cell's diffuse background). The resulting histogram shows well-separated peaks, corresponding to cells with 1, 2, . . . mRNA molecules (compare Figure 7.7 on page 165). On the horizontal axis, the observed fluorescence intensities have all been rescaled by a common value, chosen to place the first peak near the value 1. Then *all* the peaks were found to occur near integer multiples of that value. This calibration let the experimenters infer the absolute number of mRNA molecules in any cell. [From Golding et al., 2005.]

spots seen in the microscope. A histogram of the observed values of this quantity showed a chain of evenly spaced peaks (Figure 8.5c), consistent with the expectation that each peak represents an integer multiple of the lowest one.[13] Thus, to count the mRNA copies in a cell, it sufficed to measure that cell's fluorescence intensity and identify the corresponding peak.

The experimenters wanted to test the hypothesis that mRNA population dynamics reflects a simple birth-death process. To do so, they noted that such a process is specified by just two parameters, but makes more than two predictions. They determined the parameter values (β_s and $k_\varnothing$) by fitting some of the predictions to experimental data, then checked other predictions.

One such experiment involved suddenly switching on ("inducing") the production of mRNA.[14] In a birth-death process, the number of mRNA molecules, ℓ, averaged over many independent trials, follows the saturating time course given by Equation 8.5.[15] This prediction of the model yielded a reasonable-looking fit to the data. For example, the red trace in Figure 8.6a shows the prediction of the birth-death model with the values $\beta_s \approx 0.15/\text{min}$ and $k_\varnothing \approx 0.014/\text{min}$.

8.4.2 mRNA is produced in bursts of transcription

Based on averages over many cells, then, it may appear that the simple birth-death model is adequate to describe gene expression in *E. coli*. But the ability to count *single molecules in individual cells* gave Golding and coauthors the opportunity to apply a more

[13] The intensity of fluorescence per mRNA molecule had some spread, because each mRNA had a variable number of fluorescent proteins bound to it. Nevertheless, Figure 8.5c shows that this variation did not obscure the peaks in the histogram.

[14] Chapter 9 discusses gene switching in greater detail.

[15] See Problem 8.2.

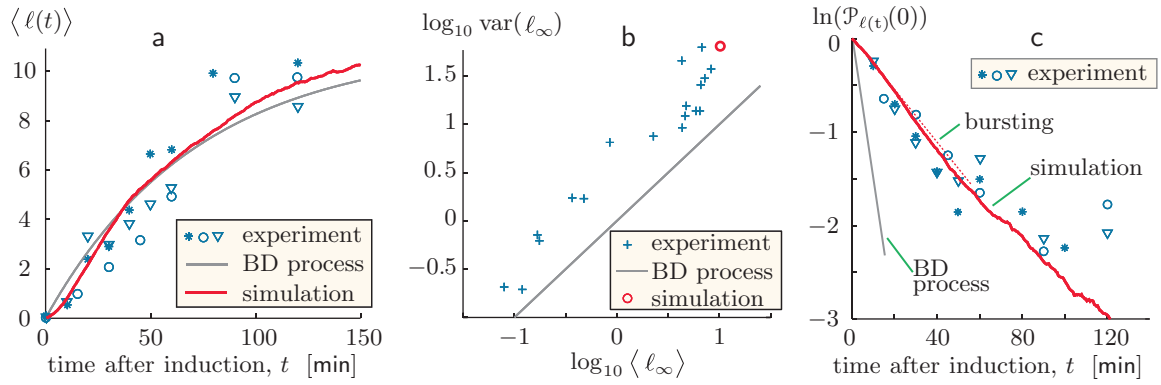


Figure 8.6 [Experimental data with fits.] **Indirect evidence for transcriptional bursting.** (a) *Symbols:* The number of mRNA transcripts in a cell, $\ell(t)$, averaged over 50 or more cells in each of three separate experiments. All of the cells were induced to begin gene expression at a common time, leading to behavior qualitatively like that shown in Figure 8.3a. The *gray curve* shows a fit of the birth-death (BD) process (Equation 8.5, page 184) to data, determining the apparent synthesis rate $\beta_s \approx 0.15/\text{min}$ and clearance rate constant $k_\varnothing \approx 0.014/\text{min}$. The *red trace* shows the corresponding result from a computer simulation of the bursting model discussed in the text (see also Section 8.4.2′b, page 198). (b) Variance of mRNA population versus sample mean, in steady state. *Crosses:* Many experiments were done, each with the gene turned "on" to different extents. This log-log plot of the data shows that they fall roughly on a line of slope 1, indicating that the Fano factor $(\text{var}\,\ell)/\langle\ell\rangle$ is roughly a constant. The simple birth-death process predicts that this constant is equal to 1 (*gray line*), but the data instead give the value ≈ 5. The *red circle* shows the result of the bursting model, which is consistent with the experimental data. (c) Semilog plot of the fraction of observed cells that have zero copies of mRNA versus elapsed time. *Symbols* show data from the same experiments as in (a). *Gray line:* The birth-death process predicts that initially $\mathcal{P}_{\ell(t)}(0)$ falls with time as $\exp(-\beta_s t)$ (see Problem 8.4). *Dotted line:* The experimental data instead yield initial slope $-0.028/\text{min}$. *Red trace:* Computer simulation of the bursting model. [Data from Golding et al., 2005; see Dataset 12.]

stringent test than the one shown in Figure 8.6a. First, we know that the steady-state mRNA counts are Poisson distributed in the birth-death model,[16] and hence that $\text{var}\,\ell_\infty = \langle\ell_\infty\rangle$. Figure 8.6b shows that the ratio of these two quantities (sample variance/sample mean) really is approximately constant over a wide range of conditions. However, contrary to the prediction of the birth-death model, the value of this ratio (called the **Fano factor**) does *not* equal 1; in this experiment it was approximately 5. The birth-death model also predicts that the fraction of cells with zero copies of the mRNA should initially decrease exponentially with time, as $e^{-\beta_s t}$ (see Problem 8.4). Figure 8.6c shows that this prediction, too, was falsified in the experiment.

These failures of the simplest birth-death model led the experimenters to propose and test a modified hypothesis:

> *Gene transcription in bacteria is a* **bursting process**, *in which the gene makes spontaneous transitions between active and inactive states at mean rates* β_{start} *and* β_{stop}. *Only the active state can be transcribed, leading to bursts of mRNA production interspersed with quiet periods.* (8.9)

More explicitly, β_{start} is the probability per unit time that the gene, initially in the "off" state, will switch to the "on" state. It defines a mean waiting time $\langle t_{\text{w,start}}\rangle = (\beta_{\text{start}})^{-1}$, and similarly for β_{stop} and $t_{\text{w,stop}}$.

[16] See Idea 8.7 (page 186).

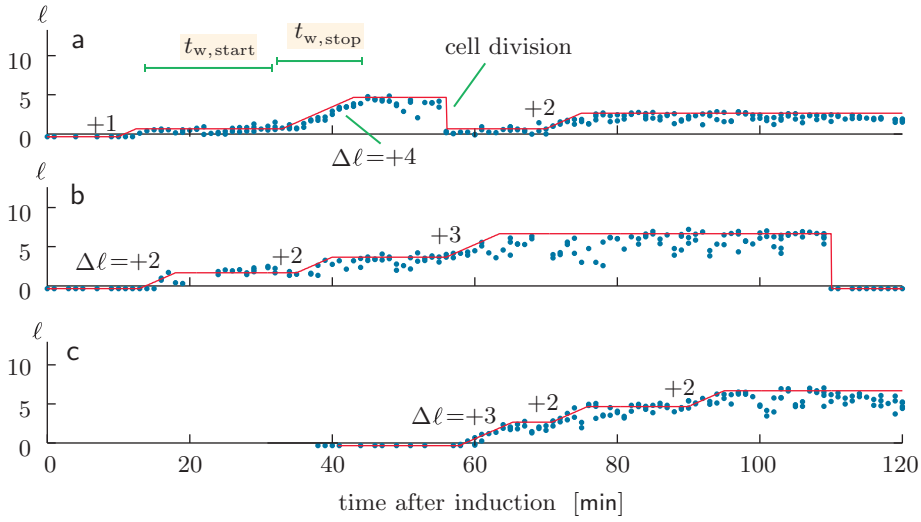


Figure 8.7 [Experimental data.] **Direct evidence for bursting in bacterial gene transcription.** The panels show time courses of ℓ, the population level of a labeled mRNA transcript, in three typical cells. (a) *Dots:* Estimated values of ℓ for one cell. This number occasionally steps downward as the cell divides, because thereafter only one of the two daughter cells' mRNA counts is shown. In this instance, cell division segregated only one of the total five transcripts into the daughter cell selected for further observation (the other four went into the other daughter). The data show episodes when ℓ holds steady (*horizontal segments*), interspersed with episodes of roughly constant production rate (*sloping segments*). The *red line* is an idealization of this behavior. Typical waiting times for transitions to the "on" ($t_{w,start}$) or "off" state ($t_{w,stop}$) are shown, along with the increment $\Delta\ell$ in mRNA population during one episode of transcriptional bursting. (b,c) Observations of two additional individual cells. [Data from Golding et al., 2005.]

The intuition behind the bursting model runs roughly as follows:

1. Each episode of gene activation leads to the synthesis of a variable number of transcripts, with some average value $m = \langle \Delta\ell \rangle$. We can roughly capture this behavior by imagining that each burst contains *exactly* m transcripts. Then the variance of ℓ will be increased by a factor of m^2 relative to an ordinary birth-death process, whereas the expectation will only increase by m. Thus, the Fano factor is larger than 1 in the bursting model, as seen in the data (Figure 8.6b).
2. In the bursting model, the cell leaves the state $\ell = 0$ almost immediately after the gene makes its first transition to the "on" state. Thus, the probability per unit time to exit the $\ell = 0$ state is given by β_{start}. But the initial growth rate of $\langle \ell(t) \rangle$ is given by $\beta_{start} m$, which is a larger number. So the observed initial slopes in panels (a,c) of Figure 8.6 need not be equal, as indeed they are not.

The experimenters tested the bursting hypothesis directly by looking at the time courses of mRNA population in individual cells. Figure 8.7 shows some typical time courses of ℓ. Indeed, in each case the cell showed episodes with no mRNA synthesis, alternating with others when the mRNA population grows at an approximately constant rate.[17] The episodes were of variable duration, so the authors then tabulated the waiting times to transition from

[17] T_2 The mRNA population in any one cell also dropped suddenly each time that cell divided, because the molecules were partitioned between two new daughter cells, only one of which was followed further. Section 8.4.2′a (page 197) discusses the role of cell division.

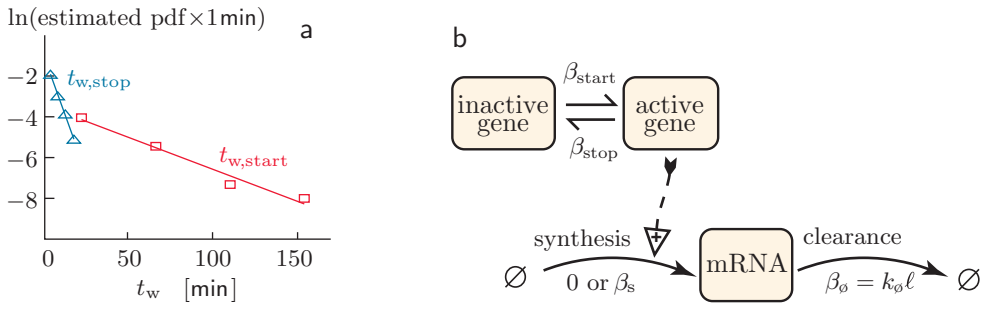


Figure 8.8 Model for transcriptional bursting. (a) [Experimental data.] Semilog plot of the estimated probability density for the durations $t_{w,stop}$ of transcription bursts (waiting times to turn off) and of waiting times $t_{w,start}$ to turn on. Fitting the data yielded $\langle t_{w,stop} \rangle \approx 6$ min and $\langle t_{w,start} \rangle \approx 37$ min. [Data from Golding et al., 2005.] (b) [Network diagram.] The bursting hypothesis proposes a modified birth-death process, in which a gene spontaneously transitions between active and inactive states with fixed probabilities per unit time (compare Figure 8.2b on page 182). The *boxes* on the top represent the populations of the gene in its two states (in this case, either 1 or 0). *Solid arrows* between these boxes represent processes that increase one population at the expense of the other. The *dashed arrow* represents an interaction in which one species (here, the gene in its active state) influences the rate of a process (here, the synthesis of mRNA).

"on" to "off" and vice versa, and made separate histograms for each. In the bursting model, when the gene is "on" the probability per unit time to switch off is a constant, β_{stop}. Thus, the model predicts that the waiting times $t_{w,stop}$ will be Exponentially distributed[18] with expectation $(\beta_{stop})^{-1}$, and indeed such behavior was observed (Figure 8.8a). The probability per unit time to switch "on," β_{start}, was similarly found by fitting the distribution of $t_{w,start}$.

The bursting model can be summarized by a network diagram; see Figure 8.8b.

Quantitative checks

The experimental data overconstrain the parameters of the bursting model, so it makes falsifiable predictions.

First, fitting the red data in Figure 8.8a to an Exponential distribution gives $\beta_{start} \approx 1/(37\,\text{min})$. Point **2** above argued that $\ln\left(\mathcal{P}_{\ell(t)}(0)\right)$ initially falls as $-\beta_{start}\,t$, and indeed the data in Figure 8.6c do show this behavior, with the same value of β_{start} as was found directly in Figure 8.8a.

Second, Figure 8.6b gives the burst size $m \approx 5$. Point **1** above argued that, if bursts of size m are generated with probability per unit time β_{start}, then we can get the expected number of transcripts by modifying Equation 8.5 to

$$\langle \ell(t) \rangle = \frac{m\beta_{start}}{k_\emptyset} \left(1 - e^{-k_\emptyset t}\right).$$

The only remaining free fitting parameter in this function is $k_\emptyset$. That is, a single choice for this parameter's value predicts the *entire curve* appearing in Figure 8.6a. The figure shows that, indeed, the value $k_\emptyset = 0.014/\text{min}$ gives a function that fits the data.[19] Thus,

[18] See Idea 7.5 (page 159).

[19] In fact, Section 8.4.2′a (page 197) will argue that the value of $k_\emptyset$ should be determined by the cells' doubling time, further overconstraining the model's parameters.

the transcriptional bursting hypothesis, unlike the simple birth-death process, can roughly explain all of the data in the experiment.

T_2 *This section argued heuristically that it is possible to reconcile all the observations in Figures 8.6a–c and 8.8a in a single model. A more careful analysis, however, requires computer simulation to make testable predictions. Section 8.4.2′b (page 198) describes such a stochastic simulation.*

8.4.3 Perspective

Golding and coauthors followed a systematic strategy for learning more about gene expression:

- Instead of studying the full complex process, they focused on just one step, mRNA transcription.
- They found an experimental technique that let them determine absolute numbers of mRNA, in living cells, in real time.
- They explored the simplest physical model (the birth-death process) based on known actors and the general behavior of molecules in cells.
- They found contact between the experiment and the model by examining reduced statistics, such as the time course of the average of the copy number, its steady-state variance, and the probability that it equals zero. Establishing that contact involved making predictions from the model.
- Comparing these predictions to experimental data was sufficient to rule out the simplest model, so they advanced to the next-simplest one, introducing a new state variable (gene on or off) reminiscent of many other discrete conformational states known elsewhere in molecular biology.
- Although the new model is surely not a complete description, it did make falsifiable predictions that could be more directly tested by experiments designed for that purpose (Figure 8.7), and it survived comparison to the resulting data.

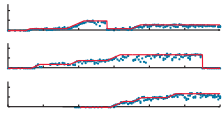


Figure 8.7 (page 191)

Many other groups subsequently documented transcriptional bursting in a wide variety of organisms, including single-cell eukaryotes and even mammals. Even within a single organism, however, some genes are observed to burst while others are not. That is, *transcriptional bursting is a controlled feature* of gene expression, at least in eukaryotes.

Several mechanisms have been proposed that may underlie bursting. Most likely, the complete picture is not simple. But already, this chapter has shown how targeted experiments and modeling succeeded in *characterizing* transcription of a particular gene in a significantly more detailed way than had been previously possible. More recent experiments have also begun to document more subtle aspects of bursting, for example, correlations between transcription bursts of different genes.

8.4.4 Vista: Randomness in protein production

Transcription is just one of many essential cell activities. The general method of fluorescence tagging has also been used to characterize the randomness inherent in protein translation, and in the overall levels of protein populations in cells. In part, protein level fluctuations track mRNA levels, but their randomness can also be increased (for example, by Poisson noise from translation) or suppressed (by averaging over the many mRNA copies created by a single gene).

THE BIG PICTURE

This chapter began by studying random walks in space, such as the trajectories of small diffusing objects in fluid suspension. We then generalized our framework from motion in ordinary space to chemical reactions, which we modeled as random walks on a *state space*. We got some experience handling probability distributions over all possible histories of such systems, and their most commonly used reduced forms. Analogously to our experience deducing a hidden step in myosin-V stepping (Section 7.5.1), we were able to deduce a hidden state transition, leading to the discovery of bursting in bacterial gene expression. Cells must either exploit, learn to live with, or overcome such randomness in their basic processes.

However, we also found situations in which the randomness of gene expression had little effect on the dynamics of mRNA levels, because the overall inventory of mRNA was high.[20] Chapters 9–11 will make this continuous, deterministic approximation as we push forward our study of cellular control networks.

KEY FORMULAS

- *Diffusion:* A small particle suspended in fluid will move in a random walk, due to its thermal motion in the fluid. The mean-square deviation of the particle's displacement, after many steps, is proportional to the elapsed time.
- *Birth-death process:* Let β_s be the synthesis rate, and $k_\emptyset$ the degradation rate constant, for a birth-death process. In the continuous, deterministic approximation the population ℓ of a species X follows $d\ell/dt = \beta_s - \ell k_\emptyset$. One solution to this equation is the one that starts with $\ell(0) = 0$: $\ell(t) = (\beta_s/k_\emptyset)\left(1 - e^{-k_\emptyset t}\right)$.
- *Stochastic simulation:* The relative propensities for a two-reaction Gillespie algorithm, with reaction rates β_1 and β_2, are $\xi = \beta_1/(\beta_1 + \beta_2)$ and $(1 - \xi)$. (See Equation 8.6.)
- T_2 *Master equation:*

$$\frac{\mathcal{P}_{\ell_{i+1}}(\ell) - \mathcal{P}_{\ell_i}(\ell)}{\Delta t} = \beta_s\left(\mathcal{P}_{\ell_i}(\ell - 1) - \mathcal{P}_{\ell_i}(\ell)\right) + k_\emptyset\left((\ell + 1)\mathcal{P}_{\ell_i}(\ell + 1) - \mathcal{P}_{\ell_i}(\ell)\right).$$

FURTHER READING

Semipopular:
Hoagland & Dodson, 1995.

Intermediate:
Klipp et al., 2009, chapt. 7; Otto & Day, 2007; Wilkinson, 2006.
mRNA dynamics: Phillips et al., 2012, chapt. 19.
T_2 Master (or Smoluchowski) equations: Nelson, 2014, chapt. 10; Schiessel, 2013, chapt. 5.

Technical:
Gillespie algorithm: Gillespie, 2007; Ingalls, 2013.
Bursting in prokaryotes: Golding et al., 2005; Paulsson, 2005; Taniguchi et al., 2010.
Transcriptional bursting in higher organisms: Raj & van Oudenaarden, 2009; Suter et al., 2011; Zenklusen et al., 2008.

[20] See Idea 8.8 (page 187).

T_2 **Track 2**

8.3.4′ The master equation

We have seen that in the birth-death process, the distribution of system states ℓ is Poisson.[21] We can confirm this observation by inventing and solving the system's "master equation." Similar formulas arise in many contexts, where they are called by other names such as "diffusion," "Fokker-Planck," or "Smoluchowski" equations.

Any random process is defined on a big sample space consisting of all possible *histories* of the state. Treating time as discrete, this means that the sample space consists of sequences $\ell_1, \ldots, \ell_j, \ldots$, where ℓ_j is the population at time t_j. As usual, we'll take $t_i = (\Delta t)i$ and recover the continuous-time version later, by taking the limit of small Δt.

The probability of a particular history, $\mathcal{P}(\ell_1, \ldots)$, is complicated, a joint distribution of many variables. We will be interested in reduced forms of this distribution, for example, $\mathcal{P}_{\ell_i}(\ell)$, the marginal distribution for there to be ℓ molecules of species X at time $(\Delta t)i$, regardless of what happens before or after that time. The Markov property implies that this probability is completely determined if we know that the system is in a definite state at time $i-1$, so we begin by assuming that.

Imagine making a large number N_{tot} of draws from this random process, always starting the system at time 0 with the same number of molecules, ℓ_{ini} (see Figure 8.9a). That is, we suppose that $\mathcal{P}_{\ell_0}(\ell) = 1$ if $\ell = \ell_{\text{ini}}$, and zero otherwise. We can summarize the notation:

ℓ_i	number of molecules at time $t_i = (\Delta t)i$, a random variable
ℓ_{ini}	initial value, a constant
N_{tot}	number of systems being observed
β_s	mean rate for synthesis
$k_\emptyset$	clearance rate constant

Each of these quantities is a constant, except the ℓ_i, each of which is a random variable.

Equivalently, we can express ℓ_{ini} in terms of ℓ in each case, and compare the result with the original distribution:

$$\frac{\mathcal{P}_{\ell_1}(\ell) - \mathcal{P}_{\ell_0}(\ell)}{\Delta t} = \begin{cases} \beta_s & \text{if } \ell = \ell_{\text{ini}} + 1; \\ (\ell+1)k_\emptyset & \text{if } \ell = \ell_{\text{ini}} - 1; \\ -(\beta_s + k_\emptyset \ell) & \text{if } \ell = \ell_{\text{ini}}; \\ 0 & \text{otherwise.} \end{cases} \tag{8.10}$$

Equation 8.10 is applicable to the special case shown in Figure 8.9a.

Next, suppose that initially a fraction q of the N_{tot} systems started out with ℓ_{ini} molecules, but the other $1-q$ instead start with some other value ℓ'_{ini} (see Figure 8.9b). Thus, the initial distribution is nonzero at just two values of ℓ, so on the next time step the distribution evolves to one that is nonzero on just those two values and their four flanking values, and

[21] See Idea 8.7 (page 186).

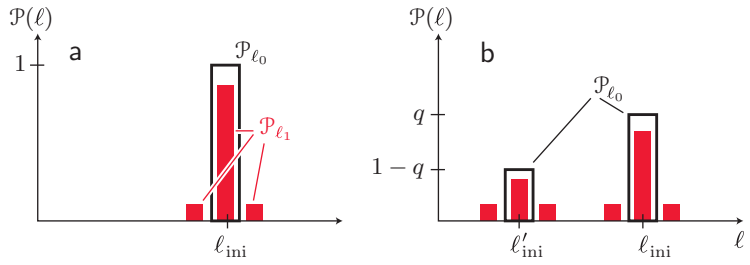


Figure 8.9 [Sketch graphs.] **Time evolution in a birth-death process.** (a) Suppose that a collection of identical systems all have the same starting value of ℓ (*black*). Each of the systems evolves in the next time slot to give a distribution with some spread (*red*). (b) This panel represents an initial distribution of states with *two* values of ℓ. This distribution evolves into one with nonzero probability at *six* values of ℓ.

so on. The six cases that must be considered can all be elegantly summarized as a single formula, called the **master equation**:

$$\frac{\mathcal{P}_{\ell_1}(\ell)-\mathcal{P}_{\ell_0}(\ell)}{\Delta t}=\beta_s\big(\mathcal{P}_{\ell_0}(\ell-1)-\mathcal{P}_{\ell_0}(\ell)\big)+k_\varnothing\big((\ell+1)\mathcal{P}_{\ell_0}(\ell+1)-\ell\mathcal{P}_{\ell_0}(\ell)\big). \quad (8.11)$$

The master equation is actually a chain of many linked equations, one for every allowed value of ℓ. Remarkably, it is no longer necessary to itemize particular cases, as was done in Equation 8.10; this is now accomplished by expressing the right-hand side of Equation 8.11 in terms of the initial distribution $\mathcal{P}_{\ell_0}$.

Example Derive Equation 8.11. Show that it also applies to the case where the initial distribution $\mathcal{P}_{\ell_0}(\ell)$ is arbitrary (not necessarily peaked at just one or two values of ℓ).

Solution As before, it is a bit easier to start by thinking of a finite set of N_{tot} specific trials. Of these, initially about $N_{*,\ell}=N_{\text{tot}}\mathcal{P}_{\ell_0}(\ell)$ had ℓ copies of X. (These statements become exact in the limit of large N_{tot}.)

For each value of ℓ, at the next time slot about $N_{*,\ell-1}(\Delta t)\beta_s$ get added to bin ℓ (and removed from bin $(\ell-1)$).

For each value of ℓ, at the next time slot another $N_{*,\ell+1}(\Delta t)k_\varnothing(\ell+1)$ get added to bin ℓ (and removed from bin $(\ell+1)$).

For each value of ℓ, at the next time slot about $N_{*,\ell}(\Delta t)(\beta_s+k_\varnothing\ell)$ get removed from bin ℓ (and added to other bins).

Altogether, then, the number of trials with exactly ℓ copies changes from $N_{*,\ell}$ at time 0 to

$$N_\ell=N_{*,\ell}+\Delta t\big(\beta_s N_{*,\ell-1}+k_\varnothing(\ell+1)N_{*,\ell+1}-(\beta_s+k_\varnothing\ell)N_{*,\ell}\big).$$

Dividing by $(\Delta t)N_{\text{tot}}$ gives the master equation. (Note, however, that for $\ell=0$ the equation must be modified by omitting its first term.)

The right side of Equation 8.11 consists of a pair of terms for each reaction. In each pair, the positive term represents influx into the state populated by a reaction; the negative

term represents the corresponding departures from the state that is depopulated by that reaction.

Our goal was to check Idea 8.7 (page 186), so we now seek a steady-state solution to the master equation. Set the left side of Equation 8.11 equal to zero, and substitute a trial solution of the form $\mathcal{P}_\infty(\ell) = \mathrm{e}^{-\mu}\mu^\ell/(\ell!)$.

Your Turn 8D

Confirm that this trial solution works, and find the value of the parameter μ.

The master equation lets us calculate other experimentally observable quantities as well, for example, the *correlation* between fluctuations at different times. To obtain its continuous-time version, we just note that the left side of Equation 8.11 becomes a derivative in the limit $\Delta t \to 0$. In this limit, it becomes a large set of coupled first-order ordinary differential equations, one for each value of ℓ. (If ℓ is a continuous variable, then the master equation becomes a partial differential equation.)

T2 **Track 2**

8.4′ More about gene expression

1. In eukaryotes, various "editing" modifications also intervene between transcription and translation.
2. Folding may also require the assistance of "chaperones," and may involve the introduction of "cofactors" (extra molecules that are not amino acids). An example is the cofactor retinal, added to an opsin protein to make the light-sensing molecules in our eyes.
3. The gene product may be a complete protein, or just a part of a protein that involves multiple amino acid chains and cofactors.
4. To create a fusion protein, it's not enough to position two genes next to each other: We must also eliminate the first one's "stop codon," so that transcription proceeds to the second one, and ensure that the two genes share the same "reading frame."

T2 **Track 2**

8.4.2′a The role of cell division

The main text mentioned two processes that could potentially offset the increase of messenger RNA counts in cells (clearance and cell division), and tacitly assumed that both could be summarized via a single rate constant $k_\emptyset$. This is a reasonable assumption if, as discussed in the main text, mRNA molecules make random encounters with an enzyme that degrades them. But in fact, Golding and coauthors found that their fluorescently labeled mRNA constructs were rarely degraded. Instead, in this experiment cell division was the main process reducing concentration.

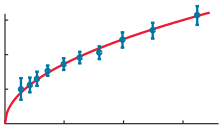


Figure 4.2 (page 73)

Upon cell division, the experimenters confirmed that each messenger RNA independently "chooses" which daughter cell it will occupy, similarly to Figure 4.2. Thus, the number passed to a particular daughter is a Binomially distributed random variable. On average,

this number is one half of the total mRNA population. The bacteria in the experiment were dividing every 50 min. Suppose that we could suddenly shut off synthesis of new mRNA molecules. After the passage of time T, then, the average number will have halved a total of $T/(50\,\text{min})$ times, reducing it by a factor of $2^{-T/(50\,\text{min})}$. Rewriting this result as $\ell(t) = \ell_{\text{ini}} \exp(-k_{\text{ø}} T)$, we find $k_{\text{ø}} = (\ln 2)/(50\,\text{min}) \approx 0.014/\text{min}$.

Making a continuous, deterministic approximation, we just found that about $k_{\text{ø}}\ell\,\mathrm{d}t$ molecules are lost in time $\mathrm{d}t$, so cell division gives rise to a "dilution" effect, similar to clearance but with the value of $k_{\text{ø}}$ given in the previous paragraph. Even if production is nonzero, we still expect that the effect of cell division can be approximated by a continuous loss at rate $k_{\text{ø}}\ell$. The main text shows that the experimental data for $\langle \ell(t) \rangle$ do roughly obey an equation with rate constant $k_{\text{ø}} \approx 0.014/\text{min}$, as was predicted above.[22]

8.4.2′b Stochastic simulation of a transcriptional bursting experiment

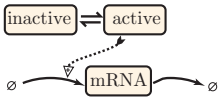


Figure 8.8b (page 192)

The main text motivated the transcriptional bursting model (represented symbolically in Figure 8.8b), then gave some predictions of the model, based on rather informal simplifications of the math. For example, cell division was approximated as a continuous, first-order process (see Section (a) above), and the variability of burst sizes was ignored. In addition, there were other real-world complications not even mentioned in the chapter:

- We have implicitly assumed that there is always exactly one copy of the gene in question in the cell. Actually, however, any given gene replicates at a particular time in the middle of a bacterial cell's division cycle. For the experiment we are studying, suppose that gene copy number doubles after about 0.3 of the cell division time, that is, after $(0.3) \times (50\,\text{min})$.
- Moreover, there may be more than one copy of the gene, even immediately after division. For the experiment we are studying, this number is about 2 (So et al., 2011). Suppose that each new copy of the gene is initially "off," and that immediately after division all copies are "off." Because the gene spends most of its time "off," these are reasonable approximations.
- Cell division does not occur precisely every 50 min; there is some randomness.

To do better than the heuristic estimates, we can incorporate every aspect of the model's formulation in a stochastic simulation, then run it many times and extract predictions for the experimentally observed quantities, for any chosen values of the model's parameters (see also So et al., 2011).

The simulation proceeds as follows. At any moment, there are state variables counting the total number of "on" and "off" copies of the gene, the number of messenger RNA molecules present in the cell, and another "clock" variable n describing progress toward division, which occurs when n reaches some preset threshold n_0. A Gillespie algorithm decides among the processes that can occur next:

1. One of the "on" copies of the gene may switch "off." The total probability per unit time for this outcome is β_{stop} times the number of "on" copies at that moment.
2. One of the "off" copies may switch "on." The total probability per unit time for this outcome is β_{start} times the number of "off" copies at that moment.

[22] You'll implement this approach to clearance in Problem 8.5. For more details about cell growth, see Section 9.4.5.

3. One of the "on" copies may create a mRNA transcript. The total probability per unit time for this outcome is a rate constant β_s times the number of "on" copies at that moment. (Note that the value of β_s needed to fit the data will not be equal to the value obtained when using the birth-death model.)
4. The "clock" variable n may increment by one unit. The probability per unit time for this outcome is $n_0/(50\,\text{min})$.

The waiting time for the next event is drawn, one of the four reaction types above is chosen according to the recipe in Section 8.3.3, and the system state is updated. Before repeating the cycle, however, the simulation checks for two situations requiring additional actions:

- If the clock variable exceeds $0.3n_0$, then the number of gene copies is doubled before proceeding (gene duplication). The new copies are assumed to be "off." No further doubling will occur prior to cell division.
- If the clock variable exceeds n_0, then the cell divides. The number of gene copies is reset to its initial value, and all are turned "off." To find the number of mRNA molecules passed on to a particular daughter cell, a random number is drawn from the Binomial distribution with $\xi = 1/2$ and M equal to the total number of molecules present.

A simulation following the procedure outlined above yielded the curves shown in Figure 8.6; see Problem 8.7.

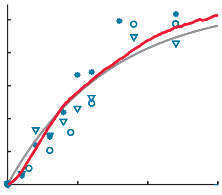


Figure 8.6a (page 190)

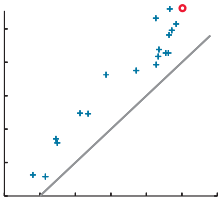


Figure 8.6b (page 190)

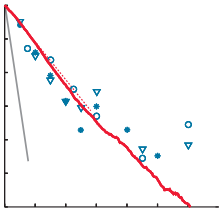


Figure 8.6c (page 190)

8.4.2′c Analytical results on the bursting process

The preceding section outlined a simulation that could be used to make predictions relevant to the experimental data shown in the text. Even more detailed information can be obtained from those data, however: Instead of finding the sample mean and variance in the steady state, one can estimate the entire probability distribution $\mathcal{P}_{\ell(t\to\infty)}$ from data (Golding et al., 2005; So et al., 2011), and compare it to the corresponding distribution found in the simulation.

If we are willing to make the idealization of treating cell division as a continuous clearance process (see Section (a) above), then there is an alternative to computer simulation: Analytic methods can also be used to predict the distribution starting from the master equation (Raj et al., 2006; Shahrezaei & Swain, 2008; Iyer-Biswas et al., 2009; Stinchcombe et al., 2012). These detailed predictions were borne out in experiments done with bacteria (Golding et al., 2005; So et al., 2011) and eukaryotes (Raj et al., 2006; Zenklusen et al., 2008).

PROBLEMS

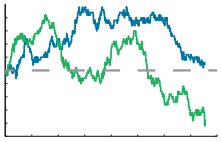


Figure 8.1 (page 181)

8.1 Random walk with random waiting times

a. Implement the strategy outlined in Section 8.2.2 to simulate a random walk in one dimension. Suppose that the steps occur in a compound Poisson process with mean rate $\beta = 1\,\mathrm{s}^{-1}$, and that each step is always of the same length $d = 1\,\mu\mathrm{m}$, but in a randomly chosen direction: $\Delta x = \pm d$ with equal probabilities for each direction. Make a graph of two typical trajectories (x versus t) with total duration $T = 400\,\mathrm{s}$, similar to Figure 8.1.

b. Run your simulation 50 times. Instead of graphing all 50 trajectories, however, just save the ending positions x_T. Then compute the sample mean and variance of these numbers. Repeat for 50 trajectories with durations 200 s, and again with 600 s.

c. Use your result in (b) to guess the complete formulas for $\langle x_T \rangle$ and $\mathrm{var}(x_T)$ as functions of d, β, and T.

d. Upgrade your simulation to two dimensions. Each step is again of length 1 μm, but in a direction that is randomly chosen with a Uniform distribution in angle. This time make a graph of x versus y. That is, don't show the time coordinate (but do join successive points by line segments).

e. T_2 An animation of the 2D walk is more informative than the picture you created in (c), so try to make one.

8.2 Average over many draws

Continuing Your Turn 8B, write a program that calls your function 150 times, always with $\ell_{\mathrm{ini}} = 0$ and for time from 0 to 300 min. At each value of time, find the average of the population over all 150 trials. Plot the time course of the averages thus found, and comment on the relation between your graph and the result of the continuous, deterministic approximation. [*Hint:* For every trial, and for every value of t from 0 to 300 min, find the step number, α, at which `ts(alpha)` first exceeds t. Then the value of ℓ after step $\alpha - 1$ is the desired position at time t.]

8.3 Burst number distribution

Consider a random process in which a gene randomly switches between "on" and "off" states, with probability per unit time β_{stop} to switch on→off and β_{start} to switch off→on. In the "off" state, the gene makes no transcripts. In the "on" state, it makes transcripts in a Poisson process with mean rate β_{s}. Simplify by assuming that both transcription and switching are sudden events (no "dead time").

Obtain analytically the expected probability distribution function for the number $\Delta\ell$ of transcript molecules created in each "on" episode, by taking these steps:

a. After an "on" episode begins, there are two kinds of event that can happen next: Either the gene switches "off," terminating the episode, or else it makes a transcript. Find the probability distribution describing which of these two outcomes happens first.

b. We can think of the events in (a) as "attempts to leave the 'on' state." Some of those attempts "succeed" (the gene switches off); others "fail" (a transcript is made and the gene stays on). The total number of transcripts made in an "on" episode, $\Delta\ell$, is the number of "failures" before the first "success." Use your answer to (a) to find the distribution of this quantity in terms of the given parameters.

c. Find the expectation of the distribution you found in (b) (the average burst size) in terms of the given parameters.

8.4 Probability of zero copies, via simulation

First work Problem 8.2. Now add a few lines to your code to tabulate the number of trials in which the number of copies, ℓ, is still zero after time t, for various values of t. Convert this result into a graph of $\ln \mathcal{P}_{\ell(t)}(0)$ versus time, and compare to the semilog plot of experimental data in Figure 8.6c.

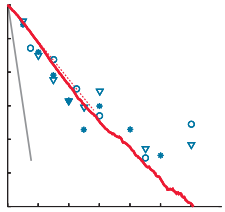


Figure 8.6c (page 190)

8.5 Simulate simplified bursting process

First work Problem 8.2. Now modify your code to implement the transcriptional bursting process (Figure 8.8b). To keep your code fairly simple, assume that (*i*) A cell contains a single gene, which transitions between "on" and "off" states. Initially the gene is "off" and there are zero copies of its mRNA. (*ii*) The cell never grows or divides, but there is a first-order clearance process[23] with rate constant $k_{\varnothing} = (\ln 2)/(50\,\text{min})$.

Take the other rates to be $\beta_{\text{start}} = 1/(37\,\text{min})$, $\beta_{\text{stop}} = 1/(6\,\text{min})$, and $\beta_{\text{s}} = 5\beta_{\text{stop}}$. Run your simulation 300 times, and make graphs of $\langle \ell(t) \rangle$ and $\ln\big(\mathcal{P}_{\ell(t)}(0)\big)$ versus time over the course of 150 minutes. Also compute the Fano factor $\text{var}(\ell_{\text{final}})/\langle \ell_{\text{final}} \rangle$, and comment.

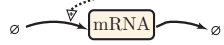


Figure 8.8b (page 192)

8.6 T_2 Probability of zero copies, via master equation

Suppose that a molecule is created (for example, a messenger RNA) in a Poisson process with mean rate β_{s}. There is no clearance process, so the population of the molecule never decreases. Initially there are zero copies, so the probability distribution for the number of molecules ℓ present at time zero is just $\mathcal{P}_{\ell(0)}(0) = 1$; all other $\mathcal{P}_{\ell(0)}(\ell)$ equal zero. Find the value of $\mathcal{P}_{\ell(t)}(0)$ at later times by solving a reduced form of the master equation.

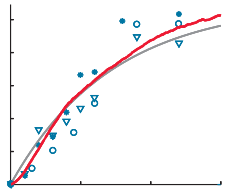


Figure 8.6a (page 190)

8.7 T_2 Simulate transcriptional bursting

Obtain Dataset 12. Use these experimental data to make graphs resembling those in Figure 8.6. Now write a computer code based on your solution to Problem 8.5, but with the additional realism outlined in Section 8.4.2′b (page 198), and see how well you can reproduce the data with reasonable choices of the model parameters. In particular, try the value $n_0 = 5$, which gives a reasonable amount of randomness in the cell division times.

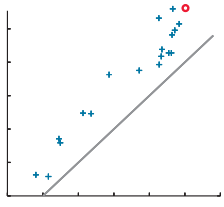


Figure 8.6b (page 190)

[23] T_2 Section 8.4.2′a (page 197) gives some justification for this approach.