Solutions to Homework Problems for the Complexity Explorer Course on Random Walks

1. `Displacement of a random walk.` Consider the Pearson random walk in any spatial dimension in which the length of each step has the fixed value a, but the direction is arbitrary. Compute the fourth moment of the displacement after N steps

$$M_4 \equiv \Big\langle \Big(\sum_{i=1}^{N} \mathbf{x}_i\Big)^4 \Big\rangle .$$

Solution. Expanding the quartic, the fourth moment can be written in the form

$$M_4 = \Big\langle \sum_{i,j,k,\ell}^{N} \mathbf{x}_i \mathbf{x}_j \mathbf{x}_k \mathbf{x}_\ell \Big\rangle .$$

The only terms that are non-zero after performing the average are those that contain even powers of $\mathbf{x}_i$. There are two such types of terms—quartic terms of the form $\mathbf{x}_i^4$ and biquadratic terms of the form $\mathbf{x}_i^2\mathbf{x}_j^2$, with $i \neq j$. Clearly, there are N quartic terms. To determine the number of biquadratic terms, imagine writing out the quartic as the product of its 4 factors. Symbolically

$$M_4 = (\ldots\ldots)(\ldots\ldots)(\ldots\ldots)(\ldots\ldots) .$$

We now need the number of distinct biquadratic groupings. If we pick $\mathbf{x}_i$ from the first factor (N ways), it must be joined with another $\mathbf{x}_i$ from one of the remaining 3 factors. Since the index i has been "used up", there are $N-1$ other available indices to group. Now this second index comes from a unique pair of the remaining two factors. To summarize, there are $6N(N-1)$ unique biquadratic pairings. Thus we obtain

$$\begin{aligned} M_4 &= N\langle \mathbf{x}_i^4\rangle + 6N(N-1)\langle \mathbf{x}_i^2\mathbf{x}_j^2\rangle \qquad\qquad i \neq j\,, \\ &= Na^4 + 6N(N-1)\langle \mathbf{x}_i^2\rangle\langle \mathbf{x}_j^2\rangle\,, \\ &= Na^4 + 6N(N-1)a^4\,. \end{aligned}$$

2. `Probability distribution of a biased random walk.` Consider a random walk in one dimension in which a step to the right occurs with probability p and a step to the left occurs with probability $q = 1-p$.

   (a) Determine the probability $P(r,\ell,t)$ that a walk of t total steps has taken r steps to the right and ℓ steps to the left.

   (b) Transform the above expression for $P(r,\ell,t)$ to $P(x,t)$, the probability that the walk is at x at time t.

   (c) Use Stirling's approximation to compute the long-time limit of the probability distribution $P(x,t)$.

   Solution. The number of distinct t-step walks with r steps to the right and ℓ steps to the left is the binomial factor

$$N(r,\ell,t) = \frac{t!}{r!\,\ell!} .$$

The probability $P(r,\ell,t)$ is then

$$P(r,\ell,t) = \frac{t!}{r!\,\ell!}\, p^r q^\ell\,.$$

Using $x = r-\ell$ and $t = r+\ell$, we eliminate r and ℓ in favor of t and x via $r = \frac{1}{2}(t+x)$ and $\ell = \frac{1}{2}(t-x)$. Using these in the above equation, we have

$$P(x,t) = \frac{t!}{\left(\frac{t+x}{2}\right)!\,\left(\frac{t-x}{2}\right)!}\, p^{(t+x)/2}\, q^{(t-x)/2}\,.$$

To obtain the long-time limit of the distribution, one needs to be careful in applying Stirling's approximation for the biased random walk. The crucial point is that the distribution is sharply peaked about the most probable value of the displacement $x_{\rm mp} = (p-q)t$, and the final Gaussian form is an expansion about this most probable displacement. Thus as a preliminary to applying Stirling's approximation, we write $x = x_{mp} + \epsilon = (p-q)t + \epsilon$, so that we can expand about $\epsilon = 0$. We have

$$\begin{aligned}
\tfrac{1}{2}(t+x) &= \tfrac{1}{2}\big[t + (p-q)t + \epsilon\big] = pt + \tfrac{1}{2}\epsilon = pt(1+\tfrac{\epsilon}{2pt})\,,\\
\tfrac{1}{2}(t-x) &= \tfrac{1}{2}\big[t - (p-q)t - \epsilon\big] = qt - \tfrac{1}{2}\epsilon = qt(1-\tfrac{\epsilon}{2qt})\,.
\end{aligned}$$

We now apply Stirling's approximation to $\ln P(x,t)$ to give

$$\begin{aligned}
\ln P(x,t) \simeq\; & t\ln t - t + \tfrac{1}{2}\ln(2\pi t)\\
& - \tfrac{1}{2}(t+x)\ln\left[\tfrac{1}{2}(t+x)\right] - \tfrac{1}{2}\ln[\pi(t+x)]\\
& - \tfrac{1}{2}(t-x)\ln\left[\tfrac{1}{2}(t-x)\right] - \tfrac{1}{2}\ln[\pi(t-x)]\\
& + \tfrac{1}{2}(t+x)\ln p + \tfrac{1}{2}(t-x)\ln q\,.
\end{aligned}$$

As an additional simplification, we write the first term as

$$t\ln t = \tfrac{1}{2}(t+x)\ln t + \tfrac{1}{2}(t-x)\ln t\,,$$

and then combine the terms with the same prefactors in front of the logarithms. We also use the alternative expressions for $\frac{1}{2}(t\pm x)$ given above to obtain, after some simple steps,

$$\ln P \simeq -pt(1+\tfrac{\epsilon}{2pt})\ln\left(1+\tfrac{\epsilon}{2pt}\right) - qt(1-\tfrac{\epsilon}{2qt})\ln\left(1-\tfrac{\epsilon}{2qt}\right) + \tfrac{1}{2}\ln\left[\frac{2\pi t}{2\pi(pt+\frac{1}{2}\epsilon)\,2\pi(qt-\frac{1}{2}\epsilon)}\right].$$

Finally, we expand the first two logarithms in a Taylor series to second order and the third logarithm to zeroth order. Clearly, all terms linear in ϵ must cancel in the first two logarithms (you can easily check this), and we are left with

$$\begin{aligned}
\ln P &\simeq -\frac{\epsilon^2}{8pt} - \frac{\epsilon^2}{8qt} + \tfrac{1}{2}\ln\left[\frac{1}{2\pi pqt}\right],\\
&= -\frac{\epsilon^2}{8pqt} - \tfrac{1}{2}\ln(2\pi pqt)\,.
\end{aligned}$$

Thus the final result in the large-time asymptotic limit is the Gaussian distribution

$$P(x,t) \simeq \frac{1}{\sqrt{2\pi pqt}}\; e^{-\epsilon^2/8pqt}\,,$$

with $\epsilon = x - (p-q)t$.

3. `Diffusion equation.` Consider a random walk that steps to the right with probability 1/3, to the left with probability 1/3, and remains in place with probability 1/3.

   (a) Write the Master equation that determines the evolution of $P(x,t)$.

   (b) Taylor expand the master equation and determine the partial differential equation that is satisfied by $P(x,t)$. Determine the diffusion coefficient D of this process and compare it with the diffusion coefficient of the nearest-neighbor random walk.

   Solution. Let us take the time increment for a single step as dt. Then the Master equation for $P(x,t)$ is

$$P(x,t+dt) = \tfrac{1}{3}P(x-1,t) + \tfrac{1}{3}P(x,t) + \tfrac{1}{3}P(x+1,t)\,.$$

   The first term on the right accounts for the contribution to $P(x,t+dt)$ due to the walk hopping from $x-1$ to x, the second term accounts for the walk remaining in place, and the third term account for the walk hopping from $x+1$ to x.

   We now Taylor expand the above Master equation to first order in time and second order in space. This gives

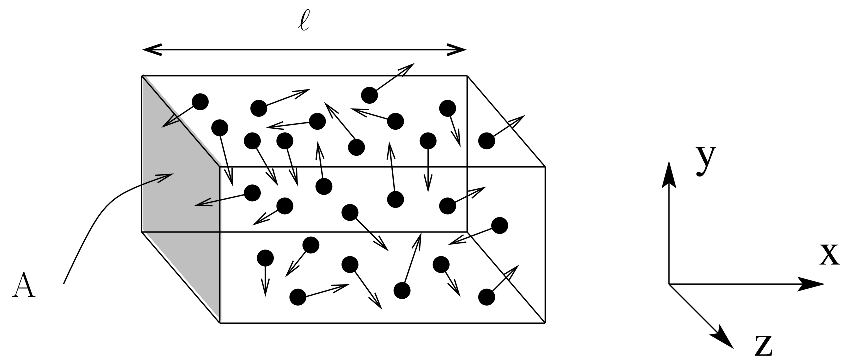
$$\begin{aligned} P(x,t) + \frac{\partial P}{\partial t}dt + \ldots = &\tfrac{1}{3}\Big[P(x,t) + \frac{\partial P}{\partial x}dx + \tfrac{1}{2}\frac{\partial^2 P}{\partial x^2}(dx)^2 + \ldots\Big] \\ &+ \tfrac{1}{3}P(x,t) \\ &+ \tfrac{1}{3}\Big[P(x,t) - \frac{\partial P}{\partial x}dx + \tfrac{1}{2}\frac{\partial^2 P}{\partial x^2}(dx)^2 + \ldots\Big]\,. \end{aligned}$$

   Gathering terms, the zeroth-order terms and the first-order spatial terms cancel and we arrive at the diffusion equation

$$\frac{\partial P}{\partial t} = \frac{1}{3}\frac{(dx)^2}{dt}\,\frac{\partial^2 P}{\partial x^2}\,,$$

   in which we identify the diffusion coefficient D with $(dx)^2/(3\,dt)$. Compared to the nearest-neighbor random walk, the diffusion coefficient is reduced by a factor of 2/3.

4. `Kinetic theory.` The density of a typical room-temperature gas is $\rho \approx 10^{20}$ molecules/cc and each molecule has a typical speed that is roughly 30000 cm/sec. Estimate the number of collisions that the ambient air makes with your body per second.



Solution. First, we estimate the number of air molecules that pass through a unit area per unit time. Consider a parallelpiped of length equal to the mean-free path ℓ and area A. The number of molecules in this parallelpiped is $\rho\ell A$. Very roughly, $\frac{1}{6}$ of the molecules are moving in each of the coordinate directions $\pm x$, $\pm y$, and $\pm z$. Thus, roughly $\frac{1}{6}$ of the molecules are moving in the $+x$ direction and will hit the shaded side of the parallelpiped in the above figure in a time that is roughly ℓ/v. Thus the number of molecules hitting the end of a parallelpiped of unit area per unit time is $\frac{1}{6}\rho v$; the exact result is $\frac{1}{4}\rho v$. I guesstimate the surface area of the body as 1 m^2, or 10^4 cm^2. Thus the number of molecules hitting somebody per unit time is

$$\tfrac{1}{6}\,\rho v \times 10^4 \approx \tfrac{1}{6}\,10^{20} \times (3\times 10^4) \times 10^4 \approx 10^{28} \text{ molecules/sec}\,.$$

5. `Extreme value statistics.` The distribution of velocities of an ideal gas of molecules of mass m at temperature T is given by the Maxwell-Boltzmann distribution

$$P(\mathbf{v})d\mathbf{v} = \left(\frac{m}{2\pi kT}\right)^{3/2}\; e^{-mv^2/2kT}\,d\mathbf{v}\,,$$

where kT, the thermal energy of each molecule is roughly $\frac{1}{40}$ eV at room temperature ($\approx 300\,$K). Thus $P(\mathbf{v})d\mathbf{v}$ is the probability that a molecule as a velocity that is in a range $d\mathbf{v}$ about $\mathbf{v}$. Using the reasoning given in the discussion of the failure of the central limit theorem, estimate the energy of the most energetic molecule in a gas at room temperature in a room of volume of 1000 cubic meters.

Solution. To simplify matters, let's directly consider the energy distribution of the molecules: $P(E)\,dE = e^{-E/kT}\,dE$, where E is the energy. Now we apply the extremal criterion that was used in lecture

$$\int_{E_{\rm max}}^{\infty} e^{-E/kT}\,dE \approx \frac{1}{N}\,.$$

Again, this criterion states that in gas of N molecules, there is one molecule whose energy is greater than or equal to $E_{\rm max}$. This statement then allows us to estimate $E_{\rm max}$. Evaluating the above integral gives

$$kT\,e^{-E_{\rm max}/kT} = \frac{1}{N} \longrightarrow E_{\rm max} = kT\ln N\,.$$

The density of a room-temperature gas is roughly 10^{20} molecules/cc. For a room of 1000 cubic meters, the number of molecules within this room is therefore roughly 10^{29}. Thus

$$E_{\rm max} \approx kT\ln 10^{29} \approx 70\,kT \approx 1.7\,\text{eV} \approx 20000\,\text{K}\,.$$

6. `First-passage in a finite interval.` Consider a biased random walk in the finite interval $[0, L]$ in which the walk hops to the right with probability p and to the left with probability $q = 1 - p$. The walk is immediately absorbed when it reaches either 0 or L.

(a) Determine the probability E_n that the walk is absorbed at L when it starts at site n.

Solution. The exit probability E_n satisfies the backward Kolmogorov equation

$$E_n = pE_{n+1} + qE_{n-1}\,.$$

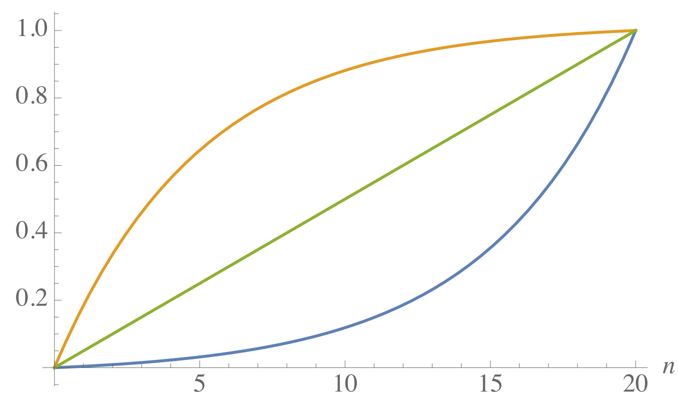
This constant-coefficient, second-order recursion generally has exponential solutions in n, except in the case where $p = \frac{1}{2}$ where the solution is a power law in n (and we know from the lecture that the solution is $E_n = \frac{n}{L}$). The statement of the problem mandates that $E_0 = 0$ and $E_L = 1$. We now assume $E_n \propto \lambda^n$, substitute into the above recursion, and cancel out a common factor λ^{n-1} to arrive at $p\lambda^2 - \lambda + 1 = 0$, with elemental solutions $\lambda_\pm = \frac{1}{2p}\big[1 \pm \sqrt{1-4pq}\big]$. Thus the general solution has the form

$$E_n = A\lambda_+^n + B\lambda_-^n\,,$$

where A and B are constants. These constants are determined by the boundary conditions $E_0 = 0$ and $E_L = 1$. The former gives $B = -A$, while the latter gives A and the solution to our problem is

$$E_n = \frac{\lambda_+^n - \lambda_-^n}{\lambda_+^L - \lambda_-^L}\,.$$

A plot of E_n versus n for the case of $L = 20$ is shown below for the cases of $p = 0.45$ (concave upward) and $p = 0.55$ (concave downward). The straight line gives the exit probability in the case of no bias.



(b) Determine the average time t_n for the walk to exit the interval either at 0 or at L when it starts at site n. Find the starting location that maximizes the exit time.

Solution. The average exit time t_n satisfies the backward Kolmogorov equation

$$t_n = p(t_{n+1} + 1) + q(t_{n-1} + 1) = pt_{n+1} + qt_{n-1} + 1\,.$$

The general solution to this equation is the sum of a particular solution plus the solution to the homogeneous equation, which is the same as the equation for E_n in part (a). With a little trial and error, the particular solution is $t_n = -\frac{n}{v}$. Thus the general solution for the average exit time is

$$t_n = -\frac{n}{v} + A\lambda_+^n + B\lambda_-^n \,.$$

The boundary condition $T_0 = 0$ immediately gives $A+B=0$, while the boundary condition $T_L = 0$ leads to

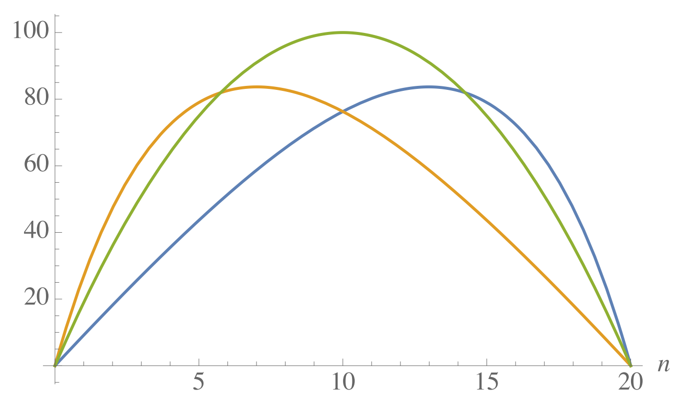
$$A = \frac{L/v}{\lambda_+^L - \lambda_-^L} \,.$$

Thus the final expression for the average exit time is

$$t_n = -\frac{n}{v} + \frac{L}{v}\left(\frac{\lambda_+^n - \lambda_-^n}{\lambda_+^L - \lambda_-^L}\right) \,.$$

This formula holds for both $v > 0$ and $v < 0$, that is for $p \neq q$.

A plot of t_n versus n for the case of $L = 20$ is shown below for the cases of $p = 0.45$ (peak at $n \approx 6$) and $p = 0.55$ (peak at $n \approx 14$). The symmetric curve is the exit time in the case of no bias, $p = \frac{1}{2}$.



7. `First-passage on the semi-infinite interval.` Consider a biased nearest-neighbor random walk with hopping probabilities p and q to the right and left, respectively, on the semi-infinite interval $[0, \infty]$. The walk is absorbed if it reaches $x = 0$.

   (a) Using the backward Kolmogorov approach, determine the probability E_n that the walk is eventually absorbed at the origin when it starts at site n. Separately consider the cases $p > q$, $p < q$, and $p = q$ (symmetric walk).

   Solution. The exit probability E_n satisfies the backward Kolmogorov equation

$$E_n = pE_{n+1} + qE_{n-1} \,.$$

   This constant-coefficient, second-order recursion generally has exponential solutions. In addition, we must have $E_0 = 1$ by definition. We now assume $E_n \propto \lambda^n$, substitute into the above recursion, and cancel out a common factor λ^{n-1} to arrive at $p\lambda^2 - \lambda + 1 = 0$. This equation has two solutions:

$$\lambda = \begin{cases} 1 & p \le \frac{1}{2}\,, \\ \frac{1}{2p}\left[1 - \sqrt{1-4pq}\right] & p > \frac{1}{2}\,. \end{cases}$$

For $p > \frac{1}{2}$, λ is always less than 1 and monotonically decays to 0 as $p \to 1$. Thus the probability of eventually reaching the origin, $E_n = \lambda^n$, decays very quickly with starting location n as $p \to 1$.

(b) For the case where $p < q$ (bias to the left), compute the average time t_n for the walk to reach the origin when it starts at site n. What is the limiting behavior of this exit time in the limit $p = q$?

Solution. Please notes that this solution uses some information that was not presented in lecture. Because we're dealing with the semi-infinite interval, I'm going to cheat a tiny bit and study the problem in the continuum limit; it's much simpler in this formulation. Thus we start with the backward Kolmogorov equation for the average exit time

$$t_n = pt_{n+1} + qt_{n-1} + 1\,,$$

replace t_n by $t(x)$ and $t_{n\pm 1}$ by $t(x \pm dx)$, define the time increment of a single step as dt, and Taylor expand the differences. This gives

$$t(x) \approx p\big[t(x) + t'dx + \tfrac{(dx)^2}{2}t'' + \ldots\big] + q\big[t(x) - t'dx + \tfrac{(dx)^2}{2}t'' + \ldots\big] + dt\,.$$

Keeping only terms up to quadratic and identifying $v = (p-q)\,\frac{dx}{dt}$ and $D = (p+q)\frac{dx^2}{2dt}$, the backward Kolmogorov equation for $t(x)$ is

$$Dt'' + vt' = -1\,.$$

We now solve this equation on the *finite* interval $[0, L]$ and then take the limit $L \to \infty$. For the problem on the finite interval, the boundary conditions are $t(0) = 0$ and (more subtly) $t'(L) = 0$. The latter reflecting boundary condition mandates that the walk is reflected if it reaches the point $x = L$.

The general solution to $Dt'' + vt' = -1$ is the sum of the particular and homogeneous solutions:

$$t(x) = -\frac{x}{v} + A\,e^{-vx/D} + B\,.$$

The boundary condition $t(0) = 0$ gives that $A + B = 0$, while the boundary condition $t'(L) = 0$ gives $A = -\frac{D}{v^2}\,e^{-vL/D}$. Thus for the finite interval

$$t(x) = -\frac{x}{v} - \frac{D}{v^2}\,e^{-vL/D}(e^{-vx/D} - 1)\,.$$

To obtain the exit time for the infinite interval we let $L \to \infty$. This gives

$$t(x) = -\frac{x}{v}\,.$$

This simple result is just the exit time for a particle that moves at constant speed v. Thus the random-walk nature of the motion plays no role for this system; all that matters is the underlying bias.