question **2** views

Daily Challenge 26.1

(**Part (a) due: Tuesday 4/2 at 12:00 noon Eastern**
Part (b) due: Tuesday 4/2 at 4:00 pm Eastern)

Let's continue working on your computational narrative in Jupyter for this chapter's tutorial.

Push a rough draft of your notebook for the presentation, including reasonable progress on the outline/suggestions for part (a), by the noon deadline. Then push a revised draft including all of the tasks in part (b) by the 4 pm deadline. I will pay when both are done.

(**Part a**) Fill in the details of the outline that you proposed in part (a) of DC 25.7, namely your plan for explaining the diffusion and random walk calculation.

Be sure to include enough examples, explanations, and details to make the exposition intelligible to a student.

Here are some suggestions for extra content, besides that in the original challenge, that you can include for pedagogical clarity:

- Show example histograms of the positions of random walkers after different numbers of steps (say 10, 100, and 200). Point out that the shape of each histogram is roughly a Gaussian with increasing width.

  Use this to motivate the definition of a probability distribution, which is roughly the limit of a histogram as the bin size gets small.

- Compute the standard deviation of the ending positions of random walkers after various numbers of steps. Show that the standard deviation tends to grow as the square root of the number of steps. Use this to introduce the calculation you do to find the variance of the Gaussian.

  (You can use any of the three ways that you used to compute the variance of a Gaussian in DC 24.4.)

- Interpret the final position of a random walker after N steps as the sum of N independent and identically distributed random variables, each of which is ± 1 with equal probabilities. Use this to explain the idea of the central limit theorem.

- Consider modifying the random walk code, either by changing the probability distribution for the step size (perhaps draw from a uniform or Gaussian distribution for the step size rather than simply choosing ± 1), or by adding drift, or by moving up in dimension (e.g. do a random walk in (x, y) space where you increment either x or y by ± 1 with equal probabilities.)

  Analyze the results of the modified experiment. Interpret what changes or doesn't change.

- Explain the following alternate way of showing that the average distance from the origin goes like $\sqrt{N}$.

  Think of each step to the left or right as a random variable x with outcomes ± 1, each of which has probability $\frac{1}{2}$. By definition, the mean (first moment) is

  $$E[x] = \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot (-1) = 0,$$

  while the second central moment (variance) is

  $$E[(x - \mu)^2] = E[x^2] = \frac{1}{2} \cdot (1^2) + \frac{1}{2} \cdot (-1)^2 = \frac{1}{4}.$$

  By the central limit theorem, when we add up some large number N of independent draws of this random variable x, it will tend to a Gaussian with mean $N\mu = N \cdot 0 = 0$ and variance $N\sigma^2 = \frac{N}{4}$. But this means $\sigma \sim \sqrt{N}$, as you showed before.

- What if the random walker moves to the right with probability p and to the left with probability $q = 1 - p$?

  Think of this like flipping an unfair coin, which we did when discussing Bayes' theorem. We saw that the probability of getting h "heads" (i.e. h steps to the right) was

  $$\mathbb{P}(h) = \binom{N}{h} p^h (1-p)^{N-h} = \frac{N!}{h!(N-h)!} p^h (1-p)^{N-h}.$$

  It's possible to show (e.g. using the steps here) that the expected value of steps to the right is pN and the expected number of steps to the left is $(1-p)N$, as you might expect. The variance is $\sigma^2 = Npq$.

  So with an unfair walk, the average distance from the origin σ still scales like $\sqrt{N}$, although with a different slope pq.

- Explain more about the connection with Fick's law of diffusion. For instance, one can consider diffusion in a semi-infinite solid rather than the entire real line as you've done here. In that case, you can integrate explicitly to show that the concentration at depth x in the solid is related to the error function. Section 5.4 of this book explains this nicely, if briefly.

The bullet-pointed tasks in part (a) are just ideas; you don't need to include all of the things I suggest. Use your own creative license, include the things that you find most interesting, and add/modify any additional demonstrations or explanations you like.

However, don't be lazy. You need to include enough content so that it takes $40 - 60$ minutes for you to go through the notebook in the tutorial presentation and explain the steps to the student.

As always, doing the bare minimum is frowned upon.

(**Part b**) Now you will analyze your code which computes the fraction of the time that a random walker spends at positions $x > 0$.

These numbered tasks of part (b) are all mandatory, not suggestions like in part (a). Do all of the following and include them with explanations in your notebook.

1. Run your `count_positive` code from yesterday, generate the list of fractions for some reasonably large number of walks and steps per walk, and plot a histogram of the results.

2. The true probability distribution for the fraction of time that a random walker spends at $x > 0$ is

$$f(x) = \frac{1}{\pi\sqrt{x(1-x)}}.$$

   Use matplotlib to plot this function on top of the histogram. How well do they agree?

3. Recall that the *beta distribution*, which we met while flipping coins, is

$$f(x; \alpha, \beta) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)},$$

   where α and β are the shape parameters.

   By plugging in $\alpha = \frac{1}{2}$, $\beta = \frac{1}{2}$, and looking up the value $B(\alpha, \beta) = \sqrt{\pi}$, we see that the probability distribution for the fraction of time a random walker spends at $x > 0$ is actually $\mathrm{Beta}\left(\frac{1}{2}, \frac{1}{2}\right)$.

   Read and understand these words, then explain them to the student in your notebook.

4. Although it is a beta distribution, this PDF is more commonly known as the arcsine distribution, since its cumulative distribution function is

$$F(x) = \int_0^x f(t)\, dt = \frac{2}{\pi}\arcsin(\sqrt{x}).$$

   Verify this by taking the derivative $\frac{dF}{dx}$ (using the fundamental theorem, of course; the variable appears in the upper bound of integration), and check that it reduces to the distribution $f(x)$ in the previous part.

   Explain all of this to the reader of your notebook. Don't assume they know anything about probability; you will need to explain what a cumulative probability distribution is.

Push a rough draft of the notebook with the (b) stuff by 4 pm.

daily_challenge

Updated 12 days ago by Christian Ferko

the students' answer, *where students collectively construct a single answer*

doop

Updated 5 days ago by Logan Pachulski

the instructors' answer, *where instructors collectively construct a single answer*

Click to start off the wiki answer

followup discussions *for lingering questions and comments*