MATH327: Statistical Physics, Spring 2022

Computer Project — Part 2

Instructions

In this second part of the computer project you will numerically analyze *anomalous diffusion* in a one-dimensional random walk, building on the numerical methods you developed for ordinary diffusion and checked against exact analytic predictions in the first part of the project.

There are two exercises below, which include some background information on the Cauchy–Lorentz distribution and anomalous diffusion. While the exercises mention some syntax specific to Python, you may use a different programming option if you prefer. This demo illustrates all the Python programming tools needed for the project. Even running slowly in the cloud via replit.com, the computing for each exercise should complete in a few minutes or less.

This part of the project is **due by 23:59 on Thursday, 24 March**, and anonymous marking is turned on. Submit it by file upload on Canvas.[1] **Both** your answers to the questions below and the code that produces your results must be submitted. These can be uploaded as separate files or in a tar/zip archive, as you prefer. With the exception of Mathematica `.nb` files, it will be quicker for me to check code submitted in its native format (for example, a `.py` file for Python code or a `.m` file for MATLAB code).

Exercise 4: Cauchy–Lorentz distribution

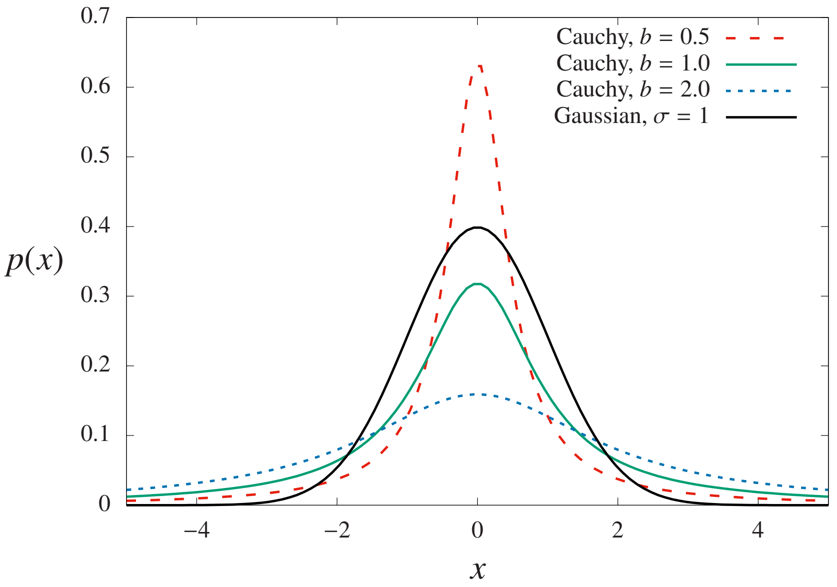
Background

The first part of this project involved the distribution $\cos(z)$ for $-\frac{\pi}{6} \leq z < \frac{\pi}{6}$, for which we were able to apply the central limit theorem based on its mean 0 and standard deviation $\frac{\sqrt{\pi^2+12\pi\sqrt{3}-72}}{6}$. We now consider a more interesting distribution,

$$p_C(x) = \left(\frac{1}{b\pi}\right)\frac{1}{1+(x/b)^2} \qquad\qquad x \in \mathbb{R} \tag{1}$$

which is known as the Cauchy (or Cauchy–Lorentz) distribution. Here b is a constant parameter that controls the width of the peak in $p_C(x)$ around $x = 0$. The figure below illustrates this by plotting the Cauchy–Lorentz distribution for each of $b = 1/2$, $b = 1$ and $b = 2$, comparing them to the normal (or gaussian) distribution $\frac{1}{\sqrt{2\pi}}e^{-x^2/2}$.

[1] By submitting solutions to this assessment you affirm that you have read and understood the Academic Integrity Policy detailed in Appendix L of the Code of Practice on Assessment and have successfully passed the Academic Integrity Tutorial and Quiz. The marks achieved on this assessment remain provisional until they are ratified by the Board of Examiners in June 2022.



The figure shows how the peak of the Cauchy–Lorentz distribution around $x = 0$ becomes higher and narrower as b decreases. Even when its peak is very narrow, as $|x|$ increases $p_C(x)$ again becomes larger than the gaussian distribution, simply because the latter decreases exponentially quickly while $p_C(x)$ decreases only $\sim 1/x^2$. These "fat tails" at large $|x|$ make the Cauchy–Lorentz distribution both interesting and challenging to analyze.

Task

Fix $b = 2$ in the Cauchy–Lorentz distribution, so that Eq. 1 becomes

$$p_C(x) = \left(\frac{1}{\pi}\right) \frac{2}{4+x^2} \qquad\qquad x \in \mathbb{R}. \tag{2}$$

What is the integral of this distribution over its full range,

$$I \equiv \int_{-\infty}^{\infty} p_C(x)\; dx = \int_{-\infty}^{\infty} \left(\frac{1}{\pi}\right) \frac{2}{4+x^2} dx?$$

The usual starting point to analyze a probability distribution is finding its mean and standard deviation, by evaluating

$$\langle x \rangle = \int x\; p(x)\; dx \qquad\qquad \langle x^2 \rangle = \int x^2\; p(x)\; dx.$$

For the Cauchy–Lorentz distribution in Eq. 2, consider instead the functions

$$f(a) = \int_{-a}^{a} x\; p_C(x)\; dx = \int_{-a}^{a} \left(\frac{1}{\pi}\right) \frac{2x}{4+x^2} dx$$
$$g(a) = \int_{-a}^{a} x^2\; p_C(x)\; dx = \int_{-a}^{a} \left(\frac{1}{\pi}\right) \frac{2x^2}{4+x^2} dx.$$

How do $f(a)$ and $g(a)$ behave in the limit $a \to \infty$?

[6 marks]

Turning to a numerical analysis of the Cauchy–Lorentz distribution, the first step is to determine the transform $F(u)$ that will map the uniformly distributed pseudo-random numbers u to $x = F(u) \in \mathbb{R}$. Recall that the uniform distribution implemented by the Python function `random.random()` is

$$p(u) = \begin{cases} 1 & \text{for } 0 \leq u < 1 \\ 0 & \text{otherwise} \end{cases} .$$

What is the transform F that provides $x = F(u)$ distributed according to $p_C(x)$ in Eq. 2?

Hints: Guided by the relation

$$p_C(x) = p(u) \frac{d}{dx} F^{-1}(x),$$

it will suffice to propose an ansatz for $F(u)$ based on integrating $p_C(x)$, and then follow the steps in Exercise 2 to confirm that this ansatz produces the desired distribution. Integrating will introduce a constant of integration, which can be chosen so that $x \to -\infty$ as $u \to 0$ and $x \to \infty$ as $u \to 1$.

`[6 marks]`

Now initialize the random number generator with seed $s = 327$. Generate $R = 1{,}000{,}000$ pseudo-random numbers $x_r = F(u_r)$ using the transform you found. Plot the histogram of these million $\{x_r\}$ and check whether it agrees with the Cauchy–Lorentz distribution shown above.

Hints: You will need to set an appropriate range for the x-axis of this histogram. A range $-10 \leq x \leq 10$ with roughly 200 bins should suffice to show all the interesting features. In Python this can be done by providing

```
bins = np.arange(-10.0, 10.0, 20.0/201.0)
```

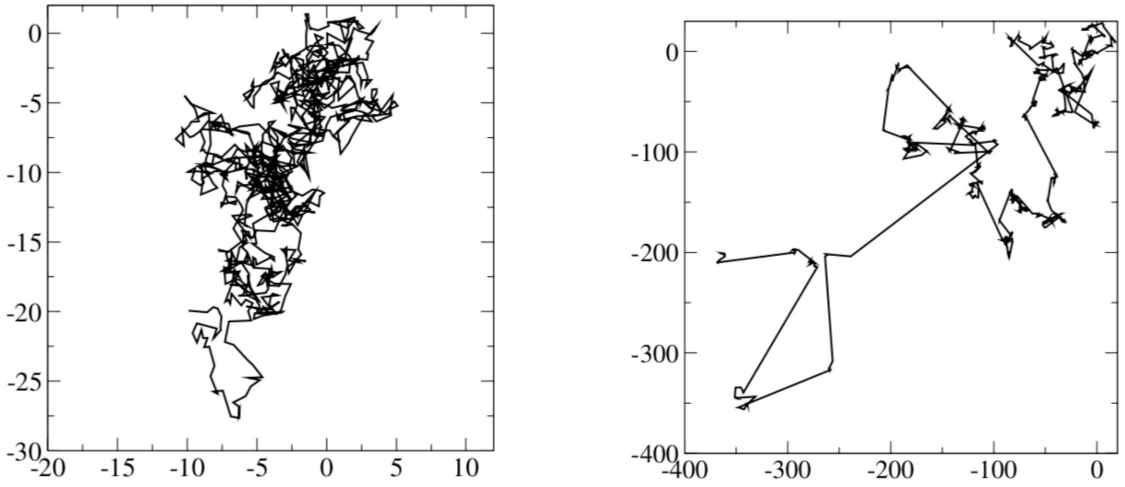
to the Matplotlib `hist` function used previously. In this exercise it is optional to plot $p_C(x)$ itself on top of this histogram—if you choose to do so, you may need to adjust its normalization (and you should think about why this is needed).

`[8 marks]`

Exercise 5: Anomalous diffusion

Background

The "fat tails" of the Cauchy–Lorentz distribution mean that $p_C(x)$ provides larger probabilities for *rare events* (with large $|x|$) to occur, compared to the gaussian distribution. This feature of the distribution is illustrated in the figures below, each of which shows a thousand-step random walk in two dimensions—randomly selecting both the size of each step and the direction $0 \leq \phi < 2\pi$ in which to step. The walk shown on the left uses step sizes drawn from a gaussian distribution. Even in two dimensions, random walks of this sort obey the law of diffusion, with a diffusion length growing proportionally to the square root of the number of steps, $\ell_2(N) = \sqrt{\left\langle [X(N)]^2 \right\rangle - \left\langle X(N) \right\rangle^2} \propto \sqrt{N}$.



The walk shown on the right instead uses step sizes drawn from a Cauchy–Lorentz distribution. Note that the axes for this figure cover a much larger range! The fat tails of the Cauchy–Lorentz distribution result in occasional very large jumps, leading to random walks that do not obey the law of diffusion.

Returning to one-dimensional random walks, some of the results from Exercise 4 motivate defining the **generalized diffusion length**

$$\ell_\theta(N) = \left\langle |X(N)|^\theta \right\rangle^{1/\theta}, \tag{3}$$

which depends on a positive real parameter $\theta > 0$. Since θ is not necessarily an integer, the absolute value is needed to ensure $\ell_\theta \in \mathbb{R}$, rather than becoming complex valued. If $\langle X(N) \rangle = 0$ and ℓ_θ is well-defined with $\theta = 2$, then this generalized diffusion length could reproduce the standard deviation ℓ_2 and exhibit the ordinary law of diffusion, $\ell_2 \propto N^{1/2}$.

For the Cauchy–Lorentz distribution, ℓ_θ is ill-defined for any $\theta \geq 1$. This parameter θ can take only values $0 < \theta < 1$. The resulting ℓ_θ exhibits **anomalous diffusion**,

$$\ell_\theta(N) \propto N^\alpha,$$

where the exponent is either $\alpha > \frac{1}{2}$ (called *super-diffusion*) or $0 < \alpha < \frac{1}{2}$ (called *sub-diffusion*). This exercise investigates the exponent α for the distribution $p_C(x)$ in Eq. 2, and checks whether or not α depends on θ.

Task a: Fixed number of steps

Reset by initializing the random number generator with seed 327. With fixed $N = 100$, generate R 100-step random walks,

$$X_r(N) = \sum_{i=1}^{N} x_i \qquad\qquad r = 1, 2, \cdots, R,$$

for each of the five $R = 10$, 100, 1000, $10{,}000$ and $100{,}000$. Use the resulting X_r to numerically estimate

$$\overline{\ell_\theta(N)}_R \approx \left[\frac{1}{R}\sum_{r=1}^{R} |X_r(N)|^\theta\right]^{1/\theta}$$

for three values of $\theta = 0.25$, 0.5 and 0.75. (**Hint:** NumPy provides both an `abs` function to take the absolute value, and a `power` function to compute non-integer powers.)

`[12 marks]`

Task b: Anomalous diffusive exponent

Reset by initializing the random number generator with seed 327. Then fix $R = 10{,}000$ and estimate $\overline{\ell_\theta(N)}_R$ for *every* $N = 1, 2, \cdots, 250$, again considering $\theta = 0.25$, 0.5 and 0.75. Instead of reporting your numerical results, plot all three $\ell_\theta(N)$ vs. N in a single figure. (**Hint:** You can ignore potential correlations between $\ell_\theta(N)$ for different values of N.)

`[8 marks]`

Now fit your numerical results for each $\theta = 0.25$, 0.5 and 0.75 to the function

$$\ell_\theta(N) = DN^\alpha.$$

Report your results for D and α, and comment on their sensitivity to the value of θ. (**Hint:** Optionally testing different values of R, N or θ may help to distinguish between real sensitivity vs. statistical fluctuations, if you are unsure whether or not an observed effect is significant.)

`[10 marks]`