**Note 1**

* In this plot, you will learn about the random walk, and see how it is a very basic but illuminating example of a stochastic process.
* After summarizing the necessary theoretical knowledge that we have already learned in the course, we will conduct computer simulations of a random walk and perform a stochastic analysis of the results.
* At the end of this plot, we will understand the close relationship between a random walk and the diffusion process.

**Note 2**

* Because a "random walk" is translated in Japanese as "drunken steps", I will use this analogy to explain the random walk process.
* Imagine that a drunken man is wandering around town after a heavy night of drinking. He is completely disoriented, with no clue as to how to get home. At each step, he loses memory of which direction he was going, and chooses a new random direction in which to move.
* For simplicity, we will consider the 1−D version of this random walk. Thus, at each step the drunken man has two options, he can either take a step to the left, or to the right. At each step he is basically flipping a coin to determine in which direction to move.
* Thus, we set si=−1or 1, be the stochastic variable that determines i-th step, where si=−1 means to take a step to the left and +1 to the right.
* After m such steps, the location of the man is given by “l” equals to a sum of si from i=0 to m.
* Assuming he was initially at the origin, i.e., l=0 at m=0, our aim is to calculate the probability distribution function P(l,m) of finding the drunken man at a location L=l after M=m steps.

**Note 3**

* As you probably realized, we can evaluate P(l,m) in terms of the binomial distribution function.
* To this end, we introduce two new variables, namely n+, the total number of steps to the right, and n−, The total number of steps to the left.
* Then the total number of steps m is represented as “n+” + “n−“, and the current location l is represented as “n+” – “n−“.
* Solving for n+ and n− in terms of m and l, we obtain Eq.(E1).
* Using  n+ or n−, the probability distribution function P(l,m) can be estimated using the binomial distribution function P\_Binomial of obtaining n+ right steps in “m” total steps or equivalently, of obtaining n− steps in “m” steps.
* We already know that such probabilities are given explicitly by the form of Eq.(E3).

**Note 4**

* When n+ and m are much larger than 1, we know that the binomial distribution function converges to the normal distribution given in Eq.(E4), with the average and variance of n+ given in Eqs.(E5) and (E6).
* Recall that n+ is equal to (m+l) divided by 2. Thus, the binomial distribution function appearing in Eq.(E4) can be rewritten as Eq.(E7), this time as a function of l.
* Therefore the desired probability distribution function P(l,m) is obtained as Eq.(E9), with the average and the variance for l given by Eqs.(E10) and (E11).

**Note 5**

* Now let us perform computer simulations for a random walk using a random number generator to generate the steps of the walk.
* First import the numerical and graphical libraries as usual.

**Note 6**

* The main part of the simulation code is shown here.
* In the beginning, we set p=0.5, which represents a drunken man with equal probability for moving right or left. Then, set the number of total steps M=1,000 and the number of independent random walkers N=100,000.
* The average and the standard deviation for the final location l after M steps are calculated using Eqs.(E10) and (E11).
* We create a location array "L", of size N, which will contain the final position for all of the N independent drunken men.
* After initializing the location array L to zero, assuming the initial position is the same for all of them, we generate a random sampling of steps for each of the men.
* The M steps for each walk are randomly chosen from -1 or +1, and stored as a variable "step" in the 10th line.
* Then, we calculate the final location l by accumulating all of the steps, and storing the result in the L array, at line 11 of the code.
* This procedure, of generating the M steps, and then adding them to obtain the final location is repeated N times, once for each drunken man. This is achieved by the for loop which starts at line 9. Notice the indentation of the code for lines 10 and 11, to indicate that they are performed inside the for loop.
* The remaining part of the code is just used to draw the plots of the histogram for l, as well as the expected Gaussian distribution.

(Run cell)

* By running this code example, you see that the probability distribution function P(l,m) is perfectly represented by Eqs.(E9) to (E11).
* You may repeat the same simulation by choosing different values of total steps, for example M=100, 1,000, 10,000, and 100,000 to see how the distribution changes with the total number of steps.

**Note 7**

* Let us next discuss how the random walk can be connected to the diffusion constant D.
* To this end, we define two constants, namely the length of a single step "a" and the time between subsequent steps ts.
* Here we consider a drift free case where p=0.5 and thus the average of l is 0.
* Using "a" and ts, we define the position of the random walker as x, which is equal to "a" multiplied with the current location "l", and the duration of time for making m steps as ts, which is equal to ts multiplied with the number of total steps "m".
* Taking care of the normalization as shown in Eq.(E12), the probability distribution function P(l,m) can be converted to P(x,t) as shown in Eqs.(E13) to (E16).
* Note that the variance of the position x is now given as a function of time t, and not in terms of the number of steps m.

**Note 8**

* Now, let us turn our attention to the diffusion equation shown in Eq.(E17) with a diffusion constant D.
* As initial condition we assume a delta distribution centered at the origin. This means that all the walkers or diffusin particles are located at the origin.
* If we solve this diffusion equation using the initial condition shown in Eq.(18), the solution is given by Eq.(E19), which has an identical form to Eq.(E14).
* You can convince yourselves that this is indeed a solution, by computing the time derivative and the second spatial derivative, and verifying that Eq. (E17) is satisfied.
* By comparing Eqs.(E14) and (E19), we finally obtain the following general formula which relates the diffusion constant D and the variance of the position of random walkers σ2, which is also referred to as the mean square displacement.