**Note 1**

* In this plot, we will introduce the "Central Limit Theorem", which is probably the single most important theorem in the theory of probability.
* The theorem helps to explain why the Gaussian or Normal distribution finds such widespread applications in natural and social sciences.
* Briefly stated, it says that the sum of N independent random variables, which can be drawn from any arbitrary, but well behaved distributions, will converge to a Gaussian distribution in the limit of large N.
* We will discuss this in more detail later, but first we will begin by working out some practical examples that show the central limit theorem in practice.

**Note 2**

* In a previous lesson, we introduced the binomial distribution P(n,M), which describes a random process with two possible outcomes. The standard example is that of a coin toss, which can come up heads with probability p, and tails with probability (1-p).
* The probability to observe n heads after M throws is given by equation (C6).
* In the limit when both n and M are very large, we have seen that the binomial distribution converges to the Gaussian distribution with average μ1=Mp and variance σ2=Mp(1−p).
* We will see that this is a trivial application of the central limit theorem.
* It works, because the result of each coin toss is given by an independent random process with a well defined average and variance.

**Note 3**

* While the proof for the equivalence has been given in the supplemental note, let us examine this by performing numerical experiments for various values of M using Python.
* As always, we begin by importing the necessary numerical and graphical libraries

**Note 4**

* In this code example, we assume we are using a fair coin, such that p = 1-p = 0.5.
* The experiment consists in performing M coin tosses and counting the number of heads.
* To get reliable statistics we repeat this for N = 100,000 times, and calculate the histogram of the data.
* The sampling is performed in one line, using numpy's built-in binomial function, and the sampled data is stored as an array X of size N.
* Remember, each element of the X array contains the number of heads after M coin-tosses.
* We plot the histogram using the "hist" function.
* Finally, we compare the results of our experiment with the corresponding Gaussian distribution shown in Eq.(C1).
* For this, we generate an array of x values in a range 5 standard deviations to the right and left of the average value, and calculate the theoretical Gaussian distribution over this array.
* If you run this code example, the results are plotted and overlayed on the histogram.
* You should repeat this experiment for several values of M=1,2,4,10,100, and 1000, to see how the histogram converges towards the Gauss distribution.
* The convergence is assured in the limit of large n and M.
* In practice, you can know how large should M be for the two distributions to match.

**Note 5**

* Let us discuss what we have seen in the experiments.
* First we define a stochastic variable "s" which is a result of single binary choice from 0 or 1.
* The total number of heads, after M tosses, is given by the sum of all "s" from the 1st to the M-th choices.
* This defines a new stochastic variable nM.
* First, consider the case of a single coin toss for M=1, for which n of M is the same as a single binary choice s as shown in Eq.(D1).
* Since this random variable is drawn from a binomial distribution, the average and variance are given by equations (D2) and (D3).
* Namely, the average number of heads, from one coin toss is equal to p, and the variance is equal to p\*(1-p).
* Now, consider the case for arbitrary M.
* As we already discussed above, n of M is given by the sum all "s" up to the M-th choices, or equivalently as the sum of M independent single coin toss results n of M=1 as shown in Eq.(D4).
* In the limit when M is very large, we have already proven that this converges to a Gaussian.
* The average and variance of this distribution, is just the average and variance of the distribution for a single coin toss multiplied by M, shown in Eqs.(D5) and (D6).

**Note 6**

* The previous relation between binomial and Gaussian distributions is in fact an example of the Central Limiting Theorem.
* It is not valid only for the sum of random variables drawn from binomial distributions, but is applicable to any distributions with finite variance.
* Thus, if we have a set of M independent random variables, n of M=1, with average and variance μ1 and σ2 of M=1.
* The sum of these random variables is itself a random variable, and in the limit when M is very large, it converges to a Gaussian distribution.
* The average and variance of n of M >> 1, which is the sum of the M independent random variables, is given by Eqs.(D8) and (D9) which are generalizations of Eq.(D5) and (D6).
* The average and variance of this sum are then M times the average and variance of the individual distribution.

**Note 7**

* One of the most useful forms of the CLT appears when considering the average of a series of stochastic variables, not just the sum.
* As before, let "n\_j" of M=1 be a series of stochastic variables.
* We assume they have average and variance of μ1 and σ2 of M=1.
* They should be independent and identically distributed, but other that that they can be drawn from any distribution with a finite variance.
* The average of these stochastic variables is given by the weighted sum in Eq.(D10).
* Apart from the factor of 1/M, this is exactly the same as we considered in the previous example.
* Now, we know that in the limit of very large M, the distribution of this variable converges to a Gaussian.
* However, because of the 1/M factor in Eq.(D10), the average and variance are now given by Eq.(D11) and (D12).
* Notice that the average value of n of M is now equal to the average value of the n of M=1.
* More importantly, the variance of n of M is the variance of n of M=1 divided by M.

**Note 8**

* Let us see another example of the CLT in practice.
* Now, instead of a binomial distribution, let us consider a continuous uniform distribution.
* We define a stochastic variable "x" of M=1, which is uniformly distributed between the unit interval.
* Thus, the probability density of observing any value x is constant and equal to one, if x lies in the unit interval, and zero if it is outside.
* The average and the variance can be easily calculated analytically, and are given by 1/2 and 1/12 respectively.
* Now, define a new stochastic variable x of M, which is the sum of M of these uniform random variables.
* In the limit when M is very large, x of M is described by a Gaussian distribution with average and mean given by Eqs. (D8) and (D9).
* In this case, the average is just M/2 and the variance M/12.

**Note 9**

* In this numerical experiment, we will verify that the sum of uniformly distributed random numbers converges to a Gaussian distribution with average and variance defined in Eqs.(D17) and (D18).
* In the code example, we first set M, the number of random variables to add.
* You should try with several values of M, such as M=1, 2, 4, 10, and 100 to see how the distribution converges to a Gaussian.
* Thus, one experimental run consists of generating M random variables, and adding them to generate a sample for the cumulative variable.
* We store these in the array X.
* To obtain reliable statistics, we repeat this experiment N = 100,000 times.
* Then, we plot the histogram of X, together with the corresponding theoretical Gaussian distribution.
* The two should be equal in the limit when M is very large.
* In practice, how many uniform random variables must you add to obtain a Gaussian random variable?
* As you should see for yourselves, adding just 10 uniform random variables is already enough to reproduce a Gaussian distribution.