This assignment is due **August 23, 2024**. You may work in teams of up to four people. If you work in teams, make sure you include everyone's name on the assignment and there is only one submission per team. Please submit one html or pdf file. Show all your work. Use as few Python packages as possible to complete the coding portion of this assignment. Email your solutions to **charles.tutoring@gmail.com**.

1. We will create a toy model to illustrate how difficult it is to determine an investor's skill. This model assumes that investors are either skilled or unskilled, each has control of an equal amount of capital, and that market beats are independent. Let us say that skilled investors have a 55% chance of beating the market, and unskilled investors have a 46% chance of beating the market. Use the class `binom` within `scipy.stats` to solve this problem.

(a) Because of transaction costs and fees, the average investor should beat the market slightly less than 50% of the time. If we assume that the average investor beats the market 47% of the time, what fraction of investors must be skilled?

(b) Given that an investor beat the market last year, use Bayes' rule to determine the probability that she is skilled.

(c) Use Bayes' rule to determine the probability that an investor is skilled given that she beats the market in at least two of the last three years.

(d) Verify (c) using a simulation. Suppose there are 1,000,000 investors and determine the number of skilled and unskilled investors. Then simulate the number of times each investor beats the market over the last three years for each skill level. Use these results to find the ratio of skilled investors that beat the market at least twice to the total number of investors that beat the market at least twice.

(e) Using either Bayes' rule or a simulation create a graph of the number of years examined to the probability of being skilled. Let the horizontal axis represent the number of years n and consider $2 \leq n \leq 10$. Then use `plot` in `matplotlib.pyplot` to graph the probabilities of being skilled given $n-2$, $n-1$, and n market beats in the last n years. Add a legend to help distinguish your three plots.

2. Assume that R is the random variable representing the monthly return of a stock and $X = \ln(1+R) \sim \mathcal{N}(\mu, \sigma^2)$, i.e. the corresponding continuous return X follows a normal distribution.

(a) Suppose $X_k \sim \mathcal{N}(\mu, \sigma^2)$ for $k = 1, 2, \ldots, 12$. Assume X_i and X_j are independent for $i \neq j$. Because you can add continuous returns to find the total return over multiple periods, the total return over a twelve-month period is

$$Y = X_1 + X_2 + \ldots + X_{12}.$$

Find formulas for $E[Y]$ and $\mathrm{Var}(Y)$ using our assumptions about X_k.

Month	April	May	June
Return	-4%	5%	3%

We would like to determine the annualized mean and standard deviation using the monthly returns in the table. The most common way to annualize monthly returns and variances is via $\bar{R} \mapsto 12 \cdot \bar{R}$ and $s^2 \mapsto 12 \cdot s^2$, where $\bar{R}$ and s^2 denote the respective sample mean and variance of R. This is not the best method from a mathematical perspective because it ignores compounding. However, this technique is more reasonable for continuous returns due to (a). With that in mind, given the discrete expected return and variance, we can calculate μ and σ^2, annualize them, and then convert them to discrete annual returns. To accomplish this, note that

$$E[R] = E\left[e^X\right] - 1 = e^{\mu+\sigma^2/2} - 1 \qquad \text{and} \qquad \mathrm{Var}(R) = \mathrm{Var}(e^X) = (e^{\sigma^2} - 1)e^{2\mu+\sigma^2}.$$

(b) Use the data in the table to estimate $E[R]$ and $\mathrm{Var}(R)$ using sample statistics. Then use algebra to find the corresponding estimates of μ and σ^2. Call the respective estimates $\hat{\mu}$ and $\hat{\sigma}^2$. Annualize $\hat{\mu}$ and $\hat{\sigma}^2$ via $\hat{\mu} \mapsto 12 \cdot \hat{\mu}$ and $\hat{\sigma}^2 \mapsto 12 \cdot \hat{\sigma}^2$.

(c) Calculate the annualized discrete return and variance using your annualized results from (b).

3. Within the Black-Scholes framework, the continuous return over T periods of a non-dividend paying stock is $X_T \sim \mathcal{N}\left((r - \frac{\sigma^2}{2})T, T\sigma^2\right)$, where r is the continuous risk-free rate of return. If the initial price of the stock is S_0, it follows that the price at time T is the random variable $S_T = S_0 e^{X_T}$. Let us suppose $S_0 = 90$, $T = 2$, $r = 4\%$, and $\sigma = 45\%$.

(a) Generate 100,000 values of X_T using `scipy.stats.norm.rvs`. Use `np.random.seed` to set the random seed to a particular number to maintain reproducibility. Then find the corresponding values of S_T using the formula above. Use `plt.subplots` to create a figure with subplots of the histograms of X_T and S_T. For both histograms, set `bins = 20` and `density = True`. Place a title above each subplot.

(b) We can use your results in part (a) to price a European call option for this stock. This option gives the owner the right, but not the obligation, to convert to equity for a price of K at maturity. The fair value of this option—given our assumptions—is

$$C = e^{-rT} E\left[\max(S_T - K, 0)\right].$$

Compute the payoff $\max(S_T - K, 0)$ for each of the values of S_T from (a). Assume $K = 100$ in your calculations. Take the mean to estimate the expected payoff. Lastly, discount the expected payoff to the present value using e^{-rT}.

(c) The Black-Scholes formula for the value of a European call option is

$$C = S_0\Phi(d_1) - Ke^{-rT}\Phi(d_2),$$

where

$$d_1 = \frac{\ln\left(\frac{S_0}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}} \qquad \text{and} \qquad d_2 = d_1 - \sigma\sqrt{T}.$$

Use the formula to verify your result in (b). Because of the stochastic nature of the calculation in (b), your result in (c) may not be exactly the same.