BUFN402 Homework 1

James Zhang*

February 18, 2024

#1. (Large Deviation/Extreme events, 10 pts) Consider an experiment that you a fair coin for 1000 times. Let the number of heads from the experiment be S_n. Answer the following questions.

a) Moderate deviation (6 pts):
   a. What is the **exact** probability of $\mathbb{P}(485 \le S_n \le 516)$ not using the central limit theorem (CLT)?
   b. What is the approximated probability using the CLT?
   c. What is the percentage error between results a) and b)?

b) Large deviation (tail/extreme behavior) [2 pts]:
   a. What is the **exact** probability of $\mathbb{P}(S_n \ge 564)$ not using the CLT? Note that $\lfloor \mu + 4\sigma \rfloor = 564$.
   b. What is the approximated probability using the CLT?
   c. What is the percentage error between them?

c) [2 pts] Are the results expected from the CLT? If so, explain. If not, explain why not. Complete proof is not required. Simple statements are enough.

Solution.

a. a. Note that we can model repeated coin flips using a binomial random variable. Let us denote this $X \sim \text{Binomial}(n = 1000, p = 0.5)$. The pmf of X is

$$p(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

$$p(x) = \binom{1000}{x} (\frac{1}{2})^{1000}$$

$$\mathbb{P}(485 \le S_n \le 516) = \sum_{i=485}^{516} \binom{1000}{i} (\frac{1}{2})^{1000}$$

$$\mathbb{P}(485 \le S_n \le 516) \approx 0.8516 - 0.1795 = 0.6721$$

b. Note that $\mathbb{E}(X) = np = 500$ and $\text{Var}(X) = np(1-p) = 250 \implies \sigma(X) = \sqrt{250}$. Using CLT

$$\mathbb{P}(485 \le S_n \le 516) = \mathbb{P}(\frac{485 - 500}{\sqrt{250}} \le \frac{S_n - 500}{\sqrt{250}} \le \frac{516 - 500}{\sqrt{250}})$$

*Email: jzhang72@terpmail.umd.edu

$$= \phi(\frac{516 - 500}{\sqrt{250}}) - \phi(\frac{485 - 500}{\sqrt{250}}) = \phi(1.0119) - \phi(-0.9487)$$
$$\approx 0.8442 - 0.1714 = 0.6728$$

c. The percentage error is approximately

$$|\frac{0.6721 - 0.6728}{0.6721}| \approx 0.1402\% \text{ error}$$

b. a. The probability should be extremely low given the fact that it is 4 standard deviations away. The exact probability of $\mathbb{P}(S_n \geq 564)$ is

$$\mathbb{P}(S_n \geq 564) = 1 - \mathbb{P}(S_n < 564) = 1 - 0.99997 = 0.00003$$

b. Using CLT,

$$1 - \mathbb{P}(S_n < 564) = 1 - \mathbb{P}(\frac{S_n - 500}{\sqrt{250}} < \frac{564 - 500}{\sqrt{250}}) = \phi(\frac{564 - 500}{\sqrt{250}})$$
$$= 1 - \phi(4.0477) = 1 - 0.99997 = 0.00003$$

c. There is 0% percentage error between the two calculations.

c. The results are expected from the CLT because the given experiment satisfies the conditions of the CLT. There are a 1000 samples, which is sufficient number of samples and the underlying Binomial distribution has finite variance, which means the CLT can be applied. Thus, this Binomial distribution over the course of experiment converges in distribution to a normal distribution.

□

#2 (Limit of diversification) [20 pt]. In class, we discussed an argument about the diversification effect. You will show the details of the argument. Recall the return of a portfolio (PF) of n assets is

$$R_p = \sum_{i=1}^{n} w_i R_i \equiv w^t R$$

where w_i denotes a weight given to an asset i and R_i a return of an asset i. Assume that the variances of assets are finite: $Var[R_i] < \infty$ for all i. Answer the following questions. Show your work for the full credit.

a) Show that the variance of the PF is

$$Var[R_p] = \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \sigma_{ij} \equiv w^t \Omega w$$

where $\Omega = [\sigma_{ij}]_{ij}$, and $\sigma_{ij} := cov[R_i, R_j]$.

In parts b), c), and d), assume that the portfolio is equally weighted – that is, $w_i = 1/n$ for all $i = \{1, \cdots, n\}$.

b) Suppose $\sigma_{ij} = 0$ for all $j \neq i$ (i.e., all assets are uncorrelated). Then, show that the portfolio becomes risk-free as $n \to \infty$.

c) Suppose $\sigma_{ij} \neq 0$ for all $j \neq i$. Then, show that

$$Var[R_p] = \frac{1}{n}(\text{Average variance of the individual stocks}) + \left(1 - \frac{1}{n}\right)(\text{Average covariance between the individual stocks})$$

Solution.

a. To calculate the total variance of the portfolio R_p,

$$\text{Var}(R_p) = \text{Var}(\sum_{i=1}^{n} w_i R_i)$$

Recall that variance does not follow linearity, and instead follows the property that

$$\text{Var}(aX + bY) = a^2\text{Var}(X) + b^2\text{Var}(Y) + 2ab\text{Cov}(X, Y)$$

For the case of n assets, we write

$$\text{Var}(\sum_{i=1}^{n} w_i R_i) = \sum_{i=1}^{n} w_i^2 \text{Var}(R_i) + 2 \sum_{1 \le i < j \le n} w_i w_j \text{Cov}(R_i, R_j)$$

Recall that $\text{Var}(R_i) = \text{Cov}(R_i, R_i)$ and so we can rewrite it as

$$\text{Var}(R_p) = \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \text{Cov}(R_i, R_j) = \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \sigma_{ij}$$

which is the first defined form. Note that this equivalent to the quadratic form, by definition of quadratic form, and I will show this by considering the vector of weights and the covariance matrix.

$$\text{Var}(R_p) = w^T \Omega w = \begin{bmatrix} w_1 & \cdots & w_n \end{bmatrix} \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1n} \\ \vdots & \ddots & \vdots \\ \sigma_{n1} & \cdots & \sigma_{nn} \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix}$$

$$= \begin{bmatrix} w_1 & \cdots & w_n \end{bmatrix} \begin{bmatrix} w_1\sigma_{11} + \cdots + w_n\sigma_{1n} \\ \vdots \\ w_1\sigma_{n1} + \cdots w_n\sigma_{nn} \end{bmatrix} = \begin{bmatrix} w_1 & \cdots & w_n \end{bmatrix} \begin{bmatrix} \sum_{i=1}^n w_i\sigma_{1i} \\ \vdots \\ \sum_{i=1}^n w_i\sigma_{ni} \end{bmatrix}$$

$$= w_1 \sum_{i=1}^n w_i\sigma_{1i} + \cdots + w_n \sum_{i=1}^n w_i\sigma_{ni} = \sum_{i=1}^n \sum_{j=1}^n w_i w_j \sigma_{ij}$$

as desired.

b. Assuming equal weightings in all assets and covariances between all assets are exactly 0, then

$$\text{Var}(R_p) = \text{Var}(\sum_{i=1}^n \frac{1}{n} R_i) = \frac{1}{n^2}\text{Var}(\sum_{i=1}^n R_i)$$

by a property of covariance. Now, taking the limit as $n \to \infty$,

$$\lim_{n\to\infty} \frac{1}{n^2} \sum_{i=1}^n \text{Var}(R_i) = 0$$

which means the portfolio essentially becomes risk-free, as desired.

c. Now, assuming equal weightings in all assets but nonzero covariances,

$$\text{Var}(R_p) = \sum_{i=1}^n \frac{1}{n^2}\text{Var}(R_i) + 2 \sum_{1\le i<j\le n} \frac{1}{n^2}\sigma_{ij}$$

$$\text{Var}(R_p) = \left(\frac{1}{n}\frac{1}{n} \sum_{i=1}^n \text{Var}(R_i) \right) + \left(2 \sum_{i=1}^n \sum_{1\le i<j\le n} \frac{1}{n^2}\sigma_{ij} \right)$$

$$\text{Var}(R_p) = \frac{1}{n}(\text{Average Variance}) + \frac{2}{n^2} * \binom{n}{2} * \frac{\sum_{i=1}^n \sum_{1\le i<j\le n}^n \sigma_{ij}}{\binom{n}{2}}$$

where the N choose 2 follows from the fact that there are $\binom{N}{2}$ combinations of covariances between 2 assets. Note that therefore, the last term is simply the average covariance. Thus,

$$\text{Var}(R_p) = \frac{1}{n}(\text{Average Variance}) + \frac{2}{n^2} * \frac{n(n-1)}{2}(\text{Average Covariance})$$

$$V(R_p) = \frac{1}{n}(\text{Average Variance}) + (1 - \frac{1}{n})(\text{Average Covariance})$$

as desired.

□

#3 (Data manipulation) [20 pt]. Access WRDS and download the daily returns from 12/19/2010 – 6/28/2020 of (i) HSBC Holdings, (ii) Microsoft, (iii) Pfizer, (iv) Exxon Mobil, and (v) SP500 index.

a) Convert daily returns to weekly returns.

b) What are the sample average and standard deviation (volatilities) of weekly returns? Do the patterns of the sample averages and volatilities agree with or against your intuition? Justify your answer. Note that the firms are sampled from different industries.

c) Assume that the true return distributions are normal. What is the 99% confidence interval for the true mean of returns of SP500?

d) Assume that the true return distributions are normal. Test the null hypothesis if the true mean of returns of SP500 is equal to 0. Given a 1% significance level, what is your recommendation?

Solution.
Solution is written in the attached notebook.

□

#4 (Simplified random walk)[1] [20 pt]. Brownian motion (BM, henceforth) has many applications in finance, including option pricing (Black-Sholes), fixed income, and real option. Let's construct a simple random walk. First, divide the time interval $[0, t]$ into n subintervals (equal length δ) such that $t = n\delta$. Each step X_i is an independent random variable such that

$$X_i = \begin{cases} \sqrt{\delta} & \text{with probability } 1/2 \\ -\sqrt{\delta} & \text{with probability } 1/2 \end{cases}$$

After time t, the location of the particle is

$$W_t = \sum_{i=1}^{n} X_i$$

a) What are $\mathbb{E}[X_i]$ and $\text{Var}[X_i]$?

b) What are $\mathbb{E}[W_t]$ and $\text{Var}[W_t]$?

c) What is the (asymptotic) distribution of the random location W_t as $n \to \infty$?

d) Based on the answer c), what's the probability that W_1 (the random location at time $t = 1$) falls below -1.5?

Solution.

a. $\mathbb{E}(X_i) = \frac{1}{2}\sqrt{\delta} + \frac{1}{2}(-\sqrt{\delta}) = 0$
   $\text{Var}(X_i) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \mathbb{E}(X^2) - 0 = \frac{1}{2}\delta + \frac{1}{2}\delta = \delta$

b. $\mathbb{E}(W_t) = \mathbb{E}(\sum_{i=1}^n X_i) = \sum_{i=1}^n \mathbb{E}(X_i) = 0$
   $\text{Var}(W_t) = \text{Var}(\sum_{i=1}^n X_i) = \sum_{i=1}^n \text{Var}(X_i)$, since note that each X_i is not scaled by a scalar and each X_i is also an independent random variable, meaning $\text{Cov}(X_i, X_j) = 0 \ \forall \ i \neq j$. Therefore,

$$Var(W_t) = n\delta$$

c. Note that t is fixed such that $t = n\delta$. Therefore, as $n \to \infty \implies \delta \to 0$. Thus, by the Central Limit Theorem, considering that we are interested in the asymptotic distribution of the sum and not a standardized sum, the asymptotic distribution $W_t = \sum_{i=1}^n X_i$ is

$$W_t \sim N(0, n\delta) = N(0, t)$$

d. At time $t = 1, W_t \sim N(0, 1)$, so we just seek

$$\phi(-1.5) \approx 0.0668$$

e. Shown in attached notebook.

f. Shown in attached notebook.

g. Shown in attached notebook.

h. Shown in attached notebook.

$\square$

All parts of problem 5 are answered in the attached notebook.