Econ 432 Homework 3

Richard Grigorian (UID: 505-088-797)

2/17/23

Contents

1 Part I: Review Questions

1.1 Problem 1

Let Y_1, Y_2, Y_3, and Y_4 be iid(μ, σ^2). Let $\bar{Y} = \frac{1}{4}\sum_{t=1}^{4} Y_t$.

Part (a)

What are the expected value and variance of $\bar{Y}$?

Solution

We can simply apply the linear proprerty of expectations to find:

$$\mathbb{E}[\bar{Y}] = \mathbb{E}\left[\frac{1}{4}\sum_{t=1}^{4} Y_t\right] = \frac{1}{4}\sum_{t=1}^{4}\mathbb{E}[Y_t] = \frac{1}{4}4\mu = \mu$$

We can also find the variance of the estimator using

$$\text{Var}(\bar{Y}) = \frac{\sigma^2}{n} = \frac{\sigma^2}{4}$$

Part (b)

Now, consider a different estimator of μ:

$$W = \frac{1}{8}Y_1 + \frac{1}{8}Y_2 + \frac{1}{4}Y_3 + \frac{1}{2}Y_4$$

What are the expected value and variance of W?

Solution

Once again, we can simply apply the linear property of expectations to show that:

$$\mathbb{E}[W] = \mathbb{E}\left[\sum_{t=1}^{4} a_t Y_t\right] = \sum_{t=1}^{4} a_t \mathbb{E}[Y_t]$$

where a_t is weights on each Y_t. Substituting for our weights we get

$$\mathbb{E}[W] = \frac{1}{8}\mu + \frac{1}{8}\mu + \frac{1}{4}\mu + \frac{1}{2}\mu = \mu$$

Now for variance of our estimator, we apply the variance property of addition to get

$$\text{Var}(W) = \left(\frac{1}{8}\right)^2 \text{Var}(Y_1) + \left(\frac{1}{8}\right)^2 \text{Var}(Y_2) + \left(\frac{1}{4}\right)^2 \text{Var}(Y_3) + \left(\frac{1}{2}\right)^2 \text{Var}(Y_4)$$

where we would typically have a series of covariance terms; however, since the variables are iid, then they are independent and hence the covariances are 0. With this we express the above as

$$\text{Var}(W) = \frac{\sigma^2}{64} + \frac{\sigma^2}{64} + \frac{\sigma^2}{16} + \frac{\sigma^2}{4} = \frac{11}{32}\sigma^2$$

Part (c)

Which estimator μ do you prefer? Fully justify your answer.

Solution

Both estimators $\bar{Y}$ and W are unbiased since the their expectations are equal to μ. Therefore, to minimize Mean Squared Error, we want the estimator with the smaller variance. Between the two estimators, $\bar{Y}$ has a smaller variance since $\frac{1}{4}\sigma^2 < \frac{11}{32}\sigma^2$. Therefore, we prefer $\bar{Y}$ as our estimator of μ.

1.2 Problem 2

Let $Y_1, Y_2, Y_3, \dots, Y_n$ be iid (μ, σ^2) and let $\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$.

Part (a)

Define the class of *linear estimator* of μ by

$$W_a = \sum_{i=1}^{n} a_i Y_i$$

where a_i' s are constants. What restriction on the a_i' s is needed for W_a to be an unbiased estimator of μ?

Solution

W_a is an unbiased estimator of μ if

$$\sum_{i=1}^{n} a_i = 1$$

In other words, the weights of the random variables need to sum to 1.

Part (b)

Find Var(W_a).

Solution

The variance of our estimator is simply

$$\text{Var}(W_a) = \sum_{i=1}^{n} a_i^2 \, \text{Y}_\text{i} = \sum_{i=1}^{n} a_i^2 \sigma^2$$

Part (c)

For any numbers a_i, $i = 1, \ldots, n$, the following inequality holds

$$\left(\sum_{i=1}^{n} a_i \right)^2 \leq n \sum_{i=1}^{n} a_i^2$$

Use this (and above results) to show that $\bar{Y}$ is the *best linear unbiased estimator* (BLUE).

Solution

Recall that $\bar{Y}$ is an evenly weighted average of all of our Y_i. Notably, the sum of its weights a_i is equal to one. Hence we call on the variance expression for $\bar{Y}$ found earlier and use the Markov Inequality to prove the above claim.

$$\begin{aligned}
\text{Var}(\bar{Y}) &= \frac{\sigma^2}{n} \\
&= \frac{1}{n} \left(\sum_{i=1}^{n} a_i \right)^2 \sigma^2 \\
&= \frac{1}{n} (1)^2 \sigma^2 \\
&\leq \left(\sum_{i=1}^{n} a_i^2 \right) \sigma^2 = \text{Var}(W_a)
\end{aligned} \qquad \blacksquare$$

Hence, using the Markov Inequality, we have shown that $\bar{Y}$ has a variance that is less than or equal to that of W_a despite both being unbiased. Hence, it is the *best linear unbiased estimator* of μ.

1.3 Problem 3

Consider the constant expected return model

$$r_{it} = \mu_i + \varepsilon_{it} \quad t = 1, \ldots, T; \quad i = 1\ (GS),\ 2\ (AIG)$$

$$\varepsilon_{it} \sim \text{iid } N(0, \sigma_i^2), \quad \text{cov}(\varepsilon_{1t}, \varepsilon_{2t}) = \sigma 12, \quad \text{cor}(\varepsilon_{1t}, \varepsilon_{2t}) = \rho_{12}$$

for the monthly cc returns on GS (Goldman Sachs) and AIG (American International Group). The estimates (rounded for computations) are given (T = 100):

	GS	AIG		GS & AIG
$\hat{\mu}_i$	0.01	-0.03	$\hat{\sigma}_{12}$	0.01
$\hat{\sigma}_i$	0.1	0.3	$\hat{\rho}_{12}$	0.4

Part (a)

For both GS and AIG cc returns, compute (asymptotic) 95% CI for μ_i and σ_i^2.

Solution

The asymoptotic 95% CI for μ_i is formulated as:

$$\begin{aligned} &\hat{\mu}_i \pm t_c \cdot \text{SE}(\hat{\mu}_i) \\ &\pm 1.96 \cdot \frac{\hat{\sigma}_i}{\sqrt{T}} \end{aligned}$$

We can then substitute in our values from the above table to get

$$95\%\ \text{CI}\ [\mu_{GS}] : \quad 0.01 \pm 1.96 \cdot \frac{0.1}{\sqrt{100}} = [-0.0096\ ,\ 0.0296]$$

$$95\%\ \text{CI}\ [\mu_{AIG}] : \quad -0.03 \pm 1.96 \cdot \frac{0.3}{\sqrt{100}} = [-0.0888\ ,\ 0.0288]$$

In a similar fashion, we can constructed the 95% CI for σ_i^2:

$$\begin{aligned} &\widehat{\sigma^2}_i \pm t_c \cdot \text{SE}(\hat{\sigma}_i^2) \\ &\pm 1.96 \cdot \frac{\sqrt{2}\,\hat{\sigma}_i^2}{\sqrt{T}} \end{aligned}$$

We can then substitute our values from the above table once more to get

$$95\% \text{ CI } [\sigma^2_{GS}]: \quad (0.1)^2 \pm 1.96 \cdot \frac{\sqrt{2}\,(0.1)^2}{\sqrt{100}} = [0.0072\ ,\ 0.0127]$$

$$95\% \text{ CI } [\sigma^2_{AIG}]: \quad (0.3)^2 \pm 1.96 \cdot \frac{\sqrt{2}\,(0.3)^2}{\sqrt{100}} = [0.0651\ ,\ 0.1149]$$

Part (b)

Compute the (asymptotic) 95% confidence interval for ρ_{12}.

Solution

We will use the fact that SE $(\widehat{\rho}_{12}) = \sqrt{\frac{1-\widehat{\rho}_{12}}{T}}$ to construct our interval as:

$$\begin{aligned} 95\% \text{ CI } [\rho_{12}] : \widehat{\rho}_{12} \pm t_c \cdot \text{SE}(\widehat{\rho}_{12}) \\ \pm 1.96 \cdot \sqrt{\frac{1-\widehat{\rho}_{12}}{T}} \\ 0.4 \pm 1.96 \cdot \sqrt{\frac{1-0.4}{100}} &= [0.2204\ ,\ 0.5796] \end{aligned}$$

Part (c)

Test the hypothesis (significance tests) for $i = 1, 2$ with 5% significance level,

$$H_0 \ : \ \mu_i = 0 \quad \text{v.s.} \quad H_1 \ : \ \mu_i \neq 0$$

Are expected returns of these assets (statistically) different from zero? Justify your answer.

Solution

We compute the t-statistic:

$$t_{GS} = \left| \frac{\widehat{\mu}_{GS} - \mu_i}{\text{SE}(\widehat{\mu}_{GS})} \right| = \frac{0.01}{0.01} = 1 < 1.9842 = q^{T(99)}_{0.975}$$

$$t_{AIG} = \left| \frac{\widehat{\mu}_{AIG} - \mu_i}{\text{SE}(\widehat{\mu}_{AIG})} \right| = \frac{0.03}{0.03} = 1 < 1.9842 = q^{T(99)}_{0.975}$$

Hence, we **fail to reject the null for both GS and AIG**. We cannot say that expected returns for these assets are statistically significantly different from zero since their t-statistics are smaller than the t-critical values.

Part (d)

Test the hypothesis (significance tests) for $i = 1, 2$ with 5% significance level,

$$H_0 \ : \ \sigma_i^2 = 0.0225 \quad \text{v.s.} \quad H_1 \ : \ \sigma_i^2 \neq 0.0225$$

Solution

By a very similar process, we compute the t-statistics:

$$t_{GS} = \left| \frac{\hat{\sigma}_{GS}^2 - \sigma_i^2}{\sqrt{2}\hat{\sigma}_{GS}^2/\sqrt{T}} \right| = \frac{(0.1)^2 - 0.0225}{\sqrt{2}(0.1)^2/10} = 8.84 > 1.96 = q_{0.975}^Z$$

$$t_{AIG} = \left| \frac{\hat{\sigma}_{AIG}^2 - \sigma_i^2}{\sqrt{2}\hat{\sigma}_{AIG}^2/\sqrt{T}} \right| = \frac{(0.3)^2 - 0.0225}{\sqrt{2}(0.3)^2/10} = 5.30 > 1.96 = q_{0.975}^Z$$

Therefore, we **reject the null for both GS and AIG**. Since the t-stats are greater than the necessary critical scores, we say that the variance for these assets is statistically significantly different from 0.0225.

1.4 Problem 4

Suppose we have an i.i.d sample ${X_1, …, X_T } with mean μ and finite varaince σ^2. Show that the sample variance

$$\hat{\sigma}^2 = (T-1)^{-1} \sum_{t=1}^{T} (X_t - \hat{\mu})^2$$

is an unbiased estimator of σ^2 where $\hat{\mu} = T^{-1} \sum_{t=1}^{T} X_t$.

Solution

To prove unbiasedness, we want to show that $\mathbb{E}[\hat{\sigma}^2] = \sigma^2$.

Proof

Using the expectation definition of variance $\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$, we re-write our variance expression as

$$\hat{\sigma}^2 = \frac{T}{T-1} \left(\frac{1}{T} \sum_{t=1}^{T} X_t^2 - \hat{\mu}^2 \right)$$

Then, by properties of our random variable X_t.

$$\mathbb{E}[\hat{\mu}^2] = \text{Var}(\hat{[}\mu]) + \mathbb{E}[\hat{\mu}]^2 = \frac{\sigma^2}{T} + \mu^2$$

We can now take the expectation of our sample variance expression to show that

$$
\begin{aligned}
\mathbb{E}[\widehat{\sigma}^2] &= \frac{T}{T-1}\mathbb{E}\left[\frac{1}{T}\sum_{t=1}^{T} X_t^2 - \widehat{\mu}^2\right] \\
&= \frac{T}{T-1}\left(T^{-1}\sum_{t=1}^{T}\mathbb{E}[X_i^2] - \mathbb{E}[\widehat{\mu}^2]\right) \\
&= \frac{T}{T-1}\left(\mathbb{E}[X_i^2] - \left(\frac{\sigma^2}{T} + \mu^2\right)\right) \\
&= \frac{T}{T-1}\left(\mu^2 + \sigma^2 - \frac{\sigma^2}{T} - \mu^2\right) \\
&= \frac{T}{T-1}\left(\frac{T\sigma^2 - T}{T}\right) \\
\mathbb{E}[\widehat{\sigma}^2] &= \sigma^2 \qquad \blacksquare
\end{aligned}
$$

1.5 Problem 5

Suppose that $\widehat{\theta}$ is an estimator of θ constructed using a random sample with sample size T. Which of the following statements are correct?

Part (a)

If $Bias(\widehat{\theta}, \theta) \to 0$ as $T \to \infty$, then $\widehat{\theta}$ is a consistent estimator of θ.

Solution

False! Essentially, we want to show that there exists an estimator that is unbiased but that is also inconsistent. An example of this is the first observation X_1 of $\{X_1, \dots, X_T\} \sim$ i.i.d. $N(\mu, 1)$. If we use this to estimate the unkown parameter μ, then we will define $\widehat{\theta} \equiv X_1$ and $\theta \equiv \mu$. Hence, with this example, it is clear that the expectation of X_1 is just μ. So it is unbiased. However, it is also inconsistent since

$$\lim_{T\to\infty} \mathbb{P}(|\widehat{\theta} - \theta| > \varepsilon) \neq 0 \ \forall \varepsilon > 0$$

Part (b)

If $Var(\widehat{\theta}) \to 0$ as $T \to \infty$, then $\widehat{\theta}$ is a consistent estimator of 0.

Solution

False! Using the same example as in part (a). If we let $\hat{\theta} \equiv X_1/T$, then $Var(\hat{\theta}) \to 0$ as $T \to \infty$. This is clear since

$$Var(\hat{\theta}) = \frac{Var(X_1)}{T^2}$$

which will converge to 0 as $T \to \infty$. However, $\hat{\theta}$ is not a consistent estimator of $\theta = \mu$ unless $\mu = 0$.

Part (c)

If $MSE(\hat{\theta}, \theta) \to 0$ as $T \to \infty$, then $\hat{\theta}$ is a consistent estimator of θ.

Solution

True! By the Markov Inequality

$$\mathbb{P}(|\hat{\theta} - \theta| \geq \varepsilon) \leq \frac{\mathbb{E}[|\hat{\theta} - \theta|^2]}{\varepsilon^2}$$

Since $\mathbb{E}[|\hat{\theta} - \theta|^2] = MSE(\hat{\theta}, \theta)$ and ε is a finite value, we have

$$\mathbb{P}(|\hat{\theta} - \theta| \geq \varepsilon) \to 0 \text{ as } T \to \infty$$

in view of $MSE(\hat{\theta}, \theta) \to 0$ as $T \to \infty$

2 Python Exercises

2.1 Problem 1

Let X and Y be distributed bivariate normal with

$$\mu_X = 0.01,\ \mu_Y = 0.05,\ \sigma_X = 0.25,\ \sigma_Y = 0.15$$

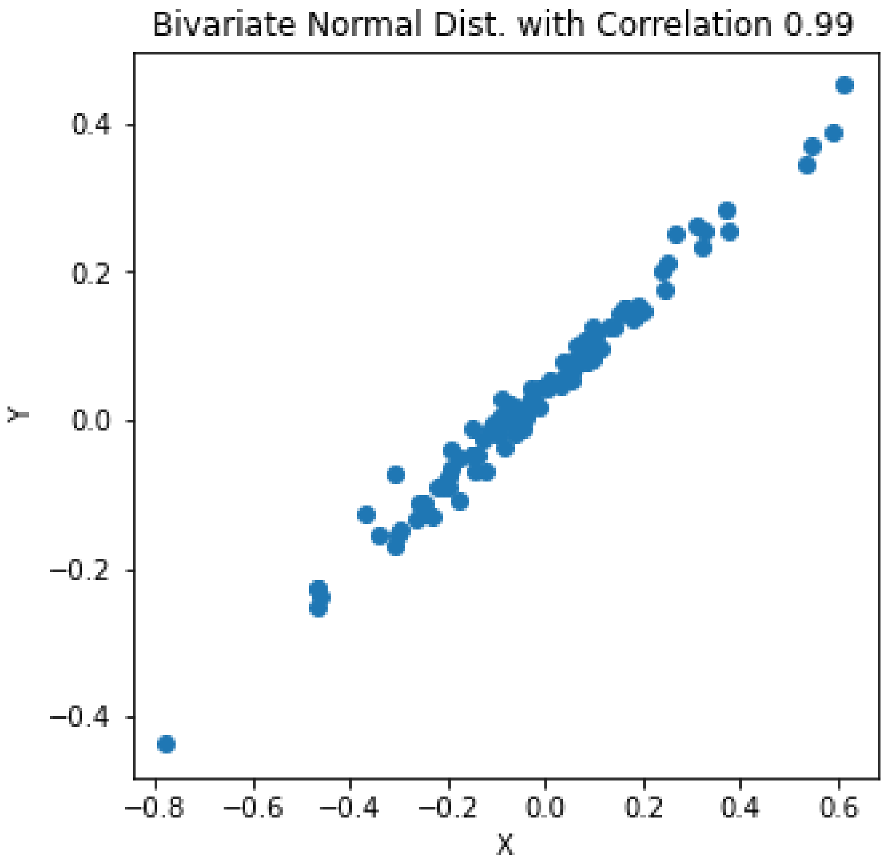
Part (a)

Using Python function `multivariate_normal.rvs()`, simulate 100 observations from the bivariate distribution with $\rho_{XY} = 0.99$. Using `plt.scatter()` function, create a scatterplot of the observations and comment on the direction and strength of the lienar association. Using the function `multivariate_normal.cdf()`, compute the joint probability $\mathbb{P}(X \leq 0, Y \leq 0)$.

Solution

```python
# Imports
import numpy as np
import pandas as pd
import scipy.stats as stats
import matplotlib.pyplot as plt
```

```python
# Simulating Bivariate Normal
mean = [0.01, 0.05]
## Symmetric cov matrix
cov = np.array([[0.25**2, 0.99*0.25*0.15], [0.99*0.25*0.15, 0.15**2]])
n = 100
XY = stats.multivariate_normal.rvs(
    mean = mean,
    cov = cov,
    size = n,
    random_state = 432
)
# Plotting
plt.figure(figsize=(5,5))
plt.scatter(x = XY[:,0], y=XY[:,1])
plt.xlabel('X')
plt.ylabel('Y')
plt.title('Bivariate Normal Dist. with Correlation 0.99')
plt.show()
```



From the above scatterplot, it is clear that X and Y have a **very strong positive** linear association. This is clear because the scatterplot of the two variables forms a near perfect 45 degree line (correlation of 0.99). Hence, the relationship is very strong and positive.

```
# P(X <= 0, Y<=0)
prob = stats.multivariate_normal.cdf(
    x = [0,0], # Vector of upper limits
    mean = mean,
    cov = cov
)
print('Pr(X<=0, Y<=0) = ', prob)
```

```
Pr(X<=0, Y<=0) =  0.36905556559852487
```

Hence, we find that $\mathbb{P}(X \leq 0, Y \leq 0) = 0.369$.

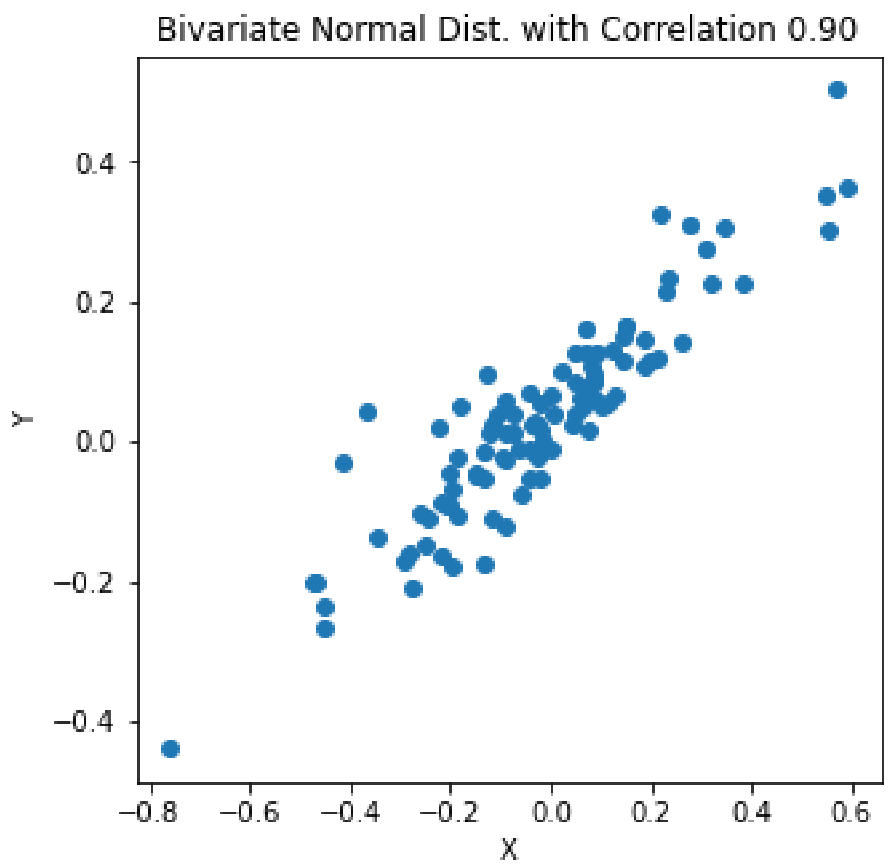
Part (b)

Do the same exercise with $\rho_{XY} = 0.9$.

Solution

We will simply repeat the above steps with this new correlation.

```python
# Simulating Bivariate Normal
mean = [0.01, 0.05]
## Symmetric cov matrix
cov = np.array([[0.25**2, 0.90*0.25*0.15], [0.90*0.25*0.15, 0.15**2]])
n = 100
XY = stats.multivariate_normal.rvs(
    mean = mean,
    cov = cov,
    size = n,
    random_state = 432
)
# Plotting
plt.figure(figsize=(5,5))
plt.scatter(x = XY[:,0], y=XY[:,1])
plt.xlabel('X')
plt.ylabel('Y')
plt.title('Bivariate Normal Dist. with Correlation 0.90')
plt.show()
```



The association is still **very strong and positive linear**; however, it is slightly less strong than before. The lower correlation means that the points will not fall as closely on a 45 degree line.

```
# P(X <= 0, Y<=0)
prob = stats.multivariate_normal.cdf(
    x = [0,0], # Vector of upper limits
    mean = mean,
    cov = cov
)
print('Pr(X<=0, Y<=0) = ', prob)
```

```
Pr(X<=0, Y<=0) =  0.3420967411639703
```

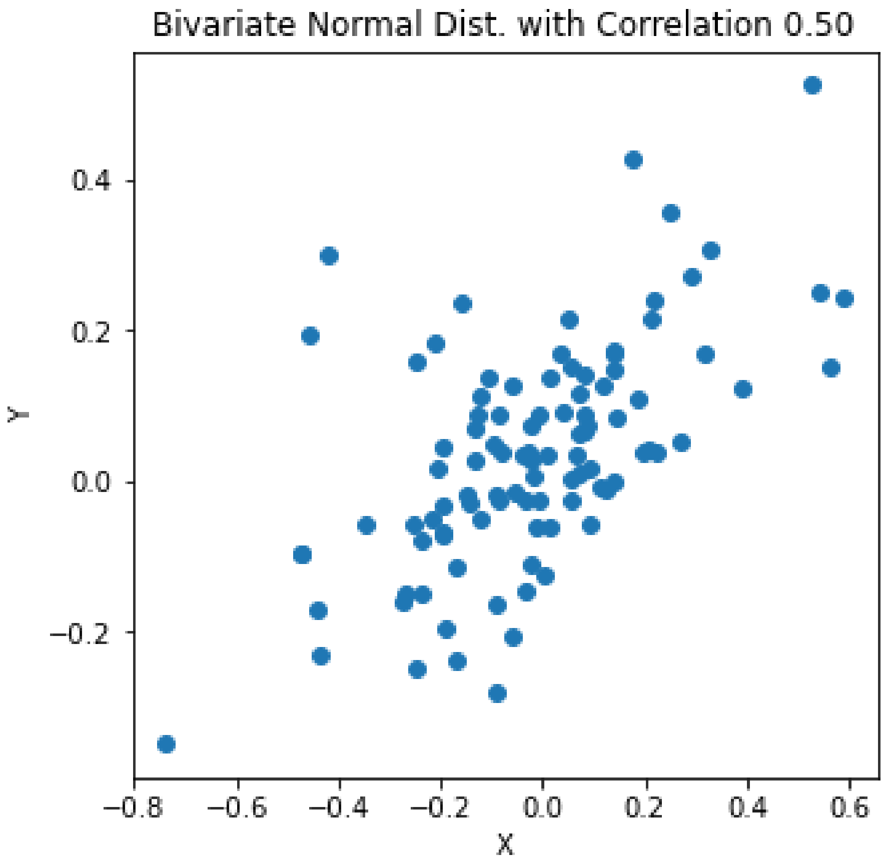
We now find that $\mathbb{P}(X \leq 0, Y \leq 0) = 0.342$ which is smaller than part (a).

Part (c)

Do the same exercise with $\rho_{XY} = 0.5$.

Solution

```python
# Simulating Bivariate Normal
mean = [0.01, 0.05]
## Symmetric cov matrix
cov = np.array([[0.25**2, 0.50*0.25*0.15], [0.50*0.25*0.15, 0.15**2]])
n = 100
XY = stats.multivariate_normal.rvs(
    mean = mean,
    cov = cov,
    size = n,
    random_state = 432
)
# Plotting
plt.figure(figsize=(5,5))
plt.scatter(x = XY[:,0], y=XY[:,1])
plt.xlabel('X')
plt.ylabel('Y')
plt.title('Bivariate Normal Dist. with Correlation 0.50')
plt.show()
```



There is still a **clear positive linear relationship; however, it is much weaker than before!** The scatter is much further away from a perfect 45 degree line. This is due to the even smaller correlation between X and Y.

```
# P(X <= 0, Y<=0)
prob = stats.multivariate_normal.cdf(
    x = [0,0], # Vector of upper limits
    mean = mean,
    cov = cov
)
print('Pr(X<=0, Y<=0) = ', prob)
```

```
Pr(X<=0, Y<=0) =  0.2574488265503332
```

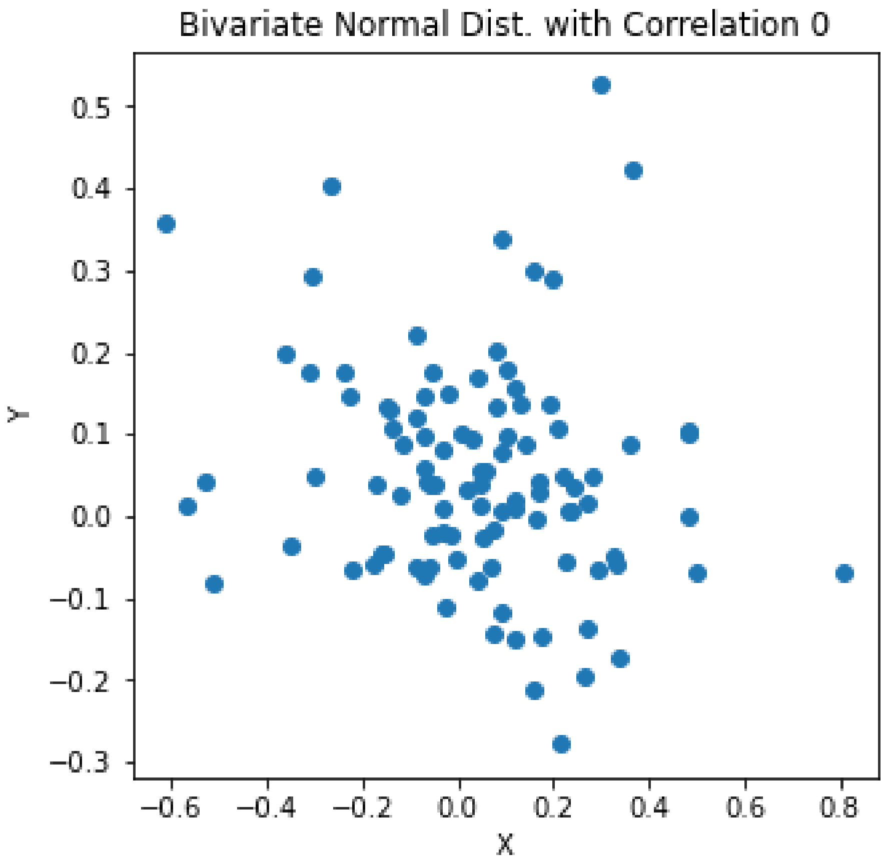
We now find that $\mathbb{P}(X \leq 0, Y \leq 0) = 0.257$ which is smaller than part (a) and (b).

Part (d)

Do the same exercise with $\rho_{XY} = 0$.

Solution

```python
# Simulating Bivariate Normal
mean = [0.01, 0.05]
## Symmetric cov matrix
cov = np.array([[0.25**2, 0*0.25*0.15], [0*0.25*0.15, 0.15**2]])
n = 100
XY = stats.multivariate_normal.rvs(
    mean = mean,
    cov = cov,
    size = n,
    random_state = 432
)
# Plotting
plt.figure(figsize=(5,5))
plt.scatter(x = XY[:,0], y=XY[:,1])
plt.xlabel('X')
plt.ylabel('Y')
plt.title('Bivariate Normal Dist. with Correlation 0')
plt.show()
```



As we now can see, there is **no linear relationship** between X and Y. Since their correlation is 0, the two will be randomly scattered with no clear pattern.

```
# P(X <= 0, Y<=0)
prob = stats.multivariate_normal.cdf(
    x = [0,0], # Vector of upper limits
    mean = mean,
    cov = cov
)
print('Pr(X<=0, Y<=0) = ', prob)
```

```
Pr(X<=0, Y<=0) =  0.17882681099946685
```

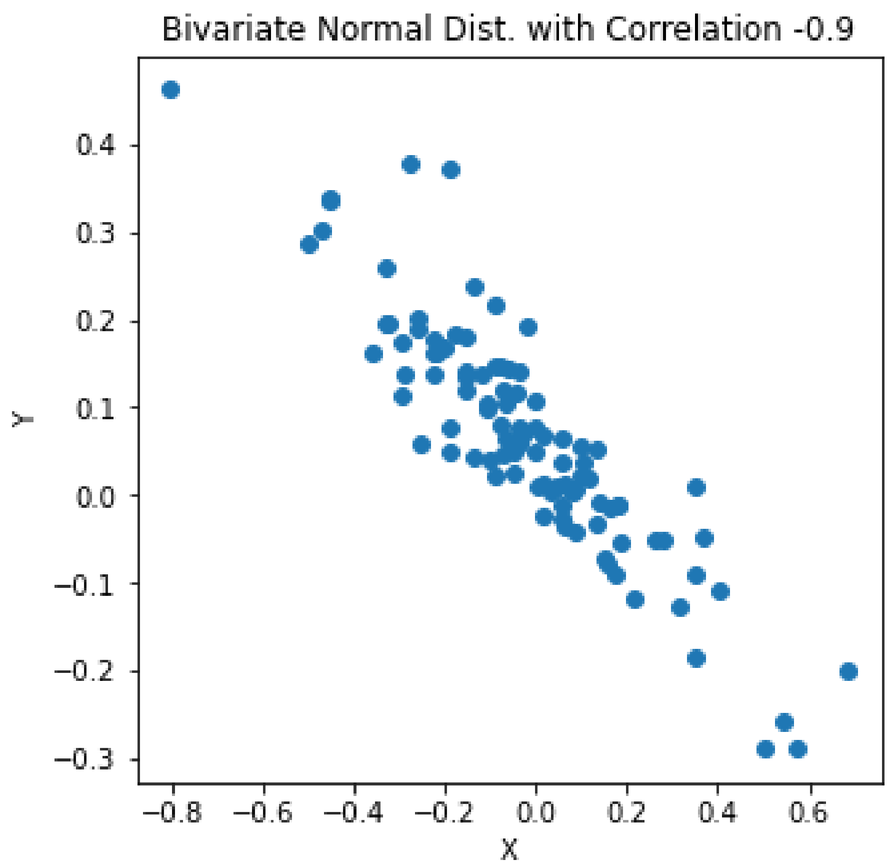
We now find that $\mathbb{P}(X \leq 0, Y \leq 0) = 0.179$ which is smaller than part (a), (b), and (c).

Part (e)

Do the same exercise with $\rho_{XY} = -0.9$.

Solution

```
# Simulating Bivariate Normal
mean = [0.01, 0.05]
## Symmetric cov matrix
cov = np.array([[0.25**2, -0.9*0.25*0.15], [-0.9*0.25*0.15, 0.15**2]])
n = 100
XY = stats.multivariate_normal.rvs(
    mean = mean,
    cov = cov,
    size = n,
    random_state = 432
)
# Plotting
plt.figure(figsize=(5,5))
plt.scatter(x = XY[:,0], y=XY[:,1])
plt.xlabel('X')
plt.ylabel('Y')
plt.title('Bivariate Normal Dist. with Correlation -0.9')
plt.show()
```



It is now clear that there is a **strong negative linear relationship** between X and Y. The high negative correlation causes this result. The above plot shows this by having the scatterplot be downwards sloping around a near negative 45 degree line.

```
# P(X <= 0, Y<=0)
prob = stats.multivariate_normal.cdf(
    x = [0,0], # Vector of upper limits
    mean = mean,
    cov = cov
)
print('Pr(X<=0, Y<=0) = ', prob)
```

```
Pr(X<=0, Y<=0) =  0.020246020323618524
```

We now find that $\mathbb{P}(X \leq 0, Y \leq 0) = 0.0202$ which is smallest of all. This is intuitive since the two variables are negatively correlated. I.e., when one is positive, the other is likely to be negative. Hence, it is very unlikely that both will be less than 0.

2.2 Problem 2: Maximum Likelihood Estimation (MLE)

Download the Microsoft daily price from Jan 02, 2010 to Feb 02, 2021 from *Yahoo* Finance, and perform the following data analysis:

Part (a)

Assume the daily cc returns are i.i.d $N(\mu, \sigma^2)$. Find the maximum likelihood estimators μ and σ, and the standard errors. Do we reject the null hypothesis $H_0 \; : \; \mu = 0$?

Solution

Recall that $\hat{\mu}_{mle}$ and $\hat{\sigma}_{mle}$ are solutions to the log-likelihood estimation problem:

$$\arg\max_{\mu,\sigma} \left\{ \sum_{t=1}^{T} \log f(r_t; \mu, \sigma) \right\}$$

where $f(r_t; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(r_t-\mu)^2}{2\sigma^2}\right)$. We then have

$$\hat{\mu}_{mle} = \frac{1}{T}\sum_{t=1}^{T} r_t$$

$$\hat{\sigma}^2_{mle} = \frac{1}{T}\sum_{t=1}^{T} (r_t - \hat{\mu})^2$$

```
# Load data
df = pd.read_csv('msftPrices.csv')
df.head() # Jan 2. Saturday
```

	Date	Open	High	Low	Close	Adj Close	Volume
0	2010-01-04	30.620001	31.100000	30.590000	30.950001	23.623903	38409100
1	2010-01-05	30.850000	31.100000	30.639999	30.959999	23.631529	49749600
2	2010-01-06	30.879999	31.080000	30.520000	30.770000	23.486511	58182400
3	2010-01-07	30.629999	30.700001	30.190001	30.450001	23.242252	50559700

	Date	Open	High	Low	Close	Adj Close	Volume
4	2010-01-08	30.280001	30.879999	30.240000	30.660000	23.402540	51197400

```
# cc Returns
df['cc'] = np.log(df['Adj Close']/df['Adj Close'].shift(1))
df.head()
```

	Date	Open	High	Low	Close	Adj Close	Volume	cc
0	2010-01-04	30.620001	31.100000	30.590000	30.950001	23.623903	38409100	NaN
1	2010-01-05	30.850000	31.100000	30.639999	30.959999	23.631529	49749600	0.000323
2	2010-01-06	30.879999	31.080000	30.520000	30.770000	23.486511	58182400	-0.006156
3	2010-01-07	30.629999	30.700001	30.190001	30.450001	23.242252	50559700	-0.010454
4	2010-01-08	30.280001	30.879999	30.240000	30.660000	23.402540	51197400	0.006873

```
from scipy.stats import norm
import scipy.optimize as optimize

cc = df['cc'].iloc[1:]

# MLE
def log_likelihood(params, data):
    mu, sigma = params
    # If the standard deviation prameter is negative, return a large value:
    if sigma < 0:
        return(1e8)
    likelihood = norm.pdf(data, loc = mu, scale = sigma)
    return -np.sum(np.log(likelihood[likelihood > 0]))

res = optimize.minimize(fun = log_likelihood,
                        x0 = [0.1, 0.5],  # initial guess
                        args = cc)
print(res)
```

```
     fun: -7573.3790028086105
hess_inv: array([[5.06638606e-08, 5.06249477e-08],
       [5.06249477e-08, 5.07222068e-08]])
     jac: array([0.        , 0.00024414])
 message: 'Desired error not necessarily achieved due to precision loss.'
```

```
    nfev: 66
     nit: 9
    njev: 22
  status: 2
 success: False
       x: array([0.00082328, 0.01601251])
```

```python
# Mean and st. dev of ML estimators
mean, std = res.x

# Computing standard errors using inverse hessian
hess = res.hess_inv
se_mean = np.sqrt(hess[0,0])
se_std = np.sqrt(hess[1,1])

print("ML estimate of mean: ", mean)
print("ML estimate of st. dev: ", std)
print("SE(mean): ", se_mean)
print("SE(st. dev): ", se_std)
```

```
ML estimate of mean:  0.0008232842658414187
ML estimate of st. dev:  0.01601250920762722
SE(mean):  0.00022508634027117912
SE(st. dev):  0.00022521591150143007
```

```python
# H0 : mu = 0
t = mean/se_mean
print('t-stat: ', t, '', '> 1.96')
print('==> Reject the null!')
```

```
t-stat:  3.657637619633176  > 1.96
==> Reject the null!
```

In this problem, we optimized the log likelihood function with respect to our μ and σ parameters. In this optimization process, `scipy.optimize()` produces an inverse Hessian matrix. By taking the square root of the diagnol entries of this inverse Hessian, we could recover the standard errors for our parameters. We then used the estimate for μ and its standard error to run the hypothesis test. Ultimately we found that the t-stat exceeded the 95% level critical score. Hence, we **reject the null!**

Part (b)

Compute the sample skewness and sample kurtosis using the estimators of μ and σ you obtained in part (a). Are the estimators of skewness and kurtosis consistent with the normal assumption obrtained in question (a)?

Solution

```
# Compute sample skewness and kurtosis
n = len(cc)
skew = np.sum((cc - mean)**3 / std**3) / n
kurt = np.sum((cc - mean)**4 / std**4) / n
print('Skewness: ', skew)
print('Kurtosis: ', kurt)
```

```
Skewness:  -0.2450878711900944
Kurtosis:  13.328635970459045
```

The skewness and kurtosis of a Normally distributed variable are 0 and 3, respectively. As we can see from our above result, the skewness looks consistent with a normal distribution; however the kurtosis is greater than 3. Therefore, we argue that our data might not be normally distributed.

This is unsurprising since financial data tends to have thicker tails and thus larger kurtosis.

Part (c)

Perform the Jarque-Bera Test on the daily cc returns and explain the test result.

Solution

```
# Jarque-Bera Test for Normality
jb = stats.jarque_bera(cc)
print('JB Stat: ', jb[0])
print('JB p-value: ', jb[1])
```

```
JB Stat:  12425.052311964031
JB p-value:  0.0
```

From the above Jarque-Bera test, we see that the p-value is 0 which means that we reject the null hypothesis that our data is normally distributed. The test statistic should be 0 if the data is normally distributed. Our test statistic is about 12425 with a p-value of 0. This means that it is statistically significantly different from 0. Hence, we say our data is **not** normally distributed.

Part (d)

In the view of the testing results in part (c), do we trust the maximum likelihood estimators of μ and σ in part (a)? Why?

Solution

No, we cannot trust our estimators! Our likelihood function used the normal distribution as its density function; hence, since our data is not actually normally distributed, the optimization procedure optimized on the wrong density function. This will give us incorrect estimators for our data. We would have been better off using a different density function (e.g., one with bigger tails).