HW2

Sitian Qian 1600011388

September 27, 2019

1 Problem 1

1.1 Subproblem 1

Proof.
Firstly we can write $\hat{\sigma}_a^2$ and $\hat{\sigma}_b^2$ as following:

$$\begin{bmatrix} \hat{\sigma}_a^2 \\ \hat{\sigma}_b^2 \end{bmatrix} = \frac{1}{nq} \sum_{k=1}^{n} \begin{bmatrix} \sum_{j=1}^{q} (\varepsilon_{q(k-1)+j}^2) \\ (\sum_{j=1}^{q} \varepsilon_{q(k-1)+j})^2 \end{bmatrix}$$

Since $\{\varepsilon_{qk}\}$ is an iid sequence, $\{\sum_{j=1}^{q}(\varepsilon_{q(k-1)+j}^2)\}$,$\{(\sum_{j=1}^{q}\varepsilon_{q(k-1)+j})^2\}$ are also iid sequences. Which enables us to use central limit theorem to obtain the asymptotic distribution.
The expectation is easy to obtain:

$$\mathbb{E} \begin{bmatrix} \sum_{j=1}^{q} (\varepsilon_{q(k-1)+j}^2) \\ (\sum_{j=1}^{q} \varepsilon_{q(k-1)+j})^2 \end{bmatrix} = \begin{bmatrix} q\sigma^2 \\ q\sigma^2 \end{bmatrix}$$

The variance can be obtained as below:

$$\begin{aligned} &Var(\sum_{j=1}^{q}(\varepsilon_{q(k-1)+j}^2)) \\ =&\sum_{j=1}^{q}(Var(\varepsilon_{q(k-1)+j}^2)) \\ =&q(A-1)\sigma^4 \end{aligned}$$

(Here I use A to represent $\mathbb{E}x^4$, x is subject to standard normal distribution, in fact $A = 3$;)

$$\begin{aligned} &Var((\sum_{j=1}^{q}\varepsilon_{q(k-1)+j})^2) \\ =&qA\sigma^4 + 6 * C_q^2\sigma^4 - q^2 * \sigma^4 \\ =&(Aq + 2q^2 - 3q)\sigma^4 \end{aligned}$$

and

$$\begin{aligned} &Cov((\sum_{j=1}^{q}\varepsilon_{q(k-1)+j})^2, \sum_{j=1}^{q}(\varepsilon_{q(k-1)+j}^2)) \\ =&Var(\sum_{j=1}^{q}(\varepsilon_{q(k-1)+j}^2)) + Cov(\sum_{i=1}^{q}\sum_{j\neq i,j=1}^{q}(\varepsilon_{q(k-1)+i}\varepsilon_{q(k-1)+j}), \sum_{j=1}^{q}(\varepsilon_{q(k-1)+j}^2)) \\ =&q(A-1)\sigma^4 \end{aligned}$$

Finally we get:

$$Var\left(\left[\begin{array}{c} \sum_{j=1}^{q}(\varepsilon_{q(k-1)+j}^{2}) \\ (\sum_{j=1}^{q}\varepsilon_{q(k-1)+j})^{2} \end{array}\right]\right)=\left[\begin{array}{cc} 2q\sigma^4 & 2q\sigma^4 \\ 2q\sigma^4 & 2q^2\sigma^4 \end{array}\right]$$

Therefore, using CLT,

$$\sqrt{n}\left(\left[\begin{array}{c} \hat{\sigma}_a^2 \\ \hat{\sigma}_b^2 \end{array}\right]-\left[\begin{array}{c} \sigma^2 \\ \sigma^2 \end{array}\right]\right)\to_d \mathbb{N}\left(0,\left[\begin{array}{cc} \frac{2\sigma^4}{q} & \frac{2\sigma^4}{q} \\ \frac{2\sigma^4}{q} & 2\sigma^4 \end{array}\right]\right)$$

With the Delta method, representing $J_r=\frac{\hat{\sigma}_b^2}{\hat{\sigma}_a^2}-1$ *by* $J_r=h(\hat{\sigma}_a^2,\hat{\sigma}_b^2)$,*we have:*

$$\sqrt{n}(h(\hat{\beta})-h(\beta))\to_d \mathbb{N}\left(0,\nabla h^T(\beta)D\nabla h(\beta)\right)$$

where $\hat{\beta}=(\hat{\sigma}_a^2,\hat{\sigma}_b^2)^T$, $\beta=\vec{0}$, *and* $D=\left[\begin{array}{cc} \frac{2\sigma^4}{q} & \frac{2\sigma^4}{q} \\ \frac{2\sigma^4}{q} & 2\sigma^4 \end{array}\right]$,

Namely, the distribution of J_r *is,*

$$\sqrt{n}\,(J_r-0)\to_d \mathbb{N}\left(0,\left[\begin{array}{cc} -\frac{1}{\sigma^2} & \frac{1}{\sigma^2} \end{array}\right]\left[\begin{array}{cc} \frac{2\sigma^4}{q} & \frac{2\sigma^4}{q} \\ \frac{2\sigma^4}{q} & 2\sigma^4 \end{array}\right]\left[\begin{array}{c} -\frac{1}{\sigma^2} \\ \frac{1}{\sigma^2} \end{array}\right]\right)=\mathbb{N}(0,\frac{2(q-1)}{q})$$

Q.E.D

2 Problem 2

The code I write to conduct VR test is shown below:

```
import numpy as np
import scipy.stats as st

def J_r_calc(data,q):
    '''
    Calculation of J_r statistic

    Parameters:

    data: 1D array,
    time series data of a certain stock (difference between each time point)

    q: int, parameter of the VR test

    Return:

    return 1: float, value of J_r statistic

    return 2: int, n of the VR test, namely Size of data devided by q

    '''
    # modify the length of data to nq
    if len(data)%q != 0:
        data = data[len(data)%q:]

    n = len(data)//q

    sigma_a_sq=np.power(data,2).sum()
```

```
    # print(q,n,len(data))
    data_q = data.reshape(q,n)
    sigma_b_sq=np.power(data_q.sum(axis=0),2).sum()

    return sigma_b_sq/sigma_a_sq - 1 ,n

def VR_test(data,q):
    '''
    Operate the VR test to the time series data

    Parameters:

    data: 1D array,
    time series data of a certain stock (difference between each time point)

    q: int, parameter of the VR test

    Return:

    return 1: float, p-value of the VR test

    '''

    J_r,n=J_r_calc(data,q)

    sigma=np.sqrt(2*(q-1)/q)

    # print(sigma,J_r)

    return (1-st.norm.cdf(np.sqrt(n)*abs(J_r)/sigma))*2
```

I choose the IBM's stock, using its monthly return data from 1970/01 to 2018/03 as input to conduct VR test, here I neglect the data pre-processing part,and the whole jupyter notebook is displayed in the appendix.
Then conduct VR test for $q = 2, 3, 6, 12$:

```
for q in [2,3,6,12]:
print("q={:d},p-value={:.4f}".format(q,VR_test(IBM_return,q)))
# the variable IBM_return is the monthly return of IBM
```

The output is shown below:

```
q=2,p-value=0.9516
q=3,p-value=0.1601
q=6,p-value=0.6176
q=12,p-value=0.3126
```

I choose the precession to be 4 digits after dot for simplicity.

3 Problem 3

It's obvious to find that since S^* is the max among those $(|\sqrt{\frac{nq}{2(q-1)}}J_r(q)|)$s, and every $\sqrt{\frac{nq}{2(q-1)}}J_r(q)$ is subject to a standard normal distribution, so if S^* is bigger than the $(1-\alpha^+/2)$-th percentile of the standard normal distribution, it is equal to that the smallest p-value from VR test to different q is smaller than α^+.
So I can write code like this:

```
alpha_plus=1-(1-0.05)**(1/4) #alpha=0.05
p_values=[VR_test(IBM_return,q) for q in [2,3,6,12]]
S_star_p=min(p_values)
if(S_star_p<alpha_plus):
    print('Reject')
else:
    print("Not Reject")
```

The output is "Not Reject". Namely, we can not reject the null hypothesis with $\alpha = 0.05$.