Exercise 1

A

The figure is below:

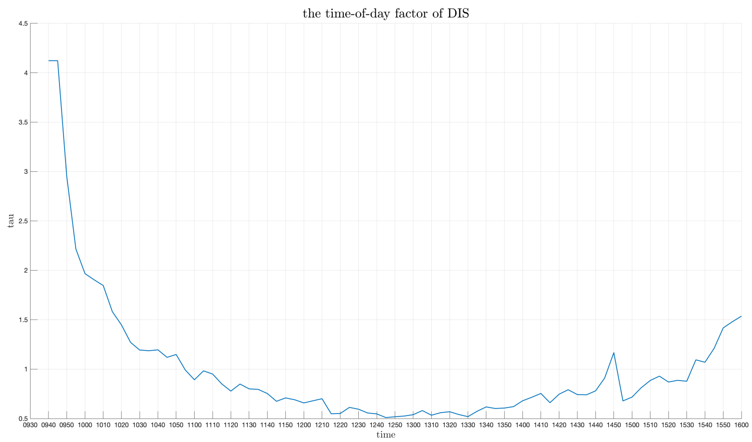

Figure 1: the time-of-day factor of DIS.

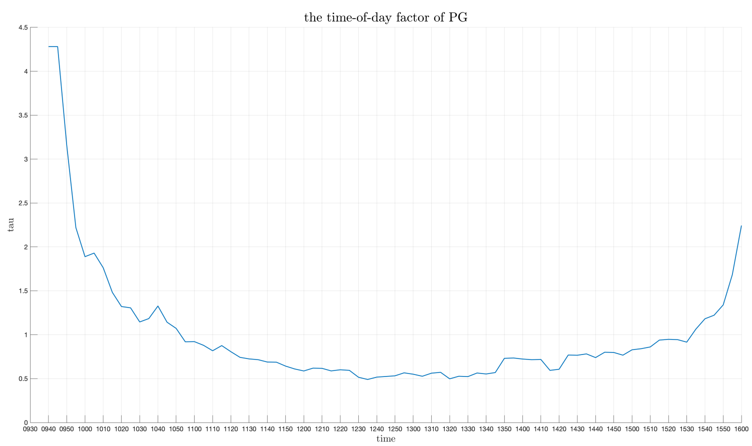

Figure 2: the time-of-day factor of PG.

Interpret:the time-of-day factor τ is the volatility weight within days as an adjustment of volatility. As we can see from the figure 1 and 2, the shape of τ is like a "U", which means the volatilities of the morning and the afternoon are bigger than the rest of the day, because the transactions in these two time periods are more frequent. And the τ in the morning is bigger than the τ in the afternoon, because after experiencing the accumulation of market information for one night, the volume of transactions was released in the morning.

C

The figure is below:

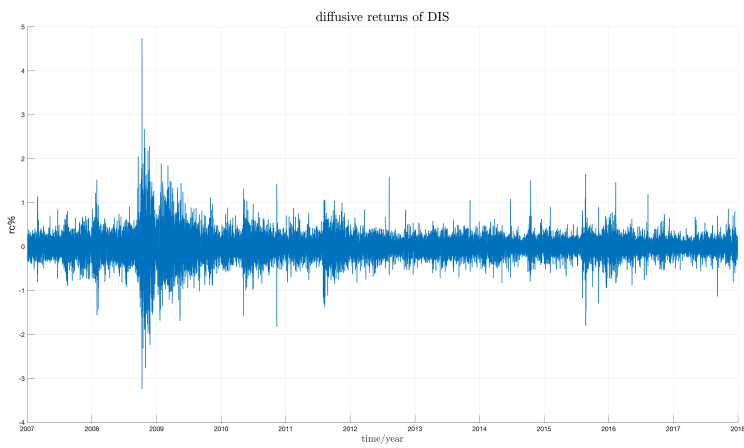

Figure 3: the diffusive returns of DIS.

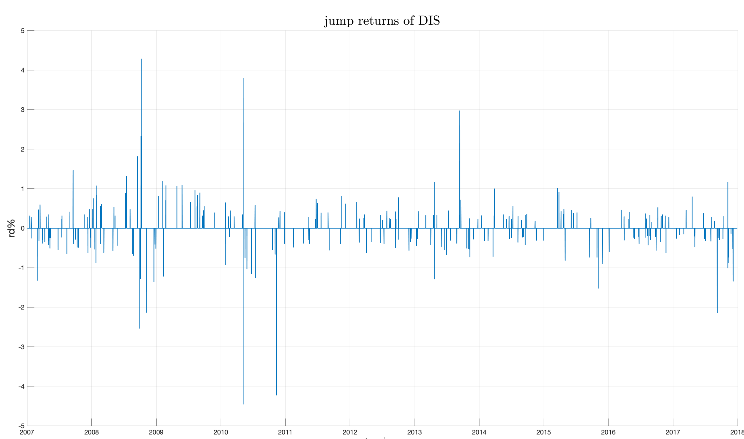

Figure 4: the jump returns of DIS.

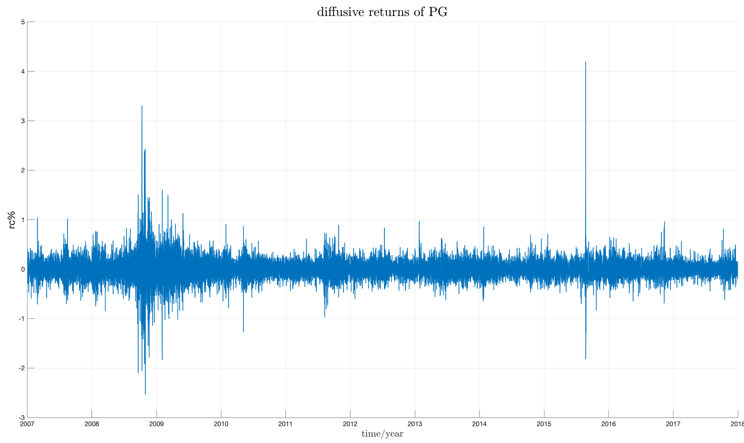

Figure 5: the diffusive returns of PG.

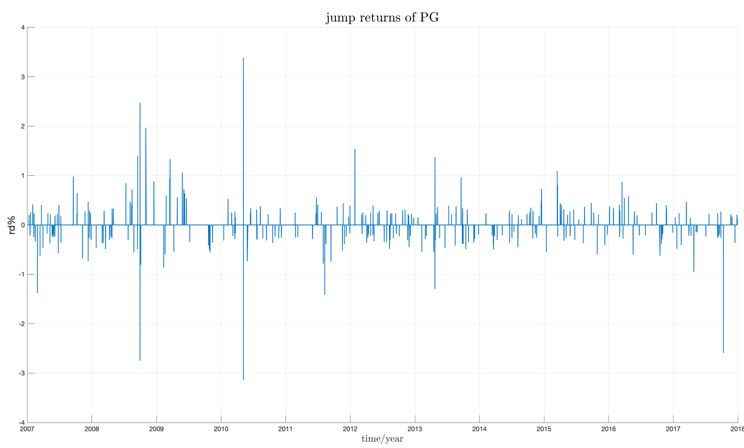

Figure 6: the jump returns of PG.

Interpret: From the figure 3 and 5, we can see the diffusive returns fluctuate wildly, especially around 2009, which is exactly after the famous 2008 financial crisis. Around 2016, they meet the second obvious volatility, which is also another stock market shock.

As for the jump returns in figure 4 and 6, although the trend is similar with the diffusive returns, they are more sparse than the diffusive returns, because the number of occurrences of jumps are much less than the number of occurrences of diffusive.

D

No, we can obviously observe there are not more jumps during periods of crisis like around 2008 and 2016. From the tables we can see, jumps are approximately evenly distributed throughout the years.
The DIS table is below:

years	jumps
2007	35
2008	28
2009	19
2010	21
2011	14
2012	24
2013	26
2014	28
2015	15
2016	30
2017	26

The PG table is below:

years	jumps
2007	38
2008	24
2009	18
2010	29
2011	20
2012	41
2013	26
2014	28
2015	23
2016	23
2017	27

E

The density function is a function whose value at any given sample (or point) in the sample space (the set of possible values taken by the random variable) can be interpreted as providing a relative likelihood that the value of the random variable would equal that sample.

I think the density of the diffusive returns would look like a normal distribution, because diffusive returns randomly oscillate around 0; the density of the jump returns would look like a normal distribution, but it would be more flat than the density of the diffusive returns, because the jumps are more scattered; the density for the non-zero jump returns would look like two mountain peaks, which sags at x equals 0, because we ignored the zeros in the jump returns.

G

The figure is below:

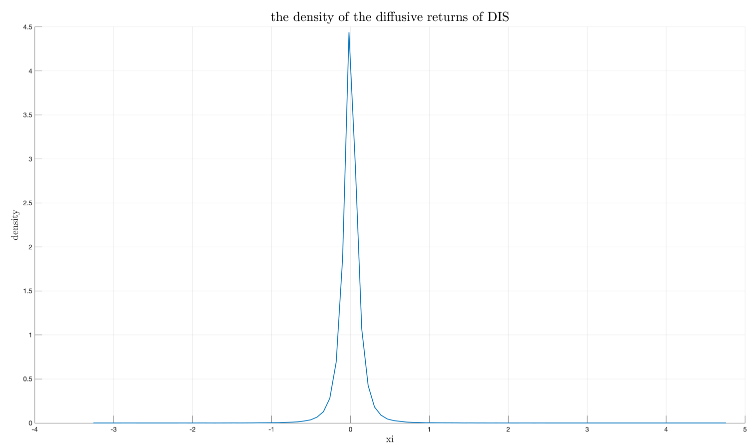

Figure 7: the density of the diffusive returns of DIS.

H

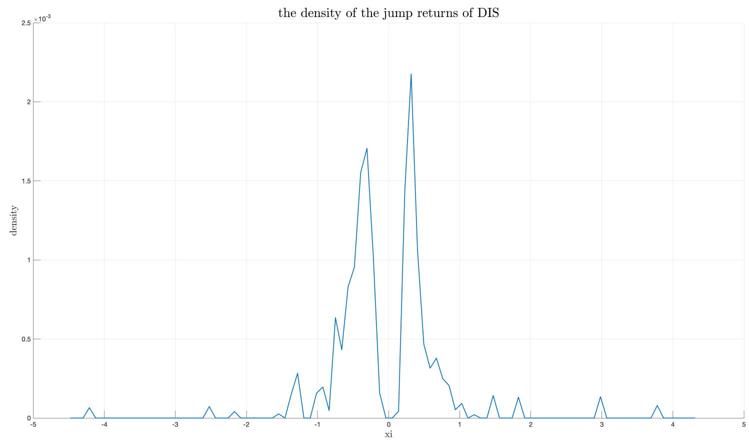

Figure 8: the density of the jump returns of DIS.

H

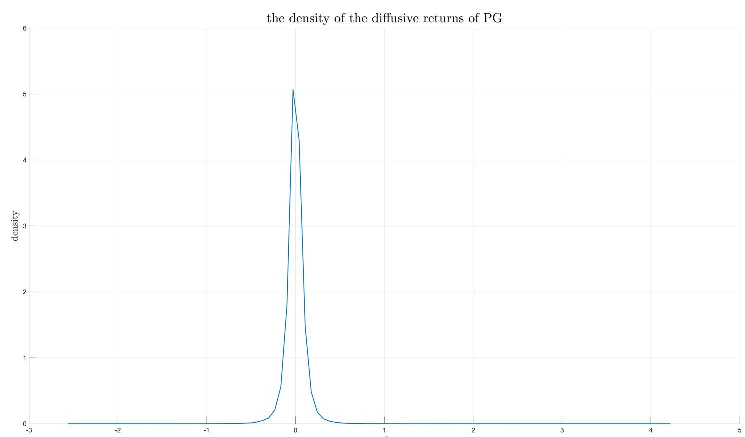

Figure 9: the density of the diffusive returns of PG.

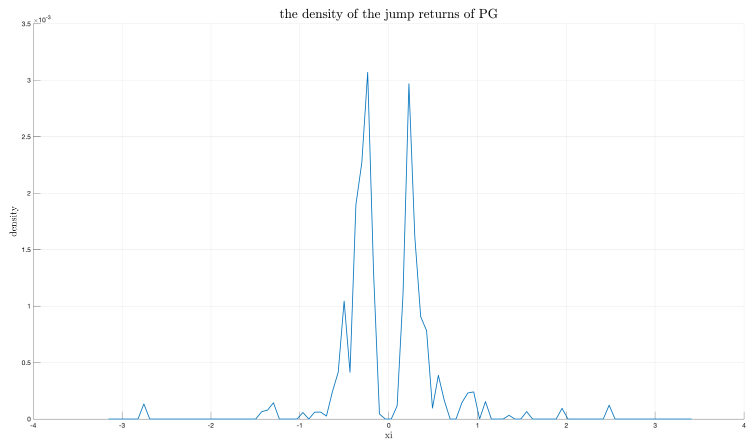

Figure 10: the density of the jump returns of PG.

Interpret: We can see from the figures, the density of the diffusive returns is highly concentrated at 0, while the density of the jump returns is like two mountain peaks with some little ups and downs, and is 0 at x equals 0. The figures meet my previous expectations.

H

I think this viewpoint is wrong. From my density plots, I can not conclude that negative jump returns are larger and occur much more often than positive jump returns.

Exercise 2

A

The figure is below:

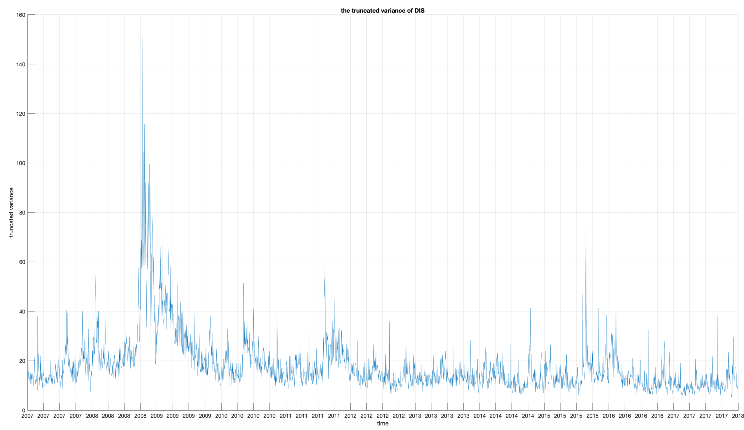

Figure 11: the truncated variance of DIS.

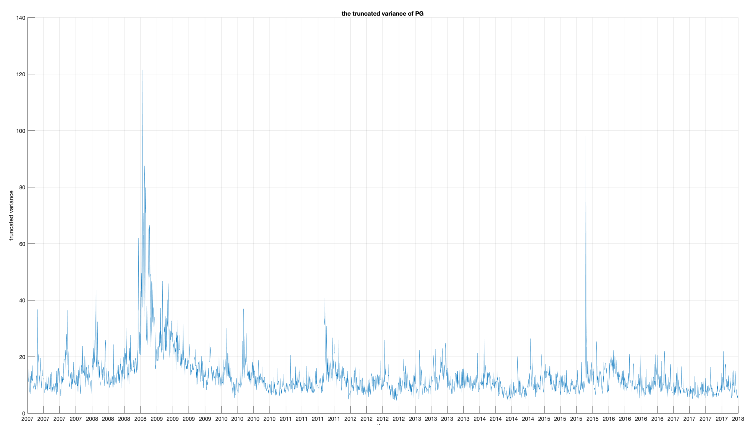

Figure 12: the truncated variance of PG.

Interpret: In addition to the figures, we can find the volatility of the truncated variance is very intense, and have obvious trend, which reached its highest point around 2009 and the second highest point around 2016. At the same time, it experienced ups and downs in the rest of time.

B

The figure is below:

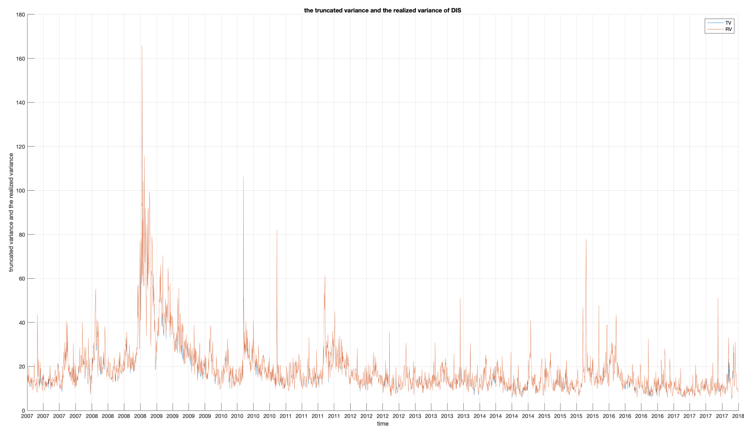

Figure 13: the truncated variance and the realized variance of DIS.

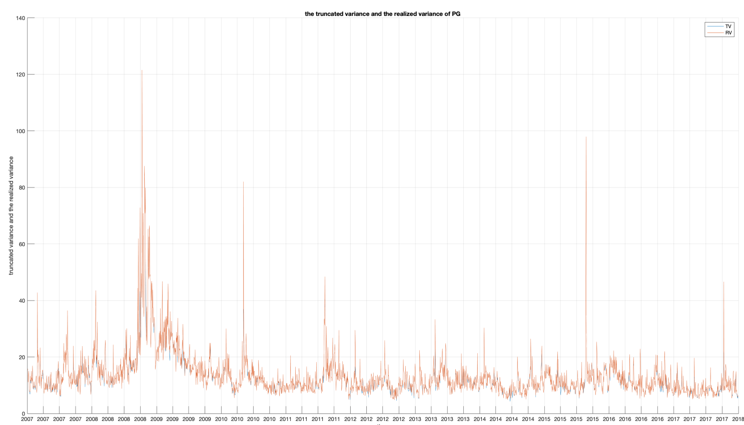

Figure 14: the truncated variance and the realized variance of PG.

Compare: From the figure 13 and 14, we can find the trend of the truncated variance and the realized variance is very similar, while the truncated variance is a little bit smaller than the realized variance some time.

C

The figure is below:

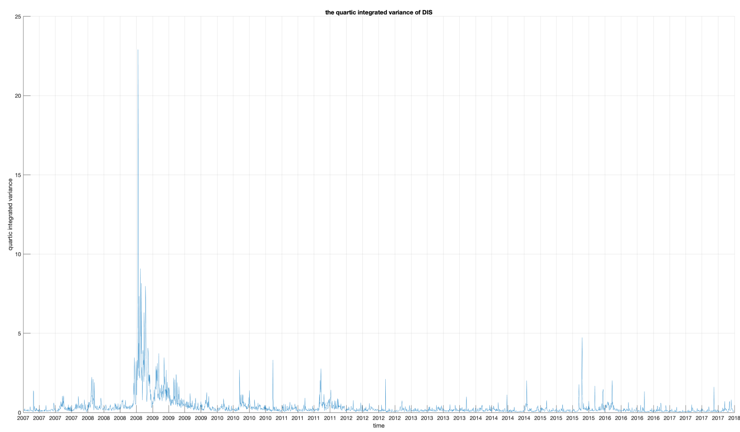

Figure 15: the quartic integrated variance of DIS.

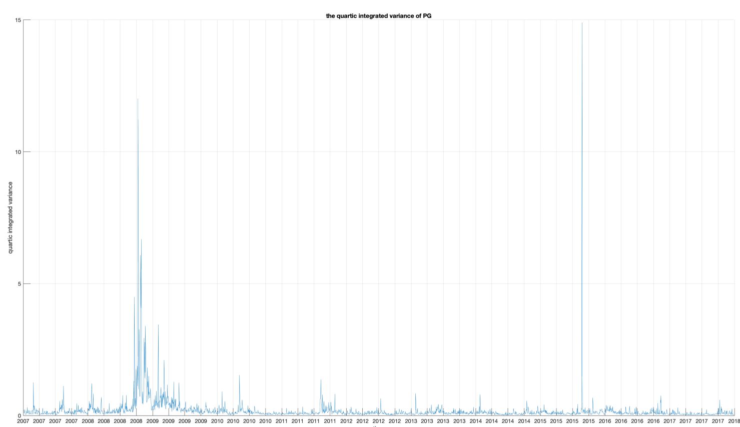

Figure 16: the quartic integrated variance of PG.

Interpret: We can see the volatility of the quartic integrated variance is relatively small from the figure 15 and 16. At the same time, it allows us to observe the great change after 2008 and before 2016 clearly.

D

The figure is below:

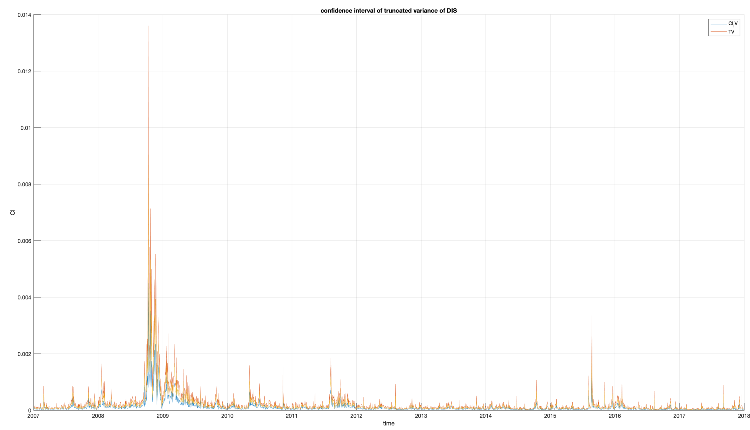

Figure 17: the confidence interval of truncated variance of DIS.

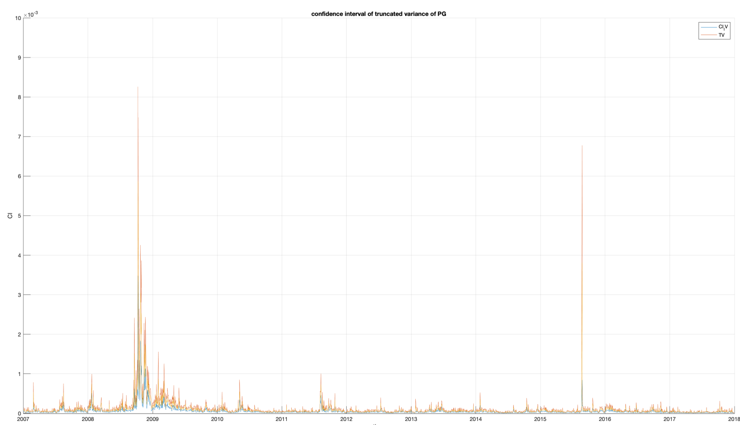

Figure 18: the confidence interval of truncated variance of PG.

E

The figure is below:

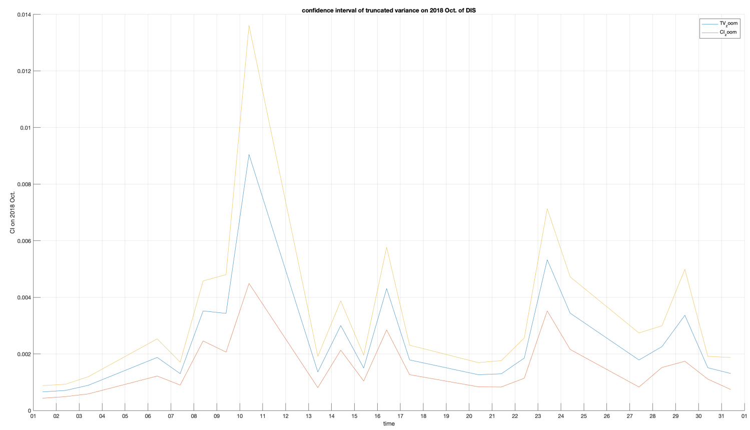

Figure 19: the confidence interval of truncated variance on October 2008 of DIS.

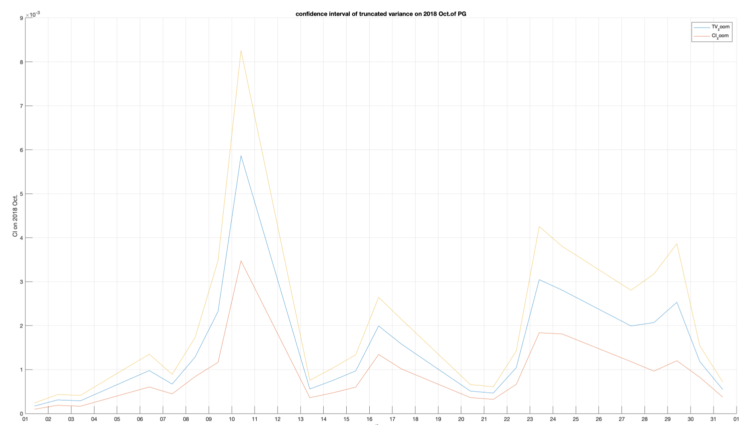

Figure 20: the confidence interval of truncated variance on October 2008 of PG.

Discuss: We can find the truncated variance experiences obvious volatility on the 11th, 16th, 23rd, 29th day of October 2008. Compared to the plots in the previous exercise, these figures can show more details clearly.

F

- Delta Method theorem is a result concerning the approximate probability distribution for a function of an asymptotically normal statistical estimator from knowledge of the limiting variance of that estimator. It allow us to change the unit of measure of the estimator and its limit distribution.

- Because we need it to get the asymptotic distribution of a rescaled estimator under the situation that we only know the variance of the original estimator and the differentiable transform function
- g should be defined as $g(x) = 100 \times \sqrt{252 \times TV_t}$
- g' should be defined as $g'(x) = 50 \times \sqrt{\frac{252}{TV_t}}$
- The asymptotic distribution is $\mathcal{N}(0, g'(TV_t) \times 2\widehat{QIV_t})$

G

The confidence intervals for the annualized integrated vaqriance:

$$CI(100\sqrt{252 \times IV_t}, 1-\alpha) = [TV_t + qz(\alpha/2)g'(TV_t)\sqrt{\Delta_n 2\widehat{QIV}_t}, TV_t - qz(\alpha/2)g'(TV_t]\sqrt{\Delta_n 2\widehat{QIV_t}}]$$

H

The figure is below:

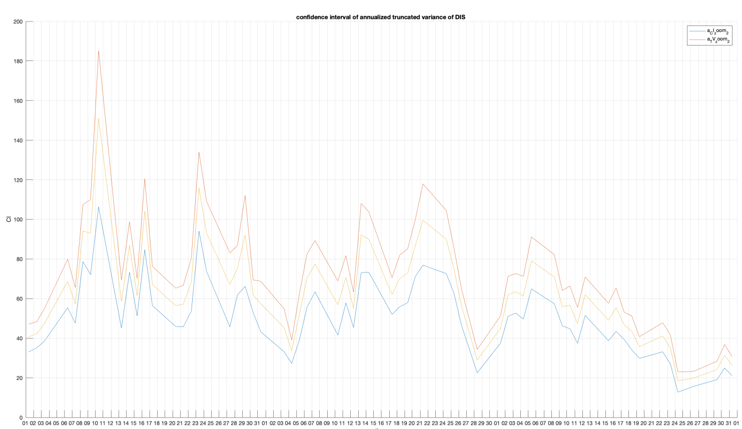

Figure 21: the confidence interval of annualized truncated variance of DIS.

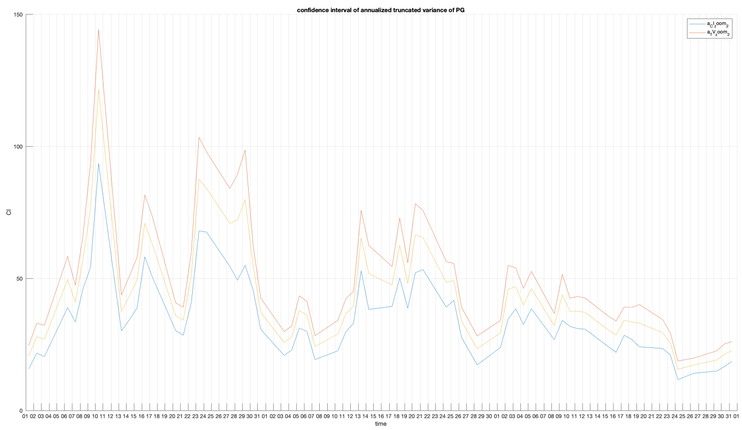

Figure 22: the confidence interval of annualized truncated variance of PG.

Interpret: I chosen three months from October to November 2008. And we can see annualized truncated variance show a similar pattern in the three months, except that it is more intense in October.At the same time, compared to the previous figures that the truncated variance is not annualized, we can see it keeps basically same on October, 2018 which means the delta method does work well.

Exercise 3

A

The figure is below:

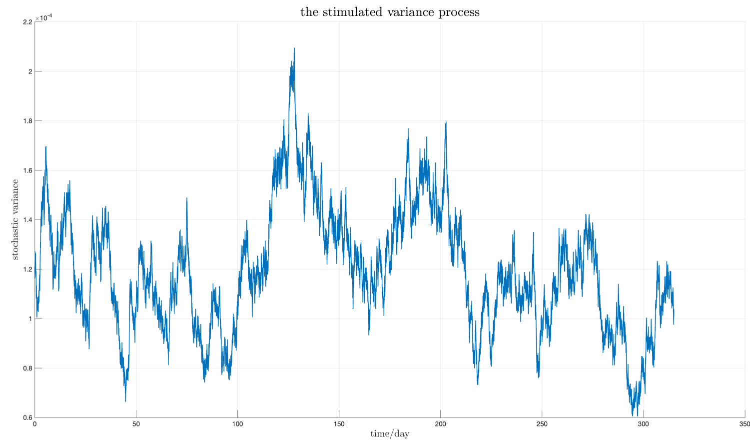

Figure 23: the stimulated variance process

B

- IV represents the variance of the variance process, so the number of it means the volatility of the variance.
- We can find the average variance of a day to get the similar c.

C

The figure is below:

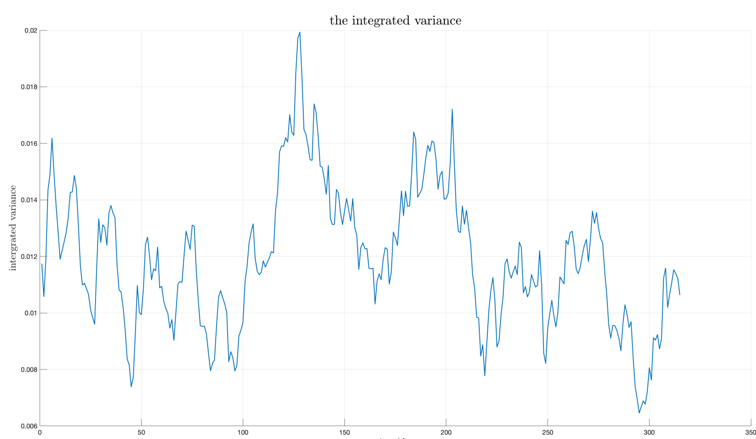

Figure 24: the integrated variance

Interpret: The trend of this figure is the same as the stimulated variance process but it is much more smooth.

D

The figure is below:

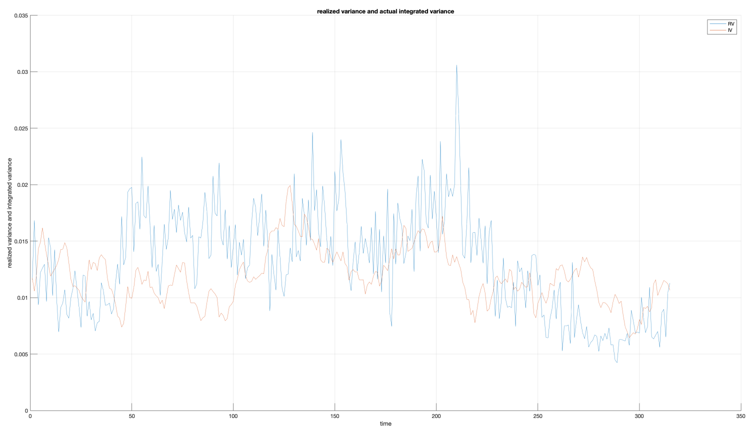

Figure 25: the realized variance and integrated variance

Discuss: The trends of the realized variance and integrated variance are not similar at the most of time.In the beginning and the end of year, integrated variance is likely bigger than the realized variance. The integrated variance is usually bigger than the realized variance at the rest of time.

E

The figure is below:

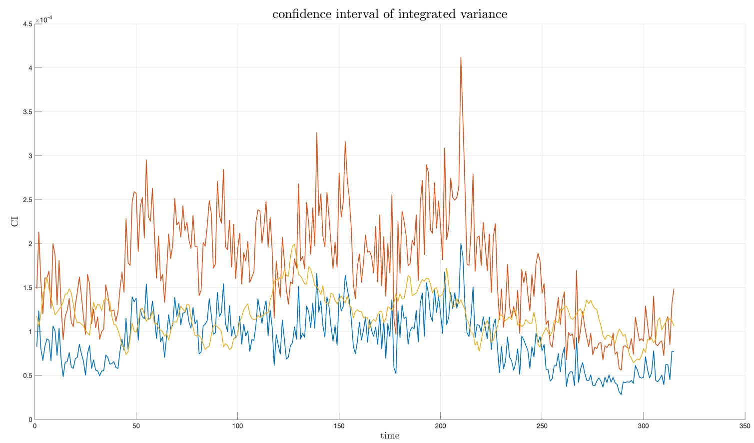

Figure 26: confidence interval of integrated variance

F

- The days that IV is within the confidence interval are 173 days.
- The average coverage rate is 54.92%
- It is not close to the expected 95%

G

The figure is below:

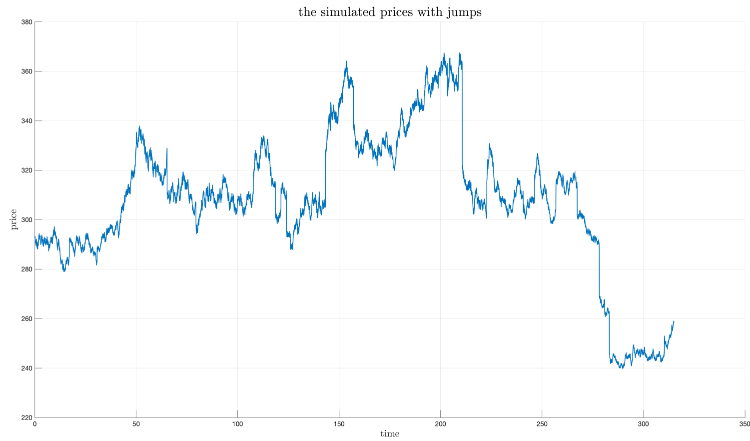

Figure 27: the simulated prices with jumps

H

- The coverage rate of the confidence intervals is 59.37%.
- It is bigger than the rate without jumps.
- Yes. Because the price volatility would be better captured with jumps.
- It would make the confidence interval be worse, and the inference is less correct.

I

The figure is below:

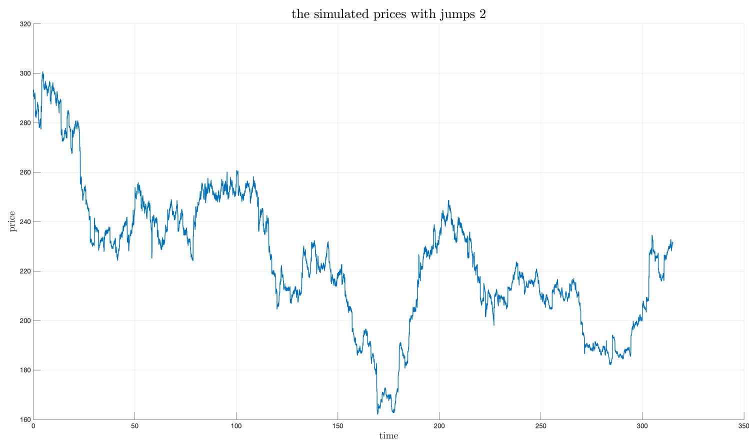

Figure 28: the simulated prices with jumps

Interpret: The coverage rate of the confidence intervals is 77.14%.It is bigger than the rate with jumps.

Code

The code is below:

Listing 1: ex_1.m

```
[dates, prices] = load_stock('DIS.csv');

% question A
[dates_return,deltax] = log_returns(dates, prices);
[n,T] = size(deltax);
deltan = 1/n;
% b_f = sum(abs(deltax(2:end,:).*deltax(1:n-1,:)),2)/T;
% b1 = b_f(1:1);
% b = [b1;b_f];
% tau = n*b/sum(b);
tau = tau_f(deltax);

f = figure(1);
set(f,'units','normalized','outerposition',[0 0 1 1]);
plot(dates_return(:,1),tau)
datetick('x','HHMM');
title('the time-of-day factor of DIS');
box off; grid on;
```

```
xlabel('time');
ylabel('tau');
print(f,'-dpng','-r200','figures/1A1');
close(f);

% question C
BV = bipower_var(deltax);
alpha = 4;
cutoff = alpha*deltan^0.49*sqrt(tau*BV);
rc = deltax;
rc(abs(deltax)>cutoff)=0;
rd = deltax;
rd(abs(deltax)<=cutoff)=0;

f = figure(2);
set(f,'units','normalized','outerposition',[0 0 1 1]);
plot(dates_return(:),100*rc(:))
datetick('x','yyyy');
title('diffusive returns of DIS');
box off; grid on;
xlabel('time/year');
ylabel('rc%');
print(f,'-dpng','-r200','figures/1C11');
close(f);

f = figure(3);
set(f,'units','normalized','outerposition',[0 0 1 1]);
plot(dates_return(:),100*rd(:))
datetick('x','yyyy');
title('jump returns of DIS');
box off; grid on;
xlabel('time/year');
ylabel('rd%');
print(f,'-dpng','-r200','figures/1C12');
close(f);

% question D
jumps = sum(rd~=0);
y = zeros(12,1);
[yyyy,mm,dd]= ymd_f('DIS.csv');
for i = 1:11
    y(i+1) = sum(ismember(yyyy,2006+i));
end
jumps_y = zeros(11,1);
for i = 1:11
    jumps_y(i) = sum(jumps(1,(sum(y(1:i))+1):sum(y(1:i
```

```
        +1))));
end
% xlswrite('DIS.csv',rd);

 % question G
[density_1,xi_1] = ksdensity(100.*rc(:),'Kernel','epanechnikov','bandwidth',0.01);

f = figure(4);
set(f,'units','normalized','outerposition',[0 0 1 1]);
plot(xi_1,density_1)
title('the density of the diffusive returns of DIS');
box off; grid on;
xlabel('xi');
ylabel('density');
print(f,'-dpng','-r200','figures/1G11');
close(f);

[density_2,xi_2] = ksdensity(100.*rd(:),'Kernel','epanechnikov','bandwidth',0.01);

f = figure(5);
set(f,'units','normalized','outerposition',[0 0 1 1]);
plot(xi_2,density_2)
title('the density of the jump returns of DIS');
box off; grid on;
xlabel('xi');
ylabel('density');
print(f,'-dpng','-r200','figures/1G12');
close(f);
```

Listing 2: ex_2.m

```
%compute the truncated variance
[dates, prices] = load_stock('DIS.csv');

% question A
[dates_return,deltax] = log_returns(dates, prices);
[n,T] = size(deltax);
deltan = 1/n;

tau = tau_f(deltax);
BV = bipower_var(deltax);
alpha = 4;
cutoff = alpha*deltan^0.49*sqrt(tau*BV);
rc = deltax;
```

```
rc(abs(deltax)>cutoff)=0;

TV = sum((rc).^2);

f = figure(11);
set(f,'units','normalized','outerposition',[0 0 1 1]);
plot(dates_return(1,:),100*sqrt(TV*252))
datetick('x','YYYY');
title('the truncated variance of DIS');
box off; grid on;
xlabel('time');
ylabel('truncated variance');
print(f,'-dpng','-r200','figures/2A1');
close(f);

RV1 = realized_var(deltax);

RV_last1 = 100*sqrt(RV1*252);

dates_RV_last1 = dates_return(1,:);

f = figure(12);
set(f,'units','normalized','outerposition',[0 0 1 1]);
plot(dates_return(1,:),100*sqrt(TV*252))
hold on
plot(dates_RV_last1,RV_last1)
p2.Color(4) = 0.25;
hold off
legend('TV','RV')
datetick('x','YYYY');
title('the truncated variance and the realized
    variance of DIS');
box off; grid on;
xlabel('time');
ylabel('truncated variance and the realized variance')
    ;
print(f,'-dpng','-r200','figures/2B1');
close(f);

%create QIV
QIV = 1/(3*deltan)*sum((rc).^4);

f = figure(13);
set(f,'units','normalized','outerposition',[0 0 1 1]);
plot(dates_return(1,:),100*sqrt(QIV*252))
datetick('x','YYYY');
```

```
title('the quartic integrated variance of DIS');
box off; grid on;
xlabel('time');
ylabel('quartic integrated variance');
print(f,'-dpng','-r200','figures/2C1');
close(f);

% CI95 = norminv([0.025 0.975]);
% TVCI95 = TV + CI95(:).*sqrt(2/n*QIV) ;
% questian D
CI_TV = ci_f(0.05,n,TV,QIV);


f = figure(14);
set(f,'units','normalized','outerposition',[0 0 1 1]);
plot(dates(1,:),CI_TV(1:2,:));
hold on
plot(dates(1,:),TV);
hold off
datetick('x','yyyy');
legend('CI_IV','TV');
title('confidence interval of truncated variance of DIS');
box off; grid on;
xlabel('time');
ylabel('CI');
print(f,'-dpng','-r200','figures/2D1');
close(f);

% question E
zoom_start = find(dates_return==datenum(2008,10,1,16,0,0))/n;
zoom_end = find(dates_return==datenum(2008,10,31,16,0,0))/n;
TV_zoom = TV(zoom_start:zoom_end);
CI_zoom = CI_TV(:,zoom_start:zoom_end);
dates_zoom = dates_return(1,zoom_start:zoom_end);

f = figure(15);
set(f,'units','normalized','outerposition',[0 0 1 1]);
plot(dates_zoom,TV_zoom)
hold on
plot(dates_zoom,CI_zoom)
hold off
datetick('x','dd');
legend('TV_zoom','CI_zoom');
```

```
title('confidence interval of truncated variance on
    2018 Oct. of DIS');
box off; grid on;
xlabel('time');
ylabel('CI on 2018 Oct.');
print(f,'-dpng','-r200','figures/2E1');
close(f);

% question h

% a_CI_TV_zoom = 100*sqrt(CI_TV_zoom*252);
zoom_start_2 = find(dates_return==datenum
    (2008,10,1,16,0,0))/n;
zoom_end_2 = find(dates_return==datenum
    (2008,12,31,16,0,0))/n;
TV_zoom_2 = TV(zoom_start_2:zoom_end_2);
CI_zoom_2 = CI_TV(:,zoom_start_2:zoom_end_2);
dates_zoom_2 = dates_return(1,zoom_start_2:zoom_end_2)
    ;
a_CI_zoom_2 = 100*sqrt(CI_zoom_2*252);
a_TV_zoom_2 = 100*sqrt(TV_zoom_2*252);

f = figure(16);
set(f,'units','normalized','outerposition',[0 0 1 1]);
plot(dates_zoom_2,a_CI_zoom_2);
hold on
plot(dates_zoom_2,a_TV_zoom_2);
hold off
datetick('x','dd');
legend('a_CI_zoom_2','a_TV_zoom_2');
title('confidence interval of annualized truncated
    variance of DIS');
box off; grid on;
xlabel('time');
ylabel('CI');
print(f,'-dpng','-r200','figures/2H1');
close(f);
```

Listing 3: ex_3.m

```
n = 80;
T = 1.25*252;
nE = 20*n;
rho = 0.03;
muc = 0.011^2;
sigmac = 0.001;
```

```
c0 = muc;
x0 = log(292.58);
deltaE = 1/nE;
delta = 1/n;

seed = 3005;
rng(seed,'twister');

c = simsv(c0,rho,muc,deltaE,sigmac,nE,T);

t = 0:deltaE:T;

f = figure(23);
set(f,'units','normalized','outerposition',[0 0 1 1]);
plot(t,c);
title('the stimulated variance process');
box off; grid on;
xlabel('time/day');
ylabel('stochastic variance');
print(f,'-dpng','-r200','figures/3A1');
close(f);

% question C
c_a = reshape(c(2:end),nE,T);
IV = (1/nE)*sum(c_a);
t_day = 1:T;

f = figure(24);
set(f,'units','normalized','outerposition',[0 0 1 1]);
plot(t_day,IV*100);
title('the integrated variance');
box off; grid on;
xlabel('time/day');
ylabel('intergrated variance');
print(f,'-dpng','-r200','figures/3C1');
close(f);

%quesition D
x = simlp(c0,x0,rho,muc,sigmac,delta,n,T);

p = exp(x);

deltax = simlr(x,n,T);
deltax_a = reshape(deltax(2:end),n,T);
RV1 = realized_var(deltax_a);

```

```
f = figure(25);
set(f,'units','normalized','outerposition',[0 0 1 1]);
plot(t_day,RV1*100)
hold on
plot(t_day,IV*100);
hold off
legend('RV','IV')
title('realized variance and actual integrated
    variance');
box off; grid on;
xlabel('time');
ylabel('realized variance and integrated variance');
print(f,'-dpng','-r200','figures/3D');
close(f);

% question E
alpha = 0.05;
z = norminv([(alpha/2) (1-alpha/2)]);
QIV = sum((deltax_a).^4)*n/3;
CI_IV = RV1 + z(:).*sqrt(2/n*QIV);

f = figure(26);
set(f,'units','normalized','outerposition',[0 0 1 1]);
plot(t_day,CI_IV)
hold on
plot(t_day,IV);
hold off
legend('CI_IV','IV');
% datetick('x','yyyy');
title('confidence interval of integrated variance');
box off; grid on;
xlabel('time');
ylabel('CI');
print(f,'-dpng','-r200','figures/3E');
close(f);

% %question F
q_1 = IV>=CI_IV(1,:);
w_1 = IV<=CI_IV(2,:);
days_1 = sum(and(q_1,w_1));
acr_1 = days_1/T;

%question G
lambda = 20/252;
N = poissrnd(lambda*T);
sigmaj = 30*sqrt(muc/N);
```

```
Y = normrnd(0,sigmaj,1,n*T);
U = unifrnd(0,1,1,n*T);

J = simcp(Y,U,N,n,T);
p = exp(x+J);
tp = 0:delta:T;

f = figure(27);
set(f,'units','normalized','outerposition',[0 0 1 1]);
plot(tp,p)
title('the simulated prices with jumps');
box off; grid on;
xlabel('time');
ylabel('price');
print(f,'-dpng','-r200','figures/3G');
close(f);

%question H
r = diff(x+J);
r = reshape(r,n,T);
RV_j = realized_var(r);
QIV_j = n/3*sum(r.^4);
CI_j = ci_f(0.05,n,RV_j,QIV_j);
q_2 = IV>=CI_j(1,:);
w_2 = IV<=CI_j(2,:);
days_2 = sum(and(q_2,w_2));
acr_2 = days_2/T;

% question I
lambda = 1;
N = poissrnd(lambda*T);
sigmaj = 30*sqrt(muc/N);
Y = normrnd(0,sigmaj,1,n*T);
U = unifrnd(0,1,1,n*T);


J = simcp(Y,U,N,n,T);
p = exp(x+J);
tp = 0:delta:T;

f = figure(28);
set(f,'units','normalized','outerposition',[0 0 1 1]);
plot(tp,p)
title('the simulated prices with jumps 2');
box off; grid on;
xlabel('time');
```

```
ylabel('price');
print(f,'-dpng','-r200','figures/3I');
close(f);

r = diff(x+J);
r = reshape(r,n,T);
RV_j = realized_var(r);
QIV_j = n/3*sum(r.^4);
CI_j = ci_f(0.05,n,RV_j,QIV_j);
q_3 = IV>=CI_j(1,:);
w_3 = IV<=CI_j(2,:);
days_3 = sum(and(q_3,w_3));
acr_3 = days_3/T;
```