

Exercise 1

- (A) Assume $X1$ is the market index, $X2$ is stock of company, then we can apply this formula to calculate the realized bate:

$$R\beta_t = \frac{RCov(X1, X2)_t}{IV_{1,t}}$$

We can estimate $RCov(X1, X2)_t$ by $RCov(\hat{X}1, X2)_t = \sum_{i=1}^n (r_{1,t,i}^c r_{2,t,i}^c)$ and estimate $IV_{1,t}$ by $\sum_{i=1}^n (r_{1,t,i}^c)^2$.

- (i) If we set $\alpha = 4.5$, the number of jumps from 2007 to 2017 for PG is

19, 13, 7, 12, 10, 23, 16, 13, 12, 11, 14

and total 150 times.

The number of jumps each for PG is

16, 13, 8, 15, 5, 14, 14, 10, 7, 13, 13

and total 128 times.

- (ii) From the jump times figures, there are some periods with a higher concentration of jumps: around 2007, 2010, 2012 and 2017. Those periods with more frequency jumps are consistent with the periods of crisis.

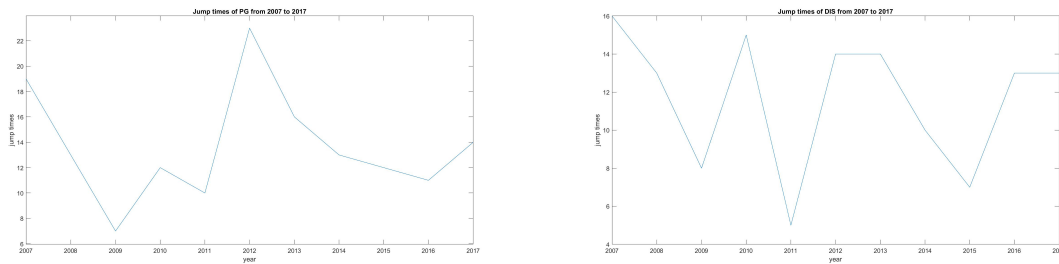


Figure 1: Jumps number of PG and DIS from 2007 to 2017

- (B) Here are the plots of truncated variance of PG and DIS in the first 2 week of 2007.

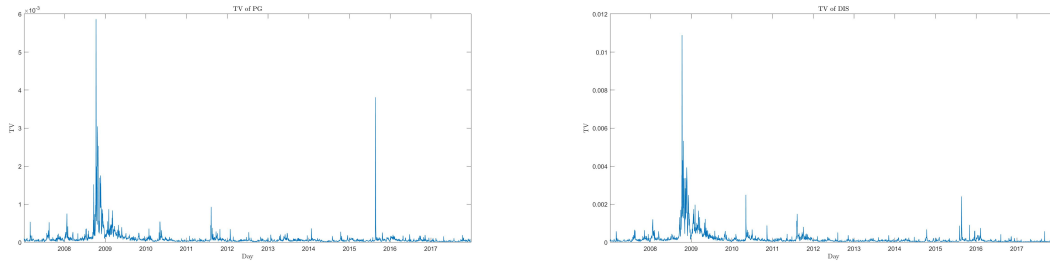


Figure 2: Truncated variance of PG and DIS

From this figure we can see that, the value of TV is largest around 2009, which means the stock price in 2009 has experienced a large volatility.

- (C) The follows are the figures of TV and estimated confidence interval of PG and DIS for the first 2 week in 2007.

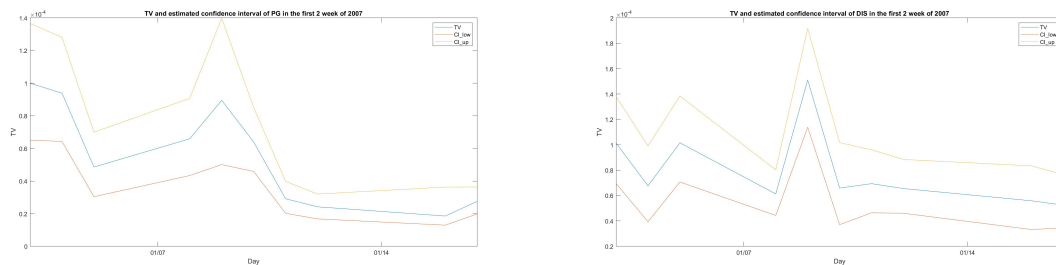


Figure 3: Truncated variance and confidence interval of PG and DIS (first 2 week of 2007)

The range of TV for PG and DIS are $(0.2 \times 10^{-4}, 1.0 \times 10^{-4})$ and $(0.5 \times 10^{-4}, 1.6 \times 10^{-4})$ respectively. According to the range of TV we can see that, the volatility of DIS' log returns' volatility is bigger than PG's, which means that the risk of DIS may be higher than PG's. We can also see that around January 10, the value of both PG and DIS' TV have experienced a big fluctuation.

As for the estimated confidence interval, it is clear that all the value of IV are bounded by the upper confidence interval and the lower confidence interval. We can calculate the cover rate of the confidence interval for PG and DIS: the cover rate for PG is 100% and the cover rate for DIS is 100%. These results are better than using asymptotic distribution(in project 3).

- (D) We do not need to apply Delta method to compute the confidence interval of annualized integral variance. Bootstrap method calculate the confidence interval for estimators by finding the quantile of samples, since transferring daily truncated variance to annualized truncated variance is a homogeneous increasing function, the rank of sample data

will not change: the 97.5% quantile and 2.5% quantile of annualized truncated variance is just the annualized 97.5% quantile and 2.5% quantile of daily truncated variance.

- (E) The follows are the figures of annualized TV and estimated confidence interval of PG and DIS for the first 2 week in 2007.

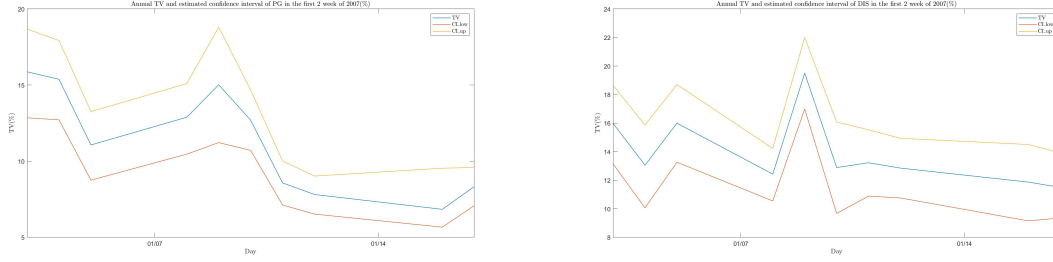


Figure 4: Annualized truncated variance and confidence interval of PG and DIS (first 2 week of 2007)

The range of annual TV for PG and DIS are (5%, 12%) and (11%, 20%) respectively. For the estimated confidence interval, we can find that all the value of IV are bounded by the confidence interval. We can calculate the cover rate of the confidence interval for annual PG and DIS' TV: the cover rate for PG is 100% and the cover rate for DIS is 100%. These results are the same as what we find in part C. According to the result, we can verify that when calculate annual TV's confidence interval, we can first transfer daily sample base into annualized data, and use the new data to do bootstrap (we do not need to use delta method).

- (F) The follows are the figures of IV and confidence interval estimated by asymptotic distribution and bootstrap method.

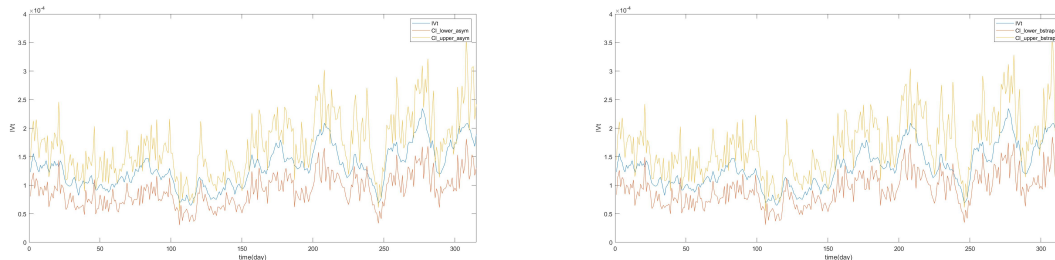


Figure 5: IV and confidence interval

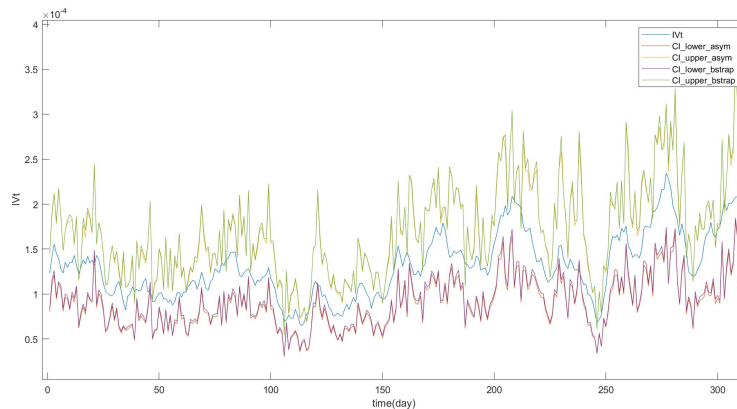


Figure 6: IV and CI estimated by asymptotic distribution and bootstrap

From these figures we can see, the shape of confidence interval estimated by asymptotic distribution and bootstrap method are very similar, and the value of lower confidence interval and upper confidence interval of bootstrap method are bigger than what we get from asymptotic distribution method.

The cover number and cover rate for these two methods in one experiment is:

$$Cover_num_{asym} = 291 \quad Cover_ratio_{asym} = 92.38\%$$

$$Cover_num_{bootstrap} = 292 \quad Cover_ratio_{bootstrap} = 92.70\%$$

The accuracy of distribution asymptotic and bootstrap method to estimate the confidence interval for IV in this case is similar.

The MATLAB:

```

1  addpath('D:\ZM-Documents\MATLAB\data','functions','scripts');
2  [dates_PG,lp_PG]=load_stock('PG.csv','m');
3  N_PG=sum(floor(dates_PG(1,1))==floor(dates_PG(:,1)));% number of observations
   per day
4  T_PG=size(dates_PG,1)/N_PG;
5  [rdates_PG,lr_PG]=log_return([dates_PG lp_PG],N_PG,1);
6  days_PG=unique(floor(rdates_PG));
7
8
9  %!!!
10 %if you want to check each scripts, you have to set the repeat time for this
   script
11 %otherwise, the code won't run successfully
12 M=11;
```

```
13 n=N_PG-1;
14 kn=(N_PG-1)/M;
15 %repeat times
16 %r=10000;
17
18 %part a
19 a=4.5;
20 [lr_c_PG,lr_d_PG]=c_d_log_returns(lr_PG,N_PG,a);
21
22 Y1={'20070101','20080101','20090101','20100101','20110101','20120101','
    '20130101','20140101','20150101','20160101','20170101','20180101'};
23 numY_PG=datenum(Y1,'yyyymmdd');
24 L=[rdates_PG(:),lr_d_PG(:)];
25 numJ_PG=zeros(1,11);
26 for i=1:11
27     numJ_PG(i)=sum(L(:,1)<=numY_PG(i+1) & L(:,2)~=0);
28 end
29 numJ_PG=diff([0 numJ_PG]);
30 J_PG=sum(numJ_PG);
31 figure;
32 plot(numY_PG(1:11),numJ_PG);
33 xlabel('year');
34 ylabel('jump times');
35 xlim([min(numY_PG(1:11)),max(numY_PG(1:11))]);
36 title('Jump times of PG from 2007 to 2017');
37 datetick('x','keeplimits');
38
39 %part b
40 TV_PG=truncated_var_day(lr_c_PG);
41
42 figure;
43 plot(days_PG,TV_PG);
44 xlabel('Day');
45 ylabel('TV');
46 title('TV of PG');
47 xlim([min(days_PG),max(days_PG)]);
48 datetick('x','keeplimits');
49
50 %part c
51 stv=zeros(r,T_PG);
52 parfor i=1:r
53     newsample=bsample(lr_PG,kn,n,T_PG);
54     stv(i,:)=sum(newsample.^2);
55 end
```

```
56
57 CI_low_PG = quantile(stv, 0.025);
58 CI_up_PG = quantile(stv, 0.975);
59 %evaluate estimated CI
60 num_in_PG_d=sum(CI_low_PG <=TV_PG & CI_up_PG >=TV_PG);
61 cover_rate_PG_d=num_in_PG_d/T_PG;
62
63 figure;
64 plot(days_PG(1:10),TV_PG(1:10));
65 hold on;
66 plot(days_PG(1:10),CI_low_PG(1:10));
67 hold on;
68 plot(days_PG(1:10),CI_up_PG(1:10));
69 hold off;
70 xlabel('Day');
71 ylabel('TV');
72 title('TV and estimated confidence interval of PG in the first 2 week of 2007
73 ');
74 xlim([min(days_PG(1:10)),max(days_PG(1:10))]);
75 legend('TV','CI_low','CI_up');
76 datetick('x','keeplimits');
77
78 %part e
79 annualTV_PG=100*sqrt(252*TV_PG);
80 %b=max(find(rdates_PG(:)>=733045 & rdates_PG(:)<733059+1)/(N_PG-1));
81 atv=zeros(r,T_PG);
82 parfor i=1:r
83     newsample=bsample(lr_PG,kn,n,T_PG);
84     atv(i,:)=100*sqrt(252*sum(newsample.^2));
85 end
86
87 annualCI_low_PG = quantile(atv, 0.025);
88 annualCI_up_PG = quantile(atv, 0.975);
89 %evaluate estimated CI
90 num_in_PG_a=sum(annualCI_low_PG <=annualTV_PG & annualCI_up_PG >=annualTV_PG)
91 ;
92 cover_rate_PG_a=num_in_PG_a/T_PG;
93
94 figure;
95 plot(days_PG(1:10),annualTV_PG(1:10));
96 hold on;
97 plot(days_PG(1:10),annualCI_low_PG(1:10));
98 hold on;
99 plot(days_PG(1:10),annualCI_up_PG(1:10));
```

```
98 hold off;
99 xlabel('Day');
100 ylabel('TV(\%)');
101 title('Annual TV and estimated confidence interval of PG in the first 2 week
      of 2007(\%)');
102 legend('TV','CI_low','CI_up');
103 xlim([min(days_PG(1:10)),max(days_PG(1:10))]);
104 datetick('x','keeplimits');
```

```
1  addpath('D:\ZM-Documents\MATLAB\data','functions','scripts');
2
3  %EXERCISE 1
4  %PART F
5
6  % SET PARAMETERS
7  n=80;
8  T=1.25*252;
9  ne=n*20;
10 l=0.03;
11 u_c=0.011^2;
12 sigma_c=0.001;
13 c0=u_c;
14 x0=log(292.58);
15
16 %!!!
17 %if you want to check each scripts, you have to set the repeat time for this
    script
18 %otherwise, the code won't run successfully
19
20 %repeat times
21 %r=10000;
22 kn=n;
23 r=10;
24
25
26
27 %Produce Ct
28 Ct=JD_sv_ct(ne,T,l,u_c,sigma_c,c0);
29 %Compute Real IV
30 Ct1=reshape(Ct(2:end),ne,[]);
31 IVt=sum(Ct1*l/ne);
32
33 %Stimulate Xt and Compute RV
34 %produce high frequency Xt
35 Xt=JD_sv_xt(ne,T,Ct,x0);
```

```
36 %compute coarser frequency log-return
37 Xt1=Xt(1:20:end);
38 lr=reshape(diff(Xt1),80,[]);
39 %compute RVt
40 RVt=realized_var_day(lr);
41
42 %Estimate the CI of IVt by asymptotic distribution
43 QIVt=QIV(lr,n);
44 c=-norminv(0.025);
45 CI_l_1=RVt-c*sqrt(2*QIVt/n);
46 CI_u_1=RVt+c*sqrt(2*QIVt/n);
47 % Evaluate
48 num_in_1=sum(CI_l_1<=IVt & IVt<=CI_u_1);
49 cover_rate_1=num_in_1/T;
50
51 %draw asymptotic CI
52 figure;
53 plot((1:T),IVt);
54 hold on;
55 plot((1:T),CI_l_1);
56 hold on;
57 plot((1:T),CI_u_1);
58 xlabel('time(day)');
59 ylabel('IVt');
60 xlim([0,T]);
61 legend('IVt','CI_lower_asym','CI_upper_asym');
62
63 %Estimate the CI of IVt by bootstrap
64
65
66 %part c
67 sRV=zeros(r,T);
68 parfor i=1:r
69     newsample=bsample(lr,kn,n,T);
70     sRV(i,:)=sum(newsample.^2);
71 end
72
73 CI_l_2= quantile(sRV, 0.025);
74 CI_u_2 = quantile(sRV, 0.975);
75
76 % Evaluate
77 num_in_2=sum(CI_l_2<=IVt & IVt<=CI_u_2);
78 cover_rate_2=num_in_2/T;
79
```



```
80 %draw bootstrap CI
81 figure;
82 plot((1:T),IVt);
83 hold on;
84 plot((1:T),CI_l_2);
85 hold on;
86 plot((1:T),CI_u_2);
87 xlabel('time(day)');
88 ylabel('IVt');
89 xlim([0,T]);
90 legend('IVt','CI_lower_bootstrap','CI_upper_bootstrap');
91
92 %draw two CI together
93 figure;
94 plot((1:T),IVt);
95 hold on;
96 plot((1:T),CI_l_1);
97 hold on;
98 plot((1:T),CI_u_1);
99 hold on;
100 plot((1:T),CI_l_2);
101 hold on;
102 plot((1:T),CI_u_2);
103 xlabel('time(day)');
104 ylabel('IVt');
105 xlim([0,T]);
106 legend('IVt','CI_lower_asym','CI_upper_asym','CI_lower_bootstrap','CI_upper_bootstrap');
```

Exercise 2

(A) The follows are the plot of realized beta of PG and DIS.

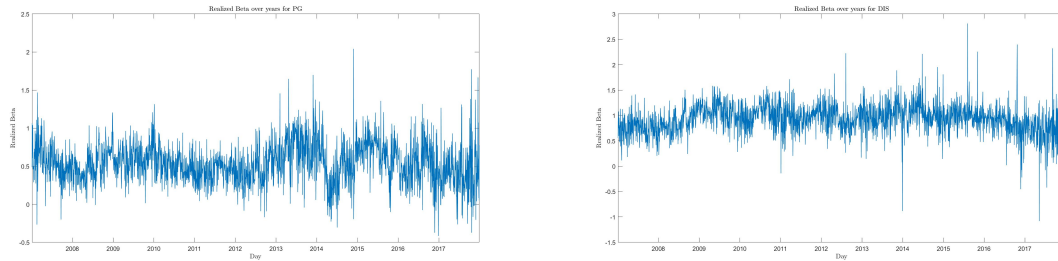


Figure 7: Realized beta of PG and DIS

From the figure, the realized beta of PG and DIS are volatility and do not show a reasonable stable trend. PG's realized beta mostly falls below the line $R\beta = 1$, which means that PG may have a lower beta than the market for most in trading day. This is verse in the case of DIS. As we can see from the figure, most of DIS' realized beta falls above or around the line $R\beta = 1$, this indicates that the beta of DIS may have a higher value than market in most of trading day.

(B) From figures in part A, the value of realized value for both stock vary throughout years. DIS stock's realized beta's volatility is smaller than PG's and we can find a value that the value of DIS' realized bate may move around. However, this isn't enough for us to assume the beta value is fixed. For PG, whose realized beta varies more, this assumption will be even more unreasonable.

(C) The follows are the plots of estimated confidence interval of PG and DIS.

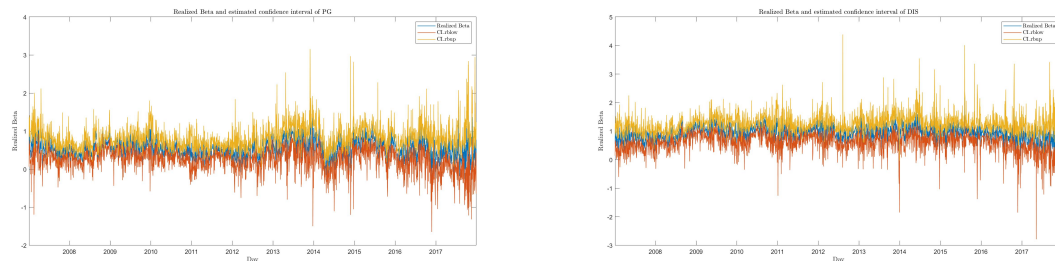


Figure 8: Realized beta and estimated confidence interval of realized beta of PG and DIS

The number of realized beta of PG falls into estimated confidence interval is 2759 and the cover rate is 99.64%. The number of realized beta of DIS falls into estimated confidence interval is 2769 and the cover rate is 100%. This result indicates that the estimated confidence interval is good and the accuracy of using bootstrap to estimate the confidence interval is higher than the method of using asymptotic distribution.

(D) The follows are the figures of realized beta of PG and DIS in January 2007.

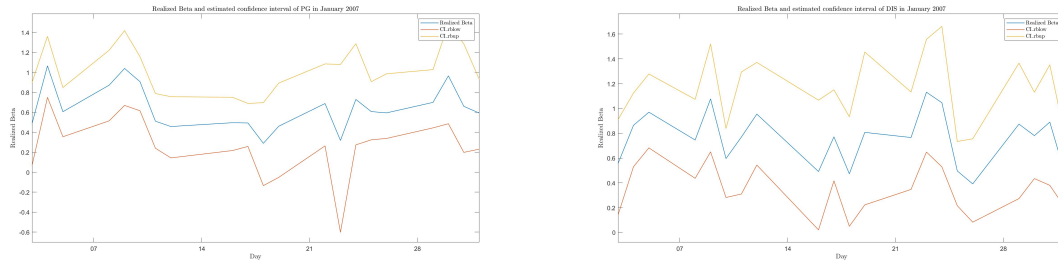


Figure 9: $R\beta$ and estimated CI of PG and DIS in January 2007

We can see from the figures that, the estimated confidence interval bounds the realized beta for both PG and DIS well. Look closer to the shape of confidence interval, we can find that the shape of upper bound and the lower bound is a little different, while the shape of upper bound and lower bound we get from asymptotic distribution method is similar.

- (E) (i) The confidence interval contains 1: $N_{1_PG} = 648$ and $N_{1_DIS} = 2091$.
(ii) The confidence interval is below 1: $N_{<1_PG} = 2118$ and $N_{<1_DIS} = 477$.
(iii) The confidence interval is above 1: $N_{>1_PG} = 3$ and $N_{>1_DIS} = 201$.

For PG stock, the number of confidence interval below 1 takes up 76.49%, which indicates PG stock will probably have a lower beta than the market for most of its trading day. A lower beta value means that the stock is less volatility than the market and has a lower risk and returns. For DIS stock, the number of confidence interval above or equal to 1 takes up 92.74%, which indicates DIS stock will probably have a higher or equal beta with the market for most of its trading day. A higher beta value means that the stock is more volatility than the market and has a higher risk and returns.

The **MATLAB**:

```

1  addpath('D:\ZM-Documents\MATLAB\data','functions','scripts');
2
3
4  [dates_PG,lp_PG]=load_stock('PG.csv','m');
5  N_PG=sum(floor(dates_PG(1,1))==floor(dates_PG(:,1)));% number of observations
   per day
6  T_PG=size(dates_PG,1)/N_PG;
7  [rdates_PG,lr_PG]=log_return([dates_PG lp_PG],N_PG,1);
8  days_PG=unique(floor(rdates_PG));
9  [lr_c_PG,lr_d_PG]=c_d_log_returns(lr_PG,N_PG,a);
10
11
12 [dates_SPY,lp_SPY]=load_stock('SPY.csv','m');
13 N_SPY=sum(floor(dates_SPY(1,1))==floor(dates_SPY(:,1)));% number of
   observations per day
14 T_SPY=size(dates_SPY,1)/N_SPY;
15 [rdates_SPY,lr_SPY]=log_return([dates_SPY lp_SPY],N_SPY,1);
16 days_SPY=unique(floor(rdates_SPY));
17 [lr_c_SPY,lr_d_SPY]=c_d_log_returns(lr_SPY,N_SPY,a);
18
19 %!!!
20 %if you want to check each scripts, you have to set the repeat time for this
   script
21 %otherwise, the code won't run successfully
22
23 a=4.5;
24 M=11;
25 n=N_PG-1;
26 kn=(N_PG-1)/M;
27 %(repeat time)
28 %r=10000;
29
30 %——part a
31 rbeta_PG=realized_beta(lr_c_PG, lr_c_SPY);
32
33 figure;
34 plot(days_PG,rbeta_PG);
35 xlabel('Day');
36 ylabel('Realized Beta');
37 title('Realized Beta over years for PG');
38 xlim([min(days_PG),max(days_PG)]);
39 datetick('x','keeplimits');
40
41 %part c

```

```
42 sbeta=zeros(r,T_PG);
43 parfor i=1:r
44     [newsample1,newsample2]=bsample_2(lr_c_PG,lr_c_SPY,kn,n,T_PG);
45     sbeta(i,:)=sum(newsample1.*newsample2)./sum(newsample2.^2);
46 end
47
48 CI_low_rbPG = quantile(sbeta, 0.025);
49 CI_up_rbPG = quantile(sbeta, 0.975);
50
51 %evaluate the estimated confidence interval
52 num_in_PG=sum(CI_low_rbPG <=rbeta_PG & CI_up_rbPG >=rbeta_PG);
53 cover_rate_PG=num_in_PG/T_PG;
54
55 figure;
56 plot(days_PG,rbeta_PG);
57 hold on;
58 plot(days_PG,CI_low_rbPG);
59 hold on;
60 plot(days_PG,CI_up_rbPG);
61 hold off;
62 xlabel('Day');
63 ylabel('Realized Beta');
64 title('Realized Beta and estimated confidence interval of PG');
65 xlim([min(days_PG),max(days_PG)]);
66 legend('Realized Beta','CI\_rb_low','CI\_rb_up');
67 datetick('x','keeplimits');
68
69 %part d(focus in a month)
70 d=21;
71 figure;
72 plot(days_PG(1:d),rbeta_PG(1:d));
73 hold on;
74 plot(days_PG(1:d),CI_low_rbPG(1:d));
75 hold on;
76 plot(days_PG(1:d),CI_up_rbPG(1:d));
77 hold off;
78 xlabel('Day');
79 ylabel('Realized Beta');
80 title('Realized Beta and estimated confidence interval of PG in January 2007'
81     );
81 xlim([min(days_PG(1:d)),max(days_PG(1:d))]);
82 ylim([min(CI_low_rbPG(1:d))-0.1,max(CI_up_rbPG(1:d))+0.1]);
83 legend('Realized Beta','CI\_rb_low','CI\_rb_up');
84 datetick('x',7,'keeplimits');
```

```
85 |  
86 | %part e  
87 | n1_PG=sum(CI_low_rbPG <=1 & CI_up_rbPG >=1);  
88 | n2_PG=sum( CI_up_rbPG < 1);  
89 | n3_PG=sum( CI_low_rbPG > 1);
```