Statistics 699 Report

Professor :ZhengJun Zhang

Name: Qiuying Li

12/22/2015

This study is to study the extreme co-movements in the stock market. In order to apply extreme co-movement measure (Zhang&Smith, 2010) into the real stock market, the first 100 stock price were chosen to be calculated into the formula below:

$$\lambda(t,T) = \lim_{\mathbf{u} \nearrow \mathbf{x}_F} \Pr\{\xi(t,T,\mathbf{u}) \geq 2 \mid \xi(0,t,\mathbf{u}) \geq 1\},$$

where $\mathbf{x}_F$ is the right end point of the distribution function F and

$$\xi(t,T,\mathbf{u}) = \max_{t \leq i \leq T} \sum_{d=1}^{D} I_{(Y_{id} > u_d)}.$$

Considering that there are more than 50000 different stocks in the data set, so it would be efficient to use one stock price as a specific example. Below is the price movement of the chosen stock.

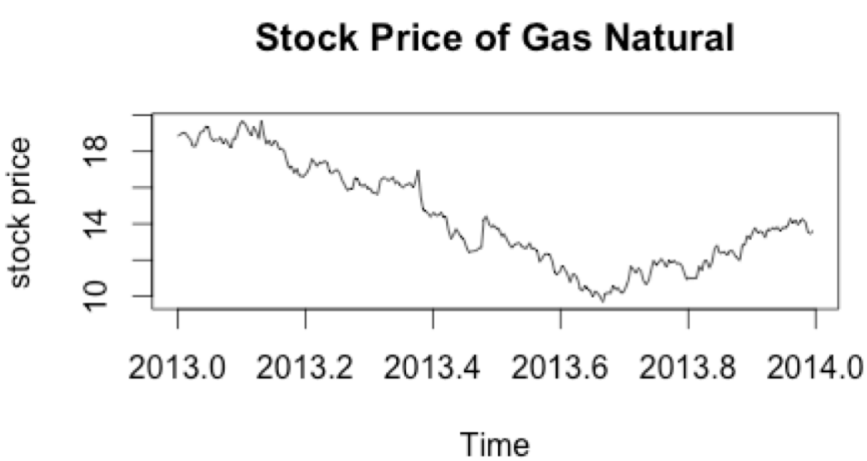

Figure 1: Daily stock price of Gas Natural

From the figure 1, the general trend of the stock price can be decided into two parts. Before June 2013, the stock prices continue decreasing, while after June 2013, the stock prices continue increasing.

The first step is to convert the daily price time series to logarithmic return series, due to most of the daily price time series are not stationary, and log returns are far superior to arithmetic returns since the sum of repeated samples from a population with a finite second moments approximates distributed.

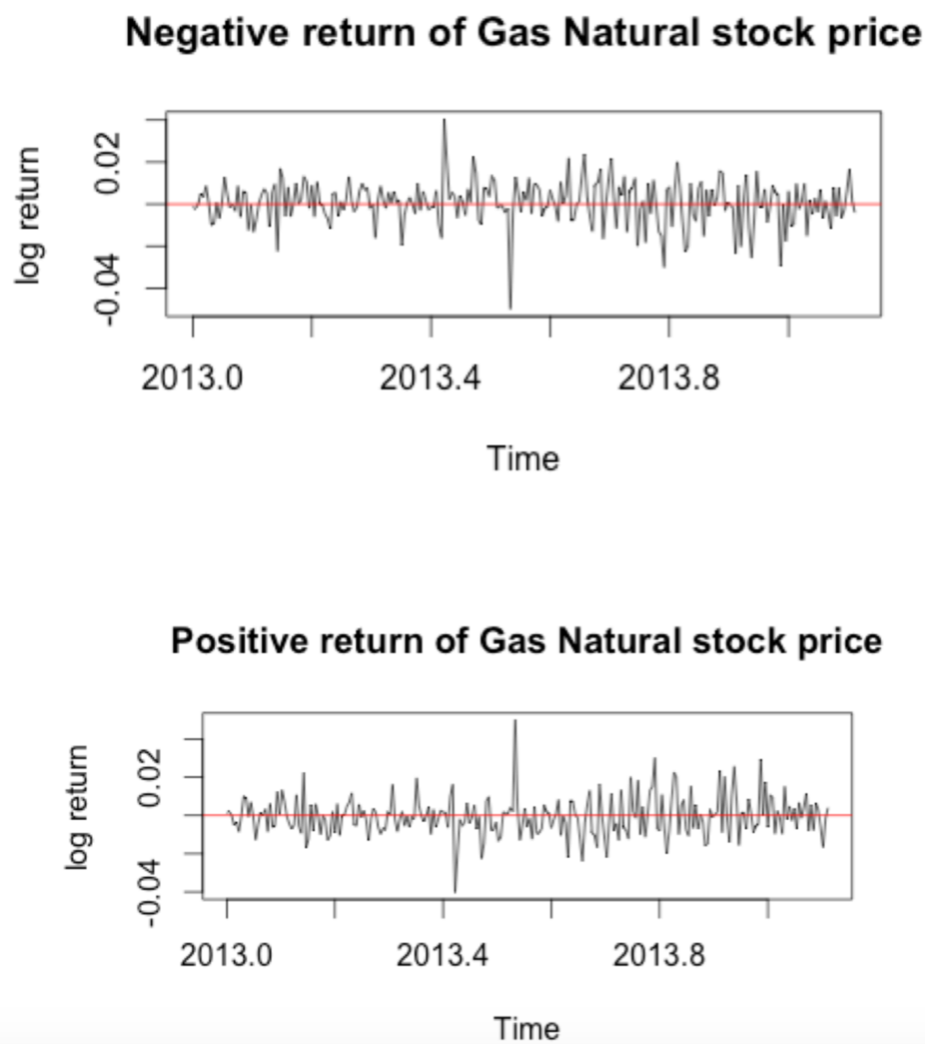

Figure 2: Log return of the stock price

From figure 2, we can notice that both negative returns and positive returns can be thought of stationary. It is known that financial return series tend to have certain degrees of autocorrelation. Moreover, we are also able to identify the extremal return observations, jumps in returns, and clustering in volatility from two return series (Zhang & Smith, 2010). There are two extreme jumps in the negative return, both appear around April 2013. Likewise, there are also two jumps appear around April 2013 in the positive return. In addition to the extreme jumps, there are also extremal observed values in each series, and some changes in volatility. Thus it is essential to propose a procedure to remove the volatility. To remove volatility, we propose a simple application of the GARCH (generalized autoregressive conditional heteroscedasticity) model. The GARCH(*1,1*) filtered conditional standard deviations in the Figure **3**.

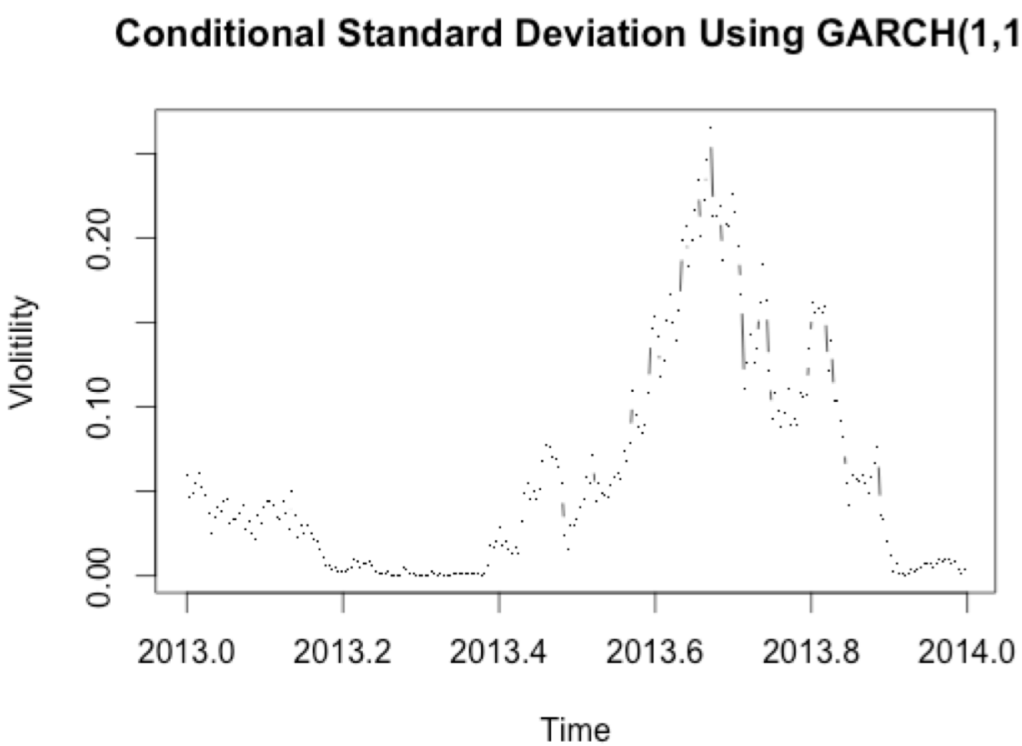

Figure 3: Conditional standard deviation under Garch(1,1)

Figure 3 shows the conditional standard deviations embedded in the raw return time series are extracted. It becomes obvious that the return time series demonstrate variations in volatility to certain degree. Moreover, it is also clear that peaks appear in the June 2013 to August 2013. According to Zhang, the conditional standard deviation time series reminds us of several familiar instances when the financial markets experience excess volatilities, such as the 1997-1998 Russian credit crisis and Asian financial crisis, and the 2001-2002 tech-bubble. Thus it would be valuable to check if there were any significant events happened that affected stock price during June 2013 to August 2013.

After getting the conditional standard deviation from GARCH(1,1) model, the next step is to get pseudo-observations by using original return time series to be divided estimated conditional standard deviations, part of the results of the first stock are as below:

1	18.859428	Yesterday	Today	negative log return	positive log return	volatility	negative pseudo- observations	positive pseudo- observations
2	18.922557	18.859428	18.922557	-0.001451306	0.001451306	0.00010494	-13.83025265	13.83025265
3	19.019359	18.922557	19.019359	-0.002216054	0.002216054	0.00010333	-21.44619846	21.44619846
4	19.023569	19.019359	19.023569	-9.61219E-05	9.61219E-05	0.00010185	-0.943786175	0.943786175
5	18.808921	19.023569	18.808921	0.004928116	-0.004928116	0.00010021	49.17981714	-49.17981714
6	18.657407	18.808921	18.657407	0.003512597	-0.003512597	9.89E-05	35.5272718	-35.5272718
7	18.295455	18.657407	18.295455	0.008508071	-0.008508071	9.74E-05	87.36589433	-87.36589433
8	18.27441	18.295455	18.27441	0.00049985	-0.00049985	9.69E-05	5.157503492	-5.157503492
9	18.699495	18.27441	18.699495	-0.009986514	0.009986514	9.54E-05	-104.725552	104.725552
10	19.090909	18.699495	19.090909	-0.008996729	0.008996729	9.63E-05	-93.45247451	93.45247451
11	19.061448	19.090909	19.061448	0.000670719	-0.000670719	9.67E-05	6.932837974	-6.932837974
12	19.368686	19.061448	19.368686	-0.00694427	0.00694427	9.52E-05	-72.95372382	72.95372382
13	19.335016	19.368686	19.335016	0.000755623	-0.000755623	9.49E-05	7.958725303	-7.958725303
14	18.77946	19.335016	18.77946	0.012661436	-0.012661436	9.34E-05	135.530281	-135.530281
15	18.57744	18.77946	18.57744	0.004697233	-0.004697233	9.46E-05	49.6439255	-49.6439255
16	18.640573	18.57744	18.640573	-0.001473391	0.001473391	9.34E-05	-15.7808743	15.7808743
17	18.636363	18.640573	18.636363	9.80971E-05	-9.80971E-05	9.20E-05	1.066218986	-1.066218986
18	18.775253	18.636363	18.775253	-0.003224637	0.003224637	9.06E-05	-35.60696789	35.60696789
19	18.413299	18.775253	18.413299	0.008454193	-0.008454193	8.95E-05	94.46110565	-94.46110565
20	18.674241	18.413299	18.674241	-0.006111354	0.006111354	8.92E-05	-68.52711732	68.52711732
21	18.413299	18.674241	18.413299	0.006111354	-0.006111354	8.88E-05	68.81381813	-68.81381813
22	18.190236	18.413299	18.190236	0.005293272	-0.005293272	8.79E-05	60.20255145	-60.20255145
23	18.712119	18.190236	18.712119	-0.012284637	0.012284637	8.69E-05	-141.3559285	141.3559285
24	18.733163	18.712119	18.733163	-0.000488141	0.000488141	8.91E-05	-5.477554455	5.477554455
25	19.309764	18.733163	19.309764	-0.013165854	0.013165854	8.78E-05	-150.0139401	150.0139401

Since we have the negative pseudo observations and positive pseudo observations, the next task is to fit the exceedance data within the pseudo-observation series over some high thresholds (u) to a generalized extreme value (GEV) distribution. In this study, threshold u is chosen to be 1.2, which has been used financial return series if gives approximately 10% of exceedances. The parametric formula of the GEV distribution is employed in a maximum likelihood estimation (MLE) to determine the three parameters, the location parameter μ, the scale parameter log psi, and the shape parameter xi, in the GEV distribution. As mentioned in the methodology section, the GEV fitting is applied to both the positive and negative pseudo-observation time series, separately (Zhang & Smith, 2010).

negative pseudo- observations						Positive pseudo- observations			
Prem Num	Num mu	mu	log psi	xi		Num mu	mu	log psi	xi
100001	105	306.4856	4.009124	-0.04913		128	333.322167	4.323776	0.028966
100025	112	347.693789	3.768367	-0.188902		138	405.581523	4.63823	-0.14844
100026	89	282.929379	3.998817	-0.021398		126	265.22761	4.058269	0.01626
100032	127	382.042368	3.884073	-0.26155		140	383.501321	3.764554	-0.224323
100044	80	289.831835	3.729996	-0.142679		126	306.159241	3.976416	-0.06758
100051	119	267.09065	3.286134	-0.268387		138	514.182926	4.659922	-0.001812
100065	108	528.248056	4.496446	-0.083275		133	398.440094	3.72833	-0.249371
100104	111	511.428082	4.375693	0.031274		124	687.299449	5.190375	0.101462
100107	126	326.511013	4.260227	0.0119		128	409.735796	4.648559	0.094207
100113	73	845.602563	5.005593	-0.056512		122	903.211578	5.306675	0.024009
100138	120	666.675039	4.745431	-0.085931		143	558.924853	4.155873	-0.213581
100145	110	629.448507	4.528212	-0.14325		133	721.810006	4.574007	-0.152706
100147	118	633.419655	4.648219	-0.093377		143	664.215285	4.810578	-0.044374
100151	126	1276.616547	5.359237	-0.094738		133	1065.076812	5.062508	-0.124129
100158	114	400.227323	4.371952	-0.013092		81	337.518955	3.648309	-0.286973
100180	108	336.943838	4.162848	-0.038012		133	421.559071	4.326161	-0.052882
100200	86	283.129539	3.645922	-0.176747		131	378.56324	4.163638	-0.071331
100201	118	196.30093	4.135921	0.202863		74	130.695778	3.357823	0.011502
100207	107	811.06856	4.87387	-0.116839		97	661.290919	4.293037	-0.232285
100220	123	567.727871	4.062014	-0.277783		133	654.367366	4.426263	-0.174953
100223	89	547.91497	4.759866	-0.018403		130	1108.866269	5.975617	0.23461
100225	127	705.983413	4.810724	-0.076691		121	582.651123	3.868807	-0.341188

We expect the shape parameter xi to be greater than 0, which is the evidence to show that the probability distributions of the underlying financial asset returns are heavy-tailed. When xi parameters greater than 0, the conventional thin-tailed distribution based modeling approach turns out to be not appropriate. However, in the table above, most of the xi parameters are negative, which indicates that the probability distributions of the underlying financial asset returns have upper bound. This is an interesting phenomenon that worths further studying.

As mentioned before, the positive and negative pseudo-observation series are fitted into GEV distributions separately. Accordingly, they are transformed into the unit Fréchet distribution separately as well, figure 5 is the positive daily return on Fréchet scale, and figure 6 is negative daily return on Fréchet scale.

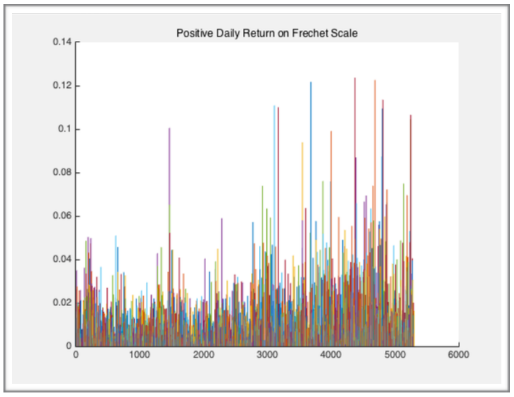

Figure 5

(mu=306.4856, log psi=4.009124, xi=-0.04913)

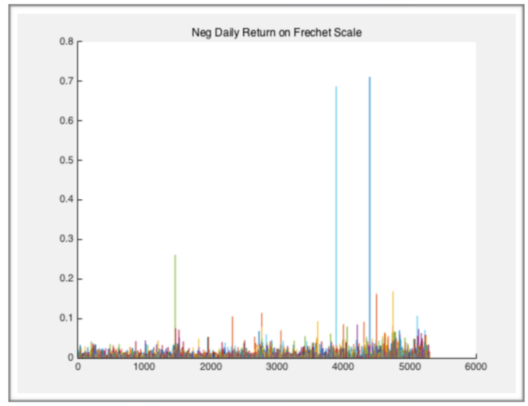

Figure 6:

(mu=333.322167,log psi=4.323776,xi=0.028966)

After calculating Fréchet Scale of 100 stocks, the next step is to rank all the Fréchet scale from order of small to large. Then the 90th percentile Ud is 0.0222, the 95th Ud is 0.0261, and the 97.5th percentile Fréchet scale is 0.0291.

According to the formula below:

$$\lambda(t,T) = \lim_{\mathbf{u} \nearrow \mathbf{x}_F} \Pr\{\xi(t,T,\mathbf{u}) \geq 2 | \xi(0,t,\mathbf{u}) \geq 1\},$$

where $\mathbf{x}_F$ is the right end point of the distribution function F and

$$\xi(t,T,\mathbf{u}) = \max_{t \leq i \leq T} \sum_{d=1}^{D} I_{(Y_{id} > u_d)}.$$

Thus the idea is to estimate the maximum number $\xi(t,T,\mathbf{u})\geq 2$ of joint exceedances in the time period t to T given at least one exceedance in (0, t) (Embrechts et al. 2003). In this study T-t =1, T-t = 2, T-t = 3, T-t = 4, T-t = 5, T-t = 6, T-t = 7.

Based on the formula, the first step is to get the number $\xi(0,t,\mathbf{u})\geq 1$ on the 3 Ud on different percentiles among 100 stocks at the same time.

Percentile	Ud	$\xi(0,t,\mathbf{u})\geq 1$
90th Percentile	0.0222	13
95th Percentile	0.0261	11
97.5th Percentile	0.0291	9

The next step is to get the number of the union between $\xi(t,T,\mathbf{u})\geq 2$ and $\xi(0,t,\mathbf{u})\geq 1$

The formula above means, given the negative return of stock price exceeds the Ud, these negative return need to also exceed Ud when T-t =1, T-t = 2, T-t =3, T-t = 4, T-t =5, T-t = 6, and T-t = 7.

Percentile	Ud	T-t =1	T-t =2	T-t =3	T-t =4	T-t =5	T-t =6	T-t =7
90th	0.0222	4	4	3	2	2	2	2
95th	0.0261	4	3	3	2	1	1	1
97.5th	0.0291	4	3	1	1	1	1	1

The final step is to get the conditional probability : $\Pr\{\xi(t,T,\mathbf{u})\geq 2|\xi(0,t,\mathbf{u})\geq 1\}$ which is simply used the results from second step decided by the results of first step.

Percentile	Ud	T-t =1	T-t =2	T-t =3	T-t =4	T-t =5	T-t =6	T-t =7
90^{th}	0.0222	0.3077	0.3077	0.2308	0.1583	0.1583	0.1583	0.1583
95^{th}	0.0261	0.3636	0.3108	0.3108	0.2727	0.1818	0.0909	0.0909
97.5^{th}	0.0291	0.4444	0.3333	0.2222	0.1111	0.1111	0.1111	0.1111

Conclusion:

By transforming the daily price into log return series, and then calculated the parameters that are necessary to Fréchet Scale transformation, and applied Fréchet Scale transformation results into the extreme co-movement measure, one would be able to find the probability of stock decrease decrease from today to a T time lag above certain threshold.