量化风险管理作业二

蔡宇超　2018000735

July 1, 2019

本次作业采用的时间序列为2001年2 月19日至2018年3月16日的沪深300 数据，数据长度为4145。首先通过收盘价计算收益率r_t，我们以$X_t = -r_t$作为研究对象，表示损失，作出X_t的图像如下：

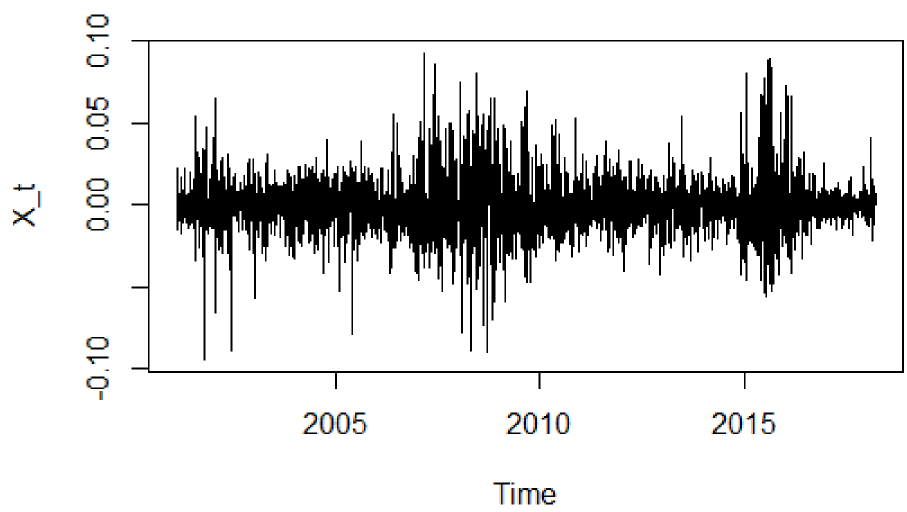

Figure 1 X_t波动图像

从图中可以看出，X_t的波动表现出明显的集群效应，考虑采用ARMA-GARCH模型进行拟合，为了确定合适的阶数，首先采用新息分布为正态的AMAR(1,1)-GARCH(1,1)模型，拟合结果如下：

GARCH Model : sGARCH(1,1)

Mean Model : ARFIMA(1,0,1)

Distribution : norm

Optimal Parameters

Estimate Std. Error t value $Pr(> |t|)$

mu -0.000234 0.000179 -1.30200 0.192918

ar1 -0.767780 0.128155 -5.99104 0.000000

ma1 0.788698 0.122602 6.43301 0.000000

omega 0.000001 0.000003 0.43482 0.663691

alpha1 0.073802 0.031582 2.33681 0.019449

beta1 0.925167 0.030324 30.50988 0.000000

从拟合结果来看，除了μ和ω参数不显著之外，其余参数均是显著的。我们观察改变ARMA-GARCH模型阶数对估计结果造成的影响，拟合新息分布为正态的ARMA(1,2)-GARCH(1,1)模型如下：

GARCH Model : sGARCH(1,1)

Mean Model : ARFIMA(1,0,2)

Distribution : norm

Optimal Parameters

Estimate Std. Error t value $Pr(>|t|)$

mu -0.000234 0.000179 -1.30615 0.191502

ar1 -0.752128 0.158466 -4.74631 0.000002

ma1 0.770589 0.159307 4.83713 0.000001

ma2 -0.003929 0.018766 -0.20935 0.834172

omega 0.000001 0.000003 0.44022 0.659781

alpha1 0.073864 0.031200 2.36741 0.017913

beta1 0.925125 0.029943 30.89606 0.000000

从拟合结果来看，增加ARMA的阶数对参数估计结果的影响很小，并且此时新增的移动平均项的估计系数ma2不显著，这表明采用低阶的ARMA(1,1)-GARCH(1,1)模型要比ARMA(1,2)-GARCH(1,1)模型更合适。

基于同样的理由，我们发现ARMA(1,1)-GARCH(1,1)模型比ARMA(2,1)-GARCH(1,1),ARMA(1,1)-GARCH(1,2),ARMA(1,1)-GARCH(2,1) 模型都合适，因此我们认为ARMA(1,1)-GARCH(1,1)的阶数是合适的。作出观测序列和残差的自相关图，右上角的Q-Q图表明新息为正态的假定不太合适，下方的残差及残差平方的自相关图说明采用的ARMA-GARCH是合理的。

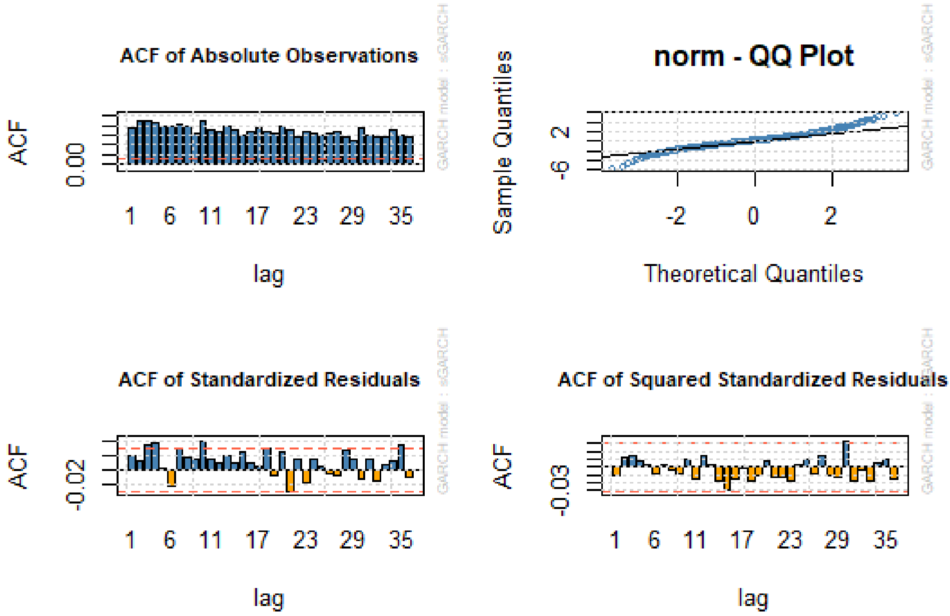

Figure 2 新息分布为正态分布的ARMA(1,1)-GARCH(1,1)模型拟合结果，左上方为$|X_t|$自相关图，右上方为Z_tQ-Q 图，左下方为$|Z_t|$ 自相关图，右下方为Z_t^2自相关图。

假定新息分布为t分布，拟合的参数如下：

GARCH Model : sGARCH(1,1)

Mean Model : ARFIMA(1,0,1)

Distribution : std

Optimal Parameters

———————————————

Estimate Std. Error t value $Pr(>|t|)$

mu -0.000361 0.000165 -2.1885 0.028631

ar1 -0.777520 0.142114 -5.4711 0.000000

ma1 0.794751 0.136946 5.8034 0.000000

omega 0.000001 0.000001 1.5434 0.122737

alpha1 0.066138 0.008091 8.1742 0.000000

beta1 0.932862 0.007725 120.7661 0.000000

shape 4.771906 0.348014 13.7118 0.000000

此时图3中的Q-Q图看上去更为合理，并且此时拟合的ARMA-GARCH模型也是合适的。计算可得新息分布为t分布的AIC和BIC为-24085.12和-24040.81，比新息分布为高斯分布的AIC(-23777.46)和BIC(-23739.48)小。因此我们最终选择新息分布为t分布的ARMA(1,1)-GARCH(1,1)模型。

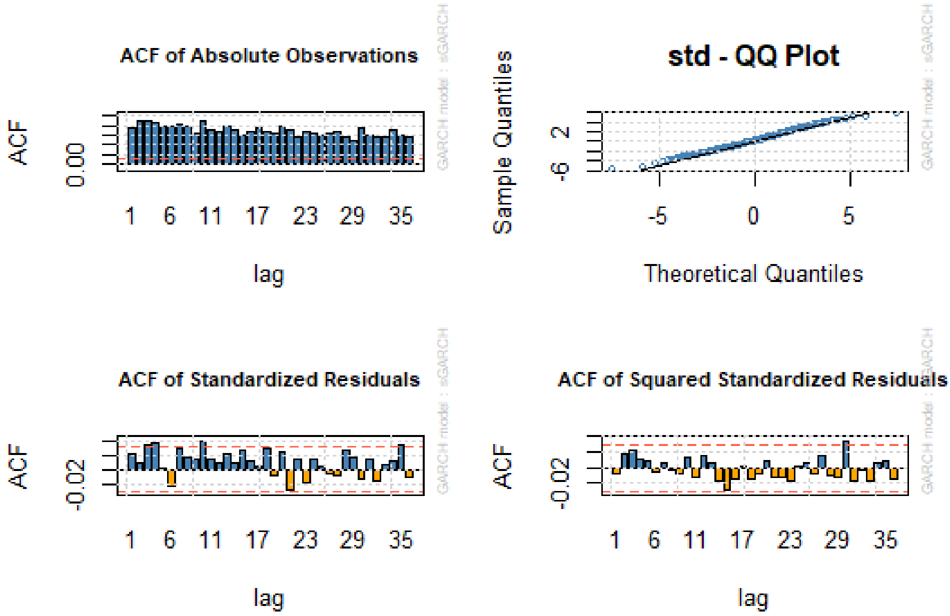

Figure 3 新息分布为t分布的ARMA(1,1)-GARCH(1,1)模型拟合结果，左上方为$|X_t|$自相关图，右上方为Z_tQ-Q 图，左下方为$|Z_t|$ 自相关图，右下方为Z_t^2自相关图。

考虑区块极值，以$n = 60, 120, 240$ 作为区块长度，对GEV分布参数ξ, μ, σ估计。

	$\hat{\xi}$	$\hat{\mu}$	$\hat{\sigma}$
60	0.1741(0.1095)	0.0280(0.0020)	0.0143(0.0015)
120	0.0353(0.2123)	0.0366(0.0035)	0.0168(0.0027)
240	-0.0662(0.2966)	0.0444(0.0052)	0.0175(0.0037)

Table 1 区块长度为60,120,240 时，参数ξ,μ,σ的估计值，括号中的数表示估计标准误。

从表中可以看出，随着区块内样本数量的增加，此时对GEV的估计更为精确，估计的偏差下降，但是这样会降低极大似然估计的准确性，导致估计的方差变大，通过三种区块大小的估计结果，我们认为$\xi > 0$，换言之，对应的GEV分布是一个Fréche分布。计算$r_{240,10} = 0.0810$，也就是平均10个区块内至少出现一次的最小极值水平为0.0810，计算可得$k_{240,0.05} = 1.9434$，这表明，平均1.9434个区块的极值比0.05来得大，注意到我们拟合出来的Z的取值范围较大，且波动剧烈，因此这样的结果是合理的。根据GEV分布拟合得到的相关参数，我们对Z的VaR指标进行估计，前面的分析表示$\xi \neq 0$，因此取区间长度$n = 120$的情形，我们得到此时的0.95分位数，也就是$\alpha = 0.95$时，

$$VaR_{0.95}^{Z-GEV} = \hat{\mu} - \frac{\hat{\sigma}}{\hat{\xi}}(1 - (-n \log \alpha)^{-\hat{\xi}}) = 0.007066971$$

从而

$$VaR_{0.95}^{X-GEV} = \mu_{t+1} + \sigma_{t+1} VaR_{0.95}^{Z-GEV} = -0.0002189993$$

右上角的“GEV”表示基于GEV方法的估计结果，下文类似的记号同理。接下来采用GPD分布族对极值进行分析，从样本关于不同阈值的超额平均图来看，在0.03 附近图像趋近于一条直线，因此我们选取0.03作为阈值。另外从阈值对GPD分布族形状参数ξ 的影响来看，在阈值大于0.03之后，估计结果趋于稳定。在阈值等于0.03的条件下，求得$\hat{\xi} = -0.1490, \hat{\beta} = 0.0185$，估计的标准误为0.0959，0.0023。另外从拟合的GPD分布来看，估计的效果是较好的。

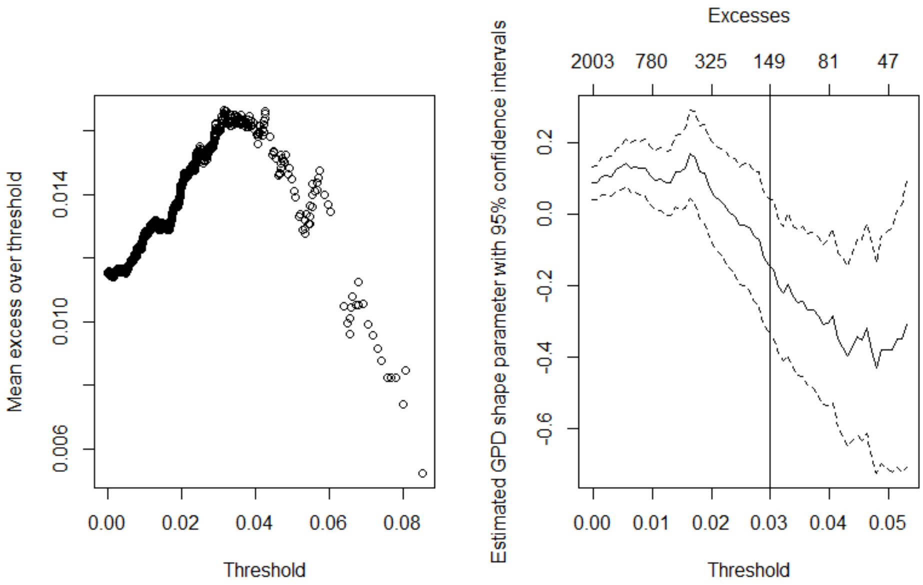

Figure 4 左图为样本关于阈值的超额平均图，右图表示不同阈值下GPD分布形状参数ξ的估计值及95%置信区间。

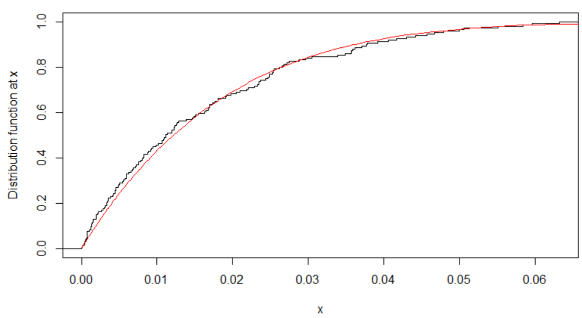

Figure 5 红色曲线表示拟合的$\hat{G}_{\xi,\beta}$，黑色折线表示样本的经验超额。

通过R中相关函数可求得$VaR_{0.95}^{Z-GPD} = 0.03$和$ES_{0.95}^{Z-GPD} = 0.0461$，利用关系

$$VaR_{0.95}^{X-GPD} = \mu_{t+1} + \sigma_{t+1} VaR_{0.95}^{Z-GPD} \quad ES_{0.95}^{X-GPD} = \mu_{t+1} + \sigma_{t+1} ES_{0.95}^{Z-GPD}$$

可得$VaR_{0.95}^{X-GPD} = 0.0000312, ES_{0.95}^{X-GPD} = 0.0002065$。

对Z_t采用hill估计方法，首先做出$\{k, \hat{\alpha}_{k,n}^{(H)}\}$ 的图像，可以看出，在$k = 150$附近，曲线趋于

稳定，在右图中将横坐标截断到250，可以发现此时$\hat{\alpha}_{k,n}^{(H)}$在2.5附近波动，因此选取$k = 150$，则

$$\hat{\alpha}_{k,n}^{(H)} = (\frac{1}{k}\sum_{i=1}^{k} \log X_{i,n} - \log X_{k,n})^{-1} = 2.614938,\ X_{1,n} \geq \ldots \geq X_{n,n}$$

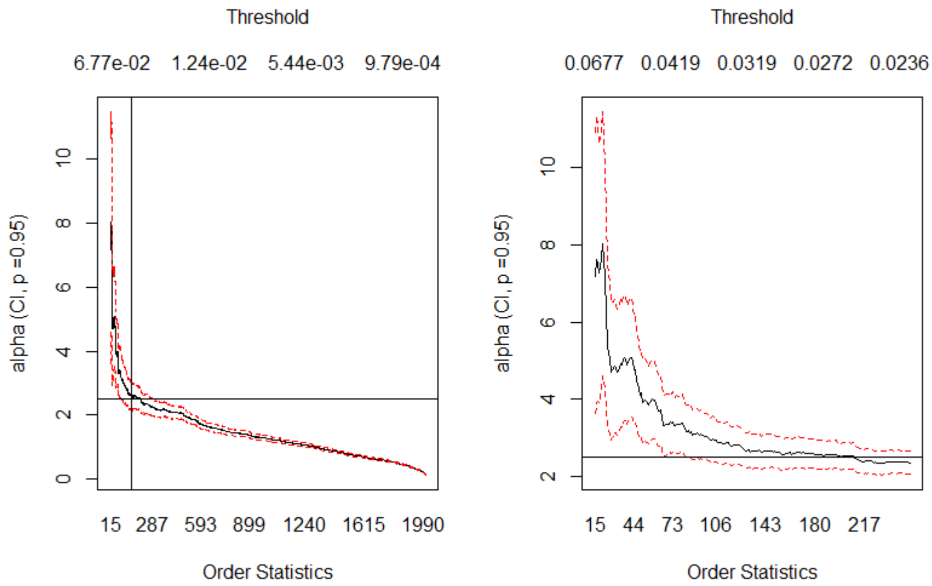

Figure 6 Hill图，左图中竖线的坐标为150，横线的坐标为2.5，右图中横线的坐标为2.5。

求得的结果与图像是吻合的，从而可求得Z的VAR和ES估计为

$$VaR_{0.95}^{Z-hill} = (\frac{n}{k}(1-\alpha))^{-\frac{1}{\hat{\alpha}_{k,n}^{(H)}}} X_{k,n} = 0.02646318$$

$$ES_{0.95}^{Z-hill} = \frac{\hat{\alpha}_{k,n}^{(H)}}{\hat{\alpha}_{k,n}^{(H)} - 1} VaR_{0.95}^{Z-hill} = 0.04284968$$

从而有$VaR_{0.95}^{X-hill} = -0.000007386657$, $ES_{0.95}^{X-hill} = 0.00017139$。可以发现Hill估计的结果要比GPD估计的结果来的更小些。