

## 同仁堂日个股回报率分析

一、下载一只股票或市场指数 3 年的日收益率数据，计算样本偏度、峰度，利用 Q-Q 图和 Jarque-Bera 进行正态性检验。

采用的数据为同仁堂股票 2016 年 3 月 23 日到 2019 年 3 月 23 日的**考虑现金红利再投资的日个股回报率**，数据来源为国泰安经济金融研究数据库 (<http://www.gtarsc.com>)。图 1 展示了收益率数据的时序图，图二展示了数据的直方图。可以发现数据大多集中在 0 附近，表现出波动聚集，数据分布略微呈右偏。

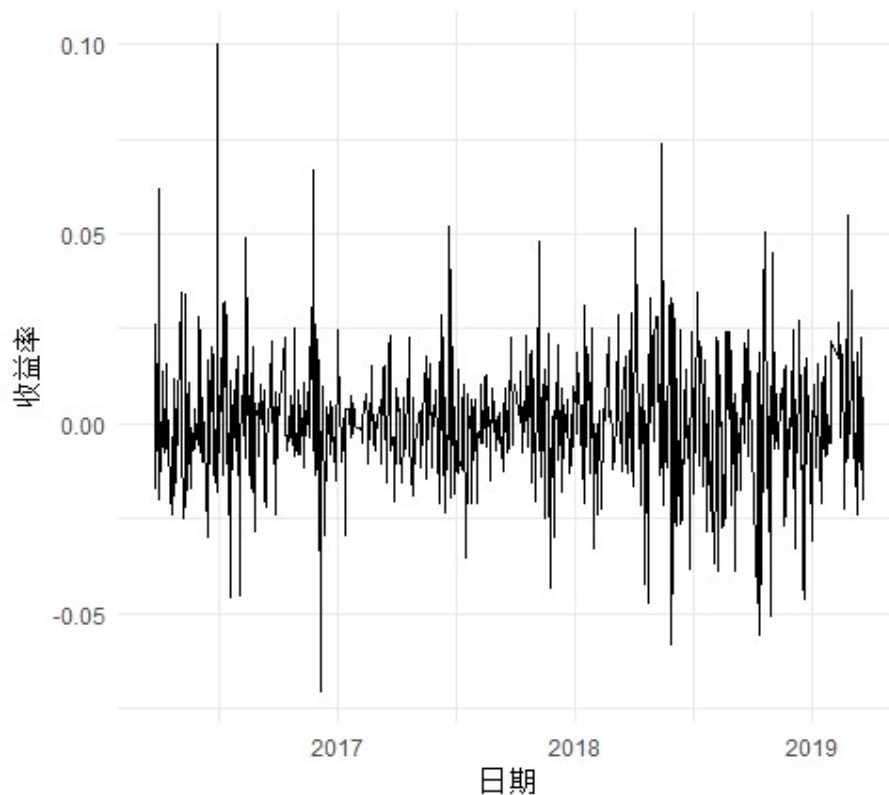


图 1 同仁堂 2016 年 3 月 23 日到 2019 年 3 月 23 日的收益率序列数据

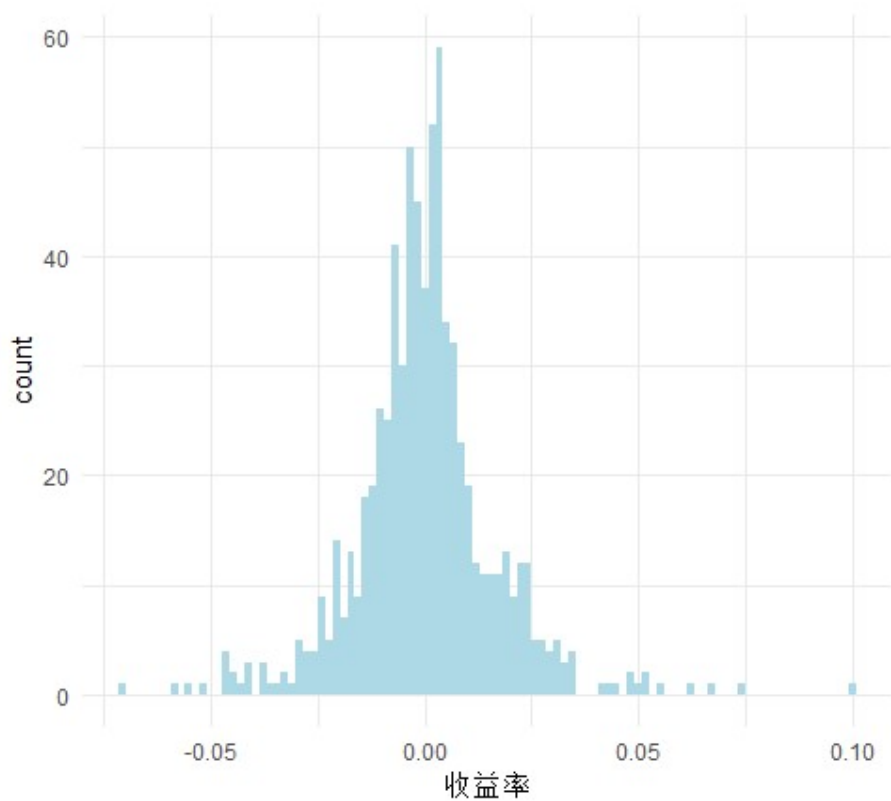


图 2 同仁堂 2016 年 3 月 23 日到 2019 年 3 月 23 日的收益率数据直方图

### 1. 计算

计算得到样本的偏度为 0.390，峰度为 7.031，峰度远大于 3，样本表现出略微的右偏和尖峰的特性。图 3 为收益率数据的 Q-Q 图，数据并不能和正态分布很好的拟合，而是呈现倒“S”型，样本相较于正态分布含有更多的极端值，表现出“厚尾”的特性。

### 2. 正态性检验

进行 Jarque-Bera 正态性检验，统计量的值为 514.27，p 值小于  $2.2e-16$ ，拒绝正态性的原假设。

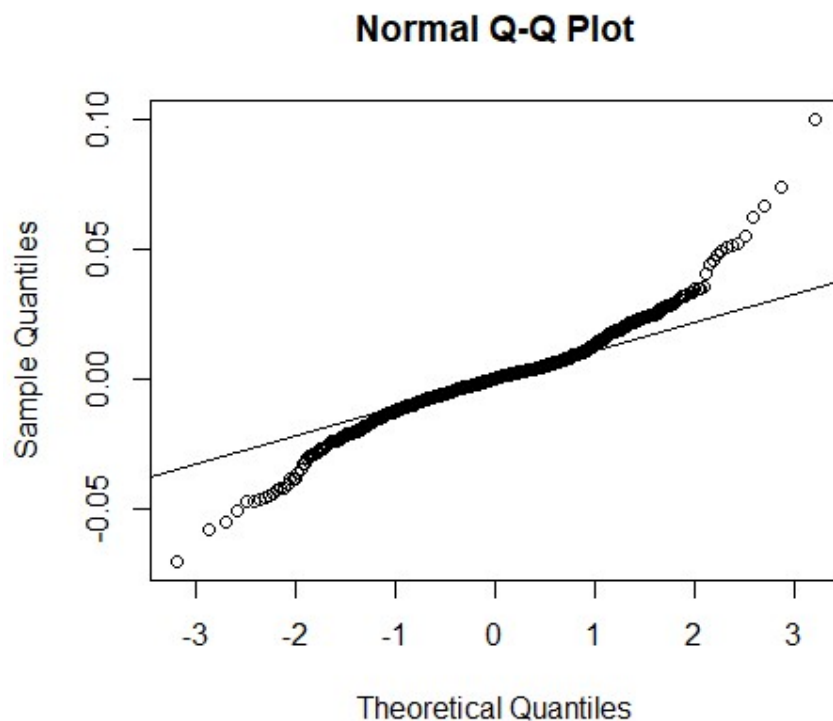


图 3 收益率 Q-Q 图

## 二、绘制收益率和收益率绝对值的 ACF 图像，进行分析和 Ljung-Box 检验。

图 4、图 5 分别展示了收益率和收益率绝对值的自相关图，对收益率而言，只有延迟 2、3 期自相关系数比较显著；对收益率绝对值而言，延迟 3-10 期自相关系数，除了延迟 8 期自相关系数外，都显著大于 0，数据表现出短期相关性。两幅图都表明数据存在一定的序列相关性。

进一步，取延迟期数分别为 2、4、6、8、10，对收益率数据和收益率绝对值做 Ljung-Box 纯随机性检验。表 1 展示了检验的 p 值。显著性水平设为 0.05，收益率数据在所有延迟期数值上都拒绝了数据是纯随机的原假设，收益率绝对值在延迟期数较小时表现出纯随机性，但随着延迟期数增大，p 值小于 0.05，拒绝纯随机性的原假设。

表 1 Ljung-Box 检验 p 值

Lag	2	4	6	8	10
收益率	0.005	9.287e-05	6.709e-05	1.658e-04	5.711e-04
绝对收益率	0.203	4.454e-04	2.071e-05	2.998e-06	1.94e-10

收益率自相关图

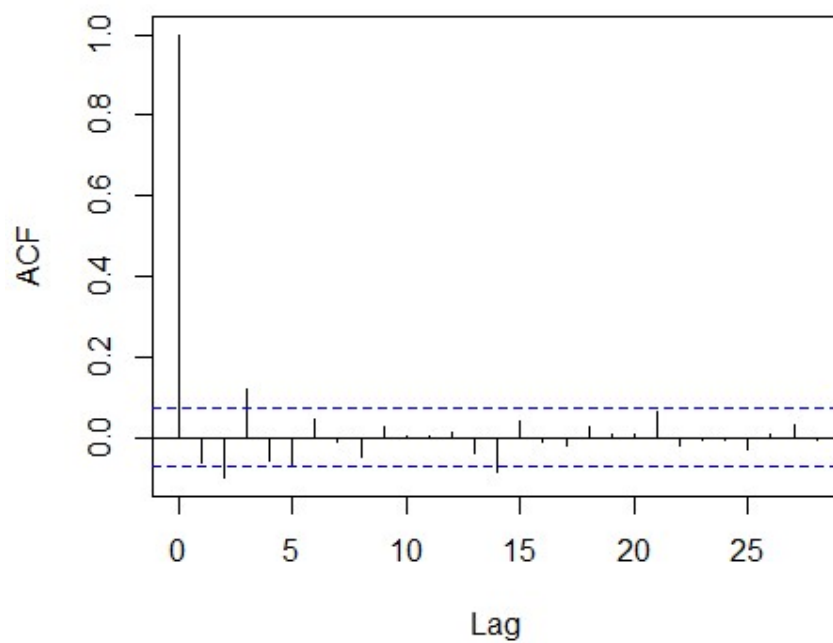


图 4 收益率自相关图

收益率绝对值自相关图

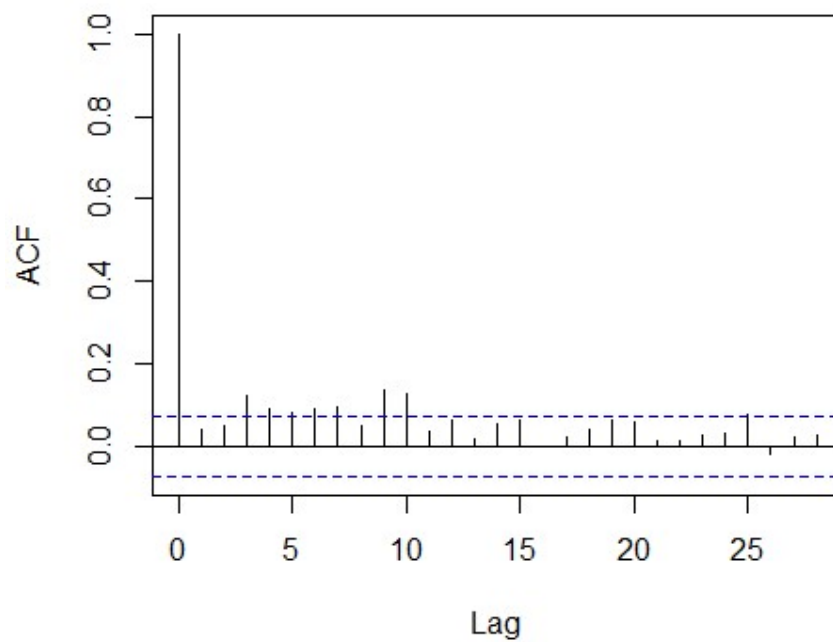


图 5 收益率绝对值自相关图

### 三、分别利用高斯新息和 $t$ 分布新息的 ARMA(1, 1)-GARCH(1, 1) 模型进行拟合，对标准化残差进行分析（ACF 和 Q-Q 图），利用 AIC 和 BIC 比较高斯新息和 $t$ 分布新息模型。

分别利用高斯新息和  $t$  分布新息进行模型拟合，参数估计值如表 2 所示，所有参数估计值在 0.05 的显著性水平下都显著。对标准化残差绘制自相关图和 Q-Q 图，结果如图 6 所示。计算的 AIC、BIC 如表 3 所示。无论从 Q-Q 图还是 AIC、BIC 的值来看，都是  $t$  分布新息拟合得更好，但是二者都没有充分提取序列的相关性，表现在残差仍然存在显著的序列相关性。

进一步，对标准化残差做 Ljung-Box 检验，取延迟期数为 6，检验  $p$  值分别为 0.049 和 0.023，在 0.05 的显著性水平下，拒绝原假设，标准化残差存在序列相关性。对标准化残差的平方做 Ljung-Box 检验， $p$  值分别为 0.724、0.811，残差序列没有表现条件异方差。

因此，根据对数据自相关图和偏自相关图的进一步观察，重新使用 ARMA(3, 3)-GARCH(1, 1) 模型进行拟合，标准化残差的自相关图和 Q-Q 图如图 7 所示。对标准化残差做 Ljung-Box 检验， $p$  值分别为 0.946，0.437。对标准化残差的平方做 Ljung-Box 检验， $p$  值分别为 0.762，0.839。标准化残差没有表现出序列相关和条件异方差。计算的 AIC、BIC 如表 4 所示。

可以看到， $t$  分布新息的拟合结果仍然优于高斯新息。与 ARMA(1, 1)-GARCH(1, 1) 模型对比，ARMA(3, 3)-GARCH(1, 1) 模型 AIC 值更低，但由于参数变多，而 BIC 准则对参数数量加罚的权重更大，导致 ARMA(3, 3)-GARCH(1, 1) 模型 BIC 值更高，但是，残差不再表现显著的序列相关性，序列相关性提取更充分。

表 2 ARMA(1, 1)-GARCH(1, 1) 模型参数估计值

	ar1	ma1	alpha0	alpha1	beta1
高斯新息	0.719 (0.239)	-0.778 (0.217)	9e-06 (1e-06)	0.088 (0.010)	0.885 (0.015)
$t$ 分布新息	0.753 (0.139)	-0.826 (0.121)	8e-06 (4e-06)	0.072 (0.007)	0.919 (0.019)

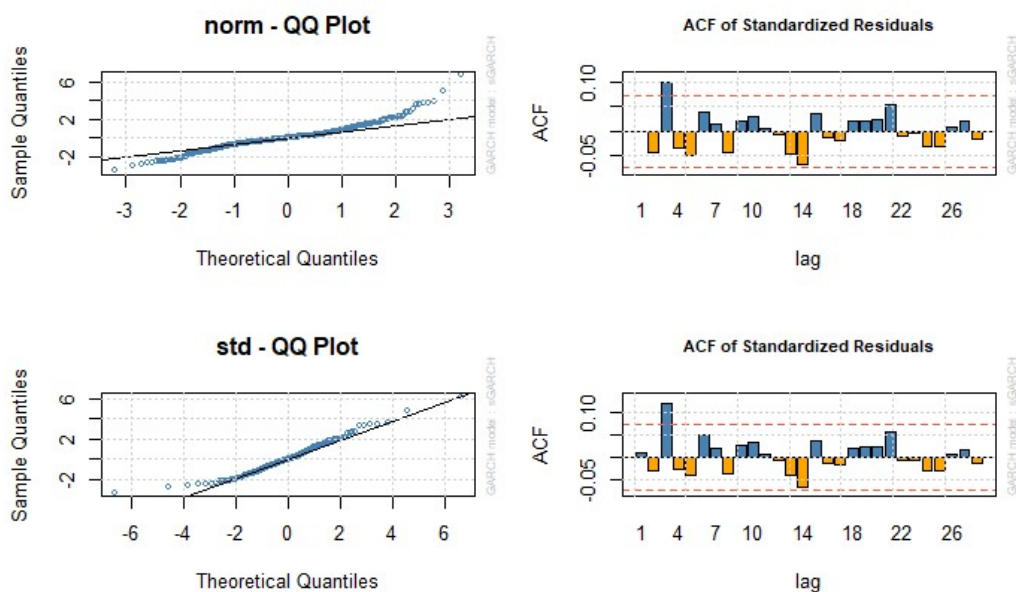


图 6 ARMA(1,1)-GARCH(1,1)模型标准化残差的自相关图和 Q-Q 图

表 3 ARMA(1,1)-GARCH(1,1)模型拟合的 AIC、BIC 值

	高斯新息	t 分布新息
AIC	-3978.095	-4097.635
BIC	-3955.116	-4070.06

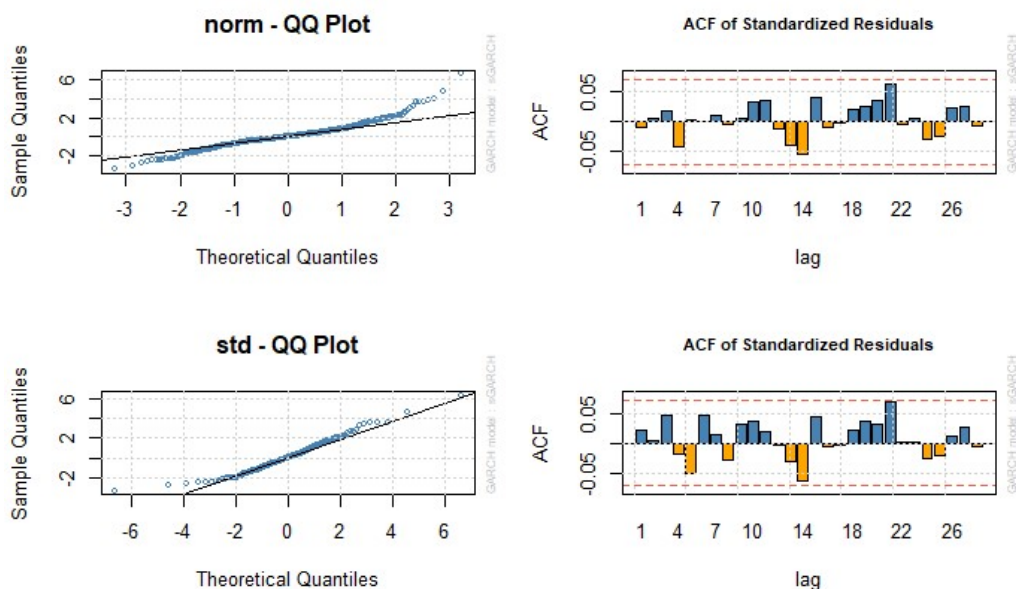


图 7 ARMA(3,3)-GARCH(1,1)模型标准化残差的自相关图和 Q-Q 图

表 4 ARMA(3,3)-GARCH(1,1)模型拟合的 AIC、BIC 值

	高斯新息	t 分布新息
AIC	-3978.095	-4097.635
BIC	-3955.116	-4070.06

AIC	-3981.824	-4099.269
BIC	-3940.462	-4053.311

四、基于 3 中的模型，预测一天后的收益率以及方差，计算 VaR 和 ES 指标 ( $\alpha=0.95$ )。

表 5、6 分别展示了 ARMA(1, 1)-GARCH(1, 1) 模型和 ARMA(3, 3)-GARCH(1, 1) 模型的预测结果和相应的 VaR 和 ES 指标。t 分布比高斯分布尾部更厚，高分位数数值更大，使得高斯新息稍微高估了 VaR 指标，大大低估了 ES 指标。ARMA(1, 1)-GARCH(1, 1) 模型和 ARMA(3, 3)-GARCH(1, 1) 模型的 VaR、ES 指标差异不大。

**表 5 ARMA(1, 1)-GARCH(1, 1) 模型预测结果及 VaR、ES 指标**

	预测值	Sigma	VaR	ES
高斯新息	7.338e-04	0.016	0.027	0.033
t 分布新息	7.042e-04	0.018	0.025	0.061

**表 6 ARMA(3, 3)-GARCH(1, 1) 模型预测结果及 VaR、ES 指标**

	预测值	Sigma	VaR	ES
高斯新息	-2.488e-03	0.016	0.024	0.033
t 分布新息	-1.288e-03	0.018	0.023	0.062

---

感谢中国人民大学统计学院吴文琦同学提供的数据、代码和报告