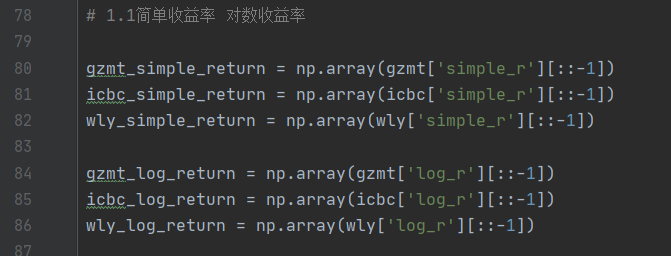
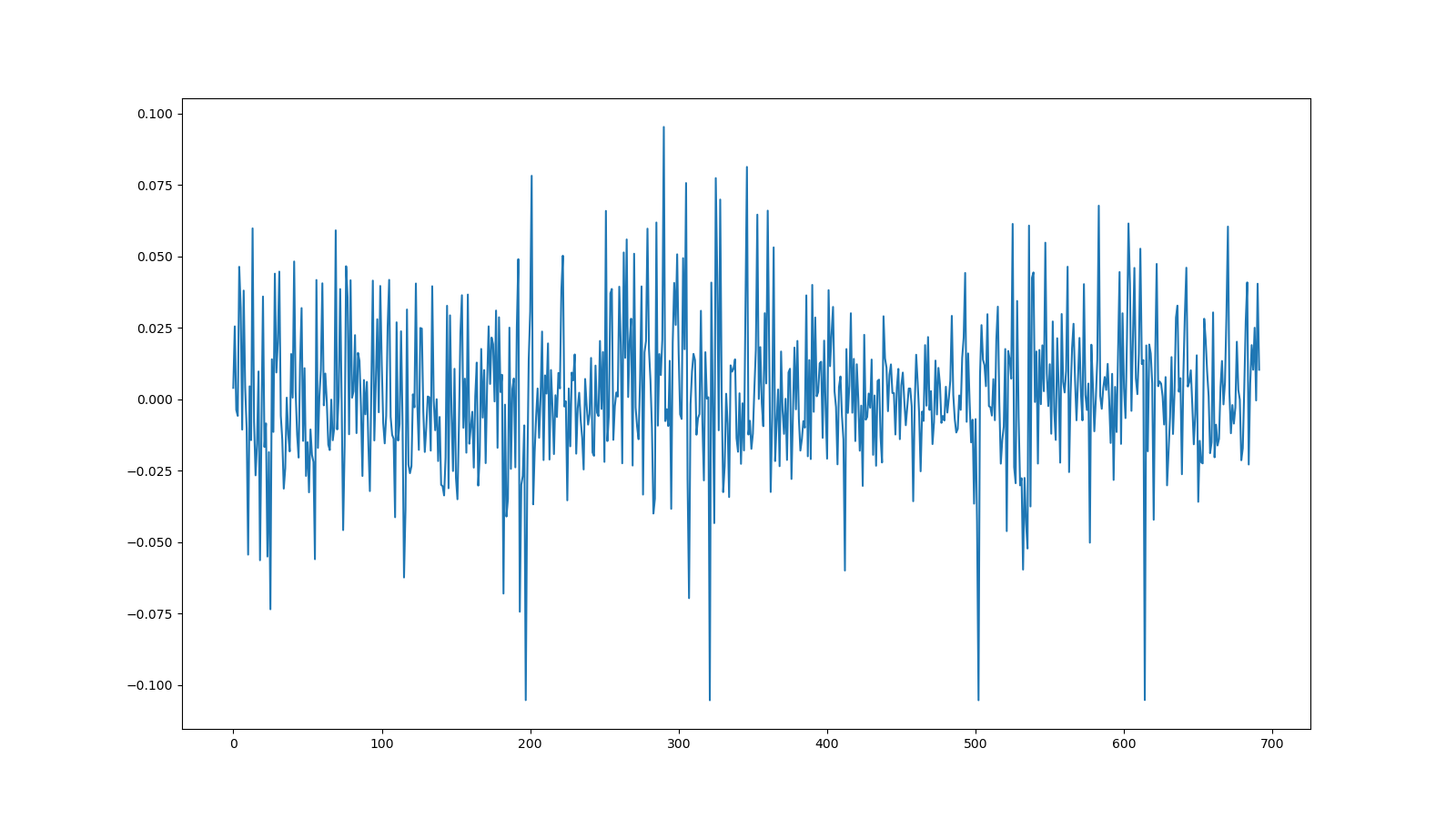
1 贵州茅台 工商银行 五粮液

1.1简单收益率 对数收益率

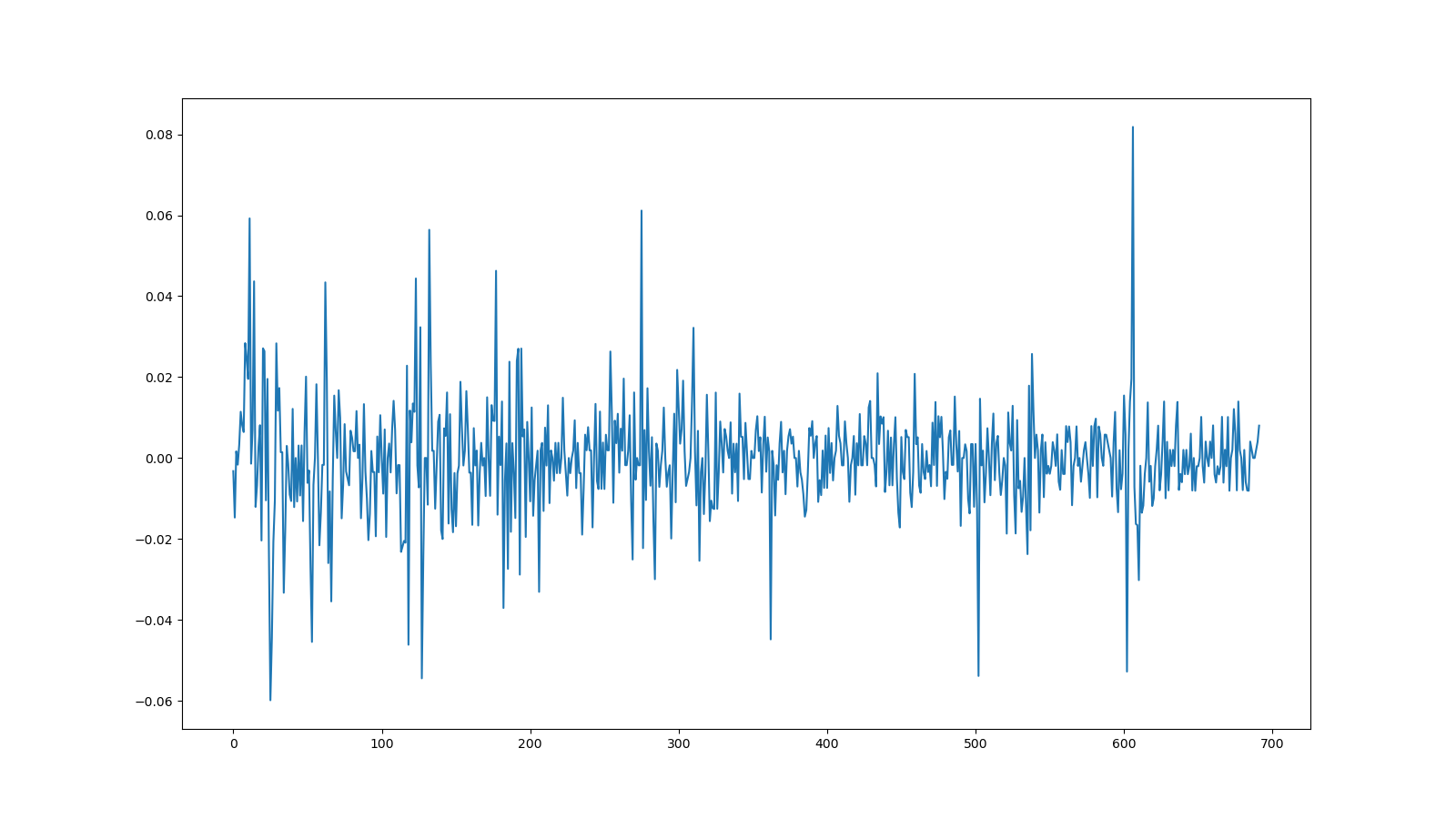


1.2对数收益率时序图

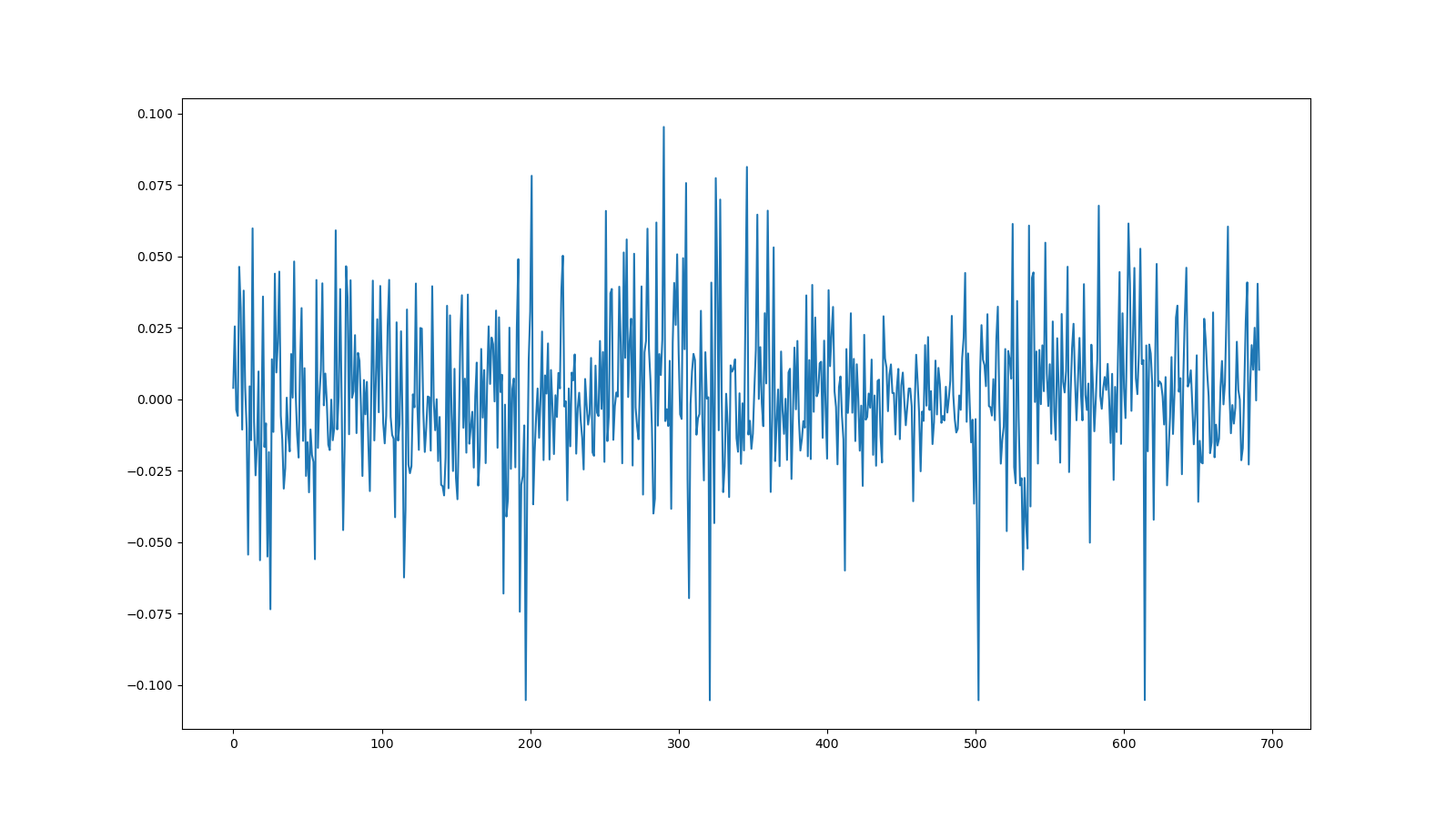




贵州茅台

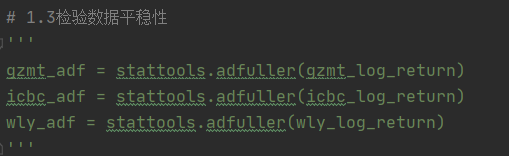


工商银行



五粮液

1.3检验数据平稳性



检验结果:贵州茅台和五粮液平稳 工商银行7阶平稳

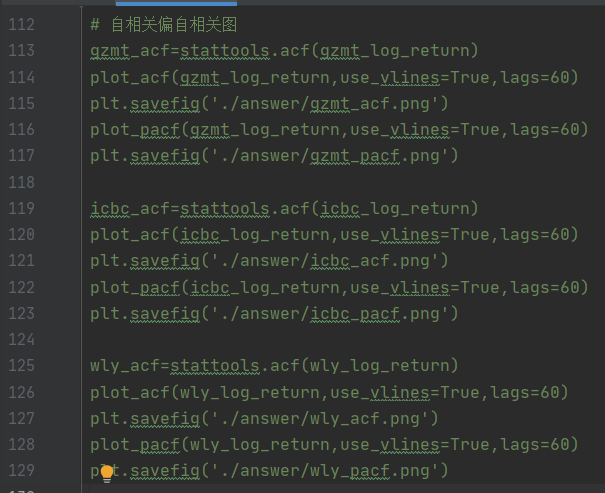
gzmt(-26.40707683339336, 0.0, 0, 691, {'1%': -3.4398489271901713, '5%': -2.8657317673059395, '10%': -2.569002221596252}, -3324.304797850458)

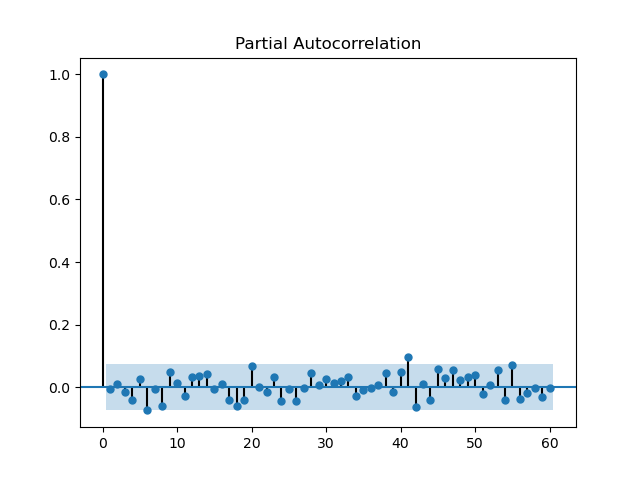
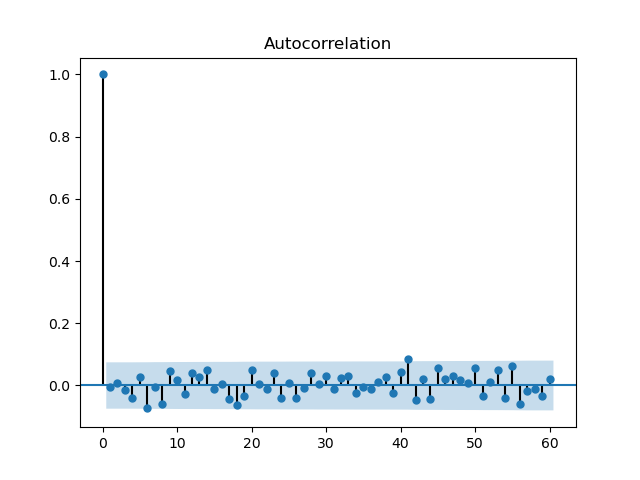
icbc(-10.667451464679653, 4.251165964717407e-19, 7, 684, {'1%': -3.439946506895534, '5%': -2.8657747597081404, '10%': -2.5690251267911495}, -3979.29930404665)

wly(-25.403506965722965, 0.0, 0, 691, {'1%': -3.4398489271901713, '5%': -2.8657317673059395, '10%': -2.569002221596252}, -3023.636593958669)

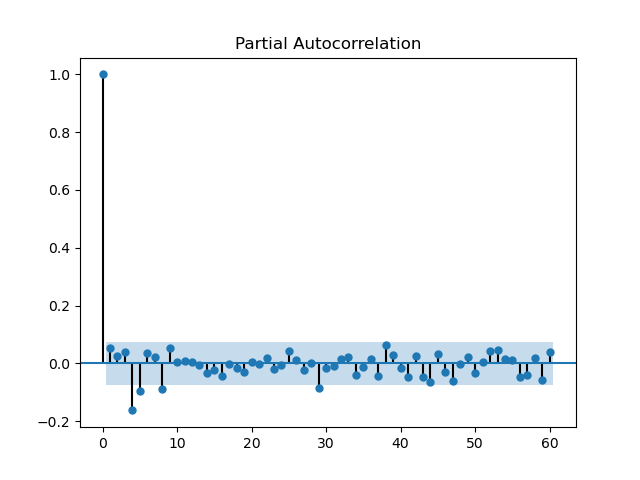
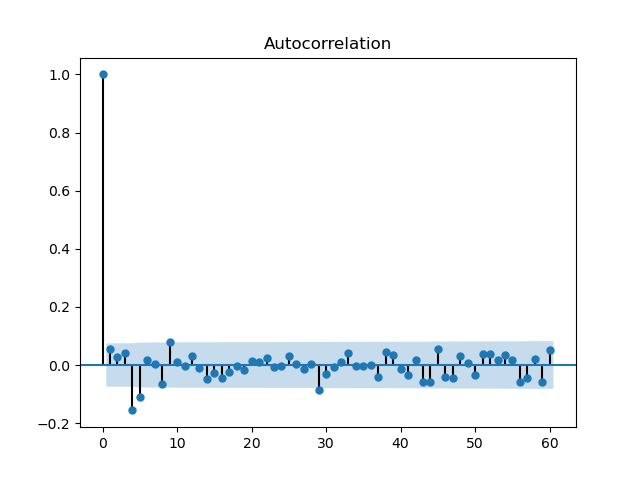
其实对于有涨跌停板限制的a股市场, 验证日收益率序列的平稳性是没有意义是 必然平稳

1.4通过信息准则建立合适的模型

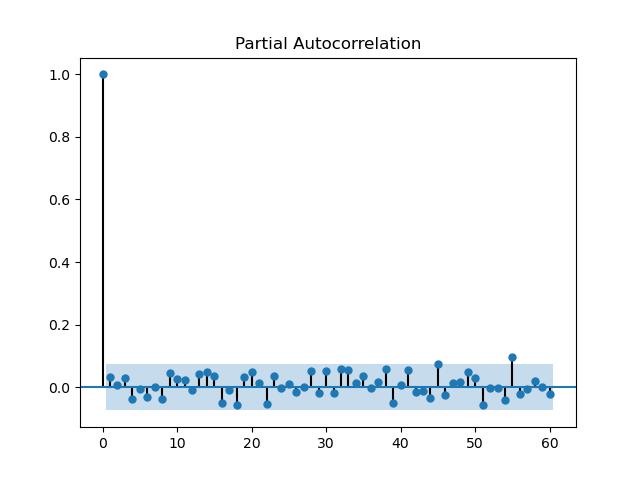
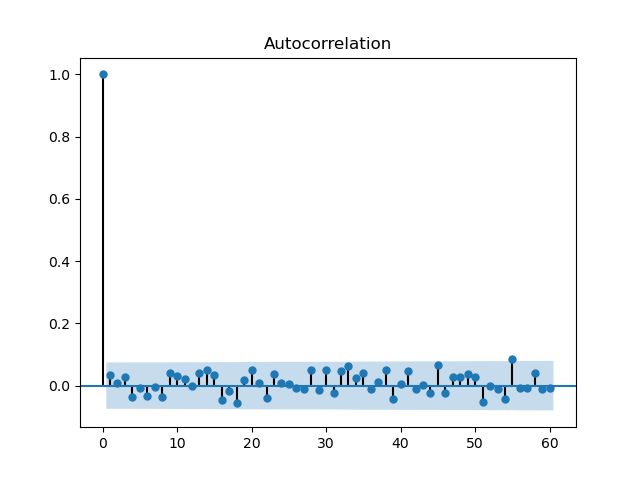




贵州茅台acf和pacf

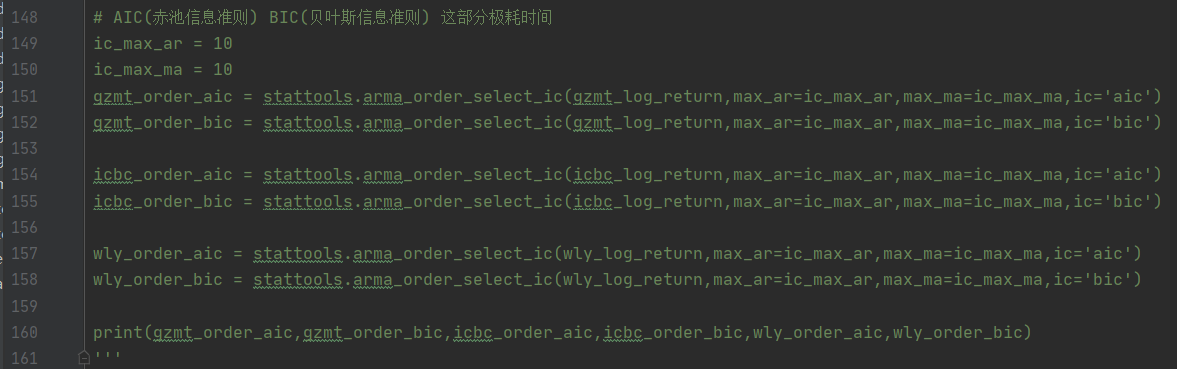


工商银行acf pacf



五粮液acf pacf

偏相关图自相关图看不出什么 继续进行信息准则定阶



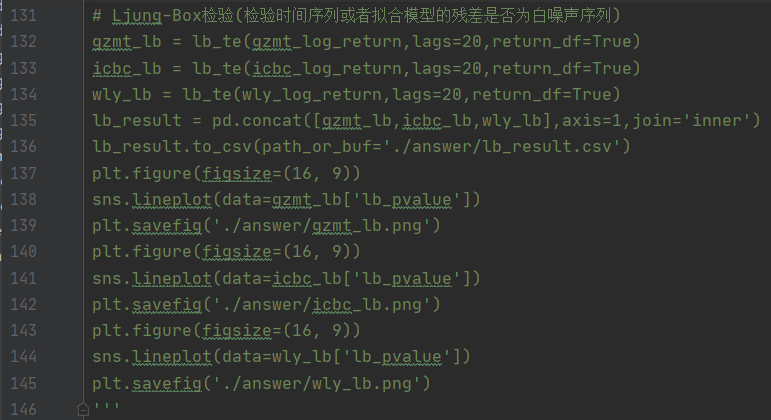
结果见文件夹下ic\_order.txt和ic\_order\_another.txt 最大定阶数为ar和ma 均10阶

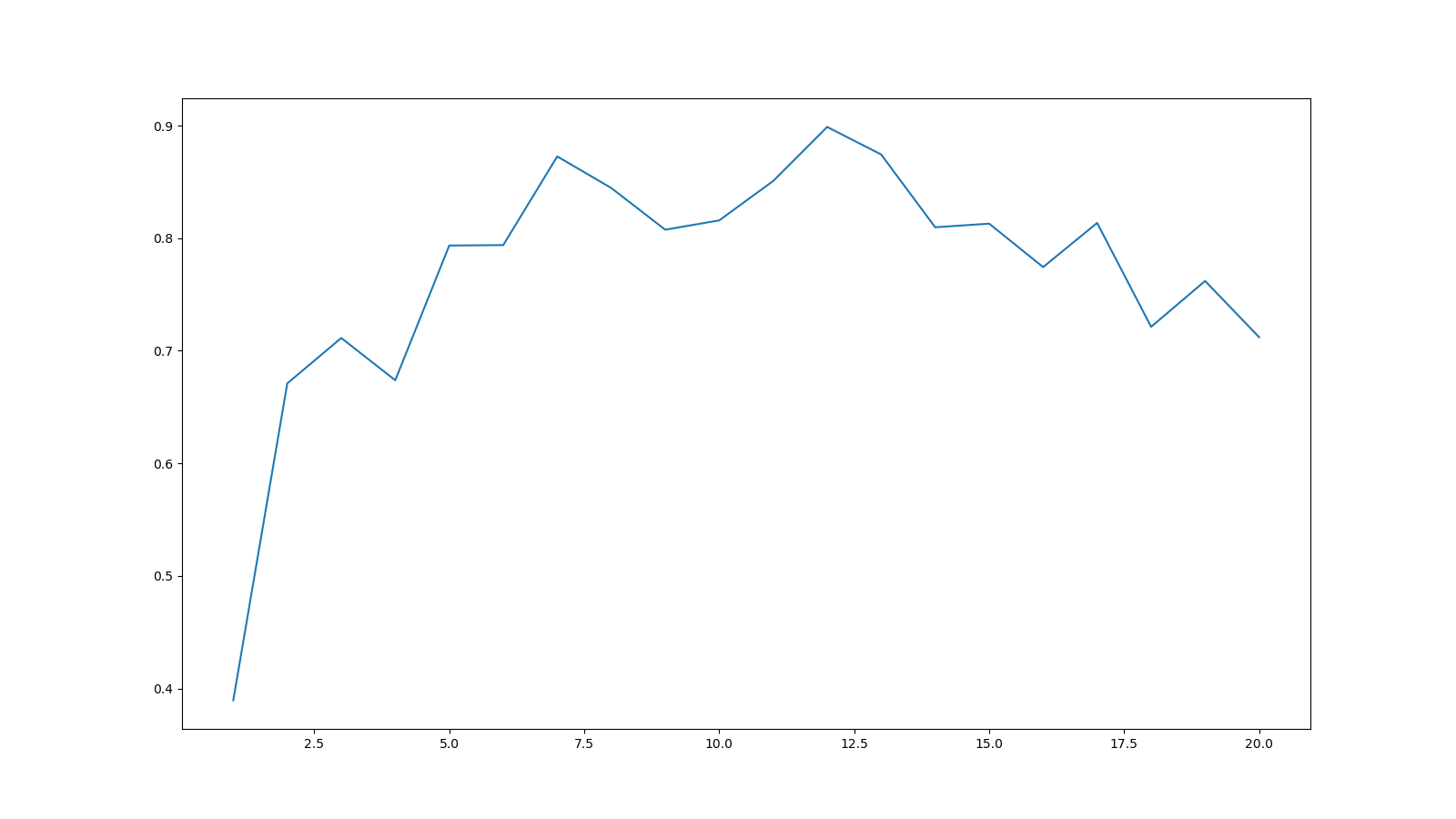
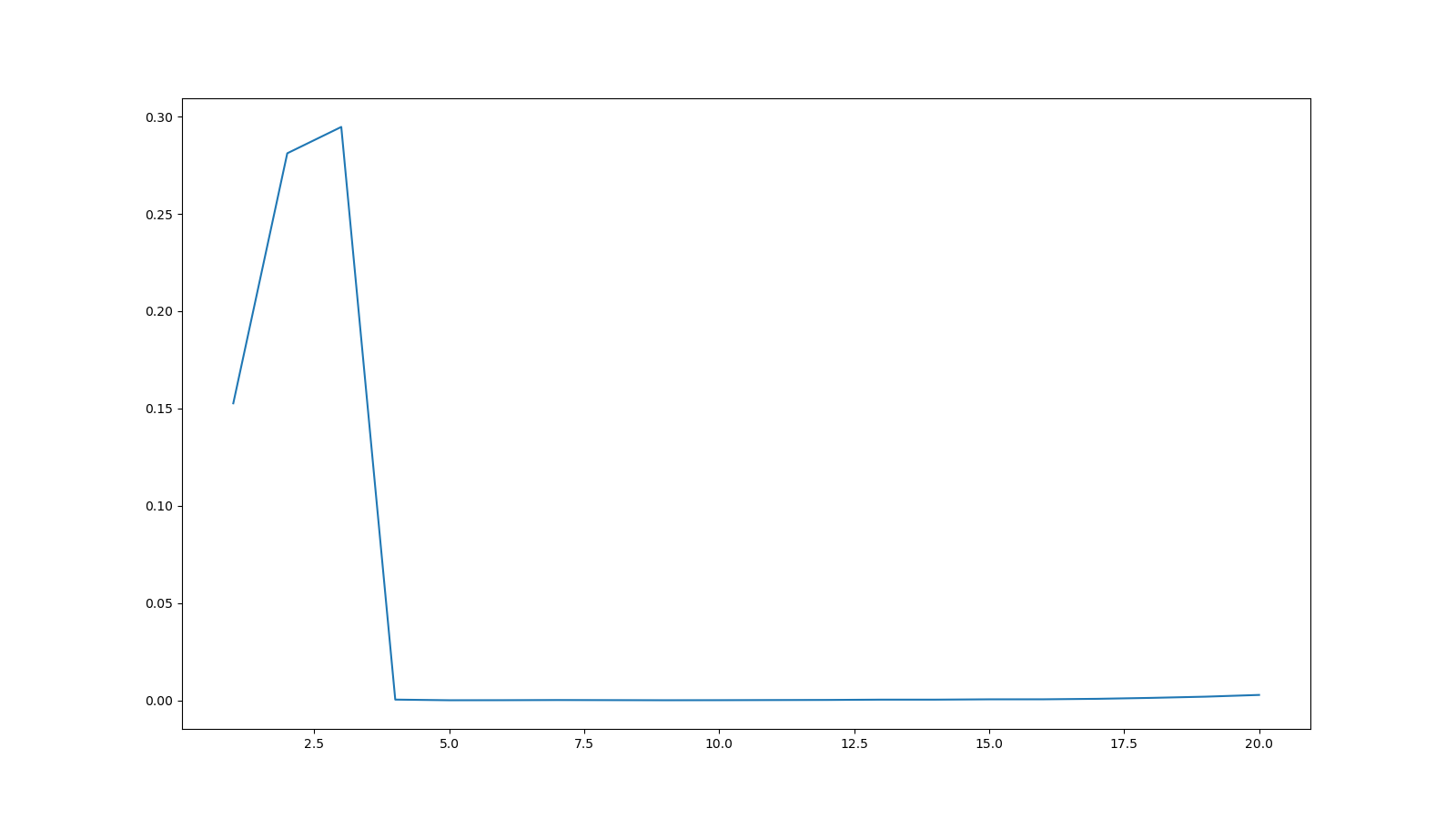
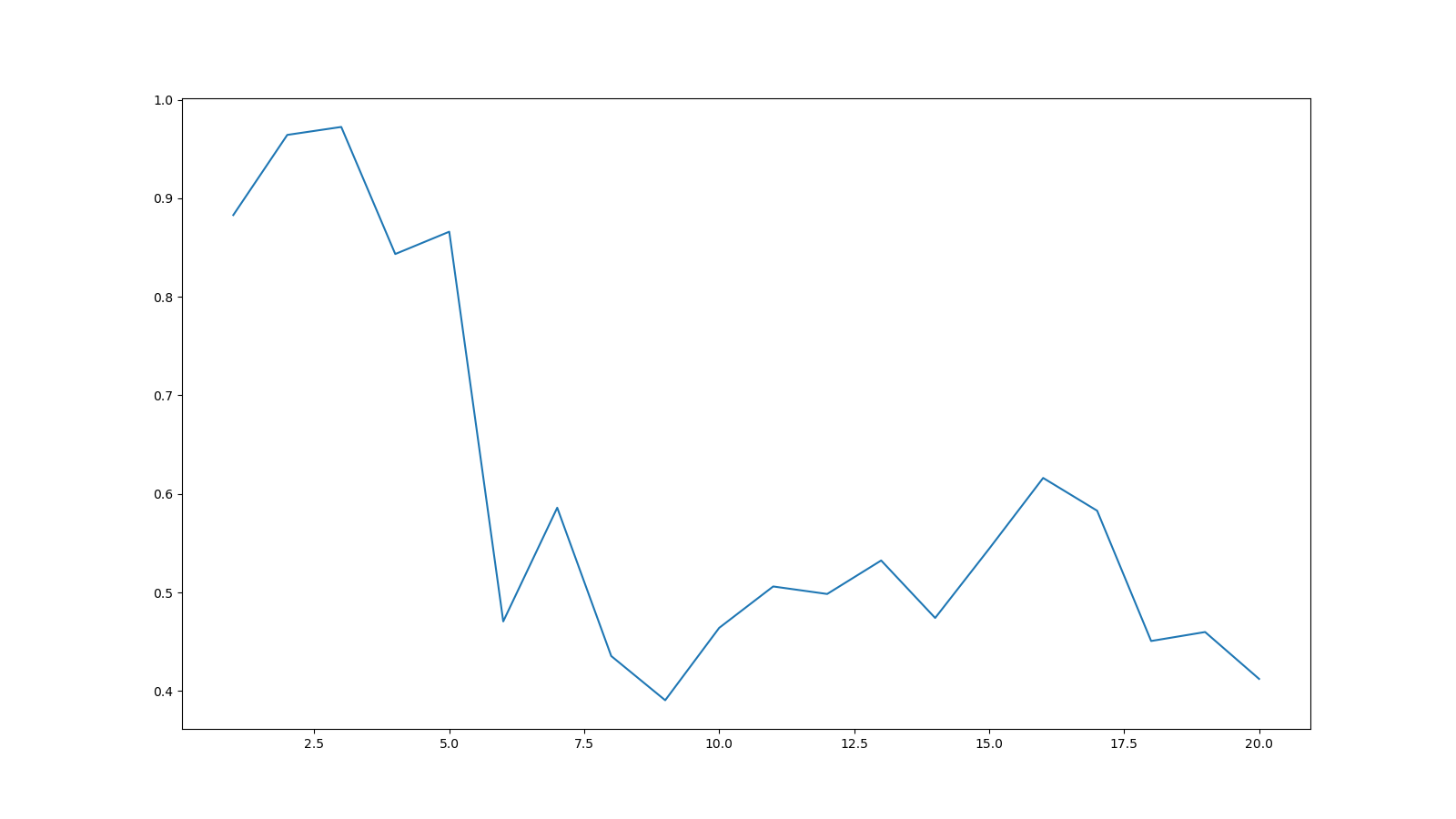
R5 3600X跑了一晚上才出结果 不要轻易运行

结论是贵州茅台在aicbic下均为0,0 工商银行在aic下6,2 bic下0,0 五粮液在aic下8,6 bic下0,0

结果很不好看 于是怀疑序列有为白噪声的可能性 白噪声就没有建模的意义了

因此直接对收益率序列进行lb检验 结果见lb\_result.csv

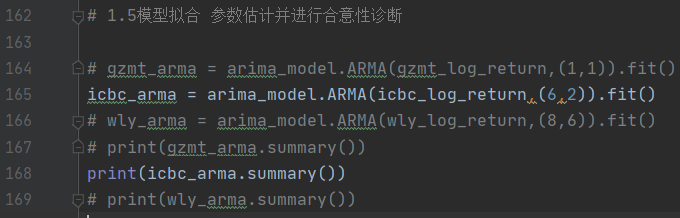




结论为贵州茅台和五粮液日收益率为白噪声序列, 没有建模的意义

因此最后只对工商银行的日收益率序列进行建模和分析ARMA(6,2)

1.5模型拟合 参数估计并进行合意性诊断



结果如下(完整结果见fit\_result\_1.txt)

Tit = total number of iterations

Tnf = total number of function evaluations

Tnint = total number of segments explored during Cauchy searches

Skip = number of BFGS updates skipped

Nact = number of active bounds at final generalized Cauchy point

Projg = norm of the final projected gradient

F = final function value

\* \* \*

N Tit Tnf Tnint Skip Nact Projg F

9 295 379 2 0 0 2.354D-05 -2.949D+00

F = -2.94918839158974

CONVERGENCE: REL\_REDUCTION\_OF\_F\_<=\_FACTR\*EPSMCH

ARMA Model Results

==============================================================================

Dep. Variable: y No. Observations: 692

Model: ARMA(6, 2) Log Likelihood 2040.838

Method: css-mle S.D. of innovations 0.013

Date: Sun, 10 Jan 2021 AIC -4061.677

Time: 14:37:41 BIC -4016.281

Sample: 0 HQIC -4044.119

==============================================================================

coef std err z P>|z| [0.025 0.975]

------------------------------------------------------------------------------

const -0.0003 0.000 -0.674 0.500 -0.001 0.001

ar.L1.y -1.1128 0.195 -5.697 0.000 -1.496 -0.730

ar.L2.y -0.5865 0.181 -3.247 0.001 -0.941 -0.233

ar.L3.y 0.1213 0.061 1.976 0.048 0.001 0.242

ar.L4.y -0.0825 0.061 -1.355 0.175 -0.202 0.037

ar.L5.y -0.2503 0.057 -4.368 0.000 -0.363 -0.138

ar.L6.y -0.1932 0.043 -4.458 0.000 -0.278 -0.108

ma.L1.y 1.1633 0.198 5.868 0.000 0.775 1.552

ma.L2.y 0.6830 0.186 3.675 0.000 0.319 1.047

Roots

=============================================================================

Real Imaginary Modulus Frequency

-----------------------------------------------------------------------------

AR.1 1.0999 -0.9268j 1.4383 -0.1114

AR.2 1.0999 +0.9268j 1.4383 0.1114

AR.3 -0.6622 -1.0150j 1.2119 -0.3420

AR.4 -0.6622 +1.0150j 1.2119 0.3420

AR.5 -1.0854 -0.7248j 1.3051 -0.4063

AR.6 -1.0854 +0.7248j 1.3051 0.4063

MA.1 -0.8516 -0.8596j 1.2100 -0.3743

MA.2 -0.8516 +0.8596j 1.2100 0.3743

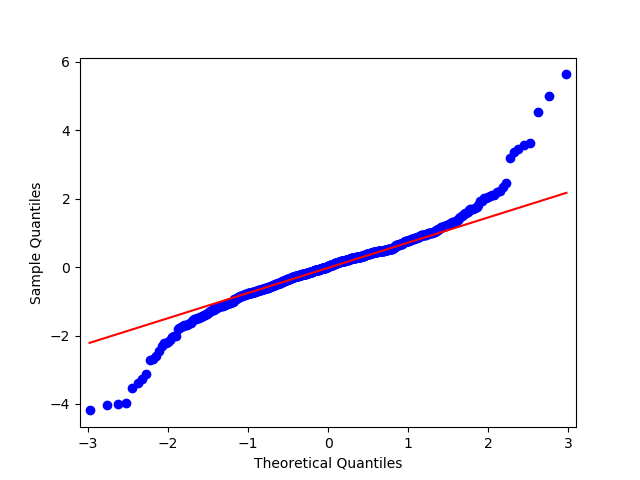
-----------------------------------------------------------------------------

得到

合意性检验 使用qq图和DW检验

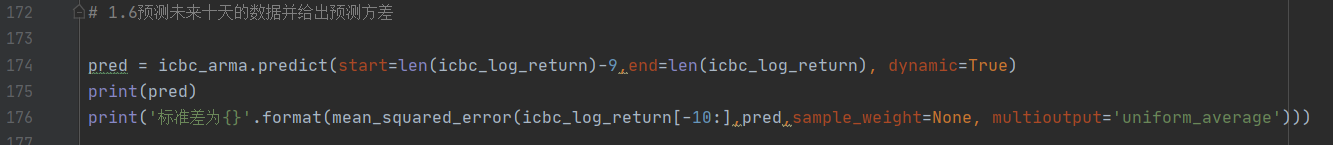
DW检验值为1.9966411564506925 说明残差不存在自相关性

qq图说明残差基本满足正态分布



1.6预测未来十天的数据并给出预测方差

由于手头没有未来数据而且之前程序没考虑分训练集和测试集所以做了样本内估计, 等有时间再进行改进



10天预期为[-0.00101387 0.00187632 -0.00132168 0.00085113 0.00032685 -0.00069343 -0.00040652 -0.00012824 -0.00061258 -0.00040647]

预测标准差为2.7934151713424298e-05

2 沪深300 上证50 中证500

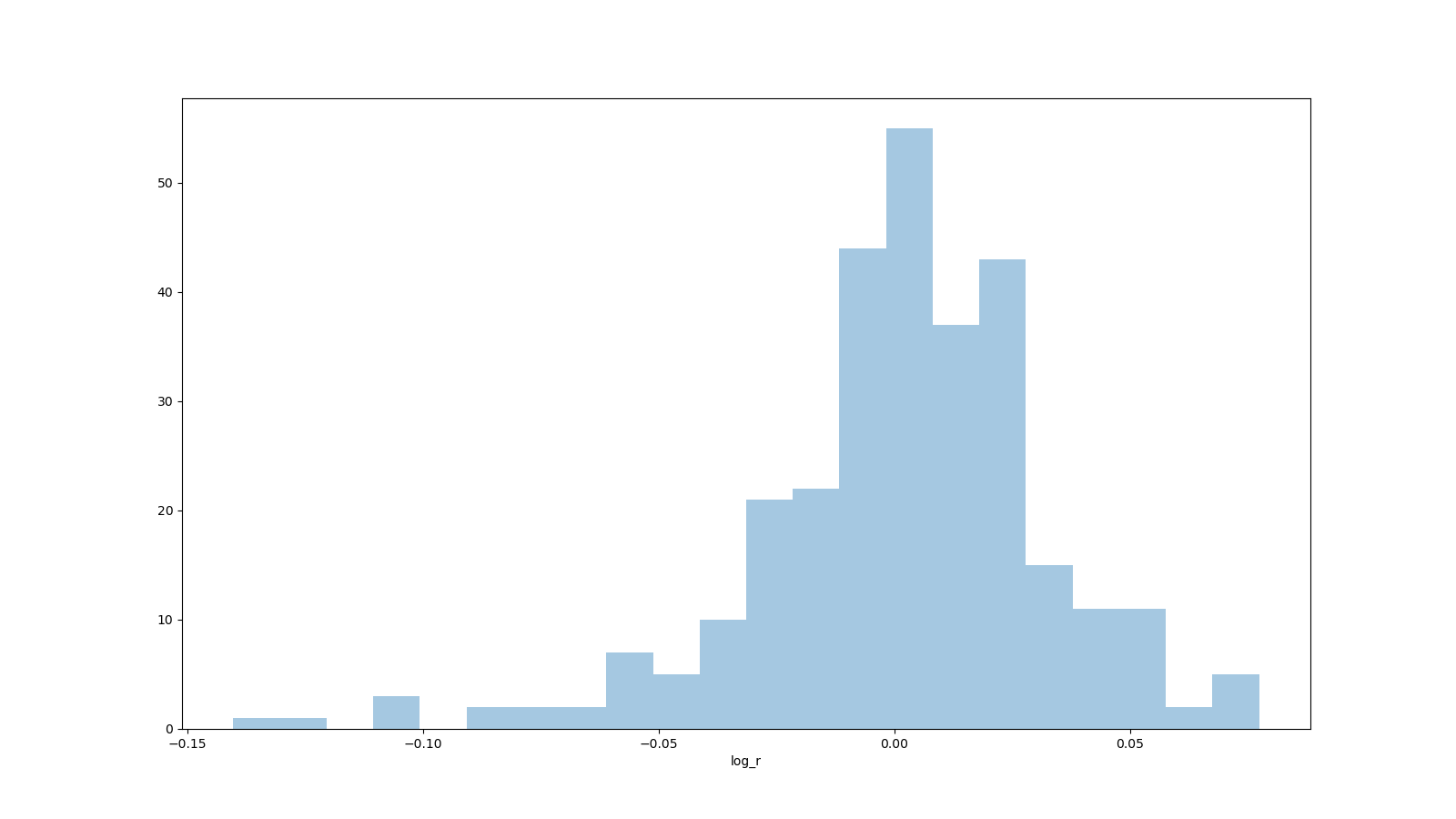
2.1 均值标准差偏度超额峰度 结果见answer\_2\_1.csv



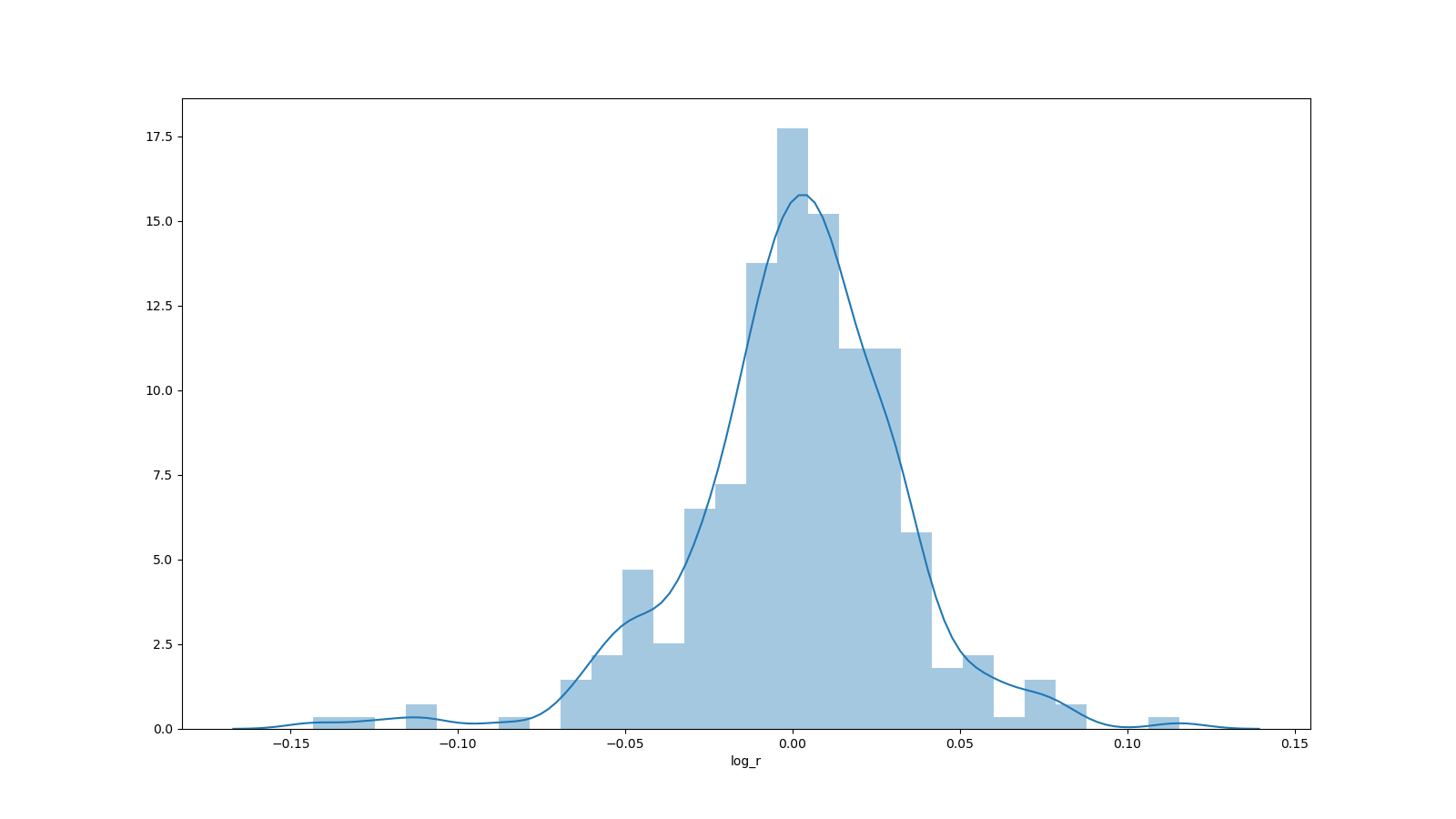
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | mean | std | skew | kurtosis | excess\_k |
| hs300 | 0.001118 | 0.031686 | -0.94658 | 2.694339 | -0.30566 |
| sz50 | 0.000953 | 0.031966 | -0.61627 | 2.78246 | -0.21754 |
| csi500 | 0.000549 | 0.040115 | -0.9821 | 3.711614 | 0.711614 |

2.2比较三者的经验分布 判断是否属于正态分布

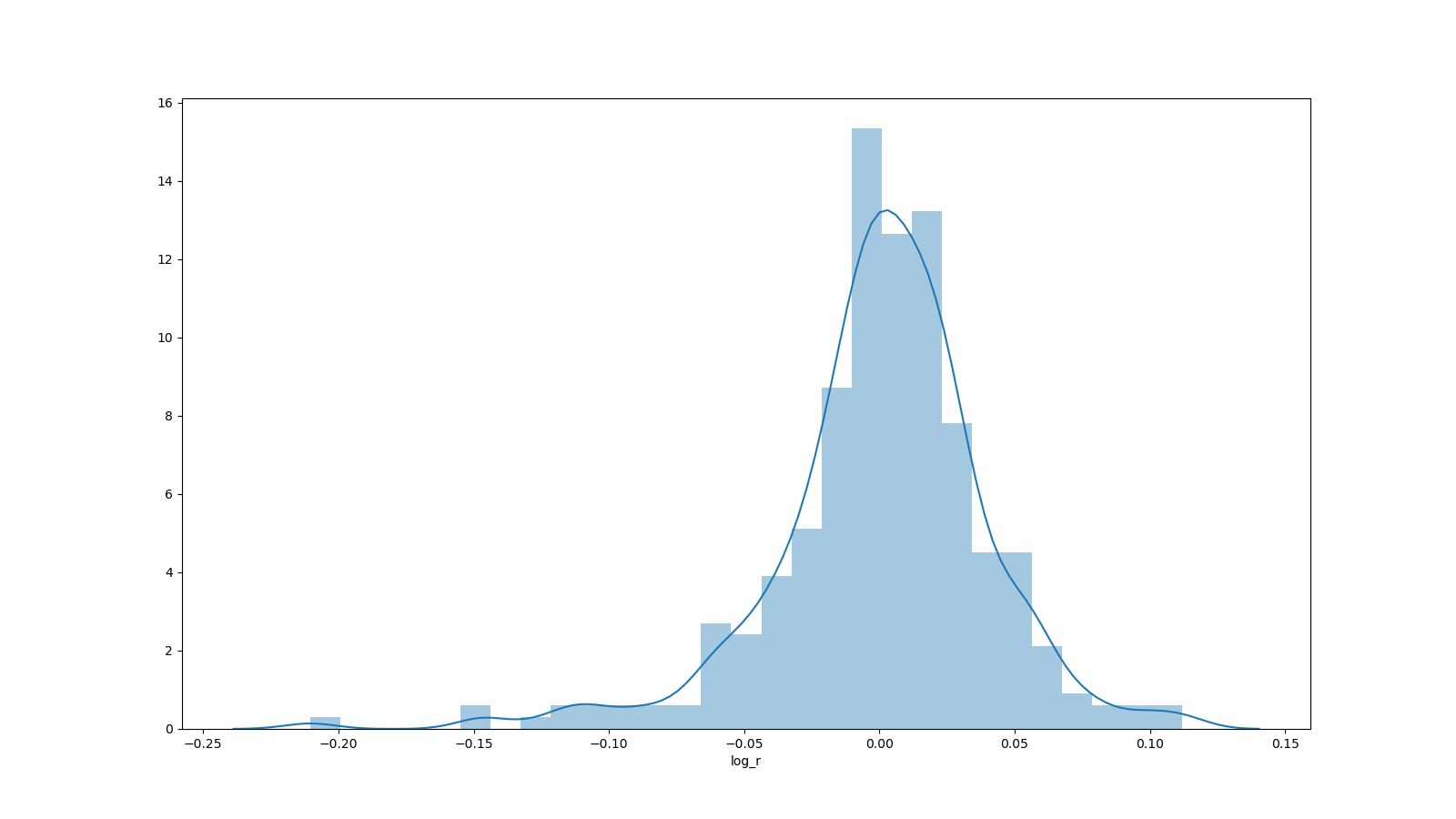




HS300



SZ50



CSI500

使用Anderson-Darling检验其是否服从正态分布 结果见2.2\_AD.txt

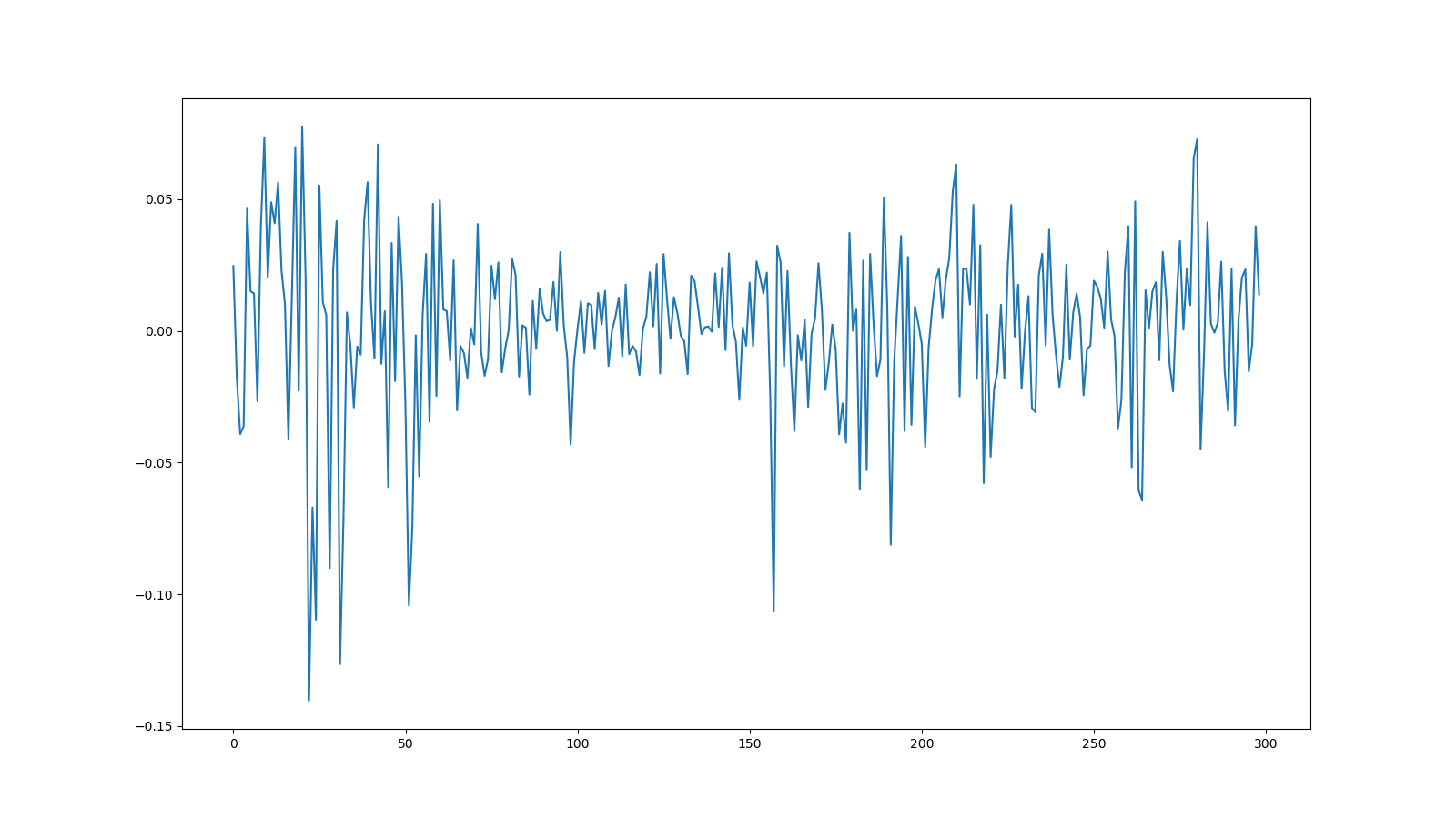
AndersonResult(statistic=3.4427821502071083, critical\_values=array([0.569, 0.648, 0.777, 0.906, 1.078]), significance\_level=array([15. , 10. , 5. , 2.5, 1. ]))

AndersonResult(statistic=2.770997695783592, critical\_values=array([0.569, 0.648, 0.777, 0.906, 1.078]), significance\_level=array([15. , 10. , 5. , 2.5, 1. ]))

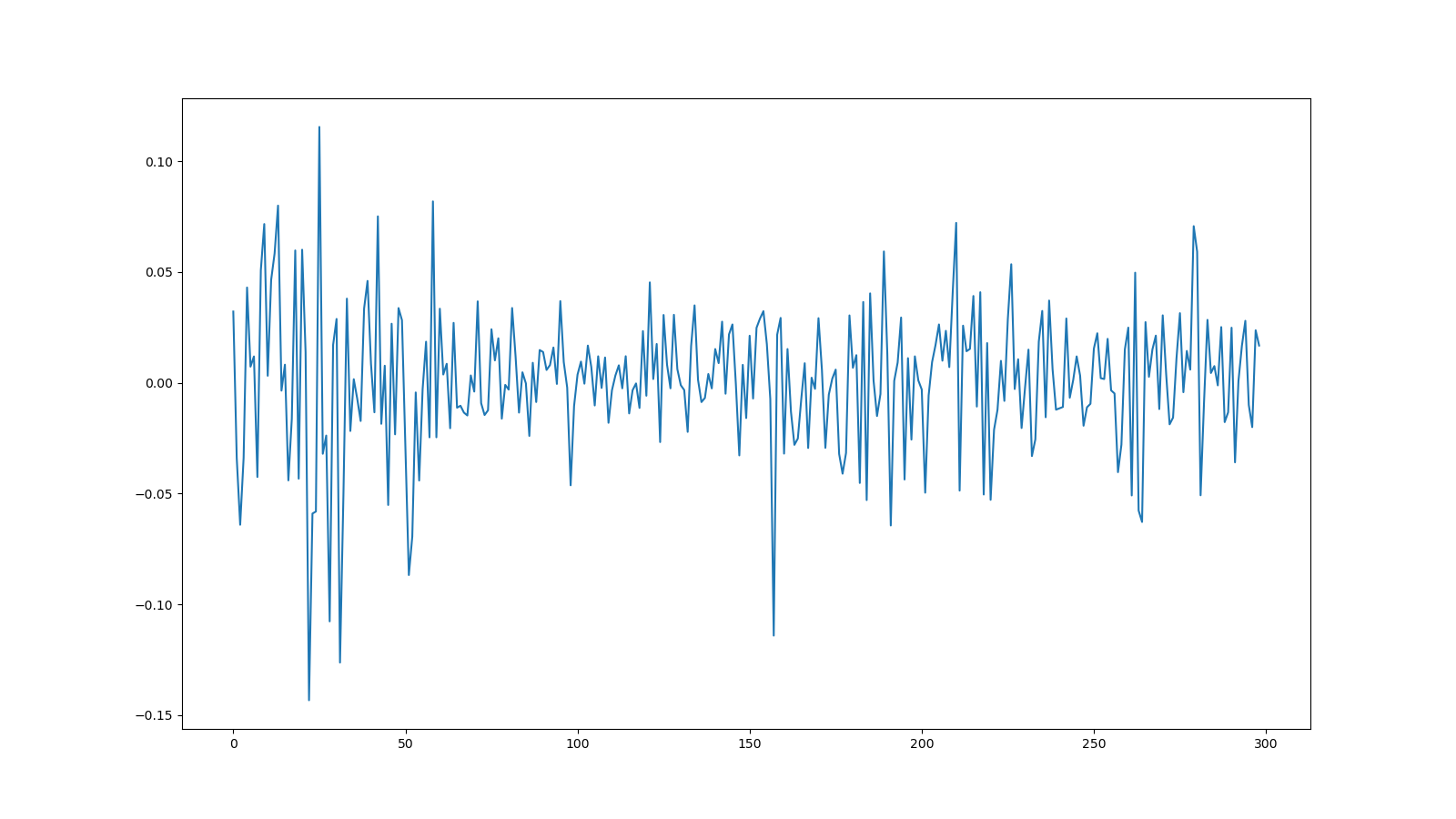
AndersonResult(statistic=4.2670460848644325, critical\_values=array([0.569, 0.648, 0.777, 0.906, 1.078]), significance\_level=array([15. , 10. , 5. , 2.5, 1. ]))

全部拒绝原假设 而原假设为H0为样本服从特定分布 因此这三者都不服从正态分布

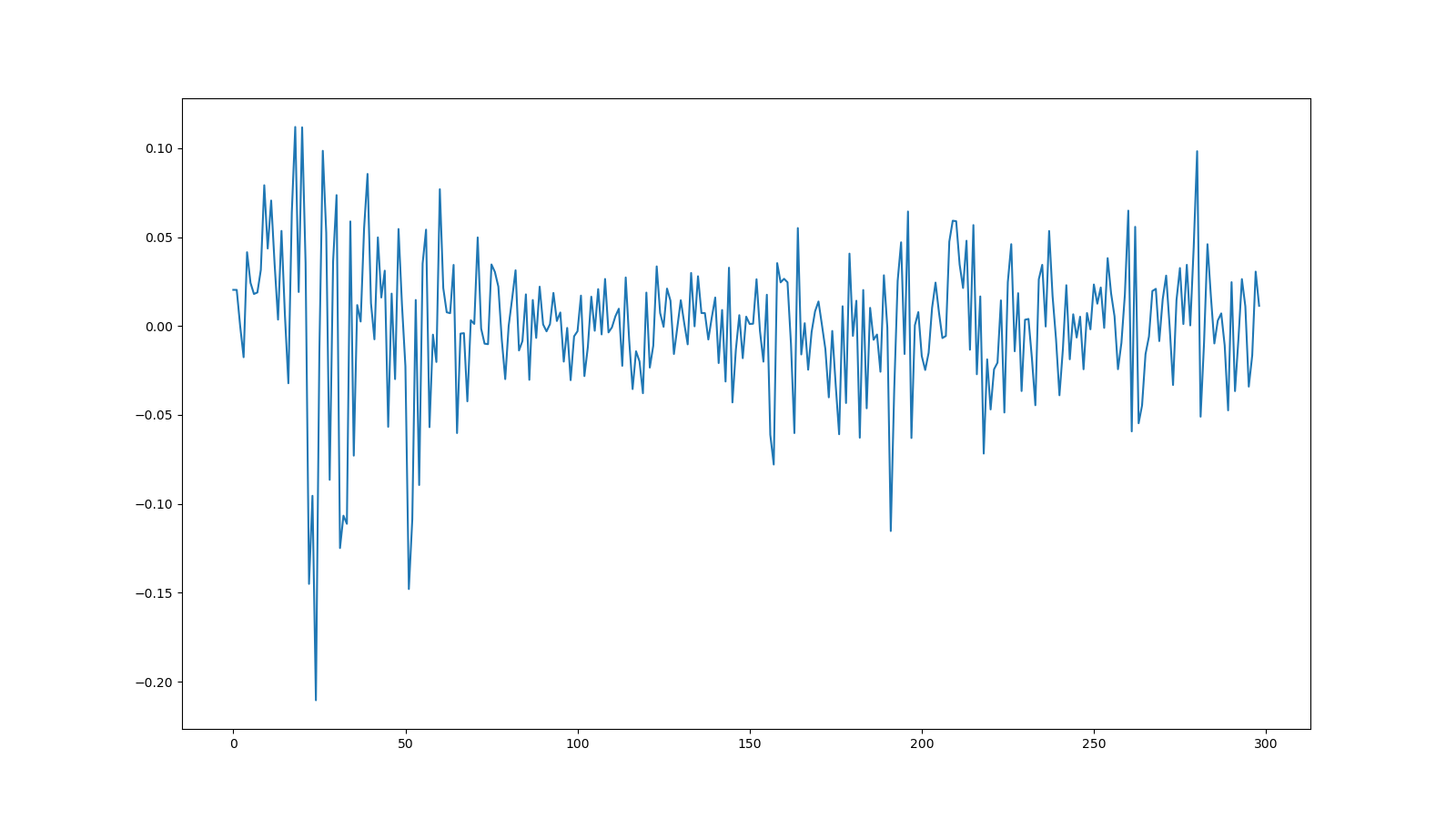
2.3时序图 判断领先滞后关系



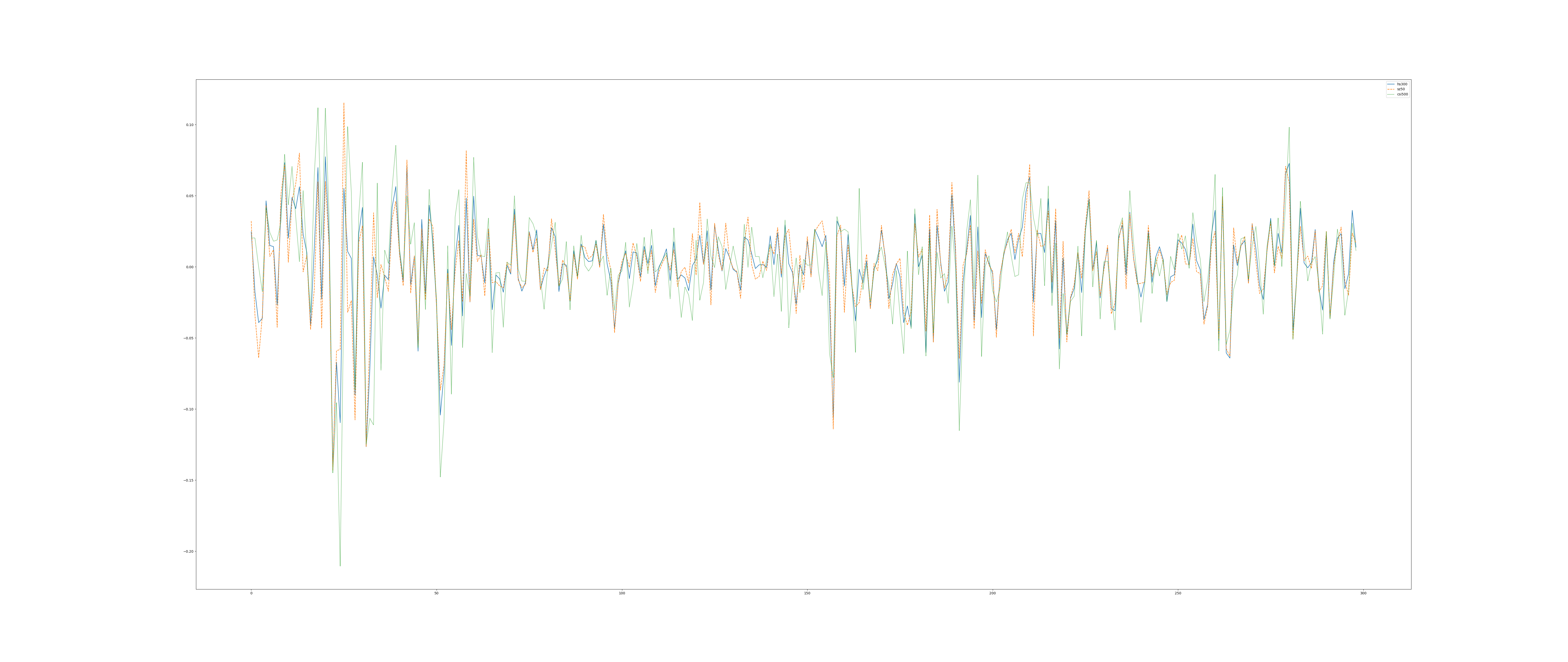
HS300



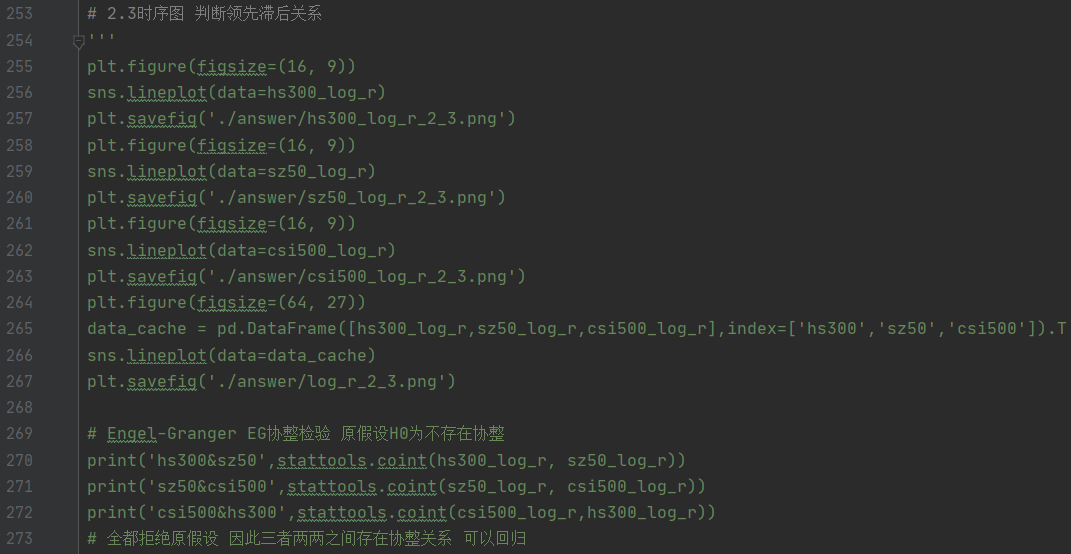
SZ50



CSI500



总结果

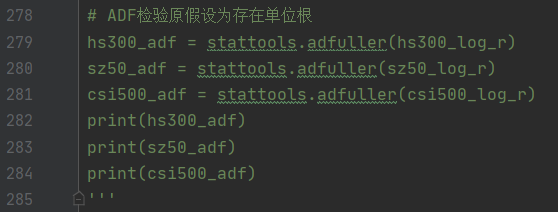


hs300&sz50 (-9.62155583131033, 2.2564790612687137e-15, array([-3.93356888, -3.35671052, -3.05871284]))

sz50&csi500 (-18.428270316413794, 1.456472354087983e-29, array([-3.93356888, -3.35671052, -3.05871284]))

csi500&hs300 (-9.468393218414628, 5.5392770044561065e-15, array([-3.93356888, -3.35671052, -3.05871284]))

2.4平稳性检验 构建指数间的预测模型



Hs300(-16.2670275894668, 3.518751730002799e-29, 0, 298, {'1%': -3.4524859843440754, '5%': -2.871288184343229, '10%': -2.571964047565425}, -1149.3061748720452)

Sz50(-17.568578603333485, 4.062400159735734e-30, 0, 298, {'1%': -3.4524859843440754, '5%': -2.871288184343229, '10%': -2.571964047565425}, -1154.0819025405708)

Csi500(-14.92711255221588, 1.3837933563750384e-27, 0, 298, {'1%': -3.4524859843440754, '5%': -2.871288184343229, '10%': -2.571964047565425}, -1011.1156385501249)

2.5对3 4的结论进行经济学或投资逻辑上的解释

沪深300是整个市场的晴雨表

上证50是优质大盘股的代表

中证500是小盘股的代表

4 2\_13

假设平稳 先定阶

{'aic': 0 1 2 3 4 5

0 2868.237901 2844.730382 2846.713061 2848.277614 2850.042404 2852.040856

1 2846.220700 2846.716850 2848.650277 2846.257077 2852.040878 2854.040853

2 2846.352145 2848.124647 2850.094653 2843.362680 2850.215631 2850.050473

3 2848.117265 2846.141261 2848.560261 2847.878828 2845.705981 2847.614745

4 2850.115138 2848.088756 2842.982343 2851.020584 2853.805239 2850.735419

5 2852.052898 2850.048281 2847.485261 2845.816687 2847.200677 2843.212128, 'aic\_min\_order': (4, 2)} {'bic': 0 1 2 3 4 5

0 2876.482886 2857.097861 2863.203033 2868.890078 2874.777361 2880.898305

1 2858.588179 2863.206821 2869.262741 2870.992034 2880.898328 2887.020796

2 2862.842116 2868.737111 2874.829610 2872.220129 2883.195573 2887.152908

3 2868.729729 2870.876218 2877.417711 2880.858770 2882.808416 2888.839673

4 2874.850095 2876.946206 2875.962285 2888.123019 2895.030167 2896.082840

5 2880.910348 2883.028224 2884.587697 2887.041615 2892.548098 2892.682042, 'bic\_min\_order': (0, 1)}

Process finished with exit code 0

aic定了4,2 bic定了0,1 按aic的来 即AR(4) MA(2)

拟合结果和预测如下

\* \* \*

Tit = total number of iterations

Tnf = total number of function evaluations

Tnint = total number of segments explored during Cauchy searches

Skip = number of BFGS updates skipped

Nact = number of active bounds at final generalized Cauchy point

Projg = norm of the final projected gradient

F = final function value

\* \* \*

N Tit Tnf Tnint Skip Nact Projg F

7 36 61 1 0 0 4.716D-05 3.100D+00

F = 3.09976134102346

CONVERGENCE: REL\_REDUCTION\_OF\_F\_<=\_FACTR\*EPSMCH

ARMA Model Results

==============================================================================

Dep. Variable: y No. Observations: 456

Model: ARMA(4, 0) Log Likelihood -1419.058

Method: css-mle S.D. of innovations 5.436

Date: Sun, 10 Jan 2021 AIC 2850.115

Time: 19:50:19 BIC 2874.850

Sample: 0 HQIC 2859.859

==============================================================================

coef std err z P>|z| [0.025 0.975]

------------------------------------------------------------------------------

const 1.0601 0.302 3.513 0.000 0.469 1.651

ar.L1.y 0.2395 0.047 5.114 0.000 0.148 0.331

ar.L2.y -0.0587 0.048 -1.218 0.223 -0.153 0.036

ar.L3.y -0.0222 0.048 -0.461 0.645 -0.117 0.072

ar.L4.y -0.0022 0.047 -0.046 0.963 -0.094 0.090

Roots

=============================================================================

Real Imaginary Modulus Frequency

-----------------------------------------------------------------------------

AR.1 1.6297 -2.1386j 2.6888 -0.1464

AR.2 1.6297 +2.1386j 2.6888 0.1464

AR.3 -6.7584 -4.2760j 7.9975 -0.4102

AR.4 -6.7584 +4.2760j 7.9975 0.4102

-----------------------------------------------------------------------------

ARMA Model Results

==============================================================================

Dep. Variable: y No. Observations: 456

Model: ARMA(0, 2) Log Likelihood -1419.357

Method: css-mle S.D. of innovations 5.439

Date: Sun, 10 Jan 2021 AIC 2846.713

Time: 19:50:19 BIC 2863.203

Sample: 0 HQIC 2853.209

==============================================================================

coef std err z P>|z| [0.025 0.975]

------------------------------------------------------------------------------

const 1.0612 0.317 3.344 0.001 0.439 1.683

ma.L1.y 0.2403 0.047 5.084 0.000 0.148 0.333

ma.L2.y 0.0064 0.048 0.132 0.895 -0.089 0.101

Roots

=============================================================================

Real Imaginary Modulus Frequency

-----------------------------------------------------------------------------

MA.1 -4.7634 +0.0000j 4.7634 0.5000

MA.2 -32.9510 +0.0000j 32.9510 0.5000

-----------------------------------------------------------------------------

D:\ProgramData\Anaconda3\lib\site-packages\statsmodels\tsa\arima\_model.py:1490: RuntimeWarning: invalid value encountered in sqrt

return np.sqrt(np.diag(-inv(hess)))

D:\ProgramData\Anaconda3\lib\site-packages\scipy\stats\\_distn\_infrastructure.py:903: RuntimeWarning: invalid value encountered in greater

return (a < x) & (x < b)

D:\ProgramData\Anaconda3\lib\site-packages\scipy\stats\\_distn\_infrastructure.py:903: RuntimeWarning: invalid value encountered in less

return (a < x) & (x < b)

D:\ProgramData\Anaconda3\lib\site-packages\scipy\stats\\_distn\_infrastructure.py:1912: RuntimeWarning: invalid value encountered in less\_equal

cond2 = cond0 & (x <= \_a)

ARMA Model Results

==============================================================================

Dep. Variable: y No. Observations: 456

Model: ARMA(4, 2) Log Likelihood -1413.491

Method: css-mle S.D. of innovations 5.347

Date: Sun, 10 Jan 2021 AIC 2842.982

Time: 19:50:19 BIC 2875.962

Sample: 0 HQIC 2855.974

==============================================================================

coef std err z P>|z| [0.025 0.975]

------------------------------------------------------------------------------

const 1.0828 0.073 14.763 0.000 0.939 1.227

ar.L1.y 0.2469 nan nan nan nan nan

ar.L2.y 0.9084 3.74e-06 2.43e+05 0.000 0.908 0.908

ar.L3.y -0.2675 0.000 -2301.356 0.000 -0.268 -0.267

ar.L4.y 0.0709 nan nan nan nan nan

ma.L1.y 8.014e-05 0.004 0.021 0.983 -0.007 0.008

ma.L2.y -0.9990 0.004 -264.581 0.000 -1.006 -0.992

Roots

=============================================================================

Real Imaginary Modulus Frequency

-----------------------------------------------------------------------------

AR.1 -1.0000 -0.0000j 1.0000 -0.5000

AR.2 1.0265 -0.0000j 1.0265 -0.0000

AR.3 1.8721 -3.1981j 3.7057 -0.1657

AR.4 1.8721 +3.1981j 3.7057 0.1657

MA.1 -1.0005 +0.0000j 1.0005 0.5000

MA.2 1.0005 +0.0000j 1.0005 0.0000

-----------------------------------------------------------------------------

[2.28826743 0.7950713 0.75794064]

[2.30570054 1.09296014 1.06124179]

Process finished with exit code 0