1. Netflix stock.

<table>
<thead>
<tr>
<th>Type</th>
<th>Mean</th>
<th>Median</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma_{0,t}^2$</td>
<td>37.94</td>
<td>15.80</td>
<td>0</td>
<td>872.94</td>
</tr>
<tr>
<td>$\sigma_{1,t}^2$</td>
<td>50.67</td>
<td>23.34</td>
<td>0.13</td>
<td>1473.6</td>
</tr>
<tr>
<td>$\sigma_{2,t}^2$</td>
<td>38.84</td>
<td>22.21</td>
<td>1.62</td>
<td>645.07</td>
</tr>
<tr>
<td>$\sigma_{3,t}^2$</td>
<td>69.52</td>
<td>39.94</td>
<td>2.84</td>
<td>1157.0</td>
</tr>
<tr>
<td>$\sigma_{4,t}^2$</td>
<td>39.05</td>
<td>22.27</td>
<td>1.50</td>
<td>401.83</td>
</tr>
<tr>
<td>$\sigma_{5,t}^2$</td>
<td>71.48</td>
<td>40.37</td>
<td>2.84</td>
<td>754.58</td>
</tr>
</tbody>
</table>

2. The time plot of volatility is in Figure 1. Let $x_t$ be the log volatility, then the model is

$$(1 - B)x_t = a_t - 0.064a_{t-1},$$

where the MA coefficient is significant. Model checking indicates the presence of possible outliers. The 1-step to 5-step predictions are 0.4245, 0.4247, 0.4249, 0.4251, and 0.4254.


   (a) The fitted logistic regression model is

   $$\text{logit}(p_t) = 0.221 + 0.278M_{t-1} - 0.093S_{t-1} - 0.022M_{t-2} - 0.081S_{t-2},$$

   where $p_t = P(M_t = 1)$. The $p$-value of the coefficient estimates are all greater than 0.05. Thus, the model is not informative. In fact, only $M_{t-1}$ contributes marginally in the model.

   (b) The fitted network model is

   $$h(o_t) = \begin{cases} 
   1 & \text{if } o_t > 0 \\
   0 & \text{if } o_t \leq 0 
   \end{cases}$$

   $$o_t = 2.83 - 2.38h_{1t} - 4.37h_{2t} + 3.46h_{3t} - 0.49M_{t-1} + 0.68S_{t-1} - 1.04M_{t-2} + 1.31S_{t-2},$$
where

\[
\begin{align*}
  h_{1t} &= \frac{\exp(4.03 + 4.63M_{t-1} - 5.56S_{t-1} + 7.51M_{t-2} - 5.79S_{t-2})}{1 + \exp(4.03 + 4.63M_{t-1} - 5.56S_{t-1} + 7.51M_{t-2} - 5.79S_{t-2})} \\
  h_{2t} &= \frac{\exp(0.53 + 0.03M_{t-1} + 0.06S_{t-1} + 7.44M_{t-2} + 4.76S_{t-2})}{1 + \exp(0.53 + 0.03M_{t-1} + 0.06S_{t-1} + 7.44M_{t-2} + 4.76S_{t-2})} \\
  h_{3t} &= \frac{\exp(1.32 + 4.19M_{t-1} - 4.23S_{t-1} + 2.02M_{t-2} - 3.24S_{t-2})}{1 + \exp(1.32 + 4.19M_{t-1} - 4.23S_{t-1} + 2.02M_{t-2} - 3.24S_{t-2})},
\end{align*}
\]

(c) In term of out-of-sample prediction, the sum of square of forecast errors for the logistic regression is 33 and that of the network is 36. The two approaches provide similar result.

4. The tick-by-tick trade data of Exxon-Mobil stock from September 1 to September 30, 2013.

- The 5-m log returns are shown in Figure 2.
- No, there are no serial correlations in the 5-m log returns as we have \( Q(10) = 12.62 \) with \( p \)-value 0.25.
- The RV is shown in the upper plot of Figure 3.
- The RV is shown in the lower plot of Figure 3.

5. The tick-by-tick trade data of Exxon-Mobil stock in September 2013.

- See Figure 4.
- Yes, there is a clear diurnal pattern in the trading intensity. See Figure 5.
Figure 1: Time plot of volatility via Yang-Zhang method

Figure 2: Time plot of intraday 5-m log returns of XOM stock in September 2013.
Figure 3: Time plots of realized volatility for XOM stock in September 2013.

Figure 4: Time plot of numbers of trade in 5 minute interval for XOM stock in September 2013.
Figure 5: ACF for the numbers of trade in 5 minute interval of XOM stock in September 2013.