**Time Series Analysis: Assignment 6**

Bowen Ma (12960780), Mianyun He (13605275), Shiyi Yang (13627295)

## *Part 1*

### a.

We know that $y_{t+1}$ has the actual conditional density $p_t\left(y_{t+1}\right)$ ; ($y_{t+1}$ is a random variable)

$$\mathrm{E}_{p_t}\left[\log p_t\left(y_{t+1}\right)-\log \hat{g}_t\left(y_{t+1}\right)\right]=\mathrm{E}_{p_t}\left[\log \frac{p_t\left(y_{t+1}\right)}{g_t\left(y_{t+1}\right)}\right]=\int_{-\infty}^{\infty} \underbrace{p_t\left(y_{t+1}\right)}_{\text{density}} \underbrace{\log \left(\frac{p_t\left(y_{t+1}\right)}{\hat{g}_t\left(y_{t+1}\right)}\right)}_{\text{r.v.}} d y_{t+1}$$

Hence, $\mathrm{E}_{p_t}\left[\log p_t\left(y_{t+1}\right)-\log \hat{g}_t\left(y_{t+1}\right)\right]=\operatorname{KLIC}\left(\hat{g}_t\right)$

### b.

We know that,

$$\mathrm{E}_{p_t}\left[\log \left(\frac{p_t\left(y_{t+1}\right)}{\hat{f}_t\left(y_{t+1}\right)}\right)\right]=\mathrm{E}_{p_t}\left[-\log \left(\frac{\hat{f}_t\left(y_{t+1}\right)}{p_t\left(y_{t+1}\right)}\right)\right]=-\mathrm{E}_{p_t}\left[\log \left(\frac{\hat{f}_t\left(y_{t+1}\right)}{p_t\left(y_{t+1}\right)}\right)\right]$$

Therefore

$$\mathrm{E}_{p_t}\left[\log \left(\frac{p_t\left(y_{t+1}\right)}{\hat{f}_t\left(y_{t+1}\right)}\right)\right] \geqslant 0 \Leftrightarrow \mathrm{E}_{p_t}\left[\log \left(\frac{\hat{f}_t\left(y_{t+1}\right)}{p_t\left(y_{t+1}\right)}\right)\right] \leq 0, \quad \forall p_t,\ \hat{g}_t \in \mathscr{P}$$

From

$$\mathrm{E}_{p_t}\left[\log \left(\frac{\hat{f}_t\left(y_{t+1}\right)}{p_t\left(y_{t+1}\right)}\right)\right] \leqslant \mathrm{E}_{p_t}\left[\frac{\hat{f}_t\left(y_{t+1}\right)}{p_t\left(y_{t+1}\right)}\right]-1=\int_{-\infty}^{\infty} p_t\left(y_{t+1}\right) \cdot \frac{\hat{f}_t\left(y_{t+1}\right)}{p_t\left(y_{t+1}\right)} d y_{t+1}-1$$

$$=\int_{-\infty}^{\infty} \hat{f}_t\left(y_{t+1}\right) d y_{t+1}-1=1-1=0, \quad \text{with equality iff } \hat{f}_t(y_{t+1})=p_t(y_{t+1})$$

Hence

$$\mathrm{E}_{p_t}\left[\log \left(\frac{p_t\left(y_{t+1}\right)}{\hat{f}_t\left(y_{t+1}\right)}\right)\right] \geq 0 \quad \text{holds}$$

### c.

We need to prove:

$$\mathrm{E}_{p_t}\left[\log \left(\hat{f}_t\left(y_{t+1}\right)\right)\right] \leq \mathrm{E}_{p_t}\left[\log \left(p_t\left(y_{t+1}\right)\right)\right], \quad \forall p_t,\ \hat{f}_t \in \mathscr{P}, \text{with equality iff } \hat{f}_t=p_t$$

Use the result from b.

$$\mathrm{E}_{p_t}\left[\log \left(\frac{p_t\left(y_{t+1}\right)}{\hat{f}_t\left(y_{t+1}\right)}\right)\right] \geq 0 \Rightarrow \mathrm{E}_{p_t}\left[\log \left(p_t\left(y_{t+1}\right)\right)-\log \left(\hat{f}_t\left(y_{t+1}\right)\right)\right] \geq 0$$

$$\Rightarrow \mathrm{E}_{p_t}\left[\log \left(\hat{f}_t\left(y_{t+1}\right)\right)\right] \leq \mathrm{E}_{p_t}\left[\log \left(p_t\left(y_{t+1}\right)\right)\right]$$

Also, we know from b, $\mathrm{E}_{p_t}\left[\log\left(\frac{p_t(y_{t+1})}{\hat{f}_t(y_{t+1})}\right)\right] = 0$ iff $\hat{f}_t(y_{t+1}) = p_t(y_{t+1})$, then it is equivalent to say

$$\mathrm{E}_{p_t}\left[\log\left(\hat{f}_t(y_{t+1})\right)\right] = \mathrm{E}_{p_t}\left[\log\left(p_t(y_{t+1})\right)\right] \text{ iff } \hat{f}_t(y_{t+1}) = p_t(y_{t+1})$$

Hence, the log-score is a strictly proper scoring rule.

**d.**

$\mathrm{H}_0$: $\mathrm{E}[d_{t+1}] = 0$ with $d_{t+1} = \mathrm{S}(\hat{f}_t; y_{t+1}) - \mathrm{S}(\hat{g}_t; y_{t+1})$ for any scoring rule S.

Under $\mathrm{H}_0$ and for the log score: $\mathrm{E}\left[\log\left(\hat{f}_t(y_{t+1})\right) - \log\left(\hat{g}_t(y_{t+1})\right)\right] = 0$

$$\begin{aligned}
&\Rightarrow \mathrm{E}\left[\log\left(\hat{f}_t(y_{t+1})\right)\right] = \mathrm{E}\left[\log\left(\hat{g}_t(y_{t+1})\right)\right] \\
&\Rightarrow \mathrm{E}\left[\log\left(p_t(y_{t+1})\right) - \log\left(\hat{f}_t(y_{t+1})\right)\right] = \mathrm{E}\left[\log\left(p_t(y_{t+1})\right) - \log\left(\hat{g}_t(y_{t+1})\right)\right] \\
&\overset{a.}{\Longrightarrow} \mathrm{KLIC}(\hat{f}_t) = \mathrm{KLIC}(\hat{g}_t)
\end{aligned}$$

The average score difference:

$$\bar{d}_{m,n} = \frac{1}{n}\sum_{t=m}^{\mathrm{T}-1} d_{t+1} = \frac{1}{n}\sum_{t=m}^{\mathrm{T}-1}\left[\mathrm{S}(\hat{f}_t; y_{t+1}) - \mathrm{S}(\hat{g}_t; y_{t+1})\right]$$

Since under $\mathrm{H}_0$ :

$$\mathrm{E}[d_{t+1}] = 0 \Rightarrow \mathrm{E}\left[\mathrm{S}(\hat{f}_t; y_{t+1}) - \mathrm{S}(\hat{g}_t; y_{t+1})\right] = 0$$

By Law of Large Numbers:

$$\frac{1}{n}\sum_{t=m}^{\mathrm{T}-1}\left[\mathrm{S}(\hat{f}_t; y_{t+1}) - \mathrm{S}(\hat{g}_t; y_{t+1})\right] \xrightarrow{\mathrm{P}} \mathrm{E}\left[\mathrm{S}(\hat{f}_t, y_{t+1}) - \mathrm{S}(\hat{g}_t; y_{t+1})\right] = 0$$

Hence

$$\bar{d}_{m,n} \xrightarrow{\mathrm{P}} \mathrm{E}[d_{t+1}] = 0$$

**e.**

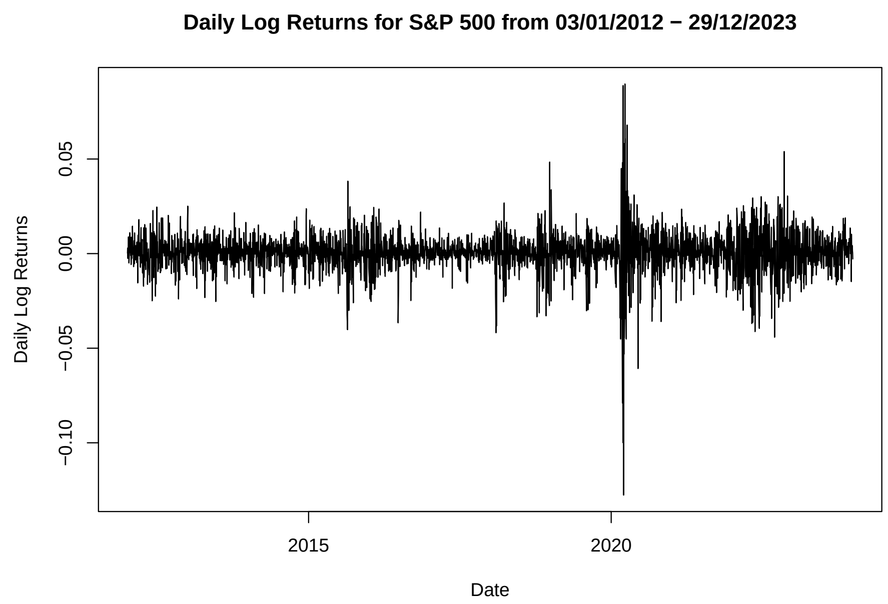
In this case, the test-statistic that is equal to the Diebold-Mariano test is the one from the Giacomini and White test. The requirement that the log score is strictly proper is satisfied. However, we cannot directly apply the Neyman-Pearson lemma to the Giacomini and White test in this scenario.

Under the assumption that $\hat{f}_t = p_t$ and $\hat{\sigma}^2$ is known, the objective of conducting the Giacomini and White test is to test whether $\hat{g}_t = p_t$, so we make a comparison between two different densities, whereas the alternative hypothesis only suggests that $\hat{g}_t \neq p_t$, without specifying the exact density of $\hat{g}_t$, so the alternative only provides a "space" of candidate densities. However, the logic behind the Neyman-Pearson lemma is that under a certain density, we test whether a certain parameter, say, $\theta$ equals $\theta_0$ under the null or equal to $\theta_1$ under the alternative, and the alternative hypothesis specifies a certain parameter value for this $\theta$ to take, instead of giving a range of possible parameter values. Hence, the Giacomini and White test is conceptually different from the NP lemma, and due to the difference in defining the alternative hypothesis, we cannot draw analogies between these two concepts. Therefore, applying the Neyman-Pearson lemma directly to the Giacomini and White test is infeasible.

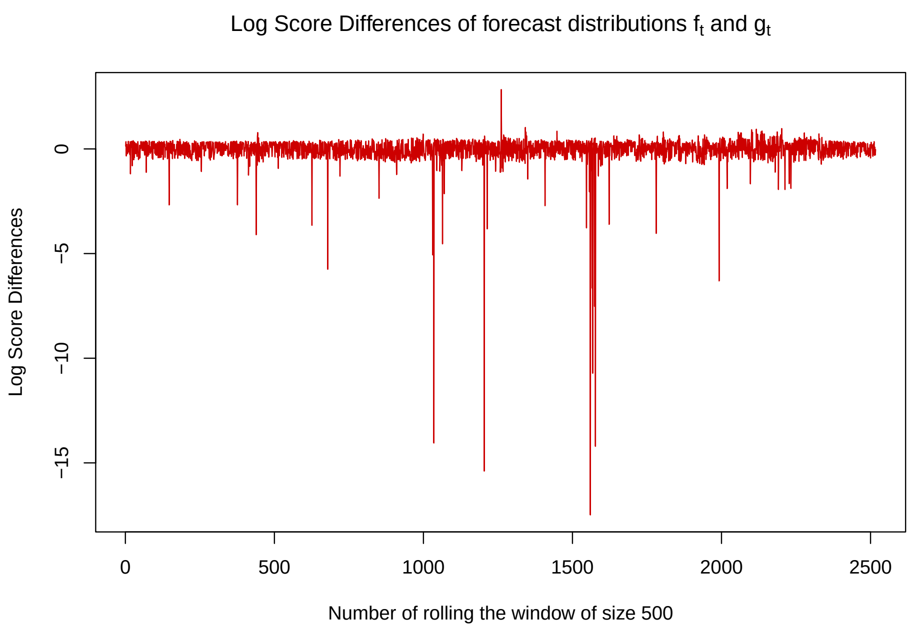
## *Part 2*

**a.**

The R script imports the daily data of the S&P 500 from January 3, 2012, up to and including December 29, 2023. Any null values in the dataset are removed. The script further processes the data to extract prices and dates, ensuring they are converted to the appropriate data types. After this, the daily log-return at time t is computed as $\log(\frac{y_t}{y_{t-1}})$, and we use them as our time series $y_t$. In part 2, we model $y_t$ as AR(5)-GARCH(1,1), and MLE is applied to estimate the parameters.




In this section, we focus on two density forecasts $f_t$ and $g_t$, which assume a standard Normal distribution and a standardized Laplace distribution respectively, for $\eta_t$. We use the rolling window method with $m = 500$ to calculate the log-likelihood for both forecast densities $f_t$ and $g_t$. Then, we use these to determine the log score differences, which are shown in the following plot:



The plot of log score differences between $f_t$ and $g_t$ suggests that, on average, both density forecasts perform equally well, with the majority of log score differences clustering around 0. However, when comparing with the plot depicting the daily log returns for S&P 500, it becomes clear that the Laplace density outperforms the Normal density when extreme high and low values of daily log-return occur. Hence, based on our observations, the Laplace density forecasts extreme gains and losses (observations in the tails) in the daily log returns more accurately than the Normal density.

**b.**

In this section, we use the log score differences from section a to perform the Diebold-Mariano test. The test statistic $t_{m,n}$ is computed via $t_{m,n} = \frac{\bar{d}_{m,n}}{\sqrt{\hat{\sigma}^2_{m,n}/n}}$, where the HAC estimator $\hat{\sigma}^2_{m,n}$ is computed using $\hat{\sigma}^2_{m,n} = \hat{\gamma}_0 + 2\sum_{k=1}^{K-1} \alpha_k \hat{\gamma}_k$, with more details provided in Diks & Vrugt (2010).

Here, we perform a one-sided test to compare the forecast accuracy of the Normal density and the Laplace density, where $H_0 : E(d_{t+1} = 0)$ vs $H_1 : E(d_{t+1}) < 0$, indicating that Laplace density outperforms the Normal density under the alternative hypothesis. The sample realization of the test statistic is -3.091294, which is smaller than the 5% critical value -1.645, suggesting that the the Laplace density outperforms the Normal density at 5% significance level.

**c.**

In this section, we calculate the coverage rate of the VaR, based on the tick-loss function, for $\alpha = 0.1$, $\alpha = 0.05$ and $\alpha = 0.01$ for both Normal and Laplace density. The coverage rates (probability) of two densities are computed via their associated hit series $x_t = \mathbf{1}_{y_t} < -\text{VaR}_{\alpha,t}$, where $\mathbf{1}_A$ is 1 if A is true and zero otherwise. Note that the negative of the $\alpha -$ th quantile of both $f_t$ and $g_t$ are positive if $\alpha$ takes 0.1, 0.05 and 0.01. The computation results are presented in the following table.

Table 1: The Coverage Rate of the VaR for both Normal and Laplace density

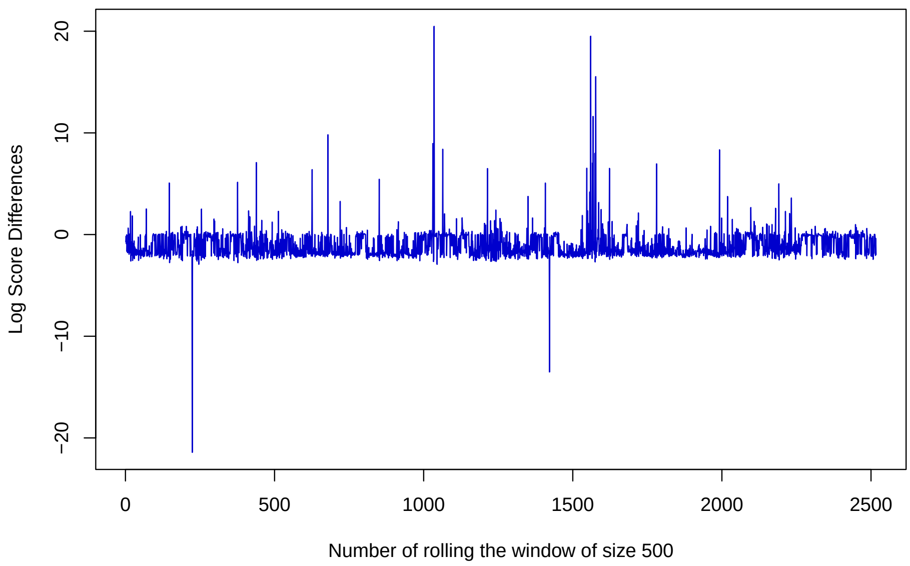
| Confidence Level | Coverage Rate for Normal density | Coverage rate for Laplace density |
|---|---|---|
| $\alpha = 0.10$ | 0.9333775 | 0.9058668 |
| $\alpha = 0.05$ | 0.9671859 | 0.9648658 |
| $\alpha = 0.01$ | 0.9897249 | 0.9933709 |

The coverage rate of the Normal density is higher than that of the Laplace density at confidence levels of $\alpha = 0.1$ and $\alpha = 0.05$, but the Laplace density shows a higher coverage rate at $\alpha = 0.01$, an really small level of confidence. These observations suggest that the Normal density outperforms in forecasting losses in daily log-returns at relatively higher confidence levels. However, when faced with extremely large losses, such as those resulting from financial crises or wars, the Laplace density has greater accuracy in capturing and forecating them. Therefore, in risk management, it is suggested to use the Normal density for forecasting losses that are not too extreme, and the Laplace density is adopted to predict extreme large losses in daily log returns.

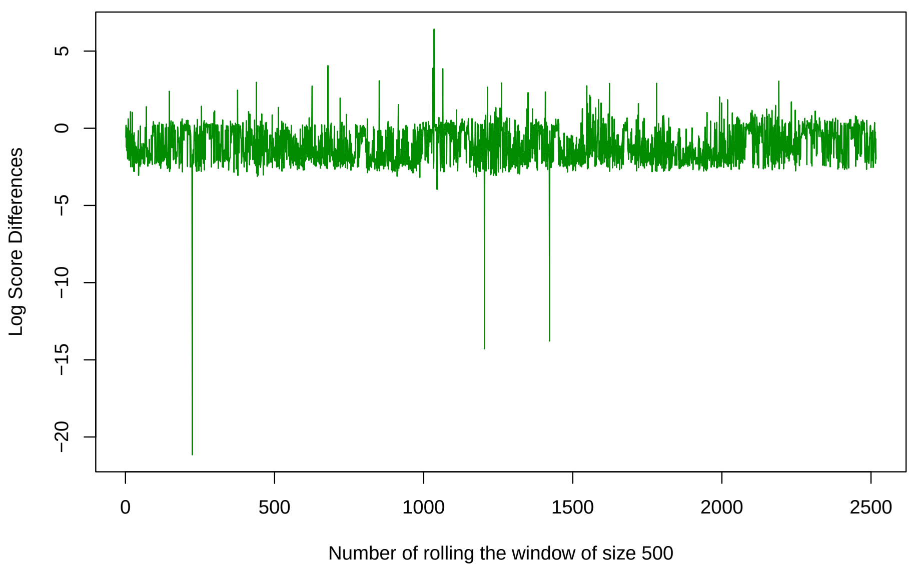
**d.**

In the section, we consider a third forecast distribution $h_t$, for which a Student-t($\nu$) distribution is taken for the innovations $\eta_t$. We repeat the calculations of the log-likelihoods with the rolling window method for $h_t$. Then calculate the log score differences relative to $f_t$ and $g_t$ and perform the Diebold-Mariano test to compare the forecast methods based on $h_t$ to $f_t$ and $g_t$.

Log Score Differences of forecast distributions $h_t$ and $f_t$




Log Score Differences of forecast distributions $h_t$ and $g_t$




Again, the three forecast densities shows comparable performance on average, with the majority of log score differences clustered around 0. The first plot indicates that the Student-t(ν) density outperforms the Normal density in forecasting extreme values, while the second plot indicates that it does not perform as well as the Laplace density in forecasting extreme gains and losses of daily log returns. The sample realizations of the two Diebold-Mariano test statistics are -28.28348 and -33.59423 respectively, and they are both smaller than the 5% critical value of -1.645. This suggests that the Normal density outperforms the Student-t(ν) density at a 5% significance level, and similarly, the Laplace density also outperforms the Student-t(ν) density at a 5% significance level.

**e.**

The iterative elimination procedure to establish the Model Confidence Set (MCS) involves first selecting a pool of candidate forecast densities, estimating the parameters within each density and evaluating each density using a consistent dataset, calculating forecast scores, and then performing pairwise statistical tests to identify models with significantly worse performance. Models failing these tests are eliminated iteratively until no further exclusions are possible.

The remaining models establish the MCS, representing the best-performing models within the specified confidence level, offering a robust selection for forecasting densities.