For this assignment, you may work in groups of two or three students. Part 1 is theoretical and part 2 is an application of your results from part 1. For each sub-question, you can earn three points. It is crucial to motivate your answers: No points will be awarded for answers without motivation. Furthermore, reports that are not typeset as a single PDF file in LaTeX will not be considered. Please also make sure to submit your R (and/or Python) codes for question 2, separately, together with a CSV file of the data. For question 2, your code should reproduce all reported results without returning errors. Results that are not returned by your code cannot be graded.

The deadline for this assignment is Mon, March 18, 15:00.

Part 1. Suppose we have two candidate density forecasts $\hat{f}_t$ and $\hat{g}_t$, living in a convex class of densities $\mathcal{P}$ that also includes the actual conditional density p_t. Define the log score differences as $d_{t+1} = \log \hat{f}_t(y_{t+1}) - \log \hat{g}_t(y_{t+1})$.

(a) Show that the expected score difference between the actual density p and (second) candidate $\hat{g}_t$ coincides for the logarithmic scoring rule with the Kullback-Leibner Information Criterion of $\hat{g}_t$, defined as

$$\mathrm{KLIC}(\hat{g}_t) = \int_{-\infty}^{\infty} p_t(y_{t+1}) \log \left(\frac{p_t(y_{t+1})}{\hat{g}_t(y_{t+1})} \right) dy_{t+1}.$$

(b) Show that

$$\mathbb{E}_{p_t} \left[\log \left(\frac{p_t(y_{t+1})}{\hat{f}_t(y_{t+1})} \right) \right] \geq 0, \quad \forall p_t, \hat{g}_t \in \mathcal{P}$$

Hint: Use the inequality $\log x \leq x - 1$.

(c) A scoring rule is proper if an incorrect density forecast $\hat{f}_t$ does, in expectation, not receive a higher score than the true conditional density. In other words, S is proper if

$$\mathbb{E}_{p_t} \left[S(\hat{f}_t; y_{t+1}) \right] \leq \mathbb{E}_{p_t} \left[S(p_t; y_{t+1}) \right], \quad \forall p_t, \hat{f}_t \in \mathcal{P}.$$

Furthermore S is strictly proper if, additionally, the weak inequality is an equality if and only if $\hat{f}_t = p_t$. Prove that the log score is a strictly proper scoring rule.

(d) The null hypothesis of the Diebold-Mariano test is defined as:

$$H_0 : \mathbb{E}[d_{t+1}] = 0$$

with $d_{t+1} = S(\hat{f}_t; y_{t+1}) - S(\hat{g}_t; y_{t+1})$ for any given scoring rule S. Prove that this null hypothesis for the log score implies that $\mathrm{KLIC}(\hat{f}_t) = \mathrm{KLIC}(\hat{g}_t)$. Also prove that the average score difference goes to $\mathbb{E}(d_{t+1}) = 0$.

(e) Now suppose that $\hat{f}_t = p_t$, $\hat{\sigma}^2$ is known and we consider the log sore difference rule again. Then the test statistic from the Diebold-Mariano test is equal to another known test statistic. Which one? Can we directly apply the Neyman Pearson lemma to this test [Hint: be careful]?

Part 2. In each of the sub-questions below, make sure to program everything from scratch. Besides base R/Python, your code from previous assignments, packages for plotting and optimization, you are not allowed to use any other packages for your final results. Here, 'base Python' refers to any basic operations captured by Numpy and other packages providing operations that translate to R.

Download daily data from the website of Yahoo Finance for the S&P 500 (ticker: ˆ GSPC) from January 3, 2012, up to and including December 29, 2023. Save as a CSV file and enable your script to read the data. Use the adjusted closing prices as our time series. Note that the data may contain days without a value (non-trading days). Make sure you remove these empty entries. Calculate the daily log-returns and use these as y_t.

We consider an AR(5) model for the conditional mean return together with a GARCH(1,1) model for the conditional variance:

$$y_t = \mu_t + \varepsilon_t = \mu_t + \sqrt{h_t}\eta_t$$

$$\mu_t = \rho_0 + \sum_{j=1}^{5} \rho_j y_{t-j}$$

$$h_t = \omega + \alpha\varepsilon_{t-1}^2 + \beta h_{t-1}$$

Furthermore, we only focus on the log score differences as scoring rule.

(a) Suppose we focus on two forecasts methods first: f_t, for which we assume a standard normal distribution for η_t, and g_t, for which we assume a standardized Laplace distribution instead.

Use the rolling window method with $m = 500$ to calculate the log-likelihood of the forecast methods for both forecast distributions f_t and g_t. Then use these to determine the log score differences. Show the log score differences in a plot.

(b) Now, use the log score differences from (a) to perform the Diebold-Mariano test. What is your conclusion based on a significance level of 0.05?

(c) 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 forecast distributions. Show the results in a table.

(d) Now, consider a third forecast distribution h_t, for which a Student-t(ν) distribution is taken for the innovations. The varying degrees of freedom ν must be determined using

MLE. 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.

(e) We now have three candidates for the density forecasts and their pairwise comparisons based on the Diebold Mariano tests above. Since multiple pairwise comparisons do not extend to valid statements regarding multiple hypotheses, it is good practice to complete the analysis by deriving the Model Confidence Set (MCS), which includes the set of best models with a confidence level of 0.90. First, explain in your own words how the iterative elimination procedure that leads to the MCS works. Second, determine which of the three forecast methods survives this iterative procedure, that is, determine the MCS of the three candidates.

