

FINANCIAL STATISTICS

2024-12-10

Time:	08.00 - 13.00
Place:	TBD
Approved aid:	available for download at the SEB-imbedded homepage of the exam; calculator

Questions about the content of the exam: The examiner (or the TA) plans to stop by the exam venue **at the start and one more time during the exam.**

To upload the R-packages use plain R. If you try to upload the packages in RStudio, your PC might become irresponsive and you need to restart the RStudio or/and computer. There is “.txt” help file that contains a function to install most of needed packages. It is easy to copy/paste. There is a separate help file for R-commands available on the homepage of the exam. After the packages are uploaded using the nearest mirror site (usually, Umeå), you can link the libraries using RStudio. FinTS package is wise to unpack only at the end, when one tests for the ARCH effect. If you need to upload some additional packages, do so using plain R. If an R-package you want to use is missing from the provided list, indicate clearly in your exam report the name of the package and for what purpose you use it. No IT help is available for this step.

Remember to save your progress in RStudio frequently in case you need to re-start RStudio/PC.

Below are the examples of possible instructions if you decide to make separate uploads:

```
install.packages("forecast"); library(forecast)
```



```
install.packages("tseries"); library(tseries)
```

```
install.packages("lmtest"); library(lmtest)
```

```
install.packages("FinTS"); FinTS::ArchTest()
```

If you ask a question during the exam, you will get a clarifying answer about the content of the exam question, not an answer how to solve it. No IT help is available during the exam.

- **Part I (Time Series, R-Studio based)**

Your task is to analyze **DAILY** stock data **provided in .csv format** using RStudio and submit a written report in an essay form. Your essay should contain the key R-commands, the output from R-Studio as well as clear interpretation of results. Submissions in .docx and .pdf formats are ok but **the .pdf format is preferred**. In case, the key R commands are not clearly incorporated in the

text of your answers, the final grade will be drastically reduced. In case there is no output or interpretation of the output, zero credit is given.

Please attach your complete R-code as a separate file for the reference purposes. The file that contains the R-code is not graded and rarely looked into for the grading purposes. All important information should be present in your report.

- **Part II (Exercises)**

You are expected to provide detailed solutions to the asked questions and write them clearly on the paper provided at the exam. Submit your solutions on time. The answers (correct or not) without solutions receive zero credit for grading purposes.

Reminder: the first page of your answers for Part1 and ALL the sheets you are submitting as solutions for Part2 should contain your anonymous code and be ordered the way you want your answers are read. If labelling of the pages is not properly done, one cannot guarantee that the pages that are not marked properly will be graded.

Upload Deadline: 13:00 (no submissions after the deadline are graded)

- **Grading, minimum requirements:**

- E: fully answer “Part I” a-d and make a fair attempt on remaining of “Part I” & at least one of the questions in “Part II”. Satisfactory answers in “Part II” or remaining of “Part I” may compensate for somewhat lacking but not wrong answers in a)-d) of “Part I”
- C: fully answer Part I a-d, answer e-g in a satisfactory way; make a fair attempt on Part II (two questions) or solve one of the problems from Part II in a satisfactory way. Satisfactory answer to Part II (two questions) may compensate for lacking in e-g of Part I
- A: answer Part I in full with minor deficiencies. Answer Part II (two questions) in a satisfactory way

These are preliminary requirements. The examiner might adjust the scale upward, if necessary.

- NOTE! Fx and F are failing grades that require re-examination. Students who receive the grade Fx or F cannot supplement for a higher grade.
- If you are not satisfied with your grade, you may contact the examiner for further instructions.

GOOD LUCK

Part I, Data

Below, you will find a selection of time series data sets. Your data corresponds to the number from the anonymous code, assigned to you for this exam. For instance, if you have the anonymous code 311-0001-XXX, your data is “aapl_us_d”, Nvidia Corp. Upload available .csv file as explained during lectures/labs.

Code	Symbol	Name
0001	aapl_us_d	Apple Inc
0002	amd_us_d	Advanced Micro Devices Inc
0003	intc_us_d	Intel Corp
0004	mara_us_d	Mara Holdings Inc
0005	nvda_us_d	Nvidia Corp
0006	rivn_us_d	Rivian Automotive Inc
0007	rklb_us_d	Rocket Lab USA Inc
0008	smci_us_d	Super Micro Computer Inc
0009	snap_us_d	Snap Inc
0010	sofi_us_d	SoFi Technologies Inc
0011	aapl_us_d	Apple Inc
0012	amd_us_d	Advanced Micro Devices Inc
0013	intc_us_d	Intel Corp
0014	mara_us_d	Mara Holdings Inc
0015	nvda_us_d	Nvidia Corp
0016	rivn_us_d	Rivian Automotive Inc
0017	rklb_us_d	Rocket Lab USA Inc
0018	smci_us_d	Super Micro Computer Inc
0019	snap_us_d	Snap Inc
0020	sofi_us_d	SoFi Technologies Inc
0021	aapl_us_d	Apple Inc
0022	amd_us_d	Advanced Micro Devices Inc
0023	intc_us_d	Intel Corp
0024	mara_us_d	Mara Holdings Inc
0025	nvda_us_d	Nvidia Corp
0026	rivn_us_d	Rivian Automotive Inc
0027	rklb_us_d	Rocket Lab USA Inc
0028	smci_us_d	Super Micro Computer Inc
0029	snap_us_d	Snap Inc
0030	sofi_us_d	SoFi Technologies Inc
0031	aapl_us_d	Apple Inc
0032	amd_us_d	Advanced Micro Devices Inc
0033	intc_us_d	Intel Corp
0034	mara_us_d	Mara Holdings Inc
0035	nvda_us_d	Nvidia Corp
0036	rivn_us_d	Rivian Automotive Inc
0037	rklb_us_d	Rocket Lab USA Inc
0038	smci_us_d	Super Micro Computer Inc
0039	snap_us_d	Snap Inc
0040	sofi_us_d	SoFi Technologies Inc
0041	aapl_us_d	Apple Inc
0042	amd_us_d	Advanced Micro Devices Inc

The goal of the assignment in Part I is to analyze assigned time series in RStudio and then write an essay answering the questions below. Make one headline in the essay for each part that you answer, e.g. “a) Describe the time series and test for trend” and “b) Stationarity.” Preferably, use a new page when you answer a new question. It is important that you include plots/tables that are outputs preceded by the key R-commands. Use “Close” column from the data. In case you need to convert the character format of the date in the first column, you might want to use the following (or similar) command: `Ticker$Date ← as.Date(Ticker$Date, format = “%d / %m / %Y”)`

- (a) Describe the time series and test for trend with an appropriate plot/diagram. Comment on the diagram and provide summary statistics. What properties of “stationarity” look like not fulfilled? Is there a visible trend? Test for the trend using simple linear regression. You can choose different time intervals to “prove” that the time series has extensive time periods of trends, if any.

For steps b, c, and d, leave out the last 10 observations that will be referred to as “the testing set”.

- (b) Stationarity. Use a formal test for stationarity and for detecting a trend in the time series. If it is clear from the diagram that your data is non-stationary, you may apply an appropriate transformation but perform both tests first. State the hypotheses and interpret the outcome of each test. Use 5% significance level. If the data does not seem to be stationary even after transformation, attempt another transformation and repeat the stationarity testing. Make sure that you have the figure containing the original and the stationary time series. In case you have to choose between mildly (visualize the time series in question) non-stationary time series and “white noise” style time series, choose the former in favor of the latter.
- (c) Plot the ACF and the PACF graphs. Choose the number of lags appropriately and justify your choice. Explain what the plots show. Specifically state and interpret the values of the first (1st) bar in both ACF and PACF plots. Give an outline (no exact calculations required) of how you would calculate them if no R-Studio is available. Compare the two values. What is the difference? Discuss. For the A-B grade, remember to return to these plots when you have chosen “the best” model. Make specific references to the models you selected for the step (d). Use these plots as the basis for your choices in (d).
- (d) Whatever the (P)ACF plots you get, choose (by performing proper statistical analysis) at least four the most promising ARIMA models that are the top candidates to fit the data (leaving out the testing set). One of the remaining models should be an ARIMA(0,d,0): indicate what “d” you have chosen and why. You have to have a detailed description of your selection process. Explain why selected by you models are better than other potential candidates are. Provide output, analysis and calculate the AIC scores for each “potentially good” model. Interpret your findings. Which model seems most promising? Should any of the checked models be disregarded completely? Why?
- (e) Select the two “best” models from (d). Use your testing set to calculate RMSE. The calculations should be explicit and easy to follow. Compare the two and interpret the difference. If needed, perform an exploratory residual analysis to select the best model. Moreover, visualize both predictions and discuss the figures. Make clear and well-motivated conclusions. Specifically discuss the values of the RMSE for the selected models and make conclusions on the quality of the models. Finally, use Holt method to have a benchmark model. Compare

the quality predictions of the selected ARIMA models and the Holt method. Discuss the differences.

- (f) Pick “the best” ARIMA model from part (e). Test it for the “ARCH effect”. State the hypothesis and explain the outcome of your test. Explain what presence of ARCH effect would mean for your results in (d), even if you do not find any. Discuss briefly, why estimating variance is important in financial forecasting. Visualize the ARCH effect for your model.
- (g) Pick “the best” model from question (e). Write down the model explicitly and discuss the coefficients. Discuss connection to the (P)ACF plots. Provide residual analysis of the model and report your conclusions. Further, provide overall conclusions for your analysis.

END OF PART I

Part II (“Pen-and-Paper-Answer” questions: one cannot provide R-Studio based solutions. Write your detailed answers on paper)

1. Let us assume that we consider an ARCH-style time series X_t with a drift, i.e.

$X_t = 0.1 + X_{t-1} + \varepsilon_t$ with $\text{Var}(\varepsilon_t) = h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2$; further, let us assume that $\alpha_0 = 0.2$, $\alpha_1 = 0.3$ and the error terms to be normally distributed.

Calculate the volatilities $\sqrt{h_3}$, $\sqrt{h_4}$ for the time stamps $t=3$, $t=4$ and

the probability $P(X_4 > 10.4)$ if we observe that $X_1 = 10.7$, $X_2 = 11.2$, $X_3 = 9.8$. Further, assume that $\alpha_0 = 0.2$, $\alpha_1 = 0$ and calculate $P(X_4 > 10.4)$.

What is the difference between two probabilities? Why? Provide “financial intuition” to the difference of the probabilities when one changes α_1 .

For a “fair” grade, one is expected to calculate both probabilities. For a “satisfactory” grade, one is expected to compare both probabilities and give a clear explanation how changing parameter α_1 influenced the result and why.

2. Provide detailed calculations for the following questions. Note that the correct answer without calculations is not accepted as a solution and will be marked “empty+”. For a “fair” attempt grade, one has to have a “fair” attempt on both parts of the exercise:

a) Derive the autocorrelation function for the following MA(1) process $Z_t = \varepsilon_t - 2\varepsilon_{t-1}$. Assume that Z_t is stationary.

b) Derive the autocorrelation function for the following MA(2) process $Z_t = \varepsilon_t - 2\varepsilon_{t-1} + 0.5\varepsilon_{t-2}$. Assume that Z_t is stationary.

For a “fair” grade, one is expected to solve part a) and clearly show how to calculate part b). For a “satisfactory” grade, one is expected to answer both questions a) and b) with minor mistakes.

GOOD LUCK
