

This question paper consists of 3 printed pages, each of which is identified by the reference **MATH5315M**.

©UNIVERSITY OF LEEDS  
(Autumn 2025)

Assessed coursework for the degree of  
MSc Financial Mathematics

**MATH5315M Applied Statistics and Probability**

Assignment (worth 30% of the module mark)

---

*This assignment consists of 3 pages and 4 main questions.*

*Hand-out date: 19th November, 2025.*

*Hand-in date: 12 noon, 17th December, 2025.*

**Note:** Your report should be readable as a standalone document and understandable to an independent reader. It should **not** simply be a collection of R commands and pasted raw R output without any further explanation. However, you can (and I recommend that you do) include important figures/plots in your main report with explanations. Your final mark will consider the quality of presentation and clarity.

---

1. A computer company wants to understand how long their hard drives last before failing. The data file `hard_drive.txt` contains the lifetimes in yrs (recorded to two decimal places) of 200 hard drives that they have tested.

Draw a histogram and a boxplot for these data. Try different break points for the histogram (see the help pages), and conclude that the hard drive lifetimes  $X$  can be modelled as an exponential or gamma distribution.

Fit an exponential distribution and a gamma distribution to the lifetimes using the method of moments. Write a report that briefly explains why the exponential distribution might be an appropriate distribution for these times and compares the two fitted distributions with the sample data. As part of the report, provide some plots that highlight the differences between the fitted distributions and their level of their agreement with the sample data. (Any calculations or raw code output from a statistical package should be added as an appendix.)

**(6 out of 30 marks, no more than 3 sides of A4)**

2. The spreadsheet `EuroIndices_Close.csv` contains daily closing price information for four stock indices over a 150-trading-day period.

*Part I* — First let us consider just the DAX data (a stock market index of 40 major German companies):

- (a) Attempt to fit at least three different classes of time series model to the DAX data.
- (b) Using appropriate criteria, comment on which model is the best fit to the data.
- (c) Using your selected model, produce estimates for the DAX index over the next 10 trading days after the period for which you have data.

*Part II* — Now consider all four indices:

- (a) Fit an appropriate vector autoregressive model to all four indices simultaneously.
- (b) Consider the value of using of all the data in the context of estimating the DAX index over the next 10 trading days.

Report on your findings for both part I and II of the question including any analysis of underlying model assumptions that you may have made. For all parts of this question, any calculations or raw code output from a statistical package should be added as an appendix.

**(9 out of 30 marks, no more than 5 sides of A4)**

3. The file `deaths.csv` contains the UK government's official total number of weekly deaths of those who previously tested positive with COVID-19. Find a suitable time series model for this dataset, and use it to predict the following two weeks, with corresponding 95% confidence intervals for your prediction. Repeat the task using the square root of the data, but back-transforming the forecast (and intervals) to the original scale. Interpret your model, and state, with reasons which model you prefer.

For all parts of this question, any calculations or raw code output from a statistical package should be added as an appendix.

**(8 out of 30 marks, no more than 4 sides of A4)**

4. The data file `USD-GBP.csv` contains the Forex daily close data for the USD to GBP exchange for a twelve year period. Use appropriate time series models to investigate the change in volatility in the exchange rate before and after the credit crisis (for simplicity, you can assume that the credit crisis struck overnight between 31st July 2007 and 1st August 2007).

Write a report on the changing volatility over this twelve year period, which includes your fitted models and any analysis of underlying model assumptions that you may have made. Once again, any calculations or raw code output from a statistical package should be added as an appendix.

**(7 out of 30 marks, no more than 3 sides of A4)**

---

All of the data files that are needed to answer the questions are available from the MATH5315M module page on Minerva. There is no word count for this assignment, but you are limited to 15 sides of A4 paper (not including any appendices). Important plots should go in the **main report** and not the appendices.

The deadline date for this assignment is **12.00 noon on 17th December 2025**.

---

An electronic copy of the assignment must be submitted to the Assignment Submission area within the MATH5315M module resource on Minerva no later than 12 noon on the deadline date.

Failure to meet this initial deadline will result in a reduction of marks. Emailed or hard copies of the assignment will not be accepted. Please ensure that the file name is saved as your student ID e.g. (212345678.pdf) and your name is not written anywhere on your assignment.

END