

UNIVERSITY OF EXETER
DEPARTMENT OF MATHEMATICS AND STATISTICS
MTH1004 Probability, Statistics and Data

Summative coursework for term 2

Publication date: Friday 28 February 2025
Submission date: Friday 28 March 2025 by noon
Return date: Friday 2 May 2025

Instructions

This coursework comprises 15% of the overall module assessment. It is an individual assessment and you must not discuss it with anyone other than the lecturer. This assessment is also AI-prohibited. Please refer to our [regulations on academic conduct](#) and the guidance below about the use of generative artificial intelligence (GenAI) tools.

Conduct the investigation outlined below and write a report of your findings, addressing each part of the investigation in a separate section. The data set needed for the investigation is contained in the file `MTH1004T2CW.RData`, which you should download from the [summative assessment tile on the module's ELE page](#).

Submit your answers on ELE by the date shown above. You must submit a Zip archive containing two files: one R script (ending `.R` or `.r`) containing the R code that you used to produce your results, and one pdf file containing your report. Your R code should be formatted to aid readability, should include brief comments to explain what each section of code does, and should run without errors. Your report (excluding your code) should comprise at most three sides of word-processed text (e.g. written with Word or LaTeX) followed by as many pages of figures or tables that you feel is appropriate. No credit will be awarded to additional pages of text. (You do not need to write three sides of text to achieve full marks—the model solution has only two sides of text and about 900 words.) You must use the page size A4 and a font size of at least 11 points, lines must be single-spaced and all margins must be at least one inch wide. Do not submit any other files, such as extra R scripts or separate figure files.

Investigation

The number of hurricanes that hit the Atlantic coast of the United States of America each year was recorded from 1870 to 2024. Temperatures in the Pacific Ocean were also recorded each year and used to calculate an El Niño Southern Oscillation (ENSO) index which measures (in degrees Celsius) how much above or below normal the ocean temperatures are. When the ENSO index is positive, ocean temperatures are warmer than normal and this is believed to decrease the chance of hurricanes forming in the Atlantic; when the ENSO index is negative, ocean temperatures are cooler than normal and this is believed to increase the chance of hurricanes forming in the Atlantic. The file `MTH1004T2CW.RData` contains a data frame called `Hurricanes` with three columns of data: `Year` (year), `Number` (number of hurricanes) and `ENSO` (ENSO index).¹

¹These data are provided by the NOAA Physical Sciences Laboratory from their website at <https://psl.noaa.gov/> and the Hurricane Research Division of the NOAA Atlantic Oceanographic and Meteorological Laboratory from their website at <https://www.aoml.noaa.gov/>.

1. Consider modelling the numbers of hurricanes as realisations of independent and identically distributed random variables. Choose a parametric model for these data from among the distributions discussed in the lecture notes. Then estimate the model parameter (or parameters), estimate the standard errors of your parameter estimates, and assess the realism of your model. Include in your answer the formulae for your parameter estimates, the numerical values of the parameter estimates and standard errors, a discussion of your model's realism and appropriate numerical and graphical evidence. You do not need to include the derivation of your formulae and your model does not need to fit the data very well to receive full marks.
2. Use point estimates and confidence intervals to estimate the expected number of hurricanes when the ENSO index is positive and when it is negative, and to assess whether the expected number of hurricanes changes with the sign of the ENSO index. Include in your answer the formulae for your confidence intervals as well as their numerical values.
3. (a) Compute an approximate 90% prediction interval for the number of hurricanes in 2025, and a probability prediction for there being three or more hurricanes in 2025, and comment on the trustworthiness of your predictions.
(b) Suppose that a climate scientist tells you that the probability of the ENSO index being positive in 2025 is 0.25. Revise your probability prediction from the previous part in light of this information. (This final part is intended to be challenging. There are no similar examples in the module but it can be answered with the insightful use of material seen during the module.)

Marking criteria

Of the 100 marks available, approximately 35 are available for part 1, 20 for part 2, 25 for part 3 and 20 for the presentation of your report. Of the 80 marks available for parts 1, 2 and 3, approximately 20 are awarded for your method, 40 for your results and 20 for your discussion of your results. To achieve a pass mark, you will have applied the main inferential methods without major errors, presented your results intelligibly, interpreted your results without major errors, and written an intelligible report. To achieve a first-class mark, you will have selected appropriate inferential methods, applied them accurately, presented your results effectively, interpreted your results accurately and with an appropriate level of detail, and written a clear report.

Use of GenAI tools

The University of Exeter is committed to the ethical and responsible use of Generative Artificial Intelligence (GenAI) tools in teaching and learning, in line with our academic integrity policies. Direct copying of AI-generated content can be an academic offence and is addressed under plagiarism, misrepresentation and contract cheating in [section 12.3](#) of the Teaching Quality Assurance manual. To support students with assessments, staff will identify whether the use of GenAI tools is integrated, supported or prohibited in each assessment. Further guidance on using GenAI tools to enhance your learning, and on referencing them appropriately, is available on [Study Zone digital](#).

This assessment is AI-prohibited. You must not use GenAI tools for this assessment. This is because you will demonstrate that you have achieved the intended learning outcomes

only if you complete the assessment without using GenAI tools. If you do use GenAI tools for parts of this assessment then you will receive no marks for those parts. If you use GenAI tools for parts of this assessment and do not declare in your submission that you have done so, or fail to reference the GenAI outputs in your work, then this would be an academic offence.

If markers think that you might have committed an academic offence then they may require you to attend a viva (oral exam) in order to establish the legitimacy of your work.