INSTITUTE AND FACULTY OF ACTUARIES

EXAMINATION

12 September 2023 (am)

Subject CP2 – Modelling Practice
Core Practices

Paper One

Time allowed: Three hours and twenty minutes

In addition to this paper you should have available the 2002 edition of the Formulae and Tables and your own electronic calculator.

If you encounter any issues during the examination please contact the Assessment Team on T. 0044 (0) 1865 268 873.

© Institute and Faculty of Actuaries

Exam requirements

1 Modelling steps and data checks

Read the background document, which describes the scenario that needs to be modelled and documented for this project.

Construct a spreadsheet model that produces the following calculations and charts. You should ensure that your spreadsheet contains appropriate self-checks and that you have performed reasonableness checks at each stage of your calculations.

Modelling steps:

(i) Perform thorough checks on the data provided to confirm that the data is complete and fit for use. This should include the following:

(a) Calculate suitable summary statistics that add value to the checking process. Both the list of weather stations and the temperature data should be checked. Categorisation of the data (for example, by year or by month) would also be useful to identify trends and perform checks.

(b) Construct suitable charts to illustrate the accuracy and suitability of the data.

Determine suitable values to use instead of the default values that are present in the data. [9]

(ii) For each weather station, for each year, calculate:

(a) average yearly temperatures,

(b) average summer temperature and

(c) average winter temperature,

and determine which group each weather station fits into within the north/south hemisphere, continent and high/low elevation categories. [5]

(iii) Summarise the overall average annual temperature, across all weather stations, for each year and in each category (north/south hemisphere, continent, high/low elevation, winter/summer).

For each category, construct a chart to illustrate how annual temperatures vary over time (from 1961 to 2010). [8]

(iv) Construct a linear regression model of average summer temperatures. Use this to predict temperatures from 2011 to 2030, and construct a chart of summer temperatures from 1961 to 2030. [7]

(v) Determine the expected cost in each year from 2021 to 2030. [3]

[Sub-total 32]

2 **Marks available for spreadsheet model and checks:**

(i) Automatic checks on the modelling completed in (1). [2]

(ii) Demonstration of good modelling technique and practice. [7]

[Sub-total 9]

3 **Audit trail**

Prepare an audit trail for your spreadsheet model that includes the following aspects:

- Purpose of the model
- Data and the reasonableness of the data
- Assumptions used
- Methodology, i.e. description of how each calculation stage in the model has been produced
- Explanation of any checks performed.

You should ensure that your audit trail is suitable for both a senior actuary, who has been asked to approve your work, and a fellow student, who has been asked to peer review and correct your model, or may be asked to continue work on it, or to use it again for a similar purpose in the future.

Marks available for audit trail:

Audit approach

(i) Communication skills. [4]

(ii) Fellow student can review and check methods used in the model. [7]

(iii) Senior actuary can scrutinise and understand what has been done. [7]

(iv) Written in clear English. [4]

(v) Written in a logical order. [3]

Audit content

(vi) All steps clearly explained. [7]

(vii) Reasonableness checks included. [5]

(viii) Clear signposting included throughout. [4]

(ix) Statement of assumptions made. [4]

(x) All model steps accurately covered. [14]

[Sub-total 59]

[Total 100]

Background

You are an actuarial student in a consulting firm. Your manager is working with a multinational corporation that is interested in the impact of climate change on their worldwide operations. The company runs refrigerated storage facilities and is concerned that rising global temperatures will have an impact on their running costs.

Your manager has sourced a dataset of suitable climate data from the National Centres for Environmental Information (NCEI) and has asked that you analyse the changes in temperatures and determine a long-term trend that could be used as an input into the client's predictive cost model. The dataset of monthly mean temperatures covers average monthly temperatures for weather stations all over the world for each month for the 50 years from 1961 to 2010. A sample of 246 weather stations has been taken, which represents the locations where the client has operations. The data consists of two tables – one with location details of each weather station, and a separate table with the average monthly temperature in degrees Celsius for each weather station. Where temperature measurements or weather station information is not available, the NCEI uses default values to populate the dataset.

Before you use the data, you should check and adjust it, where needed. You then need to:

1. construct suitable summary statistics that add value to the checking process.
2. construct suitable charts to illustrate the accuracy and suitability of the data.
3. determine suitable values or data entries to use instead of the default values that are present in the data.

To support your client's investigations:

1. calculate the annual average temperature for each weather station for each year, and then group the stations by:
   - hemisphere: northern and southern (latitude is higher or lower than zero).
   - each of the six continents.
   - elevation: high and low (above or below 500 m above sea level).

   [**Note**: see 'Additional guidance' (1 and 2).]

2. calculate the average summer and winter temperature each year, where summer in the northern hemisphere is taken as April to September, and in the southern hemisphere October to March (both inclusive).

The client has indicated that average summer temperatures have the most impact on refrigeration costs and that an increase in summer temperatures result in an increase in cost of $e^{k\Delta T}$, where k is a constant of 0.4, and ΔT is the annual change in summer temperature; that is:

$$\text{cost in year } y + 1 = \text{cost in year } y \times e^{k\Delta T}$$

where ΔT is the change in temperature from year y to year $y + 1$.

By determining a trend for overall summer temperatures (using linear regression), you will be able to predict summer temperatures from 2011 to 2030 (see 'Additional guidance' (3)) and subsequently the changes in costs over the same period.

The client is looking for the average predicted summer temperature from 2026 to 2030, and what level of increase this represents to the average actual summer temperature recorded from 2006 to 2010. They would also like to see the expected average costs between 2026 and 2030, given that refrigeration costs in 2020 were $250 million.

In addition, the model needs to:

- construct charts to show the movement in temperatures over time (1961 to 2010) for each of the categories.
- construct a chart showing the average summer and winter temperatures over time (1961 to 2010).
- construct a chart showing the actual average summer temperatures up to 2010 and the predicted temperatures from 2011 to 2030 obtained from the linear model.

Additional guidance:

1. Terminology

   - Latitude: this is a measure of distance in degrees to the north (positive values up to 90) or south (negative values down to –90) from the equator (where latitude value is 0).
   - Longitude: this is a measure of distance to the east (positive values up to 180) or west (negative values down to –180) from the line that runs through Greenwich in London (where longitude value is 0).
   - Elevation: this is a measure of height above sea level in metres.

2. You could use the AVERAGEIFS(data, condition range, condition, …) function. This takes the average of a series of numbers that comply with certain conditions. For example:

   AVERAGEIFS(H:H, B:B, "<0", C:C, "North") returns the average of all figures in column H where the value in column B is less than zero, and the value in column C is 'North'.

3. There are a number of ways to get a linear regression in Excel, and any correct method will obtain credit.

   Possible approaches include the following:

   - Using the LINEST(Known_ys, Known_xs) function: this is an array function, which returns m and c in the formula $y = mx + c$ if given a series of y and x. Because it returns two values, you need to select two cells next to each other, enter in the formula and press Ctrl+Shift+Enter to enter it as an array (you will see that Excel displays the formula enclosed in curly brackets: { }. If this is not done, the formula will return the value of m only). If you are struggling with this, use the calculated value of m and a suitable value of c and continue with the rest of the exam paper.
   - Adding a trendline to a graph: right-click the line on the graph you want to do the regression for and select 'Add Trendline'. In the dialogue box, select 'Display Equation on chart', and this will provide you with the $y = mx + c$ formula. You can then use the values of m and c elsewhere in the model.

   [**Note**: other approaches are also acceptable.]

END OF PAPER