Area of Study 1 – Unit 3

Core

Data analysis

Investigating data distributions, including:

- review of types of data

- review of representation, display and description of the distributions of categorical variables: data tables, two-way frequency tables and their associated segmented bar charts

- use of the distribution/s of one or more categorical variables to answer statistical questions
- review of representation, display and description of the distributions of numerical variables: dot plots, stem plots, histograms; the use of a log (base 10) scale to display data ranging over several orders of magnitude and their interpretation in powers of ten
- summary of the distributions of numerical variables; the five-number summary and boxplots (including the use of the lower fence (Q $-$ 1.5 × IQR) and upper fence (Q+ 1.5 × IQR) to identify and display possible outliers);
- the sample mean and standard deviation and their use in comparing data distributions in terms of centre and spread

- use of the distribution/s of one or more numerical variables to answer statistical questions
- the normal model for bell-shaped distributions and the use of the 68–95–99.7% rule to estimate percentages and to give meaning to the standard deviation; standardised values (z-scores) and their use in comparing data values across distributions
- population and sample, random numbers and their use to draw simple random samples from a population or randomly allocate subjects to groups, the difference between population parameters (e.g., μ and σ), sample statistics (e.g., x and s).

Investigating associations between two variables, including:

- response and explanatory variables and their role in investigating associations between variables
- contingency (two-way) frequency tables, two-way frequency tables and their associated bar charts (including
- percentaged segmented bar charts) and their use in identifying and describing associations between two
  categorical variables
- back-to-back stem plots, parallel dot plots and boxplots and their use in identifying and describing associations
  between a numerical and a categorical variable
- scatterplots and their use in identifying and qualitatively describing the association between two numerical
  variables in terms of direction (positive/negative), form (linear/non-linear) and strength (strong/moderate/weak)
- answering statistical questions that require a knowledge of the associations between pairs of variables
- Pearson correlation coefficient, r, its calculation and interpretation
- cause and effect; the difference between observation and experimentation when collecting data and the need
  for experimentation to definitively determine cause and effect
- non-causal explanations for an observed association including common response, confounding, and coincidence;
  discussion and communication of these explanations in a particular situation in a systematic and concise
  manner.

Investigating and modelling linear associations, including:

- least squares line of best fit y = a + bx, where x represents the explanatory variable and y represents the response
  variable; the determination of the coefficients a and b using technology, and the formulas

$$b = r\frac{s_y}{s_x} \text{ and } a = \overline{y} - b\overline{x}$$

- modelling linear association between two numerical variables, including the:
  ✓ identification of the explanatory and response variables
  ✓ use of the least squares method to fit a linear model to the data
- interpretation of the slope and intercepts of the least squares line in the context of the situation being modelled,
  including:

  ✓ use of the rule of the fitted line to make predictions being aware of the limitations of extrapolation
  ✓ use of the coefficient of determination, r^2, to assess the strength of the association in terms of explained variation
  ✓ use of residual analysis to check quality of fit

- data transformation and its use in transforming some forms of non-linear data to linearity using a square, log
  or reciprocal transformation (on one axis only)
- interpretation and use of the equation of the least squares line fitted to the transformed data to make predictions.

Investigating and modelling time series data, including:

- qualitative features of time series plots; recognition of features such as trend (long-term direction), seasonality
  (systematic, calendar related movements) and irregular fluctuations (unsystematic, short-term fluctuations);
  possible outliers and their sources, including one-off real world events, and signs of structural change such as
  a discontinuity in the time series
- numerical smoothing of time series data using moving means with consideration of the number of terms
  required (using centring when appropriate) to help identify trends in time series plot with large fluctuations
- graphical smoothing of time series plots using moving medians (involving an odd number of points only) to help
  identify long-term trends in time series with large fluctuations
- seasonal adjustment including the use and interpretation of seasonal indices and their calculation using seasonal
  and yearly means
- modelling trend by fitting a least squares line to a time series with time as the explanatory variable (data de-seasonalised where necessary), and the use of the model to make forecasts (with re-seasonalisation where necessary) including consideration of the possible limitations of fitting a linear model and the limitations of extending into the future.

Recursion and Financial Modelling

This topic covers the use of first-order linear recurrence relations and technology to model and analyse a range of financial situations, and solve related problems involving interest, appreciation and depreciation, loans, annuities and perpetuities.

Depreciation of assets, including:

- review of the use of a first-order linear recurrence relation to generate the terms of a sequence
- use of a recurrence relation to model and compare (numerically and graphically) flat rate, unit cost and reducing balance depreciation of the value of an asset with time, including the use of a recurrence relation to determine the depreciating value of an asset after n depreciation periods, including from first principles for n ≤ 5
- use of the rules for the future value of an asset after n depreciation periods for flat rate, unit cost and reducing balance depreciation and their application.

Compound interest investments and loans, including:

- review of the concepts of simple and compound interest
- use of a recurrence relation to model and analyse (numerically and graphically) a compound interest investment or loan, including the use of a recurrence relation to determine the value of the compound interest loan or investment after n compounding periods, including from first principles for n ≤ 5
- difference between nominal and effective interest rates and the use of effective interest rates to compare investment returns and the cost of loans when interest is paid or charged, for example, daily, monthly, quarterly
- rule for the future value of a compound interest investment or loan after n compounding periods and its use to solve practical problems.

Reducing balance loans (compound interest loans with periodic repayments), including:

- use of a first-order linear recurrence relation to model and analyse (numerically and graphically) the amortisation of a reducing balance loan, including the use of a recurrence relation to determine the value of the loan or investment after n payments, including from first principles for n ≤ 5
- use of a table to investigate and analyse the amortisation of a reducing balance loan on a step-by-step basis, the payment made, the amount of interest paid, the reduction in the principal and the balance of the loan
- use of technology with financial modelling functionality to solve problems involving reducing balance loans, such as repaying a personal loan or a mortgage, including the impact of a change in interest rate on repayment amount, time to repay the loan, total interest paid and the total cost of the loan.

Annuities and perpetuities (compound interest investments with periodic payments made from the investment), including:

- use of a first-order linear recurrence relation to model and analyse (numerically and graphically) the amortisation of an annuity, including the use of a recurrence relation to determine the value of the annuity after n payments, including from first principles for n ≤ 5
- use of a table to investigate and analyse the amortisation of an annuity on a step-by-step basis, the payment made, the interest earned, the reduction in the principal and the balance of the annuity
- use of technology to solve problems involving annuities including determining the amount to be invested in an annuity to provide a regular income paid, for example, monthly, quarterly
- simple perpetuity as a special case of an annuity that lasts indefinitely.

Compound interest investment with periodic and equal additions to the principal (an annuity investment), including:

- use of a first-order linear recurrence relation to model and analyse (numerically and graphically) annuity investment, including the use of a recurrence relation to determine the value of the investment after n payments have been made, including from first principles for n ≤ 5
- use of a table to investigate and analyse the growth of an annuity investment on a step-by-step basis after each payment is made, the payment made, the interest earned and the balance of the investment
- use of technology with financial modelling functionality to solve problems involving annuity investments, including determining the future value of an investment after a number of compounding periods, the number of compounding periods for the investment to exceed a given value and the interest rate or payment amount needed for an investment to exceed a given value in a given time.