

Chapter notes: 21 Summarising data

Overview

Much of this topic has been moved into Prior learning, so you should make sure that the students are familiar with mean, median, mode, quartiles and percentiles. This chapter looks at calculating those statistics from grouped data, and the formula for standard deviation. The unbiased estimate of the population standard deviation has been moved into the Statistics option, but the Extension worksheet for this chapter investigates the idea of the distribution of the sample mean and sample variance. The chapter needs approximately four hours of teaching time if the concepts are unfamiliar.

Introductory problem

The introductory problem should encourage the understanding that what might be considered a ‘small’ variation depends on the spread of data. This should lead to the introduction of measures of spread. The worked solution is given at the end of the chapter, page 701; the idea being that students should be able to answer the question using the methods covered in the chapter.

21A Some important concepts in statistics, p689

In this section we discuss different types of data and the difference between a sample and a population. These concepts are important for applications of statistics, but they do not feature prominently in the Core IB course. Supplementary sheet 11 ‘Measuring risk’ looks at some ways in which statistics can be misleading. You could also use Supplementary sheet 20 ‘Misleading statistics’ from the Standard Level CD-ROM.

The ‘Research explorer’ box on page 690 refers to quantum energy levels. ‘Ultra-violet catastrophe’ is an incorrect prediction of classical statistical mechanics that radiation at short wavelengths contains arbitrarily large amounts of energy. This follows from the Rayleigh-Jeans law for energy emitted by a black body at various temperatures. The paradox was resolved by Max Planck’s theory that electromagnetic radiation can only be emitted in discrete packets.

21B Measures of spread, p691

The students should be familiar with range and inter-quartile range from Prior learning section Y. In this section we introduce the standard deviation. Both versions of the formula are given in the formula book, but in the form involving frequencies. The syllabus states that all data can be treated as populations, so s_n and σ can be used interchangeably.

The ‘rule of thumb’ given at the top of page 693 is a rough guide for normal-like distributions. Supplementary sheet 12 ‘Significant discoveries’ examines this idea further, but it requires calculations with normal distribution. It may be interesting to ask students to come up with data sets for which this rule doesn’t apply.

The two ‘From another perspective’ boxes on page 691 highlight the importance of using precise language in mathematics.

Hints for grade 7 questions:

6. Use the formula from Key point 21.2.
7. Use the fact that the variance is positive to write an inequality.
8. (b) Use the result from part (a).

21C Frequency tables and grouped data, p695

Finding the median and quartiles by interpolation is not required.

Hints for grade 7 questions:

6. Use formulae for the mean and variance to form two equations.