Mathematical Statistics

MAS 713

Tutorial about Chapter 1

Exercise 1

Answer by yes or no, and **explain**.

1. Will the sample mean always correspond to one of the observations of the sample?
2. Will exactly half of the observations in a sample always fall below the mean?
3. Will the sample mean always be the most frequently occurring data value in the sample?
4. Can the sample standard deviation be equal to zero?
5. Can the sample median be equal to the sample mean?

Exercise 2

Answer by yes or no, and **explain**.

1. Suppose that you add +10 to all of the observations in a sample.
   How does this change the sample mean?
   How does it change the sample standard deviation?
2. Suppose that you multiply all of the observations in a sample by 2.
   How does this change the sample mean?
   How does it change the sample standard deviation?
3. A sample of temperature measurements in a furnace yielded a sample average of 446°Celsius and a sample standard deviation of 5.8°Celsius. You would like to communicate this information to an American colleague.
   What are the sample average and the sample standard deviation expressed in °Fahrenheit?
   (*Hint :* temperature in °C = (temperature in °Fahrenheit - 32)$\times 5/9$)

Exercise 3

An experiment to investigate the survival time (in hours) of an electronic component consists of placing the parts in a test cell and running them for 100 hours under elevated temperature conditions (this is called an 'accelerated life test'). Eight components were tested with the following resulting failure times :

$$75, 63, 100^{+}, 36, 51, 45, 80, 90$$

The observation 100^{+} indicates that the unit still functioned at 100 hours.

Is there any meaningful measure of location that can be calculated for these data?
What is its numerical value?

Exercise 4

Consider a sample of observations $x_1, x_2, \ldots, x_n$.

- For what value a is the quantity $\frac{1}{n-1} \sum_{i=1}^{n} (x_i - a)^2$ minimised?
- Interpret in terms of location and dispersion parameters you know.

Exercise 5

The following data is a sample of shear strength, (MPa) of a joint bonded in a particular manner :

$$22.4, 40.4, 16.4, 73.7, 36.6, 109.9, 30.0, 4.4, 33.1, 66.7, 81.5$$

- Determine the 5-number summary.
- Determine the iqr. Are there any outliers (by the $1.5\times$ iqr rule)?
- Construct a box-plot and comment on its features.
- Determine the mean $\bar{x}$ and the sample standard deviation s.
- By how much could the largest observation be decreased without affecting the iqr?

Exercise 6

Direct evidence of Newton's universal law of gravitation was provided from a renowned experiment by Henry Cavendish (1731-1810). In the experiment, masses of objects were determined by weighting, and measured force of attraction was used to calculate the density of earth. The values of the earth's density estimated by Cavendish, expressed as a multiple of the density of water (1 g/cm^3), are :

5.50 5.30 5.47 5.10 5.29 5.65 5.55 5.61 5.75 5.63 5.27 5.44 5.57 5.36
4.88 5.86 5.34 5.39 5.34 5.53 5.29 4.07 5.85 5.46 5.42 5.79 5.62
5.58 5.26

(*Source :* Philosophical Transactions, 17 (1798), 469)

1. Find the sample mean, the sample standard deviation and the sample median of these data.
2. Determine the iqr. Are there any outliers (by the 1.5$\times$ iqr rule)?
3. Construct a box-plot and comment on its features.
4. Would you suggest the sample mean or the sample median as single estimate of the density of earth from Cavendish's data?

Exercise 7

A.A. Michelson (1852-1931) made many series of measurements of the speed of light. Using a revolving mirror techniques, he obtained

12 30 30 27 30 39 18 27 48 24 18

for the differences **[(velocity of light in air) - 299,700 km/s]**.
(*Source :* The Astrophysical Journal, 65 (1927), 11.)

1. Draw a dotplot.
2. Find the median and the mean. Locate both on the dotplot.
3. Find the variance and standard deviation.
4. Find the quartiles.
5. Find the minimum, maximum, range, and interquartile range.

Exercise 8

An experimental study of the atomisation characteristics of biodiesel fuel was aimed at reducing the pollution produced by diesel engines. Biodiesel fuel is recyclable and has low emission characteristics. One aspect of the study is the droplet size (μm) injected into the engine, at a fixed distance from the nozzle. Consider the following observed droplet size :

2.1 2.2 2.2 2.3 2.3 2.4 2.5 2.5 2.5 2.8 2.9 2.9 2.9 3.0 3.1 3.1 3.2 3.3
3.3 3.3 3.4 3.5 3.6 3.6 3.6 3.7 3.7 4.0 4.2 4.5 4.9 5.1 5.2 5.3 5.7 6.0 6.1
7.1 7.8 7.9 8.9

(*Source :* Kim et al (2008), Energy and Fuels, 22, 2091–2098.)

1. Obtain a frequency table
   using $[2,3)$, $[3,4)$, $[4,5)$, $[5,7)$, and $[7,9)$ as classes,
2. Construct a density histogram.
3. Obtain the sample mean $\bar{x}$ and the sample variance s^2.