1 | Conceptual Exercises

1. A researcher is analysing individuals' relative fear of being a victim of burglary on a 1-100 scale. A random sample of 9 individuals found a mean score of 47 on the scale with a sample variance of 158.76 for fear of being burgled.

   a. What distribution would be used to calculate an 80% confidence interval around this mean?
   A t-distribution as we don't know the population standard deviation and n is small.

   b. Construct that interval.

   $\bar{x}$ = 47
   n = 9
   t from tables = 1.397

   $s = \sqrt{158.76}$
   s = 12.6

   Confidence interval formula
   $\bar{x} \pm t \times \frac{s}{\sqrt{n}}$
   Lower bound
   $47 - 1.397 \times \frac{12.6}{\sqrt{9}} = 47 - 5.867 = 41.13$

   Upper bound
   $47 + 1.397 \times \frac{12.6}{\sqrt{9}} = 47 + 5.867 = 52.87$

2. We are investigating the height of men in the UK. For this we have obtained a random sample of 100 UK men and found they had a mean height of 180cm with a standard deviation of 10cm.
   a. Construct a 95% confidence interval for the mean height of UK males.

   $\bar{x}$ = 180
   s = 10
   n = 100

   As the population standard deviation is not known, the t distribution and t need to be used.

   Find the t-score for a 95% confidence interval in the t-table with 99 df.
   t = 1.984

   Confidence interval:
   $\bar{x} \pm t \times \frac{s}{\sqrt{n}}$

   Lower bound:
   $180 - 1.984 \times \frac{10}{\sqrt{100}}$ = 180 – 1.984 = 178.02
   Upper bound:
   $180 + 1.984 \times \frac{10}{\sqrt{100}}$ = 181.98

   b. Select all true statements concerning the constructed confidence interval and justify your choice for each statement.
      i. The probability of the population mean being within the upper and lower bounds is 95%.
         FALSE - The population mean is fixed but unknown and therefore can either be inside the bounds or outside. The Probability is therefore 50%.

      ii. 95% of men's heights fall between the upper and lower bound.
         FALSE - The distribution calculated is not the distribution of men's height, but the sampling distribution of the mean male height.

      iii. 95% of the cases in the sample fall between the upper and lower bound.
         FALSE - The distribution calculated is not of men's height in this sample, but the sampling distribution of the mean male height.

      iv. On average 95% of confidence intervals constructed would contain the population mean.
         TRUE

v. On average 95% of the means of samples with 100 respondents will fall within the upper and lower bands.
FALSE - This confidence interval is not making statements about various sample means but rather about the population mean.

vi. On average 95% of the sample means equal the population mean.
FALSE - The confidence interval is a range and does not make claims about where the population mean is exactly.

2 | Applied Exercises

See RScript in the Online Companion