ST201 Data Analysis

Continuous Assessment – Assignment Sheet 4

Instructions

- Use R to Answer the question below for continuous assessment. Due: Before **5pm on 3rd December 2021**. Upload a word document with the answers to Moodle. Upload your R code as a separate file.
- Include your name, student number and tutorial time with your work.

Question 1

Suppose a Normal Distribution has $\mu = 100$ and $\sigma = 10$.

(a) Plot the probability density function for this distribution. Use values in the range 50 to 150 for creating the plot.

(b) Calculate the area (i) below 90, (ii) above 105, (iii) between 80 and 107.

(c) Plot the cumulative probability distribution function. Use the same range of values that you used in part (a).

(d) Find the 2.5th and 97.5th percentile values. Provide an interpretation.

Question 2

Drunk driving is one of the main causes of car accidents. Interviews with drunk drivers who were involved in accidents and survived revealed that one of the main problems is that drivers do not realize that they are impaired, thinking "I only had 1-2 drinks. . . I am OK to drive."

A sample of 20 drivers was chosen, and their reaction times in an obstacle course were measured before and after drinking two beers. The purpose of this study was to check whether drivers are impaired after drinking two beers. You'll find the the dataset called `beers.rds` on the Rstudio server. The dataset contains two variables, `Before` (reaction times before the beers) and `After` (reaction times after the beers).

Load this dataset into R and answer the following questions.

(a) Create a new variable which is the difference between the `After` reaction times and the `Before` reaction times. What is the mean difference based on this sample?

(b) What is the 90% confidence interval for the population mean difference? Provide an interpretation.

(c) What conclusions can you draw from (b) about the mean difference in reaction times?