MIOT H6014

Statistical Analysis for Engineers

Assignment on Hypothesis Testing

Semester 2

Instructions

- Answer all questions in this assignment. The assignment refers to data stored in the Excel document 'Assignment1-2018.xlsx' and the worksheets contained there.
- For each question, outline clearly the assumptions made about the data and the test and conditions required for the test to be rigorous. In each case, identify the limitations on the result.
- Your answers may be saved as a PDF file and uploaded on the module Moodle page.
- You answers must be handed up in hard copy, before the 2nd September, either printed or handwritten. Please include a cover sheet, giving your student number and a signed statement that your answers are entirely your own work.

Question 1 – 20 Marks

A manufacturer of breakfast cereal makes two claims concerning the weight of the packets they produce. The manufacturer claims that

a) The mean is 200g
b) The variance is $0.8g^2$.

Answer the following questions.

1. To investigate the claims, the weight of a sample of packets produced in a given shift was measured. The values found are listed in part (a) of worksheet 'Dataset1', with grams (g) as the unit of measurement. Carry out appropriate statistical tests on claims (a) and (b).
2. Several weeks later, a smaller sample of packets as taken and their weight was measured; the data is listed in part (b) of worksheet 'Dataset1', with grams (g) as the unit of measurement. Use this data to construct confidence intervals for the mean and variance of the variable 'W', the weight of a packet of cereal.
3. Identify any issues with adding the second set of data to the first and 'updating' the results of part 1.

Question 2 – 20 Marks

A study was carried out to determine whether the resistance of a control circuit in a machine is lower when the machine motor is running. To investigate this question, some of the control circuits were tested as follows. Their resistance was measured while the machine motor was not running and then again while the motor was running for a certain period of time. The values found are listed in worksheet 'Dataset2', with kilo-Ohms as the unit of measurement. Answer the following questions.

1. Carry out an appropriate statistical test to determine whether the resistances have decreased.
2. Explain why the test was done with measurements when the motor was not running first and then measurements with the motor running, in other words, the order of the measurements.

3. Create a confidence interval for the difference.
4. Setting X as the resistance without and Y as the resistance with the motor running, set up the least squares linear regression equation

$$Y = a + bX$$

predicting Y from X. Carry out an appropriate test on the coefficients and on the correlation coefficients.

5. State what values of the coefficients of a linear equation should be consistent with the model in part (1).

Question 3 – 20 Marks

A study is being conducted to monitor how second-level students progress in education and how this is influenced by the economic position of their school. A randomly selected sample of such students who complete the Junior Cycle are identified; they are then categorised by their subsequent education outcome, Category A being little or no further education, to Category E the furthest level of education. The schools they attended are put in groups 1 to 3, with group 1 schools being the most economically disadvantaged, group 3 being the least disadvantaged. The numbers of pupils in each category of school with each type of educational outcome are given in the table in 'Dataset3' in the Excel document. Carry out an appropriate statistical test to determine whether the data supports a link between category of school and educational outcome.

Question 4 – 20 Marks

A manufacturer has tested the time taken to manufacture a fixed number of working components, measuring the times in units of 10^3 seconds. The

resulting data values are shown in Dataset 4 as a frequency distribution. Carry out an appropriate statistical test to determine if this data may be said to have a normal distribution. Identify any questions around identifying the number of degrees of freedom for the relevant test.

Question 5 – 30 Marks

A component manufacturer wishes to measure the quality of five different methods of production of components. An experiment to test this was conducted with a one-way layout, where the response being measured was the lifetime of the components, when tested to destruction, measured in units of 10^6 seconds. The data is shown in Dataset 5. Carry out an appropriate statistical test to see if any one of the five methods has produced better components than the others, as measured by their lifetime.

Question 6 – 30 Marks

A trial is conducted to measure the effect of five different fertilising methods and three different pots on the fruit yield of pepper trees. The weight of fruit produced for each of 120 pepper trees was measured, and the results grouped according to fertilising method and type of pot. The data is shown in Dataset 6. Carry out an appropriate statistical test to see if any one of the five methods or the three types of pot yield more fruit than the others.