

Advanced Analysis of Scientific Data (STAT1301)

Assignment 3

The due date/time is given on Blackboard.

1. Consider the observations: 8, 6, 6, 4, 5, 6, 11, 8, 9.
 - (a) Calculate the mean and the standard deviation of these observations.
 - (b) Calculate the median and the interquartile range.
 - (c) Check your results using R software
2. Formulate a statistical model for each of the situations below, in terms of one or more iid samples. If a model has more than one parameter, specify which parameter is of primary interest.
 - (a) A ship builder buys each week hundreds of tins of paint, labeled as containing 20 liters. The builder suspects that the tins contain, on average, less than 20 liters, and decides to determine the volume of paint in nine randomly chosen tins.
 - (b) An electronics company wishes to examine if the rate of productivity differs significantly between male and female employees involved in assembly work. The time of completion of a certain component is observed for 12 men and 12 women.
 - (c) The head of a mathematics department suspects that lecturers A and B differ significantly in the way they assess student work. To test this, 12 exams are both assessed by lecturer A and B.
 - (d) To investigate if the support for the *Honest* party is the same in two different cities, one hundred voters in each city are asked if they would vote for the *Honest* party or not.
3. In Example 6.5 of the Lecture Notes, prove, using the definition of a stochastic confidence interval, that if (a, b) is a 95% numerical confidence for μ , then $(a/50, b/50)$ is a 95% numerical confidence interval for $\mu/50$.
4. In 2000, a milk vendor found that 30% of his milk sales were of a low fat variety. Last week, of his 1500 milk sales, 500 were low fat. Is there any indication of a move towards low fat milk? Give the P-value associated with the test.
5. A study of iron deficiency among infants compared breast-fed with formula-fed babies. A sample of 25 breast-fed infants gave a mean blood haemoglobin level of 13.3 and a standard deviation of 1.4, while a sample of 21 formula-fed infants gave a mean and standard deviation of 12.4 and 2.0 respectively. From previous experience, haemoglobin levels can be assumed to follow a normal distribution. Assuming equal variances, carry out a statistical test to compare the two groups.

6. Brand A batteries cost more than brand B batteries. Their life lengths are normally distributed. A sample of 16 batteries of brand A gave a mean life length of 4.6 (hours) and a standard deviation of 1.1 (hours), while a sample of 26 batteries of brand B gave a mean and standard deviation of 4.2 and 0.9 respectively. Test whether brand A is better than brand B.