MIOT H5014

Statistical Analysis for Engineers
Worksheet 4 on Hypothesis Testing

Semester 2, 2016/7

Question 1

It is to be determined whether a coin is a fair coin, that is to say, whether tossing the coin gives one side or the other with equal probability of 0.5.

1. Before gathering any data, frame the Hypotheses.
2. It has been decided that the test will be carried out with just ten throws of the coin. Identify an appropriate statistic to carry out a two-tailed test on the Null Hypothesis. Assuming the value of the level of significance α is 0.05 or less, identify the critical values of the relevant statistic.
3. If $\alpha < 0.05$, identify the minimal number of throws possible to make a meaningful test.
4. The number of throws is now to be a large number above 10^3. Identify a suitable test statistic and write an equation for it.
5. State why redoing this procedure for a large number of throws as a goodness of fit test would be essentially the same test.

Question 2

It is to be determined whether a dice is fair, that is to say, whether rolling the dice gives any one of the six sides with equal probability.

1. Before gathering any data, frame the Hypotheses.

2. The number of throws is to be a large number $N > 100$. Identify a suitable test statistic Q for the goodness-of-fit test and write an equation for it. Simplify this equation under the Null hypothesis to show that the statistic is

$$Q = \frac{6 \sum_i O_i^2}{N} - N,$$

where the O_i are the observed values.

3. It has been decided that the test will be carried out with just 50 rolls of the dice. Identify an appropriate statistic to carry out a two-tailed test on the Null Hypothesis.

Question 3

A computer program has been written to reproduce the random variable S of the sum of two fair dice.

1. Before gathering any data, set up the Null and alternative Hypothesis that the program achieves its aim.

2. The number of throws is to be a large number above 100. Identify a suitable test statistic Q for the goodness-of-fit test.

3. Assume your code has a function randomU() which produces a uniform distribution between 0 and 1. Write a piece of code or pseudo-code that would generate suitable values for S.

Question 4

A horticulturist is planting seeds, eight at a time in trays. It is to be assumed that the seeds in a tray will germinate independently and that they are all identically fed and watered.

1. Let G be the random variable of the number of seeds in a tray that germinate. Identify what distribution G should follow.

2. A sample is taken of 100 trays after two weeks and the number of seeds that have germinated are counted for each tray. State how a frequency distribution would be created from this data.

3. Identify a suitable test statistic Q for the goodness-of-fit test. Explain how to calculate a sample value of the parameter(s) for the distribution and so calculate the expected values for the test statistic

4. Explain what the observed values will be.