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.

Solution

Let p be the probability that a 'head' comes up.

Under the Null Hypothesis, H_0, $p = 0.5$, the definition of a fair coin.

The alternative is p is not 0.5.

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.

Solution

This is the binomial distribution $N = \sum_{i=1}^{n} B_i$, $P[N = r] = {}^nC_r p^r(1 - p)^{n-r}$.

Under H_0, $p = 0.5$, so $P[N = r] = {}^nC_r\, 0.5^r(1 - 0.5)^{n-r} = {}^nC_r/2^n$.

With $n = 10$, this is $P[N = r] = {}^{10}C_r/1024$.

We will do a two tailed test.

Looking at $r = 10$ or $r = 0$, which are the same and then $r = 9$ or $r = 1$, which are also the same, the probabilities are

2 x 1/1024 + 2 x 10/1024 = 2/1024 + 20/1024 = 22/1024 = 0.0215.

Bringing in $r = 2$ or 8 will clearly bring the probabilities above 0.025, so the critical values for r are 1 or 9.

3. If $\alpha < 0.05$, identify the minimal number of throws possible to make a meaningful test.

Solution

With $P[N = r] = {}^nC_r/2^n$, as the test statistic:

The extreme numbers come at $r = 0$ and $r = n$, when the value of ${}^nC_r = 1$.

The two-tailed test will not be about useful when these numbers are above 0.05.

So this happens when $2/2^n > 0.05$, so $2^{n-1} < 20$, so $2^{n-3} < 5$, $n = 5$.

The test is not useful at this point.

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.

Solution

Each of the throws of the coin is a Bernoulli trial, B.

Then for the variable B: $\mu = p$, $\sigma^2 = p(1 - p)$, where the true proportion is p.

Look at the mean number for variable B : $Z = \dfrac{\overline{B} - \mu}{\sigma/\sqrt{n}} = \dfrac{\overline{B} - p}{\sqrt{p(1-p)/n}}$.

The mean of B is an estimate of the proportion of heads.

So if our estimate is $\hat{p}$, then $Z = \dfrac{\hat{p} - p}{\sqrt{p(1-p)/n}} \sim \mathrm{N}(0,1)$.

Under the Null hypothesis $p = 0.5$, so the test statistic is $2(\hat{p} - 0.5)\sqrt{n}$.

5. State why redoing this procedure for a large number of throws as a goodness of fit test would be essentially the same test.

If the test is recast as a goodness of fit test, it would use a contingency table of size 2 by 2, thus there would be only 1 degree of freedom. The statistic would then essentially be a standard normal distribution.

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.

Solution

The Null Hypothesis is that all sides come up equally, so that the problem is modelled by a uniformly distributed discrete random variable going from 1 to 6. The alternative is that it is not. Take the level of significance as 0.05.

2. The number of throws is to be a large number 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.

Solution

For a test on a discrete uniform variable going from 0 to k, let N be the number of throws and let O_i be the number of times value i comes up.

With the uniform distribution, the expected value E_i for every i is N/k, which we can call 'O-bar' as this will be the average number of observed values.

Then the statistic is $Q = \sum_i \frac{(O_i - E_i)^2}{E_i}$.

Do a bit of algebra on this with all the E_i set as 'O-bar':

$$Q = \sum_i \frac{(O_i - \overline{O})^2}{\overline{O}} = \frac{1}{\overline{O}} \sum_i (O_i - \overline{O})^2 .$$

This is the familiar expression for the top line of a variance calculation, so:

$$Q = \frac{1}{\overline{O}} \left(\sum_i O_i^2 - k\overline{O}^2 \right) = \frac{1}{\overline{O}} \sum_i O_i^2 - k\overline{O} = \frac{k}{N} \sum_i O_i^2 - N .$$

Our case is $k = 6$, as required.

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.

Solution

This would be done with the multinomial distribution.

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.

Solution

Let X be the random variable produced by this computer program. The Null hypothesis is that X has the distribution for the sum of two numbers 1 to 6, which is:

$P[X = 2] = 1/36$, $P[X = 3] = 2/36$, $P[X = 4] = 3/36$, …

$P[X = 7] = 6/36$, $P[X = 8] = 5/36$, … $P[X = 12] = 1/36$,

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.

Solution

The statistic is $Q = \sum_i \frac{(O_i - E_i)^2}{E_i}$.

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.

Solution

Either use $F_X^{-1}(\text{random}U)$ where F_X is the CDF outlined above or generate two values for the uniform discrete distribution 1 to 6 and then add them!

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.

Solution

This will be the binomial distribution with $n = 8$ and p unknown.

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.

Solution

For the numbers 0 to 8, count how many times a tray produces each number of germinated seeds. This is the set of observed values.

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 Q.

Solution

The statistic is $Q = \sum_i \frac{(O_i - E_i)^2}{E_i}$.

The available estimate of unknown parameter p is the total number of germinated plants divided by the total originally planted, in other words 8 times the number of trays. The expected values are calculated with the binomial distribution. The degrees of freedom for the test is $9 - 1 - 1 = 7$.