Informatics 1B. Tutorial 5
Data and Analysis: Unstructured Data

Manuel Marques Pita and Helen Pain
Issued by Gaya Nadarajan

For Week 9 beginning March 5, 2007

1 Evaluation

The following hypothesis is to be tested experimentally:

Two interfaces are provided to a problem solving tool: interface X and interface Y. People using interface X solve problems in less time than those using interface Y.

The two alternative designs given below are proposed for testing this hypothesis.

Design 1: Twenty undergraduate students are selected at random in the students union. They are pre-tested on a problem solving task, using pen and paper, and divided into two matched groups based on their performance. One group uses interface X and then interface Y. The other group uses interface Y and then X. The number of errors made in completing a problem solving task (similar to the pre-test task), using each interface, is recorded and compared.

Design 2: Twenty undergraduate AI students are recruited. All the students use interface X and interface Y. Half use first X then Y, and the other half use first Y then X. Two problem solving tasks are used, and assigned randomly for use with each interface (such that each used completes both task). Time taken to complete the problem solving tasks using each interface is recorded and compared.

1. Compare the two designs, stating the strengths and limitations of each in relation to the hypothesis to be tested.
2. Propose a new design that incorporates the best features of each.

2 Correlation

A Russian herbal remedy (coming from the root of a plant similar to the famous *ginseng* plant) called *Rhodiola Rosea* is being tested as a natural agent which might enable people to keep productivity levels high under stressful conditions.

Different quantities of this root extract were given to a number of people on different days. People taking part in this experiment were all under stressful conditions, being required to produce good-quality work with not enough time for them to do so.

What follows are two linked lists relating the amount of Rhodiola Rosea given to subjects and the amount of effective work hours they did on each corresponding day.

```
rhodiola_amounts_mgr    = [50  80  60  100  70  120  130  90  110  120];
effective_hours_worked  = [6  6.5  6.1  6.6  6  7  7.2  6.1  6.8  6.8];
```

Key. In this table, for day 1, 50mg of Rhodiola were given to subjects and the average number of effective hours was 6.

We want to know if the amount of Rhodiola Rosea given has a significant correlation with the amount of effective hours worked. If there is a correlation, this might suggest that Rhodiola Rosea supports people in working more effective hours under stress.

In order to test this correlation, you will calculate the *Pearson Coefficient, r,* given by the formula

$$r = \frac{\sum_i (X_i - \overline{X})(Y_i - \overline{Y})}{N \sigma_X \sigma_Y}$$

Where X_i and Y_i correspond to each of the values in the first and second list respectively. $\overline{X}$ and $\overline{Y}$ are the *means* of each set of values. σ_X and σ_Y are the *standard deviation* of each set.

In the same way as we did in previous tutorials with the χ^2 test, we will have a *null hypothesis,* H_0 which in this case would be "The amount of Rhodiola Rosea given to subjects does not correlate with the amount of effective hours they work". Next we calculate the value of the test, in this case the correlation r. If the value is greater than the corresponding critical value (taken from a statistics table) we then reject H_0. This would mean that there is a correlation.

For this example, you will want to know that the critical value p for $N - 2 = 8$ degrees of freedom (df) is $p = .765$ with 99.99% confidence.

(a) Is the null hypothesis accepted or rejected in this case?

(b) Suppose the value of r was -0.9227? How would you interpret this?

(c) If we had, for example, 102 data entries and the value of r was .38, given that the critical value p (for 100 df) is .254, How would you interpret this result?