

# Inf1B Data and Analysis

## Tutorial 8 (week 10)

11 March 2010

- Please answer all questions on this worksheet in advance of the tutorial, and bring with you all work. Tutorials cannot function properly unless you do the work in advance.
- Data & Analysis tutorial exercises are not assessed, but are a compulsory and important part of the course. If you do not do the exercises then you are unlikely to pass the exam.
- Attendance at tutorials is obligatory; please let your tutor know if you cannot attend.
- *Background Reading:* Lecture slides on Statistical Analysis of Data III.2-III.4

### Introduction

In this tutorial we will perform statistical analysis over data on students' nationality, physical exercise and alcohol intake. The data for this tutorial was collected in the first DA lecture on 12th January using an anonymous questionnaire asking students for their: hours of physical exercise last week; units of alcohol intake last week; whether of UK, Irish or Other nationality.

### Question 1: Statistical analysis of numerical data

Refer to the on-line data file `data.pdf` on the Data & Analysis course homepage.

- (a) Extract a sample of of 10-20 students from this data.
- (b) Based on your sample, estimate the mean, and standard deviations for both weekly alcohol consumption and weekly exercise hours.

- (c) Based on your sample, estimate the Pearson correlation coefficient to compare weekly alcohol consumption and weekly exercise hours. Is there a significant correlation? Is it positive or negative?

## Question 2: Correlation between categorical data

The following are some statistics from the the file `data.pdf`.

- No. UK/Irish students who drank some alcohol: 52
- No. UK/Irish students who drank no alcohol: 14
- No. foreign students who drank some alcohol: 51
- No. foreign students who drank no alcohol: 26
  
- No. UK/Irish students who drank at least 10 units alcohol: 38
- No. UK/Irish students who drank less than 10 units alcohol: 28
- No. foreign students who drank at least 10 units alcohol: 22
- No. foreign students who drank less than 10 units alcohol alcohol: 55

Answer the following questions.

- (a) Compile the relevant contingency tables to investigate correlation between: (a) nationality and drinking some alcohol; (b) nationality and drinking at least 10 units of alcohol.
- (b) Calculate the corresponding tables of expected frequencies.
- (c) Calculate the corresponding  $\chi^2$  values.
- (d) Are there correlations? What are the significance levels?
- (e) Estimate the mean alcohol intake of UK/Irish students and the mean alcohol intake of other students.
- (f) Which information do you find more informative: the answer to question (d) or the answer to question (e)?