

## **SAP for Can you taste the difference?**

(this is one possible approach, there are others, and depends on how you designed the experiment)

### **Population**

- Statistics students (this might be more restrictive by specifying in Glasgow)

### **Primary Objective:**

- To test if Statistics students can tell the difference between caffeinated and de-caffeinated coffee

### **Secondary Objectives:**

- None additionally stated, but we could imagine differences between males and females, or between those who regularly drink coffee and those who do not

### **Data Collection:**

- Random sample of students could describe how such a random sample would be generated, and also some justification or discussion of how many students might be needed.
- Decide how many trials each student should have, e.g. 8 coffee samples (4 DC and 4C), so there is replication and would also need to discuss the order that the samples are presented in (random? Or DC followed by C)
- Definition of response, C, DC, a third possibility would be don't know

### **Variables Under Consideration:**

- Number of correct answers in total for each individual (perhaps also the number of DC and number of C correct – **Primary outcome variable**)
- Could be sex (M/F), could be whether the individual drinks coffee (Y/N) – both are binary **Primary explanatory variable**

### **Missing Data Procedures:**

None described and unlikely to be an issue

### **Summaries to be presented:**

- Table of Counts and Percentages of number of correct answers for each individual
- For the experiment, as a whole could show a frequency histogram of the number of individuals getting 0,1, 2, ..,8 correct answers
- Cross-tabulations of true answer (DC, C) against individual's answer (DC, C)

### **Models to be fitted**

- The natural model to consider here is whether or not someone is guessing. If guessing, then the model for the number of correct answers is

$\text{Bi}(n, 0.5)$ , so a Binomial model could be fitted and the probability of giving a correct answer estimated

- Under the guessing hypothesis we can then evaluate the probability of getting  $x$  answers correct
- For the cross-tabulation, we could carry out a Chi-squared test (this is formally called a McNemar test for this case)