

ST201 Data Analysis

Continuous Assessment – Assignment Sheet 2

Instructions

- Answer the question below for continuous assessment. Due: Before **5pm on 5th November 2021**. Upload a Word Document or PDF to Moodle.
- Include your name, student number and tutorial time with your work.

Question

Patients that have a certain disease in a particular location tend to go to either one of two hospitals (h1 or h2) for treatment. The outcome of the treatment given in the hospitals is characterised as a “success” or “failure”. The disease itself can be characterized as being less severe or more severe (less, more). You will find the file `hospital_dat.rds` in the ST201 folder on the Rstudio server containing data on these variables for a sample of 200 patients. Open a new R script and load the data into R.

For this assignment you should analyse these data to look at the relationships between the variables `hospital`, `outcome` and `severity`.

- The first step in the analysis should be to consider the relationship between the `hospital` and `outcome` variables only.
- The next step should be to consider the impact of adding the `severity` variable to the analysis.
- You should address the question of whether Simpson’s paradox is observed here based the steps above.
- Produce a 1-page (max) report of your findings. **Your report should be presented in a word or PDF document.**
- You should upload your Rcode as a separate file or as an appendix.

To achieve full marks the 1-page report should be organised as follows:

Aim: Briefly describe the aim of your analysis.

Data: Introduce the data. Provide a brief description of the data and the variables you intend to analyse.

Analysis: This should include points 1-4 below.

1. 2-way (hospital and outcome) and 3-way (hospital, outcome and severity) contingency tables. Discuss the chance of a failure outcome in each hospital in the 2-way and 3-way tables.
2. A Chi-squared test for testing dependence between hospital choice and outcome with interpretation.
3. Odds ratios of a failure outcome in H1 compared to H2 for the 2-way (hospital and outcome) and 3-way relationships (hospital, outcome and severity) with interpretation.
4. A comment on Simpson’s paradox.

Conclusions: Briefly mention what you have concluded from your analysis.