



THE UNIVERSITY OF  
**NEWCASTLE**  
AUSTRALIA

**FACULTY OF SCIENCE AND INFORMATION  
TECHNOLOGY**

School of Mathematical and Physical Sciences

*Student Enquiries*  
Telephone: (02) 4921-5000

## **STAT1070 Statistics for the Sciences Assignment 3**

**Semester 2, 2013**

---

**Due:**

- This assignment must be submitted electronically using Turnitin (link from Blackboard) by 10am Friday 11<sup>th</sup> October. **Detailed instructions for electronic submission are provided on Blackboard.**

**Weighting:** This assignment has a maximum mark of 22. It comprises 2 questions and contributes 9% to your final assessment mark.

**Instructions and General Marking Criteria:**

- You **MUST** use **the appropriate cover page, provided on the STAT1070 BB site under Assessment.**
  - You **MUST** also have completed the online Academic Integrity Module (AIM) and indicated this on your cover page. Please note that this is an **INDIVIDUAL** assignment, not a group assignment. **Inappropriate collaboration will be penalized.**
  - You **MUST** clearly indicate your **laboratory class (day, time and room) and your tutor's name.** We need this information to process assignments efficiently. Students who do not provide the cover sheet with the correct information will have marks deducted.
  - Assignments may be handwritten (and then scanned) but they must be neat, legible and appropriately formatted.
  - The file you submit must be appropriately formatted to be clearly legible when printed on A4 paper.
  - Any data files will be available via the link provided on the STAT1070 BB site.
  - For some questions we expect that students will use JMP to assist with calculations and preparation of appropriate graphs but raw computer output without explanatory text is not acceptable. Answers must be written in clear English sentences clearly linked to all **APPROPRIATE** working and/or supporting computer output. (Appropriate means you need to use some judgement. Extract the parts of the JMP output that are relevant to answering the question. Don't include the bits that are not.)
-

**Question 1 (Adapted from Utts & Heckard, 2012)****[Total: 14 marks]**

In the early days of compulsory seatbelt use the percentage of people wearing seatbelts was lower than it is now. As part of a Youth Risk Behaviour Surveillance System, a sample of students was asked how often they wear a seatbelt.

The following contingency table summarizes the results of the survey:

<b>Gender</b>	<b>Seatbelt use</b>		Row
	Most times/always	Rarely/never	Totals
Female	964	97	1061
Male	928	154	1082
Column Totals	1892	251	2143

(a) Calculate, using the appropriate formula, a 95% confidence interval for the proportion of female students rarely/never wearing a seatbelt. [2 marks]

(b) Calculate the row percentages for female and male students. Compare them.

[1 mark]

Interest is in whether there is a difference between males and females, in terms of seatbelt use, in the student population this sample was taken from.

(c) Specify the null and alternative hypotheses that could be tested using a chi-squared statistic in this situation. [1 mark]

(d) Calculate, using the appropriate formula, the expected count and chi-square contribution for the cell “Male, Most times/always”. [1 mark]

(e) Determine the appropriate test statistic and p-value. [2 marks]

(f) Provide a conclusion for your analysis. [2 marks]

(g) Use JMP to provide a 99% confidence interval for the difference between females and males in proportions of students who most times/always wear seatbelts.

[2 marks]

(h) Based on the p-value, at what level of confidence would you expect a confidence interval to be on the verge of excluding zero as a plausible difference between the two proportions? [1 mark]

- (i) Another way of looking at the difference in seatbelt use between female and male students is assessing an odds ratio. Use JMP to calculate a 99% confidence interval for an appropriate odds ratio and interpret. [2 marks]

**Question 2 (Adapted from Moore et al. 2012)**

**[Total: 8 marks]**

Recall the study of macaque monkeys from Assignment 1. Researchers had theorised that neurons would respond more strongly to monkey calls than to pure tones. Exploratory data analysis (as conducted in Assignment 1) suggests that this is true.

- (a) Now analyse the data in more detail by conducting an appropriate hypothesis test about the mean difference in responses to test this theory. The relevant column has been added in the dataset **monkeycalls3.jmp**. [5 marks]
- (b) Often the more relevant analysis is about estimating the size of the effect. Calculate a 90% confidence interval for the population mean difference and summarise your findings. [3 marks]

**Additional Marking Criteria for Questions 1 and 2:** You may calculate by hand using the appropriate formula (show working) or use JMP (provide copy of relevant output), except when the question specifically directs the method to be used. Marks will be awarded based on the quality of your assessment of the data and how clearly that assessment is communicated.

**End of Assignment 2**