

|  |
| --- |
| MXB344 Modelling Non-Normal Data with Generalised Linear Models **Unit Guide**  Semester 2 , 2016 |

**Unit details**

**Unit Title: Modelling Non-Normal Data with Generalised Linear Models**

**Credit points: 12**

**Semester offered: 2**

**Contact hours: 4 per week**

**Prerequisite(s)/ Corequisite(s): MXB242 or MAB414**

**Teaching team**

**Unit Coordinator: Dr James McGree**

**Teaching team: Miles McBain**

### Contact procedures

#### Consultation procedures

Staff from the teaching team will be available for consultation each *(insert day)* between *(insert time)* in *(insert venue)* .

#### Other enquiries and assistance

For more general inquiries please contact the Student Services Help Desk

**Phone (07) 3138 8822**

**Email:** [sef.enquiry@qut.edu.au](mailto:sef.enquiry@qut.edu.au)

**Visit**: Level 3, O Podium, Gardens Point Campus

**Unit aims, objectives and goals**

This unit aims to introduce modern statistical methods of data analysis to prepare you for a career in industry or government. For data that arise in, for example, science and commerce, it is often unreasonable to assume they are continuous random variables from a normal distribution. It is likewise unlikely that data are handed to an analyst in a state ready for advanced statistical techniques. This unit introduces modelling techniques and methodology for the explanation of non-normal data. It also introduces, by way of a realistic project, techniques to overcome common issues with shaping data for analysis. This will ensure appropriate statistical practice when such data are encountered.

**Learning outcomes**

After successfully completing this unit you will be able to:

On successful completion of this unit you should be able to:  
1. Expertly and critically carry out statistical analysis using statistical models in the analysis of various data sets and examples.  
2. Use R to carry out statistical analyses.  
3. Communicate statistical conclusions clearly and concisely both in written form and orally.

**Timetable and weekly unit content**

**Times and venues** *(If necessary, adapt the table)*

|  |  |  |  |
| --- | --- | --- | --- |
| **Activity** | **Time** | **Cohort** | **Room** |
| Lecture | Thursday 0900-1100 | All | GP-M303 |
| Workshop | Thursday  1100-1300 | All | GP-S637 |

**Weekly unit content** *(provide details of lectures and workshops topics and Assessment due dates)*

|  |  |  |  |
| --- | --- | --- | --- |
| **Wk** | **Date** | **Lecture** | **Workshop** |
| **1** | **28/7** | Course Admin   * WIL concept * Assessment * Course overview   + Workshop vs Lecture content * Introduction to Reproducibility * Introduction to Generalised-Linear Models   + lm vs glm   + Link function   + Distribution of data   + Linear Predictor   + Analysis of Variance vs Deviance | None.  Homework:  Create a Github account and obtain the educational discount.  Have a look at: <https://www.lynda.com/GitHub-tutorials/GitHub-Web-Designers/162276-2.html>  Login via: <https://ezp01.library.qut.edu.au/login/lynda> |
| **2** | **4/8** | Exponential Family   * Poisson, Binomial * Canonical Link * Dispersion   + Approximate estimate   Deviance   * Different for each exponential family * Test of residual deviance   Residuals   * Deviance * Pearson   + Test | Rmarkdown, Github, %>%, |
| **3** | **11/8** | Model Fitting Process   * Estimation   + MLE   + IWLS * Nested Models * AIC * Step   Possion Regression   * Count data * Rate data * Offsets/Exposure * Interpretation * Pitfalls   + Dispersion   + Too many zeroes * Negative Binomial * Pearson residual | Poisson GLM  **Assessment 1 Released** |
| **4** | **18/8** | Binomial/Logistic Regression   * Binary Data * Proportion Data * Link functions   + Logit   + Probit   + cloglog * Interpretations * Goodness of fit measures * Performance | Binomial GLM |
| **5** | **25/8** | Data Wrangling for Predictive Modelling   * Joining Tables * Munging Data * Variable selection * VIF * Feature Engineering   Missing Data   * Types * Imputation methods * Creating indicators | Dplyr, Missing Data  **Assessment 1 due 26/8** |
| **6** | **1/9** | Introduction to Credit Risk Modelling as Quantitative Analyst: Burton Wu Bank of Queensland | **WIL Assessment Released** |
| **7** | **8/9** | Comparing Predictive Models   * Cross-Validation * Gini coefficient * ROC * AUC * Psuedo R^2 * Diagnostics | Caret, ROCR  Review Assignment Plan |
| **8** | **15/9** | Poisson/Binomial Extensions:  Poisson Log Linear Models  Multinomial Regression | Review Assignment Plan |
| **9** | **22/9** | Special Industry Topics in Data Science | **Interim Report Due 23/9**  Review Assignment Plan |
|  |  | **Mid-semester break** | |
| **10** | **6/10** | Special Industry Topics in Data Science | Review Assignment Plan |
| **11** | **13/10** |  | Review Assignment Plan |
| **12** | **20/10** | **Presentation Due** | Presentation debriefs. |
| **13** | **27/10** |  | Review Assignment Plan |
| **14** | **3/11** |  | **Final Report Due4/11** |
| **Exam period starts** | | | |

## Assessment information

## Assessment summary

|  |  |  |  |
| --- | --- | --- | --- |
| **Item no.** | **Description** | **Value** | **Due date** |
| 1 | Assessment 1 (Theory) | 40% | 5pm, 26/8/2016 |
| 2 | WIL Project Presentation | 20% | TBA, 20/10/2016 |
| 3 | WIL Project Report | 40% | 5pm,  4/11/2016 |

## Assessment supplementary information

**Extensions**

Requests for extensions need to be submitted before the assessment item due date to SEF Student Services.  If you can't apply by this date due to circumstances beyond your control, contact  [SEF Student Services](https://www.student.qut.edu.au/about/contact/groups/faculty-student-offices)  to discuss your options. You need to provide supporting documentation to verify your special circumstances claim and to show how the circumstances impact your ability to submit the assignment by the due date. Examples of special circumstances that warrant an extension, and those that do not, can be found on Student Gateway: <https://www.student.qut.edu.au/studying/special-circumstances>

**Late submission of assessment item**

If you submit an assessment item after the due date without an approved extension (or after the extended date where an extension has been granted) your work will not be marked and will be awarded a grade of 1, or 0%.

If special circumstances prevent you from meeting the assessment due date, you can apply for an extension (see above). If you don’t have an approved extension you should submit the work you have completed by the due date and it will be marked against the assessment criteria.

QUT’s assessment submission requirements reflect the expectations of professional practice where you will need to meet deadlines. Further information is available on Student Gateway <https://www.student.qut.edu.au/studying/assessment/late-assignments-and-extensions>

**Remarking of your assessment item(s)**

QUT has mechanisms in place to ensure that all assessment pieces are marked consistently and fairly. During the semester or teaching period you should discuss your progress in all course work with teaching staff. You can expect a clear indication of whether you have achieved the objectives set for each assessment item. Accordingly, no individual piece of assessment will be remarked during the semester.

At the end of semester, please check your overall mark and if you are not satisfied with your final grade, you can contact relevant teaching staff to clarify the reason for your grade and if you remain dissatisfied after discussion with teaching staff, you can apply for a formal review of grade within ten working days. Further information is available on Student Gateway: <https://www.student.qut.edu.au/studying/assessment/reviews-and-appeals/review-of-grade-or-academic-ruling>

# *Provide criteria and standards rubrics for all Assessment Items*