

Probability and Statistical Inference
Continuous Assessment Part II
Semester I 2018/2019
Due Date: Sunday 2nd December 2018 @ 23.59

OVERVIEW

For this part of the assignment you are required to conduct and present appropriate statistical tests *including at least one multivariate inferential statistical technique covered in this module*. You can use either the same variables you used for part I of the CA or new variables you have chosen for the purposes of this part of the CA. You can adapt the research question you used in part I or chose a new research question. You can choose one of the following options:

Option A:

- Using either Multiple Linear Regression or Logistic Regression:
- Develop hypotheses testable by building a regression model. These should be related to your overall research question.
- Derive and present appropriate statistical evidence.
- Build a baseline regression model.
 - Assess its fit and usefulness.
 - Illustrate your findings using appropriate examples from your data.
- Build at least one additional model which extends/amends this baseline either adding or removing predictors relevant to your hypotheses.
 - Assess its fit and usefulness.
 - Illustrate your findings using appropriate examples from your data.
 - Your regression model should include at least one nominal predictor.
- Compare the fit and usefulness of your successive models.

Option B:

- Using a Dimension Reduction technique followed by either Multiple Linear Regression or Logistic Regression:
- Develop hypotheses testable by dimension reduction followed by regression. These should be related to your overall research question.
- Derive and present appropriate statistical evidence.
- Assess the suitability of the dataset for dimension reduction.
- Conduct your dimension reduction.
 - Assess the effectiveness of the dimension reduction.
- If the dimension reduction succeeds, use the outcome of the dimension reduction as part of multiple linear regression or logistic regression model (must include multiple predictors).
- If the dimension reduction does not succeed, identify an alternate mechanism to derive a measure for concept for which you conducted dimension reduction. Use this as part of a multiple linear regression or logistic regression model (must include multiple predictors).
- Your regression model should include at least one nominal predictor.
- Assess the fit and usefulness of the regression model.
 - Illustrate your findings using appropriate examples from your data.

This part of the assignment is worth 50% of the CA for module as marked out of 100% (25% of CA for the module).

Probability and Statistical Inference
Continuous Assessment Part II
Semester I 2018/2019
Due Date: Sunday 2nd December 2018 @ 23.59

DESCRIPTION

You are expected to :

- Present a summary of the variables used, critically discussing relevant issues which impact statistical analysis;
 - Include statistical summaries of the variables of interest and evidence to support relationships or difference to justify their inclusion in a dimension reduction or regression model.
- Use appropriate statistical techniques to achieve either option a or option b above;
- Present and interpret the findings;
- Briefly draw conclusions discussing your findings in terms of other related work and any implications for future work;
- Adopt the APA guidelines for reporting statistical analysis using APA citation and referencing.
- You must use R to conduct your analysis;
- You should cite appropriate sources (which are accessible) in order to support the guidelines you adopt in your decision making and interpretation of findings.

Your will need to demonstrate:

- An ability to generate and correctly state a hypothesis or hypotheses that is/are theoretically-informed;
 - The ability to correctly prepare, present, analyse and critically assess the dataset used from the perspective of statistical analysis;
 - The ability to correctly execute, present and interpret appropriate statistical tests using statistical software;
 - The ability to analyse and present the findings gained from your statistical analysis in a clear and accurate way to a standard expected of masters/PhD level academic work;
 - The ability to construct a report on a statistical inquiry.
-

Probability and Statistical Inference
Continuous Assessment Part II
Semester I 2018/2019
Due Date: Sunday 2nd December 2018 @ 23.59

DELIVERABLES

- You need to submit an R markdown file plus the HTML/PDF created from this.
- You must include the following information at the start of your RMD file:
 - Student Number: <<your student number>>
 - Student Name: <<your name>>
 - Programme Code: <<programme code>>
 - Option Chosen: <<option a/option b>>
 - **PhD students: <<dataset used>>**
 - The version of R used.
 - The R packages needed for your code to execute successfully.
- State clearly the hypotheses you intend to test.
- You must describe your variables.
 - In terms of their statistical measurement types and describe them with appropriate descriptive statistics and graphs.
 - You must address all issues which could impact on the choices when building a model.
 - You must present statistical evidence to support inclusion of variables as predictors in any model/use in a dimension reduction.
- You must build, present and illustrate your model as outlined in the overview.
 - Justify your choices based on your assessment of the dataset.
 - Illustrate how your model works using appropriate data.
- You must present and interpret your findings in paragraphs using APA style for reporting statistical results.
- Interpret your findings appropriately relevant to your hypotheses.
- A useful guide to creating a report of a statistical inquiry using APA guidelines is available at <http://www.discoveringstatistics.com/docs/writinglabreports.pdf>.

SUBMISSION

All required documents should be submitted using the **CA Part II Assignment** box in the Webcourses module.

- A HTML or PDF file that contains your report for part II. The format is at your discretion but you should name them <<Student #>>PSICAPart2.rmd and <<Student #>>PSICAPart2.html or <<Student #>>PSICAPart2.pdf where <<Student #>> is your student number e.g. D1111111PSICAPart2.rmd, D1111111PSICAPart2.html or D1111111PSICAPart1.pdf
 - A HTML or PDF file that includes all the R commands you used to conduct all the tests included in the report and outputs of these commands. This should be named <<Student #>>PSICAPart2RCommands.html or <<Student #>>PSICAPart2RCommands.pdf where <<Student #>> is your student number.
 - In order to overcome the issues with Webcourses and submission of HTML you can submit a single compressed file (.zip, .rar) which includes both of the previously described files. This should be named <<Student #>>PSICAPart2RCommands.zip (.rar is also acceptable).
-

Probability and Statistical Inference
Continuous Assessment Part II
Semester I 2018/2019
Due Date: Sunday 2nd December 2018 @ 23.59

NOTES

1. Unfair practice is a very serious offence in the DIT and you must acknowledge any material used by including a referenced bibliography in your report. Any issues will be investigated and those considered serious will be handled via the DIT Plagiarism policy (details are available in the General Assessment Regulations).
2. Assignments must be submitted via Webcourses through the assignment section. Email submissions will be ignored.
3. Extensions due to acceptable personal circumstances must be requested by email in advance of the deadline.
4. For late submissions (i.e. without an agreed extension), a penalty of 5% will be applied for every day a submission is late.
5. No submissions will be accepted after Friday December 15th 2018 @ 23:59 unless an extension has been agreed.
NB: Anything submitted later than this date without agreement will be ignored.
6. Assignments which do not adhere to the requirements or which are submitted incorrectly will attract a penalty of up to 10%.
7. No resubmission of assignments after feedback is given is allowed.

Basic Marking Scheme

	Option a	Option b
The ability to correctly prepare, present, analyse and critically assess the dataset used from the perspective of the proposed statistical analysis to justify use of chosen technique(s);	15	20
Option A		
Assessing fit and usefulness of model(s) created using appropriate statistical evidence;	30	
Assessing how well dataset/models meet assumptions of regression using appropriate statistics;	15	
Illustration of model using example data;	15	
Comparison of successive models using appropriate statistics.	15	
Option B		
Assessing the suitability of the dataset for the purposes of the dimension reduction technique chosen;		25
Description and assessment of the effectiveness of the outcomes of the dimension reduction using appropriate statistics;		25
Assessing fit and usefulness of regression model created using appropriate statistics;		10
Illustration of model using example data.		10
The ability to interpret the findings from your data within the context of your question and draw conclusions from this.	10	10
Total	100	100