



The University of Southern Queensland

Course specification

Description: Data Analysis

Subject	Cat-nbr	Class	Term	Mode	Units	Campus
STA	2300	48090	3, 2005	EXT	1.00	Toowoomba

Academic group:	FOSCI
Academic org:	FOS003
Student contribution band:	2
ASCED code:	010103

STAFFING

Examiner: Shahjahan Khan

Moderator: Taryn Swan

OTHER-REQUISITES

Recommended Pre-requisite: CSC1402 or CIS1000 or ELE1301

RATIONALE

Statistics are pervasive in work and life. They are the product of a society interested in understanding itself and the world in which it exists. To this end, data is being generated at an ever-accelerating rate. The collection and conversion of data into useful information is what the discipline of statistics is all about. Whether it is the planning and implementation of a survey to assess the market penetration of a new product, the design of an experiment to test the efficacy of a new drug, or the gathering together and summarizing of data provided by a government organization to support an argument, the discipline of statistics contributes in an essential way. It addresses questions such as how best to collect data, how much data to collect, how best to summarise and analyse data, and how to draw legitimate conclusions from data. Never before has some understanding of the discipline of statistics been so important to an educated person. Regardless of whether you ever need to initiate the collection or analysis of data in your future studies or future work, some understanding of statistical methods is highly desirable, if not essential, in being able to critically appraise the statistical methods employed by others in generated information of importance to you. This course endeavours to provide this understanding by covering basic statistical concepts and giving practice at some of the methods and skills necessary for students in business, commerce, psychology and the physical sciences to collect, appraise, present, analyse and interpret data. Because these concepts and methods are interdisciplinary in nature, students will encounter problems from many sources including their own area of interest. The statistical knowledge developed in this course forms the basis for more advanced statistical methods and concepts in specialist fields and in other courses in several programmes at USQ, as well as being of interest in their own right in contributing towards our understanding of the knowledge society in which we live.

SYNOPSIS

Students are introduced to the basic concepts involved in descriptive and inferential statistics. Topics include: methods of producing data and the impact and importance of randomness and randomisation in data generation; methods of summarising and displaying data in single and two-variable situations; modelling of data using the normal and binomial distributions; and analysis of data with a view to drawing reliable conclusions including specifically the use of hypothesis testing and confidence intervals in dealing with means and proportions, assessing associations and correlations and making predictions using simple regression. Emphasis is placed on understanding the basic concepts and principles of dealing with data. In particular, issues are covered relating to cause and effect; the nature of variability and the reliability of a sample in representing a population; the presentation of summarized data in tabular, graphical or descriptive form; the appropriate choice of parameter or parameters in meeting an objective of a study; the limitations and assumptions underpinning statistical techniques; the impact of sample size on inference; the appropriate use of language in interpreting an analysis; and the use of software in facilitating summary and analysis. The mathematical underpinning of the statistical methodologies used are not covered. Other statistics courses deal with these aspects.

OBJECTIVES

On completion of this course students will be able to:

1. demonstrate an understanding of why and how statistics as a discipline contributes in essential ways to all other disciplines;
2. choose and apply appropriate graphical and numerical tools for organising, describing and displaying summarised data in single and two variable situations as appropriate for categorical and quantitative variables and their combinations;
3. demonstrate an awareness of the impact of lurking variables in making cause and effect conclusions and be able to critically appraise studies in this regard;
4. demonstrate an understanding of the distinction between and the major advantages and disadvantages of observational and experimental studies;
5. describe several commonly-used sampling designs for surveys and be aware of the major statistical issues relating to sample surveys;
6. demonstrate an awareness of the statistical principles of good experimentation;
7. recognise the impact of conditioning in determining proportions or probabilities;
8. model appropriate real-life situations using the normal and binomial models;
9. demonstrate an understanding of and quantify the impact of variability and sample size on the precision and reliability of commonly-used statistics in random sampling;
10. make the appropriate choice of parameter or parameters and choose the appropriate procedure to employ in a range of standard single-sample and two-sample analyses;
11. demonstrate an understanding of the impact of sample size on inference;
12. demonstrate an understanding of the construction, interpretation and usefulness of a confidence interval and its relationship to hypothesis testing;
13. demonstrate an understanding of the meaning of a P-value and be able to make appropriate use of it in stating a conclusion from a hypothesis test;
14. demonstrate an understanding of the distinction between statistical significance and practical significance;
15. demonstrate an understanding of the meanings of association and correlation and be able to assess their presence appropriately;
16. make use of simple regression as a means of estimating and predicting where applicable;
17. demonstrate an awareness that statistics is not an exact science and there is generally more than one way of appropriately summarising and analysing data;
18. demonstrate an awareness of the existence of parametric and nonparametric methods;

19. demonstrate an awareness that assumptions underpin all statistical analyses and be able to check where possible those assumptions in a range of standard analyses;
20. make appropriate use of a statistical computer package with regard to data entry, data summary, graphics generation, analysis and interpretation of analyses, as relevant to this course.

TOPICS

	Description	Weighting (%)
1.	Quantitative basics (not examinable): Use of the calculator: order of operations, brackets, percentages, powers, rounding, use of the memory key, other useful keys, interpretation and substitution into formulae, statistical functions. Graphing: coordinates, straight lines, gradient, general form, special equations.	1.00
2.	Exploring and understanding data: The discipline of statistics: relevance, importance, applications. Variables and values; primary and secondary data; nominal, ordinal, quantitative data, single and multi-variable data. Important sources of secondary data. Introduction to SPSS. Displaying nominal and ordinal data: tables, frequencies, relative frequencies; bar graphs, pie charts. Two categorical variables: contingency tables, interpretation, marginal distributions, conditional distributions, clustered bar graphs, association and independence. Displaying quantitative data: histograms, stem and leaf displays. Describing distributions: shape, centre, spread, outliers. Summary statistics: mean, media, quartiles, interquartile range, five number summary, standard deviation. Standardising and applications. The normal model: 68-95-99.7 rule, standard normal, using normal tables, assessing normality.	19.00
3.	Exploring relationships between variables: Two quantitative variables: scatterplots, form, strength, direction; linearity and correlation; simple linear regression, explanatory and response variables, intercept and slope, predictions, residuals; R sq. Problems in regression: extrapolation, outliers, influential observations; lurking variables, causation. One quantitative and one categorical variable: boxplots, comparison of summary statistics.	12.00
4.	Gathering data: Sample surveys: census and sampling, populations and samples, a parameters and statistics; simple random sampling, stratified sampling, systematic sampling, non-probability sampling; cautions: under-coverage, response bias, non-response bias. Experiments: comparative experiments; experimental units, factors, levels, treatments; principles of control, random assignment, replication, blocking; blinding, placebos; causation and confounding; concept of significance.	10.00
5.	Randomness and probability: Random phenomena, long-run relative frequency, the law of large numbers, personal (subjective) probability, basic probability properties and rules, equally likely outcomes, disjoint events, independent events.	10.00

- Random variables: probability models, probability distributions; expected value and standard deviation of a probability distribution. Binomial model, binomial probabilities from tables, mean and standard deviation, normal approximation model.
6. Sampling distribution models: Distribution of a sample proportion; distribution of a sample mean; central limit theorem. Standard errors: proportions, means, other statistics; error bars, tabular and graphical representation; principle of diminishing returns. 8.00
 7. Generalising to the World at Large: Introduction to inference. Sign test: null and alternative hypotheses, test statistic, logic of hypothesis testing, P-values, one and two-sided alternatives, statistical significance. Large sample confidence intervals: proportion, mean, margin of error. Sample size determination: means, proportions. Large sample hypothesis test for a proportion, conditions. 12.00
 8. Learning about the World: Inference for means: the t distribution, tests and confidence intervals (CI), assumptions, robustness. Comparing two means: independent samples, unequal variances t procedure, tests and C.I.s, assumptions, robustness; blocking and paired samples, tests and C.I.s. Distinction between parametric and nonparametric procedures. Connection between C.I.s and hypothesis tests. 12.00
 9. Chi square testing: Inference and contingency tables: Goodness of fit testing, test of independence: expected and observed counts, chi-square distribution, robustness, residual examination, conditions. 8.00
 10. Inferences for regression: Regression model: population and sample, assumptions, residual plot, checking residuals; inference for the slope and mean and individual predictions; standard errors; cautions. 8.00

TEXT and MATERIALS required to be PURCHASED or ACCESSED

ALL textbooks and materials are available for purchase from USQ BOOKSHOP (unless otherwise stated). Orders may be placed via secure internet, free fax 1800642453, phone 07 46312742 (within Australia), or mail. Overseas students should fax +61 7 46311743, or phone +61 7 46312742. For costs, further details, and internet ordering, use the 'Textbook Search' facility at <http://bookshop.usq.edu.au> click 'Semester', then enter your 'Course Code' (no spaces).

SPSS Student Version 12.0 (Version 10.0 or later is acceptable) for Windows, Prentice Hall. (Available separately or bundled with De Veaux & Velleman)

De Veaux, RD & Velleman, PF 2004, *Intro Stats*, Pearson Addison Wesley,

REFERENCE MATERIALS

Reference materials are materials that, if accessed by students, may improve their knowledge and understanding of the material in the course and enrich their learning experience.

STUDENT WORKLOAD REQUIREMENTS

ACTIVITY	HOURS
Assessment	20.00
Examinations	3.00
Private Study	147.00

ASSESSMENT DETAILS

Description	Marks out of	Wtg(%)	Due date
ASSIGNMENT	10.00	2.00	25 Nov 2005
CMA ON TOPICS 2 TO 4	15.00	4.00	12 Dec 2005
ASSIGN ON TOPICS 2 TO 5	100.00	15.00	03 Jan 2006
CMA ON TOPICS 5 TO 7	15.00	4.00	16 Jan 2006
ASSIGN ON TOP UPTO & INC TOP 8	100.00	15.00	23 Jan 2006
EXAM PTA 3HR RESTRICTED	20.00	20.00	END S3 (see note 1)
PART B OF ABOVE 3HR EXAM	40.00	40.00	END S3

NOTES

1. Examination dates will be available during the semester. Please refer to the examination timetable when published.

IMPORTANT ASSESSMENT INFORMATION

- 1 Attendance requirements:
There are no attendance requirements for this course. However, it is the students' responsibility to study all material provided to them or required to be accessed by them to maximise their chance of meeting the objectives of the course and to be informed of course-related activities and administration.
- 2 Requirements for students to complete each assessment item satisfactorily:
To complete each of the assessment items satisfactorily, students must obtain at least 50% of the marks available for each assessment item.
- 3 Penalties for late submission of required work:
If students submit assignments after the due date without prior approval then a penalty of 10% of the total marks gained by the student for the assignment will apply for each working day late.
- 4 Requirements for student to be awarded a passing grade in the course:
To be assured of a passing grade, students must demonstrate, via the summative assessment items, that they have achieved the required minimum standards in relation

to the objectives of the course by satisfactorily completing all summative assessment items (the examination, assignments and CMA's).

- 5 Method used to combine assessment results to attain final grade:
The final grades for students will be assigned on the basis of the aggregate of the weighted marks obtained for each of the summative assessment items in the course.
- 6 Examination information:
The only materials that candidates may use in the restricted examination for this course are: writing materials (non-electronic and free from material which could give the student an unfair advantage in the examination); calculators which cannot hold textual information (students must indicate on their examination paper the make and model of any calculator(s) they use during the examination). Formula sheets will be provided by the Examiner with the examination paper. Students whose first language is not English, may, with the Examiner's approval, take an appropriate non- electronic translation dictionary into the examination. Students who wish to use a translation dictionary MUST request and receive written approval from the Examiner at least one week before the examination date. Translation dictionaries will be subject to perusal and may be removed from the candidate's possession until appropriate disciplinary action is completed if found to contain material that could give the candidate an unfair advantage.
- 7 Examination period when Deferred/Supplementary examinations will be held:
Any Deferred or Supplementary examinations for this course will be held during the examination period at the end of the semester of the next offering of this course. Students who obtain an overall passing mark, but who do not perform satisfactorily in the examination, may, at the discretion of the examiner, be granted a supplementary examination. Students will be granted a deferred examination only if they perform satisfactorily in the assignments.
- 8 University Regulations:
Students should read USQ Regulations 5.1 Definitions, 5.6. Assessment, and 5.10 Academic Misconduct for further information and to avoid actions which might contravene University Regulations. These regulations can be found at the URL <http://www.usq.edu.au/corporateservices/calendar/part5.htm> or in the current USQ Handbook.

ASSESSMENT NOTES

- 9 The due date for an assignment is the date by which a student must despatch the assignment to the USQ. The onus is on the student to provide proof of the despatch date, if requested by the Examiner. Students must retain a copy of each item submitted for assessment. This should be despatched to USQ within 24 hours of receipt of a request to do so. The examiner may grant an extension of the due date of an assignment in extenuating circumstances.