STAT830

Statistical Methods in Bioinformatics

S1 Day 2016

Dept of Statistics

Contents

General Information

Unit convenor and teaching staff
Lecturer
Nino Kordzakhia
nino.kordzakhia@mq.edu.au
Level 2, AHH
TBA

Credit points
4

Prerequisites
Admission to MBiotech or MBiotechMCom or MBioBus or MLabQAMgt or PGDipLabQAMgt or PGCertLabQAMgt or GradDipLabQAMgt or GradCertLabQAMgt or ((admission to MConsBiol or GradDipConsBiol) and BIOL603))

Corequisites

Co-badged status

Unit description
This unit introduces the statistical and probabilistic concepts that are the basis for the study of bioinformatics. Topics include an introduction to probability and conditional probability, probability distributions, sampling distributions and an introduction to Markov processes. Particular attention is paid to how they relate to specific applications in the field of bioinformatics. A basic understanding of calculus will be an advantage.

Important Academic Dates

Information about important academic dates including deadlines for withdrawing from units are available at https://www.mq.edu.au/study/calendar-of-dates

Learning Outcomes

On successful completion of this unit, you will be able to:

Understand basic notions and fundamentals of Probability and Statistics.
Familiarity with special classes of discrete and continuous random variables and their distribution functions. Being able to evaluate probabilities of events, expected values and variances of random variables.
Validation of conditions for Hardy-Weinberg Equilibrium.
Understand basic properties of Markov Chains. Being able to recognise Markov

processes and understand how they can be used in applications.

Being able to apply the Probability theory in DNA sequencing analysis.

Be familiar with basic principles of statistical data modelling using nonparametric methods.

Assessment Tasks

Name	Weighting	Due
Test 1	10%	Week 3
Test 2	10%	Week 7
Assignment	20%	Week 9
Test 3	10%	Week 10
Final Examination	50%	University exam timetable

Test 1

Due: **Week 3**
Weighting: **10%**

Test 1 will be held in the tutorial time and will be 30 minutes long.

On successful completion you will be able to:

- Understand basic notions and fundamentals of Probability and Statistics.

Test 2

Due: **Week 7**
Weighting: **10%**

Test 2 will be held in the tutorial time and will be 30 minutes long.

On successful completion you will be able to:

- Familiarity with special classes of discrete and continuous random variables and their distribution functions. Being able to evaluate probabilities of events, expected values and variances of random variables.

- Validation of conditions for Hardy-Weinberg Equilibrium.

Assignment

Due: **Week 9**
Weighting: **20%**

The assignment will be administered via iLearn.

On successful completion you will be able to:

- Familiarity with special classes of discrete and continuous random variables and their distribution functions. Being able to evaluate probabilities of events, expected values and variances of random variables.
- Validation of conditions for Hardy-Weinberg Equilibrium.

Test 3

Due: **Week 10**
Weighting: **10%**

Test 3 will be held in the tutorial time and will be 30 minutes long.

On successful completion you will be able to:

- Understand basic properties of Markov Chains. Being able to recognise Markov processes and understand how they can be used in applications.
- Being able to apply the Probability theory in DNA sequencing analysis.
- Be familiar with basic principles of statistical data modelling using nonparametric methods.

Final Examination

Due: **University exam timetable**
Weighting: **50%**

A three-hour final examination for this unit will be held during the University Examination period.

You are permitted ONE A4 page of paper containing reference material printed or handwritten on both sides. The page will not be returned at the end of the final examination.

Calculators will be needed but must not be of the text/programmable type.

You are expected to present yourself for examination at the time and place designated in the University Examination Timetable. The timetable will be available in Draft form approximately eight weeks before the commencement of the examinations and in Final form approximately four weeks before the commencement of the examinations at

http://exams.mq.edu.au/

The Macquarie university examination policy details can be viewed at

http://www.mq.edu.au/policy/docs/examination/policy.htm

On successful completion you will be able to:

- Understand basic notions and fundamentals of Probability and Statistics.
- Familiarity with special classes of discrete and continuous random variables and their distribution functions. Being able to evaluate probabilities of events, expected values and variances of random variables.
- Validation of conditions for Hardy-Weinberg Equilibrium.
- Understand basic properties of Markov Chains. Being able to recognise Markov processes and understand how they can be used in applications.
- Being able to apply the Probability theory in DNA sequencing analysis.
- Be familiar with basic principles of statistical data modelling using nonparametric methods.

Delivery and Resources

Classes

Lectures begin in Week 1. Tutorials begin in Week 2.

Students must attend two hours of lectures and two hours of tutorials per week.

The lecture notes will be made available on iLearn before the lecture.

Tutorial exercises will be set weekly and will be available on iLearn before the tutorial.

The timetable for classes can be found at http://www.timetables.mq.edu.au

iLearn

All unit related materials including lecture notes, tutorials and instructions for assessment tasks and administrative updates, will be posted on iLearn at

https://ilearn.mq.edu.au/login/MQ/

Software

The statistical software R will be used. This is a free software environment for statistical computing and graphics and can be downloaded from the website

http://www.r-project.org/

Required and recommended texts and materials

There is no required textbook for this unit.

Recommended reference sources are:

1. W. P. Krijnen Applied Statistics for Bioinformatics using R, 2009.

http://cran.r-project.org/doc/contrib/Krijnen-IntroBioInfStatistics.pdf

2. S. Draghici Statistics and Data Analysis for Microarrays Using R and Bioconductor. Chapman & Hall/CRC Mathematical and Computational Biology, 2nd Edition, 2012.

3. W. J. Ewens and G. R. Grant. Statistical Methods in Bioinformatics, an Introduction. Springer, 2005.

4. K. Lange. Mathematical and Statistical Methods for Genetic Analysis, Statistics for Biology and Health. Springer, 2002.

5. J. C. Miller and J. N. Miller. Statistics for Analytical Chemistry. Ellis Horwood PTR Prentice Hall, Analytical Chemistry Series, 3rd edition, 1993.

Unit Schedule

Weeks	Lecture Topics
W1	Introduction
W2	Random Variables, Probability Functions, Characteristics of Random Variables.
W3-W4	Hardy-Weinberg Equilibrium
W5-W6	Statistical problems in DNA sequencing
Mid –session break: 11/04/16-22/04/16 ***25/04/16 Anzac Day Public Holiday***	
W7-W8	Hypothesis testing and its applications

W9-W10	Markov Processes and their applications
W11-W12	Nonparametric statistics and its applications
W13	Review

Policies and Procedures

Macquarie University policies and procedures are accessible from Policy Central. Students should be aware of the following policies in particular with regard to Learning and Teaching:

Academic Honesty Policy http://mq.edu.au/policy/docs/academic_honesty/policy.html

New Assessment Policy in effect from Session 2 2016 http://mq.edu.au/policy/docs/assessment/policy_2016.html. For more information visit http://students.mq.edu.au/events/2016/07/19/new_assessment_policy_in_place_from_session_2/

Assessment Policy prior to Session 2 2016 http://mq.edu.au/policy/docs/assessment/policy.html

Grading Policy prior to Session 2 2016 http://mq.edu.au/policy/docs/grading/policy.html

Grade Appeal Policy http://mq.edu.au/policy/docs/gradeappeal/policy.html

Complaint Management Procedure for Students and Members of the Public http://www.mq.edu.au/policy/docs/complaint_management/procedure.html

Disruption to Studies Policy http://www.mq.edu.au/policy/docs/disruption_studies/policy.html *The Disruption to Studies Policy is effective from March 3 2014 and replaces the Special Consideration Policy.*

In addition, a number of other policies can be found in the Learning and Teaching Category of Policy Central.

Student Code of Conduct

Macquarie University students have a responsibility to be familiar with the Student Code of Conduct: https://students.mq.edu.au/support/student_conduct/

Results

Results shown in *iLearn*, or released directly by your Unit Convenor, are not confirmed as they are subject to final approval by the University. Once approved, final results will be sent to your student email address and will be made available in eStudent. For more information visit ask.mq.edu.au.

Student Support

Macquarie University provides a range of support services for students. For details, visit http://students.mq.edu.au/support/

Learning Skills

Learning Skills (mq.edu.au/learningskills) provides academic writing resources and study strategies to improve your marks and take control of your study.

- Workshops
- StudyWise
- Academic Integrity Module for Students
- Ask a Learning Adviser

Student Services and Support

Students with a disability are encouraged to contact the Disability Service who can provide appropriate help with any issues that arise during their studies.

Student Enquiries

For all student enquiries, visit Student Connect at ask.mq.edu.au

IT Help

For help with University computer systems and technology, visit http://www.mq.edu.au/about_us/offices_and_units/information_technology/help/.

When using the University's IT, you must adhere to the Acceptable Use of IT Resources Policy. The policy applies to all who connect to the MQ network including students.

Graduate Capabilities

PG - Discipline Knowledge and Skills

Our postgraduates will be able to demonstrate a significantly enhanced depth and breadth of knowledge, scholarly understanding, and specific subject content knowledge in their chosen fields.

This graduate capability is supported by:

Learning outcomes

- Understand basic notions and fundamentals of Probability and Statistics.
- Familiarity with special classes of discrete and continuous random variables and their distribution functions. Being able to evaluate probabilities of events, expected values and variances of random variables.
- Validation of conditions for Hardy-Weinberg Equilibrium.
- Understand basic properties of Markov Chains. Being able to recognise Markov processes and understand how they can be used in applications.
- Be familiar with basic principles of statistical data modelling using nonparametric

methods.

Assessment tasks

- Test 1
- Test 2
- Assignment
- Test 3
- Final Examination

PG - Critical, Analytical and Integrative Thinking

Our postgraduates will be capable of utilising and reflecting on prior knowledge and experience, of applying higher level critical thinking skills, and of integrating and synthesising learning and knowledge from a range of sources and environments. A characteristic of this form of thinking is the generation of new, professionally oriented knowledge through personal or group-based critique of practice and theory.

This graduate capability is supported by:

Learning outcomes

- Understand basic notions and fundamentals of Probability and Statistics.
- Familiarity with special classes of discrete and continuous random variables and their distribution functions. Being able to evaluate probabilities of events, expected values and variances of random variables.
- Validation of conditions for Hardy-Weinberg Equilibrium.
- Understand basic properties of Markov Chains. Being able to recognise Markov processes and understand how they can be used in applications.
- Being able to apply the Probability theory in DNA sequencing analysis.

Assessment tasks

- Test 1
- Test 2
- Assignment
- Test 3
- Final Examination

PG - Research and Problem Solving Capability

Our postgraduates will be capable of systematic enquiry; able to use research skills to create new knowledge that can be applied to real world issues, or contribute to a field of study or practice to enhance society. They will be capable of creative questioning, problem finding and problem solving.

This graduate capability is supported by:

Learning outcomes

- Familiarity with special classes of discrete and continuous random variables and their distribution functions. Being able to evaluate probabilities of events, expected values and variances of random variables.
- Validation of conditions for Hardy-Weinberg Equilibrium.
- Understand basic properties of Markov Chains. Being able to recognise Markov processes and understand how they can be used in applications.

Assessment tasks

- Test 2
- Assignment
- Test 3
- Final Examination

Changes from Previous Offering