STAT830
Statistical Methods in Bioinformatics
S1 Day 2016
Dept of Statistics

Contents

General Information .................................................. 2
Learning Outcomes ............................................... 2
Assessment Tasks .................................................. 3
Delivery and Resources ........................................... 5
Unit Schedule ....................................................... 6
Policies and Procedures .......................................... 7
Graduate Capabilities ............................................. 8
Changes from Previous Offering .................. 10

Disclaimer
Macquarie University has taken all reasonable measures to ensure the information in this publication is accurate and up-to-date. However, the information may change or become out-dated as a result of change in University policies, procedures or rules. The University reserves the right to make changes to any information in this publication without notice. Users of this publication are advised to check the website version of this publication [or the relevant faculty or department] before acting on any information in this publication.
**General Information**

<table>
<thead>
<tr>
<th>Unit convenor and teaching staff</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lecturer</td>
</tr>
<tr>
<td>Nino Kordzakhia</td>
</tr>
<tr>
<td><a href="mailto:nino.kordzakhia@mq.edu.au">nino.kordzakhia@mq.edu.au</a></td>
</tr>
<tr>
<td>Level 2, AHH</td>
</tr>
<tr>
<td>TBA</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Credit points</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Prerequisites</th>
</tr>
</thead>
<tbody>
<tr>
<td>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))</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Corequisites</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Co-badged status</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Unit description</th>
</tr>
</thead>
<tbody>
<tr>
<td>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.</td>
</tr>
</tbody>
</table>

**Important Academic Dates**

Information about important academic dates including deadlines for withdrawing from units are available at [http://students.mq.edu.au/student_admin/enrolmentguide/academicdates/](http://students.mq.edu.au/student_admin/enrolmentguide/academicdates/)

**Learning Outcomes**

1. Understand basic notions and fundamentals of Probability and Statistics.
2. 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.
3. Validation of conditions for Hardy-Weinberg Equilibrium.
4. Understand basic properties of Markov Chains. Being able to recognise Markov processes and understand how they can be used in applications.
5. Being able to apply the Probability theory in DNA sequencing analysis.
6. Be familiar with basic principles of statistical data modelling using nonparametric methods.

Assessment Tasks

<table>
<thead>
<tr>
<th>Name</th>
<th>Weighting</th>
<th>Due</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test 1</td>
<td>10%</td>
<td>Week 3</td>
</tr>
<tr>
<td>Test 2</td>
<td>10%</td>
<td>Week 7</td>
</tr>
<tr>
<td>Assignment</td>
<td>20%</td>
<td>Week 9</td>
</tr>
<tr>
<td>Test 3</td>
<td>10%</td>
<td>Week 10</td>
</tr>
<tr>
<td>Final Examination</td>
<td>50%</td>
<td>University exam timetable</td>
</tr>
</tbody>
</table>

Test 1
Due: **Week 3**
Weighting: 10%

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

This Assessment Task relates to the following Learning Outcomes:
- 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.

This Assessment Task relates to the following 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.
Assignment
Due: Week 9
Weighting: 20%

The assignment will be administered via iLearn.

This Assessment Task relates to the following 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.

Test 3
Due: Week 10
Weighting: 10%

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

This Assessment Task relates to the following Learning Outcomes:
- 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://unitguides.mq.edu.au/unit_offerings/57550/unit_guide/print
This Assessment Task relates to the following 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.
- 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](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/](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/](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  
   [http://cran.r-project.org/doc/contrib/Krijnen-IntroBioInfStatistics.pdf](http://cran.r-project.org/doc/contrib/Krijnen-IntroBioInfStatistics.pdf)

2. S. Draghici  


4. K. Lange.  

5. J. C. Miller and J. N. Miller.  

---

**Unit Schedule**

<table>
<thead>
<tr>
<th>Weeks</th>
<th>Lecture Topics</th>
</tr>
</thead>
<tbody>
<tr>
<td>W1</td>
<td>Introduction</td>
</tr>
<tr>
<td>W2</td>
<td>Random Variables, Probability Functions, Characteristics of Random Variables.</td>
</tr>
<tr>
<td>W3-W4</td>
<td>Hardy-Weinberg Equilibrium</td>
</tr>
<tr>
<td>W5-W6</td>
<td>Statistical problems in DNA sequencing</td>
</tr>
</tbody>
</table>
**Mid-session break:** 11/04/16-22/04/16

25/04/16 Anzac Day Public Holiday

<table>
<thead>
<tr>
<th>Week</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>W7-W8</td>
<td>Hypothesis testing and its applications</td>
</tr>
<tr>
<td>W9-W10</td>
<td>Markov Processes and their applications</td>
</tr>
<tr>
<td>W11-W12</td>
<td>Nonparametric statistics and its applications</td>
</tr>
<tr>
<td>W13</td>
<td>Review</td>
</tr>
</tbody>
</table>

**Policies and Procedures**

Macquarie University policies and procedures are accessible from [Policy Central](http://mq.edu.au/policy/docs/). Students should be aware of the following policies in particular with regard to Learning and Teaching:


In addition, a number of other policies can be found in the [Learning and Teaching Category](http://mq.edu.au/policy/docs/) 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/](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 Enquiry Service

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

Equity 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.

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 - 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 - 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
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**