



# STAT830

## Statistical Methods in Bioinformatics

S1 Day 2019

*Dept of Mathematics and Statistics*

### Contents

|                                       |   |
|---------------------------------------|---|
| <u>General Information</u>            | 2 |
| <u>Learning Outcomes</u>              | 2 |
| <u>General Assessment Information</u> | 3 |
| <u>Assessment Tasks</u>               | 4 |
| <u>Delivery and Resources</u>         | 5 |
| <u>Unit Schedule</u>                  | 6 |
| <u>Policies and Procedures</u>        | 7 |
| <u>Graduate Capabilities</u>          | 8 |

#### **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

Unit convenor and teaching staff

Lecturer

Nino Kordzakhia

[nino.kordzakhia@mq.edu.au](mailto:nino.kordzakhia@mq.edu.au)

Contact via Email

12 Wally's Walk Office 6.11

11am -12pm & 1pm - 2pm Tuesday

Credit points

4

Prerequisites

Admission to MBiotech or MSc or MDataSc or GradDipBioTech or MBiotechMCom or MBioBus or MLabQAMgt or GradDipLabQAMgt or GradCertLabQAMgt or MConsBiol or GradDipConsBiol or MMarScMgt or MScInnovation

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:

Knowledge of basic notions and fundamentals of Probability and Statistics.

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

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

Knowledge of basic properties of Markov Chains. Being able to recognise Markov processes and understand how they can be used in applications.

Knowledge of fundamental principles of statistical data modelling utilised in practice.

## General Assessment Information

### Small Group Teaching Activity (SGTA)

If you are unable to attend a weekly SGTA, you must apply for special consideration via <https://ask.mq.edu.au/>

The SGTA work must be submitted by the official due date and time. Late submissions will not be marked unless an extension has been granted following a successful application for .

### Late Submission of Work

All assignments and assessment tasks must be submitted by the official due date and time. No marks will be given for late work unless an extension has been granted following a successful application for [Special Consideration](#). Please contact the unit convenor for advice as soon as you become aware that you may have difficulty meeting any of the assignment deadlines.

### Mid-Semester Test

You are permitted ONE A4 page of paper containing reference material printed or handwritten on both sides. Calculators will be needed but must not be of the text/programmable type.

If a student is unable to attend the mid-semester test due to unavoidable circumstances, the student must apply for special consideration via <https://ask.mq.edu.au/>

### Practical Test

The unit material distributed via iLearn is permitted. You may bring to the practical test a hard copy of your lecture and SGTA notes.

#### **IMPORTANT**

If a student is unable to attend the Practical Test due to unavoidable circumstances, the student must apply for Special Consideration via <https://ask.mq.edu.au/>

If you apply for Special Consideration for the Practical Test, you must make yourself available for the Supplementary Examination as organised by the Faculty of Science & Engineering. If you are not available at that time, there is no guarantee that an additional examination time will be offered. Specific examination dates and times will be determined at a later date. You can check the details of Supplementary exams via iLearn page FSE101 ([bit.LY/FSESup](https://ask.mq.edu.au/))

## Assessment Tasks

| Name                     | Weighting | Hurdle | Due     |
|--------------------------|-----------|--------|---------|
| <u>SGTA Work 1</u>       | 15%       | No     | Week 4  |
| <u>Mid-Semester Test</u> | 20%       | No     | Week 8  |
| <u>SGTA Work 2</u>       | 15%       | No     | Week 11 |
| <u>Practical Test</u>    | 50%       | No     | Week 13 |

### SGTA Work 1

Due: **Week 4**

Weighting: **15%**

A worksheet will be handed out in the second hour of the class and is to be submitted before 23:55 on the day of the class via iLearn.

For the SGTA work assessment conditions see **General Assessment Information**.

On successful completion you will be able to:

- Knowledge of basic notions and fundamentals of Probability and Statistics.

### Mid-Semester Test

Due: **Week 8**

Weighting: **20%**

The Mid-Semester Test will be held in the second hour of lecture time and will be 50 minutes long.

For the test conditions see **General Assessment Information**.

On successful completion you will be able to:

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

### SGTA Work 2

Due: **Week 11**

Weighting: **15%**

A worksheet will be handed out in the second hour of the class and is to be submitted before 23:55 on the day of the class via iLearn.

For the SGTA work assessment conditions see **General Assessment Information**.

On successful completion you will be able to:

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

## Practical Test

Due: **Week 13**

Weighting: **50%**

The Practical Test will be held during the SGTA class and will be 1 hour and 40 minutes long. For the practical test conditions see **General Assessment Information**.

On successful completion you will be able to:

- Knowledge of basic notions and fundamentals of Probability and Statistics.
- Knowledge of special classes of discrete and continuous random variables and their distribution functions. Being able to evaluate the probabilities of events, expected values and variances of random variables.
- Being able to apply the Probability theory in DNA sequencing analysis.
- Knowledge of basic properties of Markov Chains. Being able to recognise Markov processes and understand how they can be used in applications.
- Knowledge of fundamental principles of statistical data modelling utilised in practice.

## Delivery and Resources

### Classes

**Lectures begin in Week 1. SGTA begin in Week 2.**

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

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

SGTA exercises will be set weekly and will be available on iLearn before each class.

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

### iLearn

All unit related materials including lecture notes, SGTA's and instructions for assessment tasks and administrative updates, will be published 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/>

### **Texts and materials**

There is no required textbook for this unit.

#### ***Recommended reference sources:***

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. P. N. Suravajhala. Your passport to a career in bioinformatics. New Delhi: Springer, 2013.

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

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

## **Unit Schedule**

| <b>Weeks</b>                                 | <b>Lecture Topics</b>  |
|--|--|
| <b>W1</b>                                    | Introduction   |
| <b>W2</b>                                    | Discrete random variables and their characteristics  |
| <b>W3 - W5</b>                               | Hardy-Weinberg Equilibrium (HWE); Departures from HWE; Statistical testing of HWE.           |
| <b>W6 - W7</b>                               | HWE for X-linked loci.<br>Introduction to continuous random variables: Uniform Distribution. |
| <b>Session 1 Recess: 15/04/19 - 26/04/19</b> |  |
| <b>W8</b>                                    | Continuous random variables and their characteristics  |

|           |   |
|-----------|---|
| W10 - W11 | Hypothesis testing and its applications |
| W12       | Markov Chains and their applications    |
| W13       | Review                                  |

## Policies and Procedures

Macquarie University policies and procedures are accessible from [Policy Central](https://staff.mq.edu.au/work/strategy-planning-and-governance/university-policies-and-procedures/policy-central) (<https://staff.mq.edu.au/work/strategy-planning-and-governance/university-policies-and-procedures/policy-central>). Students should be aware of the following policies in particular with regard to Learning and Teaching:

- [Academic Appeals Policy](#)
- [Academic Integrity Policy](#)
- [Academic Progression Policy](#)
- [Assessment Policy](#)
- [Fitness to Practice Procedure](#)
- [Grade Appeal Policy](#)
- [Complaint Management Procedure for Students and Members of the Public](#)
- [Special Consideration Policy](#) (**Note:** *The Special Consideration Policy is effective from 4 December 2017 and replaces the Disruption to Studies Policy.*)

Undergraduate students seeking more policy resources can visit the [Student Policy Gateway](https://students.mq.edu.au/support/study/student-policy-gateway) (<https://students.mq.edu.au/support/study/student-policy-gateway>). It is your one-stop-shop for the key policies you need to know about throughout your undergraduate student journey.

If you would like to see all the policies relevant to Learning and Teaching visit [Policy Central](https://staff.mq.edu.au/work/strategy-planning-and-governance/university-policies-and-procedures/policy-central) (<https://staff.mq.edu.au/work/strategy-planning-and-governance/university-policies-and-procedures/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/study/getting-started/student-conduct>

## Results

Results published on platform other than [eStudent](#), (eg. iLearn, Coursera etc.) 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](https://ask.mq.edu.au) or if you are a Global MBA student contact [globalmba.support@mq.edu.au](mailto:globalmba.support@mq.edu.au)

## Student Support

Macquarie University provides a range of support services for students. For details, visit <http://stu>

[dents.mq.edu.au/support/](https://unitguides.mq.edu.au/support/)

## Learning Skills

Learning Skills ([mq.edu.au/learningskills](https://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](https://ask.mq.edu.au)

If you are a Global MBA student contact [globalmba.support@mq.edu.au](mailto:globalmba.support@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/](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

- Knowledge of basic notions and fundamentals of Probability and Statistics.
- Knowledge of special classes of discrete and continuous random variables and their distribution functions. Being able to evaluate the probabilities of events, expected values and variances of random variables.
- Knowledge of basic properties of Markov Chains. Being able to recognise Markov processes and understand how they can be used in applications.



- Knowledge of fundamental principles of statistical data modelling utilised in practice.

## **Assessment tasks**

- SGTA Work 1
- Mid-Semester Test
- SGTA Work 2
- Practical Test

## **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**

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

## **Assessment tasks**

- SGTA Work 1
- Mid-Semester Test
- SGTA Work 2
- Practical Test

## **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**

- Knowledge of special classes of discrete and continuous random variables and their distribution functions. Being able to evaluate the probabilities of events, expected values and variances of random variables.
- Knowledge of basic properties of Markov Chains. Being able to recognise Markov processes and understand how they can be used in applications.

## **Assessment tasks**

- Mid-Semester Test
- SGTA Work 2
- Practical Test