# STAT8830

# Statistical Methods in Bioinformatics

Session 1, In person-scheduled-weekday, North Ryde 2023

*School of Mathematical and Physical Sciences*

# Contents

**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
Nino Kordzakhia
nino.kordzakhia@mq.edu.au

Credit points
10

Prerequisites
Admission to MBiotech or (MSc or MScInnovation) or GradDipBioTech or MBiotechMCom or MBioBus or MLabQAMgt or GradDipLabQAMgt or GradCertLabQAMgt or MConsBiol or GradDipConsBiol or MMarScMgt or GradDipMarScMgt or MRadiopharm

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:

**ULO1:** Communicate the knowledge of fundamentals of Probability and Statistics using specific terminology.

**ULO2:** Use relevant terminology and describe the distribution functions and characteristics of some discrete and continuous random variables.

**ULO3:** Evaluate probabilities of events, expected values andvariances of random variables.

**ULO4:** Apply statistical and probabilistic modelling approach to genetic data.

**ULO5:** Apply fundamental principles of statistical data analysis.

# General Assessment Information

**Requirements to pass this unit:**

1. Attempt all assessments and

2. Achieve a total mark equal to or greater than 50%

**ASSIGNMENT SUBMISSION**: Assignment submission will be online through the iLearn page. Submit assignments online via the appropriate assignment link on the iLearn page.

- Assignment submission is via iLearn. You should upload this as a single scanned PDF file.

- It is your responsibility to make sure your assignment submission is legible.

You may submit as often as required prior to the due date/time. Please note that each submission will completely replace any previous submissions. It is in your interests to make frequent submissions of your partially completed work as insurance against technical or other problems near the submission deadline.

## LATE SUBMISSION OF WORK:

- In the case of late submission of an assignment, if special consideration has NOT been granted, the following deductions will be applied to the awarded assessment mark: 12 to 24 hours late = 10% deduction; for each day thereafter, an additional 10% per day or part thereof (including weekends and/or public holidays) will be applied until five days beyond the due date. After this time, a mark of zero (0) will be given.

**Assignments 1 and 2:** YES, Standard Late Penalty applies

**Test:** NO, unless Special Consideration is Granted

**Practical Test:** NO, unless Special Consideration is Granted

**Special Considerations:** The Special Consideration Policy aims to support students who have been impacted by short-term circumstances or events that are serious, unavoidable, and significantly disruptive, and which may affect their performance in assessment. If you experience circumstances or events that affect your ability to complete the assessments in this unit on time, please inform the convenor and submit a Special Consideration request through ask.mq.edu.au

# Assessment Tasks

| Name | Weighting | Hurdle | Due |
| --- | --- | --- | --- |
| Assignment 1 | 10% | No | Week 4 |
| Test | 30% | No | Week 8 |

| Name | Weighting | Hurdle | Due |
|------|-----------|--------|-----|
| Assignment 2 | 10% | No | Week 11 |
| Practical Test | 50% | No | Week 12 |

# Assignment 1

Assessment Type [1]: Quantitative analysis task
Indicative Time on Task [2]: 28.5 hours
Due: **Week 4**
Weighting: **10%**

Reinforce and apply skills learned in computer labs through data analysis. The tasks given during computer lab sessions are to be completed within the allocated time and submitted via iLearn.

On successful completion you will be able to:

• Apply fundamental principles of statistical data analysis.

# Test

Assessment Type [1]: Quiz/Test
Indicative Time on Task [2]: 1 hours
Due: **Week 8**
Weighting: **30%**

This is a paper based mid-semester test. Further information will be provided in the iLearn site of the unit.

On successful completion you will be able to:

• Use relevant terminology and describe the distribution functions and characteristics of some discrete and continuous random variables.

• Evaluate probabilities of events, expected values andvariances of random variables.

• Apply statistical and probabilistic modelling approach to genetic data.

# Assignment 2

Assessment Type [1]: Quantitative analysis task
Indicative Time on Task [2]: 28.5 hours

Due: **Week 11**
Weighting: **10%**

Reinforce and apply skills learned in computer labs through data analysis. The tasks given during computer lab sessions are to be completed within the allocated time and submitted via iLearn.

On successful completion you will be able to:

- Apply fundamental principles of statistical data analysis.

## Practical Test

Assessment Type [1]: Quiz/Test
Indicative Time on Task [2]: 2 hours
Due: **Week 12**
Weighting: **50%**

The practical test is designed to examine data analysis and R output interpretation skills taught in the unit.

On successful completion you will be able to:

- Communicate the knowledge of fundamentals of Probability and Statistics using specific terminology.
- Apply fundamental principles of statistical data analysis.

[1] If you need help with your assignment, please contact:

- the academic teaching staff in your unit for guidance in understanding or completing this type of assessment
- the Writing Centre for academic skills support.

[2] Indicative time-on-task is an estimate of the time required for completion of the assessment task and is subject to individual variation

# Delivery and Resources

## Classes:

The lectures begin in Week 1. SGTAs begin in Week 2.

Students must attend two hours of lectures and 1-hour 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  https://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/

## *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.  W. J. Ewens and G. R. Grant. Statistical Methods in Bioinformatics, an Introduction. Springer, 2000

# Methods of Communication:

We will communicate with you via your university email or through announcements on iLearn. Queries to convenors can be sent through direct email to the unit convenor.

Students can access the *iLearn* page by logging on at https://ilearn.mq.edu.au. Students must log in regularly to read the Announcements and access the teaching material.

# COVID Information:

For the latest information on the University's response to COVID-19, please refer to the Coronavirus infection page on the Macquarie University website: https://www.mq.edu.au/about/coronavirus-faqs. Remember to check this page regularly in case the information and requirements change during the semester. If there are any changes to this unit in relation to COVID, these will be communicated via *iLearn*.

# Unit Schedule

| Study Weeks | Lecture Topics | Due |
|---|---|---|
| W1 | Introduction | |
| W2 | Discrete random variables and their characteristics | |
| W3 - W5 | Hardy-Weinberg Equilibrium (HWE); <br><br> Departures from HWE; Statistical testing of HWE. | Week 4 <br><br> Assignment 1 |
| W6 - W7 | HWE for X-linked loci. <br><br> Introduction to continuous random variables: Uniform Distribution. | |
| **MID-SESSION BREAK** | | |
| W8 | Continuous random variables and their characteristics | Test |
| W10 - W11 | Hypothesis testing and its applications <br> Markov Chains and their applications | Week 11 <br> Assignment 2 |
| W12 | | Practical Test |

# Policies and Procedures

Macquarie University policies and procedures are accessible from Policy Central (https://policies.mq.edu.au). 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

- Assessment Procedure

- Complaints Resolution Procedure for Students and Members of the Public

- Special Consideration Policy

Students seeking more policy resources can visit Student Policies (https://students.mq.edu.au/support/study/policies). It is your one-stop-shop for the key policies you need to know about throughout your undergraduate student journey.

To find other policies relating to Teaching and Learning, visit Policy Central (https://policies.mq.edu.au) and use the search tool.

## Student Code of Conduct

Macquarie University students have a responsibility to be familiar with the Student Code of Conduct: https://students.mq.edu.au/admin/other-resources/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 or if you are a Global MBA student contact globalmba.support@mq.edu.au

# Academic Integrity

At Macquarie, we believe academic integrity – honesty, respect, trust, responsibility, fairness and courage – is at the core of learning, teaching and research. We recognise that meeting the expectations required to complete your assessments can be challenging. So, we offer you a range of resources and services to help you reach your potential, including free online writing and maths support, academic skills development and wellbeing consultations.

# Student Support

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

## The Writing Centre

The Writing Centre provides resources to develop your English language proficiency, academic writing, and communication skills.

- Workshops

- Chat with a WriteWISE peer writing leader

- Access StudyWISE

- Upload an assignment to Studiosity

- Complete the Academic Integrity Module

The Library provides online and face to face support to help you find and use relevant information resources.

- Subject and Research Guides
- Ask a Librarian

## Student Services and Support

Macquarie University offers a range of Student Support Services including:

- IT Support
- Accessibility and disability support with study
- Mental health support
- Safety support to respond to bullying, harassment, sexual harassment and sexual assault
- Social support including information about finances, tenancy and legal issues
- Student Advocacy provides independent advice on MQ policies, procedures, and processes

## Student Enquiries

Got a question? Ask us via AskMQ, or contact Service Connect.

## 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.