Te Kura Pūtaiao Koiora
School of Biological Sciences

Ngā whakamārama | Course outline 2021

BIOL209 – Biological Data Analysis

0.125 EFTS, 15 Points.
First Semester

Whakamahuki | Description

Introductory statistics with specific examples for biologists. This course is strongly recommended for all students in BIOL.

Whāinga Mahi | Goals of the Course

The overall aim of BIOL209 is to introduce you to presentation of results, statistical analyses and interpretation of experimental data, as they apply to biological research. The biological focus applies both to the choice of relevant methods and the specific examples discussed. The examples will cover a wide range of biology, from biochemistry to ecology, so that the course is applicable across all biological disciplines. One aim of the course is to prepare students for undergraduate analytical exercises, postgraduate research and jobs in research organisations. BIOL209 progresses from concepts of central tendency probability distributions, then on to hypothesis testing of various types.

Me whakaoti i mua | Pre-requisites

STAT 101 or 15 points of 100-level MATH

Relationship to other courses

BIOL209 will be of considerable value in analytical requirements for all other science courses in which data are collected. BIOL209 is also a pre-requisite for BIOL309 (Experimental Design and Data Analysis for Biologists), which is essential for students advancing to postgraduate study in biological sciences.

Kairuruku akoranga | Course coordinator

Dr Sarah Flanagan, von Haast 520, sarah.flanagan@canterbury.ac.nz

Pūkenga | Teachers

Assoc. Prof. Daniel B. Stouffer, von Haast 518, daniel.stouffer@canterbury.ac.nz
Prof. David R. Schiel, von Haast 245, david.schiel@canterbury.ac.nz

Hua Akoranga me ngā Aromatawai | Intended Learning Outcomes and Assessment

As a student in this course, I will develop the ability to:

Learning Outcome Number 1 (LO1)
A clear understanding of basic statistical principles (*assessment: lab quizzes, midterm test, final exam*)
Related Graduate Attributes and Kaupapa: Critically competent in the core academic discipline, Employable, innovative and enterprising

Learning Outcome Number 2 (LO2)
Proficiency in the transcription and manipulation of data (*assessment: lab quizzes, midterm test, final exam*)
Related Graduate Attributes and Kaupapa: Critically competent in the core academic discipline, Employable, innovative and enterprising

Learning Outcome Number 3 (LO3)
A basic understanding of a wide range of parametric and non-parametric statistical tests (*assessment: lab quizzes, final exam*)
Related Graduate Attributes and Kaupapa: Critically competent in the core academic discipline, Biculturally Competent and Confident (kaupapa 1), Employable, innovative and enterprising

Learning Outcome Number 4 (LO4)
Proficiency in the analysis of a wide range of biological data, including the ability to place the data in an appropriate context (*assessment: lab quizzes, midterm test, final exam*)
Related Graduate Attributes and Kaupapa: Critically competent in the core academic discipline, Employable, innovative and enterprising, Biculturally Competent and Confident (kaupapa 1, 3, 5)

Learning Outcome Number 5 (LO5)
Ability to use R to process and analyze data (*assessment: lab quizzes, final exam*)
Related Graduate Attributes and Kaupapa: Critically competent in the core academic discipline, Employable, innovative and enterprising

Pūkenga ngaio | Transferable skills

As a student in this course, I will develop the following skills:

- Understand statistical results presented in research papers and technical reports. The ability to critically evaluate and interpret statistical information is not only essential in higher-level courses but is a part of everyday life. *(assessment: lab quizzes, final exam)*
  Related Graduate Attributes and Kaupapa: Critically competent in the core academic discipline, Employable, innovative and enterprising
- Ability to apply basic concepts in exploratory data analysis. This ability is important for distinguishing between different types of data, methods of summarising data both graphically and through summary statistics. *(assessment: lab quizzes, midterm test, final exam)*
  Related Graduate Attributes and Kaupapa: Critically competent in the core academic discipline, Employable, innovative and enterprising
- Knowledge of the basics of collecting data and generating descriptive statistics. This skill is essential for all higher-level courses that include laboratory or field based research activities. *(assessment: lab quizzes, midterm test, final exam)*
  Related Graduate Attributes and Kaupapa: Critically competent in the core academic discipline
- Ability to apply the appropriate test and draw appropriate conclusions from the test output. This ability is important aspect of research and its application. *(assessment: lab quizzes, final exam)*
  Related Graduate Attributes and Kaupapa: Critically competent in the core academic discipline, Employable, innovative and enterprising
- Interpretation and communication skills. The ability to describe what the results mean in the context of the problem and being able to explain the results to someone else is essential for any professional career. *(assessment: lab quizzes, final exam)*
  Related Graduate Attributes and Kaupapa: Critically competent in the core academic discipline, Employable, innovative and enterprising

Aromatawai | Assessment:

20% Lab Quizzes (one per week, end of each lab, 2% each)
30% Midterm Test (on LEARN during week 4)
50% final exam (end of semester 1)

Lab Quizzes: The lab assessments will be run at the end of every lab. Each will ask you to run an example of something that has been practiced during the labs, and get you to enter a few key answers into LEARN. Each week is worth 2%, and we will drop your worst mark (count the best 9 of the 10 assessments).

Midterm Test: This test will be administered on LEARN and will consist primarily of multiple choice questions. It will cover the first three weeks of material, focusing on summarizing various types of probability distributions. You will be given a window of ~24 hours during which you will need to find time to sit the test, which will take 2 hours to complete.

Final exam: The final exam is cumulative, therefore the material in the first 3 weeks will be assessed in both the midterm and the final. Unless stated otherwise closer to exam time, you can expect it to be comprised of a combination of multiple choice and short answer questions.

A major challenge of BIOL209 for some students is keeping up with each topic without falling behind. The topics generally build on each other throughout the course and it can be difficult for students to catch up if they fall too far behind. This is one reason that we have both the weekly lab quizzes and the midterm test, to give you continual feedback on how you are doing. If you have problems with concepts, please discuss them with lecturers or lab demonstrators as soon as possible. Don't let your problems compound by falling even further behind in the course.

See below for departmental policies on late work, illness, and work that exceeds the length limits. Note that Biology policy says that to pass the course you need at least a 40% average across the interim work (lab assessments) and at least a 40% average in the final exam, AND at least 50% overall (see departmental policies below for more detail).

Tuhinga | Textbook

The recommended textbook is:

Crawley, M. J. (2015). *Statistics: an introduction using R.* 2nd edition. Cichester: John Wiley & Sons.

This is available from the University Bookshop and Amazon,
It can also be accessed via the library:
Option 1: Connect to electronic resource
http://canterbury.eblib.com.au/patron/FullRecord.aspx?p=1784599
Option 2: Connect to electronic resource Ebrary Academic Complete International Subscription Collection
http://site.ebrary.com/lib/uofcanterbury/Top?id=10931968
This book is user-friendly and pretty good at giving the applications and numerical recipes for most basic statistical routines.

However, you should be aware that no stats book seems to be good at explaining everything that might be helpful to a statistically-minded biologist; Crawley will not cover everything discussed during the course. In topics for which Crawley is inadequate, additional references will be recommended and put on restricted loan in the library.

Wā hui | Course Contact Hours

The course runs through the first semester (Terms 1 and 2) and consists of lectures, computer labs, and optional tutorials. **Check the UC timetable for time and room allocations** as these may change before term starts.

Lectures: There are two lectures per week, and we highly recommend that you attend these so that you have the opportunity to ask questions immediately as they arise.

Labs: There is one lab per week. In weeks 1 and 2, we will have shorter 1-hour labs so that we can have smaller streams and facilitate introducing the R statistical programming language. In the subsequent weeks, each lab will last for 2 hours. All labs will be led by Assoc. Prof. Daniel Stouffer unless noted otherwise.

Optional tutorials: These tutorials are optional drop-in question-and-answer sessions hosted by the demonstrators for the course.

ISO Week	Monday date	Lecture Number and Topic	Labs
8	22 Feb	**Daniel Stouffer** 1. R, variables, sampling 2. Distributions from permutations	1. Introduction to R
9	1 Mar	3. Discrete distributions 4. Continuous distributions	2. R and statistics
10	8 Mar	**Sarah Flanagan** 5. Sampling and central tendency 6. Exploratory data analysis & statistical inference	3. Plotting probability distributions
11	15 Mar	7. Experimental design 8. Hypothesis testing (one-sample t-tests)	4. Central tendency **MIDTERM**
12	22 Mar	9. Power & p-values 10. Normality testing & transformations	5. One sample tests
13	29 Mar	11. Chi-squared, Fisher's exact test 12. Covariance and correlation	6. Normality and transformations
Easter and mid-semester break			
17	26 Apr	ANZAC day [no lecture] 13. Dependencies and regression	7. Correlations
18	3 May	14. Regression assumptions, model fit 15. Regressions with groups	8. Regression 1
19	10 May	16. Comparing groups: parametric 17. Comparing groups: non-parametric	9. Regression 2
20	17 May	**David Schiel** 18. Anova introduction: Sums of Squares 19. Anova: multiple comparisons of means	10. Comparing groups
21	24 May	20. Anova: linear models 21. Anova: meeting assumptions, transformation	11. Anova 1 DRS
22	31 May	**Sarah Flanagan** 22. Revision and overview 1 23. Revision and overview 2	12. Anova 2 DRS
midyear break (**final exam**)			

Feedback from previous Course Surveys

The course has not had a proper course survey in the last couple of years (we hope to this year), but last year we received feedback in other formats. Below we summarize the changes implemented this year to respond to that feedback. During the semester, feel free to bring any positive or negative feedback to the lecturer/demonstrator, Sarah Flanagan as course coordinator, or anonymously through the Student Representative. Don't hesitate to get in touch if you have any problems about your work, or any personal difficulties. If a problem does arise, it is important to let one of us know as soon as possible so that corrective action can be taken quickly.

Issue identified from 2020	Change for 2021	Intention of change
The weekly lab quizzes implemented several years ago seem to work well, but they are worth too much of the final grade, so they caused the students undue stress (despite automatically dropping quiz grades).	Introducing a midterm worth 30%, reducing the final exam to 50% from 60% of the overall grade and reducing lab quizzes to 20% of the overall grade from 40% (each quiz will be worth 2%).	Make the quizzes a lower-stress incentive for the students to keep up with the work. The midterm will also serve as a better point of comparison for special considerations on the final exam than the quizzes.
One of the major criticisms students had for the course were related to the labs and the challenge presented by introducing R to students in a large computer lab without the ability to demonstrate the code in the computer lab. The 3 demonstrators and 1 academic staff in attendance at the labs are in high demand, so not all students felt as though they received adequate help. Students also did not seem to realize that many of the code solutions were in the lecture notes.	Have smaller but shorter labs the first 2 weeks (e.g., in weeks 1 and 2 have 6 1-hour streams of 30 students each rather than 3 2-hour streams of 60 students each).	These first two weeks will be focused on introducing students to R, and will be in classrooms where the teacher can live-code while demonstrators can walk through the class to help students who are struggling. The smaller class size will provide a lower student-teacher ratio, which should help facilitate student learning. In these initial labs the instructors can also demonstrate the expected approach to completing lab exercises (e.g., through troubleshooting).
The order of topics caused some student confusion, as hypothesis testing was introduced before the concept of distributions and probabilities.	Re-order the topics such that Daniel Stouffer will introduce probability distributions in weeks 1 and 2 and Sarah Flanagan will introduce experimental design and hypothesis testing after that.	A basic understanding of probability distributions is required to really understand the concept of p-values and hypothesis testing, so this order should improve learning.
Essentially no kaupapa were addressed in the course in 2020.	Incorporate more kaupapa (addressing #1, 3, 5, and 7) mainly through labs and use of additional datasets but also somewhat in lectures. Quizzes will feature reflective questions to assess this material.	Incorporating these kaupapa, associated with several learning outcomes and assessments in the course, will build on the bicultural competence and confidence concepts introduced in SCIE101 and provide second touchpoints for these kaupapa in the biological sciences curriculum.

RULES, REGULATIONS, AND WHAT TO DO WHEN THINGS GO WRONG

[updated 3 April 2020]

If in doubt: ASK! The course coordinator is happy to answer questions at any time. All staff involved in the course are available for advice on specific issues.

What do I do if I have to miss something or if my performance was impaired?

If you feel that **illness, injury, bereavement or other extenuating circumstances beyond your control** prevented you from completing an item of assessment worth 10% or more of the total course assessment or if these circumstances affected your performance in such assessments, you should apply for Special Consideration. Applications for Special Consideration should be submitted via the Special Consideration website http://www.canterbury.ac.nz/study/special-consideration/ *within five working days* of the assessment or its due date. You will also need to notify the course coordinator. If you apply for Special Consideration because of medical reasons, you should visit a doctor within a reasonable timeframe (application form available on the website above or from the Student Health Centre).

The Special Consideration provisions are intended to assist students who have covered the work of a course but have been prevented by illness or other critical circumstances from demonstrating their mastery of the material or skills at the time of assessment – they do not excuse you from doing the assessment within a reasonable time agreed with the course coordinator. You should expect to be required to submit additional work if you miss a major assignment (e.g. a field trip for which a major write-up is required).

You should also apply for Special Consideration if you are not be able to complete an assessment or attend a field trip because of **involvement in international or national representative sport or cultural groups**. Please review the Special Considerations policy, because very few kinds of activities will be eligible for such consideration (e.g. holiday trips, birthday parties etc. are not given special status in the University policy).

Students prevented by extenuating circumstances from completing the course after the final date for withdrawing, may apply for Special Consideration for late discontinuation of the course. Applications must be submitted via http://www.canterbury.ac.nz/study/special-consideration/ no later than five working days after the examination period has finished.

Plagiarism

It is essential that you are aware that plagiarism is considered a very serious offence by the academic community, the University and the School of Biological Sciences. Plagiarism is defined as taking content from another work or author and presenting it, without attribution, as if it is your own work. Content here includes text (sentences or major parts of sentences), display items (graphs and tables), and overall structure (the detailed sequence of ideas). Plagiarism includes:

- re-use of previous assignments (even if each individual sentence has been rephrased to say the same thing in different words, if the overall structure is re-used).
- copying of another student's work (with or without their consent).
- the unreferenced use of published material or material from the internet, e.g. cutting and pasting of paragraphs or pages into an essay.

For most pieces of in-term assessment you will be given information concerning the use of direct and indirect quotes from previously published work. If you have any doubt about the appropriate use of published material, please speak with an academic staff member. If you are unsure what plagiarism is, seek advice.

It is a School policy that courses may request that you submit work electronically for subsequent analysis of originality using *Turnitin*. Students agree that by taking courses in BIOL, assessments may be submitted to Turnitin.com for textual similarity review. All submitted papers will be included as source documents in the Turnitin.com reference database solely for the purpose of detecting plagiarism of such papers. Use of the Turnitin.com service is subject to the Terms and Conditions of Use as posted on the Turnitin.com site.

Where do I hand in assignments and then collect them once marked?

All assignments should be placed in the designated collection boxes in the foyer of the 2nd floor of the School of Biological Sciences (Julius von Haast building, at the top of the stairs), unless directed otherwise by the course coordinator. All assignments must be accompanied by a cover sheet signed by you stating that the submitted work is not plagiarised. Cover sheets are available on top of the collection boxes, or you can download one from the Biology website (http://www.canterbury.ac.nz/media/documents/science-documents/assignment-coversheet.pdf). In addition, you may also be asked to submit your work electronically (via Learn) for analysis in *Turnitin*.

Marked assignments can be collected from the School of Biological Sciences reception, unless directed otherwise by the course coordinator. Teaching staff will endeavour to return work as soon as possible, and should contact you if there are likely to be any delays that will prevent return within the maximum 4-week timeframe.

What if I can't get it finished in time?

Reports and assignments should be handed in on time. Extensions may be granted if you have a valid reason. **If you require an extension, you should request one from the course coordinator** (or the lecturer responsible for marking the work), with as much notice as possible. Please do this BEFORE the deadline for the assignment. **If you have been given an extension you should hand the work DIRECTLY to the course coordinator** (do not put it in the drop box as it may not be cleared after the due date).

If an extension has not been granted:

- work handed in within 1 hour of the deadline: penalty of up to 5 percentage points of the mark for the assignment (e.g., a mark of 75% might be reduced to 70%).
- work handed in 1 – 24 hours after the deadline: penalty of 10 percentage points of the mark for the assignment (e.g., a mark of 75% is reduced to 65%).
- work handed in 1 – 7 days after the deadline: penalty of 15 percentage points of the mark for the assignment (e.g., a mark of 75% is reduced to 60%).
- work handed in more than 7 days after the deadline will not be marked or earn credit.

What if I have written more than the word or page limit?

If there is a word limit on an assignment, it is usually there to stop you doing too much work and to encourage you to write succinctly. You can be up to 10% over without too much worry, but if the length increases beyond that your mark may suffer due to failure to follow the requirements. If you find yourself way over the word limit talk to the lecturer concerned about how to get your assignment to an acceptable length.

What if I fail part of the course?

In Biological Sciences, we require a satisfactory level of achievement in both the theoretical aspects of the discipline and in practical activities. This means you must attend all class activities and submit all items of assessment unless you have a very good reason not to (e.g. medical reasons). **A student must attain an average score of at least 40% for in-course assessments (e.g. assignments, reports) and an average score of at least 40% in the exam and/or tests, AND score at least 50% overall for the course, to be awarded a passing grade. See the course outlines for clarification of the assessment items included in each category and ask the coordinator if you are still unsure.**

What's the best way to give feedback?

We welcome constructive feedback at all times – help us to make this a valuable course for you. We endeavour to remain approachable at all times. If you would rather give feedback anonymously, please use the online course survey or talk to lab demonstrators, or your class rep (who will all report back to the staff-student liaison committee that includes a representative from each of the undergraduate classes). Class representatives will be selected from each class at the start of course.

What's the best way to complain?

If you feel you have not been fairly treated during this course, please raise the issue with the lecturer or course coordinator in the first instance. Other avenues include your class rep., who can raise issues anonymously, or the UCSA education coordinator.

Grading

A+	90% or above
A	85 – 90
A-	80 – 84
B+	75 – 79
B	70 – 74
B-	65 – 69
C+	60 – 64
C	55 – 59
C-	50 – 54

A restricted pass (R) **may** be awarded to those who are close to a pass (i.e. an overall score of 48-49.9%) AND who have achieved at least a 40% overall score in both in-course assessment and tests/exams. If an R grade is awarded you gain credit for the course but **cannot continue into papers that require this course as a pre-requisite**. NB. The R grade is only available at 100 and 200 level - it cannot be awarded for third year papers.

Failing grades: D 40-49 E 0–39