

STAT 221/231 - Statistics

Fall 2016

Course Syllabus

The Course Information Page:

learn.uwaterloo.ca You are expected to regularly read your UWaterloo email and visit the course website on Learn for announcements.

Instructors:

Sec	Instructor	Office	Email	Lecture Times
001	Cyntha Struthers	M3 3129	castruth	1:30-2:20MWF STP 105
002	Surya Banerjee	M3 3112	surya.banerjee	9:30-10:20MWF DC 1351

Course Description:

This course provides a systematic approach to empirical problem solving which will enable students to critically assess the protocol and conclusions of an empirical study including the possible sources of error in the study and whether evidence of a causal relationship can be concluded. The connection between the attributes of a population and the parameters in the named distributions covered in STAT 230 will be emphasized. Numerical and graphical techniques for summarizing data and checking the fit of a statistical model will be discussed. The method of maximum likelihood will be used to obtain point and interval estimates for the parameters of interest as well as testing hypotheses. The interpretation of confidence intervals and p-values will be emphasized. The Chi-squared and t distributions will be introduced and used to construct confidence intervals and tests of hypotheses including likelihood ratio tests. Contingency tables and Gaussian response models including the two sample Gaussian and simple linear regression will be used as examples.

Course Objectives:

- To provide students, who already have a basic understanding of probability, with an understanding of the role of variation in empirical problem solving and statistical concepts.
- To provide students with basic statistical methods to draw inferences from observed data.
- To provide students with the skills needed to understand, interpret, and critically evaluate statistical studies reported in the media, on the internet and in scientific articles.

Pre-requisites:

To take this course you need to have taken MATH 138 or 148 or 128 as well as STAT 230 or 240 or STAT 220 with a grade of at least 70%.

Background:

It is assumed that you have a solid understanding of Calculus 1 and 2 (i.e. MATH 128/138/148). This course relies heavily on the material from STAT 220/230/240. You are responsible for reviewing this material on your own. If your final mark in STAT 230/240 was below 60% you will find this course very challenging.

Learning Outcomes:

Upon successful completion of this course, you will be able to:

- use numerical and graphical summaries of a data set to describe the characteristics of a variate and to check the fit of a statistical model to the data;
- use the steps of PPDAC to identify both the objectives and possible sources of error in an empirical study, and to critically evaluate the conclusions;

- identify the connection between attributes of a population and the parameters in the named distributions (Binomial, Poisson, Multinomial, Exponential, Normal);
- define and use the likelihood function to obtain point and interval estimates of the unknown parameters in a model particularly for the Binomial, Poisson, Multinomial, Exponential and Gaussian (Normal) models;
- use a pivotal quantity to construct a confidence interval for a parameter and interpret the confidence interval;
- use the likelihood function to construct and conduct a test of hypothesis for an unknown parameter in a model;
- interpret p-values and describe the connection with confidence intervals;
- define the properties of the Chi-squared and t distribution;
- define a Gaussian response model including simple linear regression; determine point and interval estimates and conduct tests of hypotheses for the parameters in a Gaussian model;
- describe the importance of randomization and pairing in experimental design and be able to recognize whether a study design allows the researcher to conclude cause and effect;
- use a Goodness of Fit test to test the fit of a model, independence in a two-way table and equality of proportions for two or more groups

Course Notes:

STAT 221/231/241 Course Notes, Fall 2016 Edition are posted on Learn and are available at the University Bookstore. These Course Notes are designed to complement the material covered in lectures.

iClickers:

We will be using iClickers in lectures to engage students and provide instant feedback. New iClickers can be purchased at the UW bookstore in South Campus Hall for around \$40. You can also buy a used iClicker which is cheaper. STAT 231 does not support iClicker Go or clickers for mobile devices.

When using your iClicker in the classroom, make sure you have set the frequency correctly. The frequency **must** be set at the beginning of **every** class. The frequency code for the classroom is a two letter combination that is usually posted on the wall near the front of the room. After you cast your vote, the *Vote Status* light on the clicker should flash green.

You must attend the lecture section you are registered in to receive clicker marks. If you do not attend your registered lecture section, you will not get any clicker marks. You must register your iClicker through the course website on Learn (Course Materials->Course Content->iClicker Registration) so that your clicker grades can be assigned. If you have problems with your clicker marks, please email the Clicker TA (email information posted on Learn).

Answering clicker questions for another student is an academic offence.

Discussion Forum:

A great place to get help with course material that you do not understand is the discussion forum for this course on Piazza. It enables you to learn from the questions of others, and to avoid asking questions that have already been asked and answered. To join the Piazza forum for this course, go to <http://piazza.com/uwaterloo.ca/fall2016/stat221231>. Please post questions about course material to Piazza rather than emailing the instructors or the TA's.

If you have a question or concern of a personal nature (e.g. missed test due to illness), you should contact your instructor directly.

Information about office hours, material covered by tests, and other course information is posted on Learn. Contact information for questions about remarking of tests, clicker grades, etc. will also be posted on Learn. Therefore, if you post a question such as *"When are the TA office hours?"* or *"What material is covered by Midterm 1?"* or *"How did I get my test remarked?"*, the response will always be *"Please see information posted on Learn."*

Here are some guidelines for posting questions to Piazza:

- 1) Please remember that everything you post is public. Follow the Core Rules of Netiquette (<http://www.albion.com/netiquette/corerules.html>).
- 2) You can post questions anonymously to your classmates. To do so, click the Show up as Anonymous option. Note, however, that your instructors will be able to see your name.
- 3) Piazza is not a complaint forum, so please do not post complaints there. If you have a concern about anything to do with the course, the best way to deal with it, and to get results, is to contact your instructor directly.
- 4) Make it easy for other students to find your question in case they have the same problem. For example, in this course, if you have a question about end-of-chapter problem 4.8 (d), then please put *"Problem 4.8 (d)"* in the subject line and be sure to click on the appropriate folder. Posts are organized by chapters in the Course Notes and by tests.
- 5) Please search for an answer to your question before you post it. Remember that questions are organized in folders by chapters in the Course Notes and by tests. Piazza also has a very good search tool.
- 6) Keep your questions concise and specific. For example:
 - a) A post like *"Can someone walk me through the solution to Problem 4.3."* is not specific enough.
 - b) Please do not cut and paste large sections of text from the Course Notes or posted test solutions and post *"I do not understand this."*
 - c) Always indicate what you have already tried in order to solve the problem (e.g. *"I tried Problem 4.4 (c) and thought the solution should be blah, blah, blah but my solution is different from the solution in the Course Notes. Can anyone please tell me what is wrong with my solution? Thanks."*)
 - d) If you post a question, you should not expect a full solution to be posted in return. Instead, you may be given a hint to start the question or a suggestion to continue the question. The most learning occurs if you struggle a bit to solve a problem.

You can also use Piazza from your smart phone or tablet – see <https://piazza.com/product/mobile> for the iOS and Android Apps.

Office Hours:

Instructor and TA office hours will be posted in the Calendar on Learn. TA office hours take place in the SAS Tutorial Center. When there are many students waiting to ask questions, a time limit of 10 minutes per student will be used. Students should come to office hours with clear and well-organized questions.

Missed Lectures:

If you miss lectures then you are responsible for finding out from a classmate what you missed. Your instructor will not reteach the lectures you have missed.

Out-of-Class Workload:

As in any university course much of your learning in this course will take place outside of class time. Each week you have 3/4 hours of lectures/tutorial. Therefore you should plan to spend between 3 and 6 hours each week in out-of-class learning. This learning consists mostly of making sure you understand the concepts and steps that were used in class to solve problems and then solving problems from the Course Notes on your own.

Grading Scheme:

Every student in every section is treated the same way according to the grading scheme below. We cannot modify final grades to give you an extra percent – this would be unfair to the other students.

There are 2 grading schemes in order to minimize the impact of a poor performance on either of the 2 midterms:

Scheme 1

Clicker Questions – 5%
Assignments – 5%
Tutorial Tests – 15%
Midterm Test 1 – 15%
Midterm Test 2 – 15%
Final Exam – 50%

Scheme 2

Clicker Questions – 5%
Assignments – 5%
Tutorial Tests – 15%
Best Midterm Test – 15%
Worst Midterm Test – 5%
Final Exam – 60%

Students must write both midterm tests in order to qualify for this maximum scheme. (See below for how your final grade is determined if you miss any tests.) A student's final grade is the maximum of the two grades calculated using Scheme 1 and Scheme 2. If a grade of 0 is assigned to a test due to an academic offense then only Scheme 1 will be used. The maximum grade possible in this course is 100.

Assignments:

There will be 5 assignments in this course worth 1% each. The purpose of these assignments is to introduce you to the free statistical software R. Please see the course schedule below for due dates. Assignments will be submitted and returned using Crowdmark. Follow the steps in the *Introduction to R and RStudio* posted on Learn to install the software needed for this course (see Section 1 - Introduction).

Tutorial Tests:

Please check your schedule for your assigned tutorial time. There will be 3 Tutorial Tests worth 5% each held during your scheduled tutorial time. See the course schedule below. **You may only write your Tutorial Test in your assigned tutorial time.**

Tutorials:

During the weeks when there is no Tutorial Test, scheduled tutorials will be conducted by one of the course instructors. See the course schedule below. Examples will be presented in these tutorials to supplement the material covered in lectures. **Students are urged to attend these tutorials in order to be prepared for the tests and final exam.**

Midterm Tests:

There are two midterm tests:

Midterm Test 1: Thursday October 13, 2016, 4:40-6:10 pm.

Midterm Test 2: Tuesday November 15, 2016, 4:40-6:10 pm.

Final Examination:

A 2.5 hour final cumulative examination will be held during the final exam period, December 8-22.

STAT 221 registered students will have different tests from STAT 231 registered students. Details regarding the tests and final exam (e.g. material covered and locations) will be announced in class and posted on Learn.

Students must present a valid Student ID card to write all tests and final exam.

Calculator Policy:

For the tests and final exam, only a non-programmable, non-graphical, math faculty approved calculator with a **pink-tie** or **blue-goggles** sticker will be allowed.

Marking of Tests:

The tutorial and midterm tests will be graded and returned electronically using Crowdmark. The paper copies will be destroyed at the end of the examination period.

Remarking of Tests:

Information about how submit a request for remarking will be posted on Learn when your test papers are returned. **You will have 10 days to file a test remark request.**

Missed Tutorial Tests, Midterm Tests or Final Exam:

If you miss a test or final exam due to illness or extenuating circumstances you must provide proper documentation to your instructor within 48 hours of the missed test or final exam. In the case of illness you must provide a completed University of Waterloo Verification of Illness form.

See: <https://uwaterloo.ca/health-services/sites/ca.health-services/files/uploads/files/VIF-online.pdf>

In the case of extenuating circumstances you must provide sufficient documentation to your instructor to verify the circumstances.

Missed tutorial and midterm tests without proper documentation are automatically awarded a grade of 0.

If you miss both midterms (regardless of documentation) you will automatically receive a grade of DNW. Therefore you should withdraw from the course.

See <https://math.uwaterloo.ca/math/current-undergraduates/undergraduate-faq/dropping-courses>.

If you miss Tutorial Test 1 due to illness/extenuating circumstances with proper documentation then your Tutorial Test 1 grade out of 5 equals your Midterm Test 1 grade out of 5.

If you miss Tutorial Test 2 due to illness/extenuating circumstances with proper documentation then your Tutorial Test 2 grade out of 5 equals your Midterm Test 2 grade out of 5.

If you miss Tutorial Test 3 due to illness/extenuating circumstances with proper documentation then your Tutorial Test 3 grade out of 5 equals your Final Exam grade out of 5.

If you miss Midterm Test 1 or 2 (but not both) due to illness/extenuating circumstances with proper documentation then your Final Exam grade will count for 70% of your final grade.

If you miss Tutorial Test 1 and Midterm 1 (but not Midterm 2) due to illness/extenuating circumstances with proper documentation then your Final Exam will count for 75% of your final grade.

If you miss Tutorial Test 2 and Midterm 2 (but not Midterm 1) due to illness/extenuating circumstances with proper documentation then your Final Exam will count for 75% of your final grade.

If you miss the final exam due to illness/extenuating circumstances with proper documentation then the Mathematics Faculty INC Grade Policy (see below) will apply. Normally if you have not earned a passing grade on your term work and you do not write the final exam then you will receive a mark of DNW for the course.

University of Waterloo and Mathematics Faculty Policies

All instructors and students must follow the following academic policies:

Academic Integrity: In order to maintain a culture of academic integrity, member of the University of Waterloo community are expected to promote honesty, trust, fairness, respect and responsibility.

See: www.uwaterloo.ca/academicintegrity/ for more information.

Discipline: A student is expected to know what constitutes academic integrity to avoid committing an academic offence, and to take responsibility for his/her actions. A student who is unsure whether an action constitutes an offence, or who needs help in learning how to avoid offences (e.g., plagiarism, cheating) or about “rules” for group work/collaboration should seek guidance from the course instructor, academic advisor, or the undergraduate Associate Dean. For information on categories of offences and types of penalties, students should refer to Policy 71, Student Discipline, www.adm.uwaterloo.ca/infosec/Policies/policy71.htm.

See: www.adm.uwaterloo.ca/infosec/guidelines/penaltyguidelines.htm for guidelines for the assessment of penalties.

Avoiding Academic Offenses: For more information on commonly misunderstood academic offenses and how to avoid them, students should refer to the Faculty of Mathematics Cheating and Student Academic Discipline Policy. See: http://www.math.uwaterloo.ca/navigation/Current/cheating_policy.shtml.

Grievance: A student who believes that a decision affecting some aspect of his/her university life has been unfair or unreasonable may have grounds for initiating a grievance. See Policy 70, Student Petitions and Grievances, Section 4: <https://uwaterloo.ca/secretariat-general-counsel/policies-procedures-guidelines/policy-70>. When in doubt, please contact the department's administrative assistant who will provide further assistance.

Appeals: A decision made or penalty imposed under Policy 70 (Student Petitions and Grievances) (other than a petition) or Policy 71 (Student Discipline) may be appealed if there is a ground. A student who believes he/she has a ground for an appeal should refer to Policy 72 (Student Appeals).

See: www.adm.uwaterloo.ca/infosec/Policies/policy72.htm.

Mathematics Faculty INC Grade Policy: A grade of INC is awarded to a student who has completed course work during the term well enough that they could reasonably be expected to earn a passing mark in the course, but who was unable to complete end-of-term course requirements (usually the final exam) for reasons beyond his or her control. See: <http://www.math.uwaterloo.ca/navigation/Current/inc.procedure.shtml>.

AccessAbility Services:

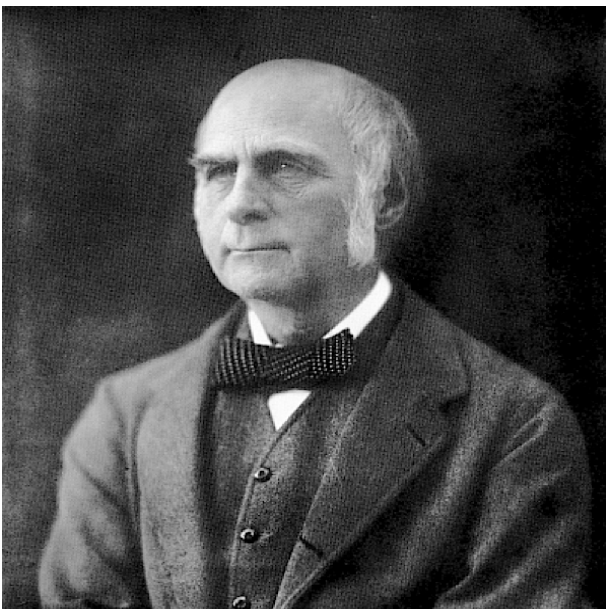
AccessAbility Services, located in Needles Hall, Room 1132, collaborates with all academic departments to arrange appropriate accommodations for students with disabilities without compromising the academic integrity of the curriculum. If you require academic accommodations to lessen the impact of your disability, please register with the AccessAbility Services at the beginning of each academic term.

Course Schedule:

The following table gives a tentative schedule for the material covered with sections in Course Notes indicated. The weeks for tutorials, tutorial tests, and midterm tests are also indicated. Reading Week is February 15-19. There is no class on Fri. Mar. 25 (Good Friday).

Week	Topics	Sec.	Notes
1. Sept. 8-14	Introduction to Statistical Sciences Collecting Data, Types of Studies Numerical Summaries	1.1 1.2 1.3	No Tutorial Sept.14 (STAT 230 Review)
2. Sept. 15-21	Graphical Data Summaries Probability Distributions and Statistical Models Data Analysis and Statistical Inference Statistical Software and R Chapter 1 Problems Choosing a Statistical Model Estimation of Parameters and the Method of Maximum Likelihood	1.3 1.4 1.5 1.6 1.7 2.1 2.2	Tutorial Wed. Sept. 21 Assignment 1 due Fri. Sept. 23 at noon
3. Sept. 22-28	Method of Maximum Likelihood Cont'd Likelihood Functions for Continuous Distributions Likelihood Functions for Multinomial Models Invariance Property of Maximum Likelihood Estimates Checking the Model	2.2 2.3 2.4 2.5 2.6	Tutorial Test 1 Wed. Sept. 28
4. Sept. 29-Oct. 5	Checking the Model Cont'd Chapter 2 Problems Empirical Studies The Steps of PPDAC Case Study Chapter 3 Problems	2.6 2.7 3.1 3.2 3.3 3.4	Tutorial Wed. Oct.5
5. Oct. 6-17	Statistical Models and Estimation Estimators and Sampling Distributions Interval Estimation Using the Likelihood Function	4.1 4.2 4.3	Study Break Oct. 11-12 Midterm 1 Thurs. Oct. 13
6. Oct. 18-24	Confidence Intervals and Pivotal Quantities The Chi-Squared and t Distributions Likelihood-Based Confidence Intervals	4.4 4.5 4.6	No Tutorial Wed. Oct.19 Assignment 2 due Fri. Oct. 21 at noon
7. Oct. 25-31	Confidence Intervals for Parameters in the $G(\mu, \sigma)$ Model Chapter 4 Problems Introduction to Tests of Hypotheses: Test Statistics and p-values	4.7 4.9 5.1	Tutorial Wed. Oct. 26
8. Nov. 1-Nov. 7	Tests of Hypotheses for Parameters in the $G(\mu, \sigma)$ Model Relationship between Hypothesis Testing and Confidence Intervals Likelihood Ratio Tests of Hypotheses – One Parameter Chapter 5 Problems	5.2 5.3 5.5	Tutorial Test 2 Wed. Nov. 2

9. Nov. 8-14	Gaussian Response Models Simple Linear Regression: Maximum Likelihood and Least Squares Estimates Confidence Intervals for Parameters in the Model	6.1 6.2	Tutorial Wed. Nov. 9 Assignment 3 due Friday Nov. 11 at noon
10. Nov. 15-21	Simple Linear Regression Cont'd: Prediction Interval for Future Response Checking Model Assumptions and Residual Plots Comparing the Means of Two Populations: Independent Samples from Two Gaussian Populations	6.2 6.3	Midterm 2 Tues. Nov. 15 No Tutorial Wed. Nov. 16
11. Nov. 22-28	Comparing the Means of Two Populations Cont'd: Paired Data Randomization, Pairing and Experimental Design Chapter 6 Problems Likelihood Ratio Test for the Multinomial Model Goodness of Fit Tests Two Way (Contingency) Tables	6.3 6.5 7.1 7.2 7.3	Tutorial Wed. Nov. 23 Assignment 4 due Fri. Nov. 25 at noon.
12. Nov. 29-Dec. 5	Testing Equality of Multinomial Parameters Chapter 7 Problems Establishing Causation Experimental Studies and Randomization Observational Studies and Simpson's Paradox Clofibrate Study Chapter 8 Problems	7.3 7.4 8.1 8.2 8.3 8.4 8.5	Tutorial Test 3 Wednesday Nov. 30 Assignment 5 due Mon. Dec. 5 at noon.



"Some people hate the very name of statistics, but I find them full of beauty and interest. Whenever they are not brutalized, but delicately handled by the higher methods, and are warily interpreted, their power of dealing with complicated phenomena is extraordinary. They are the only tools by which an opening can be cut through the formidable thicket of difficulties that bars the path of those who pursue the Science of man."

Francis Galton (1822-1911)