

# ARTSSCI 2R03: Applied Statistical Inference

## McMaster University

### Term 2, 2018-19

#### Course Objectives:

“Statistics is the science, the art, the philosophy, and the technique of making inferences from the particular to the general,” is a well-known quote attributed to John W. Tukey, one of the greatest statisticians of all times. With a strong footing on applications, this course demonstrates the statement repeatedly all along its path. The course objectives are to learn, exemplify and train the students on: (a) methods to summarize data numerically and graphically, (b) the most widely used statistical methods to draw inferences from observed data, (c) some of the mathematical details behind the methods, and (d) handling data and carrying out basic statistical analyses using the freeware statistical package R.

**Course Webpage:** We will use **Avenue To Learn** (<http://avenue.mcmaster.ca/>) as the main site for posting course-related material. In Avenue we will post assignments, midterms, solutions, R code, marks and other relevant information.

**Instructor:** Dr. Román Viveros-Aguilera, HH/212, ext. 23425, [rviveros@ms.mcmaster.ca](mailto:rviveros@ms.mcmaster.ca).

**Lectures:** Monday, Wednesday & Thursday, 1:30-2:20 pm, KTH/B132.

**Office Hours:** Wednesday, Thursday & Friday, 9:00-10:00 am.

#### Tutorials:

**T01:** Thursdays 12:30-1:20 pm in T34/104, *Tutor:* Ivana Massabki ([MASSABKI@MCMASTER.CA](mailto:MASSABKI@MCMASTER.CA)).

**T02:** Thursdays 11:30 am-12:20 pm in T34/103, *Tutor:* Adam Sapa ([SAPAA@MCMASTER.CA](mailto:SAPAA@MCMASTER.CA)).

**Software:** The computing work will be done in R, a freeware statistical software down-loadable for Windows and MacIntosh platforms from **The R Project for Statistical Computing**, webpage at <http://www.r-project.org/>. You are strongly advised to download and install the software at your earliest convenience.

**Textbook:** *Introductory Statistics - A Problem-Solving Approach*, by Stephen Kokoska, 2<sup>nd</sup> edition, 2015.

**Assignments:** There will be six assignments in the course. Make sure you write your complete name and student number on each page of every assignment. You are required to write well-organized and readable solutions. After marking, your assignments will be returned to you,

approximately one week after the due date. Solutions to the assignments will be posted in the course webpage. Assignment due dates:

**Assignment #1:** Thursday January 17, 3:30 pm.

**Assignment #2:** Thursday January 31, 3:30 pm.

**Assignment #3:** Thursday February 14, 3:30 pm.

**Assignment #4:** Thursday March 7, 3:30 pm.

**Assignment #5:** Thursday March 21, 3:30 pm.

**Assignment #6:** Thursday April 4, 3:30 pm.

The assignment questions will be posted in the course webpage one week prior to the deadline.

**Penalty for Late Assignments:** To be fair to all, no late assignments will be accepted. However, only the best 4 assignments will be used for final assignment average calculation.

**Midterm Tests:** There will be two midterm tests. Only the McMaster standard calculators **Casio FX-991 MS** or **MS Plus** will be allowed. The midterms are scheduled as follows:

**Test #1:** Thursday February 7, class hour, 50 minutes, room: UH/213.

**Test #2:** Thursday March 14, class hour, 50 minutes, room: UH/213.

Solutions to the midterm tests will be posted in the course webpage.

**Final Exam:** In April, 2.5 hours, scheduled by the Registrar's Office. The final exam period runs from Thursday April 11 to Monday April 29. The final exam will be based on all the material covered in the course. Only the McMaster standard calculators **Casio FX-991 MS** or **MS Plus** will be allowed.

**Practical Project:** Working in groups of 2 to 3 students, a required project will be conducted. Five datasets will be offered for statistical analysis and each team will pick one of them. General project parameters:

- (i) The datasets along with some description will be made available in advance.
- (ii) Each data set will contain at least 3 variables measured on each case and at least 50 cases.
- (iii) Your analysis will contain at least 3 plots.
- (iv) At least two statistical methodologies will be involved in your analysis.
- (v) Your report should contain at least 10 1.5-spacing pages (including plots).

The deadline to submit the project proposal is **3:30 pm on Friday March 1**, and for the submission of the final report is **3:30 pm on Friday March 29**. To be fair to all, no late project submissions will be accepted.

**McMaster University Research Ethics Board Statement:** Any student in the course, either individually or as part of a team, gathering and handling survey data within the McMaster community shall seek approval of their procedures from the McMaster University Research Ethics Board (MREB). The MREB is responsible for reviewing all non-medical research involving human participants carried out by McMaster faculty, students and staff to ensure the safety and wellbeing of human participants. MREB ensures compliance with Canada’s Tri-Council Policy Statement on the Ethical Conduct for Research Involving Humans. In addition the Board is responsible for educating the University community and setting University policies with respect to research involving human participants. Board members represent a broad range of disciplines and faculties, especially those in which non-medical research with humans takes place.

**Marking Scheme:** The final mark will be the highest between **Option 1** and **Option 2** below:

<b>Assignments</b> (6, due-dates: Jan. 17, 31, . . . . . 18% Feb. 14, Mar. 7, Mar. 21, Apr. 4)	18%
<b>Midterms</b> (2, dates: Feb. 7, Mar. 14) . . . . . 20%	30%
<b>Project</b> (deadlines – proposal: Mar. 1, . . . . . 10% report: Mar. 29)	10%
<b>Final Exam</b> . . . . . 52%	42%
<b>TOTAL</b> . . . . . 100%	100%

**Expectations:**

(a) **Come to class.** I know you will think that this is an obvious one but it is important nonetheless. Learning statistics is a cumulative experience; what we learn in one class builds upon what was covered in previous sessions. Missing class creates gaps in your knowledge base that will make hard to build on, and the degree of difficulty will increase rapidly.

(b) **Complete all assignments.** Much like the practice required in learning a foreign language, you will have to practice the techniques we learn in each class to get them to take hold and the homework is a good way to accomplish that. These exercises also will help you familiarize yourself with the software necessary to perform efficiently even simple analyses.

(c) **Ask questions.** Office and tutorial hours provide a great opportunity to ask for clarifications of the material and questions about your own attempts to solve a problem. When you are stuck, try to explain someone else (classmate friend, tutor, instructor) what you are trying to do. This can often help you identify for yourself the problem you are having. Often the step we are missing is readily obvious to someone who is not trying to solve the problem.

(d) **Be respectful of others.** Show your respect to your classmates by keeping a quiet and pleasant atmosphere in the classroom so that everyone focus their attention on the presentations and discussions.

These are my secrets for success about learning statistics, they come from my own personal experience as a student and as a teacher of the subject for over 25 years.

**McMaster Policy on Academic Integrity.** You are expected to exhibit honesty and use ethical behaviour in all aspects of the learning process. Academic credentials you earn are rooted in principles of honesty and academic integrity. Academic dishonesty is to knowingly act or fail to act in a way that results or could result in unearned academic credit or advantage. This behaviour can result in serious consequences—e.g., the grade of zero on an assignment, loss of credit with a notation on the transcript (notation reads: “Grade of F assigned for academic dishonesty”), and/or suspension or expulsion from the university. It is your responsibility to understand what constitutes academic dishonesty. For information on the various types of academic dishonesty, please refer to the [Academic Integrity Policy](#), located at: <http://www.mcmaster.ca/academicintegrity>. The following illustrates only three forms of academic dishonesty: 1) Plagiarism—e.g., the submission of work that is not one’s own or for which other credit has been obtained. 2) Improper collaboration in group work. 3) Copying or using unauthorized aids in tests and examinations.

**McMaster Student Absence Form (MSAF).** In the event of an absence, students should review and follow the Academic Regulations in the Undergraduate Calendar “Requests for Relief for Missed Academic Term Work.” Please consult the MSAF statement on our website (<https://arts.mcmaster.ca/forms-requests/>) and direct any questions or concerns to Shelley Anderson or Rebecca Bishop in the Arts & Science Program Office.

**Academic Accommodation of Students with Disabilities.** Students who require academic accommodation must contact [Student Accessibility Services \(SAS\)](#) to make arrangements with a Program Coordinator. SAS can be contacted by phone 905-525-9140 ext. 28652 or email [sas@mcmaster.ca](mailto:sas@mcmaster.ca). For further information, consult McMaster University’s [Academic Accommodation of Students with Disabilities](#) policy.

**Academic Accommodation for Religious, Indigenous, or Spiritual Observances (RISO).** Students requiring academic accommodation based on religious, indigenous, or spiritual observances should follow the procedures set out in the RISO policy. Students requiring a RISO accommodation should submit their request to their Faculty Office (i.e. to Shelley Anderson or Rebecca Bishop in the Arts & Science Program Office) normally within 10 working days of the beginning of term in which they anticipate a need for accommodation or to the Registrar’s Office prior to their examinations. Students should also contact their instructors as soon as possible to make alternative arrangements for classes, assignments, and tests.

**Email Contact and Student Responsibility Statement.** *Please Note:* The instructor and university reserve the right to modify elements of the course during the term. The university may change the dates and deadlines for any or all courses in extreme circumstances (e.g., severe weather, labour disruptions, etc.). Changes will be communicated through regular McMaster communication channels, such as McMaster Daily News, A2L, and/or McMaster email. It is the responsibility of students to check **their McMaster email** and course websites regularly during the term and to note any changes. Announcements will be made in class and by using the course email distribution list.

**Authenticity/Plagiarism Detection Statement.** In this course we will be using a web-based service (Turnitin.com) to reveal authenticity and ownership of student-submitted work.

Students will be expected to submit their work electronically either directly to Turnitin.com or via Avenue to Learn (A2L) plagiarism detection (a service supported by Turnitin.com) so it can be checked for academic dishonesty. Students who do not wish to submit their work through A2L and/or Turnitin.com must still submit an electronic and/or hardcopy to the instructor. No penalty will be assigned to a student who does not submit work to Turnitin.com or A2L. All submitted work is subject to normal verification that standards of academic integrity have been upheld (e.g., on-line search, other software, etc.). To see the Turnitin.com Policy, please go to [www.mcmaster.ca/academicintegrity](http://www.mcmaster.ca/academicintegrity).

## Course Schedule

Week 1: Jan. 7, 9, 10	Populations, samples, probability, statistics, observational and experimental studies random samples (Ch. 1); data types: categorical/qualitative and numerical/quantitative (Ch.2)
Week 2: Jan. 14, 16, 17	Bar charts, pie charts, stem-and-leaf plots, frequency distributions, histograms; skewness (Ch. 2); Introduction to R <b>Assignment 1 – Due Jan. 17</b>
Week 3: Jan. 21, 23, 24	measures of central tendency: sample mean and median; sample proportion, measures of variability: sample range, variance and standard deviation; quartiles, empirical rules, 5-point summary, standard and modified box plots (Ch. 3)
Week 4: Jan. 28, 30, 31	Random experiments, sample spaces and events; probability, counting methods, conditional probability, independence (Ch. 4) <b>Assignment 2 – Due Jan. 31</b>
Week 5: Feb. 4, 6, 7	Random variables, probability distributions for discrete random variables; mean, variance and standard deviation of a discrete random variable; binomial, geometric and Poisson distributions (Ch. 5) <b>Midterm 1 – Feb. 7 (Room: UH/213, class time)</b>
Week 6: Feb. 11, 13, 14	Probability distributions for continuous random variables, the normal distribution and normal probability calculations, checking for normality: the normal probability plot, the exponential distribution (Ch. 6) <b>Assignment 3 – Due Feb. 14</b>  <b>Mid-term Recess – Feb. 18-24</b>
Week 7: Feb. 25, 27, 28	Statistics, population parameters and sampling distributions; the sampling distribution of the sample mean and the Central Limit Theorem (CLT); the distribution of the sample proportion (Ch. 7) <b>Project proposal – Due Mar. 1</b>

Week 8: Mar. 4, 6, 7	Point estimation of population parameters, confidence intervals for a population mean $\mu$ when $\sigma$ is known, the case when $\sigma$ is unknown, confidence intervals for a population variance or standard deviation (Ch. 8) <b>Assignment 4 – Due Mar. 7</b>
Week 9: Mar. 11, 13, 14	The structure of hypotheses tests, choosing alternative hypotheses, hypothesis test errors, hypothesis tests about a population mean $\mu$ when $\sigma$ is known (Ch. 9) <b>Midterm 2 – Feb. 14 (Room: UH/213, class time)</b>
Week 10: Mar. 18, 20, 21	p-values, hypothesis tests about a population mean $\mu$ when $\sigma$ is unknown, large-sample tests concerning a population proportion, hypothesis tests about a population variance or standard deviation (Ch. 9) <b>Assignment 5 – Due Mar. 21</b>
Week. 11: Mar. 25, 27, 28	Simple linear regression, hypothesis tests and correlation, inferences about the mean value and an observed value of $y$ for given $x$ (Ch. 12) <b>Project report – Due Mar. 29</b>
Week 12: Apr. 1, 3, 4	Multiple linear regression, polynomial and qualitative regression models, model selection procedures (Ch. 12) <b>Assignment 6 – Due Apr. 4</b>
Week 13: Apr. 8	Review