JOHN ABBOTT CEGEP/COLLEGE

Statistical Methods

SCIENCE PROGRAM
COURSE OUTLINE
FALL 2024

General Information.
Discipline: Mathematics *Course code*: 201-DDD-05
Ponderation: 3-2-3 *Credits*: $2\frac{2}{3}$ *Prerequisite*: 201-NYA-05
Objectives:

- OOUV: To apply a scientific or technological approach to a field in the natural sciences
- OOUU: To apply knowledge and skills that have already been acquired to one or more topics in the natural sciences

Students are strongly advised to seek help from their instructor as soon as they encounter difficulties in the course.

Introduction. Statistical Methods is an option course in the Science Program. Normally taken in the fourth semester, this course, which is a fundamental branch of mathematics in its own right, introduces students to the collection, description and analysis of data.

One of the primary purposes of this course is the attainment of Objectives OOUV ("To apply a scientific or technological approach to a field in the natural sciences.), OOUU ("To apply knowledge and skills that have already been acquired to one or more topics in the natural sciences.) . To achieve this goal, the course will instruct the student how to apply the techniques of descriptive and inferential statistics to analyse data. The student will be introduced to grouped and ungrouped frequency distributions, probability distributions and sampling distributions. This will lead to the two main areas of inference: estimation and tests of hypothesis.

Statistical methods are used in a variety of disciplines. This course will emphasize application in the natural sciences.

This course can contribute to the Environmental Studies certificate. For more information, talk to the teacher or contact the certificate coordinator.

Methodology: This course meets three times a week for a total of five hours. The main techniques used will be the lecture and laboratory approaches. Other methods that may be used are: problem-solving sessions, class discussions and assigned reading for independent study. Regular homework should be expected. Students are responsible for all problems and exercises assigned by their teacher. The Math Study Area (H-200A and H-200B) functions both as a study area and as a centre where students may seek help with their mathematics courses. There are several computers equipped with statistical software and programs available for student use.

Comprehensive Assessment. The science program requires that every student passes a *Comprehensive Assessment*, which integrates the knowledge and skills acquired in the program. This is done through the science option courses.

This course includes a comprehensive assessment project. By obtaining a passing grade of 60% or higher on this project, the student will have successfully met the program requirement. Otherwise, the student will need to pass the comprehensive assessment in another option course.

Reference. There is no required textbook for this course. A set of exercises will be provided by your teacher.
A good reference for the course material is

> *Probability and Statistics for Engineering and the Sciences (9th edition)*, by Jay L. Devore
> http://www.nelsonbrain.com

Note that this book is not available for purchase at the bookstore.

Course Costs. A scientific, non-graphing, non-programmable calculator ($15-$20) is necessary.

Evaluation. A student's Final Grade for this course is a combination of the following components.

Class Tests	30%
Minor assessments (such as labs, assignments, quizzes)	15%
Project (*Comprehensive*)	15%
FINAL EXAM	40%

Students must be available until the end of the final examination period to write exams.

Note that in the event of unexpected changes to the academic calendar, the evaluation plan may be modified.

Other Resources.
Math Website.
http://departments.johnabbott.qc.ca/departments/mathematics

Math Study Area. Located in H-200A and H-200B; the common area is usually open from 8:30 to 17:30 on weekdays as a quiet study space. Computers and printers are available for math-related assignments. It is also possible to borrow course materials when the attendant is present.

Math Help Centre. Located in H-216; teachers are on duty from 8:30 until 15:30 to give math help on a drop-in basis.

Academic Success Centre. The Academic Success Centre, located in H-139, offers study skills workshops and individual tutoring.

College Policies.
Policy No. 7 - IPESA, Institutional Policy on the Evaluation of Student Achievement: https://www.johnabbott.qc.ca/wp-content/uploads/2021/05/Policy-No.-7-IPESA-FINAL.pdf.

Religious Holidays (Article 3.2.13 and 4.1.6). Students who wish to miss classes in order to observe religious holidays must inform their teacher of their intent in writing within the first two weeks of the semester.

Student Rights and Responsibilities: (Article 3.2.18). It is the responsibility of students to keep all assessed material returned to them and/or all digital work submitted to the teacher in the event of a grade review. (The deadline for a Grade Review is 4 weeks after the start of the next regular semester.)

Student Rights and Responsibilities: (Article 3.3.6). Students have the right to receive graded evaluations, for regular day division courses, within two weeks after the due date or exam/test date, except in extenuating circumstances. A maximum of three (3) weeks may apply in certain circumstances (ex. major essays) if approved by the department and stated on the course outline. For evaluations at the end of the semester/course, the results must be given to the student by the grade submission deadline (see current Academic Calendar). For intensive courses (i.e.: intersession, abridged courses) and AEC courses, timely feedback must be adjusted accordingly.

Academic Procedure: Academic Integrity, Cheating and Plagiarism (Article 9.1 and 9.2). Cheating and plagiarism are unacceptable at John Abbott College. They represent infractions against academic integrity. Students are expected to conduct themselves accordingly and must be responsible for all of their actions.

College definition of Cheating: Cheating means any dishonest or deceptive practice relative to examinations, tests, quizzes, lab assignments, research papers or other forms of evaluation tasks. Cheating includes, but is not restricted to, making use of or being in possession of unauthorized material or devices and/or obtaining or providing unauthorized assistance in writing examinations, papers or any other evaluation task and submitting the same work in more than one course without the teacher's permission. It is incumbent upon the department through the teacher to ensure students are forewarned about unauthorized material, devices or practices that are not permitted.

College definition of Plagiarism: Plagiarism is a form of cheating. It includes copying or paraphrasing (expressing the ideas of someone else in one's own words), of another person's work or the use of another person's work or ideas without acknowledgement of its source. Plagiarism can be from any source including books, magazines, electronic or photographic media or another student's paper or work.

OBJECTIVES	STANDARDS
Statement of the competency To apply statistical methods to collect and analyse data.	**General Performance Criteria** • Appropriate use of concepts • Correct algebraic operations. • Correct choice and application of statistical techniques. • Correct interpretation of results • Accurate calculations. • Proper justification of steps in a solution. • Appropriate use of terminology • Appropriate use of the computer.
Elements of the Competency 1. To describe data. 2. To calculate the probability of an event. 3. To compute probabilities using random variables and their distributions. 4. To derive sampling distributions. 5. To estimate parameters. 6. To test hypotheses. 7. To undertake an interdisciplinary project which integrates current learning and demonstrates competence in three specific goals of the exit profile at an advanced level (00UU).	**Specific Performance Criteria** *[Specific performance criteria for each of these elements of the competency are shown below with the corresponding intermediate learning objectives. For the items in the list of learning objectives, it is understood that each is preceded by: "The student is expected to . . . ".]*

Specific Performance Criteria	**Intermediate Learning Objectives**
1. *Description of a data set*	
1.1 Description of a Population, Sample, Parameter, Statistic	1.1.1. State the definition of a Population. 1.1.2. State the definition of a Sample. 1.1.3. State the definition of a Parameter. 1.1.4. State the definition of a Statistic.
1.2 Description of a variable	1.2.1. State the definition of a variable. 1.2.2. Differentiate between a discrete and a continuous variable. 1.2.3. Differentiate between a dependent variable and an independent variable. 1.2.4. Differentiate between a qualitative variable and a quantitative variable.
1.3 Description of data collection methods	1.3.1. State the definition of Sampling. 1.3.2. State the definition of an experiment. 1.3.3. Describe other data collection methods.
1.4 Description of types of Samples	1.4.1. Describe a simple random Sample. 1.4.2. Describe a stratified Sample 1.4.3. Describe a systematic Sample. 1.4.4. Describe a cluster Sample.

Specific Performance Criteria	**Intermediate Learning Objectives**
1.5 Graphical description of data	1.5.1. Construct – in tabular form – the distribution of a data set. 1.5.2. Construct a stem and leaf plot. 1.5.3. Construct a box plot. 1.5.4. Construct a frequency and relative frequency histogram. 1.5.5. Construct frequency, relative frequency and cumulative frequency polygons. 1.5.6. Construct Bar and Pie graphs.
1.6 Calculation of measures of central tendency (raw data)	1.6.1. Define mean, median, mode, midquartile and midrange. 1.6.2. Calculate the mean, median, mode, midquartile and midrange.
1.7 Calculation of measures of dispersion (raw data)	1.7.1. State definitions of and compute the range, mean absolute deviation, variance, standard deviation (std.), coefficient of variation and interquartile range
1.8 Computation of measures of location	1.8.1. Compute percentiles, deciles and quartiles. 1.8.2. Calculate the std. score (z-score).
1.9 Computations with grouped data	1.9.1. Approximate (estimate) the std. deviation of a sample.
1.10 Calculation of the least squares (regression) equation (bivariate data)	1.10.1. Plot a scatter diagram. 1.10.2. Calculate the regression equation. 1.10.3. Plot a graph of the regression equation. 1.10.4. Use the regression equation to predict a value of the dependent variable. 1.10.5. Analyze the residuals.
1.11 Calculation of the linear correlation coefficient (r)	1.11.1. State the definition of the linear correlation coefficient r . 1.11.2. Calculate the linear correlation coefficient.
1.12 Calculation of measures for a linear function of a variable	1.12.1. Define a linear function of a variable. 1.12.2. Calculate the mean of a linear function of a variable. 1.12.3. Calculate the variance and std. deviation of a linear function of a variable.
2. *To calculate the probability of an event*	
2.1 Definition of basic terminology	2.1.1. State the definition of probability. 2.1.2. Differentiate between classical, relative frequency and subjective probabilities. 2.1.3. Define outcomes, sample space and events.
2.2 Use of counting methods	2.2.1. State and apply the fundamental counting principle. 2.2.2. State and apply the Permutation and Combination rules.
2.3 Probability Rules	2.3.1. State and apply the conditional probability rule. 2.3.2. State and apply the multiplication rule. 2.3.3. State and apply the addition rule. 2.3.4. State and apply Bayes' Rule.
3. *Computation of Probabilities using random variables and their distributions*	
3.1 Description of a random variable	3.1.1. State the definition of a discrete random variable. 3.1.2. State the definition of a continuous random variable.
3.2 Computation of probabilities using a discrete random variable.	3.2.1. Define and compute the probability of a discrete random variable.
3.3 Computation and interpretation of the mean, variance and std. deviation of a discrete random variable (r.v.).	3.3.1. Define and calculate the mean of a discrete random variable. 3.3.2. Define and calculate the expected value of a discrete random variable. 3.3.3. Define and calculate the variance and std. deviation of a discrete r.v.
3.4 Determination of a mean, variance and std. deviation of a linear function of a discrete r.v.	3.4.1. Define a linear function of a discrete r.v. 3.4.2. Calculate and interpret the mean and variance of a linear function of a discrete r.v.
3.5 Explanation and application of Tchebychev's Theorem.	3.5.1. State and prove Tchebychev's Theorem. 3.5.2. Apply Tchebychev's Theorem to any arbitrary data set.
3.6 Calculation of probabilities, mean and variance of a binomial r.v.	3.6.1. Define a binomial r.v. 3.6.2. Define a binomial probability mass function (p.m.f.). 3.6.3. Calculate probabilities using the binomial p.m.f. 3.6.4. Compute the mean and variance of the binomial r.v.
3.7 Determination of probabilities, mean and variance of a hypergeometric r.v.	3.7.1. Define a hypergeometric r.v. 3.7.2. Define a hypergeometric p.m.f. 3.7.3. Compute probabilities using the hypergeometric p.m.f. 3.7.4. Compute the mean and variance of a hypergeometric r.v.
3.8 Determination of probabilities, mean and variance of a Poisson r.v.	3.8.1. Define a Poisson r.v. 3.8.2. Define a Poisson p.m.f. 3.8.3. Calculate probabilities using the Poisson p.m.f. 3.8.4. Compute the mean and variance of the Poisson r.v.

Specific Performance Criteria	**Intermediate Learning Objectives**
3.9 Determination of probabilities, mean and variance of a continuous r.v.	3.9.1. Define and compute the mean of a continuous r.v. 3.9.2. Define and compute the variance of a continuous r.v. 3.9.3. Calculate the probability of an event described in terms of a continuous r.v.
3.10 Calculation and application of probabilities for a normal distribution.	3.10.1. State the probability density function (p.d.f.) of a normal r.v. 3.10.2. State the mean, std. deviation and resulting p.d.f. 3.10.3. Use the std. normal tables to compute probabilities for a normal r.v. 3.10.4. Use the normal distribution to solve science-related problems. 3.10.5. State the conditions under which the normal distribution can be used as an approximation of the binomial/Poisson distributions. 3.10.6. Calculate probabilities using the normal approximation.
4. *Derivation and analysis of sampling distributions.*	
4.1 Determination of probabilities for a sampling distribution.	4.1.1. State the Central Limit Theorem (C.L.T.). 4.1.2. Determine – intuitively – the results of the C.L.T. 4.1.3. Use the C.L.T. to calculate probabilities of an event described in terms of the distribution of the sample means. 4.1.4. State the distribution of sample proportions. 4.1.5. Calculate the probability of an event described in terms of the distribution of sample proportions. 4.1.6. Use the t–distribution to calculate the probability of an event described in terms of the distribution of sample means calculated from small samples (population std. deviation unknown). 4.1.7. Use the chi–squared distribution to calculate the probability of an event described in terms of the distribution of the chi–squared statistic.
5. *Estimation of Parameters*	
5.1 Determination of point estimators.	5.1.1. State the definition of a consistent estimator. 5.1.2. State the definition of an unbiased minimum variance estimator (U.M.V.).
5.2 Calculation of a point estimate (single population).	5.2.1. Compute a point estimate for the mean of a population. 5.2.2. Compute a point estimate for the proportion of successes in a binomial population. 5.2.3. Compute point estimates for the variance and std. deviation of a population.
5.3 Calculation of a point estimate (two populations).	5.3.1. Determine a point estimate for the difference of two population means. 5.3.2. Determine a point estimate for the difference
5.4 Determination of confidence interval estimates (one population).	5.4.1. State the definition of the level of confidence $(1-\alpha)$. 5.4.2. Determine a confidence interval estimate for the population mean. 5.4.3. Determine a confidence interval estimate for the population proportion. 5.4.4. Determine a confidence interval estimate for the population variance.
5.5 Determination of confidence interval estimates (two populations).	5.5.1. Calculate a confidence interval estimate for the difference of two population means. 5.5.2. Calculate a confidence interval estimate for the difference of two population proportions. 5.5.3. Calculate a confidence interval estimate for a quotient of two population variances.
5.6 Determination of sample size.	5.6.1. Calculate the margin of error. 5.6.2. Compute the minimum sample size required to estimate the population mean. 5.6.3. Calculate the minimum sample size required to estimate the population proportion.
6. *Test of Hypothesis*	
6.1 Definition of basic terms.	6.1.1. Define the following terms – used in a test of hypothesis: Null hypothesis ; Alternative hypothesis ; Type I and Type II errors ; Test criteria ; Test statistic ; Level of significance $P-$value
6.2 Test of hypothesis about the population mean.	6.2.1. Perform a hypothesis test about the population mean (population std. deviation known). 6.2.2. Perform a hypothesis test about the population mean (population std. deviation unknown).
6.3 Test of hypothesis about the proportion of successes in a binomial population.	6.3.1. Perform a test of hypothesis about the population proportion (small sample). 6.3.2. Perform a test of hypothesis about the population proportion (large sample).

Specific Performance Criteria	**Intermediate Learning Objectives**
6.4 Test of hypothesis concerning the population variance/std. deviation.	6.4.1. Perform a test of hypothesis about the variance of a normal population. 6.4.2. Perform a hypothesis test concerning the std. deviation of a normal population.
6.5 Test of hypothesis about the difference of two population means.	6.5.1. Perform a hypothesis test about the difference of two population means – using two independent random samples. 6.5.2. Perform a hypothesis test about the difference of two population means – using two dependent samples.
6.6 Test of hypothesis about the quotient of two population variances.	6.6.1. Perform a hypothesis test concerning the quotient of two population variances – using independent random samples.
6.7 Test of hypothesis about the difference of two population proportions.	6.7.1. Perform a hypothesis test about the difference in two population proportions – using large independent random samples.
6.8 Test of hypothesis concerning multinomial proportions.	6.8.1. Perform a test of hypothesis about population proportions – using independent random samples.
6.9 Test of hypothesis about the regression coefficients.	6.9.1. Perform a hypothesis test about the slope of the regression line. 6.9.2. Perform a test of hypothesis about the intercept of the regression line.
6.10 Test of hypothesis about the linear correlation coefficient.	6.10.1. Perform a test of hypothesis about the linear correlation coefficient.
7. *Integration, Comprehensive Assessment and Exit Profile Goals*	
7.1 Recognition of the links between science, technology and the evolution of society.	7.1.1. Discuss the application of Statistical Methods to a relevant problem from science.
7.2 Development of a personal system of values.	7.2.1. Discuss any social or ethical aspect of the specific problem used in your Comprehensive Assessment.
7.3 Application of acquired knowledge to a new situation.	7.3.1. Demonstrate clearly the specific statistical techniques used in some problem from science.
7.4 Clear demonstration of the links between Statistics and at least one other science discipline.	7.4.1. Apply knowledge or skills that have been acquired to topic(s) in Physics, Chemistry or Biology.