STAT 2000

Basic Statistical Analysis II

Summer 2014

Calendar Description

(Formerly 005.200) The study of estimation and hypothesis testing procedures for means and proportions in one, two, and multiple sample situations, introduction to the analysis of variance; regression and correlation analysis; optional topics may include nonparametric procedures, design of experiments, probability models. Not to be held with *STAT 2001*. Prerequisite: *STAT 1000* (C) or *STAT 1001* (C) (or *005.100* (C)).

Teaching Philosophy and Goals

It is the desire of the Department of Statistics to present this course in a manner that emphasizes and illustrates the "real-world" aspects of statistical analysis. Whenever possible, we will attempt to bring real-life examples and data into the classroom. This will be done using, as appropriate, videotape clips showing statisticians at work, newspaper articles, in-class demonstrations and experiments, and the like. It should be noted that this is a 2000 level course. As such, a certain level of maturity is expected.

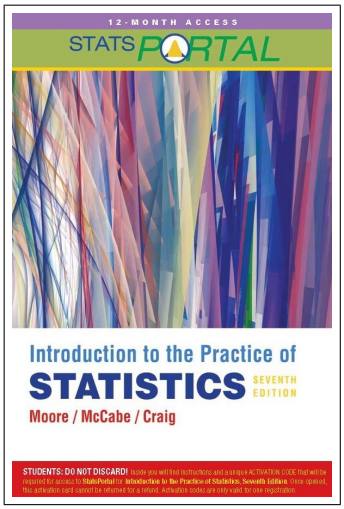
The course is designed to include those topics deemed crucial for an understanding of the foundations of statistical thinking and reasoning. The concepts of statistical analysis will be stressed rather than mathematical or probabilistic derivations or extensive numerical calculations. The course will place an emphasis on the development of critical thinking skills.

To aid in the analysis of data, extensive use will be made of the computer — with virtually every assignment involving the computer in some fashion. The computer package that has been selected for this course, JMP, is easy to use and is available for use with Macintosh or Windows systems. The package also has many advanced statistical features that you will find useful in subsequent courses.

We are interested in feedback from you. If you can think of ways in which this course could be improved, please let us know.

Text & Supplementary Material (Required)

Introduction to the Practice of Statistics (IPS), David S. Moore, George P. McCabe and Bruce A. Craig, 7th edition, W.H. Freeman, New York, 2012. There are two options for purchasing the required material:



Option 1: (ISBN 1-4292-2532-4) This option includes all required and supplementary materials for this course in electronic form, including the textbook. This may be a good option if you happen to have a copy of the book, or do not wish to have a hard copy, but would like access to the electronic supplements and the JMP software. This option must be asked for at the check-out counters in the bookstore.

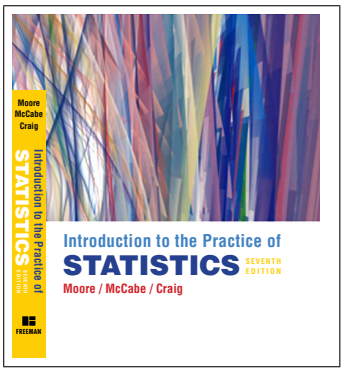
This includes the StatsPortal access card, which gives you access to the electronic version of the book, with associated tools such as StatTutor, the *Study Guide*, the JMP *Manual*, and access to the JMP software).

Option 2: (ISBN 1-4292-5231-6) This option includes all materials from Option 1 above, plus a hard copy of the textbook and a CD to accompany the book (containing statistical applets, tables, data sets, supplementary material and companion chapters).

Note that JMP software (included in both options above) is required for this course. There are many computers on campus that can be used for running JMP. In particular, the Department of Statistics has a number of Macintosh computers in the Statistics Lab (Room 311 Machray Hall) that you may use and the software is also available on the computers in the "open area" ACN computer labs. More details will be given in class.

Web Registration

This course requires you to register online with the Department of Statistics by going to http://www.stats.umanitoba.ca/register.



Here you will complete the online honesty declaration for this course, register to your i▶clicker, and signup for StatsPortal.

You will need the code in the StatsPortal package shown at the top right of this page. If you previously registered for StatsPortal (for either STAT 1000 or STAT 2000), you do not need to purchase StatsPortal again.

If you do not sign up for StatsPortal through the Department of Statistics, you **cannot** receive marks for your assignments.

i▸clickers

Throughout the course, extensive use of the i▸clicker classroom response system will be made in order to enhance your understanding of the material and promote classroom participation. Note that i▸clicker participation constitutes a portion of your grade in this course and as such you are required to bring your i▸clicker to each class.

The use of another student's i▸clicker constitutes impersonation and is strictly forbidden under the University of Manitoba's academic dishonesty policy. (See page 4.)

Mark Breakdown

Assignments	10%
i▸clicker Questions / Participation	5%
Term Test	35%
Final Examination	50%

Subject to the caveat below, the following are the minimum percentage grades required to receive each of the various letter grades: A+ (≥ 90), A (≥ 80), B+ (≥ 75), B (≥ 70), C+ (≥ 65), C (≥ 60), D (≥ 50).

To obtain a grade of C or better, you must obtain at least 50% on the final examination.

Assignments

The assignments in this course will be done in `StatsPortal`, using their online assignment system. There will be a total of five assignments. Your lowest assignment grade will not count towards your final grade. All assignments will be **due at 10:00 p.m.** on the respective due dates.

Assignment extensions will not be given to individual students.

i▸clicker / Participation

For every i▸clicker response that you give, you will be awarded 1 point. For questions with a correct answer, an additional point will be awarded for selecting the correct response. Full marks (5/5) will be given if you receive at least 75% of the total possible i▸clicker points. Partial marks (3/5) will be given if you receive between 50% and 75%. No marks (0/5) will be given if you receive less than 50%. You are responsible for bringing your i▸clicker to class and ensuring that it has functional batteries.

Test and Examination

The term test will be two hours in duration and will be held on Tuesday, May 20 from 8:30 a.m. to 10:30 a.m., as noted in the `Aurora Student Online Class Schedule`. The material to be covered will be announced in class. The final examination will be on Friday, May 30 from 9:00 am - 12:00 pm. three hours in duration as it is posted in Aurora system. It will cover the entire course with emphasis on material covered after the midterm test. Members of the University's (Provincial or National) athletic teams should normally give three weeks notice for special arrangements to be made for writing the term test or final examination off campus.

The term test will be all multiple-choice. The final examination will contain both multiple-choice questions and a written component, in an approximate 70:30 ratio. For the test and examination: (i) non-programmable hand-held calculators are permitted (graphing calculators are not permitted), (ii) electronic devices, such as cell phones and head phones, are prohibited, (iii) statistical tables will be provided, if required, (iv) selected formulae will be provided.

Demonstrator Hours

In the Statistics Lab in Room 311 Machray Hall (which contains a number of computers), graduate students and senior undergraduate students in statistics are available to help you at the following times (from May 05 until June 18):

Monday	10:30 a.m. – 12:30 p.m.
Tuesday	6:00 p.m. – 7:00 p.m.
Wednesday	10:30 a.m. – 12:30 p.m.
Thursday	6:00 p.m. – 7:00 p.m.
Friday	10:30 a.m. – 12:30 p.m.

Voluntary Withdrawal

Note that the voluntary withdrawal date is May 22, 2014 (by which time you will have received your marks for the test and several assignments).

Academic Dishonesty

It is important that you understand what constitutes academic dishonesty and that you are familiar with the very serious consequences. Links to resources that describe academic dishonesty (including plagiarism, cheating, inappropriate collaboration and examination impersonation) can be found at:
`http://umanitoba.ca/faculties/science/undergrad/resources/webdisciplinedocuments.html`.

Typical penalties imposed within the Faculty of Science for academic dishonesty are also described. See also:
`http://crscalprod1.cc.umanitoba.ca/Catalog/ViewCatalog.aspx?pageid=viewcatalog&catalogid=60&chapterid=227&topicgroupid=4056&loaduseredits=False`.

Course Content

The following is a non-exhaustive list of topics. Most of these are covered in the text.

Module I: Inference for the Mean of a Single Population when σ is Known or the Sample Size is Large; Inference for the Mean of a Single Population when σ is Not Known Chapter 6 (§6.1–§6.4) and §7.1

- Review of principles of statistical inference: testing and estimation, confidence intervals
- Statistical decisions: Type I and Type II errors and their probabilities, power of a test
- Review of t-distribution (comparison with normal distribution), tests and confidence intervals, robustness of t-procedures

Module II: Inference for the Means of Two Populations Chapter 7 (§7.1–§7.2)

- Review of procedures for matched pairs procedures
- Inference for the equality of means in two populations: assumptions of normality, independence and equality of variances

Module III: Inference for the Means of Two or More Populations Chapter 12 (§12.1)

- Inference for the equality of means in two or more populations: introduction to ANOVA, basic ideas of multiple comparisons
- The F-distribution
- Equivalence of pooled 2-sample t-test and F-test
- Graphical comparison of distributions

Module IV: Probability and Discrete Distributions Chapter 4 (§4.1–§4.5), §5.1

- Review of probability concepts and rules
- Conditional probability
- Random variables, probability distributions, mean and variance of a random variable
- Mean, variance and distribution of the sum and difference of two independent random variables
- Review of binomial distribution
- Poisson distribution

The material to be covered in the Term Test will be announced in class. The test is on
Tuesday, May 20
from 8:30 a.m. to 10:30 a.m.

Module V: Study of Attribute Data §8.2, Chapter 9 (§9.1–§9.4)

- Inference for a population proportion
- Inference for comparing two population proportions
- Inference for ($r \times c$) two-way tables: tests of independence and homogeneity of proportions, chi-square test, expected values, degrees of freedom
- Equivalence of Z-test and chi-square test
- Goodness of fit tests—binomial and Poisson

Module VI: Regression and Correlation Chapter 10 (§10.1–§10.2), Chapter 11 (§11.1–§11.2)

- Inference in simple linear regression (slope, y-intercept, confidence intervals, prediction intervals)
- Correlation: inference, correlation vs. regression
- Equivalence of testing for zero correlation and testing for zero slope
- Analysis of residuals and use of diagnostic tools
- Multiple regression

Module VII: Nonparametric Statistics

- Sign tests for median and paired data

* * * * * * * * * * * * * *

After STAT 2000

After you have completed STAT 2000, you may want to take further courses in statistics, or possibly become a statistician! So, what courses should you take? Here are some options for your next course:

If you are interested in entering an honours or major program in statistics, then you should take **STAT 2400 (Introduction to Probability)**, if you have not already done so. This course introduces the basic concepts of probability and provides a solid foundation for further courses in mathematical statistics. Applications of probability in many areas are provided, including forensic science, games of chance, genetics, reliability and statistical inference. The prerequisites for STAT 2400 are a grade of C or better in STAT 1000 and in one of MATH 1700 or MATH 1690. STAT 2400 is a required course in any honours or major program in statistics.

Without taking STAT 2400, you can also take a number of applied statistics courses. These are listed in the *University Calendar*. The only requirement to take these courses is a grade of C or better in STAT 2000.

2013–2014 REGISTRATION ADVISORY

Important Note from the Dean of Science:

It is your responsibility to ensure that you are entitled to be registered in this course. This means that you have:

- the appropriate prerequisites, as noted in the calendar description, or have permission from the instructor to waive these prerequisites;
- not previously taken, or are concurrently registered in, this course and another that has been identified as "not to be held with" in the course description. For example, STAT 1000 cannot be held for credit with STAT 2220.

The registration system may have allowed you to register in this course, but it is your responsibility to check. If you are not entitled to be in this course, you will be withdrawn, or the course may not be used in your degree program. There will be no fee adjustment. This is not appealable. Please be sure to read the course description for this and every course in which you are registered.

Selected Formulae for STAT 2000

1. $SE\left(\overline{x}_1 - \overline{x}_2\right) = \sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}$ with $df = \dfrac{\left(\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}\right)^2}{\dfrac{1}{n_1 - 1}\left(\dfrac{s_1^2}{n_1}\right)^2 + \dfrac{1}{n_2 - 1}\left(\dfrac{s_2^2}{n_2}\right)^2}$

2. $SE\left(\overline{x}_1 - \overline{x}_2\right) = \sqrt{s_p^2\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}$ with $df = n_1 + n_2 - 2$

   where $s_p^2 = \dfrac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$

3. $SSG = \sum_{i=1}^{I} n_i \left(\overline{X}_i - \overline{X}\right)^2$

4. $P(X = k) = \dfrac{e^{-\lambda}\lambda^k}{k!}, \quad k = 0, 1, 2, \ldots$

5. $SE\left(\hat{p}_1 - \hat{p}_2\right) = \sqrt{\hat{p}(1 - \hat{p})\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}$ if $p_1 = p_2$ where $\hat{p} = \dfrac{x_1 + x_2}{n_1 + n_2}$

   $SE\left(\hat{p}_1 - \hat{p}_2\right) = \sqrt{\dfrac{\hat{p}_1(1 - \hat{p}_1)}{n_1} + \dfrac{\hat{p}_2(1 - \hat{p}_2)}{n_2}}$ if $p_1 \neq p_2$

6. $SE_{b_1} = \dfrac{s_e}{\sqrt{\sum \left(x_i - \overline{x}\right)^2}}, \quad s_e = \sqrt{MSE} = \sqrt{\dfrac{\sum (y_i - \hat{y})^2}{n - 2}}$

7. $t = \dfrac{r\sqrt{n - 2}}{\sqrt{1 - r^2}}$

8. $SE_{\hat{\mu}} = s_e\sqrt{\dfrac{1}{n} + \dfrac{\left(x^* - \overline{x}\right)^2}{\sum \left(x_i - \overline{x}\right)^2}}$

9. $SE_{\hat{y}} = s_e\sqrt{1 + \dfrac{1}{n} + \dfrac{\left(x^* - \overline{x}\right)^2}{\sum \left(x_i - \overline{x}\right)^2}}$