STAT 1000

Basic Statistical Analysis I

Winter 2011

Calendar Description

(Formerly 005.100) An introduction to the basic principles of statistics and procedures used for data analysis. Topics to be covered include: gathering data, displaying and summarizing data, examining relationships between variables, sampling distributions, estimation and significance tests, inference for means. *Not to be held with STAT 1001, STAT 2220 (or the former 005.222)*. Prerequisite: Any grade 12 or 40S Mathematics, or equivalent.

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 statistical analysis arising from "real-world" applications. Whenever possible, we will attempt to bring real-life examples and data into the classroom. Upon completion of this course students can proceed in many directions: to further intensive study of statistics, to one or more additional courses in statistics, to the use of statistical methods in other fields of study, to being a consumer of statistical information in daily life. It is our objective to serve all of these diverse directions.

The course is designed to include basic topics deemed crucial for problem formulation and understanding of the foundations of statistical thinking and reasoning. The concepts of statistical analysis will be stressed. The course will place an emphasis on the development of critical thinking skills.

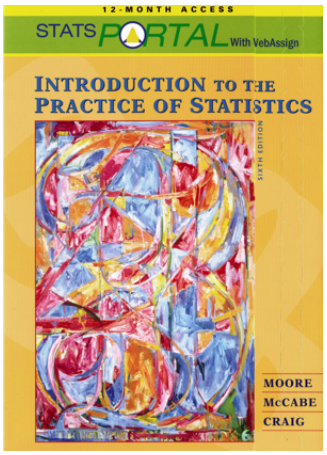
To aid in the analysis of data, extensive and intelligent use will be made of the computer — with virtually every assignment involving the computer in some fashion. The computer package that has been selected for this course, `JMP`, is easy to use and is available for use with Macintosh or Windows systems. 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.

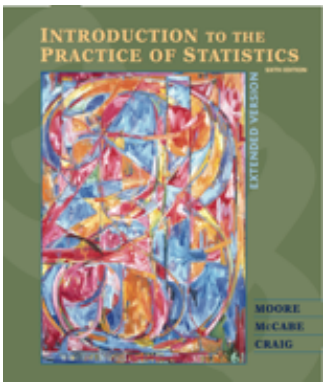
Text & Supplementary Material (Required)

Introduction to the Practice of Statistics (IPS), David S. Moore, George P. McCabe and Bruce A. Craig, 6^{th} edition, W.H. Freeman, New York, 2009. 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 following components: (i) the WebAssign access card (which gives you access to the WebAssign online homework system for one class within a 6-month period), (ii) the StatsPortal access card (which gives you 1-year 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).



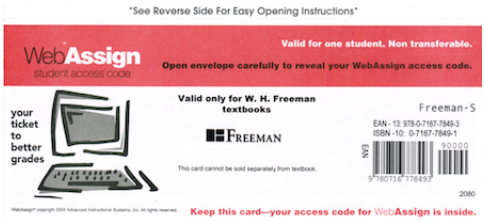
Option 2: (ISBN 1-4292-5231-6) This option includes all materials from Option 1 above, plus a hard copy of the textbook and a CD to accompany the book (containing statistical applets, tables, data sets, supplementary material and companion chapters).

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 access StatsPortal and WebAssign.

To signup for StatsPortal, go to http://courses.bfwpub.com/ips6e and click "Register". You will need the code in the StatsPortal package shown at the top right of this page. If you have previously registered in this course and you already have a StatsPortal account, you do not need to purchase a new one (select "Need to switch courses").



To sign up for WebAssign and register your i▸clicker, go to http://www.stats.umanitoba.ca/register. After your account has been created, you will require the code from the WebAssign ticket shown above. You must sign up for WebAssign *each semester* in which it is required.

For detailed signup instructions, please go to http://umanitoba.ca/statistics/signup.

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.)

Optional Material

Multiple-Choice Problems Set for Basic Statistical Analysis I, STAT 1000, Compiled by Dr. Smiley W. Cheng. (2010 Edition) The Department of Statistics, University of Manitoba; Winnipeg, Manitoba, Canada (ISBN: 0-5583-0526-1). You should find this "problems set" to be useful when studying for the term test and final exam. It contains term tests and final exams for this course from recent years. It also contains a collection of multiple-choice questions from previous introductory courses in statistics.

Mark Breakdown

Assignments	10%
i▸clicker Questions / Participation	5%
Term Test I	20%
Term Test II	20%
Final Examination	45%

Grade Cut-offs

Subject to the caveat in the paragraph 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%).

There is an **additional requirement** for obtaining a C or a D in the course: **to obtain a grade of C or better, you must obtain at least 50% on the final examination; to obtain a D you must obtain at least 40% on the final examination.**

Assignments

The assignments in this course will be done in `WebAssign`, an on-line assignment system. There will be 9 assignments. Assignments will be **due at 11:59 p.m.** on the due dates.

It is very important that you complete all 9 assignments. However, only the **best 8 of the 9** assignments will be used to determine your Assignment Mark. **Consequently, assignment extensions will not be given.**

It is important that you do lots of problems on a regular basis. For extra practice you should do questions from the text; the answers to most of the odd-numbered questions are given in the back of the book and many detailed solutions are given in the *Study Guide*.

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.

Tests and Examination

The first term test will be held **Saturday, February 12, 2011, from 9:30 – 11:30 a.m.** and will cover Units 1-4 in the course outline. The second term test will be held **Saturday, March 12, 2011, from 9:30 – 11:30 a.m.** and will cover Units 5-7 in the course outline. The second term test is **not** cummulative. The final exam will be 2 hours in duration and will be scheduled by the Student Records Office. The final exam will cover Units 1-11, with emphasis on Units 8-11.

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

Test and Exam Review Sessions

Three review sessions will be conducted prior to each test and the final exam. A course instructor will work through questions from sample tests and exams, which will be made available to you ahead of time. These sessions are free of charge and past students have found them very useful in preparing for tests and exams. More information will be given by your instructor in class.

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 January 10 until April 8):

Monday & Thursday	9:30 a.m.–4:00 p.m.
Tuesday & Wednesday	9:30 a.m.–7:00 p.m.
Friday	9:30 a.m.–11:00 a.m.

Note: The lab will be closed during reading week (Feb. 20–26).

Voluntary Withdrawal

Note that the voluntary withdrawal date is March 18, 2011 (by which time you will have received your marks for the first 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://www.umanitoba.ca/faculties/science/student/webdisciplinedocuments.html`.

Typical penalties imposed within the Faculty of Science for academic dishonesty are also described. See also:
`http://webapps.cc.umanitoba.ca/calendar10/regulations/plagiarism.asp`.

Course Content

The following is a non-exhaustive list of topics. Most of these are covered in the text. The course covers the material in Chapters 1–8. However, some material is omitted: decomposing time series (pp. 19–21), density estimation (p. 71), scatterplot smoothers (pp. 92–93), transforming relationships (pp. 119–121), residual plots (pp. 128–129), data mining (pp. 136–137), §2.5, capture-recapture sampling (pp. 220–221), §4.4, conditional probability and Bayes' rule (pp. 293–302) the continuity correction (pp. 326–327), the Weibull distributions (pp. 344–346), the bootstrap (p. 368), §6.4, the power of the t-test (pp. 433–435), inference for non-normal populations (pp. 435–440), inference for small samples (pp. 457–460), §7.3, plus-four confidence interval for a single proportion (pp. 491–493), §8.2.

Unit 1 – Examining Distributions — Chapter 1 (§1.1 and §1.2)

- types of variables: quantitative, categorical, continuous, nominal, ordinal
- graphs for categorical variables: bar charts, pie charts
- graphs for quantitative variables: stemplots, histograms
- examining distributions, dealing with outliers
- time plots
- describing distributions with numbers: mean, weighted mean, median, quartiles, percentiles, interquartile range, range, variance and standard deviation
- five-number summary and boxplots
- the $1.5 \times$ IQR rule for suspected outliers, modified boxplots
- resistant measures
- changing the units of measurement
- Introduction to `JMP` — a Statistical Computer Package

Unit 2 – Scatterplots, Correlation and Regression — Chapter 2 (omit §2.5)

- association, response variable, explanatory variable
- examining scatterplots
- adding categorical variables to scatterplots
- categorical explanatory variables
- correlation, properties of correlation
- least-squares criterion and least-squares regression line
- r^2
- residuals, outliers, influential observations
- cautions about correlation and regression
- association vs. causation, lurking variables
- extrapolation, causation, confounding, establishing causation

Unit 3 – Design of Experiments Chapter 3 (§3.1and §3.4)

- observations vs. experiment
- experimental units, subjects, treatments
- factors, levels
- placebo effect, control group, bias
- randomization, principles of experimental design
- statistical significance
- randomized comparative experiments
- matched pairs design, block design
- ethics

Unit 4 – Sampling Design Chapter 3 (§3.2)

- populations and samples
- voluntary response sample
- simple random sample
- probability sample, stratified random sample, multistage sample
- undercoverage, nonresponse, response bias

Term Test I covers material in Units 1-4
The test is on **Saturday, February 12, 2011**
from 9:30 – 11:30 a.m.

Unit 5 – Density Curves and Normal Distributions Chapter 1 (§1.3)

- continuous random variables, density curves
- normal distributions
- 68–95–99.7 rule
- standardizing observations (z-scores)
- normal distribution calculations
- use of normal quantile plots

Unit 6 – Randomness and Probability Chapter 4 (omit §4.4) & Chapter 5 (§5.1)

- randomness, the language of probability
- probability models, sample space, events, unions, intersections
- some probability rules, independence, general addition rule
- discrete random variables
- binomial setting and binomial distribution

Unit 7 – Sampling Distributions Chapter 3 (§3.3) & Chapter 5

- sampling distribution of a sample mean
- bias and variability
- Central Limit Theorem
- sampling distributions for proportions

Term Test II covers material in Units 5-7
The test is on **Saturday, March 12, 2011**
from 9:30 – 11:30 a.m.

Unit 8 – Confidence Intervals for a Population Mean (σ known) Chapter 6 (§6.1)

- margin of error
- effect of sample size, confidence level, standard deviation
- effect of population size
- assumptions
- choosing the sample size

Unit 9 – Tests of Significance Chapter 6 (§6.2 and §6.3)

- hypotheses, test statistic, P-value, statistical significance
- tests for a population mean (σ known)
- two-sided tests and confidence intervals
- use and abuse of tests

Unit 10 – Inference for One or Two Population Means (σ unknown) Chapter 7 (§7.1 and §7.2)

- one-sample t procedures — confidence intervals and tests
- matched pairs t procedures
- two-sample t procedure for independent samples
- robustness of the t procedures

Unit 11 – Inference for a Population Proportion Chapter 8 (§8.1)

- sampling distribution of the sample proportion
- confidence intervals and tests
- choosing the sample size

Final Examination covers material in Units 1-11, with emphasis on Units 8-11.

Unit 12 – Wrap-up & Looking Ahead

- warnings (non-random samples and inference, descriptive vs. inferential statistics, practical vs. statistical significance, etc.)
- re-iteration of some of the important ideas
- further courses in statistics, programs in statistics, careers in statistics
- a recently-introduced course in probability (STAT 2400)

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

2010–2011 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.

After STAT 1000

After you have completed STAT 1000, you may want to take further courses in statistics, or possibly become a statistician! So, what courses should you take? Here are the options for your next course:

STAT 2000 (Basic Statistical Analysis II) is a continuation of STAT 1000. This course, which is taken by students in many disciplines, covers more advanced statistical methods and leads to further courses in applied statistics. The only requirement to take this course is a grade of C or better in STAT 1000.

STAT 2400 (Introduction to Probability) 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 2000 and STAT 2400 are required courses for the honours or major program in statistics.