Number of ECTS credits: 6

One ECTS credit represents about 27.5 hours of work, so the workload for a 6 ECTS credits course is about 165 hours, or —spread over 15 weeks— 11 hours per week. This includes the class meetings (three hours per week), so I expect you to work for this class eight hours per week outside class.

Contact

Instructor: Luc Hens
Telephone: +32 2 629 11 92
e-mail: luc.hens@vub.ac.be. If you mail me concerning STA101, write STA101 Yourfirstname Yourlastname in the subject line.
Web site: http://homepages.vub.ac.be/~lmahens/. Follow the link to STA101. Contains the course syllabus, the assignments, the rules for written work, and more.
Office hours: Tuesdays 10:00–11:00, dialogue room 1, 2, 3, or 4, or meeting room (whichever is available). During office hours I will be pleased to speak with students without an appointment. If you would like to see me at another time, please arrange for an appointment before or after class or by e-mail.
Class: Tuesdays, Thursdays 08:30-10:00, room VeCo 2.

Course description

Statistics is the art of using data to make numerical conjectures about problems. Descriptive statistics is the art of summarizing data. Topics include: histograms, the average, the standard deviation, the normal curve, correlation. Much statistical reasoning depends on the theory of probability. Topics include: chance models, expected value, standard error, probability histograms, convergence to the normal curve. Statistical inference is the art of making valid generalizations from samples. Topics include: confidence interval for a mean and a percentage, tests of statistical significance.
Course prerequisites

STA101 has no college-level prerequisites. Math prerequisites do not go beyond the elementary algebra covered in Jacques (2009, or a more recent edition, Chapter 1: Linear Equations, pp. 13–92) (the textbook for MTH201 Mathematics for Business and Economics): graphs of linear equations (section 1.1); algebra (section 1.4); transposition of formulae (section 1.5). If you don’t master these concepts and skills, you should not take this course.

Learning objectives

This course aims at providing you with an understanding of inferential statistics (making valid generalizations from sample data), including multiple regression. At the end of the course, you should be able to:

- critically analyze data sets and apply the tools of statistics to data in order to improve decision making;

- use statistical software and a scientific calculator to do statistical computations (enter data, generate descriptive statistics and graphs, estimate population parameters, perform hypothesis tests)

- communicate the results of statistical work, and more specifically write up the results of statistical analysis in a report consisting of a non-technical abstract aimed at decision makers, so that they can improve their decisions, and a main section aimed at peers explaining the technical details and exact interpretation of the results. The report is formatted in APA Style.

Course materials


The Texas Instruments TI-84 calculator is required for this course.

For the assignments, you need access a computer with a word processor (such as LibreOffice Write, Pages, Microsoft Word, or the word processor of Google Drive). Download, print, and read the document Typesetting_math.pdf from the course web site. The document explains how to typeset mathematical copy in a word processor. Written work for my courses has to follow APA style. See my rules for written work (posted on the course home page) for details on how to format a paper for my classes. You can find most of what you have to know about APA Style in the textbook for the core Critical Thinking & Academic Writing classes: Bullock et al. (2014, pp. 158–191) covers APA Style, including an extended example.

You also need the open-source software R with the R Commander package. See Fox (2005), available from:

http://www.jstatsoft.org/v14/i09/paper

For more information about the hardware and software requirements, see:

STA101_Getting_started.pdf (posted on the course home page).
Additional Sources

Practice is important to learn statistics. Students who wish to work additional exercises can find hundreds of solved exercises in Kazmier (1995) (or a more recent edition). Moore et al. (2012) and De Veaux et al. (2014) cover the same ground as Freedman et al. (2007) and have many exercises; the answers to the odd-numbered exercises are provided in the book. These textbooks (sometimes in older but still useful editions) are available in the VUB library.

Dalgaard (2008) is a helpful guide to the R Statistical Environment beyond what is available in R Commander. You can do statistical calculations with R or with a statistical calculator like the TI-84, but also on the web site WolframAlpha. I posted a document that explains how to do statistical calculations in WolframAlpha on the course web site (http://homepages.vub.ac.be/~lmahens/STA201-wolfram-alpha.html).

Course assessment

STA101 uses a mix of interactive classroom lectures and in-class problem sets. Students are assessed on basis of four problem sets or short research papers and two written exams. The final grade for the course will be calculated on the following weighting:

- Assignments (four, each carrying 7.5 %) 30 %
- Midterm exam 35 %
- Final exam 35 %

Grading scale

The grading scale at Vesalius College, in line with the Flemish educational norms, is as follows:

<table>
<thead>
<tr>
<th>Grade</th>
<th>Grade points</th>
<th>Scale of 20</th>
<th>Scale of 100</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>4.00</td>
<td>17.0 to 20.0</td>
<td>85 to 100</td>
</tr>
<tr>
<td>A-</td>
<td>3.67</td>
<td>16.1 to 16.9</td>
<td>81 to 84</td>
</tr>
<tr>
<td>B+</td>
<td>3.33</td>
<td>15.3 to 16.0</td>
<td>77 to 80</td>
</tr>
<tr>
<td>B</td>
<td>3.00</td>
<td>14.5 to 15.2</td>
<td>73 to 76</td>
</tr>
<tr>
<td>B-</td>
<td>2.67</td>
<td>13.7 to 14.4</td>
<td>69 to 72</td>
</tr>
<tr>
<td>C+</td>
<td>2.33</td>
<td>13.1 to 13.6</td>
<td>66 to 68</td>
</tr>
<tr>
<td>C</td>
<td>2.00</td>
<td>12.3 to 13.0</td>
<td>62 to 65</td>
</tr>
<tr>
<td>C-</td>
<td>1.67</td>
<td>11.5 to 12.2</td>
<td>58 to 61</td>
</tr>
<tr>
<td>D+</td>
<td>1.33</td>
<td>10.7 to 11.4</td>
<td>54 to 57</td>
</tr>
<tr>
<td>D</td>
<td>1.00</td>
<td>10.0 to 10.6</td>
<td>50 to 53</td>
</tr>
<tr>
<td>F</td>
<td>0.00</td>
<td>00.0 to 9.9</td>
<td>00 to 49</td>
</tr>
</tbody>
</table>

Description of activities and grading criteria

**Attendance and preparedness:** Class attendance is crucial: statistics is a difficult subject, and students missing several classes rarely pass this course. In class, concepts are reviewed, you can ask questions, and you get to practice problems with immediate feedback. I take attendance at the beginning of
each class. If you can’t attend class for a valid reason, let me or the College administration know before class. Excused absences (doctors notice, or acceptable excuse at my discretion) will be counted as being present. Doctor’s notes should be handed in at the reception desk; the receptionist will inform all your professors.

**Written assignments:** Written work should follow APA Style. See my formatting rules for written work for details on how to format a paper for my class: [http://homepages.vub.ac.be/~lmahens//written_work.html](http://homepages.vub.ac.be/~lmahens//written_work.html)

The assignments are posted on the STA101 home page, with a detailed description of each assignment and the criteria used to assess your work (the rubric).

**Examinations:** Examinations provide the occasion for you to revise and consolidate your work during the term, to clear up any problems that have remained unresolved. They give you the chance to test whether you have really understood what you have read and done in class. The exams in STA101 typically consist four or five exercises from the textbook (Freedman et al., 2007), or exercises similar to those from the textbook. In the examinations, you have to show that you understand which assumptions were made when developing the techniques you use, and what the limitations of the techniques are. You should be able to use the TI-84 calculator to do statistical calculations. You can use the laminated formula sheets. Makeup examinations will be allowed only in extreme emergency, which must be documented by a physician or college official, in advance when possible.

**Further description of assessment activities**

The assignments are posted on the STA101 home page, with a detailed description of each assignment and the criteria used to assess your work (the rubric).

**Grading form for written assignment**

The assignments are posted on the STA101 home page, with a detailed description of each assignment and the criteria used to assess your work (the rubric).

**Additional Course Policies**

**Preparation for class:** Consult the STA101 home page regularly (at least once a week). Carefully read the materials indicated in the course schedule before coming to class. Statistics is a sequential subject: new topics build on concepts introduced before, so it is crucial to keep up with the material as we go along and to regularly review concepts. From time to time, you will work problems in class, sometimes individually, sometimes in small teams. I expect you to actively work the problems, and be prepared to briefly present the results of your work to the other students. You should bring to class: a mechanical pencil; an eraser; a pen; a ruler with a centimeter scale; the TI-84 calculator; the STA101 formula sheets (laminated); an A4 notebook with 5 mm squared paper for graphing on scale (notebooks of the *Atoma* brand allow you to easily
add, remove, and reorganize pages). You don’t have to bring your textbook to class.

**Late papers:** Written work should be submitted on paper at the beginning of class on the date specified in the course schedule. I don’t accept submission by e-mail. In accordance with Murphy’s Law, computers are likely to crash and printers are likely to run out of ink the evening before the due date. Start in time. Make backups of your work. Computer or printer problems are no valid reason for handing in work late. Late written work is allowed only in emergency cases, which must be documented by a physician or college official, in advance when possible. Otherwise the penalty for late work is 2 points (on a 20–point scale) per calendar day late. Hand in late assignments at the reception desk. The receptionist will put a date stamp on the paper.

**Returning the originals of written work:** During the semester, you should make photocopies of your graded written work (papers and midterm examinations) and return the originals to me (needed for inspection by the external examiners and the accreditation body).

**Final grades:** Don’t mail or call me to inquire about your final grade—I don’t communicate final grades by e-mail or by telephone. The college administration communicates the final grades.

**Academic Honesty**

Academic dishonesty is not tolerated in this course. Academic honesty is not only an ethical issue but also the foundation of scholarship. Cheating and plagiarism are serious breaches of academic integrity.

Following College policy, cheating and plagiarism cases will be communicated in writing to the Associate Dean for Students and submitted to the Student Conduct Committee for disciplinary action.

If you refer to someone else’s work, provide appropriate references and citations. Grammar, spelling and punctuation count, so use the tools necessary to correct before handing in assignments.

**Course schedule**

Consult the STA101 home page before every class meeting for announcements and possible schedule changes. Review the chapters that we already covered before coming to class. Work the listed exercises after we covered the chapter. I strongly encourage you to practice statistics by **working additional exercises** at home, individually or in group. The back of the textbook has solutions to all exercises except the review exercises. I’d be glad to help you when you get stuck: come and see me during office hours and bring your written preparation. The course schedule contains for each chapter a list of problems you should work on your own.
Course schedule

Consult the STA101 home page before every class meeting for announcements and possible schedule changes. Chapter numbers refer to Freedman et al. (2007). For each chapter, I added which exercises you should try to work at home after class (minimum and recommended exercises). We’ll skip chapters 10, 11, 12, 22, 24, 25, 27, and 28.

Week 1 (starting Monday 16 January 2017)


Controlled experiments and observational studies. Preface, Ch. 1, Ch. 2. Reading questions: (1) Give an example of treatment and response from your major. (2) How do the treatment group and the control group ideally differ? (3) What is a confounding factor? Give an example, referring to the case you cited for question (1). (4) What does it mean that an experiment is a randomized controlled experiment? (5) In your example from question (1), is it possible to give the control group a placebo? Why (not)? If not, why may this be a problem? (6) In your example from question (1), is it possible to make the experiment double blind? Why (not)? If so, why is this a good thing? If not, why may this be a problem? (7) Give an example of an an observational study from your major. (8) How does an observational study differ from a controlled experiment? (9) Give an example of a confounder in the observational study from question (7). (10) Can you control for the confounder from question (9)? If so, explain how. If not, explain why.

Download and print STA101_Getting_started.pdf (posted on the course home page) and follow the instructions.

(Chapter 2: Recommended: set A: 1, 5, 7, 14, 15. Review Exercises: 4, 9)

Week 2 (starting Monday 23 January 2017)

The histogram. Ch. 3. Reading questions: (1) How does a density histogram differ from a frequency histogram? (2) Why is a density histogram preferable over a frequency histogram? (3) Give examples from your major of: a qualitative variable that is nominal; a qualitative variable that is ordinal; a discrete quantitative variable; a continuous quantitative variable.

Thursday: An introduction to R and R Commander. Bring your laptop computer (if you have one) to class. Make sure R and R Commander are installed on your laptop computer, and that you have downloaded the file students.csv to your hard disk.

(Chapter 3: Minimum: Set B: 1 Set C: 1, 4. Recommended: Set A: all, Set B: 1, 2, Set C: all. Review Exercises: 1, 2, 4, 8, 10, 12.)

Week 3 (starting Monday 30 January 2017)

The average and the standard deviation. Ch. 4. Reading question: Give an example from your major of a situation where the median is a better measure of central tendency of data than the average.
Week 4 (starting Monday 6 February 2017)

Homework assignment 1 due on Thursday 9 February 2017 at the beginning of class.
The normal approximation for data. Ch. 5. Reading questions: (1) Billy’s height is (compared to the students in his class) equal to −2 standard units. Explain what this means. (2) Is Billy exceptionally short or tall, compared to the students in his class? If the heights of the students in Billy’s class are approximately normally distributed, which percentile is his height approximately? Explain what this means.

Week 5 (starting Monday 13 February 2017)

Measures of linear relationship: correlation. Ch. 7 (only sections 1 and 2, to be reviewed on your own), Ch. 8 (skip section 3 “The SD line”), Ch. 9. Reading questions: (1) Sketch a scatter plot in which the coefficient of correlation not a good measure of positive or negative association. (2) Give an example from your major of an ecological correlation. Explain why ecological correlations may be misleading. (Chapter 8: Minimum: Set B: 2, 6, 7 Set D: 1. Recommended: Set A: 5, 6. Set B: 1, 3, 4, 5. Review Exercises : 1, 3, 4, 5, 11.)

Week 6 (starting Monday 20 February 2017)

Homework assignment 2 due on Thursday 23 February 2017 at the beginning of class.
Probability. Ch. 13, 14. Reading question: Give an example from your major of a conditional chance.
(Chapter 13: Minimum: Set A: all Set B: all. Set C: 1–4, 7. Set D: 1, 2, 3, 4, 8. Recommended: Review Exercises : 2-9.)

Week 7 (starting Monday 27 February 2017)—midterm exam week

Midterm examination on Tuesday 28 February 2017. Covers all material covered to date. Bring your student ID, a mechanical pencil, an eraser, a pen, a ruler with a centimeter scale, the TI-84 calculator, a set of spare batteries for the TI-84, and the laminated formula sheets. Put everything (except the calculator and the formula sheet) in a transparent plastic resealable (Ziploc) bag.
(Thursday: I’ll take attendance, but you don’t have to read the chapter in
Week 8 (starting Monday 6 March 2017)

The law of averages. Ch. 16. Expected value and standard error. Ch. 17.
(Chapter 17: Minimum: Set A: 1, 3 Set B: 1, 2 Set C: 2 Set D: 1, 3. Set E: 1, 3. Recommended: Set A: 1–4 Set B: 1, 2, 3, 5. Set C: 1, 2, 3, 8 Set D: 1–4 Set E: 1–3, 9. Review Exercises : 1–4, 8, 11, 9.)

Week 9 (starting Monday 13 March 2017)

Friday of week 9 is the last day to drop a course.
The normal approximation for probability histograms. Ch. 18. Reading questions: What is the difference between an average and an expected value? Between a standard deviation and a standard error? Between a data histogram and a probability histogram?
(Chapter 18: Minimum: Set A: 1, 2 Set B: 2, 5 Set C: 2. Recommended: Set A: 1, 2 Set B: 1–3, 5. Set C: 2, 5–8. Review Exercises : 1, 2, 4, 6, 9, 10, 12-14.)

Week 10 (starting Monday 20 March 2017)

Homework assignment 3 due on Thursday 23 March 2017 at the beginning of class.
Sample surveys. Ch. 19. Also watch “Leading questions,” a fragment from Yes Prime Minister, series 1 episode 2 (The Ministerial Broadcast), BBC, 1986 (link on course home page) (video and transcript of the dialogue).
Chance errors in sampling. Ch. 20. Reading questions: (1) Using a problem from your major, make a Venn diagram to illustrate the concepts of population and sample. For this problem, give an example of a parameter and the corresponding statistic. (2) For the problem from question (1), give an example of selection bias; give an example of non-response bias. (3) For the problem from question (1), describe a procedure to draw a simple random sample. (4) Explain in plain English what a chance error in sampling is. Refer to the example from the previous questions. (5) Referring to your answer to (4), explain in plain English what the standard error measures. Refer to the example from the previous questions.
(Chapter 19: Minimum: Set A: 1-5 11, 12.)
(Chapter 20: Minimum: Set A: 1-5 Set B: 1-4.)

Week 11 (starting Monday 27 March 2017)

Friday 11 November 2016 is a legal holiday (Armistice day)—no class.
Spring Recess (3–14 April 2017): academic holiday—no class

Week 12 (starting Monday 17 April 2017)

Homework assignment 4 due on Thursday 20 April 2017 at the beginning of class.
The accuracy of averages. Ch. 23.

Week 13 (starting Monday 24 April 2017)
Tests of significance. Ch. 26, 29. Reading questions: (1) If an observed difference is statistically significant, does that necessarily mean that the difference is large for practical purposes? Explain using an example from your major. (2) If an observed difference is large for practical purposes, will it necessarily be statistically significant? Explain using an example from your major.

Week 14 (starting Monday 1 May 2017)
Tests of significance (continued). Ch. 26, 29.

Week 15 (starting Monday 8 May 2017) - Final exam week
Final exam (date and room to be announced). Covers all material since the midterm exam. Bring your student ID, a mechanical pencil, an eraser, a pen, a ruler with a centimeter scale, the TI-84 calculator, a set of spare batteries for the TI-84, and the laminated formula sheets. Put everything (except the calculator and the formula sheet) in a transparent plastic resealable (Ziploc) bag.

References
