Instructor: James Matis, matis@stat.tamu.edu
Blocker 404C, 845-3187

Office Hours: MWF 11-12; TuTh 8-9; or by appointment

Grader: TBA

Prerequisite: Math 141 or 166 or equivalent
(i.e., course in finite math with some probability).

Text: Principle of Biostatistics, 2nd ed., M. Pagano and K. Gauvreau

Notes: Notes are available on the web at stat.tamu.edu/~matis

Software: The course will use a statistical software package called SPSS. You are entitled to receive a copy of this package at no additional cost. Details are given at http://software.tamu.edu.

Exams: Two in-class exams will be given, tentatively on October 13 and November 10. The final exam will be given as scheduled on Monday, December 11, from 8-10 am. All exams will be closed-book, but you may bring in a page of formulas.

Make-up: A student may be excused from an exam only for one of the University-approved reasons (see www.tamu.edu). There will be a general, comprehensive make-up exam for students with approved absences at the end of the semester.

Homework: Homework will be assigned on a regular basis. It will be graded and returned. Late homework will receive half credit.

Project: An individual project is an integral part of this course. The objectives of this project include:
1) obtaining data, either from your own research or from some other scientific paper from the published literature,
2) applying some statistical procedure introduced in the text,
3) writing an article which addresses the scientific issue and presents your statistical analysis, and
4) giving an oral report of your findings.

Grading: The weighting to determine the grade is:

<table>
<thead>
<tr>
<th>Component</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project</td>
<td>25%</td>
</tr>
<tr>
<td>3 Exams @ 20%</td>
<td>60%</td>
</tr>
<tr>
<td>Homework</td>
<td>15%</td>
</tr>
</tbody>
</table>

However, a “bounce-back” weighting, where the final accounts for 40% of the grade, with the project, regular exams and homework counting for 20%, 30%, and 10%, respectively, will also be calculated. You will receive the higher of the two weighted scores.


• **ADA, Plagiarism, and Academic Integrity Statement:**

TAMU STATEMENT ON DISABILITIES: The Americans with Disabilities Act (ADA) is a federal anti-discrimination statute that provides comprehensive civil rights protection for persons with disabilities. Among other things, this legislation requires that all students with disabilities be guaranteed a learning environment that provides for reasonable accommodation for their disabilities. If you believe you have a disability requiring an accommodation, please contact the Office of Support Services for Students with Disabilities in Room 126 of the Koldus Student Services Building. The phone number is 845-1637.

TAMU STATEMENT ON PLAGIARISM: The handouts used in this course are copyrighted. By “handouts,” I mean all materials generated for this class, which include but are not limited to syllabi, quizzes, exams, lab problems, in-class materials, review sheets, and additional problem sets. Because these materials are copyrighted, you do not have the right to copy the handouts, unless I expressly grant permission. As commonly defined, plagiarism consists of passing off as one’s own ideas, words, writing, etc., which belong to another. In accordance with this definition, you are committing plagiarism if you copy the work of another person and turn it in as your own, even if you should have the permission of that person. Plagiarism is one of the worst academic sins, for the plagiarist destroys the trust among colleagues without which research cannot be safely communicated. If you have any questions regarding plagiarism, please consult the latest issue of the Texas A&M University Student Rules, under the section “Scholastic Dishonesty.”

ACADEMIC INTEGRITY STATEMENT: “An Aggie does not lie, cheat, or steal or tolerate those who do.” For further information on academic integrity, see Honor Council Rules and Procedures at [www.tamu.edu/aggiehonor](http://www.tamu.edu/aggiehonor).
<table>
<thead>
<tr>
<th>Topic</th>
<th># Lectures</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Data Presentation – data types, tables, graphs (2.1–2.4)</td>
<td>2</td>
</tr>
<tr>
<td>2. Summary Statistics – central tendency, dispersion, Chebychev’s inequality and empirical rule (3.1, 3.2, 3.4)</td>
<td>2</td>
</tr>
<tr>
<td>3. Probability Distributions – theory, binomial, normal (7.1, 7.2 7.4, 7.5)</td>
<td>2</td>
</tr>
<tr>
<td>4. Sampling Distribution of the Mean – concept, central limit theorem, application (8.1–8.4)</td>
<td>2</td>
</tr>
<tr>
<td>5. Confidence Intervals – two-sided, one-sided, Student’s (9.1–9.4)</td>
<td>2</td>
</tr>
<tr>
<td>6. Hypothesis Testing – general, types of errors, power, sample size (10.1–10.7)</td>
<td>3</td>
</tr>
<tr>
<td>7. Comparing Two Means – paired t, two-sample t (11.1–11.3)</td>
<td>2</td>
</tr>
<tr>
<td>8. Correlation – scatter plots, correlation coefficients (17.1–17.4)</td>
<td>2</td>
</tr>
<tr>
<td>9. Simple Linear Regression – concepts, model, evaluation (18.1–18.4)</td>
<td>4</td>
</tr>
<tr>
<td>10. Inference on Proportions – normal approximation, single proportion concepts, comparing two proportions (14.1-14.7)</td>
<td>3</td>
</tr>
<tr>
<td>11. Contingency Tables – chi-square test, odds ratio (15.1, 15.3)</td>
<td>3</td>
</tr>
<tr>
<td>12. Analysis of Variance – one-way ANOVA, multiple comparisons (12.1-12.5)</td>
<td>3</td>
</tr>
<tr>
<td>14. Multiple Regression – model, indicator variables, interactions</td>
<td>2</td>
</tr>
<tr>
<td>15. Rates – (4.1–4.3)</td>
<td>2</td>
</tr>
<tr>
<td>16. Life Tables – computation, application, years lost (5.1–5.4)</td>
<td>2</td>
</tr>
</tbody>
</table>
17. Sampling Theory – simple random sampling, stratified sampling, cluster sampling (21.1)

Milestones for Course Project

I. Pre-proposal, due October 4.
   1.1 Select some topic and clear it with the instructor.
   1.2 Obtain copy of the scientific paper.
   1.3 Read paper to decide if you want to pursue the topic.
   1.4 Write a paragraph describing the general scientific problem.
   1.5 Hand in the paragraph with a copy of the first page of the article as an attachment. Alternatively, describe your original research problem and data.

II. Proposal, due November 1. (20% of grade). This is a written document (not to exceed 5 pages) that:
   2.1 Introduces the scientific problem,
   2.2 Discusses the previous literature (at least one citation),
   2.3 Discusses claims of the paper you selected,
   2.4 Describes data (using at least one numerical descriptive measure and at least one graph), and
   2.5 Discusses specific aspect of the problem that you will address using some statistical inference.

III. Review Paper, due November 29. (50% of grade). This paper should be in the form of a scientific paper, not to exceed 10 pages exclusive of Appendix. One common outline you could follow is:
   3.1 Introduction – scientific problem, previous literature, claims of paper, your scientific question.
   3.2 Methods – data, statistical methodology.
   3.3 Results – your statistical analysis.
   3.4 Discussion – interpretation of your findings.
   3.5 Conclusions (optional).
   3.6 References.
   3.7 Appendix (with at least a copy of the original paper).

IV. Oral Presentation. (30% of grade). This should be a brief (not to exceed 7 minutes) oral presentation of your results.
Possible Project Topics

1. Any suitable topic associated with an original data set that is amendable to suitable statistical analysis. Many students are now, or have recently been, involved with research problems and have access to original data.

2. Any scientific article that you may be familiar with which has suitable data and statistical analysis.


4. Other possible papers from reference list supplied by the instructor.

5. Projects described on forthcoming DVD’s, including 1) density of Africanized bees in Texas and 2) aphid infestation in Texas.
Additional Articles


Statistics 302 H

Introduction

• Statistics is a body of concepts and methods to
  1) collect and describe (numerical) data, (Ch. 2-5)
  2) make inferences from data (Ch. 8-22)

• Statistical inferences are based on laws of probability (Ch. 6-7)
  e.g. birthday game

• Basic Concepts:
  population – complete set of individuals, objects or measurements (i.e. ‘elements’) of interest.
  e.g. all undergrads at TAMU
  variable – characteristic of interest concerning an element
  e.g. $X = \text{wt}$, $Y = \# \text{ dogs}$, $Z = \text{gender}$
  sample – a subset (usually small) of the population.
  e.g. students in Stat 302 H.

• Objectives:
  1) describe elements in population (or sample)
  2) make inferences from sample to population.
Assignment 1
Due Monday, September 4

Statistics has been defined as a body of concepts and methods to 1) collect and describe numerical data and 2) to make inferences from data. Chapters 2-5 of our text addresses data description and Chapters 8-22 discuss statistical inference.

An outstanding article which illustrates both concepts is Cobb, G. and Gehlbach, S. (2005). Statistics in the Courtroom: United States v. Kristen Gilbert, in Statistics: A Guide to the Unknown, 4th ed. R. Peck and others, eds. Duxbury, Belmont, CA. A copy is attached. The article illustrates how one sets up hypotheses, and tests them through the use of \( p \)-values. We will conduct many such tests through the semester, and you may want compare this article to the discussion of Hypothesis Testing and \( p \)-value on p. 232-234 of our text.

Your assignment is to read this fascinating article and answer questions 1-4 at the end of the article. We will encounter numerous other very relevant hypotheses of interest throughout the semester.