

Statistics 601 – Statistical Analysis

Course Information Fall 2019, Sections 600 and 700

STAT 601 is designed to introduce students to statistical methods and thinking, particularly in ways useful to research in engineering and science. No prior statistical experience is required. Emphasis is placed on calculus-based probability, modeling data, choosing the analysis appropriate for the data, understanding sampling error and how to manage it, knowing the assumptions made and checking their validity. The methods covered in depth include multiple regression, basic experimental designs and the analysis of categorical data.

Time and Place: MWF 12:40pm–1:30pm and Tue 12:45pm–1:45pm, Blocker 457. *Section 700 students will be able to view lectures online, usually within a few hours although it could be the following day. Class time will be 12:45pm–2:00pm for midterm exams.*

Instructor: Prof. Daren Cline, Blocker 459D, 979-845-1443.
e-mail: dcline@stat.tamu.edu

Office Hours: MWF 10:20am–11:20am, or by appointment. ([my schedule](#))

Distance Section: Section 700 students are *required* to view lectures online, turn homework in online, arrange a proctor for exams, and take exams under the supervision of the proctor. The deadline for arranging a proctor is Friday, 6 September. Details will be posted in eCampus.

An online resource (probably eCampus) will be used to upload completed homework and to download/upload exams.

Please see the Statistics Department [Online Learning Program](#) for further details.

This includes all local students (not in an online degree program) registered for the distance section.

eCampus: <http://ecampus.tamu.edu> Course materials including recorded lectures, lecture notes and homework assignments will be available to *all students* at this site. A discussion board open to all students will also be maintained in eCampus.

Grader: Abhisek Chakraborty. e-mail: tba@stat.tamu.edu, office hours: see eCampus. You may contact the grader for help on homework or questions on graded homework. Requests for late homework (with valid university excuses) and questions about the exams should be made directly to me.

The grader will not accept homework from students under any circumstances; please bring it to me.

Help Sessions: TBA. TAs (not only your grader) will be available to assist students attending this and other similar courses.

The computers in the room are for statistics class use only.

Statistics 601 – Statistical Analysis

- Textbook:** D.C. Montgomery and G.C. Runger, *Applied Statistics and Probability for Engineers*, 7th ed., Wiley. ([supplemental material](#))
Please see the [course outline](#) below to correlate the textbook with the lectures.
- References:** ([on reserve in Evans Library](#))
J.S. Milton and J.C. Arnold, *Introduction to Probability and Statistics: Principles and Applications for Engineering and the Computing Sciences*, McGraw-Hill.
A.C. Tamhane and D.D. Dunlop, *Statistics and Data Analysis: from Elementary to Intermediate*, Prentice-Hall.
F.L. Ramsey and D.W. Shafer, *The Statistical Sleuth: A Course in Methods of Data Analysis*.
- Notes and Handouts:** You are expected to download the lecture notes, handouts and examples as they become available on the [eCampus](#) web site, and to bring these to class.
- Discussion Board:** A discussion board, monitored by the grader and myself, is available at the [eCampus](#) web site. I will originate topics as I see fit or by request. Its purpose is to allow you to converse freely with the class about issues/problems, especially for homework. I only ask that you give each other help in the form of hints and suggestions, but not complete solutions. Courtesy and discretion are of course required.
The discussion board will be off-limits for new posts during exams and for 24 hours afterward.
- Computing:** I will use JMP software for examples and you may wish to use it for homework. You will be able to get the JMP software free (for *your own* computer) from the Statistics Department at the beginning of the semester. Also, the computers in the [Open Access Labs](#) (such as the one in Blocker) have JMP installed.
Even if you have had JMP installed previously, you will need to update the software and license.
- Prerequisite:** Three semesters of calculus, including multiple integration. Elementary linear algebra is helpful.
- Disabilities Help:** The Americans with Disabilities Act ensures that students with disabilities have reasonable accommodation in their learning environment. If you have a disability and need help, please contact me and [Disability Services](#) in the Student Services at White Creek complex, 845-1637.
- Academic Integrity:** You are expected to follow the [Aggie Honor Code](#) and maintain the highest integrity in your work for this class. This includes not passing off anyone else's work as your own, even with their permission.
Please see the [homework](#) and [exam policies](#) below for specifics.
- Copyright:** All the resources I provide for this course are copyrighted and may not be copied, sold or distributed without my express, written permission.
- Homework:** Homework will be assigned on the [eCampus](#) website and collected regularly and is worth 20% of the total term score. *Please see the [homework policy](#) below.*

Statistics 601 – Statistical Analysis

- Exams:** Two midterm quizzes worth 22.5% each and a final worth 35%. *Please see the [exam policy](#) below.*
- Exam Dates:** Exam I: Tuesday, 1 October, 12:45pm–2:00pm. *Keep these dates open!*
Exam II: Tuesday, 12 November, 12:45pm–2:00pm.
Note the extended class time for the midterm exams.
Final Exam: Wednesday, 11 December, 8:00am–10:00am. Students may not take the final exam early; so please do not even ask.
Distance students must make arrangements in advance with their proctors, who of course must be in attendance while the students take the exam. If you use a testing center be sure to reserve your time early!
Section 700 students not in an online degree program living elsewhere must take each exam at the same time as the on-campus section.
Students in an *online degree program* will have flexible start times. (More details later.)
- Grading Scale:** A 85% – 100%. (These are firm.)
B 70% – 85%.
C 60% – 70%.
D 50% – 60%.
- Homework Policy:** Your homework solutions must be your own work, not from outside sources, consistent with the university rules on [academic integrity](#). I expect you to follow this policy scrupulously. *Your performance on the exams is much more likely to be better if you do.*
- You may use:
- Your textbook and notes from class.
 - Your notes, homework, etc., from a related class that you took or are taking.
 - References listed on the syllabus.
 - Discussion with the instructor or grader.
 - *Voluntary, mutual and cooperative* discussion with other students currently taking the class (including on the class discussion board). Do not post complete solutions. Suggestions, descriptions and partial explanations are ok.
- You may not use:
- Solutions manuals (printed or electronic) or copies of pages from solutions manuals, except those designated by the publisher as “student manuals”.
 - Solutions from previous classes or classes taught elsewhere or at another time.
 - Solutions, notes, homework, etc., from students who took the class previously.
 - Copying from students in this class, including expecting them to reveal their solutions in “discussion”. That is, you may work together as indicated above as long as you work out and prepare your own solutions.
- Homework is to be submitted by the end of class time on its due date unless I specify otherwise. *Late homework is not acceptable without a university sanctioned excuse.*

Statistics 601 – Statistical Analysis

Exam Policy:

The discussion board will be off-limits during exams and for 24 hours afterward.

Distance students *must be proctored* according to the directions you will be given. See [Homework and Exam Submissions](#) for the online proctor application. Even if you assigned a proctor for previous classes, you will need to assign or reactivate one for this semester.

Your exam solutions must be your own work, using only resources I explicitly allow, consistent with the university rules on [academic integrity](#).

No exam may be taken early or made up, except if you provide a university excused absence with appropriate documentation.

Each exam will be comprehensive, cumulative and closed book. They will have a mix of computations, use and interpretation of computer output, concept questions and mathematical derivations.

Acceptable resources are:

- A calculator for numerical calculations only. The calculator *may not be part of, associated with or connected to any communication device*, such as a cell phone, iPod, tablet or laptop.
- Statistical tables. (Obtain your own copies. I will have versions available online.)
- One page (8½×11, both sides) of notes for the first exam, two pages for the second exam and four pages for the final exam. These *must be of your own construction*, not copied from somewhere else.

No other resources are acceptable.

Copies of some old exams will be available on the [eCampus](#) web site.

Answers should be complete, but not lengthy, with the solution clearly marked. Most questions are designed to be answerable with just a few sentences.

Classroom:

Please turn off all communication devices (cell phones, iPods, etc.) while in the classroom. You can have a calculator for in-class work. A laptop or tablet is ok as long as you only use it to take notes or to view notes and handouts for this course.

Questions are encouraged, especially to help clarify points in the lectures and homework. No question is "bad" or "dumb" if it is relevant (although I do appreciate it if you listen and avoid asking a question just answered).

Help Session Lab Room:

You are encouraged to take advantage of the help sessions in Blocker 162. Various TAs will be available, depending on the time, but all are familiar with the topics of this and similar courses.

The computers in the room are for statistics class use only. Please be considerate of others.

Missed Work and Incompletes:

This is based on [university policy](#).

- If you must miss an exam due to illness or circumstances beyond your control, notify me or the Statistics Department, in writing or by email (before, if feasible, otherwise within two working days after you return). See me as soon as possible to schedule a make-up exam.
- An Incomplete grade will be given only in the event that circumstances beyond your control cause prolonged absence from class and the work cannot be made up.

Statistics 601 – Statistical Analysis

Topic	Course Outline	Textbook Section
<i>I. Modeling Data</i>		
A. Data and Model, Randomness and Inference		1.1–1.4
B. Types of Data		3.1, 4.1
C. Displaying Data: cross-tabulation, histogram, scatterplot		6.2, 6.3, 6.6, 11.1
D. Describing Data: mean, standard deviation, quantiles, boxplot		6.1, 6.3, 6.4
E. Basic Probability: interpretations, rules, conditional probability		2.1–2.6
F. Probability Models: density, mean, variance		2.8, 3.1–3.4, 4.1–4.4
G. Special Distributions: normal, gamma, Weibull, lognormal		4.5, 4.6, 4.8–4.12
H. Density Estimates and Quantile Plots		6.3, 6.7
I. Random Sampling and Simulation		
<i>II. Statistical Inference</i>		
A. Estimation of Parameters: mean, standard deviation		7.1, 7.4
B. Sampling Behavior: bias, standard error, normal approx.		7.2, 7.3, 8.3
C. Confidence Intervals: interval for a mean, confidence, delta method		8.1, 8.2, 8.6
D. Hypothesis Testing: t-test, hypotheses, p -value, Type I and II errors		9.1–9.3, 9.6
E. Designing a Sample: sample size determination		8.1, 9.2, 9.3
F. Inference for Other Parameters: two-sample t-test		8.3, 9.4, 10.1–10.5, 10.7
<i>III. Correlation and Regression</i>		
A. Joint Distributions: correlation and independence		5.1–5.5, 11.8
B. Conditional Expectation and Regression		5.1, 5.2, 11.1
C. Estimating Correlation		11.8
D. Straight Line Regression: least squares fit, inference		11.1–11.6, 11.9
E. Residuals Diagnostics		11.7
F. Multiple Linear Regression: estimation, prediction		12.1–12.4
G. Outliers and Influential Values: Studentized residuals, Cook's D		12.5, 12.6
H. Model Fitting and Selection: model selection criteria, multicollinearity		12.2, 12.6
I. Nonlinear Regression		
<i>IV. Design and Analysis of Experiments</i>		
A. Comparing Group Means: analysis of variance, Tukey's test		13.1, 3.2
B. Handling Assumptions: tests for equal variance, Kruskal-Wallis test		10.3, 13.2
C. Contrasts		
D. Random Effects Model		13.3
E. Factorial Models: interaction		14.1–14.5
F. Randomized Block Design: purpose, inference		13.4
G. General Linear Models: covariate analysis		
<i>V. Analysis of Categorical and Count Data</i>		
A. Distributions for Counts: binomial, Poisson and multinomial		3.1–3.9, 4.7, 5.1
B. Inference for Proportions		8.4, 9.5, 10.6
C. Categorical Data Analysis: chi-square tests, logistic regression		9.7, 9.8, 11.10
D. Generalized Models for Count Data		
E. Inference for Percentiles		9.9