**Stat 202: Probability** **2018**

***Probability and Statistics***, Wackerly, Chapters 1-4,5-9

Dr. Roland Minton, Trexler 270-C, 375-2358,
**office hours** by appointment at calendly.com/minton/15min
minton@roanoke.edu

**Course Objectives**: *Continue to learn how to do mathematics!* Mathematics is a problem-solving discipline, and we all have room to improve. To develop as problem-solvers, we must focus on technique and not on memorization. In this course, we focus on the basics of the fields of probability and statistics, mathematical areas that are increasingly important in our digital world. A common complaint about probability and statistics is that there are too many formulas and the formulas are confusing. The way to avoid this problem is to not memorize formulas! If you spend time learning the thought process and concepts behind the formulas, they will make sense and be "obvious" enough that no memorization is needed. Like all math courses, thinking is the key!

**Intended Learning Outcomes:** At the end of the course, successful students will be able to

- Apply techniques for discrete and continuous distributions to solve problems of interest
- Understand the role of randomness in decision-making processes
- Distinguish among different statistical tests, and know which to use in a given situation
- Distinguish among different distributions, and know which to use in a given situation
- Use integration and sums as appropriate to determine probabilities

**Attendance Policy**: You are responsible for everything done in class, through attendance and sharing class notes with classmates. If you miss a class, e-mail or call me before class is over and explain why. **If you have two unexplained absences, you will be dropped from the course.**

**Equipment**: We will use Mathematica in class, on homework assignments and on tests. You are encouraged to get a copy installed on a laptop or desktop. In addition, we will use Excel and the statistical program R for class projects. R is a free program with outstanding open-source materials. If you are interested in taking further Stat classes, download R studio for free!

**Study Problems:** You should attempt as many of the exercises in the book as you can. In general, practical computation questions are more likely to appear on tests than theoretical questions, so focus on those questions. Ask questions about any problems that you do not understand completely. I may delay answering a question until after class, but I will always get you an answer.

**Homework:** Every other week, problems will be assigned to hand in with due dates specified. Work these problems neatly and completely. You will be graded on accuracy and ability to communicate your ideas. At least one problem will require Mathematica – once you get comfortable with its syntax, you will find it to be a wonderful resource. Each assignment will include an exploration problem that will count in your exploration grade. They will be different from book problems, and you should get started on them as soon as possible. They are very doable but may require some thought and some conversation with me to see what to do.

**Exploration:** For this class to be useful for you in the future, it is vital that you do more than just work book problems and study for tests. We will spend class time doing statistical studies and writing up the results. These reports plus the exploration problems in the homework sets will add up to more than 100 points, so you have some slack on working these problems (and a chance to get extra points!).

**I expect you to spend at least 12 hours of work each week inside and outside of class.**

**Co-Curricular**: During the course of the semester, you must attend at least two approved co-curricular events offered by the MCSP department. For each, write a two-paragraph reflection paper, giving a brief summary of the talk and expanding on some aspect of particular interest to you. Reports are due within a week of the talk.

**Tests**: There will be six full test days, six partial test days, and a final exam. Each test will cover all topics covered to that date. Test dates are F 9/14, F 9/28, F 10/12, F 11/2, F 11/16 and W 12/5. The exam is Tuesday, 12/11, 2-5. Grading is mastery based, meaning that there is no partial credit. For this reason, there are six partial test days where you can attempt one or two problems. Once you have mastered a topic, you do not need to test on it further. The intent of this system is to encourage deeper understanding of topics and perseverance at mastering topics from earlier in the semester. On a full test day, the first 30 minutes is reserved for the four new topics. After finishing those, you can retest on whatever topics you wish. We will talk about strategy and status in the course at various times during the semester – the most important message is to keep working and studying!

**Make-ups**: In case of sickness or scheduling conflicts, get in touch with me ASAP.

**Academic Integrity**: The college policy is fully supported. Tests are closed notes, closed book unless noted. Electronic devices other than computers are not allowed in test situations, and computers may only be used for computation purposes in Mathematica.

**Extra Credit:** You may earn extra credit in a number of ways. My intent is to encourage you to have fun with mathematics, and that is the grading criterion that I will use – so have fun learning! You may check out from the Roanoke College library and report on "popular" statistics or probability books. You may report on mathematical web sites that have good stat demonstrations or extra material. You may do one of the extra credit options posted on Inquire during the semester. The main rule here is to do this now; waiting until the end of the semester will annoy me and (more importantly) distract you from the end-of-semester studying that you need to do.

**Grading**: The tests-exam mastery grade counts 60% of the final average.
If you master $m$ of the 28 topics your mastery grade is $16+3m$.
Homework and co-curricular reports count 20% of the final grade.
The exploration reports and class participation count 20% of the final grade.
Grades may be curved up based on extenuating circumstances, including improvement as the semester goes on.

A: 93-100  A-: 90-92
B+: 87-89  B : 83-86  B-: 80-82
C+: 77-79  C: 73-76  C-: 70-72
D+: 67-69  D: 63-67   D-: 60-62
F: 59 and below

**Stat 202 Schedule**

| Date | Sections | Topics | Due |
|---|---|---|---|
| W 8/29 | intro | Probability Puzzles | |
| F 8/31 | 1.1 | Lacrosse stats | |
| M 9/3 | 1.2-1.4 | What is Statistics? | |
| W 9/5 | 2.4-2.5 | Sample-Point Method | Exp 1 due |
| F 9/7 | 2.6 | Counting: Permutations | pre-test |
| M 9/10 | 2.7 | Conditional Probability | HW1 due |
| W 9/12 | 2.8 | Two Laws of Probability | |
| F 9/14 | TEST 1 | | TEST |
| M 9/17 | 2.9 | Event-Composition Method | |
| W 9/19 | 2.10 | Bayes' Rule | HW 2 due |
| F 9/21 | 3.1-2 | Probability Distributions | re-test |
| M 9/24 | 3.3 | Expected Value | |
| W 9/26 | 3.4 | Binomial Distribution | |
| F 9/28 | TEST 2 | | TEST |
| M 10/1 | 3.5 | Geometric Distribution | |
| W 10/3 | 3.9 | Moment-Generating Functions | HW 3 due |
| F 10/5 | 4.2 | Continuous Distributions | re-test |
| M 10/8 | 4.3 | Expected Value | |
| W 10/10 | TEST 3 | | TEST |
| F 10/12 | | Lacrosse stats | Exp 2 due |

**FALL BREAK**

**Stat 202 Schedule**

| Date | Sections | Topics | Due |
|---|---|---|---|
| M 10/22 | 4.4 | Uniform Distribution | |
| W 10/24 | 4.5 | Normal Distribution | HW 4 due |
| F 10/26 | 4.6-7 | Gamma and Beta | re-test |
| M 10/29 | 5.2 | Bivariate Distributions | |
| W 10/31 | 5.3 | Marginal Distributions | |
| F 11/2 | TEST 4 | | TEST |
| M 11/5 | 5.4 | Independence | |
| W 11/7 | 5.7 | Covariance, Correlation | HW 5 due |
| F 11/9 | 10.2 | Statistical Tests | re-test |
| M 11/12 | 7.2-3 | Central Limit Theorem | |
| W 11/14 | 10.3 | Large-Sample Tests | |
| F 11/16 | TEST 5 | | TEST |
| M 11/19 | 8.5-6 | Confidence Intervals | Exp 3 due |
| Thanksgiving Week | | | |
| M 11/26 | 10.6 | Significance Levels | |
| W 11/28 | 10.8 | Small-Sample Tests | HW 6 due |
| F 11/30 | 10.9 | Tests of Variance | re-test |
| M 12/3 | 10.10 | Power of Tests | |
| W 12/5 | TEST 6 | | TEST |
| F 12/7 | Review | | |
| T 12/11 | EXAM 2:00-5:00 | | |

# Mastery Topics Stat 202

1. Apply the empirical rule
2. Compute probabilities with the sample-point method
3. Compute combinations and permutations
4. Compute conditional probabilities
5. Compute probabilities with the event-composition method
6. Compute probabilities with Bayes' Rule
7. Compute and give graphical displays of probability distributions
8. Compute expected values for discrete distributions
9. Solve problems with binomial distributions
10. Solve problems with geometric distributions
11. Solve problems with moment-generating functions
12. Compute probabilities for continuous distributions
13. Compute probabilities with uniform distributions
14. Compute probabilities with normal distributions
15. Compute probabilities with various distributions
16. Compute probabilities with bivariate distributions
17. Compute marginal distributions
18. Determine independence
19. Compute covariances and interpret correlations
20. Use the Central Limit Theorem to apply statistical tests
21. Complete statistical tests for large samples
22. Compute confidence intervals
23. Complete statistical tests for small samples
24. Complete statistical tests on variances
25. Exam problem #1: identify distributions
26. Exam problem #2: interpret definitions and theorems
27. Exam problem #3: comprehensive problem
28. Exam problem #4: comprehensive problem

**Grading scale**

**28 mastered**: 100
**27 mastered**: 97
**26 mastered**: 94
**25 mastered**: 91
**24 mastered**: 88
**23 mastered**: 85
**22 mastered**: 82
**21 mastered**: 79
**20 mastered**: 76
**19 mastered**: 73
**18 mastered**: 70
**17 mastered**: 67
**16 mastered**: 64
**15 mastered**: 61
**14 mastered**: 58
**13**: 55 **12**: 52 **11**: 49 **10**: 46 **9**: 43 **8**:40 **7**: 37 **6**: 34 **5**: 31 **4**: 28 **3**: 25 **2**: 22 **1**: 19

# Model Reflection Paper

(This is made up, but shows what I'd like to get from you. The two main elements are (1) brief summary of talk and (2) some original thought on the subject.)

The talk on September 7th was by Dr. Sue Dokoo of Pseudo Duke University. Her research is in the game of Sudoku and discussed different aspects of this game. I have seen other people playing it, but did not know the rules or any of the mathematics behind it.

In this game, a 9x9 playing space is provided. An example given was:

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | 6 | 2 | | | 5 | 8 | |
| 4 | | 2 | 5 | | | | | 7 |
| | | | | 7 | 8 | 6 | | 3 |
| 5 | | 1 | | 6 | 7 | | | 8 |
| | 3 | | | | | | 6 | |
| 6 | | | 8 | 2 | | 9 | | 1 |
| 7 | | 4 | 3 | 9 | | | | |
| 9 | | | | | 5 | 2 | | 6 |
| | 5 | 3 | | | 1 | 4 | | |

To "solve" the puzzle, one could just enter numbers in a brute-force kind of way to see if they could get a working configuration. However, sitting in a room full of mathematicians, taking a more analytical approach seemed to be the dominant strategy. Treating this as a constraint-satisfaction problem, you can identify that certain cells must contain specific values. This leads to the conclusion that there is exactly one solution to a "well-formed" Sudoku.

This got me thinking about well-formed Sudoku, and how they are generated in the first place. It seems unlikely that the seeds are randomly assigned, you run the risk of violating set-up rules. A bigger problem is that the seeds may not constrain the possibilities enough to make a unique solution. Another naïve approach might be to take a completed grid and start taking away numbers, but I suspect that you might have a similar issue in terms of necessary constraints.

One that I want to think about is: In forming a viable Sudoku, is it the number of seeds or the placement of seeds that is more critical? I suspect the latter. Also,

- What is the maximum number of seed numbers that can be provided and still result in an ambiguous (unsolvable) puzzle?
- What is the minimum number of seed numbers that can be provided to generate a (uniquely) solvable puzzle?

We were provided two puzzles – one was rated "Easy" the other "Difficult".

- What goes into the rating system?
- Does a difficult puzzle necessarily have fewer seed numbers?
- Is the rating of the complexity somehow determined by the deductive skills required?

# Stat 202 Information Sheet

Name:

Email:

Cell phone:

Intended Major:

Hometown:

List any other statistics courses you have taken.

What are your expectations and goals for this course?

What are some of the co-curricular or other campus activities you would like to participate in this year?

Pick a (random) integer between 1 and 9 (inclusive):

Pick a random 3-digit integer:

List a sequence of 20 random H's and T's that would model coin-flipping: