

MATH 351 - Elementary Statistical Theory - Spring 2021

Class and Instructor Details

MATH 351-003 - Elementary Statistical Theory (3 credits)

Instructor

Virgil U Pierce

Email: virgil.pierce@unco.edu

Cell: 956-249-0566

Office: Ross Hall 2239

Class Meetings

Tuesday: 9:30 to 10:45 in person except for first two weeks

Thursday: 9:30 to 10:45 online Ross Hall 2261

COVID-19 Adjustments

We will start the first two weeks of Spring 2021 with online classes on both Tuesdays and Thursdays and then transition to in person classes on Tuesdays and online classes on Thursdays.

However, the situation is fluid and additional adjustments may be needed. Please watch Canvas/Email for updates.

Student Hours

Online and To be Determined.

Appointments can be made via: <https://virgilpierce.youcanbook.me>

Course Description

From the catalogue: A continuation of MATH 350. Learn about jointly distributed random variables, central limit theorem, sampling distributions, properties of estimation, confidence intervals and tests of hypothesis.

Topics will include:

- Sampling distributions and other methods of experimentation;
- Random variables and jointly distributed random variables;

- Population samples and the central limit theorem;
- Point and Interval Estimations
- Hypothesis Tests
- Bayesian Methods

Programing

Statistics is a subject that is best treated as a science and in particular extensive experimentation with the problems being considered is essential to understanding the results. As a result we will be using Python (specifically Numpy, Sympy, and Scipy) to run experiments and do computations of results. You will be provided with example code that can be modified for the problems we do in class, and really this is meant partly to build upon your experience from the Introductory Programming class. That said, CS 120 is not a prerequisite for the course and so we will have a gentle introduction to using Python and extensive help built into the class.

The good news is that you will not be asked to do Calculus computations by hand but instead will be able to use Sympy or Scipy algorithms for any integration results we need.

The course materials for class are being written in Jupyter/Python and are available in Github as .ipynb files.

You will need access to Python, here are some options:

- To install it on your own computer or laptop I recommend downloading Anaconda: <https://www.anaconda.com/products/individual> which includes Jupyter and Python.
- Google Colabs provides a Jupyter-like environment and Python (including the packages we will use), it also is integrated with Github. You will need a gmail account. <https://colab.research.google.com>

Course Prerequisites

MATH 350 and (possibly concurrent) MATH 233.

While we will be using Python as part of class, CS 120 is not a prerequisite. We will have example programs to get started and will only use a small part of what Python offers as a programming language. This course in fact will serve as a nice introduction to Python programming.

Textbook and Software

We will be following the textbook: **Mathematical Statistics with Applications** by Wackerly, Mendenhall, and Scheaffer. While it is highly recommended, any problems or other work assigned will be provided separately and much of the course material can be found on the internet.

Jupyter Notebooks and Github

The notes for class and the homework is written in Jupyter notebooks. [We will use github to distribute these here](#). I have made it a public repository so it should be easy to find. You have two choices for using this link, you can download individual files from the repository, or you can set up your Github account and then download or clone the repository to your own computer to then open the files in Jupyter. You can use the "Issues" tab in the Github website to let me know about problems you have seen in the notes (typos or things that don't run).

Python References

You may need a Python reference book, depending on whether you prefer a book or the web for referencing. Let me know if you would like me to recommend one.

That said, while I like to have a book to refer to for programing, most of the time it does not have the detail I need, and I end up using in-line help or on-line help anyway. By the way, here are some examples of how you can get some inline help in Python:

```
In [1]: # Help page describing the function or object
# I use this to get some idea of the syntax for the function - that said, I us
# end up looking on a website.

list?
```

```
Init signature: list(iterable=(), /)
Docstring:
Built-in mutable sequence.
```

If no argument is given, the constructor creates a new empty list.
The argument must be an iterable if specified.

```
Type:          type
Subclasses:    _HashedSeq, StackSummary, DeferredConfigList, SList, _ImmutableL
ineList, FormattedText, NodeList, _ExplodedList, Stack, _Accumulator, ...
```

```
In [2]: # A list of the non system attributes of a function or object
# I use this to get some idea of what functions or actions are typically appl
# to a Python object.

# It is useful to drop the special actions with the leading or trailing '__'
# they are not usually ones we care about, which is why I use this expression
# usually just a dir() would work fine.

[x for x in dir(list) if not '__' in x]
```

```
Out[2]: ['append',
'clear',
'copy',
'count',
'extend',
'index',
'insert',
'pop',
'remove',
'reverse',
'sort']
```

Learning Objectives / Outcomes for the Course

As a result of this course, students will expand their skill set to work and think with data. They will

- 1. be able to do computations of expected value and variation of discrete and continuous random variables;
- 2. Simmulate a sampling of a discrete and continuous random variable;
- 3. Be able to evaluate assumptions used in a statistical model and critique conclusions;
- 4. Develop point and interval estimatics of a random variable;
- 5. Conduct hypothesis tests to make conclusions about a population from a sample;
- 6. Explain and analyze ethical considerations for various real statistical problems.

Course Outline

As I said: Python is quickly replacing spreadsheets in my daily work. Example: I am using it to make the schedule for our class.

In [3]:

```
import MATH_351_Schedule
MATH_351_Schedule.schedule
```

Out[3]:

	Day	Week_Day	Title	Description	Notes
0	2020-01-12	Tuesday	Discrete Random Variables	Discrete random variables and their probability distributions, simulating discrete random variables	
1	2020-01-14	Thursday	Expected Value of Discrete Random Variable	Computing the expected value of a discrete random variable	
2	2020-01-19	Tuesday	Binomial and Geometric Probability Distributions		
3	2020-01-21	Thursday	Poisson Probability Distribution		
4	2020-01-26	Tuesday	Continuous Random Variables	probability density, cummulative distribution function, expected values, uniform distribution, and simulating a continuous random variable	
5	2020-01-28	Thursday	Normal Probability Distribution		
6	2020-02-02	Tuesday	Gamma and Beta Probability Distributions		
7	2020-02-04	Thursday	Other expected values and momemnt		
8	2020-02-09	Tuesday	Multivariate Probability Distributions	marginal and conditional probability distributions	
9	2020-02-11	Thursday	Independent Random Variables		
10	2020-02-16	Tuesday	Expected Value of Multivariate Random Variables and some Special Theorems		

	Day	Week_Day	Title	Description	Notes
11	2020-02-18	Thursday	Sampling dependent random variables	Sampling dependent random variables is much harder than sampling single or independent random variables	
12	2020-02-23	Tuesday	The Covariance of Two Random Variables		
13	2020-02-25	Thursday	Expected values and variance of linear functions of random variables		
14	2020-03-01	Tuesday	The multinomial probability distribution		
15	2020-03-03	Thursday	Conditional Expectations		
16	2020-03-08	Tuesday	Finding the Probability Distribution of a Function of Random Variables		
17	2020-03-10	Thursday	Order Statistics		
18	2020-03-15	Tuesday	No Class - University Holiday	NaN	NaN
19	2020-03-17	Thursday	No Class - University Holiday	NaN	NaN
20	2020-03-22	Tuesday	Sampling Distributions Related to the Normal Distribution		
21	2020-03-24	Thursday	The Central Limit Theorem		
22	2020-03-29	Tuesday	Normal Approximation to the Binomial Distribution		
23	2020-03-31	Thursday	Point Estimators		
24	2020-04-05	Tuesday	Confidence Intervals		
25	2020-04-07	Thursday	Sample Size Concerns		
26	2020-04-12	Tuesday	Methods of Estimation		
27	2020-04-14	Thursday	Common Large Sample Tests		
28	2020-04-19	Tuesday	Type II Error Probabilities		
29	2020-04-21	Thursday	Hypothesis Testing and Confidence Intervals		
30	2020-04-26	Tuesday	More Hypothesis Testing		

	Day	Week_Day	Title	Description	Notes
31	2020-04-28	Thursday	Considerations in Experimental Designs		
32	2020-05-07	Saturday	Alternative Approaches	In our last class we will take a look ahead to the methods of Data Science and resampling for hypothesis testing and confidence intervals	

Connections with Industrial Mathematics

This course is directly connected with the growing mathematics field of Data Science. There are many opportunities for people with an expertise in this field. However the field is very accessible so you have competition.

Some thoughts on how to rise to the top:

1. Have some projects you have done that you can talk about (we will do this),
2. Be able to talk about working with a team i.e. how do you help your team be effective, how did you get your team over a challenge, how do you lead?,
3. Be able to talk about challenges in the field (we will do this),
4. Have a breadth of experience in science or business generally, and mathematics, statistics, and computer science specifically (we will do this), and
5. The more you have done the better - I suggest devoting some of your play time to working with data (we will explore Kaggle in this class, and my suggestion is that it become one of your hobbies).

I have some literature on building a career in Data Science available if you like.

Look for the STAT 411 course next Fall that is our Fundamentals of Data Science.

Assessments

As with any mathematics class, doing the work is essentially to processing and learning the material. Furthermore the concepts covered in this class are not meant to be purely abstract things, but instead are actual tools deployable in any human endeavor, even your hobbies. You must do the activities in order to make progress.

<https://www.youtube.com/watch?v=91G5SdW9tiQ>

Homework

There will be some homework assignments asking you to explore topics from class, and expand on some of the work we do. These are intended to be done with a partner or two - I do not suggest more than four people working together. If you have done them you will be able to actively participate in our next class. If you have not done them, you will find our next class

overwhelming. There will roughly be one or two homework assignments per week, and I would budget 3 - 4 hours per week for them.

They are intended to be done with others, however I understand you are busy, and ultimately if you need to do the homework alone, that will be fine.

Also note that you do not necessarily need to have a perfect solution - for one thing there rarely is a perfect solution for this class, but rather you have spent some time thinking about how to get started.

Checking your understanding

We will do a number of activities during class, many as teams, to help you check on your understanding and progress.

Class Presentations and Flipgrid

You will give presentations to the class about the work we are doing.

We will use Flipgrid almost every class to share ideas and strategies on problems with each other.

Projects and Final Project

There will be two midsemester projects that you will work on to demonstrate that you are processing the information from class and incorporating it into your mathematical and statistical practice. These will be largely self selected, and so I expect each of you to have your own problems and submissions. You are permitted to seek help or advice, but you will have to submit something that you can explain.

Project Videos

Each project will have an accompanying video where you explain your methods, results, and any issues that you have found.

Grading Policies

Your grade in the class will be based on homework, team activities, the team project, and class presentations.

Grading and Feedback

I will be providing you feedback on the Homework, Class Activities and Presentations, the Projects and Project Videos. I am interested in seeing you make progress as class continues. As such, you can revise and resubmit any assignment up until the announced deadline at the end of class.

Your final grade in class will be assessed primarily from the Projects and whether I see evidence that you have met the objectives for the class. Specifically: A grade of A will mean that you have met all six course objectives, and exceeded the standard on at least 4 of them; a grade of B will mean that you have met all six course objectives, and exceed the standard on at least 2 of them; a grade of C will mean that you have met all six course objects. Finally a grade of D will mean that at least one of the standards was not met, and a grade of F will mean that multiple standards were not met or that too many assignments were missing to allow me to assess a standard.

UNCO Policy Statements

Disability Resources

It is the policy and practice of the University of Northern Colorado to create inclusive learning environments. If there are aspects of the instruction or design of this course that present barriers to your inclusion or to an accurate assessment of your achievement such as time-limited exams, inaccessible web content, or use of videos without caption, please communicate this with your professor and contact Disability Support Services (DSS) at (970) 351-2289, Michener Library L-80 to request accommodations. Students can learn more about the accommodation process at <https://www.unco.edu/disability-resource-center/accommodations/>

Food Insecurity and Basic Needs

Research shows that college students experience food insecurity at higher rates than the American household rate, and that food insecurity can negatively impact academic performance and persistence. In recognition of this problem, UNC offers assistance to students facing food insecurity through an on- campus food pantry. The Bear Pantry is located in University Center 2166A, and is open for regular hours throughout the semester.

Please visit www.unco.edu/bear-pantry for more information.

Any student who faces challenges securing their food or housing and believes this may affect their performance in the course is also urged to contact Student Outreach and Support (SOS) for assistance. SOS can assist students during difficult circumstances which may include medical, mental health, personal or family crisis, illness or injury. SOS can be reached at sos@unco.edu or via phone at 970-351-2796.

Honor Code

All members of the University of Northern Colorado community are entrusted with the responsibility to uphold and promote five fundamental values: Honesty, Trust, Respect, Fairness, and Responsibility. These core elements foster an atmosphere, inside and outside of the classroom, which serves as a foundation and guides the UNC community's academic, professional, and personal growth. Endorsement of these core elements by students, faculty, staff, administration, and trustees strengthens the integrity and value of our academic climate.

UNC's Policies

UNC's policies and recommendations for academic misconduct will be followed. For additional information, please see the Student Code of Conduct at the Dean of Student's website <http://www.unco.edu/dos/Conduct/codeofconduct.html>. In the case of academic appeals, university procedures will be followed. For information on academic appeals, see <https://www.unco.edu/trustees/pdf/board-policy-manual.pdf>.

In []: