MA 2612 – Applied Statistics II

Atwater Kent 116
M/T/R/F 10:00 AM – 10:50 AM

Professor Carly Siegel Thorp (she/her/hers)

Email: cthorp@wpi.edu **Office**: Stratton Hall 413
"Stop By My Office & Ask Questions" Hours: Mondays / Thursdays 11am-12pm and Tuesdays 1pm-2pm

Why is this course important?

In Applied Statistics I, you were introduced to statistical inference (e.g., hypothesis testing), which is, really, what the field of statistics is all about. We collect sample data to infer statements about our target population. It is rare, if not impossible in certain situations, to conduct a census and we are often required to use a sample to draw conclusions about an entire target population. Throughout this course, we will delve deeper into applications of statistical inference, building on what you learned in Applied Statistics I or other introductory statistics courses.

What does this course cover?

This course is designed to be a continuation of your introduction to applied statistics in MA 2610 / MA 2611, and a working knowledge of the content covered in MA 2610 / MA 2611 is assumed.

By the end of this course, you should be able to:

- ◊ Conduct a hypothesis test and interpret the results for comparing two samples; execute and interpret confidence intervals for comparing two samples.
- ◊ Execute and interpret the results of a chi-squared test for independence; conduct a hypothesis test and interpret the results for comparing two proportions.
- ◊ Conduct a one-way ANOVA and interpret the results for comparing more than two groups of data; understand the usage of pairwise comparison tests.
- ◊ Recognize when to use nonparametric tests when comparing samples of nonparametric data.
- ◊ Understand the steps for modeling with linear regression; execute and interpret prediction intervals for a linear regression model.

Statistics Reference Materials

While this course does not have a *required* textbook, if you are interested in delving a bit more into statistics or want a reference for the course, I recommend the following:

- ◊ Books on learning R that will, hopefully, not become coffee table books:
  - *The Book of R: A First Course in Programming and Statistics* by Tilman M. Davies
  - *Learning R: A Step-by-Step Function Guide to Data Analysis* by Richard Cotton
  - *R Cookbook: Proven Recipes for Data Analysis, Statistics, & Graphics* by J.D. Long & Paul Teetor
  - *R Graphics Cookbook: Practical Recipes for Visualizing Data* by Winston Chang
- ◊ Books on statistics for a bit of lighter, bedtime reading:
  - *Naked Statistics: Stripping the Dread from the Data* by Charles Wheelan
  - *The Lady Tasting Tea: How Statistics Revolutionized Science in the 20th Century* by David Salsburg

Grading Methodology

This course will be using an alternative grading methodology known as standards-based grading (SBG). Some of you may have encountered it before in other courses or in high school, while some of you may not have. Regardless of your experience with alternative grading, the intent of SBG is to create a learning environment that allows you to successfully learn the material.

The course standards are separated into 3 categories: Concept (C), Application (A), and Lab (L). Each category has a set number of standards, found in further detail in the following "Course Standards" section, and most standards will ***need to be met more than once*** to receive completion credit. Each standard can be completed through either an assignment or exam (or both). Refer to the "Course Standards Check Sheet," provided on the Canvas site under the "Course Need-to-Knows" module, for the completion counts required for each standard.

Your final grade will be dependent on meeting these standards, which are relative to the material you will be learning throughout the course, rather than being based on meeting X% in your homework assignments and exams, as is done with traditional grading methodology.

For further details on your final course grade, see the "Course Grade Breakdown" section.

Course Standards

Concept (C) Standards represent the core concepts you will be learning. Some of these standards will need to be met more than once to be considered completed and are listed in the following table.

Number	Standard Description
C.1	Identify the null and alternative hypotheses for comparing two samples relative to a provided problem scenario.
C.2	Identify when to check for equal variances when conducting inference for comparing two samples.
C.3	Identify the correct wording for the hypothesis test conclusion relative to the predetermined level of significance (α).
C.4	Calculate a confidence interval for comparing the difference between two samples.
C.5	Find the critical value relative a provided level of confidence or significance.
C.6	Identify if a statistical error occurred and, if applicable, its type relative to a provided problem scenario.
C.7	Calculate the expected counts for the chi-squared test of independence relative to a provided problem scenario.
C.8	Identify and calculate the components of a one-way ANOVA table.
C.9	Identify the required assumption checks for an ANOVA.
C.10	Identify when to use a nonparametric test when comparing 2 or more groups of nonparametric data.
C.11	Identify the type of association depicted on an x-y scatterplot.
C.12	Calculate Pearson's correlation coefficient and interpret it relative to a provided problem scenario.
C.13	Calculate the R^2 value for a simple linear regression model and interpret it relative to a provided problem scenario.
C.14	Calculate the model parameters of a simple linear regression model using the method of least squares estimation.
C.15	Calculate a prediction interval for a linear regression model.
C.16	Identify when to remove insignificant terms from a multiple linear regression model.

Application (A) Standards represent your ability to take the base content learned and successfully apply it to realistic problem scenarios, as you will likely do in your higher-level courses and careers. Like the *Concept (C) Standards*, some will need to be met more than once to be considered completed. They are listed in the following table.

Number	Standard Description
A.1	Apply the steps and interpret the results for a hypothesis test comparing two samples relative to a provided problem scenario.
A.2	Interpret how the confidence interval, relative to the level of significance, supports the results of a hypothesis test relative to the provided problem scenario.
A.3	Apply the steps and interpret the results for a chi-squared test of independence relative to a provided problem scenario.
A.4	Apply the steps and interpret the results for a one-way ANOVA relative to a provided problem scenario.
A.5	Interpret the results for a pairwise comparison test relative to a provided problem scenario for a one-way ANOVA.
A.6	Apply the steps and interpret the results for developing a simple linear regression model relative to a provided problem scenario; add the fitted model to an x-y scatterplot of the data and interpret.
A.7	Interpret a prediction interval for a linear regression model relative to a provided problem scenario.
A.8	Apply the steps and interpret the results for developing a multiple linear regression model relative to a provided problem scenario.

Lab (L) Standards represent the skills and tools you will be learning in *R* and *RStudio* and will not be evaluated on the exams, only in your weekly lab assignments. Most of these standards, listed in the following table, will also need to be met more than once to receive completion credit.

Number	Standard Description
L.1	Effectively use the R Markdown tool in RStudio to generate and output the lab assignment into HTML, PDF, or Word.
L.2	Import a provided ".csv" file of data and save as a dataframe for usage in a lab activity or assignment.
L.3	Generate the code and interpret the output for comparing two samples relative to a provided problem scenario.
L.4	Generate the code and interpret the results for comparing the equivalency of two variances.
L.5	Generate the code and interpret the output for conducting a chi-squared test of independence relative to a provided problem scenario.
L.6	Using the "ggplot()" function, generate a plot which contains multiple groups of data with different colors and/or shapes for the groups.
L.7	Generate the code for conducting the Shapiro-Wilk test to check the normality assumption for sample data and interpret the output.
L.8	Generate the code and interpret the results for determining the equivalency of more than two variances using the appropriate test.
L.9	Generate the code and interpret the output for conducting a one-way ANOVA relative to a provided problem scenario.
L.10	Generate the code and interpret the output for conducting a pairwise comparison relative to a provided problem scenario.
L.11	Using the "ggplot()" function, generate and interpret an x-y scatterplot, with colors and appropriate labels, relative to a provided problem scenario.
L.12	Generate the code and interpret the results for calculating Pearson's correlation coefficient relative to a provided problem scenario.
L.13	Generate the code and interpret an estimated linear regression model relative to a provided problem scenario.
L.14	Using the "ggplot()" function, add a fitted line to an x-y scatterplot, using a different color than the points on the plot, for your estimated simple linear regression model relative to a provided problem scenario.
L.15	Using the "ggplot()" function, add a fitted line with the prediction interval to an x-y scatterplot and interpret where a specific point will fall, for your estimated simple linear regression model relative to a provided problem scenario.
L.16	Generate the code and interpret a prediction interval for a specified point relative to your estimate simple linear regression model for a provided problem scenario.
L.17	Using the "ggplot()" function, generate and interpret a normal QQ plot, with an added QQ line with added details (i.e. colors and symbols) to check normality of studentized residuals.
L.18	Using the "ggplot()" function, generate and interpret an x-y scatterplot that compares the ordinary residuals to the fitted values for an estimated simple linear regression model.

Assignments, Exams, Final Individual Mini Project, and Metacognitive Reflections

Throughout the term, there will be a variety of opportunities to meet the required course standards for your desired grade in the course. Additionally, you will be required to complete a series of self-

reflections throughout the term, which use metacognitive techniques to help enhance your awareness and learning capabilities.

The first group of assignments are those which will be used to meet most of the standards:

- Concept Checks: These will focus solely on the *Concept (C) Standards* and be available to take on Canvas, with a total of 6 over the course of the term. Once available, you will have until the end of the term to submit for credit. However, I ***strongly*** encourage you to complete each one within the first week of its availability to stay on track with the learning material and complete standards in a timely manner, rather than waiting until the end of the term to meet most of the required concept standards for your intended grade!

  These will be graded as "Standards Met" or "Try Again." You will be allowed **2 attempts** for each concept check, but, once started, you will have an hour to submit, as you would a quiz, unless an approved accommodation applies. Additional attempts may be requested via the usage of a token (see the "Tokens" section below for further details).

- Homework Assignments: These will mostly focus on the *Application (A) Standards* but will have a few of the *Concept (C) Standards* included. There will be 4 of these to complete over the course of the term and, to be accepted for credit, must be submitted **within 3 days of the due date**. The due date on Canvas is the "soft" deadline, while the "hard" deadline is noted within the assignment criteria.

  If you are unable to submit on time, you may use a token for an additional 24-hour extension (see the "Tokens" section below for further details). You will be provided feedback as to which standards were met, and which ones need additional work with suggestions on how to meet those standard(s). To revise standards within a homework assignment, you will need to submit a token (see the "Tokens" section below for further details).

- Lab Assignments: These will focus almost entirely on the *Lab (L) Standards* and be connected to your weekly lab activities, with a total of 6 over the course of the term. To be accepted for credit, these must be submitted **within 3 days of the due date**. The due date on Canvas is the "soft" deadline, while the "hard" deadline is noted within the assignment criteria.

  If you are unable to submit on time, you may use a token for an additional 24-hour extension (see the "Tokens" section for further details). You will be provided feedback as to which standards were met, and which ones need additional work with suggestions on how to meet those standard(s). To revise

the standards within a lab assignment, you will need to submit a token (see the "Tokens" section below for further details).

The second group of assignments consist of your exams. There will be two comprehensive midterm exams, held during class, on the following dates.

Exam	Date
Midterm Exam 1	Thursday April 3, 2025
Midterm Exam 2	Tuesday April 29, 2025

Each exam will focus on the Concept and Application standards being assessed, so your grade will reflect which standards were met and which ones were not, rather than X% problems correct. Following each midterm exam, you will be provided the **optional** opportunity to revise and redo your understanding of the standards not completed. Further details on the revision process will be provided after your midterm exams are graded.

In lieu of a final exam, you will instead have a final individual project to complete. You will be tasked with creating a **cumulative** final exam (plus answer key) that can be completed in 50 minutes, as if you are the professor, which contains unique questions that address a variety of concept and application standards. Further details and the grading rubric will be provided on Canvas.

The last group of assignments are metacognitive self-reflections and there will be 3 of these over the course of the term, at the start, middle, and end. The intent of these is to help you assess your progress in the course by using a variety of metacognitive tools to improve your learning experience. These will not be graded but will instead be assessed as "Complete" or "Incomplete" and are due **within one week of becoming available** on Canvas.

Course Grade Breakdown

Here is a grade breakdown of the number of standards that must be completed in each category:

	Concept (C) Standards	Application (A) Standards	Lab (L) Standards
	Out of 16	Out of 8	Out of 17*
A	14	7	14
B	12	5	11
C	9	3	8

*for the Lab (L) Standards, you **must meet L.1 to receive a passing grade**, which is not included in this count

I recognize this structure leaves a bit of "gray" area, such as, if you meet the Concept and Lab standards for an A but only have 6 out of 8 Application standards, will you get an A or B? This is where something known as the "Gimmie/Plea" clause comes into play.

If you have determined that you are missing one (or more) standards in one category, which is preventing you from receiving the final grade you want, you must reach out to me as soon as possible, prior to the term's end. We can then make arrangements that will be fair and equitable to all those in the class and still meet the university's regulations on conduct and workload in classes. However, this does not allow you to ignore any work expected of everyone in the class and it is expected you will meet the course requirements in some way.

Additionally, to pass the course and be able to revise any standards, every homework and lab assignment must be completed and submitted by its "hard" deadline (see the respective section for each assignment type). Metacognitive self-reflections must also be completed and submitted.

There's a bit more flexibility on the concept checks. By the end of term, to receive an/a:

- A: must attempt all 6 concept checks
- B: must attempt at least 5 (out of 6) concept checks
- C: must attempt at least 4 (out of 6) concept checks

For the final individual mini project, it must be completed to receive a passing grade in this course.

Throughout the term, I encourage you to keep track of the standards you have met using the Excel spreadsheet linked on the course Canvas page under the module "Course Need-To-Knows." This will not only help you with tracking your grade in the class, but it will also allow you to maintain awareness of what topics or material you may need additional guidance understanding. You have also been provided the code for a "Grade Calculator," which was kindly created by one of your peers in a prior course, which you can use to check where your grade stands throughout the term.

To aid in helping keep track of your completed standards, you will be provided access to a live spreadsheet, updated weekly, that provides tabulations of which standards you have completed and which ones you have not for each concept check, homework assignment, lab assignment, and midterm exam. Canvas is not conducive for standards-based grading, so you will need to rely on the

spreadsheets provided to evaluate your progress in the course and if you are on track to receive the grade you want. Further details on grading will be provided on the first day of class.

I reserve the right to modify the above grading criteria, but not to lower the grade you would have received. If your grade does not meet the minimum criteria for a "C," then you will receive an "NR" as your final grade.

Tokens

To allow a bit of flexibility throughout the term, you will be granted 4 tokens at the beginning of the term, which can be put towards the following:

- ◊ Revisions of incomplete standards on any homework or lab assignment. These must be submitted **within 72 hours** after your grade is returned.
  - Revisions of standards on the midterm exams are separate and do not require tokens.
- ◊ 24-hour extensions on the "hard" deadline for any of your assignments, *except* for the exam revisions and final individual project. Requests for extensions must be submitted **before** the assignment is due.
- ◊ An additional attempt on a concept check.
- ◊ Any other bending of the course rules you might want – just discuss with me first!

There will be additional opportunities to earn tokens throughout the term. To use a token, requests **must be submitted** via the Google form found here and on Canvas under the module "Course Need-To-Knows." Requests submitted via email will **not** be accepted.

Assignment Revisions

After submitting a token for revising standards on a particular homework or lab assignment, you will be provided access to a *separate* revision submission on Canvas and have **one week** to submit your revised assignment. Your revision does not need to include your entire assignment, but it must meet the revision criteria provided with your revision assignment to be accepted for assessment.

Lab Software Requirements

For your lab activities and assignments, the software you will be using is *R* with *RStudio*, which was introduced in your previous applied statistics course.

Why *R* and *RStudio*? Well, for a couple of reasons…

First, it's free and open source, which makes it easily accessible for anybody, no matter what field they are studying or where they are working. Secondly, because it is open source, it is the primary software used by researchers and statisticians, so many of you will find it useful in your future studies and careers. Plus, it never hurts to add a new skill your resume!

Other statistical software packages are used within a variety of industries, some of which you may encounter in your careers, such as Minitab, JMP, and SAS. Additionally, Microsoft Excel can be used to conduct basic statistical calculations and create graphical visualizations but is generally not recommended if other statistical programs are available.

Refer to the "Course Need-To-Knows" module for instructions on updating / installing R and RStudio.

Course Communication

The best method of communication for directly contacting me or your PLAs/TAs is **through email**. However, any announcements, assignments, or any other information you need will be found on the course Canvas page. Lecture materials, under each of the respective modules, will be posted prior to class and any notes added to the lecture materials will be posted within 24 hours after class.

If I forget, at any point, to post the lecture slides or notes, please do not hesitate to send me an email (cthorp@wpi.edu) to remind me. I do occasionally forget ☺

Class Attendance and Echo360 Class Recordings

If you are unable to attend a class or would like to review a particular class for studying, there will be recordings available via Echo360.Unless an approved accommodation applies, it is ***expected*** you will attend class and not attempt to utilize the Echo360 recordings as your primary source material. While attendance is not mandatory, I encourage all of you to attend class to participate in discussions and hands-on activities to facilitate your learning.

In the event of inclement weather (or any other reason as to why class cannot be conducted in person), classes will be held via Zoom and an announcement will be provided on Canvas.

Tips for Success

To be successful in this course, I encourage you do to the following:

- ◊ **Attend class regularly**. In addition to the lectures, there will be regular in-class discussions and hands-on activities to help facilitate your learning.
- ◊ **Ask questions**. This can be done during class, office hours, or via email. If there is a topic or problem in an assignment you need additional guidance understanding, please let me know.
- ◊ **Engage and participate (in the way in which you feel most comfortable)**. While it may not seem like it, I'm an introvert by nature but I have a passion for education and want to help you learn a solid foundation in statistics. I am not the type of professor who calls out their students, as demonstrated in *Legally Blonde*, but I would also appreciate not being the one talking the whole class.
- ◊ **Be open-minded**. My approach to teaching and grading is entirely about ensuring you learn the material to the best of your abilities. If it's not the typical approach you have seen by other professors, I encourage you to give it a chance, but if there is ever a point in the course when it is not working for you, please let me know!

Statement of Respect

As your professor, I expect you to treat your peers with respect and engage in considerate communication. Everybody comes to the table with different strengths, and weaknesses, so I encourage you to focus not only on your own strengths but also those of your peers.

To ensure you feel respected and part of a safe and inclusive learning environment, please do not hesitate to reach out to me if any of the following are relevant to you:

- ◊ If you have a name and/or pronouns that differ from those in your official WPI records.
- ◊ If something was said in class (by anybody) that made you feel uncomfortable and/or you feel like your performance is being negatively impacted by experiences outside of the course.

Academic Honesty

You are expected to be familiar with WPI's Academic Honesty policies.

All acts of fabrication, plagiarism, cheating, and facilitation can be prosecuted according to the WPI's policies. If you are ever unsure as to whether your intended actions are considered academically honest, please contact me directly.

Student Resources

Mental Health & Physical Wellbeing

Your mental health and physical wellbeing are of utmost importance. If you are struggling with your health or wellbeing, please reach out to the Wellness Center or Student Development & Counseling Center (SDCC). Resources can be found here.

Accommodations

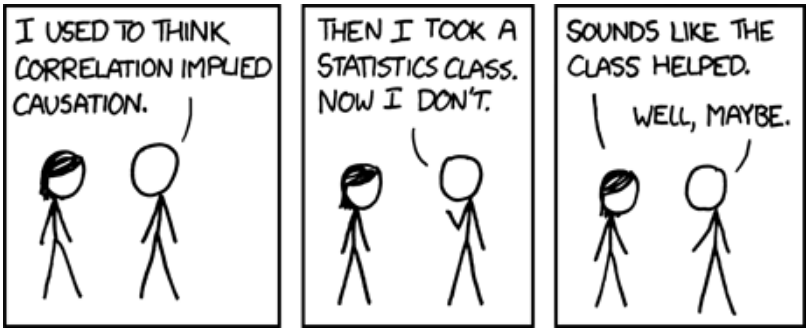
If you need accommodations or support throughout this course, you are encouraged to contact the Office of Accessibility Services (OAS) as soon as possible to ensure that such accommodations are implemented in a timely fashion. The OAS is in Unity Hall and can be reached via phone (508-831-4908) and/or email (accessibilityservices@wpi.edu).

Math Tutoring Center (MTC)

In-person math tutoring will be offered in Stratton Hall 206. Please use wpi.edu/+mtc to sign up for individual tutoring appointments and check MTC hours of operation.

Academic Resources Center (ARC) Tutoring

Peer tutoring and Math and Science Help (MASH) will be offered in-person by the Academic Resources Center (arc@wpi.edu) tutors in D Term. If you are looking for a tutor, please reach out to the Academic Resources Center (arc@wpi.edu) or submit your availability at Bit.ly/ARCTutor to request additional tutoring. No appointments are needed for MASH group sessions.



Source: https://xkcd.com/552/