# STAT 145: Biostatistics
## Syllabus, Spring 2024

**Course Information**.

- 4 semester hours
- Instructor: Thomas Scofield
- Texts:

    - *Introductory Statistics for the Life and Biomedical Sciences*, 1st Ed., by Vu and Harrington
    - *Social Justice Fallacies*, by Thomas Sowell

- Class meetings: MWF, 12:15–1:20 pm, NH 253
- Class homepage:

    https://www.myopenmath.com/course/course.php?cid=211957&folder=0

    to see contact information and office hours. The first time you visit the class homepage, you will be prompted to enroll. The **course id** and **enrollment key** will be supplied in the first class meeting.

**Catalog Description**.
An introduction to the concepts and methods of probability and statistics for students in life science programs. Topics include descriptive statistics, probability theory, random variables and probability distributions, experimental design, sampling distributions, confidence intervals and hypothesis tests, analysis of variance, and correlation and regression. This course is an alternative to STAT 143 for students in certain life science programs. It is a required course for biology and public health majors and is open to others. No student may receive credit for both STAT 143 and STAT 145.

The course meets the Mathematical Sciences Core requirement.

**Student Learning Outcomes**. Upon completion of this course, students will be able to

- Explain basic principles of study design, describing their role in answering research questions.
- Produce appropriate graphical and numerical summaries of one or two variables (categorical and/or quantitative).
- Use confidence intervals and hypothesis tests to make inferences about a population based on a sample drawn from the population.
- Choose an appropriate statistical model to analyze data in certain common situations.
- Verify whether underlying assumptions justifying the use of a statistical model are met.
- Critically evaluate presentations of statistical results (for example, in journal articles, media pieces, case studies, etc.)

- Use statistical software in the pursuit of the outcomes listed above.

As this is a core course, we formally expect students to

- Apply algorithmic, statistical, and/or mathematical methods to solve problems, broadly defined to find the answers to questions in various domains (as appropriate).
- Identify and use appropriate mathematical and statistical tools for solving a given problem; implementing solutions using the R software package, but with an ability to explain the algorithms used
- Represent, interpret, and process information in graphical, numeric, and/or symbolic forms.

Student achievement with regard to these outcomes will be assessed via homework and test questions.

**Topics** include

1. Structure and organization of data
2. Sampling and study/experimental design
3. Graphical and numerical summaries of data
4. Basic probability theory
5. Probability distributions, and methods for simulating them; Central Limit Theorem
6. Statistical inference via resampling methods
7. Parameter estimation/confidence intervals, taken from settings such as 1-proportion, 2-proportion, single mean, difference of means, correlation/slope
8. Null-hypothesis significance testing, taken from settings such as univariate (goodness-of-fit, 1-sample $t$) and bivariate data (2-sample $t$, chi-square, 1-way ANOVA, model utility)
9. Simple linear regression
10. Multiple regression (if time allows)

**Methods of Evaluation**.

| Assessment | Pct |
|---|---|
| Homework assignments (generally 1-3 per week) | 20% |
| Midterms (3 of them; refer to the class calendar for dates) | 54% |
| Final (May 1, at 1:30 pm) | 26% |

**Expectations and Advice**.

- Do **not** get behind. Keep up with the material and any/all related work.
- Come to class on time each day, equipped with a browser-capable device that you should not access unless directed, and ready to be engaged.
- Read your email during the day to see if there are any notices from me, such as assignment details, things to do to prepare for class, or changes in class procedure.

In addition, you are expected to

- submit homework on time, and managing well the late-passes allotted to you for times of need. As a general rule, homework will be due 1-2 times per week.
- take tests on the specified dates at the specified times.
- follow up with questions over ideas you are not getting before the next class meeting. This implies giving meaningful attention to course content between each class.
- determine for yourself the level of importance of various topics we encounter. Information prior to tests will indicate what chapters/sections are fodder for an upcoming test, but your instructor will make no attempt to "boil things down" to just a few ideas or exercises.
- monitor your own progress. Your scores are always accessible. If, at any point in the course, you wish to estimate your grade, use the app at

  https://scofield.site/courses/s145/wtAverager.html

  to calculate your approximate grade as a weighted average, leaving out those components that have not yet occurred. Be aware that if your performance on some component of the grade produces a negative gpa for that component, this grade averager does not handle that situation accurately and overestimates your grade.

**Accommodations**. Calvin University is committed to providing reasonable accommodations for students with documented disabilities. Students with disabilities requiring special assistance to facilitate participation in this class are urged to contact Disability Services in the Center for Student Success (disabilityservices@calvin.edu) as soon as possible to explore arrangements. Homework is scored on the basis of correctness, except if otherwise stated. Assignments have a due date you are to observe. Nevertheless, I have provided you an allotment of **late passes**, which can buy you extra time:

- Usually 72 hours is offered, but be advised to take note the new due date on any occasion you use a late pass, as the amount of time can vary. Whatever the amount of time offered, there is a hard deadline no one is allowed to surpass. (So, if on some particular assignment, 24 hours of extra time is afforded through a late pass, and 11 pm Monday is the official deadline, then when Bob uses a late pass at 11 pm Monday and Sharon uses one at 9 am on Tuesday, both will need to submit their answers by 11 pm Tuesday.)
- There is no deduction for using a late pass, but you generally will not be able to submit an assignment at all (it remains in "practice mode" only) for credit once the added time expires. (Submitting a partial homework set for some points is preferred over getting a zero for having missed the deadline.)
- Late passes in MyOpenMath are not meant to allow to view answers after the deadline, and then submit those answers once you've enabled an extended deadline. In fact, there are safeguards against this sort of use. Be warned that, should you make the attempt, even inadvertently, **you will lose the late pass without gaining a new deadline**.

- I believe the clock begins when you "use" the late pass, so you can optimize your extended time by waiting until the assignment comes due.
- You begin with 10 late passes. Treat this allotment like a bank account. Once they are all used up, there can be no more spending. Saving them except in cases of dire need takes discipline, but it lessens the likelihood that you'll find yourself in the hospital with pneumonia but with no remaining late passes.

Assignments provide an opportunity to rehearse the ideas you are to master: thinking (and writing) in the language of statistics, working with data, doing calculations, and making a statistical argument. Students who approach it that way generally get the most out of the experience. And, yes, they usually earn good homework scores, too. In contrast, it is possible to "stress out" over points to the extent that form takes precedence over substance, performance (or end-product) seems more important than the rehearsal. Guard against replacing the right emphasis with a wrong, short-sighted one.

I hope you will collaborate with others to learn the material of this course. Doing homework together is encouraged, but you must write it up in your own words to make it your own. To do otherwise is one of several things that constitutes "academic dishonesty." I define cheating in a manner that includes, but is not limited to, use of unauthorized sources, notes or devices, and copying from the work of another student or knowingly allowing a student to copy from your work. Such behavior is unacceptable, and will, in the first instance, result in a score of zero. A second offense will result in a course grade of "F". Read more in Calvin's Student Conduct Code.

**Technology**. As indicated in the SLOs, technology plays a big role in the course. You are expected to develop a facility in the use of R, an industry-standard statistical software package. Alongside this, you are expected to learn to use Quarto, a report-writing language that dovetails with R. Many examples of the use of both will be given in coordination with class, where I will work within an integrated development environment called RStudio. By the date of our first class meeting, an account for you shall have been created on the Calvin RStudio server; your login name and password will be your Calvin login and password. Note: R and RStudio can be installed for free on your own computer, but the installation process can be tricky, and I offer you no assistance in configuring it.

While some questions on tests will call upon your ability to write R commands with a specific purpose, the real benefit to learning R and Quarto is in the ease they provide for drawing on large data sets, conducting analysis of the data quickly and easily, and writing reports that facilitate understanding and repeatability of your methods to other practitioners. Technology does not remove the responsibility to *own* the statistical methods taught in the course. While calculators (not phone-apps, but real, dedicated calculators) are allowed for in-class tests, and you are allowed to employ the push of a button to carry out *elementary mathematical operations*

such as multiplying or taking a square root, when a calculation specific to some statistical method is called for, no "magically-appearing answer" arising merely from button pushes on a calculator will suffice.

**Attendance**. Please forgive the pedantic tone, as much of What I say, here, is probably assumed by most of you. Still, some person or committee is insisting that I state a policy.

On the subject of attendance, I espouse a natural consequences approach. Generally speaking, those who have a weak record attending class for whatever reason—don't like the time of day it meets, have conflicts with work schedules, have a test in another class to study for, get sick—end up with a poor understanding of course material, and get poor grades as a natural consequence. (I do not pile on artifical consequences.) Avoid that path, if possible. If missing class is unavoidable, have a friend in the class you can turn to, and a plan for how to catch each other up.

Applying similar reasoning, showing up late for class, or leaving early, should be avoided if possible. Hold high standards for yourself so that, when you meet them, you will know yourself to be taking care of the business-at-hand for a serious student. That the rest of us will think you a serious student, as well, seeing you exercise good habits, will be a natural consequence.

And, while showing up is Step 1 toward being a serious student, there remain all kinds of ways a student can get derailed. Trying to attend to social media, remaining passive when the instructor gives you an in-class activity, arriving without sufficient rest to remain attentive—these are habits that erode the ability to become knowledgeable in the subject matter. Give yourself the best foundation for learning by exercising healthy habits.

**Exceptions**. I reserve the right to make changes or exceptions to course policies, including those described in this document, either for the entire class or for individuals. The ultimate goal in this course is **learning**, and formal requirements should not unnecessarily stand in the way of that. Thus, if you think that any of the conditions of the course are interfering with learning, please speak with me about this, and we will consider what can be done.