Welcome to
Meaningful Statistics

Neil J. Hatfield

Department of Statistics,
Pennsylvania State University

January 2020

Dear Student/Teacher/Researcher,

Welcome! I'm glad that you are starting your journey into the field of Statistics. I firmly believe that Statistics is one of the most useful fields in universe, and this isn't hyperbole. I want you to think about any field for a moment...go on pick one[1]. Whatever field you chose, there is at least one application of Statistics in that field. Whether you picked business, accounting, sports, medicine and nursing, biology, chemistry, sociology, psychology, physics, mathematics, education, English, history, fashion design or any of the other fields I've not mentioned, Statistics is there. While I cannot catalog each and every way, here are few examples:

- Business: the development and testing of products.

- Accounting: detecting fraudulent entries in a business's books

- Sports: helping teams and individuals improve–this is a specialized sub-field of Statistics called sabermetrics.

- Medicine and Nursing: the development of new treatments for diseases and improving patient care.

- English: there is a growing trend in using Statistics to verify who authored what; Shakespeare didn't necessarily write everything his name is on.

- Fashion: ever heard of or used StitchFix? A statistician developed the tinder-like swiping for outfits that helps the stylists pick out the perfect outfit for you.

- and the list goes on.

Statistics touches everything. John W. Tukey, one of the greatest statisticians immortalized this with his statement that "The best thing about being a statistician is that you get to play in everyone's backyard."

---

[1]Yes, I will speak directly to you throughout this work. Imagine me sitting across from you and us having a conversation; that is what I'm doing as I write this. I will also use footnotes throughout for additional commentary.

Now that I have talked up Statistics, I must now admit that Statistics is a challenging field. Not only does Statistics inherent all of the problems and challenges present in Mathematics, we also face many of the challenges present in every other field, plus we have challenges that are unique to Statistics. Now, this does not mean that you have to be a genius in order to "get" Statistics. There are plenty of individuals who are brilliant but struggle with Statistics. There are also individuals who feel that they struggle with mathematics but find themselves thriving with Statistics. In my experience, there are several key things that you can do to ensure that you have success in Statistics:

1. **Believe in Yourself.** Yes, I know that this sounds cliche, but if you believe that you are bad at Statistics, *you will make yourself bad at Statistics*. Statistics (and mathematics) ability is not some trait you're born with. Statistics ability follows the notion of a self-fulling prophecy, especially when you are discounting yourself. Believe in yourself and that you can develop your Statistical ability.

2. **Practice, Practice, Practice.** While thinking you're bad at Statistics will cause you problems, thinking that you're good at Statistics does not guarantee that you will succeed. Statistical thinking is often something that most people don't come by naturally. There are plenty of pitfalls that come with statistical thinking[2].

   Something that students often don't realize is that you need to practice outside of class. A credit hour is defined as a one hour of direct instruction and a ***minimum*** of two hours of student work outside of class **each week** for the semester. Thus, for a 3-credit hour class, this means that you should spend around 6 hours each week on the course...not counting actual class time. For a 15-week semester, you should spend at least 90 hours working on that course. Unfortunately, there is some evidence that as much as 85% of students will not make the 90 hour benchmark and 60% will not even put make the 45 hour benchmark[3]. There is a direct, positive relationship between the amount of time you spend outside of class working on that material and your course grade.

3. **Ask Questions.** Knowledge is not transmitted from teacher to student; rather the student builds their knowledge piece by piece. Asking questions serves a dual purpose: first, questions help you build your knowledge and second, your questions help your teacher build a model for how you are thinking[4].

   Two of the key opportunities for asking questions that students often miss out on for posing questions is student hours and readings. "I wish that students would use student/office hours less." said no instructor ever. Rather, what we say is "You're sure kidnapping students to have them come into student hours is against the law? Are you sure that there isn't any leeway?" Students often avoid going to see their teachers until things are absolutely dire. The real power is to go in early, often, and ask questions...even when you think you have

---

[2]We'll explore several of these pitfalls in the coming days.
[3]That's one hour outside of class for each hour in class for the semester.
[4]Which helps them better help you

a handle on the concepts.

The second key opportunity for asking questions is when you are reading. Sadly, most students have not been taught how to read technical writing...which is what every textbook is. The approach you take with a nonfiction or fiction book often does not work with textbooks. When you are faced with a reading assignment, start early...don't procrastinate. Additionally, you should "arm yourself" for reading. This means that you not only get the reading, but you make sure that you have access to a dictionary[5], highlighters, and most importantly, a writing utensil. As you read, actively engage with the author. Make notes in the margins or on separate page about what you're questions are, what connections you're seeing. This will help you build a better understanding of the reading and give you spots where you can ask questions. If you encounter a word that you don't know, stop and look that word up; do not move on until you understand how that word is being used by the author[6]. When you're finished reading, write out 1) a summary of what you read, 2) what you understood the author to be saying, 3) any connections you've made between that reading and others, and 4) what questions you have[7].

4. **Be Flexible and Persevere.** When students struggle in Statistics, my experiences have shown me several root causes. The most common root cause is that the student comes in with a strong preconception of how a "stats" class or textbook should look. These students are also the most likely to also believe that you are either good at math or you aren't. This fixed mindset causes the student to be rigid in their thinking. Unfortunately, this rigid mentality is anathema to Statistics. Statistical thinking demands that you be open to multiple ideas and approaches. Just because a class or textbook looks different does not look like another one does not mean that something is wrong.

The second root cause is that students often persist rather than persevere. There is a subtle distinction between these two terms that most people aren't aware of; I wasn't aware of the distinction until well into my career. To highlight this, I want you think of a time when you were working on an assignment or project when all of a sudden you encountered an obstacle that you didn't anticipate when you first started out. You have three choices: 1) you can stick with your original plan and attempt to overcome the obstacle, 2) you can alter your plan and seek out additional resources to help you overcome the obstacle, or 3) you can quit.

The third option is by far the easiest but often comes with the abandonment of your broader goals or additional set backs you now have to overcome. The first option, sticking with your original plan, is what most students do, and is the definition of *persistence*. However, this approach can have a student feel

---

[5]This includes not only the traditional dictionary but also any course dictionaries like the Statistics Dictionary that comes with this text.

[6]This is one of the most important differences between reading non-fiction/fiction and reading technical works.

[7]This is what I refer to as a prècis, which is much more useful than a plain summary.

like they are spinning their wheels and are perpetually stuck and/or struggling. Persisting is not a bad thing but can create problems when the student's original plan *is* the obstacle blocking their progress. This leads us to the second option: altering your plan and seeking out resources. *Perseverance* is the continuance of effort, carried out in a thorough and diligent manner, towards some perceived goal while overcoming difficulties, obstacles, or discouragement along the way by amending your plan of attack. By moving away from persisting and towards persevering, a student can be successful in Statistics (and any other course). This can be challenging, but you can do this if you put in the effort.

5. **Don't Be Afraid to Get Help...But Be Cautious.** If you feel yourself getting overwhelmed, come in and get help from your instructor. Make use of student/office hours; if those times don't work, email your questions and ask about setting up an additional appointment. When you get help, have specific questions in mind: this will help you get the most out of time. Seek out help from additional sources including all course materials and your classmates.

   However, you do need to have some care when seeking help from outside sources such as non-institution affiliated tutoring, the internet, and even non-instructor affiliated tutoring. In particular, be careful with the internet. While the internet is a boon, it is also a bane in that many people can post "help" that is actually factually incorrect and/or supports problematic ways of thinking. While true in many fields, this definitely holds true in Statistics. For example, Kahn Academy has a plethora of videos purporting to explain ideas. While these videos do explain ideas, they often get things wrong or support students in developing meanings that are counterproductive.

6. **Don't Give Up.** As I've said, Statistics can and *will* challenge you. Statistical thinking requires that you confront ways of thinking that appear intuitive and easy but are anything but. Additionally, you will have to confront the fact that there are many statistical words used in everyday conversations that are divorced from their original meanings in Statistics. For example, "random", "correlation", and "average" are three such words. However, if you stick with this, you'll be able to construct powerful ways of thinking that will support you through this course and beyond.

By keeping the above points in mind throughout the course and following through with them, you will have success in Statistics and you'll find Statistics meaningful, rather than just a collection of procedures.

Course Design

I've built this course around two key principles: that the course should reflect the inherent nature of Statistics and that the course should be *meaningful*. The core purpose of Statistics is to answer questions about collections of objects or living beings so that we can better understand our world. To this end, each unit of the course revolves around a central question (or several interrelated questions). Throughout each unit, I will guide you through an exploration where we will seek out an answer (or multiple) to each guiding question.

The second principle is the much more challenging one and is the one that will throw most people off balance. While the first principle helps establish an ordering of topics, the second principle is the one that drives the development of the materials. Far too often, introductory statistics courses proclaim themselves "conceptual" but upon close examination, they are purely procedural. When I say that this course is conceptual, I mean that each element of this course has been chosen and presented in a way to help you build a solid foundation for your thinking. The meanings and ways of thinking that I have chosen to target through the curriculum are on the most productive end of the spectrum. However, this does come at a cost: you will be challenged. There is a pragmatic consequence to my decisions: you will be better equipped to succeed in life.

I want you to imagine two individuals who are equivalent in every way but one: one individual learned statistics in the traditional, procedural way, and the other individual learned statistics through this course. Both are up for the same job, which individual does the employer choose? The one who has ways of thinking and productive meanings. The student from the traditional course can calculate values for the statistics, but nothing more. In essence, the individual who learned statistics in a typical course is nothing more than a slower, more error-prone calculator. The individual who has productive meanings for Statistics can explain and interpret the values they find as well as make decisions about what to do when faced with an open ended situation...which is what you'll often find in life. (You aren't told what statistic to calculate or what test to run, you have to decide for yourself.) Be this second individual. While the first individual will have a much easier time learning statistics, they have been set up for failure later on in life.

A Work in Progress

Please keep in mind that much of what you'll read from me is a work in progress. I'm constantly adapting the text to better suit the development of student thinking. You might encounter typos in the readings, or changes in the way in which I approach a certain topic at two different moments in time. Please let me know of these issues so that I can address them. While writing a text on Statistics is a passion project of mine, there are many other things that take priority[8]. This does mean that I don't have chapters for every topic and sometimes the chapters might abruptly end. I ask that you bear with me in this endeavor as this text is the convergence of multiple research programs into student thinking, building an online homework data bank, the development of apps, the integration of technology, and personal growth.

This all being said, I wish you the brightest of futures as you begin your journey into *Meaningful Statistics*.

---

[8]For example, you.