Welcome to Math 161 A - Applied Probability and Statistics

Let me introduce myself: My name is Martina Bremer. The best way to reach me is through Canvas or per e-mail (martina.bremer@sjsu.edu). I will try my best to reply to you within 24 hours (36 hours on weekends).

For prerequisites, grading policies and course information please refer to the course syllabus. The schedule as well as many other materials will be available throughout the semester on Canvas

```
https://sjsu.instructure.com/
```

Canvas is the place to go to find lecture notes, videos, homework assignments (and later their solutions), quiz solutions, review materials, help with technology used in this course, my office hour schedule and more. Your assignment and quiz scores will be posted on Canvas.

If you should ever find yourself falling behind in this course please contact me (per e-mail or on Canvas) as soon as possible so that we can discuss all the options you have of catching up. Math and Statistics classes tend to be very sequential. If you miss a section you will likely have difficulties following the subsequent sections.

The Goals of this Course

Math 161 A will provide you with an overview of the most important concepts in elementary probability and statistics. You can think of probability as the more theoretical "groundwork" and statistics as the application of this theory.

What exactly is the difference between Probability and Statistics?

Probability is the science of studying the make-up of samples randomly drawn from a known population. Statistics is the science of drawing conclusions about the population from a sample.
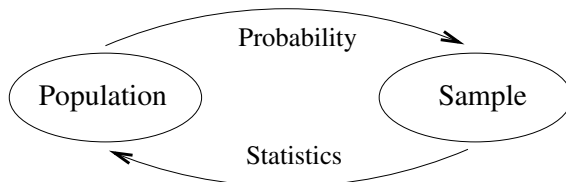
Example: Consider a large jar full of colored marbles. Somebody draws a handful of five marbles from the jar. Define the population and the sample in this context.

Phrase a meaningful question that a probabilist might ask. What are the assumptions you have to make to be able to answer this question?

Conversely, a statistician would make an observation on a sample, and then ask a question about the population.

Phrase a meaningful question on the population in the marble jar context. How could the sample data be used to answer the question?

This Course

The context of Math 161 A will fall mostly ($\sim 80\%$) into the "Probability" category. We will work our way up to the most important statement in Probability, the Central Limit Theorem. At the end of the semester, we will discuss confidence intervals and the fundamentals of hypothesis testing which are statistical applications of the Central Limit Theorem. Math 161B (the continuation of this course) will cover elemental concepts in Statistics, such as regression and analysis of variance (ANOVA). In order to understand these statistical models it is important to first obtain a solid background in probability.

Like calculus, probability allows us to formulate and solve a certain category of problems. Unlike calculus, probability models can take uncertainty into account. Most real life phenomena (e.g., games, the stock market, the weather) are not absolutely predictable. The outcomes are partially random - that means they can depend on chance. We are interested in *events* such as

- winning a game;

- whether or not a particular stock goes up by more than 5% next year;

- whether companies will be looking to hire more applied mathematicians or more statisticians and data scientists next year?

Knowing more about the chances that these events will happen will help us make smart decisions like

- changing your game strategy;

- buying more stock;

- change your major.

2

One important goal of this course is to learn how to describe randomness and to compute the probabilities of certain events.

Example: Suppose your instructor offers you the following bet:

"I bet that there are at least two people in this room that are born on the same day of the year."

Please note that this is a hypothetical example as I'm not actually allowed to gamble with my students for money :-)

Would you take that bet? Would you take it if I up the ante (i.e., if I win, you'll pay me \$1, but if you win I'll pay you \$2)? If not, what do you think would be a "fair" amount to bet? Does it depend on the number of people in the class?

Answer:

Every year, US News - World report publishes a list of the 100 "best jobs" of the year. I am including an excerpt of this list of jobs that would require students to know some probability and statistics.

Data Scientist
#22 in 100 Best Jobs
Data scientists use technology to glean insights from large amounts of data they collect. Read More »
Projected Jobs 40,500
Median Salary \$100,910
Education Needed Bachelor's

Actuary
#27 in 100 Best Jobs
Are you more of a risk calculator than a risk taker? Consider working as an actuary. These professionals are experts in uncertainty, using mathematics, statistics and financial theory to measure, manage and mitigate financial risk. Read More »
Projected Jobs 5,900
Median Salary \$105,900
Education Needed Bachelor's

Statistician
#30 in 100 Best Jobs
Statistics is the science of using data to make decisions. This is relevant in almost all fields of work and there are many opportunities for employment. Read More »
Projected Jobs 11,200
Median Salary \$95,570
Education Needed Master's

What are some other reasons why it makes sense to study probability and/or statistics (other than that you may need it eventually for your major)?

- Reason Number 1: Be informed

  Whenever a report (TV news, Facebook, Twitter (X?), research paper) contains a lot of numerical information, the information is often condensed in the form of numerical statistics. You need to know what those mean in order to be able to correctly interpret the presented information. Note that this will also enable you to detect (not so uncommon and sometimes even intentional) mistakes in an argument.

  Come up with at least one example of statistical information that you have recently come across in the media.

- Reason Number 2: Make informed judgments

  Is the information you have collected to answer a scientific question sufficient or does more information need to be collected to reliably answer the question? Does the method with which the information is collected matter?

  Can you think of an example where you have (or will) collect data in the context of your studies, work, hobbies or research?

- Reason Number 3: Evaluate the decisions that others make for you

  We're in another election year. Soon, politicians will be running ads that claim that their proposed policies are based on data driven decision making.

  Name one context area in which laws or regulations are made based on data that affect you.

NOTE: The book *"How to lie with Statistics"* by Darrell Huff is one of the most popular among all statistics (text-) books.