# Introduction
# Class 1a, TF, AID-M
Jeremy Orloff and Jonathan Bloom

## 1 Probability vs. Statistics

In this introduction we will preview what we will be studying in 18.05. Don't worry if many of the terms are unfamiliar, they will be explained as the course proceeds.

Probability and statistics are deeply connected because all statistical statements are at bottom statements about probability. Despite this the two sometimes feel like very different subjects. Probability is logically self-contained; there are a few rules and answers all follow logically from the rules, though computations can be tricky. In statistics we apply probability to draw conclusions from data. This can be messy and usually involves as much art as science.

**Probability example**
You have a fair coin (equal probability of heads or tails). You will toss it 100 times. What is the probability of 60 or more heads? There is only one answer (about 0.028444) and we will learn how to compute it.

**Statistics example**
You have a coin of unknown provenance. To investigate whether it is fair you toss it 100 times and count the number of heads. Let's say you count 60 heads. Your job as a statistician is to draw a conclusion (inference) from this data. There are many ways to proceed, both in terms of the form the conclusion takes and the probability computations used to justify the conclusion. In fact, different statisticians might draw different conclusions.

Note that in the first example the random process is fully known (probability of heads = 0.5). The objective is to find the probability of a certain outcome (at least 60 heads) arising from the random process. In the second example, the outcome is known (60 heads) and the objective is to illuminate the unknown random process (the probability of heads).

## 2 Frequentist vs. Bayesian Interpretations

There are two prominent and sometimes conflicting schools of statistics: Bayesian and frequentist. Their approaches are rooted in differing interpretations of the meaning of probability.

**Frequentists** say that probability measures the frequency of various outcomes of an experiment. For example, saying a fair coin has a 50% probability of heads means that if we toss it many times then we expect about half the tosses to land heads.

**Bayesians** say that probability is an abstract concept that measures a state of knowledge or a degree of belief in a given proposition. In practice Bayesians do not assign a single value for the probability of a coin coming up heads. Rather they consider a range of values each with its own probability of being true.

In 18.05 we will study and compare these approaches. The frequentist approach has long been dominant in fields like biology, medicine, public health and social sciences. The Bayesian approach has enjoyed a resurgence in the era of powerful computers and big data. It is especially useful when incorporating new data into an existing statistical model, for example, when training a speech or face recognition system. Today, statisticians are creating powerful tools by using both approaches in complementary ways.

# 3 Applications, Toy Models, and Simulation

Probability and statistics are used widely in the physical sciences, engineering, medicine, the social sciences, the life sciences, economics and computer science. The list of applications is essentially endless: tests of one medical treatment against another (or a placebo), measures of genetic linkage, the search for elementary particles, machine learning for vision or speech, gambling probabilities and strategies, climate modeling, economic forecasting, epidemiology, marketing, googling... We will draw on examples from many of these fields during this course.

Given so many exciting applications, you may wonder why we will spend so much time thinking about toy models like coins and dice. By understanding these thoroughly we will develop a good feel for the simple essence inside many complex real-world problems. In fact, the modest coin is a realistic model for any situations with two possible outcomes: success or failure of a treatment, an airplane engine, a bet, or even a class.

Sometimes a problem is so complicated that the best way to understand it is through computer simulation. Here we use software to run *virtual* experiments many times in order to estimate probabilities. In this class we will use R for simulation as well as computation and visualization. Don't worry if you're new to R; we will teach you all you need to know.

---

Note by Prof. Mayer: This script is the work of the original authors (Jeremy Orloff and Jonathan Bloom), which appear in the title and all praise for this work must be to them. The only modifications done to fit it to "Theoretical fundamentals of AI and Data Science" (DIT, Faculty AI, Study course AIN-B) are leaving out certain topics and changing the class numbering.

The complete original script (and more: slides, exams, question sheets, and solutions) can be downloaded at MIT OpenCourseWare: Introduction To Probability And Statistics, 18.05, Spring 2014, Undergraduate



Prof. Mayer thanks Jeremy Orloff and Jonathan Bloom not only for making the course material open to the public, but also for giving him access to the original LaTeX sources to tailor the material better to the DITs needs. Without their help, this course would by far not have the quality of its current state.