Math 335: Course review

We are half of the way through our course in Probability and Statistics. The first semester was mostly about Probability.

Probability is about understanding the relative frequency of certain events happening in repeatable situations resulting in chance outcomes. Example: Flip a fair coin 100 times; find the probability of obtaining 65 heads.

Statistics is the science of finding information in data, data collected to solve some question from some "scientific" endeavor. We look for patterns, signals, information within the data, but the patterns are obscured by random variation or "noise." To find the patterns we need an understanding of how "noise" behaves, which requires a knowledge of probability. Example: You do not know the probability, p, of "heads" for a coin (it may be biased). You flip the coin 100 times and obtain 65 heads. Estimate p; provide a 95% confidence interval.

We have learned **3 views of probability**:

(1) empirical

(2) classical

(3) axiomatic

The empirical view is sort of the definition. The classical view is the simplest way to imagine computing probabilities. The classical approach is very limited if we are restricted to just listing out the sample space, but combinatorics (sections 2.6 and 2.7) allows us to count large sets without physically listing them out.

We introduce an important other concept in chapter 2 called **conditional probability**, which is about probability situations where we are given partial information about an event and asked to re-consider the probability of the event. A special case is when the partial information is unrelated to the calculation, which gives the concept of **independent events**. Repeated independent trials is a special situation of the latter that allows us to compute complicated probabilites by multiplying simpler ones together.

The **axiomatic view codifies basic properties** of the other views and leads to a body of methods that expands our ability to calculate probabilities in more complicated situations. In particular, we get important methods such as the law of total probability (theorem 2.4.1), Bayes's Theorem (theorem 2.4.2), and multiplying probabilites in a sequence of independent trials (section 2.5; page 78 and on).

As stated above, statistical work will require us to understand how variation occurs around patterns. To understand this, we need the concept of a **random variable**, one view of which is that a random variable is the mathematical concept that describes variation about a pattern. In chapter 3, we worked our way through **various concepts related to random variables**: pdf's and CDF's to describe the variation, joint distributions when there are two random variables defined on the same sample space (pdf's only; we ignore joint CDF's), combining and transforming existing random variables into new ones (we focussed on the CDF method), order statistics (a special kind of transformation useful in ranked-based methods of statistics), conditional pdfs, and moment-generating functions.

Some types of random variables occur so frequently in applications that they are worth careful study and they are "named." These are the hypergeometric, binomial, Poisson, normal, negative binomial (geometric a special case), and gamma (exponential a special case). For these, we learned their pdfs, their means and variances, their MGFs, and the type of problems they applied to.

Normal distributions have a special place in classical probability and in the classical statistical methods we will study. This special place is conferred largely because of the amazing Central Limit Theorem (theorem 4.3.2), of which the Demoivre-Laplace theorem (theorem 4.3.1) is a special case. Our last topics were beginning statistics, topics in estimation (chapter 5). We learned what an estimator for a parameter is, and learned two methods for finding them: maximum likelihood and method of moments. We also learned the formula to compute a confidence interval for a binomial success probability from a random sample (Theorem 5.3.1); this formula is called the Wald Formula. We will pick up chapter 5 after break.