4 Basics of Probability

Most of this material is covered quite nicely in SW, so I plan to stick with the text very closely.

Populations and Samples

- SW Section 2.8. covers populations.
- SW Section 3.2 covers simple random sampling (SRS). We will use Minitab to illustrate random sampling for Example 3.1 p.74
- Bias in sampling is illustrated in Examples 3.2, and 3.3. The other examples are worth studying also.
- Sampling human populations often involves stratification and/or clustering of individuals into groups. We'll look at this, but for now just note that it is more complicated.

Probability

We will focus on the relative frequency interpretation of probability on p. 80, and the examples.

Probability Rules (Section 3.5) is really a huge topic, requiring more time than the value it adds to your understanding. While it is on the syllabus, we are going to skip it. The material on probability trees is more accessible and gets you most of what you need in probability calculation.

Probability Trees

SW Section 3.4. Trees provide a device for organizing probability calculations. Let us examine in detail examples 15 and 17, and do problem 3.9. We want to get the terms *sensitivity* and *specificity* out of this section, and understand how little information a test may have even with very good values of both.

Density Curves

SW Section 3.6. These are basically histograms of populations standardized to have and area of 1 under them, so that area can be related to sampling probability. We will do some simulations in Minitab to see how histograms of huge sets of numbers can look like smooth curves. We will cover Example 3.30.

Random Variables

SW Section 3.7. All we really want to cover is the definition. This is a mathematical model for a population. Populations have means (μ) and standard deviations (σ) , and the idea is identical for random variables. We will skip that part of this section.

Binomial Distribution

SW Section 3.8. This is our model for binary outcomes. We really want to understand well the *Independent-Trials Model* on p. 103. We will use Minitab to do the calculation for Example 3.45, and we want to understand how the model breaks down for Example 3.50.

The Normal Distribution

SW Sections 4.1-4.3. We will just get a good start on this and continue next time. We need to see how to use a table like Table 3 p. 675-6, although we will see how to get the answers more easily out of Minitab. A great deal of what we do in statistics involves normal distribution calculations, and while those usually are done within software we need to understand what is being done behind the scenes.