Week 4: Discussion

David Diez

Probability :: a review of how to approach problems including

- independence :: when one event has no influence on another. ex: flipping two coins (they are independent of one another).
- disjoint probability :: if one event happens, the other cannot (and vice versa). ex: rolling a 2 or a 3 with a die.
- conditional probability :: the most important formula is

$$P(A|B) = \frac{P(A\&B)}{P(B)}$$

  Rearranging this slightly, the following equation is also important: $P(A\&B) = P(A|B)P(B)$.
- binomial problems (These problems are when the order of events doesn't matter. For instance, if you flip a coin and want to know the probability of 4 out of 10 flips being heads, then you don't care whether the first flip was heads, just whether the total number of heads was 4. When order doesn't matter, think binomial probability.) If a problem is identified as a binomial problem, find the number of trials (n), the number of 'successes' (x), and the probability of a single success (p). Then the probability of exactly x successes in n trials is

$$P(\text{exactly } x \text{ of } n \text{ trials}) = \left(\begin{array}{c} n \\ x \end{array} \right) p^x (1-p)^{n-x} = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$$

- when you see the words "at least", think about using the complement.
  for example: P(at least 1 of 3 M&M's is orange) = 1 − P(none of the 3 M&M's is orange)

Review of quiz 4, attempt 1 ::

Remarks on labs ::

- **compare statistics** :: example :: "The median for females is 2.56, which is 0.18 higher than the male median of 2.38." Listing statistics is unimportant (since the tables offer that information explicitly) and the goal is to think about the similarities and differences in the distributions.
- biomodal :: gave points for either way since it was a bit inconclusive
- outliers :: Large data sets will always produce outliers using the 1.5*IQR rule, which may mean that we can deviate a bit from this rule when we have a lot of data (like this case) or can view something like a histogram.

- express what statistics mean :: if males have a larger standard deviation than females, what does that mean? "The distribution for males is more spread out than females – there is more variability for a higher standard deviation."

Worksheet on simulation :: worksheet handed out and worked on in class.