Probabilities

This will be fun

Ryan Erickson

Learning Goals

- 1. Use the multiplication and addition rules to calculate probabilities
- 2. Understanding Bayes' theorem and the applications in problem solving
- 3. Write your first for loop in R

Calculating Probability

Cards

- Probability is simply the chances that something will happen
    - P(Event) = number of times event occur / total number of events
- A deck of cards has 52 cards. It's separated into 4 suits (Diamonds, Clubs, Hearts, and Spades). Each suit has 13 cards labeled 2-10, A, J, Q, and K.
    - What is the probability that a King is pulled from the deck?
    - What is the probability that the Ace of Spades is pulled from the deck?

What is the probability that a King is pulled from the deck?

- 52 cards
- 4 kings in the deck
- P(K) = 4/52 = 1/13 = 0.0769 = 0.08

What is the probability that the Ace of Spades is pulled from the deck?

- There is only 1 Ace of Spades in the deck
- P(A) = 1/52 = 0.0192 = 0.01

Now some more difficult probabilities

What happens when we want more than one card?

- Things get a little weird when using ‘and/or’
  - What is the probability we draw a K or an A?
  - What is the probability we draw a 7 and a 3?
  - What is the probability of drawing a 3 of a kind?
- ‘and’ denotes multiplication probabilities usually denoted as
- ‘or’ denotes addition of probabilities

What is the probability we draw a K or an A?

- There are 4 kings and 4 aces
- P(K) = 4/52 = 0.0769 = 0.08
- P(A) = 4/52 = 0.0769 = 0.08
- P(K or A) = P(K) + P(A) = 0.08 + 0.08 = 0.16

What is the probability of drawing a 7 and a 3?
2 cards

- The probability of drawing a 7 is still 4/52
- The probability of drawing a 3 is still 4/52
- P(7 and 3) = 4/52*4/52 = 1/13*1/13 = 1/169 = 0.0059 = 0.006

What is the probability of getting a 3 of a kind?

- The probability of getting one number on the first draw is 4/52
- Getting the same number on the second draw is 3/52
- Finally the odds of getting the same number on the third draw is 2/52
- Therefore:
    - P(3x) = 4/52*3/52*2/52 = 0.00017 or 0.0002.

In terms of percentage

Just multiply by 100!

- Probability of getting a Q
    - 0.08 * 100 = 8%
- Probability we draw a K or an A
    - 0.16 * 100 = 16%
- Probability we draw a 7 and a 3
    - 0.006 * 100 = 0.6%
- Probability of drawing a 3 of a kind
    - 0.0002 * 100 = 0.02%

Bayes Theorem

This can be a little complicated

- P(A | B) = P(B | A)*P(A)/P(B)
- We don't always have P(B) directly, but we can calculate it from the law of total probability:
    - P(B) = P(B|A)P(A) + P(B|Not-A)P(Not-A)
- Remember:
    - P(B|Not-A) = 1-(P(Not-B|Not-A))
    - P(Not-A) = 1-P(A)

Sensitivity and Specificity

- **Sensitivity** is the true positive rate. It is a measure of the proportion of correctly identified positives.
    - Usually P(B|A)
- **Specificity** is the true negative rate. It measures the proportion of correctly identified negatives.
    - Usually P(Not-B|Not-A)

Bayes Example

Drug Testing

- For example, consider a drug test that is 99 percent sensitive and 95 percent specific. If half a percent (0.5 percent) of people use a drug, what is the probability a random person with a positive test actually is a user?
    - Bayes: $P(A \mid B) = P(B \mid A)P(A) / P(B)$
- In the context of the problem, we are looking for $P(\text{user} \mid +)$ so:
    - $P(\text{user} \mid +) = P(+ \mid \text{user})P(\text{user}) / P(+)$

- P(user) = 0.5% or 0.005
- P(non-user) = 0.995
- P(+ | user) = 0.99 (Sensitivity)
- P(- | non-user) = 0.95 (Specificity)
    - We can re-write this P(+ | non-user) = 1 - P(- | non-user)

- Final Form
    - P(user | +) = P(+ | user)P(user) / [P(+ | user)P(user) + P(+ | non-user)P(non-user)]
    - P(user | +) = (0.99 * 0.005) / (0.99 * 0.005+0.05 * 0.995)

- P(user | +) ≈ 9%

What's going on for the lab?

The Dataset

- The data set has 4 columns:
    - ref_dice1
    - ref_dice2
    - new_dice1
    - new_dice2
- The ref_dice are reference dice that you know are fair
- The new_dice are new dice that you are checking