April 13, 2015 –

- **Bayes' Rule -**
  - P(A|B) = P(A) * P(B|A) / P(B)
    - A = some statement
    - B = data; new information
      - Example -
        - A = accused is guilty
        - B = evidence at trial
    - Example -
      - DNA matches
      - **P(A) = 1/10,00,000**
      - **P(B |~A) = 1/1,000,000 → false positive**
      - Assumption – if accused is guilty, than **P(B|A) = 1**
        - There will be a match in the DNA evidence
      - Solve for P(B) -
        - P(B) = P(B and A) or P(B and ~A)
          - DNA matches and person is guilty *or* DNA matches and person is not match
          - Mutually exclusive statement
        - Product rule…
        - = P(A)*P(B|A) + P(~A)*P(B|~A)
        - = 1/10,000,000 * 1 + 9,999,999/10,000,000 * 1/1,000,000
        - **P(B) = 11/ 10,000,000**
      - Solve for P(A|B) –
        - P(A|B) = [(1/10,000,000) * 1] / (11/10,000,000)
        - = 1/11
        - **In light of the evidence, there is a 1/11 chance that the accused is guilty**
    - There is a very common problem of assuming P(B|~A) = P(~A|B) which is *false*
    - Think about Bayes' rule in terms of cohorts of people
      - Example:
        - The 10,000,000 people in New York
        - 1 person is guilty → 1 positive test result
        - 9,999,999 people are not guilty → 1/1,000,000 chance that a person is guilty so 1/1,000,000*9,999,999 = 10 positive test results
        - 11 positive test results for the entire city
  - **Discrete random variables –**
    - X = random variable; anything we are sure about
    - Space; probability space; support of random variable = set of possible outcomes
      - Example: trying to guess the name of the person walking by and the space would be all possible names
    - Probability distribution; P(X) = complete description of your uncertainty of the random variable

            - x = random variable
            - X = possible random variable
            - How do we specify this?
                - **1. Brute force list (not standard terminology)**
                    - Table with potential x's on the right and the P(X=x) on the left
                    - Example – price of copper
                        - x= .99, 1.00, 1.01
                        - P(X=x) = .3, .5, .2
                    - Example – the chance of a no show on a 400 person flight

x	P(X=x)
0	0.004
1	0.005
2	0.0062
…	
400	

                    - *Very tedious process so we turn to probability mass function*
                - **2. Probability mass function (PMF)**
                    - P(X=x) = f(x)
                    - Example of PMF –
                        - 1. Binomial distribution (parametric probability model)
                            - X ~ binomial (N, w) → n possibilities, with w probability of each possibility
                            - P(X=x) = (N over x) * W^x * (1-w)^(N-X) = f(x)
                        - 2. Poisson (parametric probability model)
                            - X ~ poisson (λ)
                            - P(X=x) = (λ^x / x!) * e^(-x) = f(x)
                - **3. Probability density function**
- **Expected value –**
    - Expected value = "center" of your distribution; weighted average of the possible outcomes
    - E(X) = Sum of X=Ω of P(X=x) * x
        - Sum of X=Ω → summing over all possible events (assuming Ω is discrete)
    - Example –
        - X = stock price of Apple 12 months from today

x	**P(X=x)**
$100	.2
$120	.4
$140	.4

        - E(X) = .02*100 + .4*120 + .4*140
        - =132
- **Decision tree –**
    - Ordered temporally from left to right
    - "Chose your own adventure book"
    - Represents how a sequence of events may unfold in time

- Square and Circle = nodes
  - Square = decision node; place where you have to make a decision
  - Circle = stochastic node; place where a random event's consequences unfold
- Defined by terminal outcomes at the leaves of the tree and need to know the probability distribution at of each of the stochastic nodes
- **Solving a decision tree steps –**
  - 1. Work backwards in time (right to left on the decision tree)
  - 2. Reduce every stochastic node to an expected value
  - 3. At every decision node, take the option of the highest expected value

- Example –
  - Hurricane and house insurance for a $1,000,000 home
  - Buy insurance? –
    - Yes buy insurance
      - Yes hurricane (w) – $1,000,000-c
      - No hurricane (1-w) - $1,000,000-c
        - **E(B**) = w(1mil – c) + (1-w)(1mil-c)
        - = 1mil - c
    - No buy insurance
      - Yes hurricane (w) - $0
      - No hurricane (1-w) - $1,000,000
        - **E(C)** = w*0 + (1-w)*1mil
        - = 1mil – w*1mil
  - **What is a rule for whether the person should buy insurance or not?**
    - E(B) = 1 mil – c
    - E(C) = 1mil – w*1 mil
    - *Buy whenever E(B) > E(C)*
    - E(B) > E(C) = 1mil – c > 1 mil – w*1mil
    - = -c > -w*1mil
    - **= c < w*1mil**