Discrete Probability Distributions

• Last time
  – Random Variables
  – Probability Mass Function
  – Probability vs Frequency Distributions
  – Expected value and variance
  – Cumulative distribution function
  – Permutations and Combinations
Binomial Distribution

• For all cases with this common structure:
  – A sample of $n$ independent trials, each of which have only two possible outcomes
    $P(\text{success}) = p$
    $P(\text{not success}) = 1-p = q$
Examples of Binomial Distributions

- Probability of getting X number of heads when flipping 4 coins
  
n independent trials = 4
  
  For each trial:
  
  \[
P(\text{success}) = p = P(H) = 0.5
  
  \[
P(\text{not success}) = 1-p = q = P(T) = 0.5
  
- Probability of rolling snake eyes
  
n independent trials = 2
  
  For each trial:
  
  \[
P(\text{success}) = p = P(X=1) = 1/6
  
  \[
P(\text{not success}) = P(X\neq1) = 5/6
Bio Example

• White blood cells can be categorized as neutrophils and nonneutrophils.
• A medical exam looks at the sample percentage of neutrophils white blood cells to determine if a patient has an infectious disease.
• For example, 5 cells are examined. A success occurs when a cell is neutrophilis.
• The probability of any one cell being neutrophilis is 0.6
Question 1

• What are the number of combinations of 5 cells taken 2 at a time?
  \[ nCk = 5C2 = 10 \]

• What does this represent?
  The number of ways to select 2 of the five cells to be neutrophils
  Or, think of it as the number of ways to order 5 cells where 2 of them are neutrophils
Question 2

• What is the probability that just the 2nd and 5th cells will be neutrophils (x) and the remaining nonneutrophils (o)?
• The outcome would be oxoox
• Is the cell type independent from other cells?
• \[ P(\text{oxoox}) = q^2p^3 \]
  \[ = (0.6)^2(0.4)^3 = 0.023 \]
Question 3

• What is the probability that any 2 out of the 5 cells will be neutrophils?
  – Remember there are 10 ways of ordering
    \[ S(X=2) = \{xxooo, xoxoo, xoxxo \ldots \} \]
  – The probability of getting 2 neutrophils is the total probability of all the orderings
  – Remember total probability is the sum of the individual event probabilities
  – \[ P(X=2) = 10*p^2q^3 = 10*0.023 = 0.23 \]
Binomial Distribution

• The binomial distribution of the number of successes in $n$ independent trials where the probability of success on each trial is $p$ and has a probability-mass function given by

$$P(X=k) = \binom{n}{k} p^k q^{n-k} \quad k=1,2,...,n$$
Pulmonary Disease

• An investigator notices that children develop chronic bronchitis in the first year of life in 3 of 20 households where both parents are chronic bronchitis, as compared with the national incident rate which is 5% in the first year of life.
Questions

• Is this difference “real” or can it be attributed to chance?
• What is the expected value for the number of incidences in 20 households?
• How likely are infants to develop bronchitis in at least 3 out of 20 households if the probability of developing the disease in any one household is .05?
See example in book for solution details
Binomial Distribution

- The expected value and the variance of a binomial distribution are $np$ and $npq$, respectively.
Binomial Distribution

- 4 coin toss example
  N=4, p = .5