David Youngberg
BSAD 210—Montgomery College

LECTURE 14: DISCRETE PROBABILITY FUNCTIONS I

I. Discrete probability distributions
  a. A *distribution* lists all the possible results with the frequency of each result. It's typically presented in a graph form.
  b. A *discrete probability distribution* is distribution of data made up the results from a discrete random variable (which has outcomes of a small range of whole numbers).
  c. Later we will discuss a *continuous probability distribution*: a distribution of data made up the results from a continuous random variable (which has outcomes of any numerical value or a large range of whole numbers).
  d. Some examples of a discrete random variable:
    i. The result of a flip of a coin.
    ii. The number of customers in a line in a minute.
    iii. If a bullet hits its target.

II. Binomial Distributions
  a. As the prefix "bi" suggests, binomial distributions deal with the number two. Each observation in such a distribution can be one of two results: success or failure.
    i. "Success" and "failure" is just nomenclature and does not suggest something good or bad happening. A "success" in a test for a disease can be a confirmation that the disease is present.
    ii. p is the probability of a success
    iii. q is the probability of a failure
    iv. Because there are only two outcomes, $p = 1 - q$.
    v. <u>Examples</u>: determining if a person needs or does not need corrective lens; testing if a peach is ripe or not; recording if a household has a pet or not.
  b. Binomial distributions assume the probability of success is constant. This means each trial is independent (e.g. each customer has a 10% chance of redeeming a coupon). Other ways to achieve independent trials:
    i. You are replacing each selection after a trial (chances of pulling a poker chip from a bag and replacing it each time); or
    ii. Your sample size is so small compared to the population, you don't affect the probability when you perform a trial. The

threshold for “big enough” is if your sample is less than 5% of your population (probability of finding money in any of 100 trash bags, selecting from the trash bags at the dump).

c. Mean

$$\mu = np$$

   i. *n* is the number of trials
   ii. *p* is the probability of success

d. Standard Deviation

$$\sigma = \sqrt{npq}$$

   i. *n* is the number of trials
   ii. *p* is the probability of success
   iii. *q* is the probability of failure

e. When we’ve determined probabilities, it’s good to know how often you’ll get three, four, or any other number of successes.

f. *Combination*—selecting X units from Y possibilities, where different orders are treated as the same possibility. In other words, order doesn’t matter (e.g. a poker hand).
   i. *n* is the number of options.
   ii. *x* is the number of “slots” you can put those options in.

$${}_nC_x = \frac{n!}{(n-x)!\,x!}$$

      1. Note that 0!=1.

   iii. Example: Drawing two aces from the same deck. First, how many different ways are there to draw two cards?

$$\frac{52!}{2!\,(52-2)!} = 1{,}326\ ways$$

Now, how many different ways are there to draw two aces (keeping mind there are four aces)?

$$\frac{4!}{2!\,(4-2)!} = 6\ ways$$

Now we just divide:

$$\frac{6}{1{,}326} = \frac{1}{221}$$

Note if we did this with raw probability, the answer is:

$$\left(\frac{4}{52}\right)\left(\frac{3}{51}\right) = \frac{12}{2652} = \frac{6}{1{,}326} = \frac{1}{221}$$

Raw probability is typically easier when k is small, but if you were wondering the probability of drawing all hearts (k=13), combinations would be easier.

g. The combinations equation is used to determine the probability of exactly *x* successes after *n* trials.

$$\mathrm{P}(x, n) = \frac{n!}{(n-x)!\,x!}(p^x q^{n-x}) = \ {}_nC_x(p^x q^{n-x})$$

i. Where *n* is the number of trials,
ii. *x* is the number of successes,
iii. *p* is the probability of success,
iv. *q* is the probability of failure.