3. Examples:

a. In die toss, event of even and less than 3 are independent.

b. In card draw,
   i. event of face card and club are independent.
   ii. event of one-eyed jack and club are dependent.

c. In the coin toss example,
   i. expect
      • Heads and tails on each trial to be equally likely.
      • heads on the first trial and heads on the second trial to be ind.
   ii. Hence the probability of two heads in a row is the product of the probability of heads on the first toss times the probability of heads on the second toss, or \( \frac{1}{2} \times \frac{1}{2} = \frac{1}{4} \).
   iii. If one believes in runs of luck however, one favorable toss might be more likely to be followed by a more favorable toss; in this case one wouldn’t believe in independence.

d. In cards,
   i. suppose you examine top and bottom cards in deck.
   ii. Consider event that cards are aces.
   iii. If you shuffle between looking at top card and bottom
card, events of seeing ace are independent: \textit{sampling with replacement}.

iv. If you don’t shuffle between looking at top card and bottom card, events of seeing ace are not independent, since if you know the top card is an ace, it’s harder for the bottom card to be an ace: \textit{sampling without replacement}.

e. In a medical example, one might believe that recovery of one person might be ind. of recovery of another person in study. Invalid, for ex, if

i. disease is infectious, and

ii. subjects are assigned beds in the same room

f. In a study on hypertension diastolic and systolic blood pressures are likely related and an improvement in one is likely related to an improvement in the other.

D&P: 6.3

F. How to calculate probabilities?

1. A course in the mathematics behind probability.

2. Simulation:
   a. Recall notion of probability as long run frequency.
Lecture 8

b. Turn this around. Calculate long run frequencies and treat these as approximate probabilities.

D&P: 7.1

G. Usually summarize outcome of experiment numerically with

*random variable*

1. Takes on different values for different people in the study
2. Same taxonomy as before: Quantitative vs. Qualitative, Continuous vs Discrete, etc.
3. Has quality that if experiment were to be performed over and over again the various
   a. outcomes would be observed with fixed frequencies: *probability distribution*.
   b. For continuous variables, is ideal histogram representing many results of many separate runs of experiment, or *density*

i. Recall for sample histogram:
   a. If
      ▶ variable is continuous
      ▶ width of bins is small
      ▶ number of bins is small relative to number of individuals
contained in data set

- histogram is on density scale

- Then

- histogram will not depend much on exact bin size or cutoffs.

- Mean and variance of data will not be much different if we assume that all data are at middle of bins.

- Can suppress divisions between bins

- Can smooth out tops

- got ideal histogram of density.

ii. For population, density is limit of proportion in a small bin, divided by bin width.

4. Might ask a question about a numerical summary;

a. most frequently about the mean

i. Definition:

- Discrete: Multiply possible variable values times their probability and add

- Continuous: Approximate by discrete, and do as above.

- Still is point of balance.

ii. Examples
the mean income of people in the United States, defined by adding all of the people’s incomes and dividing by the population.

- the proportion of registered voters who favor candidate A:
  - Define a variable taking the value 1 if person favors candidate A, and 0 otherwise
  - Take average.

b. less frequently about the variance

- Discrete: Multiply possible variable values minus mean squared times their probability and add
- Continuous: Approximate by discrete, and do as above.

c. Quantity is called a parameter.