two different sample proportions:
    different due to chance?
    different due to reason?

cases (individuals)
variables
    numerical (discrete vs. continuous) vs. categorical
    levels (categories)
explanatory variable
response variable
confounding (lurking) variable
blocking variable

data matrix
frequency table
relative frequency table
contingency table

sample vs. population
sample statistic
    sample mean
    sample median
    sample standard deviation
    sample IQR
population parameter
    population mean
    population median
    population standard deviation
    population IQR
A population parameter is estimated from a sample statistic.

random (or representative) sample -> generalizable results
biased sample -> not generalizable
experiment -> causal relationships
observational study -> associations (not causal relationships)
blinding, double-blinding

scatterplot
boxplot
histogram
relative-frequency histogram
bar chart
mosaic plot

positive association
negative association
independence (no association)

Estimating Q1, median, and Q3 from histogram
skewness
    right skewed -> mean greater than median
    left skewed -> mean less than median
    symmetric -> mean approximately equal to median
modality
    unimodal, bimodal, multimodal
uniform distributions

Standard deviation is approximately the average distance from mean.
Robustness to ouliers:
    Median and IQR are more robust to outliers, should be used for skewed data.

Hypotheses:
    H0: difference due to chance
    HA: difference due to not chance
Hypothesis testing:
    Simulate H0 to indicate how unusual the difference is under H0.
    If very unusual, H0 is rejected.

Probability:
    determine equally likely outcomes
    events are sets of those outcomes
    P(event) = [num outcomes in event] / [total num of outcomes]

marginal (simple) probability
joint probability
conditional probability

Law of large numbers:
    With large sample, the sample statistic approximately equals population parameter.
    Small samples have more variability, the sample statistic may be far from pop parameter.

Two events can be:
    disjoint (mutually exclusive)
    exhaustive
    complements
    independent

general addition rule
special addition rule for disjoint events
general multiplication rule
special multiplication rule for independent events

complements rule

Venn diagrams
tree diagrams
    tree diagrams for applying Bayes' Theorem

probability distributions
random variables
mean of (discrete) probability distribution
standard deviation of (discrete) probability distribution

sampling with replacement
sampling without replacement

probability density function (density or distribution)
estimating probabilities from probability density functions