Course 1 - Week 3/4 - Probability and Distribution

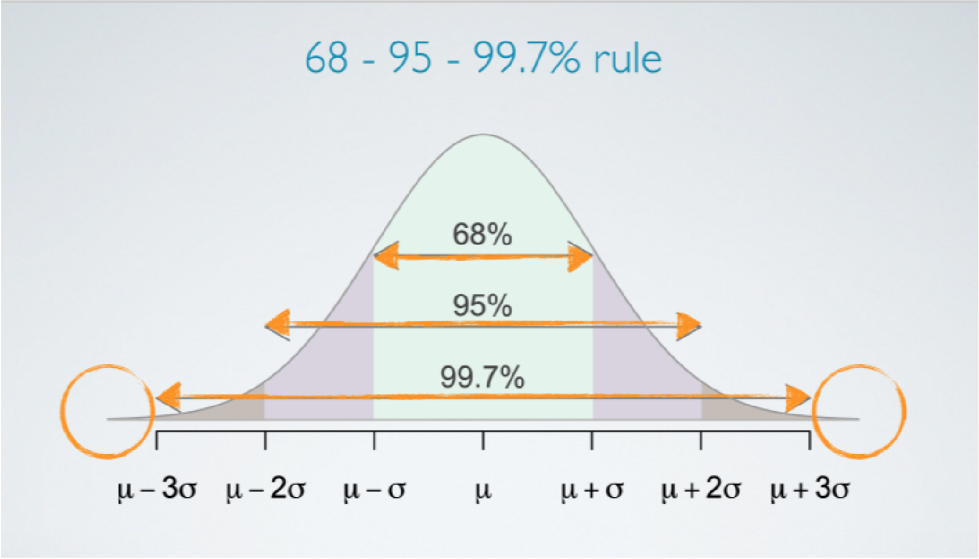
Probability

Some basics such as conditional/marginal/joint probability, product rule, add rule will not be included.

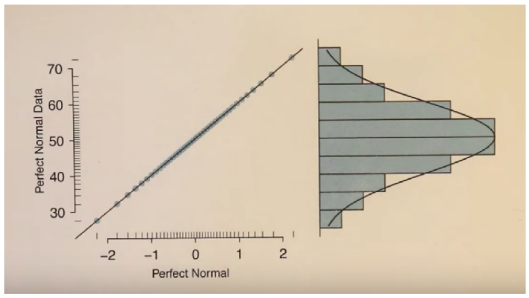
1. What's the difference between disjoint events and complementary events?
    a. Disjoint events means any two of these events cannot happen at the same time
    b. The sum of the probabilities of two complementary events must equal one
    c. The sum of the probabilities of two disjoint events may be necessarily to be one
2. Steps of Bayesian Inference
    a. Procedure
        i. Set up prior probability
        ii. Construct a probability tree listing all possible outcomes in each leaf
        iii. Perform one experiment and see which leaf did the outcome fall into
        iv. Compute posterior probability; P(outcome | experiment) = P(hypothesis | data)
        v. Update prior = posterior
        vi. Repeat from b
    b. There are two probabilities involved here: first is the one you ultimately want to calculate, another is a probability that happens during the experiment is performed
3. Re-do the Bayesian Inference practice problem!

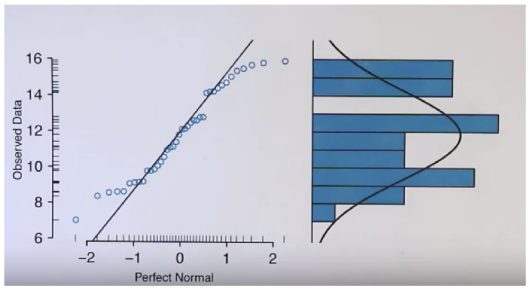
Distributions

4. What is 68 - 95 - 99.7% rule?

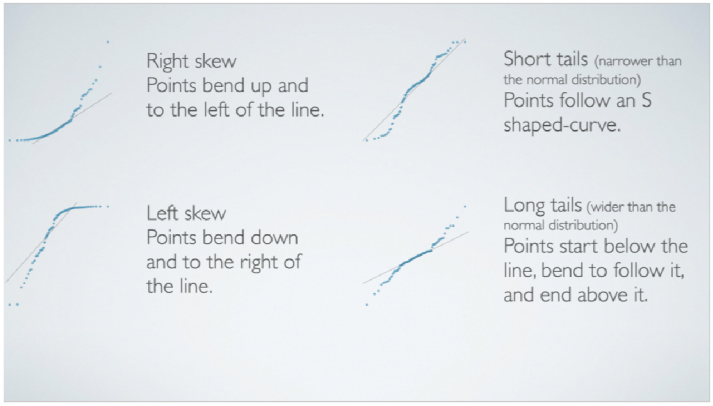


5. What is standardized Z score and why do we use it?
    a. Z score is defined as (data point - **mean**) / standard deviation; simply speaking, how many SD's are there between the observed data and the mean of the distribution.
    b. It is defined for distribution of **any** shape (not just normal distribution)
    c. Unusual observation: |Z| > 2
    d. Motivation: To compare two data points from scales (e.g., compare student's exam performance on GRE and SAT)
6. What is percentile in normal distribution and how do we use Z score to calculate percentile or vice visa?
    a. Define as the percentage of observations that fall below a data point (area of curve before a point)
    b. Use standard normal table to do the mutual conversion
7. How to evaluate whether a distribution is normal distribution?
    a. Normal probability plot: https://www.youtube.com/watch?v=smJBsZ4YQZw
    b. Procedure

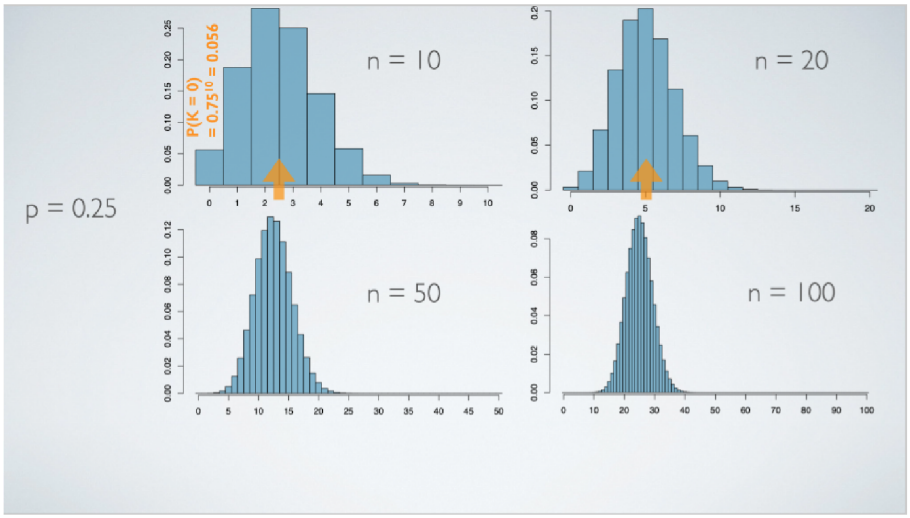




    i. Set up number of samples you want to work on (say 40 points)
    ii. Set up x-axis as the "theoretical" normal distribution quantiles, which is a perfect normal distribution. Generate 40 points on the x-axis that follow a perfect normal distribution.
    iii. Set up y-axis as the real data distribution. Plot the 40 data points (different from the prior 40 points) on the y-axis.
    iv. Map each point on the x-axis to each point on the y-axis **in order**



8. Conditions for binomial distribution
    a. Each trial must be independent from each other
    b. Each trial must have a binary outcome
    c. The probability of success must be the same for every trial
    d. Number of trails n must be fixed
9. What's the expectation and standard deviation of a binomial distribution?
    a. Expectation: μ = np
    b. Standard deviation: σ = sqrt(np(1-p))
10. What's the relationship between binomial distribution and normal distribution?
    a. First plot success probability versus trials barplot (calculate the probability of success)
    b. Increase number of trials and then we find binomial distribution approaches normal distribution
    c. Success-failure Rule: np>=10 and n(1-p)>=10, which means both the expectation of success and expectation of failure must be greater than or equal to 10



Success-failure rule: A binomial distribution with at least 10 expected successes and 10 expected failures closely follows a normal distribution.

$$np \geq 10$$
$$n(1-p) \geq 10$$

Normal approximation to the binomial: If the success-failure condition holds,

$$\text{Binomial}(n,p) \sim \text{Normal}(\mu,\sigma)$$

where $\mu = np$ and $\sigma = \sqrt{np(1-p)}$