Lunch & Learn
Intro Statistics

Heather Schroeder

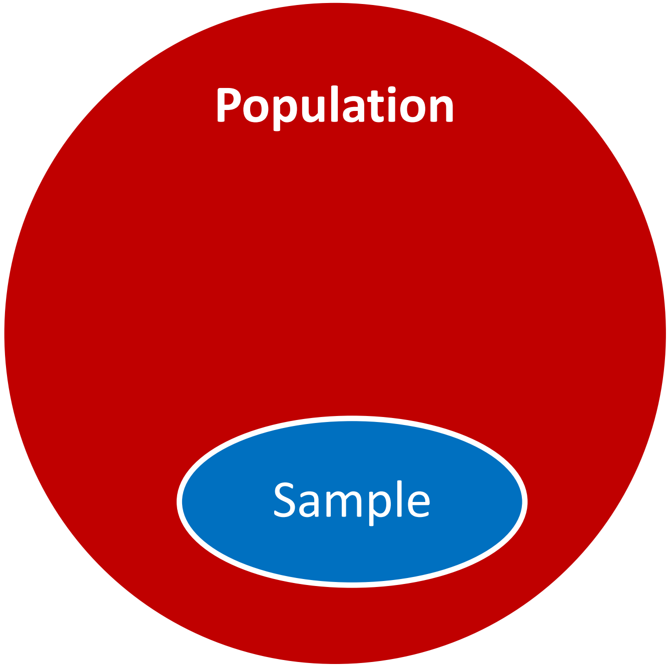
Nov. 14, 2016

Why do we do statistics?

- In one sentence

To answer questions using facts.

- What do we want to know?
  - A population
  - We only have a sample



Overview

- Text:
  - Mind on Statistics 3rd Edition (2007)
- Object is to understand terms
- We won't calculate anything
  - Discuss interpretation and not calculation
- Broad overview of topics
- Use an example throughout presentation

Example Data

- ID, Gender, Height, NPR Listenership
- Categorical, Continuous

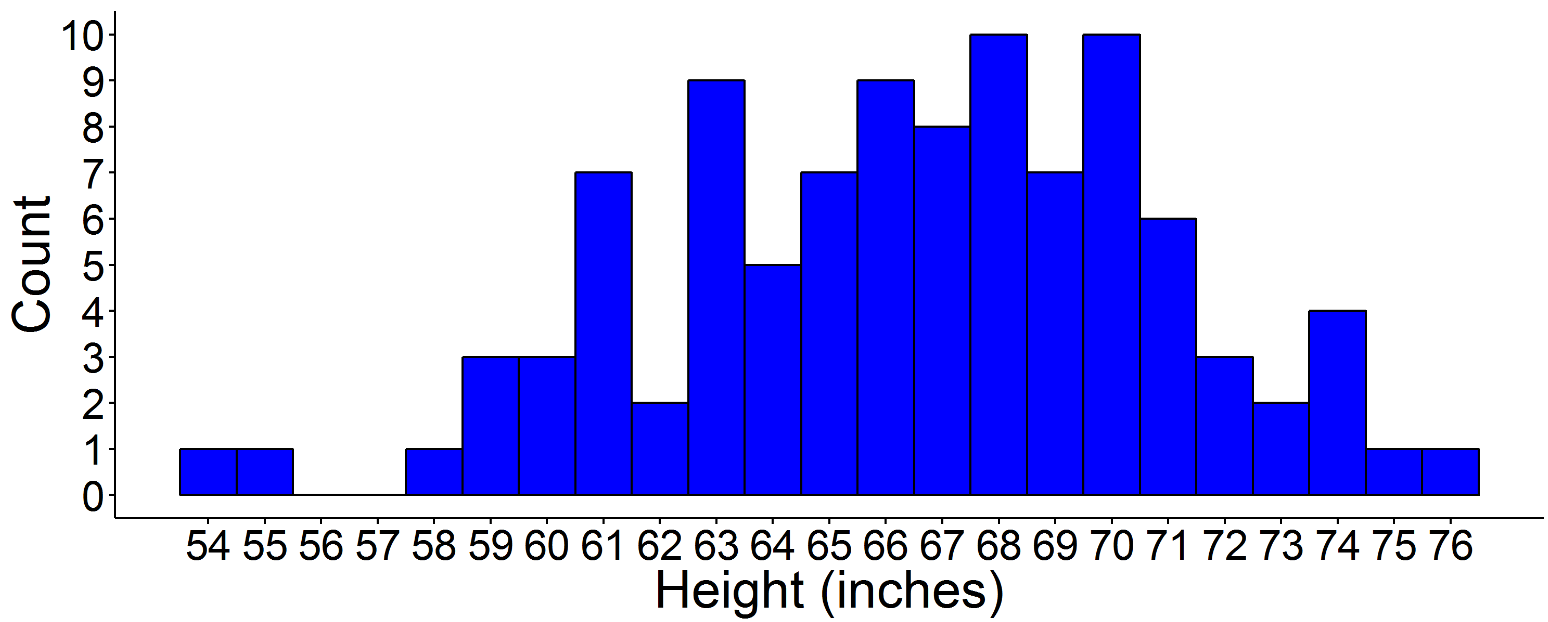
ID	Gender	Height (inches)	NPR?
1	F	73	Yes
2	M	70	Yes
…	…	…	
99	F	74	No
100	M	67	What is NPR?

Describing Continuous Variables

- Mean
  - average
- Median
  - middle number when data is sorted
- Mode
  - Value seen MOst often in the data
- Histogram
  - Visual representation of data values
- Standard Deviation
  - Variability measure

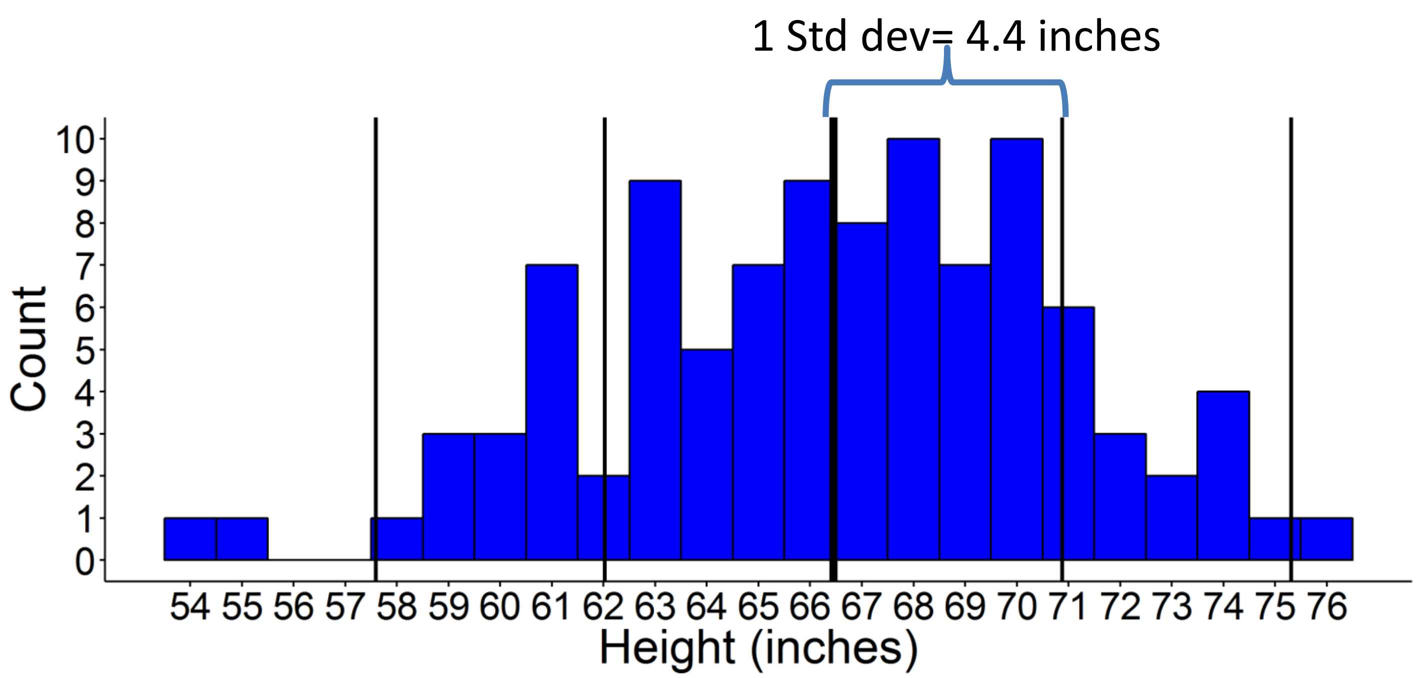
Mean, Median, Mode, Histogram

- Height example:
  - Mean= 66.5 in (5’ 6.5”)
  - Median= 67 in (5’ 7”)
  - Mode= 68 in (5’ 8”)



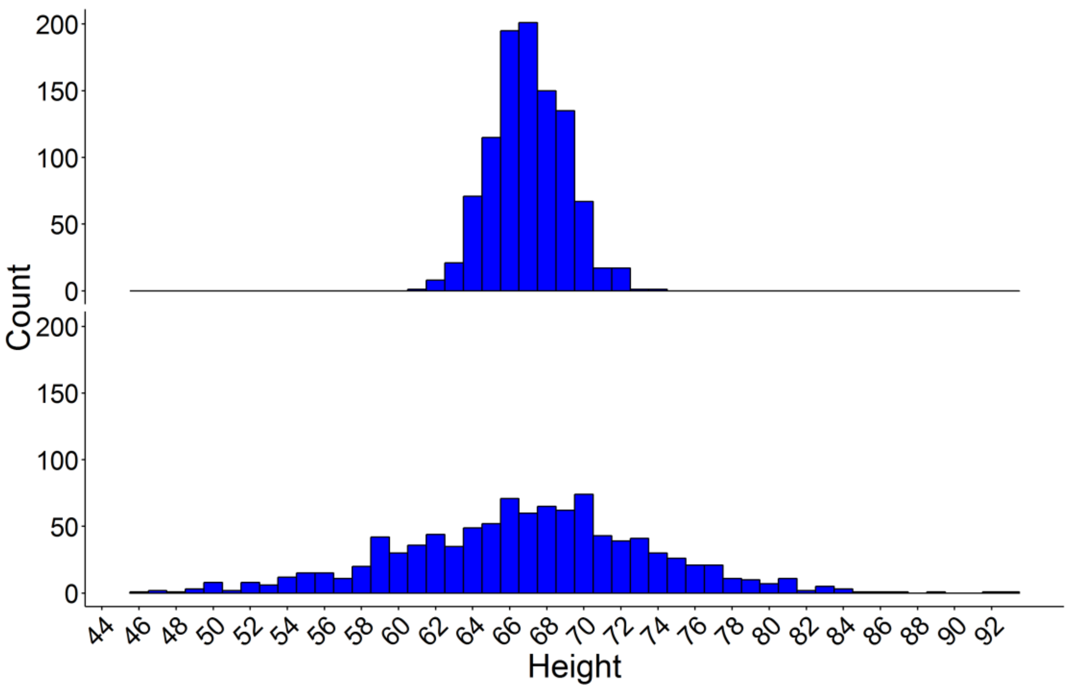
Standard Deviation

- Describes the spread of the data values
- Summarizing how far individual data values are from the mean
- If the data are normally distributed (height)
  - ~68% of the data is within +/- 1 standard deviation from the mean
  - ~95% of data points are captured within +/- 2 standard deviations from the mean
- Height example: Standard deviation= 4.4 inches



Small vs. Large Std. Deviation

- A small standard deviation: data points tend to be close to the mean
- A larger standard deviation: data points are more spread out over a range of values



Standard Error

- How good (accurate) is our estimate?
- Statistics (like mean) estimate trait of a population
- This estimate can have error
  - It can be (slightly) different from the true mean of the population
    - True mean of population = mean if we had measured every single member of the population
- We can estimate this error using a special kind of standard deviation: ***STANDARD ERROR***
  - Standard Error is an estimate of the standard deviation of a ***statistic***

Standard Error

- Roughly the average distance between our estimate and the true population parameter (mean)
- $Standard\ Error = \frac{Std.Deviation}{\sqrt{n}}$
- Takes sample size (n) into account
  - Small n -> larger error
  - Large n -> smaller error
  - Estimates get more accurate as n gets larger
- Different calculation for each kind of estimate (proportion, mean, difference between 2 means, etc)

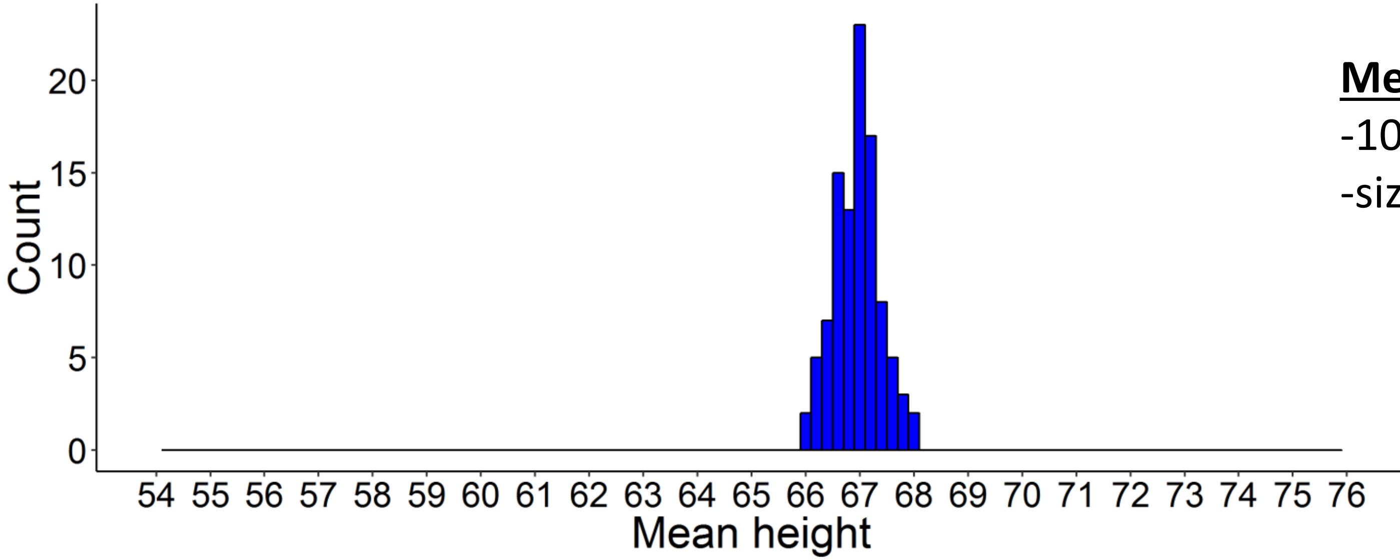
Standard Error- Ex. Mean Height

- Our example: Mean height=66.5 inches, n=100, Std. Dev = 4.4
- Stand error of mean height $\frac{\text{Std.Dev}}{\sqrt{n}} = \frac{4.4}{\sqrt{100}} =$ 0.44 inches

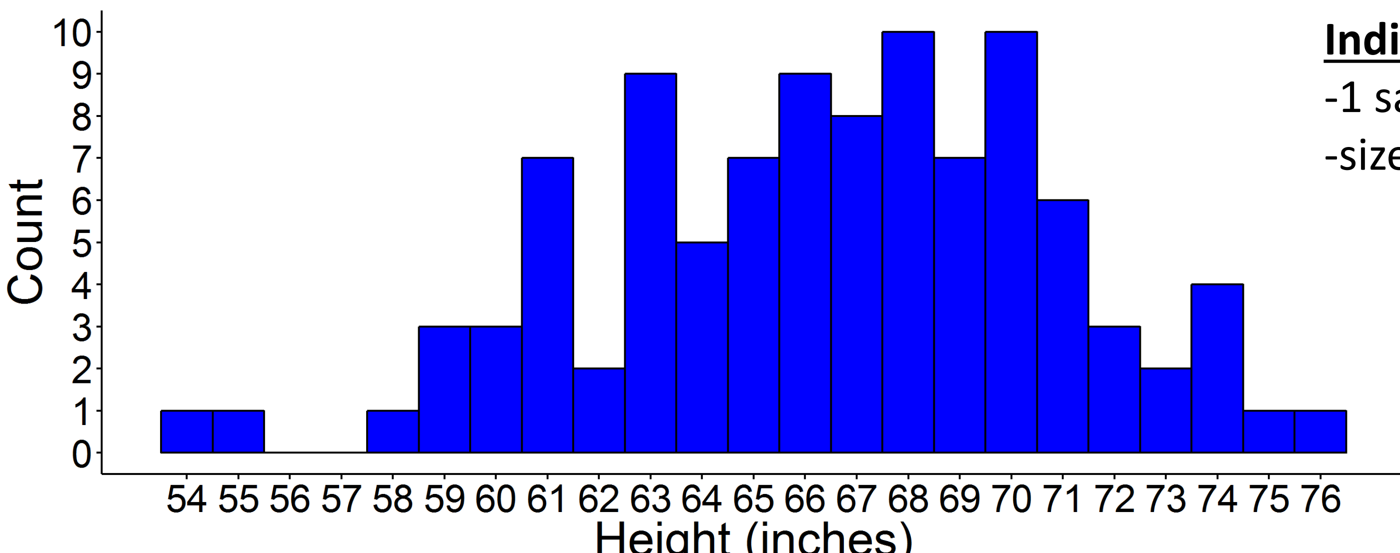
Std. Deviation vs Std. Error

- How to remember difference between standard ***deviation*** and standard ***error***:
  - Standard Deviation:
    - Spread of individual height measurements
    - No one has a "wrong" height -> not ERROR, a DEVIATION
    - Ex. Shaq is 7'1" (85 inches)
      - 4.2 std. deviations away from the mean
  - Standard Error:
    - Spread of mean heights
    - You CAN have error in your estimate

Std. Deviation vs Std. Error



Mean height
-100 samples
-size n=100



Individual heights
-1 sample
-size n=100

One sample t-test

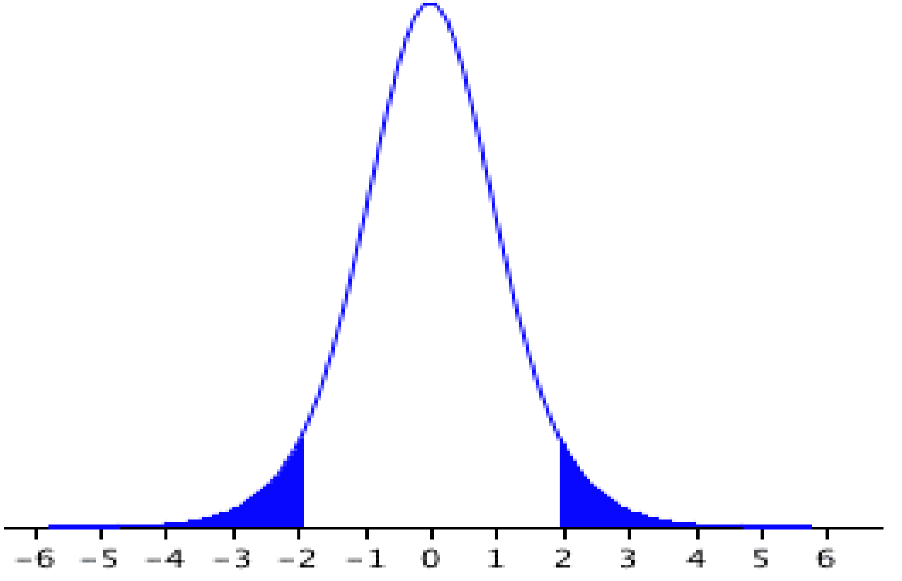
- Is the mean (population) height for females 64 inches?
- We use our data (sample) to try to answer this question about the population
- Steps
  1. Get our estimate of mean female height
  2. Get a value for the t test statistic
  3. Get a p-value
  4. Use p-value to see if the difference has "statistical significance"

One sample t-test

- 1. Estimate mean height from our sample
- 2. Get a value for the t test statistic
  - T-test statistic - a standardized value of the difference between our estimate and the value we are comparing it to (we remove the units)
  - $\left(\frac{Our\ estimate - Comparable}{Std.Error}\right)$
- 3. T-test statistic lets us look up a p-value
- 4. Draw a conclusion

One sample t-test

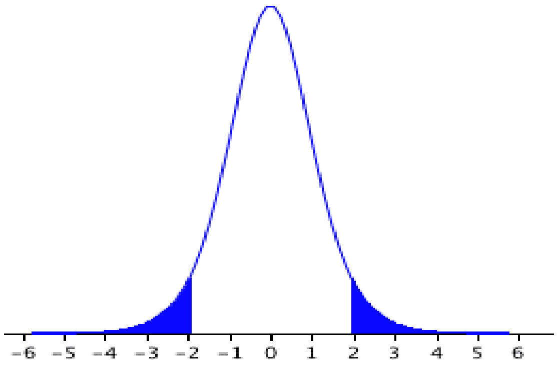
- Test statistics have known distributions
  - t, z, chi square, F, etc.
  - What the distribution looks like depends on:
    - test statistic, sample size, assumptions



One sample t-test

- 1. Estimate mean height from our sample
- 2. Get a value for the t test statistic
  - T-test statistic - a standardized value of the difference between our estimate and the value we are comparing it to (we remove the units)
  - $\left(\frac{Our\ estimate - Comparable}{Std.Error}\right)$
- 3. T-test statistic lets us look up a p-value
- 4. Draw a conclusion

P-value



- Use distribution of test statistic to
  - ***determine how likely it is we saw our result or an even more extreme result by random chance***
  - "How likely" probability (range = 0 to 1)
  - If that probability is small enough, we call it…

Statistically significant!

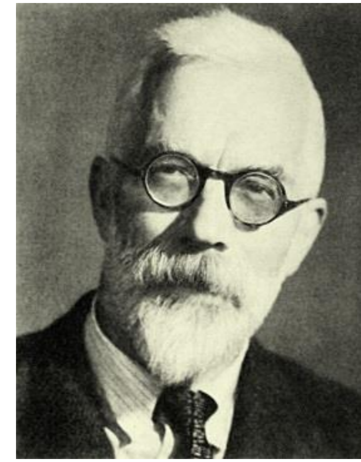
One sample t-test

- 1. Estimate mean height from our sample
- 2. Get a value for the t test statistic
  - T-test statistic - a standardized value of the difference between our estimate and the value we are comparing it to (we remove the units)
  - $(\frac{Our\ estimate - Comparable}{Std.Error})$
- 3. T-test statistic lets us look up a p-value
- 4. Draw a conclusion

Statistical Significance

- If what we are asking is true…
  - If the true mean female height is 64 inches …
- A small p-value means it’s not very likely that our result happened by chance
  - Therefore it must be statistically significant!
- A not so small p-value means our result isn’t statistically significant.
  - Doesn’t mean it’s not practically significant (in real life), just that we don’t have enough evidence to say it is statistically significant

Statistical Significance



- Unwritten standard:
  - 0.05 as a cut off for “small enough”
- Note: Sample size has a big effect on ability to detect statistical significance.
  - A large n makes it easier to have statistical significance
  - A small n makes it hard to have statistical significance

One sample t-test (Ex.)

Is the population mean female height 64 inches?

1. Estimate of mean female height from our sample
   - 65.6 inches (5’ 6”)
   - n=51
   - Std. Deviation= 5.1
   - Std. Error= 0.7206 ($\frac{Std.Dev.}{\sqrt{n}}$=$\frac{5.1}{\sqrt{51}}$)
2. Get a value for the test statistic
   - T-test statistic= 2.23 ($\frac{Our\ estimate\ -64}{Std.Error}$ = $\frac{65.6-64}{0.72}$)
3. Look up a p-value
   - P-value= 0.03
4. Use p-value to see if there is “statistical significance”
   - Since the p-value is smaller than 0.05 it is “statistically significant”
   - We have evidence that the mean height for females is not 64 inches.

95% Confidence Interval (CI)

- Statistic (ex. Mean) is a point estimate
  - Probably not 100% correct
- Therefore also report a range of likely values
  - Calculated from the sample data
  - Likely to include the true population statistic
- 95% CI should contain the “true” mean 95% of the time
  - This method should work 95% of the time
  - If we took 100 different samples and calculated the CI for mean height each time we would capture the true mean 95% of the time
    - 95 of the 100 CI would contain the true mean
    - 5 would not

95% Confidence Interval (CI)

- Formula: Point estimate ± margin of error
  - Margin of error= Multiplier * Standard error
- Width of CI depends on:
  - Spread of the data (standard deviation)
  - n (Higher n = more confidence = narrower CI)
  - Level of confidence
    - Most people use 95% CI

Confidence Interval (Ex.)

- CI for mean female height
  - (64.2, 67.1) inches
    - 65.6 ± 1.4 inches
- Based on our data it's almost certain that the mean height for females is between 64.2 inches and 67.1 inches
- Note: The interval does not include 64
  - This matches the result from the t-test!
    - T-test result: no evidence that the true mean is 64 inches
    - CI doesn't include 64 inches

Differences in Means

- Is mean height the same for males and females?
- Question: about the population
  - Answer: using our data
- We have two independent groups
  - Individuals from one population are not related in anyway to those in the other population
- How many groups do you have?
  - 2 Groups: Use two sample t-test
  - 3+ Groups: Use ANOVA

Two sample t-test

- Used when comparing two independent samples
- Are the two means equal?
  - Mean(Males) – Mean(Females) = 0
- T distribution and t statistic
- $\frac{Sample\ statistic\ - comparison}{standard\ error} = \frac{Mean(M) - Mean\ (F) - 0}{standard\ error}$
  - Standard error pools the variance for each group

Difference in Means (Ex.)

- Is mean height the same for males and females?
  1. Our estimates
     - Mean Height males: 67.5 inches (n=49)
     - Mean Height females: 65.6 inches (n=51)
     - Difference: 1.9 inches
  2. Get a value for the two sample t-test statistic
     - 2.16
  3. Get a p-value
     - P-value= 0.03
  4. Use p-value to see if there is "statistical significance"
     - The p-value is smaller than 0.05 it's "statistically significant"
     - Therefore, we have evidence that the mean heights for males and females are different

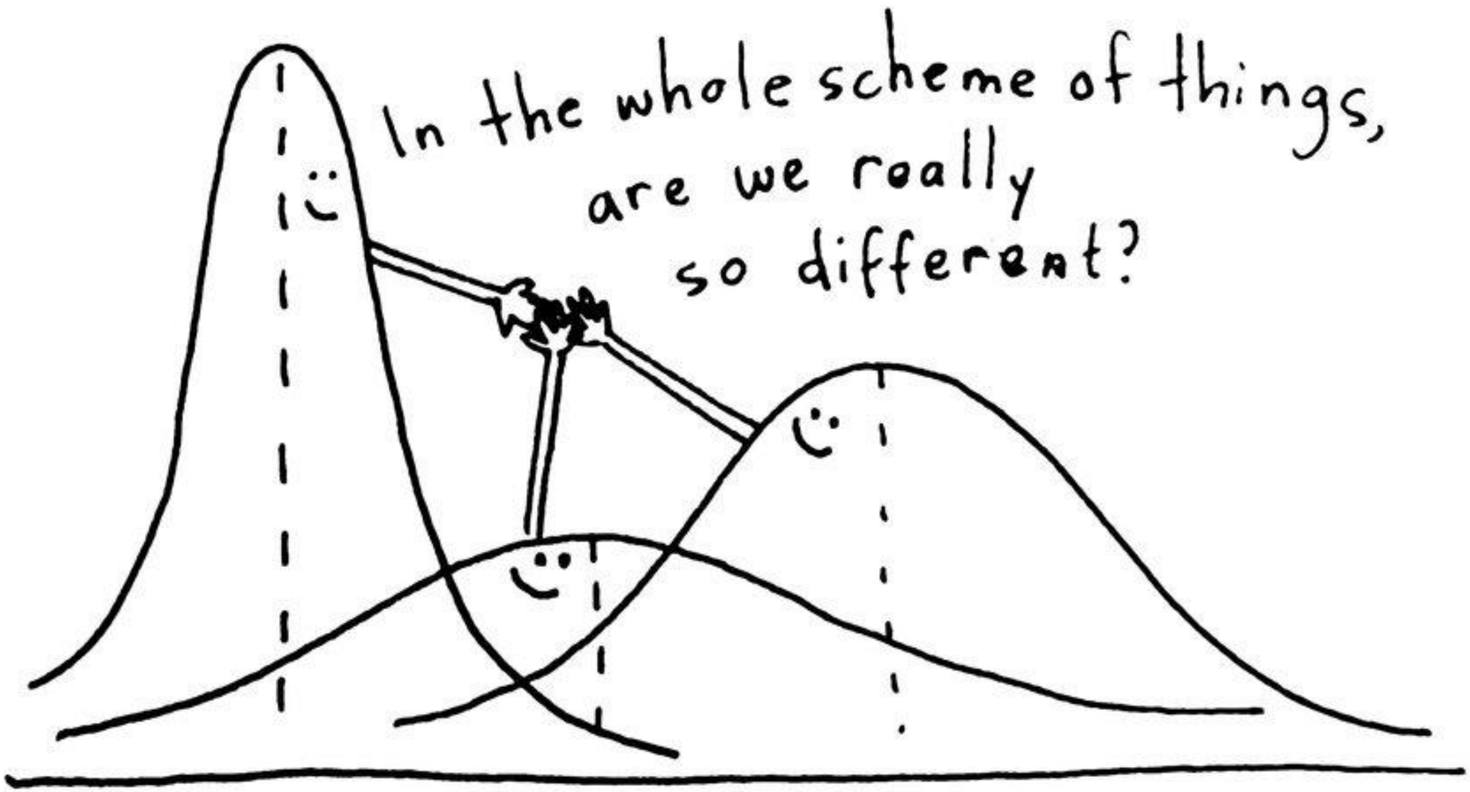
Difference in 3+ Means

- Is mean height the same for NPR listener groups?
  - Listen, Don’t listen, What is NPR?
- Question: about the population
  - Answer: using our data
- We have independent groups
  - Individuals from one population are not related in anyway to those in the other population
- How many groups do you have?
  - 2 Groups: Use two sample t-test
  - 3+ Groups: Use ANOVA
- Our example has 3 groups: we’ll use ANOVA

ANOVA (**AN**alysis **O**f **VA**riance)

- Extension of the two sample t-test
- Can handle 3+ groups
- Test if all the means are the same value
  - Or at least one different
- F distribution and F statistic
- F statistic= $\frac{Variation\ among\ sample\ means}{Natural\ variation\ within\ groups}$
  - Denominator pools variance info from each group

ANOVA (**AN**alysis **O**f **VA**riance)



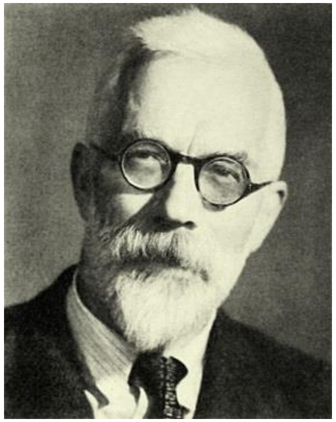
ANOVA (Ex.)

- Is mean height the same for NPR listener groups?
  1. Our estimates
     - Mean height “Yes”: 66.4 inches (n=32)
     - Mean height “No”: 67.3 inches (n=33)
     - Mean height “What is NPR?”: 65.8 inches (n=35)
  2. Get a value for the ANOVA F-test statistic
     - 0.98
  3. Get a p-value
     - P-value= 0.37
  4. Use p-value to see if there is “statistical significance”
     - The p-value is bigger than 0.05. It is not statistically significant
     - Therefore, we don’t have evidence to say there is a difference in mean heights between NPR listener groups

Other kinds of data

- We could have used similar methods with proportions
  - Formulas change, same basic methods
    1. Plan analysis
    2. Calculate estimates from sample
    3. Get a value for appropriate test statistic
    4. Get a p-value
    5. Use p-value to see if there is "statistical significance" and interpret results

“Unwritten Standard”



- Sir R. A. Fischer
  - Biologist and statistician
  - ANOVA and experimental design
- *Statistical Methods for Research Workers* (1925), he states
- **... it is convenient to draw the line at about the level at which we can say: "Either there is something in the treatment, or a coincidence has occurred such as does not occur more than once in twenty trials."... If** one in twenty does not seem high enough odds, we may, if we prefer it, draw the line at one in fifty (the 2 per cent point), or one in a hundred (the 1 per cent point). Personally, the writer prefers to set a low standard of significance at the 5 per cent point, and ignore entirely all results which fail to reach this level. A scientific fact should be regarded as experimentally established only if a properly designed experiment rarely fails to give this level of significance.
- http://www.p-value.info/2013/01/whats-significance-of-005-significance_6.html