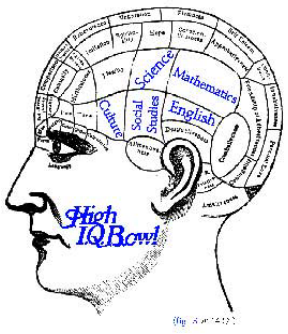


Chapter 6 – The Normal Model

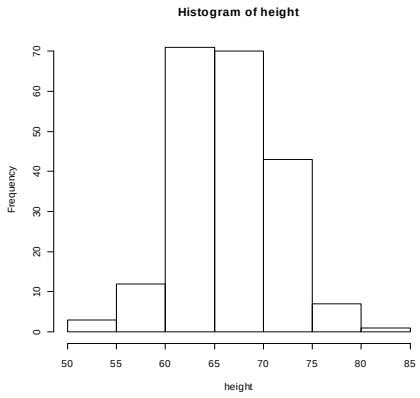
http://medicineworld.org/cancer/lead/5-2006/does-iq-drop-with-age.html

Project 1
Analyze Data

- Due the meeting after Exam 1
- Groups Randomly Assigned (Size 3-4)
  - Collaboration: http://docs.google.com/
- Use our Math 140 Survey Data – All Variables
  - http://www.canyons.edu/faculty/morrowa/140/datasets/
- Grades will be assigned based on insight, accuracy, and analysis in answering the posed question
  - Dig into the data and ascertain which variables are most relevant to your selected variable
  - Start with the original dataset. If you choose to remove any observations, be specific as to which and why.
  - Include evidence to support your claims (graphs, reasoning, etc)
  - Don't forget to introduce the data (the W'S)
  - Your individual grade will be a scaled version of the group's grade based on your group's review of your performance.
- Maximum: 5 Pages
- If you are not strong with writing, be sure to visit the English tutors in the TLC lab…
  - Papers should include intro (and thesis statement), body, and conclusion

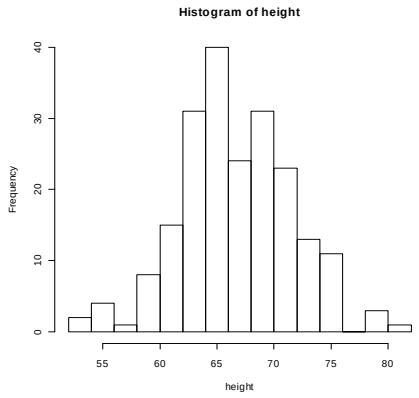
A Look at Math 140 Student Heights

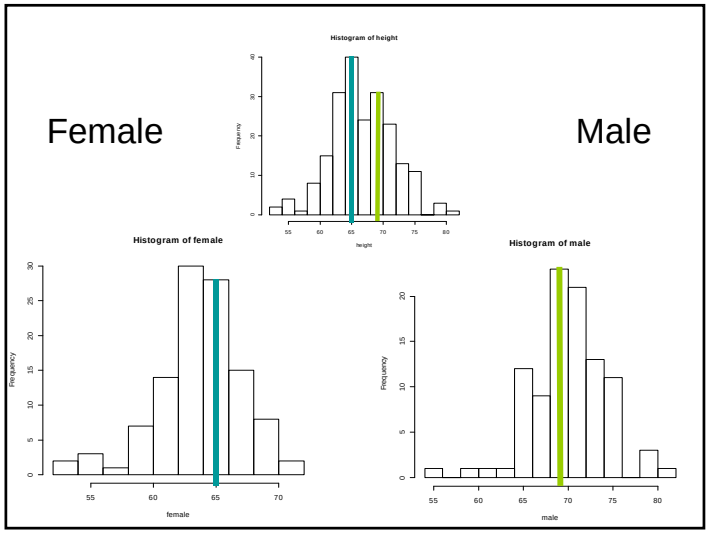
- Note: Graphs generated with R
- Fall 2009 Data
  - Suspicious heights removed
- Describe the Distribution
  - Shape, Center, Spread

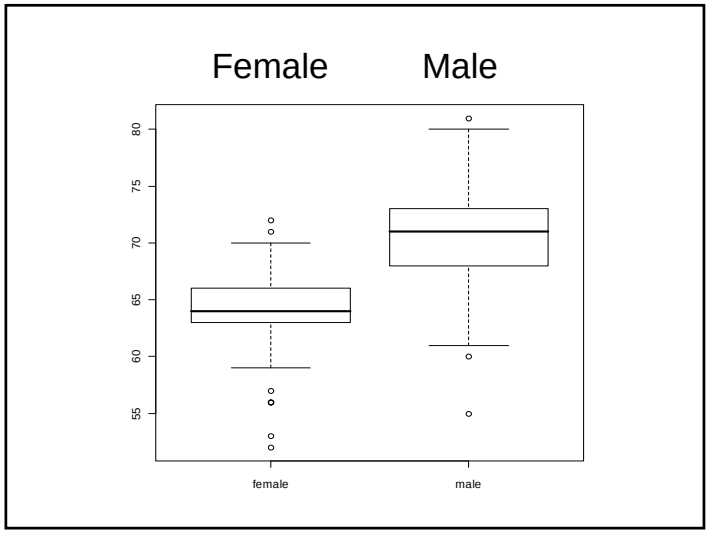


A Closer Look…

- Change the bin size…
- Now Describe the Distribution







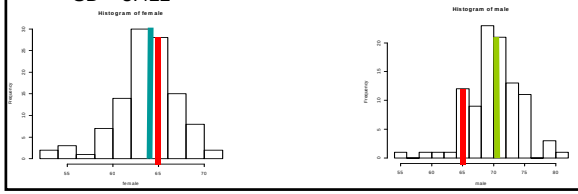
Z-Scores: Compare Apples to Oranges

- Who is 'taller'? A 65" female or a 65" male?
- Standardize the Results
  - Compute the **z-score**…

$$z = \frac{y - \bar{y}}{s}$$

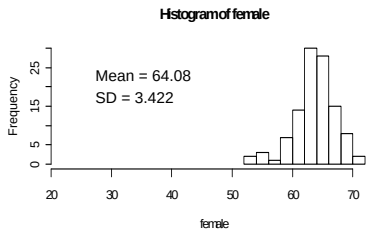
  - Calculates the distance of a value from the mean in standard deviations

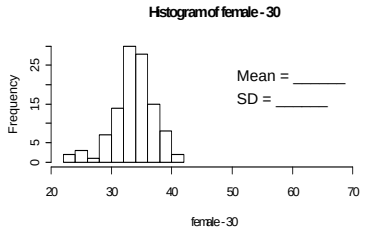
- Females: Mean = 64.08, SD = 3.422
- Males: Mean = 70.49, SD = 4.080



Effect of Shifting Data (Add/Subtract)

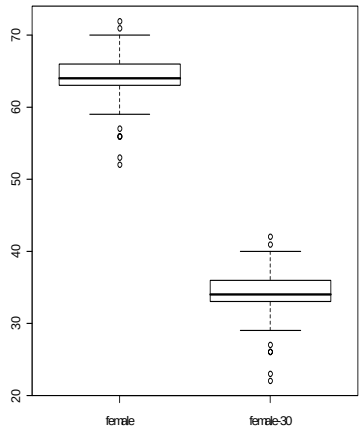
- Subtract 30 from heights of females
- What changes in the histogram?





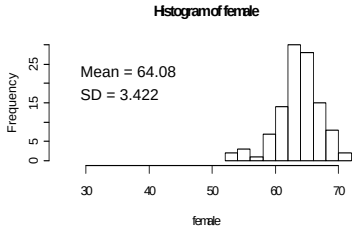
Effect of Shifting Data (Add/Subtract)

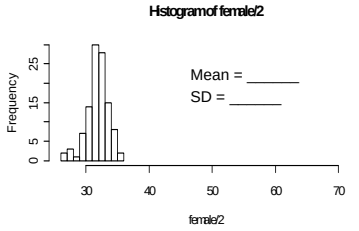
- Subtract 30 from heights of females
- What changes in the boxplot? In the distribution?



Effect of Rescaling Data (Mult/Divide)

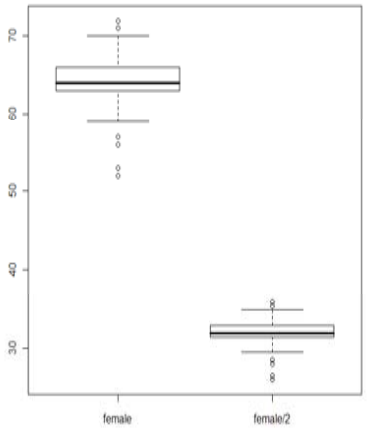
- Examples… Change from inches to feet (or divide out SD)
- Here, female heights divided by 2
- What changes?





Effect of Rescaling Data (Mult/Divide)

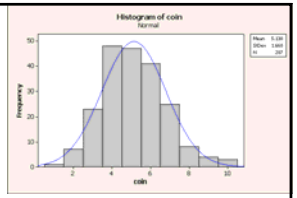
- Examples… Change from inches to feet (or divide out SD)
- Here, female heights divided by 2
- What changes?



Recap: Shifting/Rescaling Data

- If we shift the data (add or subtract the same constant to each data point), then
  - All measures of position change the same way
  - All measures of spread remain unchanged
- If we rescale the data (multiply or divide by the same constant), then
  - All measures of position change the same way
  - All measures of spread change the same way
- So… Standardizing to z-scores
  - Changes the shape by: _____
  - Changes the center by making the mean: _____
  - Changes the spread by making the SD: _____

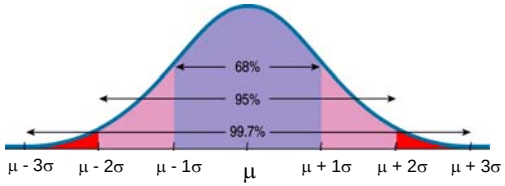
The Normal Model



- George Box:
  - *"All models are wrong, but some are useful."*
- If data are unimodal and roughly symmetric, then the **normal model** is appropriate.
  - Written N(μ, σ)
    - μ is the mean for the model
    - σ is the SD for the model
- Z-scores for appropriate data follow _____

The 68-95-99.7 Rule

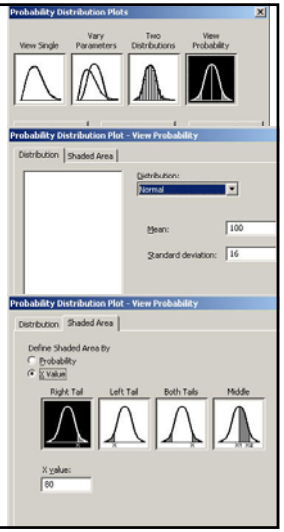
- 68% of values fall within 1 SD of the mean
- 95% of the values fall within 2 SDs of the mean
- 99.7% of the values fall within 3 SDs of the mean



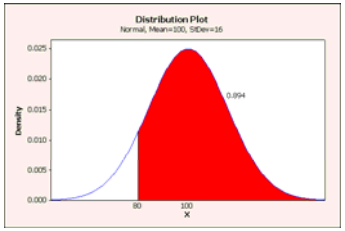
- So… Only need to sketch out to 3 SD
  - Centered at the mean, and the inflection point happens at 1 SD
- Values outside 3 SD are **unusual**.

Finding Normal Percentiles

- Given a particular range of values, what percent of values fall in that range
  - MAKE A PICTURE FIRST!!!!!
- Minitab > Graph > Probability Distribution Plot > View Probability
  - Distribution: Normal, Enter Mean, Enter SD
  - Shaded Area:
    - 'Probability' if you're given the %
    - 'X Value' if you want the %
    - Select the appropriate picture
    - Enter the value
- **Ex**: IQ scores are nearly normal with mean 100 and SD 16. Find the percent of IQ scores over 80.



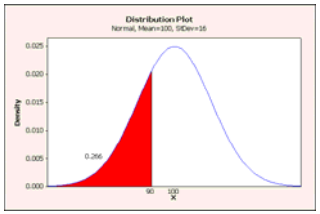
Just Checking: 0.894



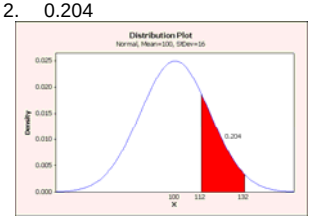
- IQ Scores Again… Mean 100, SD 16
  1. Find the percent of scores below 90
  2. Find the percent of scores between 112 and 132
  3. Changing it up… Now find the 25th percentile (Q1).
     - *This is the IQ score so that 25% of the scores are lower than it.*

Just Checking…

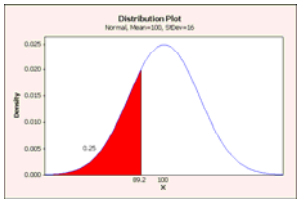
1. 0.266



2. 0.204



3. 89.2

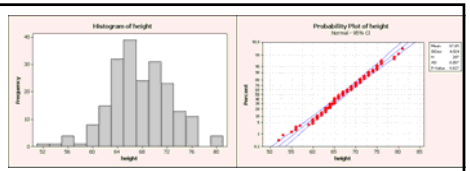


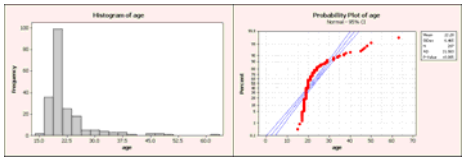
Does the Normal Model Apply?

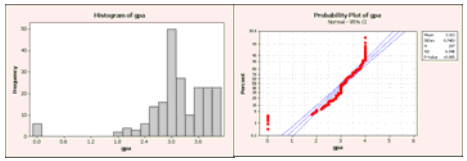
- Check #1: Histogram
  - Make sure unimodal, roughly symmetric

- Check #2: Normal Probability Plot
  - Minitab > Graph > Probability Plot
    - Single
    - Enter the variable
  - The Normal Probability Plot plots the data values against what we would expect under a perfect normal model
  - Look for: Points that form a roughly straight, diagonal line
  - *Note: This graph is slightly different from your book, but same general idea. You can do a lot of work in the options to get it to look identical (not worth it).*
- http://www.canyons.edu/faculty/morrowa/140/datasets
  - Use the Fall 2009 COC Math 140 Survey Results (cleaned heights)
  - Determine if the normal model applies for height, age, and gpa

Just Checking (Fall 2009)







Class Work

- To get credit, it is your responsibility to get checked off.
  1. Chapter 6 Handout
     - Rules for checking answers: No Pens in the Front!!!

Homework

- Textbook/Routine Homework
  - Due Next Time (25% chance of collection)
  1. Read Chapter 6
  2. Pg 147-153 #3, 7, 9, 15, 19, 25, 27, 31, 43, 45
     - Note: Minitab 15 will be helpful/needed for #27, 31(a), 43, 45
- Project/Exploration Homework
  - Variable Selection and Group Selection – Due Next Time!!!
  - Project #1 – Due meeting after Exam 1