Chapter 4 Homework

Due Wednesday, January 28

Problems: 4-1, 4-3, 4-6, 4-8, 4-9,

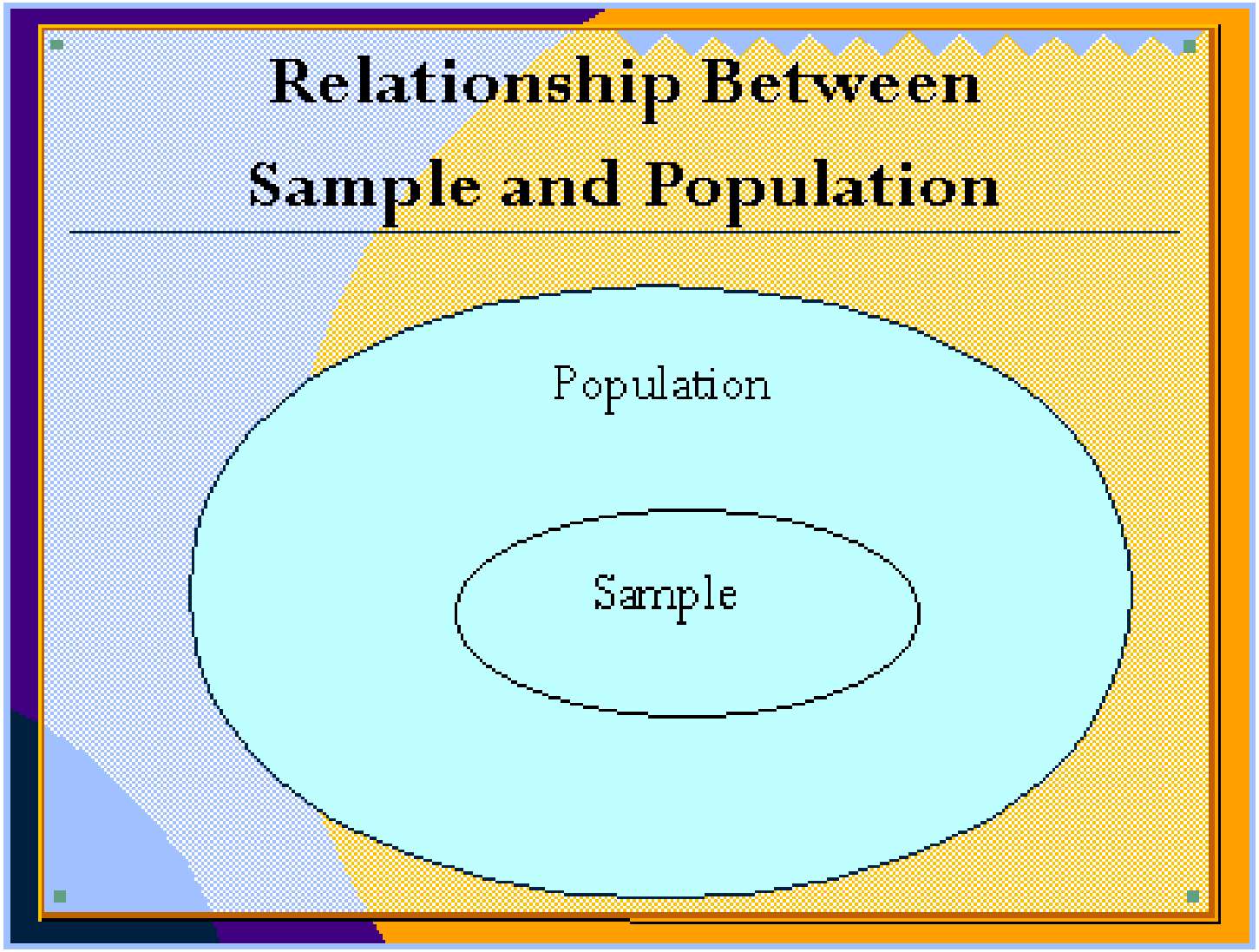
4-11, 4-13, 4-14, 4-20, 4-22

Chapter 4 - Statistics

Statistics - Provides a means to make conclusions from data. It is a tool that must be used with caution. Statistics do not make the conclusions -- the chemists do.

Basic Statistical Concepts

- ***Population:*** the universe or whole under consideration
- ***Sample:*** a subset of the universe or whole which is obtained in order to draw some inference about the population
- ***Indices of location:*** average, mean, median
- ***Indices of dispersion:*** standard deviation, variance



http://www.wabash.edu/depart/psych/Courses/Psych_2-3/Definitions/Sample_%26_Population.htm

Basic terms and equations

- **Average ($\bar{x}$):** for a sample
- **Mean (μ):** for a population

$\mu \approx \bar{x} = \sum_i x_i / n$ where x_i is measured value and n is the number of measurements

- **Median:** middle value when the data are ordered from lowest====>highest

Example - 34, 38, 41, 55, 58, 61, 64

median

- **Standard Deviation** (σ or s): how closely the data are clustered about the mean.

Computing the st. dev. (σ) - population

Standard deviation:

- Review of the standard deviation computation:
  - definitional formula: **root-mean-square of the deviation scores**:
  - example:
    - **step 1**: compute the deviation scores:

$$X - \overline{X}$$

    - **step 2**: obtain the sum of squares:

$$\Sigma(X - \overline{X})^2 = SS$$

    - **step 3**: obtain the mean squared deviation, the variance:

$$\frac{\Sigma(X - \overline{X})^2}{n} = \frac{SS}{n}$$

    - **step 4**: take the square root of the variance:

$$\sqrt{\frac{\Sigma(X - \overline{X})^2}{n}} = \sqrt{\frac{SS}{n}} = \sqrt{variance} = s$$

http://www.nd.edu/~aventer/statistics/chapter4/ppframe.htm

Computing the sample standard deviation (S)

- This is done in the same way as for σ, only (N-1) is used in the denominator!
- Be sure to check what your software and calculator are doing!
- Most of the time, we will want to use the sample standard deviation (S, N-1 formula) when describing dispersion of our experimental measurements

Example: Computing average and standard deviation

Student	Test Score	(Test Score – ave)	$(dif)^2$
Jordan	90	90 – 88.4 = 1.6	2.56
Joshua	89	89 – 88.4 = 0.6	0.36
Jalisa	68	68 – 88.4 = -20.4	416.16
James	95	95 – 88.4 = 6.6	43.56
Jani	100	100 – 88.4 = 11.6	134.56

average = (90+89+68+95+100)/5 = 88.4

st. dev. = $((2.56+0.36+416.16+43.56+134.56)/(5-1))^{1/2}$

= 12.2

What does the average tell you?

The average gives the "mid-range" value. It may not be useful if your measured results are skewed high or low (such as lots of high readings with a couple of very low readings).

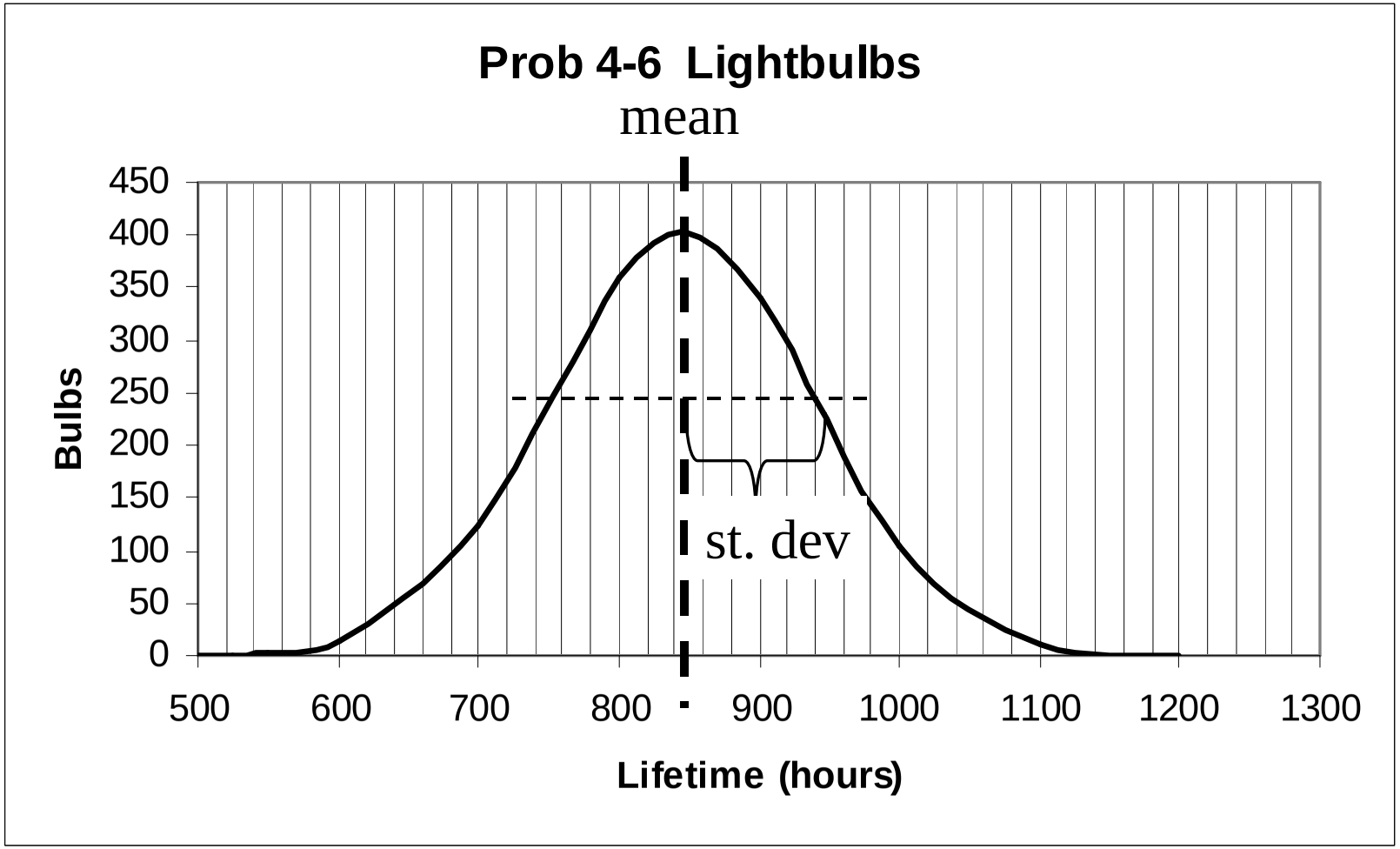
From example: Drop Jalisa's score of 68 and….
average = 93.5 (compared to 88.4)

What does the standard deviation tell you?

The standard deviation tells you the spread in your measurements. The larger the value - the larger the spread. The Gaussian curve gets tall and skinny with small standard deviations – the curve gets short and fat with large standard deviations.

Gaussian (normal) distribution:

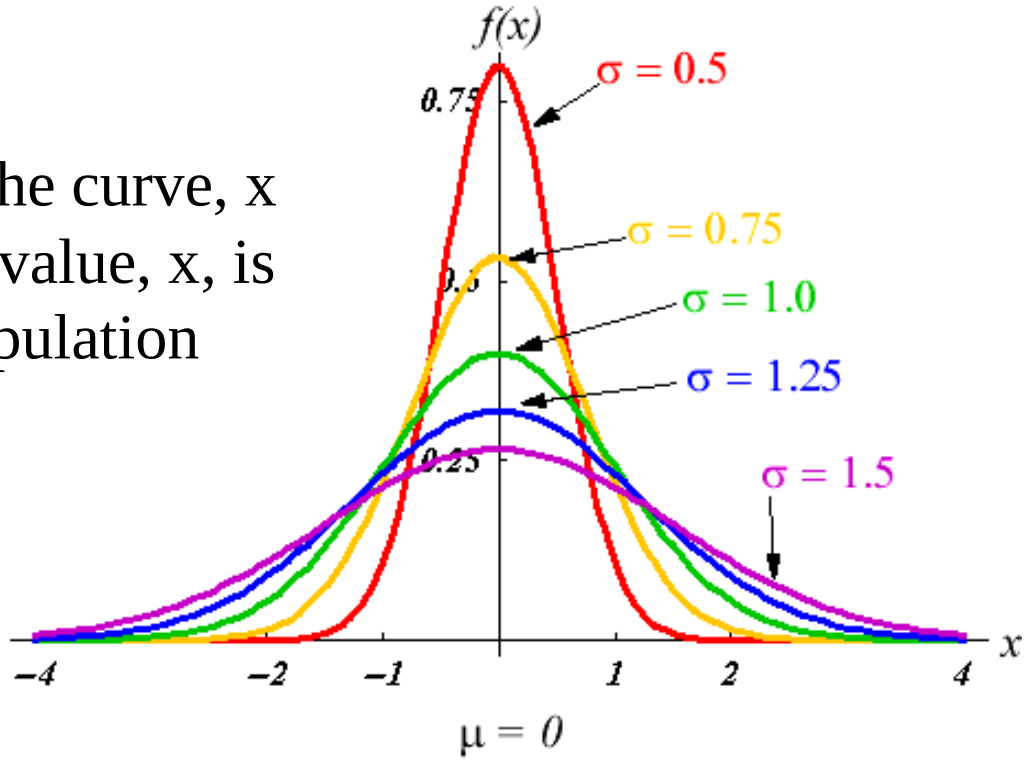
Curve that describes what a measurement might be within a large population of measurements.



Gaussian distribution

$$f(x) = \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad -\infty < x < \infty$$

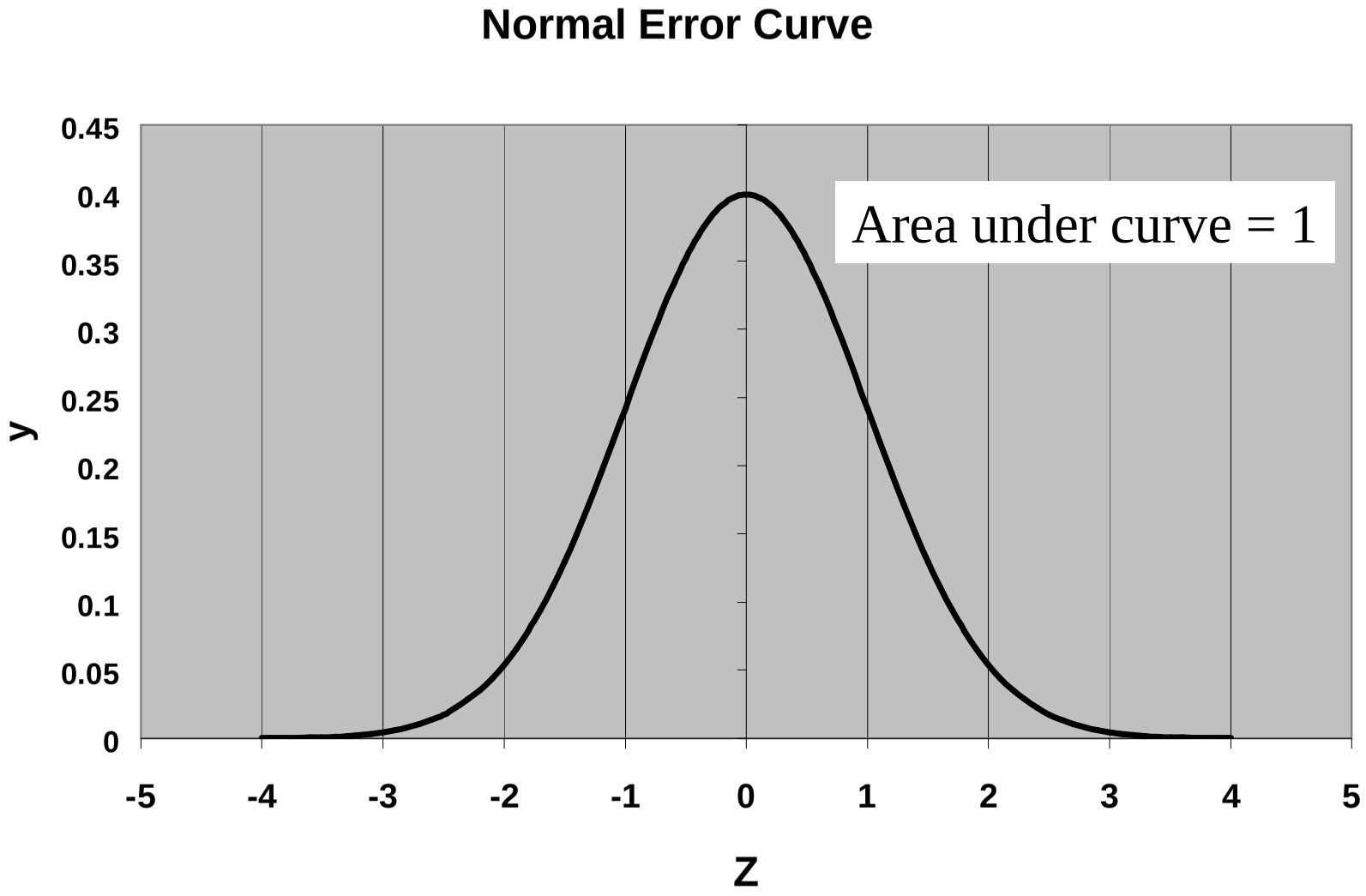
NOTE:
At the max in the curve, x = μ (measured value, x, is equal to the population mean, μ).



http://www.efunda.com/math/distributions/NormalDistPlot.cfm

Z-scoring (autoscaling, standardizing)

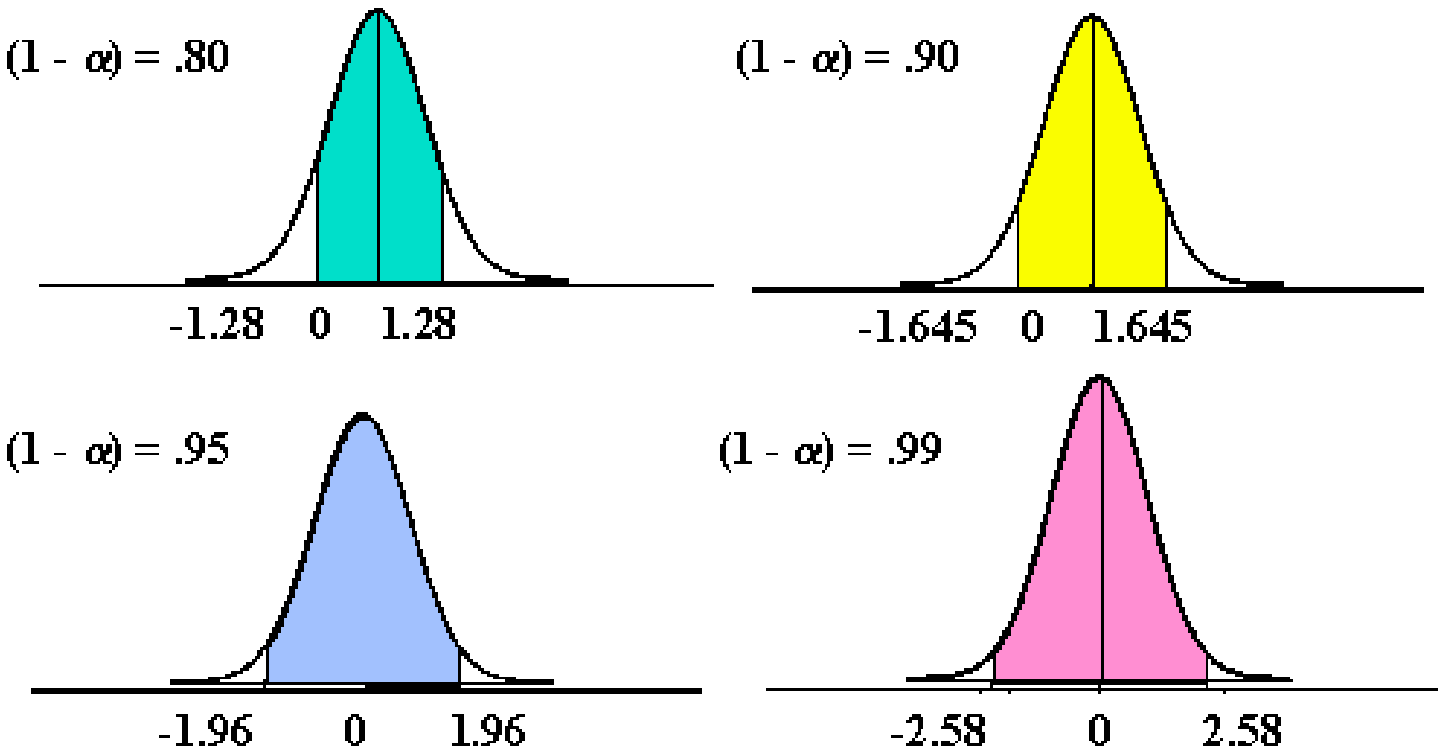
$$Z = (x - x_{avg})/s$$



Z-scoring (autoscaling, standardizing)

$$Z = (x - x_{avg})/s$$

Graphs of the Various Z-values



In class statistical experiment: M&M Statistics

Step 1: Question – What is the color distribution of M&M's in a package?

Step 2: Hypothesis – Educated guess

Step 3: Experiment – group M&M's by color & count

Total number of M&M's =

Total number of M&M's in each color group =

Step 4: Results – statistical analysis

Step 5: Conclusions – how do results compare to hypothesis?