STAT 217 – Handout 4
Normal Distributions

The simulation analyses that we have performed all resulted in a common and familiar shape:

- Many real-world phenomena follow a bell-shaped curve called a normal distribution.
  - Many sample statistics also follow a normal distribution when repeated samples are taken.
- A normal distribution is characterized by two values: its mean (μ) and standard deviation (σ).
  - The mean μ indicates where the center and peak of the distribution are.
  - The standard deviation σ indicates how spread out the distribution is.
    - σ is the distance from the mean to where the curvature of the “bell” changes.

**Example 4-1: Normal Practice**
For each of the three normal curves graphed below, take a guess for its mean and standard deviation:

- The total area under every normal curve is one (or 100%). The area under the curve over a certain interval of values indicates:
  - The proportion of values in that interval
  - The probability that a randomly selected observational unit will be in that interval
To find the area under the curve over a certain interval, we:
  o **Standardize** by subtracting the mean and dividing by the standard deviation
    o This **z-score** indicates how many SDs above or below the mean the value is.
  o Use a table of standard (mean 0, SD 1) normal probabilities, which gives the area
    under the curve to the left of that value (see link on online table from our course
    webpage)
    o Or you can use the Java applet called “normal probability calculations” (see
      link under “Data and Applets” on our course webpage)
    o Or you can use Excel (Insert function> NORMDIST)
  o **The empirical rule** states that with any normal distribution:
    o 68% of the data falls within 1 SD of the mean
    o 95% of the data falls within 2 SDs of the mean
    o 99.7% of the data falls within 3 SDs of the mean

**Example 4-2: Comparing SAT and ACT**
Suppose that the distribution of scores on the SAT exam is symmetric and mound-shaped with
mean 1500 and std dev 250, while the distribution of scores on the ACT exam is symmetric and
mound-shaped with mean 21 and std dev 6.

a) About 95% of SAT scores fall between what two values?

b) About 95% of ACT scores fall between what two values?

c) Suppose that Bobby scores 1750 on the SAT. About what percentage of SAT takers scored
   higher? [**Hint:** Draw a sketch and use the empirical rule.]

d) Suppose that Kathy scores 30 on the ACT. Who has done better compared to their peers –
   Bobby or Kathy? Explain.

e) Suppose that Peter scores 1000 on the SAT. About what percentage of SAT takers scored
   higher?

f) Suppose that Kelly scores 12 on the ACT. Who has done better compared to their peers – Peter
   or Kelly? Explain.
g) Calculate the $z$-scores for Bobby, Kathy, Peter, and Kelly. Who has the highest? Who has the lowest?

Bobby: 
Kathy: 
Peter: 
Kelly: 

**Example 4-3: Birthweights**
Birthweights of babies born in the United States can be modeled by a normal distribution with mean 3250 grams and standard deviation 550 grams.

a) Draw a sketch of this distribution.

b) Babies who weigh less than 2500 grams are classified as “low birth weight.” Shade the area corresponding to the probability that a randomly selected baby is low birth weight. Then make an educated guess for the value of this probability.

c) Calculate the $z$-score corresponding to 2500 grams. Also interpret what this $z$-score reveals.

d) Use this $z$-score and the table to determine the probability that a randomly selected baby is of low birth weight. Is the probability close to what you guessed earlier?

e) Use the Normal Probability Calculator applet and/or Minitab and/or Excel to confirm this probability calculation.
f) Determine the proportion of babies who weigh more than 10 pounds (4536 grams) at birth.

g) Describe two different ways that you could have used the table to answer the previous question.

h) Determine the probability that a randomly selected baby weighs between 3000 and 4000 grams at birth. [*Hint: Decide what to do with the tabled values for the two relevant z-scores.*]

i) How little would a baby have to weigh to be among the lightest 2.5% of all newborns? Draw a sketch to support your answer, and explain how you must read the table differently here.

j) How much would a baby have to weigh to be among the heaviest 10% of all newborns?