CPS 5310 Spring 2014
Shirley Moore, Instructor
January 28 Class

Descriptive Statistics

First we'll use the descriptive statistics techniques we learned in class today to do some descriptive analysis of climate data.

1. Go to the World Resources Institute (WRI) Climate Analysis Indicators Tool (CAIT) website at http://cait2.wri.org/ . Download comma separated value (CSV) files for greenhouse gas (GHG) emissions and socio-economic data for the US fifty states and District of Columbia. Combine these statistics into a single Excel or Calc file. Use Calc/Excel to carry out the tasks below.
   a. Create a new column with the per capita GHG per state (multiply by 10^6 to express the per capita figures in metric tons).
   b. Below the column you created for a, compute the following:
      i) mean per capita GHG
      ii) min, max, and range of the per capita GHG
      iii) standard deviation of the per capita GHG
   c. Compute and interpret Pearson correlation coefficients for the following:
      i) population vs. total GHG emissions
      ii) GDP vs. total GHG emissions
      iii) total energy usage vs. total GHG emissions

2. Go to the Berkeley Earth website at http://berkeleyearth.lbl.gov/ and retrieve the raw data for land temperature for El Paso, Texas. Copy the part of the file for which there are sufficient values to compute the moving averages – e.g., from 1838 to 2003. Compute the 5- and 10-year moving averages in new columns and see if your computed values agree with those already computed in the table.

3. Read the webpage on the NASA GISS Surface Temperature Analysis at http://data.giss.nasa.gov/gistemp/ . What are the various types of biases that can occur in the raw data? What techniques are used to adjust the data so as to correct for these biases?

4. Read the article Global Temperature Update Through 2013 by Hansen, Sato, and Ruedy available from the course website. How are phenomenological and mechanistic modeling being used together in the analysis?

Next we'll work some problems using R and what we learned about probability density and probability distribution functions.

5. For what constant *k* is $f(x) = k\,e^{-x}$ a probability density function on [0, 1]?

6. If $f(x) = k$ is a uniform density function on [a, b], what is *k*? Sketch the graph of *f(x).*

7. Pressure gauges manufactured by Precision Corp. must be checked for accuracy before being placed on the market. To test a pressure gauge, a worker uses it to measure the pressure of a sample of compressed air known to be at a pressure of exactly 50 pounds per square inch. If the gauge reading is off by more than 1% (0.5 pounds), the gauge is rejected. Assuming that the reading of a pressure gauge under these circumstances is a normal random variable with mean 50 and standard deviation 0.5, find the percentage of gauges rejected.

8. Plutonium 239 decays continuously at a rate of 0.00284% per year. If X is the time a randomly chosen plutonium atom will decay, write down the associated probability density function and use it to compute the probability that a plutonium atom will decay between 100 and 500 years from now.