# ST518 - Homework #1

1. Consider the 928 bivariate measurements of height ($y$) and midparent height ($x$) available on moodle as either "`heights.txt`" or "`heightfreqs.txt`."

   (a) Consider the population from which this simple random sample of adult males was drawn. Let $\mu$ denote the mean of the population. Use statistical software to compute $\overline{y}$ as an estimate of $\mu$. Obtain a 95% confidence interval for $\mu$. If another person is sampled at random, would you expect this interval to capture this person's height with high confidence?

   (b) Report the following summary statistics:

      i. $\bar{y} = (1/n)\sum y_i = (1/n)(y_1 + y_2 + \cdots + y_n)$
      ii. $\bar{x} = (1/n)\sum x_i = (1/n)(x_1 + x_2 + \cdots + x_n)$
      iii. $S_{xy} = \sum(x_i - \bar{x})(y_i - \bar{y})$
      iv. $s_{xy} = (1/(n-1))\sum(x_i - \bar{x})(y_i - \bar{y})$
      v. $(1/(n-1))\sum(x_i - \bar{x})y_i$
      vi. $(1/(n-1))\sum x_i(y_i - \bar{y})$
      vii. $S_{xx} = \sum(x_i - \bar{x})^2$
      viii. $s_x^2 = \frac{1}{n-1}\sum(x_i - \bar{x})^2$
      ix. $S_{yy} = \sum(y_i - \bar{y})^2$
      x. $s_y^2 = \frac{1}{n-1}\sum(y_i - \bar{y})^2$
      xi. $r$

   (c) Express the slope from the least squares regression line as a function of $r, s_x$ and $s_y$ above.

2. Obtain an approximate 95% confidence interval for the population correlation coefficient $\rho$ when a bivariate random sample of size $n = 20$ results in a sample correlation coefficient of $r_{xy} = -0.45$. Also, conduct a test of $H_0 : \rho = 0$.

3. Suppose that two random variables $X$ and $Y$ have correlation $\rho = 0.6$. (That is, the correlation among two quantities in an entire population is $E[(X - \mu_x)(Y - \mu_y)/(\sigma_X \sigma_Y)] = 0.6$.) What is the probability that a random sample of $n = 30$ bivariate observations will yield a sample correlation coefficient that exceeds 0.7. Find $\Pr(R > 0.7; \rho = 0.6)$.