Homework # 1, Stat 526

1. The data "star.txt" contains the light and the temperature of a few stars in the universe. Fit a simple linear regression model by taking "light" as the response variable and "temp" as the independent variable. Report the estimates of parameters (i.e., the intercept and slope), and their corresponding t-values and p-values and the covariance matrix.

2. Let
$$A = \begin{pmatrix} 5 & 4 & 0 \\ -2 & 3 & 7 \\ 0 & -1 & 1 \end{pmatrix}$$
and
$$b = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}.$$
Compute A^{-1}, Ab, $A^t b$ and $A^{-1}b$.

3. The t-confidence interval for the median (which is also the mean) based on an iid normal sample $x_1, \cdots, x_n$ is given by
$$[\bar{x} - t_{\alpha/2,n-1}\frac{s}{\sqrt{n}}, \bar{x} + t_{\alpha/2,n-1}\frac{s}{\sqrt{n}}],$$
where $t_{\alpha/2}$ is the upper $\alpha/2$ quantile of the t distribution with $n-1$ degrees of freedom. Use R to do the following simulation for the case when $x_1, \cdots, x_n$ are not iid normal but iid t distributed with degrees of freedom r, with $r = 1, 2, 5, 20, 50$ respectively: first generate $n = 10$ iid random samples from t_r distribution and then calculate the confidence interval for the median based on the formula above. Repeat the procedure $10,000$ times and calculate the percentage of the 95% confidence intervals that contain the median value 0.