ST505/607R: Fall 2012. Homework 3. Due, Friday: Oct. 5 at start of class.

First, I suggest you try problems 2.10 and 4.5 which review concepts we covered a little earlier, but these are not to hand in. I may ask about them in class.

1. ("Recent reports have emphasized that increased neopterin levels are found in almost 100 percent of AIDS and ARC patients and that neopterin excretion is closely correlated with risk factors ..." from Hutterer et. al, in *Wiener klinishce Wochenschrift*; well this was recent when it first came out). The analysis below is from a calibration of an assay to measure neopterin. A set of 28 standard were run with true concentrations between 5 and 320. where $Y =$ is a binding count. A plot of Y versus $X = log(conc)$ (where conc = true concentration) showed that a simple linear regression model for Y on X seemed reasonable. (the X value is denoted lconc below). A person comes in for analysis and the machine returns a count of 3000.

   First, estimate that person's X value (which is the log of their neopterin concentration) by giving a point estimate and finding a 95% confidence interval. Use the approximate interval which uses the standard error; see Section 4.6. Also, show how numbers would enter into the calculation of the Fieller interval (see the notes) but you don't need to calculate out. Then convert your confidence interval (which is for log-concentration) to confidence intervals for concentration by exponentiating the end points.

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	61945450	61945450	902.08	<.0001
Error	26	1785416	68670		

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	6683.97035	140.77621	47.48	<.0001
lconc	1	-1072.92715	35.72305	-30.03	<.0001

Covariance of Estimates

Variable	Intercept	lconc
Intercept	19817.939942	-4707.512482
lconc	-4707.512482	1276.1361656

2. On the web (under data) find the cow carcass pH data and

   (a) Run a linear regression analysis of pH (response) on time (predictor) getting estimated coefficients, the MSE and the variance-covariance matrix of the estimated coefficients (which contains $s^2\{b_0\}$, etc.) Also, get a plot with data and the individual confidence intervals plotted versus time. [This part should just be a review. You should be able to use your old code with just a change in data an variable names. Note if running SAS you should delete the first row or in the infile statement say firstobs=2]

(b) Estimate the time at which the expected pH is 5.5. Give a point estimate and find a 95% confidence interval two ways (one an approximate interval, the other based on Fieller's method from class). Indicate on your plot from a) how the Fieller interval arises graphically.

(c) **Additional problem for ST697R students:** The real problem is that they want the pH less than 6 to begin processing. If we took a time so that the mean was exactly 6, then (assuming a symmetric error term) at that time half of the cows would still have pH greater than 6 (and so couldn't be processed). That is why I chose (arbitrarily) to shoot for a mean of 5.5 in the question I asked you. Discuss how you would approach the problem if the investigators said that they wanted to estimate the time at which the probability of a pH greater than 6 is equal to .05, assuming $Y|X$ is normally distributed. Discuss the general problem then find an estimate and confidence interval assuming that you knew that $\sigma = .01$. Hint: with σ known you can reformulate it to a problem you know how to solve.

3. 2.39 When they say to state the characteristics, what they mean is to give the name of the distribution and the mean and variance of the distribution.

4. Problems 2.31 (a, c and d; skip b), 2.48 and 2.49. These are all related. Note that we will discuss how to do the correlation analyses in class (the SAS analysis is in the notes; I'll go over the corresponding R method).