Homework 4 Due April 7, Friday

Prob-1 KNNL 3.7. Do not worry about obtaining the correlation between the ordered residuals and normal scores in part d). Just construct the plot and comment. Instead of part e), generate a histogram of the residuals with added density curves, and comment on it.

Prob-2 KNNL 3.15 (CH03PR15)
Prob-3 KNNL 3.16

Prob-4
Consider the following data set that describes the relationship between "velocity" of an enzymatic reaction (V) and the substrate concentration (C).

```
Concentration Velocity
0.02 76
0.02 47
0.06 97
0.06 107
0.11 123
0.11 139
0.22 159
0.22 152
0.56 191
0.56 201
1.10 207
1.10 200
```

A common model used to describe the relationship between "velocity" and concentration is the Michaelis-Menten model,

$$V = \frac{\theta_1 C}{\theta_2 + C}$$

where $\theta_1$ is the maximum velocity of the reaction and $\theta_2$ describes how quickly (in terms of increasing concentration) the reaction will reach maximum velocity. With this model, $1/V$ can be written as a linear model with explanatory variable $1/C$.

$$\frac{1}{V} = \frac{1}{\theta_1} + \frac{\theta_2}{\theta_1}\frac{1}{C} = \beta_0 + \beta_1 \frac{1}{C}$$

You are asked to investigate whether this linear transformation results in a good or poor fit by doing the following steps:

a). Generate a scatterplot of $V$ vs $C$. Comment on the shape.
b). Define new variables for $1/V$ and $1/C$ and generate a scatterplot. These new variables can be defined in SAS as follows:

```
data new;
input c v;
vinv = 1/v;
cinv = 1/c;
datalines;
PUT DATA HERE
;
```

Does the fit appear linear? Does there appear to be any violation of assumptions?
c). Is the distribution of $1/C$ different than $C$? Are there any points that may be more influential in determining the fit?
d). Determine the least squares regression line for $1/V$ vs $1/C$. Save the residuals and predicted values and generate a residual plot. Does the residual plot suggest any problems with assumptions?
e). Generate a scatterplot of V versus C and add the the predicted values of V from your linear regression model (i.e., transform back to the original units). Comment on the fit. If you wanted to do this within SAS, simply take the predicted values from the regression model and "re-inverse" them. For example, consider new1 to be the data set containing the residuals and predicted values.

```
data new1;
set new1;
pred1 = 1/pred;

symbol2 v=plus i=sm5;
symbol1 v=circle;

proc gplot;
plot v*c pred1*c / overlay;
```

Prob-5 KNNL Problem 4.6
Prob-6 KNNL Problem 4.10
Prob-7 KNNL Problem 4.21