Quiz Practice Questions 6 (Attendance 12) for Statistics 512
Applied Regression Analysis
Material Covered: Chapter 13

These are practice questions for the quiz. The quiz (not the practice questions) is worth 5% and marked out of 5 points. One or more questions is closely, but not necessarily exactly, related to one or more of these questions will appear on the quiz. These practice questions are *not* to be handed in. Quizzes are to be done *using Vista* on the Internet **before** 4am (West Lafayette time!) of the date of the quiz. Vista will *not* allow any quiz to be done late. It is *highly* recommended that you complete this practice quiz, by hand, *before* logging onto Vista. The quiz is an **individual** one which means that each student does this quiz by themselves without help from others.

Applied Linear Statistical Models (Neter et al.) Questions.
All hints are taken from Kuhn's Workbook, unless otherwise specified.

Chapter	Problem(s)	hints
13, pages 559–566	13.1, 13.2	intrinsically linear?
	13.5,13.6,13.7,13.8,13.9	home computers

(13.1) intrinsically linear?

- **(a)** $f = \exp(\gamma_0 + \gamma_1 X)$
  intrinsically linear because

$$\begin{aligned} \ln Y &= \ln \exp(\gamma_0 + \gamma_1 X) \\ Y' &= \gamma_0 + \gamma_1 X \end{aligned}$$

  where $Y' = \ln f$

- **(b)** $f = \gamma_0 + \gamma_1 (\gamma_2)^{X_1} \gamma_3 X_2$
  nonlinear because, there does not appear to be a transformation to convert f into the form,

$$Y = \gamma_0 + \gamma_1 X_1' + \gamma_2 X_2'$$

- **(c)** $f = \gamma_0 + \frac{\gamma_1}{\gamma_0} X$
  nonlinear because, although

$$Y = \gamma_0 + \gamma_1' X$$

  where $Y = f$ and $\gamma_1' = \frac{\gamma_1}{\gamma_0}$, this is linear in the variables, *not* the parameters, where two parameters have been condensed into one.

(13.2) intrinsically linear?

(a) $f = \exp(\gamma_0 + \gamma_1 \log_e X)$
intrinsically linear because

$$\begin{aligned} \ln Y &= \ln \exp(\gamma_0 + \gamma_1 \log_e X) \\ Y' &= \gamma_0 + \gamma_1 \log_e X \\ &= \gamma_0 + \gamma_1 X' \end{aligned}$$

where $Y' = \ln f$ and $X' = \log_e X$. The resulting transformed function is linear in the parameters, although not linear in either of the variables, (X, Y).

(b) $f = \gamma_0 (X_1)^{\gamma_1} (X_2)^{\gamma_2}$
intrinsically linear because

$$\begin{aligned} \ln Y &= \ln \gamma_0 (X_1)^{\gamma_1} (X_2)^{\gamma_2} \\ Y' &= \ln \gamma_0 + \gamma_1 \ln X_1 + \gamma_2 \ln X_2 \\ Y' &= \gamma_0' + \gamma_1 X_1' + \gamma_2 X_2' \end{aligned}$$

where $Y' = \ln f$, $\gamma_0' = \ln \gamma_0$, $X_1' = \ln X_1$ and $X_2' = \ln X_2$

(c) $f = \gamma_0 - \gamma_1 (\gamma_2)^X$
nonlinear there does not appear to be any transformation to make it linear in the variables.

(13.5) Home Computers: qz6-13-5-computer-nonlinear

(a) *Initial Estimates.*
Initial estimates for the parameters, γ_0, γ_1 and γ_2, of the exponential model,

$$Y_i = \gamma_0 + \gamma_2 \exp(-\gamma_1 X_i) + \varepsilon_i$$

are obtained by using the linear regression,

$$Y_i' = \beta_0 + \beta_2 X_i$$

where $Y_i' = \ln Y_i$, $\beta_0 = \ln \gamma_2$ and $\beta_1 = -\gamma_1$, $\gamma_0 = 0$
and so, from the SAS output, the initial estimates for
γ_0 is: $g_0^{(0)} = 0$
γ_1 is: $g_1^{(0)} = -b_1 = -(-0.03020) = 0.03020$
γ_2 is: $g_2^{(0)} = \exp(b_0) = \exp(-0.34698) = 0.70682$

(b) *Nonlinear Least Squares Analysis.*
From SAS, using the starting values
$g_0^{(0)} = 0$, $g_1^{(0)} = 0.03020$ and $g_2^{(0)} = 0.70682$, gives

$$g_0 = 0.08360, \quad g_1 = 0.06412, \quad g_2 = 0.83053$$

In other words, the estimated nonlinear function is

$$\hat{Y}_i = 0.08360 + 0.83053 \exp(-0.06412 X_i)$$

(13.6) Home Computers: qz6-13-6-computer-nonlinear-plot

(a) *Nonlinear Regression Plot*
From the SAS output, the following nonlinear regression

$$\hat{Y}_i = 0.08360 + 0.83053\exp(-0.06412X_i)$$

does, in fact, appear to fit the data well.

(b) *Various Residual Plots.*

(i) Both residual plots (versus predicted, $\hat{Y}$ and versus price, X) seem to indicate a sinusodal pattern and so seem to indicate non–constant variance.

(ii) The normal probability plot, though, seems to be near to a straight line and so indicates normality in the residual data.

(13.7) Home Computers: qz6-13-7-computer-nonlinear-lof

1. *Statement.*
   The statement of the test is
   $H_0 : E\{Y\} = \gamma_0 + \gamma_2 \exp(-\gamma_1 X)$ versus $H_0 : E\{Y\} \neq \gamma_0 + \gamma_2 \exp(-\gamma_1 X)$

2. *Test.*
   The test statistic[1] is

$$\begin{aligned} F^* &= \frac{SSE(R) - SSE(F)}{df_R - df_F} \div \frac{SSE(F)}{df_F} \\ &= \frac{SSE - SSPE}{(n-2)-(n-c)} \div \frac{SSPE}{n-c} \\ &= \frac{SSLF}{c-2} \div \frac{SSPE}{n-c} \\ &= \frac{0.02794 - 0.0038}{7} \div \frac{0.0038}{10} \\ &= \end{aligned}$$

   (circle one) **9.075** / **45.82** / **58.57**.
   The critical value at $\alpha = 0.01$, with 3 and 10 degrees of freedom, is
   (circle one) **4.83** / **5.20** / **7.32**
   (Use PRGM INVF ENTER 7 ENTER 10 ENTER 0.99 ENTER)

3. *Conclusion.*
   Since the test statistic, 9.075, is larger than the critical value, 5.20, we (circle one) **accept** / **reject** the null hypothesis that the regression function is $E\{Y\} = \gamma_0 + \gamma_2 \exp(-\gamma_1 X)$.

[1] From SAS,
$SSE = 0.0279$ from the "reduced" nonlinear model, $Y_i = \gamma_0 + \gamma_2 \exp(-\gamma_1 X_i) + \varepsilon_i$,
and $SSPE = 0.00380$ from the 'full" ANOVA model, $Y_{ij} = \mu_j + \varepsilon_{ij}$.

(13.8) Home Computers: qz6-13-8-computer-nonlinear-CI

The 90% Bonferroni simultaneous interval[2] for γ_0 (knowing that it is one of three ($m = 3$) parameters, γ_0, γ_1 and γ_3) is given by

$$\begin{aligned} g_0 \pm t(1-\alpha/2m; n-p)s\{g_0\} &= 0.08360 \pm t(1-0.10/2(3); 20-3)0.01836 \\ &= 0.08360 \pm 2.3156(0.01836) \\ &= (0.041086, 0.126114) \end{aligned}$$

In a similar way, the other intervals are
for γ_1: $0.06412 \pm 2.3156(0.006024) = (0.050171, 0.078069)$
for γ_2: $0.83053 \pm 2.3156(0.02603) = (0.77025, 0.89081)$

[2]The CIs calculated here are not the same as are calculated on the SAS output. The two calculations are the same, except, here, we use $t(1-\alpha/2m; n-p) = t(1-0.10/2(3); 20-3) = 2.3156$ at $\alpha = 0.10$, whereas SAS calculates at $\alpha = 0.05$.

(13.9) Home Computers: qz6-13-9-computer-nonlinear2

(a) *Nonlinear Least Squares Analysis.*
From SAS, using the starting values
$g_0^{(0)} = 0$, $g_1^{(0)} = 0.03020$, $g_2^{(0)} = 0.70682$ and $g_3^{(0)} = 0$ gives

$$g_0 = 0.08460, \quad g_1 = 0.06412, \quad g_2 = 0.83053, \quad g_3 = -0.00200$$

In other words, the estimated nonlinear function in this case is

$$\hat{Y}_i = 0.08460 + 0.83053 X_{i2} - 0.00200 \exp(-0.06412 X_i)$$

(b) *90% Confidence Interval.*
The 90% Bonferroni simultaneous interval[3] for γ_0 (knowing that it is one of three ($m = 4$) parameters, γ_0, γ_1, γ_3 and γ_4) is given by

$$\begin{aligned} g_0 \pm t(1 - \alpha/2m; n - p)s\{g_3\} &= 0.08360 \pm t(1 - 0.10/2(4); 20 - 4)0.01836 \\ &= 0.01868 \pm 1.96(0.01836) \\ &= (-0.386, 0.0346) \end{aligned}$$

Since this interval includes zero (0), this indicates that there is *no* city (location) effect.
The previous test indicated that the model (without city effect) did not fit the data well. In other words, the previous test indicated that another model should be fit to the data. The present result indicates that a revised model that includes city effect also does not seem to fit the data well. In other words, other parameters should be added to the model or the model should be changed in some other way to attempt to fit the data better.

[3] The CIs calculated here are not the same as are calculated on the SAS output, as explained above.