# STA 6207 – Practice Problems – Nonlinear Regression

Q.1. A study was conducted to measure the effects of pea density ($X_1$, in plants/m$^2$) and volunteer barley density ($X_2$, in plants/m$^2$) on pea seed yield (Y). The researcher fit a nonlinear regression model:

$$E(Y) = \frac{\beta_1 X_1}{1 + \beta_2 X_1 + \beta_3 X_2}$$

p.1.a. Assuming $\beta_1, \beta_2, \beta_3 > 0$, what is E(Y) as pea density goes to infinity?

p.1.b. Assuming $\beta_1, \beta_2, \beta_3 > 0$, what is E(Y) as volunteer barley density goes to infinity?

The following table gives the estimated regression coefficients, standard errors, z-stats, and P-values (the sample size was huge):

| Coefficient | Estimate | Std Error | Z | P-value |
|---|---|---|---|---|
| B1 | 7.6 | 2 | 3.80 | 0.00014 |
| B2 | 0.019 | 0.007 | 2.71 | 0.00664 |
| B3 | 0.17 | 0.05 | 3.40 | 0.00067 |

p.1.c. What can we say about mean pea seed yield as volunteer barley density increases, controlling for pea density ($\alpha$=0.05)?

a) Increases b) Decreases c) Not Related to barley density

p.1.d. What can we say about mean pea seed yield as pea density increases, controlling for volunteer barley density ($\alpha$=0.05)?

a) Increases b) Decreases c) Not Related to pea density

p.1.e. Give the estimated pea yield for the following combinations of pea density and barley density: (pd=100,bd= 0), (100, 200), (200, 0), (200,200).

100 , 0:

100 , 200:

200 , 0:

200 , 200:

Q.2. A nonlinear regression model is fit, relating available chlorine (Y, proportion) to the length of time since it was produced (X, in weeks). Based on data considerations, the model fit is of the form with output below:

$$y_i = \alpha + (0.49 - \alpha)\exp\left[-\beta(x_i - 8)\right] + \varepsilon_i$$

```
 Nonlinear Regression Summary Statistics     Dependent Variable CHLOR

   Source                 DF  Sum of Squares  Mean Square

   Regression              2        7.98200        3.99100
   Residual               42   5.001680E-03   1.190876E-04
   Uncorrected Total      44        7.98700

   (Corrected Total)      43         .03950

   R squared = 1 - Residual SS / Corrected SS =     .87338

                                              Asymptotic 95 %
                            Asymptotic     Confidence Interval
   Parameter   Estimate     Std. Error     Lower         Upper

   ALPHA       .390140032   .005044840   .379959134   .400320931
   BETA        .101632757   .013360412
```

p.2.a. Give the predicted proportion of chlorine remaining at each of the following times since production:

p.2.a.i. 10 weeks:

p.2.a.ii. 20 weeks:

p.2.a.iii. 30 weeks:

p.2.b. Compute a 95% Confidence Interval for $\beta$

p.2.c. Give the estimated rate of change in chlorine remaining as a function of age (using point estimates of parameters in the derivative $\partial Y/\partial X$

Q.3. A nonlinear regression model is to be fit, relating Area (Y, in $m^2$) of palm trees to age (X, in years) by the Gompertz model: $E(Y) = \alpha + \exp[-\beta * \exp(-\gamma X)]$ for $\alpha > 0, \beta > 0, \gamma > 0$.

p.3.a. What is E(Y), in terms of the model parameters when X=0? When $X \rightarrow \infty$?

p.3.b. What is the instantaneous growth rate of Area in terms of the model parameters and X?

Q.4. An enzyme kinetics study of the velocity of reaction (Y) is expected to be related to the concentration of the chemical (X) by the following model (based on n=18 observations):

$$Y_i = \frac{\beta_0 X_i}{\beta_1 + X_i} + \varepsilon_i \quad \varepsilon_i \sim N\left(0, \sigma^2\right)$$

The following results are obtained.

```
                              The NLIN Procedure

                                                 Approx
        Parameter       Estimate          Std Error
        b0               28.1               0.73
        b1               12.6               0.76
```

p.4.a. Give a 95% Confidence Interval for the Maximum Velocity of Reaction

p.4.b. . Give a 95% Confidence Interval for the dose needed to reach 50% of Maximum Velocity of Reaction

p.4.c. Give the predicted velocity when X=0, 10, 20and difference between each

$\boldsymbol{Y}_0 =$ $\boldsymbol{Y}_{10} =$ $\boldsymbol{Y}_{20} =$

$\boldsymbol{Y}_{10} - \boldsymbol{Y}_0 =$ $\boldsymbol{Y}_{20} - \boldsymbol{Y}_{10} =$

Q.5. A study was conducted to compare growth rates of Thai and Commercial Tilapia. The authors fit separate models for the two strains, with samples of 20 fish of each type selected at 5 time points and weighed (then sacrificed ☹). Thus 100 fish of each strain were measured. The model fit by unweighted nonlinear least squares was (where Y=Weight, and X=Age, in days):

$$Y_i = \mu_i + \varepsilon_i = \beta_0 e^{\beta_1 X_i} + \varepsilon_i$$

p.5.a. How would you interpret $\beta_0$ and $e^{\beta_1}$?

p.5.b. Obtain $\dfrac{\partial \mu_i}{\partial \beta_0}$ and $\dfrac{\partial \mu_i}{\partial \beta_1}$

p.5.c. The estimates and standard errors for $\hat{\beta}_0$ and $\hat{\beta}_1$ are given in the following table for each strain. Obtain 95% CI's for all model parameters.

| Strain | Thai | Thai | Commercial | Commercial |
|---|---|---|---|---|
| **Parameter** | **Estimate** | **StdError** | **Estimate** | **StdError** |
| **Beta0** | **30.5771** | **3.2614** | **21.6209** | **2.6555** |
| **Beta1** | **0.0132** | **0.0008** | **0.0168** | **0.0010** |

$\beta_0^T$ : ____________________ $\beta_1^T$ : ____________________ $\beta_0^C$ : ____________________ $\beta_1^C$ : ____________________

p.5.d As a crude test of equality of Strains, based on the confidence intervals, do you reject the null hypothesis:

$H_0 : \beta_0^T = \beta_0^C, \quad \beta_1^T = \beta_1^C$ Note that their estimates are independent as models were fit separately. **Yes / No**

p.5.e. At what age (if any) do the two strains' curves cross?

Q.6. A study was conducted to study the viscoelastic behavior of polypropylene fibers. They fit nonlinear regressions for 2 fiber types (Blue and Black). The Model fit (separately for each type) was, where Y = Relaxation Modulus and X = Time (seconds):

$$Y_i = \mu_i + \varepsilon_i = \beta_0 + \beta_1 e^{-X_i/\beta_2} + \varepsilon_i$$

p.6.a. How would you interpret $\beta_0$ and $\beta_1$?

p.6.b. Obtain $\frac{\partial \mu_i}{\partial \beta_0}, \quad \frac{\partial \mu_i}{\partial \beta_1}, \quad \text{and} \quad \frac{\partial \mu_i}{\partial \beta_2}$

p.6.c. The estimates and standard errors for $\hat{\beta}_0$, $\hat{\beta}_1$, and $\hat{\beta}_2$ are given in the following table for each strain. Obtain 95% CI's for all model parameters (for each model, n = 9).

| | Blue | Blue | Black | Black |
|---|---|---|---|---|
| **Parameter** | **Estimate** | **Std Error** | **Estimate** | **Std Err** |
| **Beta0** | **3.95** | **0.15** | **4.51** | **0.17** |
| **Beta1** | **1.63** | **0.17** | **1.79** | **0.20** |
| **Beta2** | **21.69** | **6.47** | **21.51** | **6.70** |

$\beta_0^{BLUE}$ : ____________________ $\beta_1^{BLUE}$ : ____________________ $\beta_2^{BLUE}$ : ____________________

$\beta_0^{BLACK}$ : ____________________ $\beta_1^{BLACK}$ : ____________________ $\beta_2^{BLACK}$ : ____________________

p.6.d As a crude test of equality of Strains, based on the confidence intervals, do you reject the null hypothesis:

$H_0 : \beta_0^{BLUE} = \beta_0^{BLACK}, \quad \beta_1^{BLUE} = \beta_1^{BLACK}, \quad \beta_2^{BLUE} = \beta_2^{BLACK}$ Note that their estimates are independent as models were fit separately. **Yes / No**

p.6.e. For each fiber type give the difference between the fitted maximum and minimum Relaxation modulus.

BLUE ____________________________________________ BLACK ____________________________________________

Q.7. A nonlinear regression model was fit, relating cumulative cellular phones in Greece (Y, in millions) to year since 1994 (X=Year-1994). The authors considered various models, including the following Gompertz model (n can be determined from the graph):

$$E\{Y\} = \beta_0 e^{-e^{\{-\beta_1 - (\beta_2 X)\}}} \qquad \beta_0, \beta_2 > 0 \qquad e = 2.718...$$

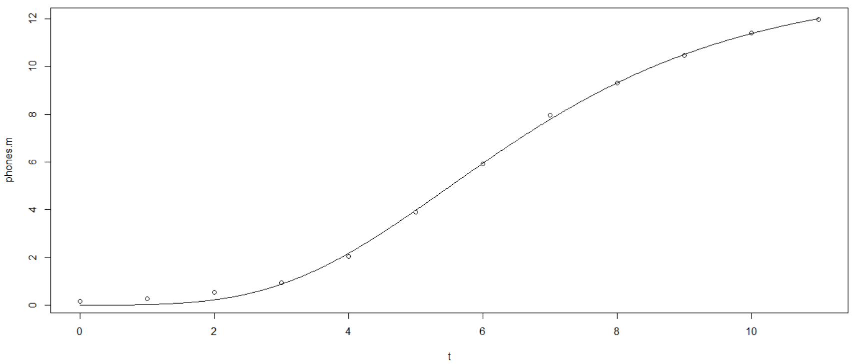
```
Formula: phones.m ~ b0 * exp(-exp(-b1 - b2 * X))

Parameters:
   Estimate Std. Error t value Pr(>|t|)
b0 13.374      0.292     45.78 5.66e-12 ***
b1 -2.208      0.083    -26.46 7.60e-10 ***
b2  0.403      0.018     21.52 4.76e-09 ***
```

p.7.a. Give the fitted values (predicted cumulative sales) based on the model for years 1994 and 2004.

1994 ____________________________________ 2004 ____________________________________

p.7.b. As X →∞, What parameter does E{Y} approach? Compute a 95% Confidence Interval for that parameter.

Parameter: _____________ Lower Bound _________________ Upper Bound __________________ p.7.c. The following plot shows the fitted equation and data. Give the approximate time when sales cross 6 (million units sold). Note that the horizontal axis is X=Year-1994.




Q.8. A nonlinear regression model was fit for steel sheets, relating plain strain form limit (Y) to scaled ultimate tensile stress (X = UTS-200). The relationship is modelled with following equation and was based on an experiment with n = 56 steel sheets (with following expected signs for the the regression coefficients).

**Answer all parts in terms of X, not UTS.**

$$Y = \beta_0 + \beta_1 e^{-\beta_2 X} + \varepsilon \quad \varepsilon \sim NID\left(0,\sigma^2\right) \quad \beta_0 > 0,\ \beta_1 > 0,\ \beta_2 > 0$$

p.8.a. Give the intercept and asymptote (limit as X → ∞) in terms of the model parameters.

Intercept ____________________ Asymptote _______________________

p.8.b. Obtain the the derivatives of E{Y} with respect to $\beta_0$, $\beta_1$, $\beta_2$

p.8.c. At what point $X^*$ does the model reach half-way between the intercept and asymptote?

p.8.d. The estimated regression coefficients and variance-covariance matrix are given below. Obtain approximate 95% Confidence Intervals for $\beta_0$, $\beta_1$, $\beta_0+\beta_1$

| Parameter | Estimate | | V{B} | | |
|---|---|---|---|---|---|
| B0 | 0.271242 | | 0.000191 | 0.001077 | 0.000032 |
| B1 | 0.543144 | | 0.001077 | 0.030845 | 0.000600 |
| B2 | 0.012461 | | 0.000032 | 0.000600 | 0.000013 |

$\beta_0$ ___________________________ $\beta_1$ ___________________________ $\beta_0 + \beta_1$ ___________________________

p.8.e. Consider the function $h(\beta_2) = 1/\beta_2$. The estimated estimated standard error of $h(\beta_2)$ is given below.

$$\hat{SE}\left\{h\left(\hat{\beta}_2\right)\right\} = \left|H\left(\hat{\beta}_2\right)\right| \hat{SE}\left\{\hat{\beta}_2\right\} \quad \text{where} \quad H(\beta_2) = \frac{\partial h(\beta_2)}{\beta_2}$$

Obtain an approximate 95% Confidence Interval for $h(\beta_2) = 1/\beta_2$.

Q.9. A paper published in 1918 based on previously published data, fit a model relating heat output (H, in calories) to body mass (M, in kg) and work equivalent (W, in calories) for n = 37 subjects performing on a bicycle ergometer. The author considered two models, one linear and one nonlinear. They are given below.

Model 1 (Linear): $H = \beta_0 + \beta_M M + \beta_W W + \varepsilon \qquad \varepsilon \sim NID\left(0, \sigma_1^2\right)$

Model 2 (Nonlinear): $H = \beta_0 + \beta_1 M + \dfrac{W}{\beta_2 + \beta_3 M} + \varepsilon \qquad \varepsilon \sim NID\left(0, \sigma_2^2\right)$

p.9.a. For model 2, give $\dfrac{\partial E\{H_i\}}{\partial \beta_3}$ where $E\{H_i\} = \beta_0 + \beta_1 M_i + \dfrac{W_i}{\beta_2 + \beta_3 M_i}$

p.9.b. The summary outputs are given below. Give the fitted values and residuals for subject 1 for each model. ($H_1$=3398, $M_1$=76.2, $W_1$=156.8)

| Model1 | Estimate |
|---|---|
| (Intercept) | 977.425 |
| mass | 17.778 |
| work | 6.244 |
| | |
| SSE1 | 734835.8 |

| Model2 | Estimate |
|---|---|
| b0 | 1180.00 |
| b1 | 14.78 |
| b2 | 0.2004 |
| b3 | -0.0006 |
| SSE2 | 736023.6 |

Model 1: $\hat{Y}_1 =$ __________ $e_1 =$ __________ Model 2: $\hat{Y}_2 =$ __________ $e_2 =$ __________

p.9.c. Give point estimates for for the error standard deviations: $\sigma_1$ and $\sigma_2$

Model 1: $\hat{\sigma}_1 =$ __________ Model 2: $\hat{\sigma}_2 =$ __________

p.9.d. Which model would you prefer to describe in a crowded academic seminar? **Model 1** or **Model 2**