Nonlinear Modeling

Nonlinear Relations wrt X – Linear wrt βs

1) Polynomial Models: $E\{Y_i\} = \beta_0 + \beta_1 X_i + \beta_2 X_i^2$

$$\frac{\partial E\{Y_i\}}{\partial X_i} = \frac{\partial}{\partial X_i}\left[\beta_0 + \beta_1 X_i + \beta_2 X_i^2\right] = 0 + \beta_1 + 2\beta_2 X_i = h(X_i)$$

$$\frac{\partial E\{Y_i\}}{\partial \beta_0} = 1 \qquad \frac{\partial E\{Y_i\}}{\partial \beta_1} = X_i \qquad \frac{\partial E\{Y_i\}}{\partial \beta_2} = X_i^2 \qquad \text{None are functions of } \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix}$$

2) Transformed Variable Models: $E\left\{\sqrt{Y_i}\right\} = \beta_0 + \beta_1 \ln(X_{i1}) + \beta_2 \left(\frac{1}{X_{i2}}\right)$

$$\frac{\partial E\left\{\sqrt{Y_i}\right\}}{\partial X_{i1}} = \beta_1 \left(\frac{1}{X_{i1}}\right) = h_1(X_{i1}) \qquad \frac{\partial E\left\{\sqrt{Y_i}\right\}}{\partial X_{i2}} = -\beta_2 \left(\frac{1}{X_{i2}^2}\right) = h_2(X_{i2})$$

$$\frac{\partial E\{Y_i\}}{\partial \beta_0} = 1 \qquad \frac{\partial E\{Y_i\}}{\partial \beta_1} = \ln(X_{i1}) \qquad \frac{\partial E\{Y_i\}}{\partial \beta_2} = \left(\frac{1}{X_{i2}}\right) \qquad \text{None are functions of } \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix}$$

In each case: $E\{Y_i\} = f(\mathbf{X_i}, \boldsymbol{\beta}) = \mathbf{X_i'}\boldsymbol{\beta}$

Case 1: $\mathbf{X}_i' = \begin{bmatrix} 1 & X_i & X_i^2 \end{bmatrix}$ Case 2: $\mathbf{X}_i' = \begin{bmatrix} 1 & \ln(X_{i1}) & \frac{1}{X_{i2}} \end{bmatrix}$

Nonlinear Regression Models

Nonlinear Regression models often use $\boldsymbol{\gamma}$ as vector of coefficients to distinguish from linear models:

Exponential Regression Models (Often used for modeling growth, where rate of growth changes):

$$E\left\{Y_i\right\} = \gamma_0 \exp\left(\gamma_1 X_i\right) \quad \Rightarrow \quad \frac{\partial E\left\{Y_i\right\}}{\partial \gamma_0} = \exp\left(\gamma_1 X_i\right) \qquad \frac{\partial E\left\{Y_i\right\}}{\partial \gamma_1} = \gamma_0 X_i \exp\left(\gamma_1 X_i\right) \qquad \text{functions of } \boldsymbol{\gamma}$$

$$f\left(\mathbf{X_i}, \boldsymbol{\gamma}\right) = \gamma_0 \exp\left(\gamma_1 X_i\right) \neq \mathbf{X_i}'\boldsymbol{\gamma}$$

More general exponential model (with errors independent and $N\left(0,\sigma^2\right)$):

$$Y_i = \gamma_0 + \gamma_1 \exp\left(\gamma_2 X_i\right) + \varepsilon_i \quad \text{Typically, } \gamma_0 > 0, \gamma_1 < 0, \gamma_2 < 0$$

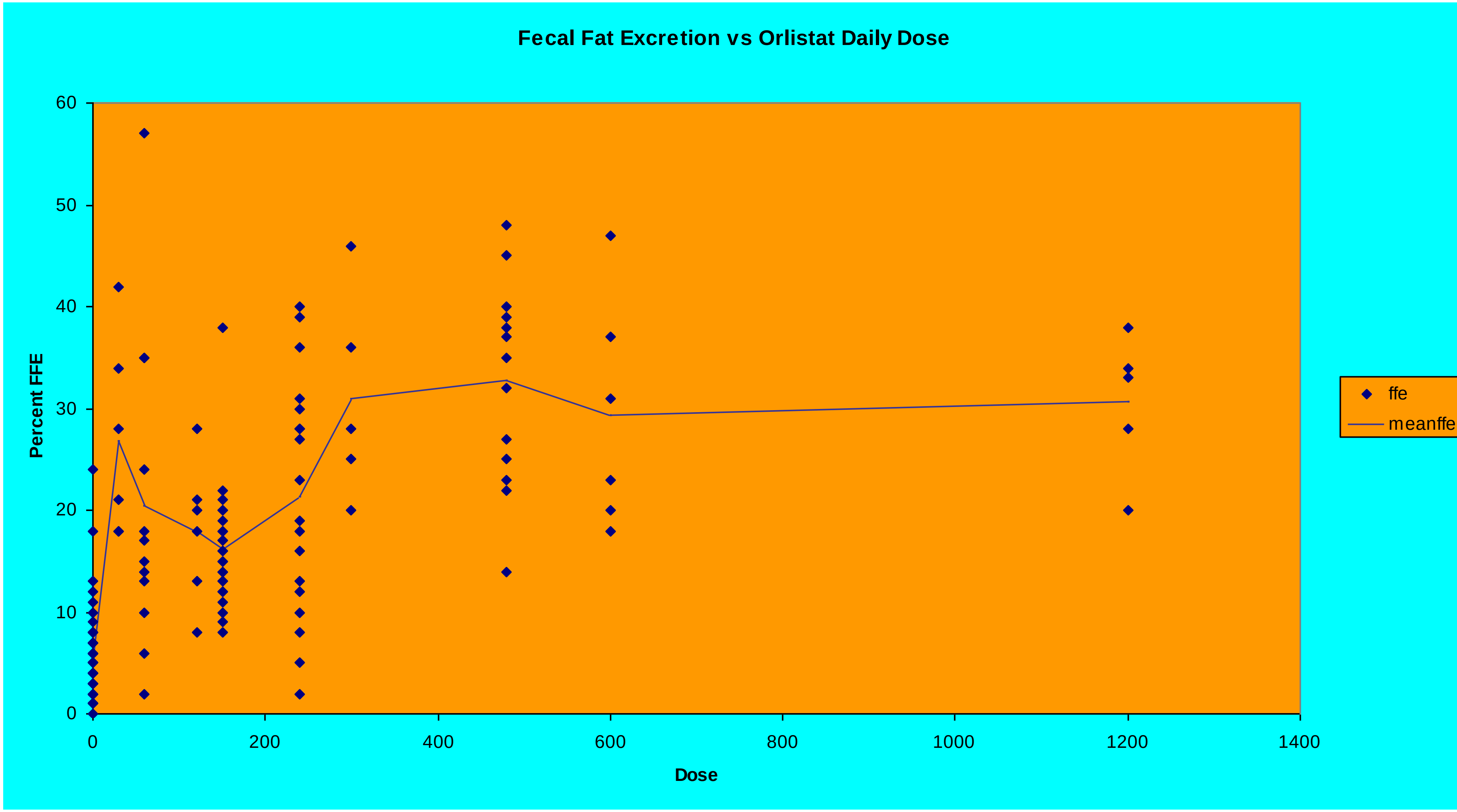
$\Rightarrow$ Intercept: $E\left(Y_i \mid X_i = 0\right) = \gamma_0 + \gamma_1(1) = \gamma_0 + \gamma_1$

$\Rightarrow$ Asymptote: $E\left(Y_i \mid X_i \rightarrow \infty\right) = \gamma_0 + \gamma_1(1) = \gamma_0$

$\Rightarrow$ "Half-way" Point: $E\left(Y_i \mid X_i = \dfrac{0.693}{\left|\gamma_2\right|}\right) = \gamma_0 + \gamma_1 \exp\left(\gamma_2\left(\dfrac{0.693}{\left|\gamma_2\right|}\right)\right) = \gamma_0 + \gamma_1 \exp\left(-0.693\right) = \gamma_0 + \left(\dfrac{\gamma_1}{2}\right)$

Data Description - Orlistat

- 163 Patients assigned to one of the following doses (mg/day) of orlistat: 0, 60,120,150,240,300,480,600,1200
- Response measured was fecal fat excretion (purpose is to inhibit fat absorption, so higher levels of response are considered favorable).
- Plot of raw data displays a generally increasing but nonlinear pattern and large amount of variation across subjects.


Fecal Fat Excretion vs Orlistat Daily Dose
Percent FFE
0
10
20
30
40
50
60
0
200
400
600
800
1000
1200
1400
Dose
ffe
meanffe

Nonlinear Regression Model - Example

$$Y = \gamma_0 + \frac{\gamma_1 x}{\gamma_2 + x} + \varepsilon$$

Simple Maximum Effect model:

 γ_0 **≡ Mean Response at Dose 0**

- γ_1 **≡ Maximal Effect of Orlistat** ($\gamma_0 + \gamma_1$ = Maximum Mean Response)
- γ_2 **≡ Dose providing 50% of maximal effect (ED$_{50}$)**

Nonlinear Least Squares

$$f\left(\mathbf{X_i},\boldsymbol{\gamma}\right)= f_i\left(\boldsymbol{\gamma}\right)= f\left(\gamma_0,\gamma_1,\gamma_2\right)=\gamma_0+\frac{\gamma_1 X_i}{\gamma_2+X_i}$$

$$\frac{\partial f\left(\mathbf{X_i},\boldsymbol{\gamma}\right)}{\partial\boldsymbol{\gamma}'}=F_i\left(\boldsymbol{\gamma}\right)=\frac{\partial f_i\left(\boldsymbol{\gamma}\right)}{\partial\boldsymbol{\gamma}'}=\begin{bmatrix}1 & \dfrac{X_i}{\gamma_2+X_i} & \dfrac{-\gamma_1 X_i}{\left(\gamma_2+X_i\right)^2}\end{bmatrix}$$

$$\mathbf{Y}=\begin{bmatrix}Y_1\\ \vdots\\ Y_n\end{bmatrix}\qquad \mathbf{f}\left(\boldsymbol{\gamma}\right)=\begin{bmatrix}f_1\left(\gamma\right)\\ \vdots\\ f_n\left(\gamma\right)\end{bmatrix}=\begin{bmatrix}\gamma_0+\dfrac{\gamma_1 X_1}{\gamma_2+X_1}\\ \vdots\\ \gamma_0+\dfrac{\gamma_1 X_n}{\gamma_2+X_n}\end{bmatrix}$$

$$\mathbf{F}\left(\boldsymbol{\gamma}\right)=\begin{bmatrix}F_1\left(\gamma\right)\\ \vdots\\ F_n\left(\gamma\right)\end{bmatrix}=\begin{bmatrix}1 & \dfrac{X_1}{\gamma_2+X_1} & \dfrac{-\gamma_1 X_1}{\left(\gamma_2+X_n\right)^2}\\ \vdots & \vdots & \vdots\\ 1 & \dfrac{X_n}{\gamma_2+X_n} & \dfrac{-\gamma_1 X_n}{\left(\gamma_2+X_n\right)^2}\end{bmatrix}$$

F acts like the **X** matrix in linear regression (but depends on parameters)

Nonlinear Least Squares

Goal: Choose $\hat{\gamma}_0, \hat{\gamma}_1, \hat{\gamma}_2$ that minimize error sum of squares:

$$Q = SSE(\boldsymbol{\gamma}) = \sum_{i=1}^{n}\left(Y_i - \left[\gamma_0 + \frac{\gamma_1 X_i}{\gamma_2 + X_i}\right]\right)^2 =$$

$$= \left(\mathbf{Y} - \mathbf{f}(\boldsymbol{\gamma})\right)'\left(\mathbf{Y} - \mathbf{f}(\boldsymbol{\gamma})\right)$$

$$\frac{\partial Q}{\partial \gamma_j} = -2\sum_{i=1}^{n}\left(Y_i - \left[\gamma_0 + \frac{\gamma_1 X_i}{\gamma_2 + X_i}\right]\right)F_i(\gamma_j) \quad j = 0,1,2$$

$$\frac{\partial Q}{\partial \boldsymbol{\gamma}'} = -2\left[\mathbf{Y} - \mathbf{f}(\boldsymbol{\gamma})\right]^{\mathrm{T}}\mathbf{F}(\boldsymbol{\gamma}) \stackrel{set}{=} \begin{bmatrix}0 & 0 & 0\end{bmatrix}$$

Estimated Variance-Covariance Matrix

$$s^2\left\{\hat{\boldsymbol{\gamma}}\right\} = s^2\left(\hat{\mathbf{F}}'\hat{\mathbf{F}}\right)^{-1}$$

$$s^2 = \frac{\left(\mathbf{Y}-\hat{\mathbf{f}}\right)^{\mathbf{T}}\left(\mathbf{Y}-\hat{\mathbf{f}}\right)}{n-p}$$

$$s\left\{\hat{\gamma}_i\right\} = s\sqrt{\left(\hat{\mathbf{F}}'\hat{\mathbf{F}}\right)^{-1}_{(i+1,i+1)}}$$

Note: KNNL uses $\mathbf{g}$ for $\hat{\boldsymbol{\gamma}}$ and $\mathbf{D}$ for $\hat{\mathbf{F}}$

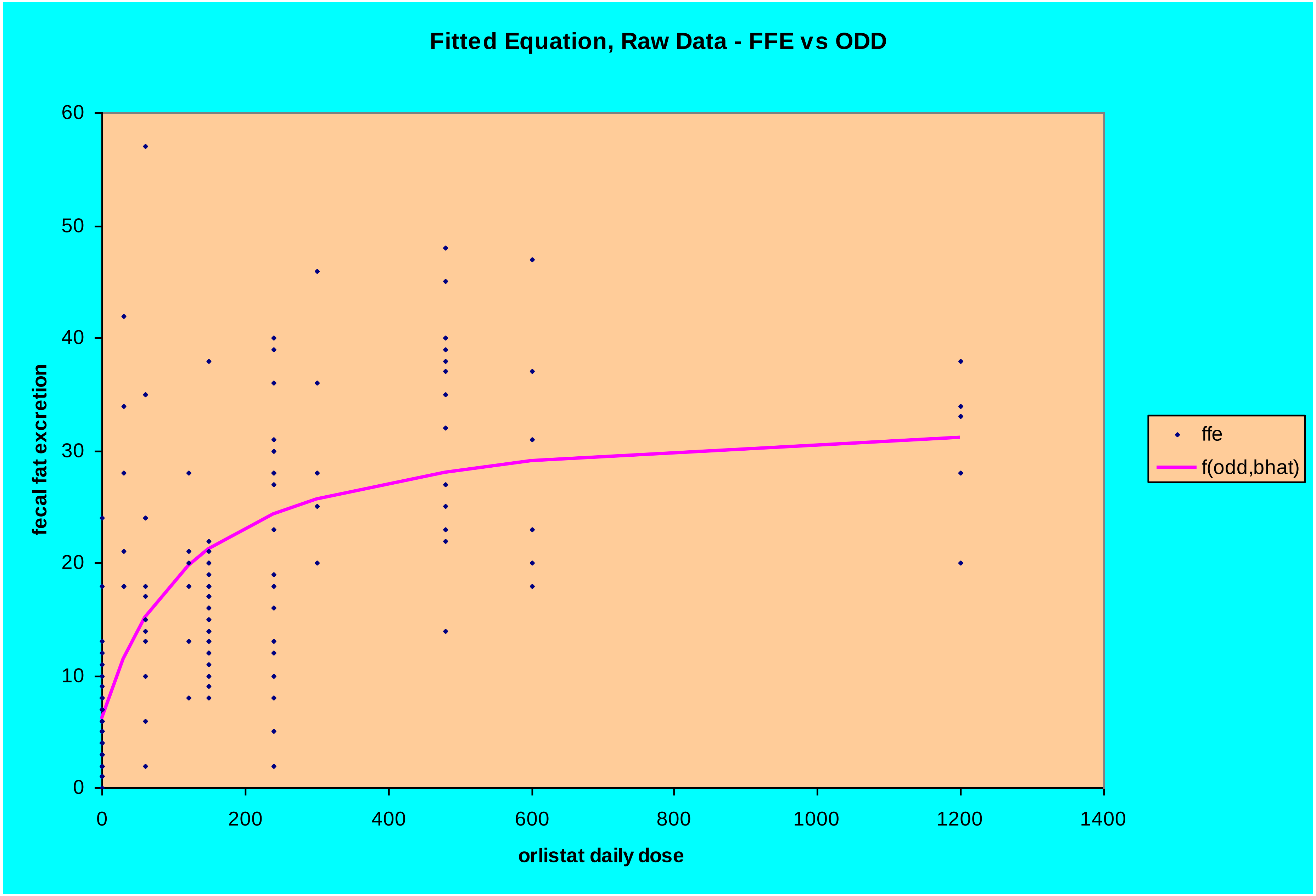
Orlistat Example

- Reasonable Starting Values:
  - γ_0: Mean of 0 Dose Group: 5
  - γ_1: Difference between highest mean and dose 0 mean: 33-5=28
  - γ_2: Dose with mean halfway between 5 and 33: 160
- Create Vectors $\boldsymbol{Y}$ and $f\,(\boldsymbol{\gamma}^{\mathbf{0}})$
- Generate matrix $\boldsymbol{F}(\boldsymbol{\gamma}^{\mathbf{0}})$
- Obtain first "new" estimate of $\boldsymbol{\gamma}$
- Continue to Convergence

Orlistat Example – Iteration History (Tolerance = .0001)

iteration	g0	g1	g2	SSE	Delta(g)
0	5.0000	28.0000	160.0	13541.6	
1	6.2379	28.5863	140.9	12945.5	365.5745418
2	6.1771	28.1281	133.7	12942.9	52.82513814
3	6.1507	27.9163	129.9	12942.2	14.44158887
4	6.1361	27.7967	127.8	12942.0	4.506448063
5	6.1277	27.7272	126.5	12941.9	1.510150161
6	6.1227	27.6861	125.8	12941.9	0.526393989
7	6.1197	27.6615	125.4	12941.9	0.187692352
8	6.1180	27.6467	125.1	12941.9	0.067822683
9	6.1169	27.6377	125.0	12941.9	0.024703325
10	6.1162	27.6323	124.9	12941.9	0.009040833
11	6.1158	27.6291	124.8	12941.9	0.003318268
12	6.1156	27.6271	124.8	12941.9	0.001220029
13	6.1155	27.6259	124.8	12941.9	0.000449042
14	6.1154	27.6251	124.7	12941.9	0.000165379
15	6.1153	27.6247	124.7	12941.9	6.09317E-05

$$\hat{Y} = 6.12 + \frac{27.62X}{124.7 + X}$$



Fitted Equation, Raw Data - FFE vs ODD
fecal fat excretion
orlistat daily dose
0
10
20
30
40
50
60
0
200
400
600
800
1000
1200
1400
ffe
f(odd,bhat)

Variance Estimates/Confidence Intervals

$$s^2 = \frac{\sum_{i=1}^{163}\left(Y_i - f_i\left(\hat{\boldsymbol{\gamma}}\right)\right)^2}{163-3} = 80.89$$

$$\mathbf{s}^2\left\{\hat{\boldsymbol{\gamma}}\right\} = s^2\left(\hat{\mathbf{F}}'\hat{\mathbf{F}}\right)^{-1} = \begin{bmatrix} 1.1594 & -0.7219 & 15.609 \\ -0.7219 & 12.081 & 130.14 \\ 15.609 & 130.14 & 2238.76 \end{bmatrix}$$

Parameter	Estimate	Std. Error	95% CI
γ_0	6.12	1.08	(3.96 , 8.28)
γ_1	27.62	3.48	(20.66 , 34.58)
γ_2	124.7	47.31	(30.08 , 219.32)

Notes on Nonlinear Least Squares

Large-Sample Theory:

When $\varepsilon_i \sim N\left(0,\sigma^2\right)$ independent , for large *n*: $\hat{\boldsymbol{\gamma}}$ is approximately normal

$E\left\{\hat{\boldsymbol{\gamma}}\right\} \approx \boldsymbol{\gamma}$ Approximate $\sigma^2\left\{\hat{\boldsymbol{\gamma}}\right\}$ estimated by $\mathbf{s}^2\left\{\hat{\boldsymbol{\gamma}}\right\} = MSE\left(\hat{\mathbf{F}}'\hat{\mathbf{F}}\right)^{-1}$

$$\Rightarrow \quad \hat{\boldsymbol{\gamma}} \sim N\left(\boldsymbol{\gamma}, \sigma^2\left(\mathbf{F}'\mathbf{F}\right)^{-1}\right) \quad \text{(Approximately)}$$

For small samples:

- When errors are normal, independent, with constant variance, we can often use the t-distribution for tests and confidence intervals (software packages do this implicitly)
- When the extent of nonlinearity is extreme, or normality assumptions do not hold, should use bootstrap to estimate standard errors of regression coefficients

Example

Many patients get concerned when a medical test involves an injection of radioactive material. For example, to scan a gallbladder, a few drops of Technetium-99m isotope are used.

Half of the Technetium-99m would be gone in about 6 hours. However, it takes about 24 hours for the radiation levels to reach what we are exposed to in day-to-day activities.

Below is given the relative intensity of radiation as a function of time.

t(hrs)	0	1	3	5	7	9
γ	1.000	0.891	0.708	0.562	0.447	0.355

If the level of the relative intensity of radiation is related to time via an exponential formula

$$\gamma = Ae^{\lambda t}$$

The value of the regression constants A and λ,

Example contd...

i	t_i	y_i	$y_i = \ln y_i$	$t_i y_i$	t_i^2
1	0	1	0.00000	0.0000	**0.0000**
2	1	0.891	−0.11541	−0.11541	**1.0000**
3	3	0.708	−0.34531	−1.0359	**9.0000**
4	5	0.562	−0.57625	−2.8813	25.0000
5	7	0.447	−0.80520	−5.6364	49.0000
6	9	0.355	−1.0356	−9.3207	81.0000
$\sum_{i=1}^{6}$	25.0000		−2.8778	−18.990	**165.00**

Example contd…

$$
\begin{aligned}
A &= e^{a_0} \\
&= e^{-2.6150\times 10^{-4}} \\
&= 0.99974
\end{aligned}
$$

$$\lambda = a_1 = -0.11505$$