Nonlinear Models
So far the models we have studied this semester have been linear in the sense that our model for the mean has been a linear function of the parameters.
We have assumed

$$E(Y) = X \beta$$

which is equivalent to

$$E(Y_i) = X_i' \beta \quad \text{for} \quad i = 1, \ldots, n$$

where

$$X_i' = [X_{i1}, X_{i2}, \ldots, X_{ip}]$$

is the $i$th row of $X$. 
We can write our model for the mean response as

$$E(y_i) = f(x_i, \beta)$$

where

$$f(x_i, \beta) = x_i \beta.$$
\( f(x_i, \beta) = x'_i \beta \) is said to be linear in the parameters of \( \beta \) because 
\[ x'_i \beta = x_{i1} \beta_1 + x_{i2} \beta_2 + \cdots + x_{ip} \beta_p \]
is a linear combination of \( \beta_1, \beta_2, \ldots, \beta_p \).

\( f(x_i, \beta) = x'_i \beta \) is linear in \( \beta \) even if the predictor variables are nonlinear functions of other variables.
For example, if

\[ X_{i1} = 1 \]
\[ X_{i2} = \text{Amount of fertilizer applied to plot } i \]
\[ X_{i3} = (\text{Amount of fertilizer applied to plot } i)^2 \]
\[ X_{i4} = \log(\text{Concentration of fungicide on plot } i) \]

\[ f(X_i, \beta) = X_i^\beta = X_{i1} \beta_1 + X_{i2} \beta_2 + X_{i3} \beta_3 + X_{i4} \beta_4 \]

\[ = \beta_1 + \text{fert}_i \beta_2 + \text{fert}_i^2 \beta_3 + \log(\text{fung}_i) \beta_4 \]

is still linear in the parameters \( \beta_1, \beta_2, \beta_3, \beta_4 \).
Now we want to consider nonlinear models for the mean $E(y_i)$.

We want to consider cases where

$$f(x_i; \beta)$$

cannot be written as a linear combination of

$$\beta_1, \beta_2, \ldots, \beta_p.$$
Why would we want to consider a non-linear model?

Pinheiro and Bates (2000) give some reasons:

• Mechanistic - based on theoretical considerations about the mechanism producing the response

• Often interpretable and parsimonious

• Can be valid beyond the range of the observed data.
Example 1:

Logistic Growth Model

$y_i$ is height of a tree at age $x_i$ ($i = 1, \ldots, n$).

$$E(y_i) = f(x_i, \beta) = \frac{\beta_1}{1 + \exp \frac{\beta_2}{\beta_3} - (x_i - \beta_2)/\beta_3}$$

$y_i = f(x_i, \beta) + \varepsilon_i$, $E(\varepsilon_i) = 0$, $\text{Var}(\varepsilon_i) = \sigma^2$

for $i = 1, \ldots, n$. 
\[ E(Y_i) = f(x_i, \beta) = \frac{\beta_1}{1 + e^{\xi \cdot (x_i - \beta_2) / \beta_3}} \]

Note that as age \( \to \infty \), \( E(Y_i) \to \beta_1 \)

At age = \( \beta_2 \), \( E(Y_i) = \beta_1/2 \), half the asymptotic height.

\( \beta_3 \) is the time it takes a tree to grow from height \( \beta_1/2 \) to about \( 3\beta_1/4 \).

\[ \left( \frac{3/4}{1 + e^{-1}} \right) \]
Example Logistic Curves

\[ \beta_1 = 20, \beta_3 = 7, \beta_3 = 3 \]

\[ \beta_1 = 18, \beta_3 = 10, \beta_3 = 5 \]
Least Squares Estimation

\[ y_i = f(x_i, \beta) + \varepsilon_i \quad i = 1, \ldots, n \]

Find \( \hat{\beta} \) that minimizes

\[ g(\beta) = \sum_{i=1}^{n} [y_i - f(x_i, \beta)]^2. \]

We start with the same strategy used in the linear case to obtain the normal equations.
\[
\frac{dg(b)}{db} = 0 
\]

\[
\frac{dg(b)}{db_i} = 2 \sum_{i=1}^{n} \left[ y_i - f(x_i, b) \right] \frac{df(x_i, b)}{db_i} 
\]

\[
\frac{dg(b)}{db_p} = 2 \sum_{i=1}^{n} \left[ y_i - f(x_i, b) \right] \frac{df(x_i, b)}{db_p} 
\]

(Estimating Equations)
If we define

\[ f(x, \beta) = \begin{bmatrix} f(x_1, \beta) \\ \vdots \\ f(x_n, \beta) \end{bmatrix} \quad \text{and} \]

\[ \mathbf{D}' = \begin{bmatrix} \frac{\partial f(x_1, \beta)}{\partial b_1} & \cdots & \frac{\partial f(x_1, \beta)}{\partial b_p} \\ \vdots & \ddots & \vdots \\ \frac{\partial f(x_n, \beta)}{\partial b_1} & \cdots & \frac{\partial f(x_n, \beta)}{\partial b_p} \end{bmatrix} \]
then \( \frac{dg(b)}{db} = 0 \) is equivalent to

\[ D'[x - f(x, b)] = 0. \]

In the linear case, \( D' = X' \) and

\[ D'[x - f(x, b)] = 0 \] becomes

\[ x'[y - xb] = 0 \iff x'xb = x'y. \]
In the nonlinear case, $D'$ depends on $\beta$ so that the equation

$$D'[y - f(x, b)] = 0$$

is not so easy to solve for $b$. 
For example, for the logistic model,

\[
\frac{df(x_i, \beta)}{d\beta_1} = \frac{2}{d\beta_1} \frac{\beta_1}{1 + \exp\left\{-(x_i - \beta_2)/\beta_3\right\}}
\]

\[
= \frac{1}{1 + \exp\left\{-(x_i - \beta_2)/\beta_3\right\}}
\]

\[
\frac{df(x_i, \beta)}{d\beta_2} = \frac{-\beta_1 \exp\left\{-(x_i - \beta_2)/\beta_3\right\}}{\left[1 + \exp\left\{-(x_i - \beta_2)/\beta_3\right\}\right]^2} \beta_3
\]
\[
\frac{df(x_i, \beta)}{d\beta_3} = \frac{-\beta_1 \exp\left\{-(x_i - \beta_2)/\beta_3\right\} \left(x_i - \beta_2\right)}{\left[1 + \exp\left\{-(x_i - \beta_2)/\beta_3\right\}\right]^2 \beta_3^2}
\]

Thus, the \(i^{th}\) column of \(D'\) is

\[
\begin{pmatrix}
1 \\
1 + \exp\left\{-(x_i - \beta_2)/\beta_3\right\} \\
-\beta_1 \exp\left\{-(x_i - \beta_2)/\beta_3\right\} \\
\left[1 + \exp\left\{-(x_i - \beta_2)/\beta_3\right\}\right]^2 \beta_3 \\
-\beta_1 \exp\left\{-(x_i - \beta_2)/\beta_3\right\} \left(x_i - \beta_2\right) \\
\left[1 + \exp\left\{-(x_i - \beta_2)/\beta_3\right\}\right]^2 \beta_3^2
\end{pmatrix}
\]
The Gauss Newton Algorithm for obtaining a solution to the estimating equations

\[ D'[y - f(x, k)] = 0 \]
First note that by Taylor's Theorem, 

\[ f(x_i, b) \approx f(x_i, b^*) + \left( \frac{df(x_i, b)}{db} \right)_{b=b^*} (b - b^*) . \]

Thus, 

\[ f(x, b) \approx f(x, b^*) + D^* (b - b^*) , \]

where \( D^* \) is \( D \) evaluated at \( b = b^* \).
Now let $b^{(0)}$ denote an initial guess for a solution to

$$D'[x - f(x, b)] = 0$$

and let $b^{(r)}$ denote an approximation to a solution obtained at step $r$ of the iterative algorithm to be defined as follows.
Let $D^{(r)}$ denote $D$ evaluated at $b = b^{(r)}$.

Consider

$$\left[D^{(r-1)}\right]' \left[ y - \{ f(X, b^{(r-1)}) + D^{(r-1)} (b - b^{(r-1)}) \} \right]$$

$$= \left[D^{(r-1)}\right]' \left[ \{ y - f(X, b^{(r-1)}) \} - D^{(r-1)} (b - b^{(r-1)}) \right]$$
Equating to 0 we have

$$\left[ D^{(r-1)} \right] \begin{bmatrix} \exists y - f(x, b^{(r-1)}) \end{bmatrix} - D^{(r-1)}(b - b^{(r-1)}) = 0$$

$$\iff \left[ D^{(r-1)} \right] \begin{bmatrix} D^{(r-1)}(b - b^{(r-1)}) \end{bmatrix} = \left[ D^{(r-1)} \right] \begin{bmatrix} \exists y - f(x, b^{(r-1)}) \end{bmatrix}$$

$$\begin{bmatrix} X^* \end{bmatrix}' X^* b^* = \begin{bmatrix} x^* \end{bmatrix}' y^*$$

Solving for $b$, we have

$$b - b^{(r-1)} = \left( \left[ D^{(r-1)} \right]' D^{(r-1)} \right)^{-1} \left[ D^{(r-1)} \right]' \begin{bmatrix} \exists y - f(x, b^{(r-1)}) \end{bmatrix}$$

$$\Rightarrow b = b^{(r-1)} + \left( \left[ D^{(r-1)} \right]' D^{(r-1)} \right)^{-1} \left[ D^{(r-1)} \right]' \begin{bmatrix} \exists y - f(x, b^{(r-1)}) \end{bmatrix}$$
Thus, we define
\[ \hat{b}^{(r)} = \hat{b}^{(r-1)} + ([D^{(r-1)}]'D^{(r-1)})^{-1}[D^{(r-1)}]'\{y - f(x, \hat{b}^{(r-1)})\} \]

We continue iterating until some convergence criterion is met.

For example, we could continue until
\[ g(\hat{b}^{(r-1)}) - g(\hat{b}^{(r)}) < \text{small constant} \]
or
\[ \max_{j=1,\ldots,p} \frac{|b_j^{(r)} - b_j^{(r-1)}|}{|b_j^{(r-1)} + \varepsilon|} < \text{small constant} \]
Selection of starting value \( b^{(0)} \): We'd like to choose a value close to the overall minimizer so that convergence will be quick and so that we won't get stuck at some local minimum.

It is often a good idea to try multiple starting values.
Normal Theory Inference:

\[ Y_i = f(x_i, \beta) + \varepsilon_i \quad i = 1, \ldots, n \]

\[ \varepsilon_1, \ldots, \varepsilon_n \overset{iid}{\sim} N(0, \sigma^2) \]

Let \( \hat{\beta} \) be the least squares estimate of \( \beta \) (obtained by finding \( \lim_{n \to \infty} \hat{\beta}^{(n)} \), e.g.).

Let \( SSE = g(\hat{\beta}) = \sum_{i=1}^{n} \left[ Y_i - f(x_i, \hat{\beta}) \right]^2 \)

Let \( MSE = \frac{SSE}{n-p} \).
For $n$ sufficiently large,

$$\frac{(n-p)\text{MSE}}{\sigma^2} = \frac{\text{SSE}}{\sigma^2} \sim \chi^2_{n-p}.$$ 

$$\hat{\beta} \sim N(\beta, \sigma^2 (D'D)^{-1})$$

$\sigma^2 (D'D)^{-1}$ can be estimated by

$\text{MSE} (D'D)^{-1}$, where $\hat{D}$ is $D$ evaluated at $\hat{\beta}$. 

An approximate F-test of $H_0: C\beta = d$ rejects $H_0$ at level $\alpha$ if and only if

$$ F = \frac{\hat{\beta}'C'[C(D'D)^{-1}C']^{-1}C\hat{\beta}}{\text{MSE}} $$

where $q = \text{rank}(C) = \text{number of rows of } C.$
An approximate $100(1-\alpha)\%$ confidence interval for $\beta$ is

$$
\hat{\beta} \pm t_{n-p} \sqrt{MSE \cdot (\hat{D}'\hat{D})^{-1}}
$$
We also have approximate F tests for reduced vs. full model comparison:

\[
F = \frac{(SSE_{\text{reduced}} - SSE_{\text{full}})}{(df_{\text{reduced}} - df_{\text{full}})} \frac{SSE_{\text{full}}}{df_{\text{full}}}
\]

\[H_0: F \sim F_{df_{\text{reduced}} - df_{\text{full}}, df_{\text{full}}}
\]
For example, consider a test of

\[ H_0 : \beta_1 = \beta_{10} \ vs. \ H_A : \beta_1 \neq \beta_{10} \] for some

fixed \( \beta_{10} \).

Let \( \beta_2 = \begin{bmatrix} \beta_2 \\ \vdots \\ \beta_p \end{bmatrix} \). Let \( f_0(\mathbf{x}, \beta_2) = f(\mathbf{x}, [\beta_{10}]) \).

Then the reduced model is

\[ y_i = f_0(x_i, \beta_2) + \epsilon_i, \quad i = 1, \ldots, n \]

\[ \epsilon_1, \ldots, \epsilon_n \overset{iid}{\sim} N(0, \sigma^2) \]
Then \( F(\beta_{10}) = \frac{\text{SSE}_{\text{reduced}} - \text{SSE}_{\text{full}}}{\text{MSE}_{\text{full}}} \) \( \overset{\text{H}_0}{\sim} F_{1, n-p} \)

The set

\[ \{ \beta_{10} : F(\beta_{10}) \leq F(\alpha) \} \]

is an approximate 100(1-\alpha)% confidence set for \( \beta_1 \).

An equivalent set is

\[ \{ \beta_{10} : -t_{n-p}^{(\alpha/2)} \leq \tau(\beta_{10}) \leq t_{n-p}^{(\alpha/2)} \} \]

where

\( \tau(\beta_{10}) = \text{sign}(\beta_{10} - \hat{\beta}_1) \sqrt{F(\beta_{10})} \)
Example 2:

Researchers were interested in studying the cold tolerance of different grasses.

18 plants of genotype 1 and 18 plants of genotype 2 were assigned to 6 temperature treatments using a completely randomized design with 3 plants per combination of genotype and temperature.

Following temperature treatment, a measure of ion leakage was obtained from each plant.

High ion leakage suggests that a plant sustained high damage due to the temperature treatment while low ion leakage suggests little damage.
```r
d = read.delim("http://www.public.iastate.edu/~dnett/S511/IonLeakage.txt")

d
   geno temp   y
 1     1     4  6.97
 2     1     4  5.66
 3     1     4  6.16
 4     1    -5 19.39
 5     1    -5 15.94
 6     1    -5 13.90
 7     1     -7 29.08
 8     1     -7 25.03
 9     1     -7 37.69
10    1     -9 50.94
11    1     -9 48.67
12    1     -9 41.48
13    1    -11 54.49
14    1    -11 58.12
15    1    -11 55.48
16    1    -15 57.63
```
<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>17</td>
<td>1</td>
<td>-15</td>
<td>61.05</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>1</td>
<td>-15</td>
<td>56.59</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>2</td>
<td>4</td>
<td>9.03</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>2</td>
<td>4</td>
<td>7.36</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>2</td>
<td>4</td>
<td>6.97</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>2</td>
<td>-5</td>
<td>22.13</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>2</td>
<td>-5</td>
<td>15.76</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>2</td>
<td>-5</td>
<td>23.70</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>2</td>
<td>-7</td>
<td>37.75</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>2</td>
<td>-7</td>
<td>48.81</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>2</td>
<td>-7</td>
<td>38.09</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>2</td>
<td>-9</td>
<td>57.33</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>2</td>
<td>-9</td>
<td>52.77</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>2</td>
<td>-9</td>
<td>58.97</td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>2</td>
<td>-11</td>
<td>62.85</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>2</td>
<td>-11</td>
<td>56.36</td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>2</td>
<td>-11</td>
<td>61.62</td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>2</td>
<td>-15</td>
<td>67.19</td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>2</td>
<td>-15</td>
<td>63.55</td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>2</td>
<td>-15</td>
<td>65.51</td>
<td></td>
</tr>
</tbody>
</table>
d$geno=as.factor(d$geno)

plot(d$temp,d$y,col=4,pch=19,
     xlab="Temperature (degrees Celsius)",
     ylab="Measure of Ion Leakage")
Scientific considerations suggest the following 4-parameter Gompertz curve as a potentially appropriate function for the relationship between temperature and the measure of ion leakage.

\[ f(x, \beta) = \beta_1 + \beta_2 e^{-(x-\beta_3)/\beta_4} \]
#Fit the 4-parameter Gompertz model
#to the data. Initially, we will ignore
#the genotype factor.

o=nls(y~b1+b2*exp(-exp(-((temp-b3)/b4))),
data=d,
    start=c(b1=7,b2=55,b3=-5,b4=-2))
summary(o)

Formula: y ~ b1 + b2 * exp(-exp(-((temp - b3)/b4)))

Parameters:

|     | Estimate | Std. Error | t value | Pr(>|t|) |
|-----|----------|------------|---------|----------|
| b1  | 7.1454   | 2.0089     | 3.557   | 0.00119  ** |
| b2  | 55.8565  | 3.0501     | 18.313  | < 2e-16  *** |
| b3  | -6.0301  | 0.2331     | -25.871 | < 2e-16  *** |
| b4  | -2.0950  | 0.2936     | -7.135  | 4.26e-08  *** |

Residual standard error: 4.948 on 32 degrees of freedom

Number of iterations to convergence: 5
Achieved convergence tolerance: 1.995e-06
#By default, a numerical algorithm is used to compute the derivatives needed for the Gauss-Newton algorithm. In some cases, there is an advantage to providing the derivatives along with the function. This can be done using the deriv function.

```r
f = deriv(y ~ b1 + b2 * exp(-exp(-((x - b3) / b4))),
          c("b1", "b2", "b3", "b4"),
          function(b1, b2, b3, b4, x) { })
```
f
function (b1, b2, b3, b4, x)
{
  .expr1 <- x - b3
  .expr4 <- exp(-(.expr1/b4))
  .expr6 <- exp(-.expr4)
  .value <- b1 + b2 * .expr6
  .grad <- array(0, c(length(.value), 4L),
    list(NULL, c("b1","b2","b3","b4")))
  .grad[,"b1"] <- 1
  .grad[,"b2"] <- .expr6
  .grad[,"b3"] <- -(b2*(.expr6* (.expr4*(1/b4))))
  .grad[,"b4"] <- -b2*(.expr6* (.expr4* (.expr1/b4^2)))
attr(.value, "gradient") <- .grad
.value
}
```r
o = nls(y ~ f(b1, b2, b3, b4, temp),
       data = d,
       start = c(b1 = 6, b2 = 55, b3 = -5, b4 = -2))
summary(o)
```

**Formula:** \( y \sim f(b_1, b_2, b_3, b_4, \text{temp}) \)

**Parameters:**

| Parameter | Estimate | Std. Error | t value | Pr(>|t|)   |
|-----------|----------|------------|---------|-----------|
| b1        | 7.1454   | 2.0089     | 3.557   | 0.00119 **|
| b2        | 55.8565  | 3.0501     | 18.313  | < 2e-16 ***|
| b3        | -6.0301  | 0.2331     | -25.871 | < 2e-16 ***|
| b4        | -2.0950  | 0.2936     | -7.135  | 4.26e-08 ***|

Residual standard error: 4.948 on 32 degrees of freedom

Number of iterations to convergence: 5
Achieved convergence tolerance: 1.993e-06
# Store the least squares estimator of beta.

```r
b = coef(o)
b
```

```
  b1   b2   b3   b4
7.145431 55.856466 -6.030129 -2.094961
```

# Add the estimated mean curve to the plot.

```r
x = seq(-16, 6, by = .1)
lines(x, f(b[1], b[2], b[3], b[4], x))
```
#Examine a plot of standardized residuals versus fitted values.

plot(o)
#Get an estimate of the variance of for the squares estimator bhat: 
#MSE(Dhat'Dhat)^{-1}

v=vcov(o)
round(v,3)

<table>
<thead>
<tr>
<th></th>
<th>b1</th>
<th>b2</th>
<th>b3</th>
<th>b4</th>
</tr>
</thead>
<tbody>
<tr>
<td>b1</td>
<td>4.036</td>
<td>-4.372</td>
<td>-0.287</td>
<td>0.141</td>
</tr>
<tr>
<td>b2</td>
<td>-4.372</td>
<td>9.303</td>
<td>0.112</td>
<td>-0.590</td>
</tr>
<tr>
<td>b3</td>
<td>-0.287</td>
<td>0.112</td>
<td>0.054</td>
<td>-0.002</td>
</tr>
<tr>
<td>b4</td>
<td>0.141</td>
<td>-0.590</td>
<td>-0.002</td>
<td>0.086</td>
</tr>
</tbody>
</table>

#Obtain the error degrees of freedom.

summary(o)$df

[1]  4 32

df=summary(o)$df[2]
df

[1]  32
# Form a 95% confidence interval for $\beta_3$.

\[ b[3] - qt(.975, df) \times \sqrt{v[3,3]} \]

\[ b3 \]

-6.504898

\[ b[3] + qt(.975, df) \times \sqrt{v[3,3]} \]

\[ b3 \]

-5.555361
The following function can be used to obtain approximate 100(1-a)% confidence intervals for each element of $C \times \beta$.

ci=function(nlsout,C,a=0.05) {
  b=coef(nlsout)
  V=vcov(nlsout)
  df=summary(nlsout)$df[2]
  Cb=C times b
  se=sqrt(diag(C times V times t(C)))
  tval=qt(1-a/2,df)
  low=Cb-tval*se
  up=Cb+tval*se
  m=cbind(C,Cb,se,low,up)
  dimnames(m)[[2]]=c(paste("c",1:ncol(C),sep=""),
      "estimate","se",
      paste(100*(1-a),"% Conf.",sep=""),
      "limits")
  m
}
ci(o,matrix(c(1,-1,0,0,
0,0,1,0),byrow=T,nrow=2))

c1 c2 c3 c4  estimate    se      95% Conf.    limits
[2,] 0 0 1 0  -6.03013 0.2330801  -6.504898  -5.555361

#We can also use the profile function
#to invert the reduced versus full model F test to obtain an approximate 100(1-alpha)%
#confidence interval for any one of the #components of beta.
```r
op = profile(o, which = 3, alpha = 0.05)
b3 = as.matrix(op$b3) [, c(1, 4)]
b3

<table>
<thead>
<tr>
<th>tau</th>
<th>par.vals.b3</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2.4216397</td>
<td>-6.613587</td>
</tr>
<tr>
<td>-2.0181718</td>
<td>-6.514136</td>
</tr>
<tr>
<td>-1.6146002</td>
<td>-6.415617</td>
</tr>
<tr>
<td>-1.2108769</td>
<td>-6.317984</td>
</tr>
<tr>
<td>-0.8069124</td>
<td>-6.221170</td>
</tr>
<tr>
<td>-0.4025805</td>
<td>-6.125083</td>
</tr>
<tr>
<td>0.0000000</td>
<td>-6.030129</td>
</tr>
<tr>
<td>0.4050864</td>
<td>-5.935176</td>
</tr>
<tr>
<td>0.8143166</td>
<td>-5.839683</td>
</tr>
<tr>
<td>1.2225874</td>
<td>-5.744620</td>
</tr>
<tr>
<td>1.6296353</td>
<td>-5.649763</td>
</tr>
<tr>
<td>2.0351862</td>
<td>-5.554827</td>
</tr>
<tr>
<td>2.4389609</td>
<td>-5.459462</td>
</tr>
</tbody>
</table>
```
cbind(pt(b3[,1],df),b3)

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.01064759</td>
<td>-2.4216397</td>
<td>-6.613587</td>
</tr>
<tr>
<td>2</td>
<td>0.02601096</td>
<td>-2.0181718</td>
<td>-6.514136</td>
</tr>
<tr>
<td>3</td>
<td>0.05810807</td>
<td>-1.6146002</td>
<td>-6.415617</td>
</tr>
<tr>
<td>4</td>
<td>0.11740501</td>
<td>-1.2108769</td>
<td>-6.317984</td>
</tr>
<tr>
<td>5</td>
<td>0.21283697</td>
<td>-0.8069124</td>
<td>-6.221170</td>
</tr>
<tr>
<td>6</td>
<td>0.34496687</td>
<td>-0.4025805</td>
<td>-6.125083</td>
</tr>
<tr>
<td>7</td>
<td>0.50000000</td>
<td>0.0000000</td>
<td>-6.030129</td>
</tr>
<tr>
<td>8</td>
<td>0.65594529</td>
<td>0.4050864</td>
<td>-5.935176</td>
</tr>
<tr>
<td>9</td>
<td>0.78925876</td>
<td>0.8143166</td>
<td>-5.839683</td>
</tr>
<tr>
<td>10</td>
<td>0.88479292</td>
<td>1.2225874</td>
<td>-5.744620</td>
</tr>
<tr>
<td>11</td>
<td>0.94350769</td>
<td>1.6296353</td>
<td>-5.649763</td>
</tr>
<tr>
<td>12</td>
<td>0.97490734</td>
<td>2.0351862</td>
<td>-5.554827</td>
</tr>
<tr>
<td>13</td>
<td>0.98977153</td>
<td>2.4389609</td>
<td>-5.459462</td>
</tr>
</tbody>
</table>
Now suppose we want to estimate the temperature at which the measure of ion leakage is 50.
If we equate \( f(x, \beta) \) to \( 50 \) and solve for \( x \), we get

\[
x = -\beta_4 \log (-\log \left( \frac{50 - \beta_1}{\beta_2} \right)) + \beta_3.
\]

Thus, an estimate of the temperature of interest is

\[
\hat{x} = -\hat{\beta}_4 \log (-\log \left( \frac{50 - \hat{\beta}_1}{\hat{\beta}_2} \right)) + \hat{\beta}_3.
\]
How do we obtain a standard error for \( \hat{X} \)?

\( \hat{X} \) is a nonlinear function of \( \hat{\beta} \).

One approach is to use the Delta Method.
We need to find

\[
\frac{dx}{d\beta} = \begin{bmatrix}
\frac{dx}{d\beta_1} \\
\frac{dx}{d\beta_2} \\
\frac{dx}{d\beta_3} \\
\frac{dx}{d\beta_4}
\end{bmatrix}
= \begin{bmatrix}
\frac{\beta_4}{(50-\beta_1) \log \left( \frac{50-\beta_1}{\beta_2} \right)} \\
\frac{\beta_4}{\beta_2 \log \left( \frac{50-\beta_1}{\beta_2} \right)} \\
1 \\
-\log \left( -\log \left( \frac{50-\beta_1}{\beta_2} \right) \right)
\end{bmatrix}
\]

Let \( \hat{\beta} = \frac{dx}{d\beta} \bigg|_{\beta = \beta} \).
The variance of \( \hat{\beta} \) is then approximated by

\[
\hat{\beta}' \text{Var}(\hat{\beta}) \hat{\beta} = \text{MSE} \hat{\beta}'(D'D)^{-1} \hat{\beta}.
\]

An approximate standard error for \( \hat{\beta} \) is then

\[
\sqrt{\text{MSE} \hat{\beta}'(D'D)^{-1} \hat{\beta}} = \text{SE}(\hat{\beta}).
\]

Approximate t-based inferences would use

\[
\frac{\hat{\beta} - \beta}{\text{SE}(\hat{\beta})} \sim t_{n-p}.
\]
x50=deriv(y~-b4*log(-log((50-b1)/b2))+b3,
    c("b1","b2","b3","b4"),
    function(b1,b2,b3,b4){})

x50
function (b1, b2, b3, b4)
{
    .expr2 <- 50 - b1
    .expr3 <- .expr2/b2
    .expr4 <- log(.expr3)
    .expr6 <- log(-.expr4)
    .value <- -b4 * .expr6 + b3
    .grad <- array(0, c(length(.value), 4L), list(NULL,
        c("b1", "b2", "b3", "b4")))
    .grad[, "b1"] <- b4 * (1/b2/.expr3/.expr4)
    .grad[, "b2"] <- b4 * (.expr2/b2^2/.expr3/.expr4)
    .grad[, "b3"] <- 1
    .grad[, "b4"] <- -.expr6
    attr(.value, "gradient") <- .grad
    .value
}
x50(b[1],b[2],b[3],b[4])
  b4
-8.812504
attr("gradient")
      b1   b2   b3   b4
[1,]  0.1844918 0.1415470  1 1.328127

est=as.numeric(x50(b[1],b[2],b[3],b[4]))
est
[1] -8.812504
#Use the delta method to obtain an approximate standard error associated with the estimate.

der = attr(x50(b[1], b[2], b[3], b[4]), "gradient")
der

```
           b1       b2       b3       b4
[1,] 0.1844918 0.1415470 1 1.328127
```

se = drop(sqrt(der %*% v %*% t(der)))
se

```
[1] 0.2643512
```

#Obtain an approximate 95% confidence interval for the temperature at which the measure of ion leakage is 50.

est = qt(.975, df) * se

```
est = qt(.975, df) * se
```

```
est = qt(.975, df) * se
[1] -8.274038
```
# Now let's consider the genotype variable.

plot(d$temp, d$y, col = as.numeric(d$geno), pch = as.numeric(d$geno),
     xlab = "Temperature (degrees Celsius)", ylab = "Measure of Ion Leakage")

legend("topright",
       c("Genotype 1", "Genotype 2"), col = 1:2, pch = 1:2)
# Let's fit a model that allows a separate 4-parameter Gompertz model for each genotype.

```r
geno = nls(y ~ b1[geno] + b2[geno] * 
           exp(-exp(-((temp - b3[geno])/b4[geno]))),
           data = d,
           start = list(b1 = c(7,7), b2 = c(55,55),
                        b3 = c(-5,-5), b4 = c(-2,-2)))

summary(ogeno)
```

Formula: y ~ b1[geno] + b2[geno] * 
           exp(-exp(-((temp - b3[geno])/b4[geno])))
### Parameters:

|   | Estimate | Std. Error | t value | Pr(>|t|) |
|---|----------|------------|---------|----------|
| b11 | 6.6605   | 2.0741     | 3.211   | 0.003309 ** |
| b12 | 7.7669   | 2.0879     | 3.720   | 0.000886 *** |
| b21 | 53.9658  | 3.3329     | 16.192  | 9.50e-16 *** |
| b22 | 57.4662  | 2.9887     | 19.228  | < 2e-16 *** |
| b31 | -6.3494  | 0.2655     | -23.919 | < 2e-16 *** |
| b32 | -5.7771  | 0.2153     | -26.838 | < 2e-16 *** |
| b41 | -2.2082  | 0.3345     | -6.601  | 3.69e-07 *** |
| b42 | -1.9115  | 0.2683     | -7.123  | 9.45e-08 *** |

Residual standard error: 3.628 on 28 degrees of freedom

Number of iterations to convergence: 7
Achieved convergence tolerance: 2.47e-06

\[ b = \text{coef(ogeno)} \]
#Add the estimated mean curves to the plot.

```r
lines(x, f(b[1], b[3], b[5], b[7], x), col=1)
lines(x, f(b[2], b[4], b[6], b[8], x), col=2)
```
# Compare reduced versus full model.

```r
anova(o, ogeno)
```

Analysis of Variance Table

Model 1: y ~ f(b1, b2, b3, b4, temp)
Model 2: y ~ b1[geno] + b2[geno] * 
\[ \exp(-\exp(-((\text{temp} - b3[\text{geno}])/b4[\text{geno}]))) \]

<table>
<thead>
<tr>
<th>Res.Df</th>
<th>Res.Sum Sq</th>
<th>Df</th>
<th>Sum Sq</th>
<th>F value</th>
<th>Pr(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>32</td>
<td>783.34</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>28</td>
<td>368.63</td>
<td>4</td>
<td>414.71</td>
<td>7.8751</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.0002197</td>
</tr>
</tbody>
</table>
#Test for a lack of fit by comparing to
#a cell means model with different means
#for each combination of genotype and temperature

ocellmeans=lm(y~geno*as.factor(temp),data=d)

anova(ofull,ocellmeans)

Analysis of Variance Table

Model 1: y ~ b1[geno] + b2[geno] *
    exp(-exp(-((temp - b3[geno])/b4[geno])))

Model 2: y ~ geno * as.factor(temp)

<table>
<thead>
<tr>
<th>Res.Df</th>
<th>Res.Sum Sq</th>
<th>Df</th>
<th>Sum Sq</th>
<th>F value</th>
<th>Pr(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>368.63</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>334.51</td>
<td>4</td>
<td>34.116</td>
<td>0.6119</td>
<td>0.6581</td>
</tr>
</tbody>
</table>

69