

Nonlinear Models

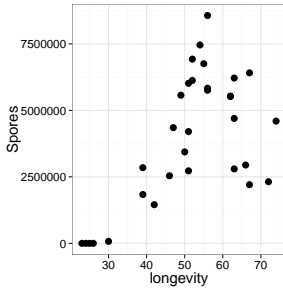
1/30

Daphnia: Purveyors of Fine Fungus



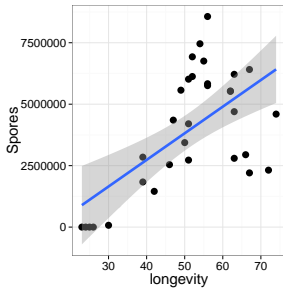
2/30

What do you do when you don't have a straight line?



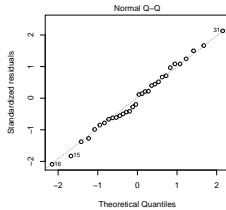
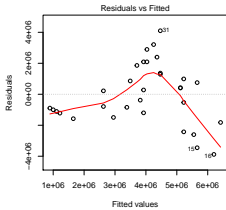
3/30

What do you do when you don't have a straight line?



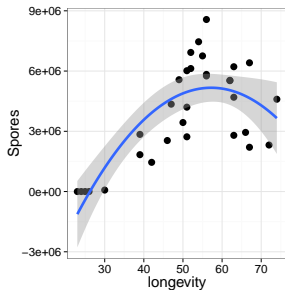
4/30

What do you do when you don't have a line?



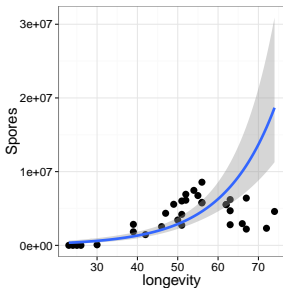
5/30

Possibility: Quadratic



6/30

Possibility: Exponential



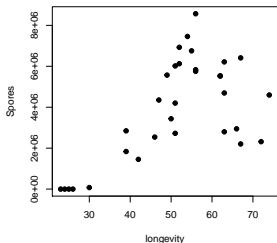
7/30

What do you do when you don't have a line?

1. If nonlinear terms are additive fit with OLS
2. Transform? But think about what it will do to error.
3. Nonlinear Least Squares
4. Generalized Linear Models

8/30

A Quadratic Adventure



$$Spores = b_0 + b_1 * Longevity + b_2 * Longevity^2 + error$$

9/30

Putting Nonlinear Terms into an Additive Model

```
fungus.lmsq <- lm(Spores ~ longevity + I(longevity^2), data=fungus)
```

10/30

Parameters are the Same as Ever

```
summary(fungus.lmsq)

#
# Call:
# lm(formula = Spores ~ longevity + I(longevity^2), data = fungus)
#
# Residuals:
#      Min       1Q   Median       3Q      Max
# -2467932 -1474558  444878  1068889  3407021
#
# Coefficients:
#              Estimate Std. Error t value Pr(>|t|)
# (Intercept)  -12432138   2883786  -4.311 0.000171
# longevity          615120    126837   4.850 3.85e-05
# I(longevity^2)   -5374         1328  -4.048 0.000351
#
# Residual standard error: 1593000 on 29 degrees of freedom
# Multiple R-squared:  0.6072, Adjusted R-squared:  0.5801
# F-statistic: 22.42 on 2 and 29 DF,  p-value: 1.303e-06
```

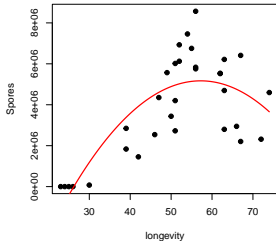
11/30

We Can't use abline Anymore

```
plot(Spores ~ longevity, data=fungus, pch=19)
#
fungusFun <- function(x) coef(fungus.lmsq)[1] +
  coef(fungus.lmsq)[2]*x +
  coef(fungus.lmsq)[3]*x^2
#
curve(fungusFun, add=T, col="red", lwd=2)
```

12/30

We Can't use abline Anymore



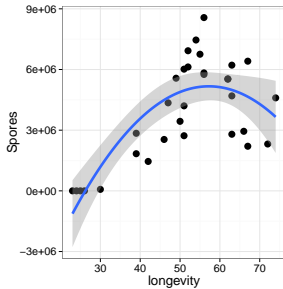
13/30

We Can't use abline Anymore

```
qplot(longevity, Spores, data=fungus, size=I(4)) +  
  theme_bw(base_size=16) +  
  stat_smooth(method="lm", size=1.5, formula = y ~ poly(x,2))
```

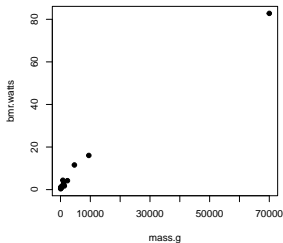
14/30

We Can't use abline Anymore



15/30

What if It's not a Linear Combination of Terms?



$$\text{MetabolicRate} = a * \text{mass}^b$$

16/30

Common Transformations

- ▶ $\log(y)$
- ▶ $\log(y)$ and $\log(x)$ (power function)
- ▶ $\arcsin(\sqrt{y})$ for bounded data
- ▶ logit for bounded data (more well behaved)
- ▶ Box-Cox Transform

May have to add 0.01, 0.5, or 1 to many of these in cases with 0s

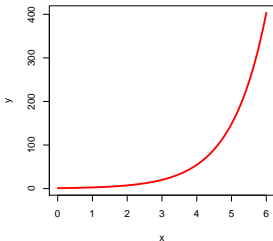
You must ask yourself, what do the transformed variables mean?

17/30

Log and Logit

$$Y = e^{a+bX}$$

$$\log(Y) = a + bX$$

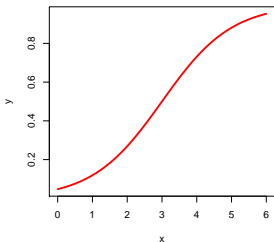


18/30

Logit Transform

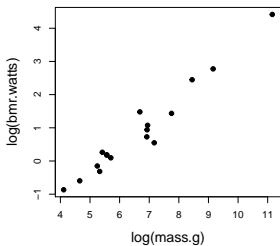
$$Y = \frac{1}{1+e^{-(a+bX)}} \text{ (bounded sigmoid curve)}$$

$$\text{logit}(Y) = a + bX$$



19/30

But Where does Error Come In?



$$\log(\text{MetabolicRate}) = \log(a) + b * \log(\text{mass}) + \text{error}$$

20/30

But Where does Error Come In?

$$\log(\text{MetabolicRate}) = \log(a) + b * \log(\text{mass}) + \text{error}$$

implies

$$\text{MetabolicRate} = a * \text{mass}^b * e^{\text{error}}$$

but we often want

$$\text{MetabolicRate} = a * \text{mass}^b + \text{error}$$

21/30

Nonlinear Least Squares Fitting

```
primate.nls <- nls(bmr.watts ~ a*mass.g^b, data=primates,  
start=list(a = 0.0172858, b = 0.74160))
```

Uses algorithm for fitting. Very flexible. Must specify start values.

22/30

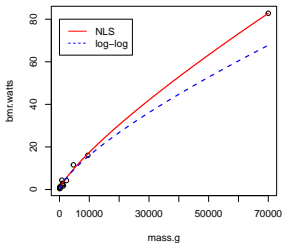
Nonlinear Least Squares Fitting

```
summary(primate.nls)

#
# Formula: bmr.watts ~ a * mass.g^b
#
# Parameters:
#   Estimate Std. Error t value Pr(>|t|)
# a 0.011063  0.002246  4.925 0.000183
# b 0.799559  0.018420  43.408 < 2e-16
#
# Residual standard error: 0.9822 on 15 degrees of freedom
#
# Number of iterations to convergence: 4
# Achieved convergence tolerance: 7.008e-07
```

23/30

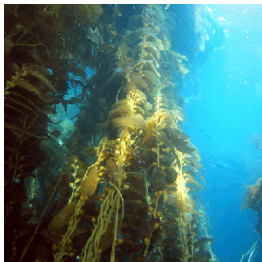
NLS Performs Better



24/30

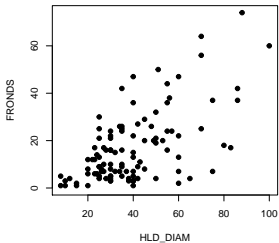
Exercise: Kelp!

- ▶ Evaluate the Frond \sim Holdfast relationship
- ▶ Fit a model with a log transformation
- ▶ Fit a model with a nls model
- ▶ Compare
- ▶ Check the diagnostics - see anything?



25/30

The Kelp Data



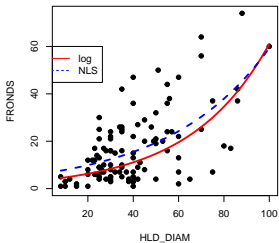
FRONDS are a count variable, cannot be < 0

26/30

Functions to Fit

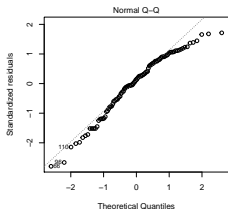
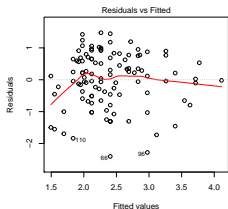
27/30

Visual Comparison of Fits



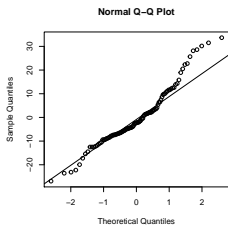
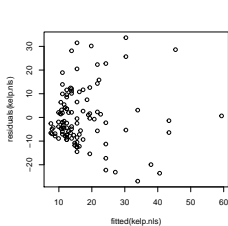
28/30

Envelope Residuals from Log Transform



29/30

Only QQ Violations even in NLS



Maybe the error is wrong...NEXT TIME!

30/30