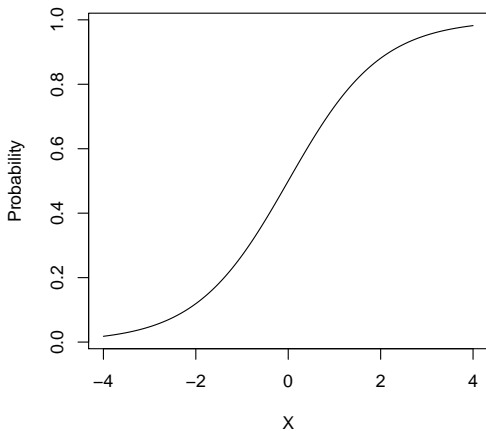


# Logistic Regression

# The Logistic Curve for Probabilities



# The Logistic Function

$$p = \frac{e^{(a+bx)}}{1 + e^{(a+bx)}}$$

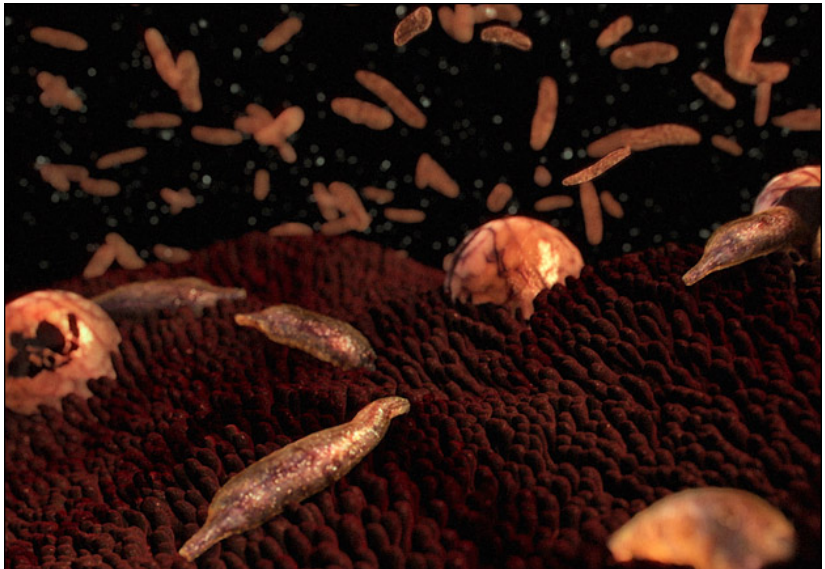
## The Log Odds

$$a + bx = \text{Log} \frac{p}{1 - p}$$

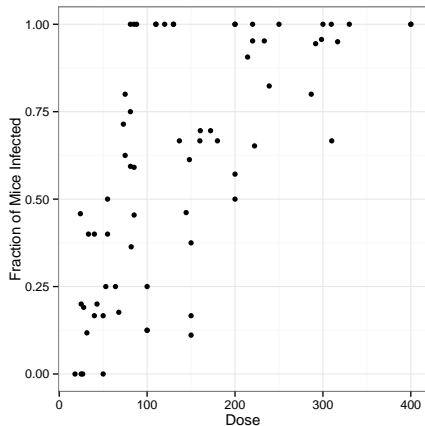
# The Odds

$$Odds = \frac{p}{1 - p}$$

# Cryptosporidium



# Fraction of Mice Infected = Probability of Infection

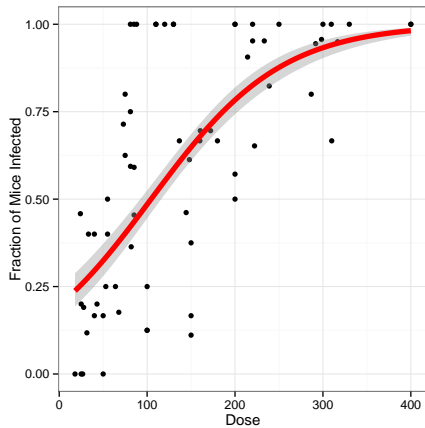


## Two Different Ways of Writing the Model

```
# 1) using Heads, Tails
glm(cbind(Y, N-Y) ~ Dose, data=crypto, family=binomial)
#
#
# 2) using weights as size parameter for Binomial
glm(Y/N ~ Dose, weights=N, data=crypto, family=binomial)
```



# The Fit Model



# The Fit Model

```
#  
# Call:  
# glm(formula = cbind(Y, N - Y) ~ Dose, family = binomial, data = crypt  
#  
# Deviance Residuals:  
#      Min        1Q    Median        3Q        Max   
# -3.953  -1.244   0.233   1.553   3.601   
#  
# Coefficients:  
#              Estimate Std. Error z value Pr(>|z|)      
# (Intercept) -1.40777    0.14848   -9.48  <2e-16      
# Dose         0.01347    0.00105   12.87  <2e-16      
#  
# (Dispersion parameter for binomial family taken to be 1)  
#  
#      Null deviance: 434.34  on 67  degrees of freedom  
# Residual deviance: 200.51  on 66  degrees of freedom  
# AIC: 327  
#
```

# The Meaning of a Logit Coefficient

Logit Coefficient: A 1 unit increase in a predictor = an increase of  $\beta$  increase in the log-odds ratio of the response.

# The Meaning of a Logit Coefficient

Logit Coefficient: A 1 unit increase in a predictor = an increase of  $\beta$  increase in the log-odds ratio of the response.

$$\text{Log Odds Ratio} = \text{Log} \left( \frac{\frac{p_1}{1-p_1}}{\frac{p_2}{1-p_2}} \right)$$

# The Meaning of a Logit Coefficient

Logit Coefficient: A 1 unit increase in a predictor = an increase of  $\beta$  increase in the log-odds ratio of the response.

$$\text{Log Odds Ratio} = \text{Log} \left( \frac{\frac{p_1}{1-p_1}}{\frac{p_2}{1-p_2}} \right)$$

We need to know both  $p_1$  and  $\beta$  to interpret this.

# The Meaning of a Logit Coefficient

$$\text{Log Odds Ratio} = \text{Log} \left( \frac{\frac{p_1}{1-p_1}}{\frac{p_2}{1-p_2}} \right)$$

If  $p_1 = 0.5$ , Odds Ratio = 1

# The Meaning of a Logit Coefficient

$$\text{Log Odds Ratio} = \text{Log} \left( \frac{\frac{p_1}{1-p_1}}{\frac{p_2}{1-p_2}} \right)$$

If  $p_1 = 0.5$ , Odds Ratio = 1

With  $\beta = 0.01347$ , we multiply 1 by  $e^{0.01347}$ .

# The Meaning of a Logit Coefficient

$$\text{Log Odds Ratio} = \text{Log} \left( \frac{\frac{p_1}{1-p_1}}{\frac{p_2}{1-p_2}} \right)$$

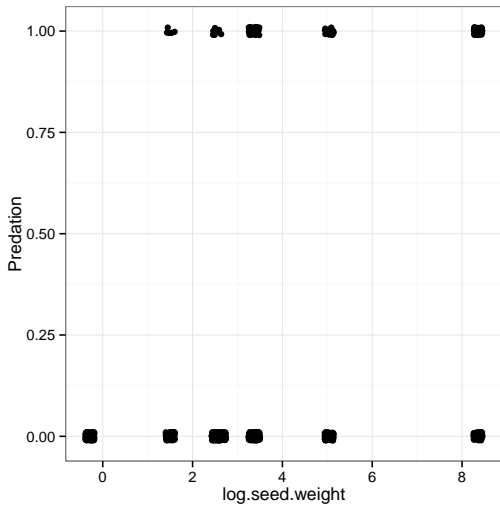
If  $p_1 = 0.5$ , Odds Ratio = 1

With  $\beta = 0.01347$ , we multiply 1 by  $e^{0.01347}$ .

The new odds ratio is 1.013561, which means  $p=0.5033674$



## What if we Only Have 1's and 0's?



# Seed Predators



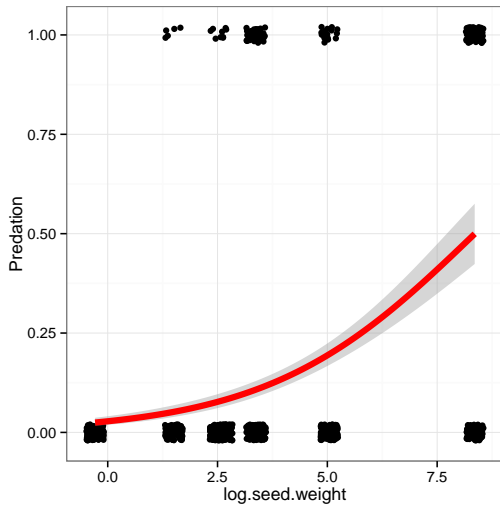
## A Quick Note on Within and Transformation

```
seeds <- within(seeds, {  
  log.seed.weight <- log(seed.weight)  
})
```

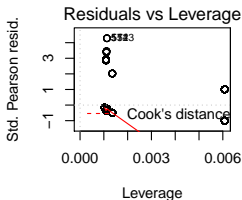
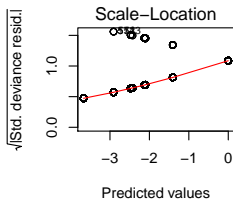
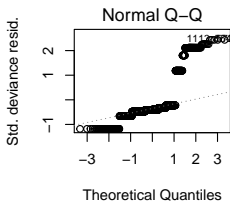
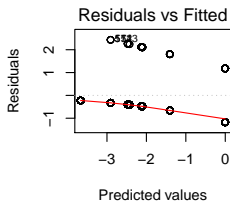
# The GLM

```
seed.glm <- glm(Predation ~ log.seed.weight,  
                data=seeds, family=binomial)
```

# Fitted Seed Predation Plot



# Diagnostics Look Odd Due to Binned Nature of the Data



## Binned Residuals Should Look Spread Out

