Interpreting Logistic Regression Models

- we want to create a spam filter based on 3921 observations/emails
- simple model, one predictor: ‘`to_multiple`‘

```
Call:
glm(formula = spam ~ to_multiple, family = binomial, data = email)

Deviance Residuals:
   Min      1Q  Median      3Q     Max
-0.477  -0.477  -0.477  -0.477   2.809

Coefficients:
               Estimate Std. Error z value Pr(>|z|)
(Intercept)    -2.11609    0.05618 -37.665  < 2e-16 ***
to_multipleyes -1.80918    0.29685  -6.095  1.1e-09 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 2437.2  on 3920  degrees of freedom
Residual deviance: 2372.0  on 3919  degrees of freedom
AIC: 2376

Number of Fisher Scoring iterations: 6
```

Interpreting Coefficients

Probability p	**Odds** p/(1-p)	**Log Odds** log[p/(1-p)]
0.1	0.1111	-2.1972
0.5	1	0
0.9	9	2.1972