

The second case study in Chapter 21 describes an experiment in which frozen dead moths of two colors were placed on trees at locations an increasing distance from Liverpool, England. This species of moth rests during the day on tree trunks and is active at night. Trees near Liverpool are darkened by smoke to a greater extent than those farther away in the Welsh countryside. At each location, the number of moths of each color that are placed and removed 24 hours later are recorded. One might expect that lighter moths are more likely to be removed near Liverpool and that darker moths are more likely to be removed farther away, as the color of the trees provides more camouflage when the color of the moth is closer.

Binomial regression is much like logistic regression except that there are multiple observations at each set of explanatory variable values. (The analysis would be identical if we used a single row for each individual observation.) To analyze this data in R, the response variable should be a two-column matrix where the first column is the count of successes (the outcome which probability we model) and the second column is the count of failures. We need to manipulate the data we read in a bit to put it into this format.

The code below uses the function `str` which provides a brief summary of the names, data types, and values of each variable in a data frame. It then uses the function `cbind` to bind together two numerical vectors as columns of a new matrix. The function `glm` with `family=binomial` is again used to fit the generalized linear model with a binomial response (and the logit link function, by default).

```
> case2102 <- read.table("sleuth/case2102.csv", header = T, sep = ",")
> attach(case2102)
> str(case2102)

'data.frame':      14 obs. of  4 variables:
 $ MORPH      : Factor w/ 2 levels "dark","light": 2 1 2 1 2 1 2 1 2 1 ...
 $ DISTANCE: num  0.0  0.0  7.2  7.2 24.1 ...
 $ PLACED   : int  56 56 80 80 52 52 60 60 60 60 ...
 $ REMOVED  : int  17 14 28 20 18 22 9 16 16 23 ...

> LEFT <- PLACED - REMOVED
> full <- glm(cbind(REMOVED, LEFT) ~ MORPH * DISTANCE, family = binomial)
> reduced <- glm(cbind(REMOVED, LEFT) ~ MORPH + DISTANCE, family = binomial)
> x2 <- deviance(reduced) - deviance(full)
> x2

[1] 11.93148

> df0 <- (summary(reduced))$df.residual - (summary(full))$df.residual
> df0

[1] 1

> p0 <- 1 - pchisq(x2, df0)
> p0

[1] 0.0005519333
```

This analysis shows strong evidence for an interaction between distance and color (MORPH) on the removal probability.

If we summarize the results, we can interpret the coefficients.

```
> summary(full)
```

Call:

```
glm(formula = cbind(REMOVED, LEFT) ~ MORPH * DISTANCE, family = binomial)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-2.21183	-0.39883	0.01155	0.68292	1.31242

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	-1.128987	0.197905	-5.705	1.17e-08	***
MORPHlight	0.411257	0.274488	1.498	0.134063	
DISTANCE	0.018502	0.005645	3.277	0.001047	**
MORPHlight:DISTANCE	-0.027789	0.008085	-3.437	0.000588	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 35.385 on 13 degrees of freedom

Residual deviance: 13.230 on 10 degrees of freedom

AIC: 83.904

Number of Fisher Scoring iterations: 3

```
> cf <- coef(full)
> slopeLight <- cf[3] + cf[4]
> slopeDark <- cf[3]
> oddsLight <- exp(slopeLight)
> oddsDark <- exp(slopeDark)
> oddsLight
```

```
DISTANCE
0.9907562
```

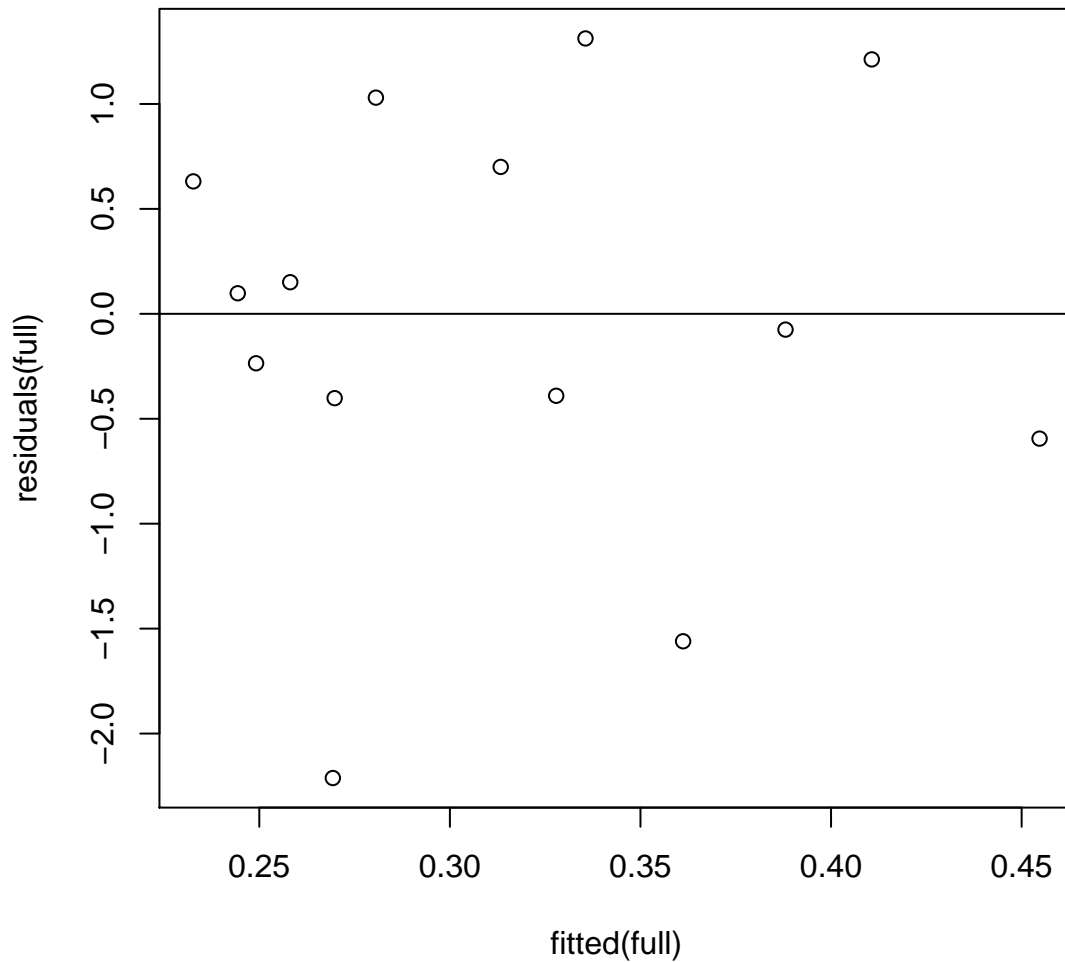
```
> oddsDark
```

```
DISTANCE
1.018674
```

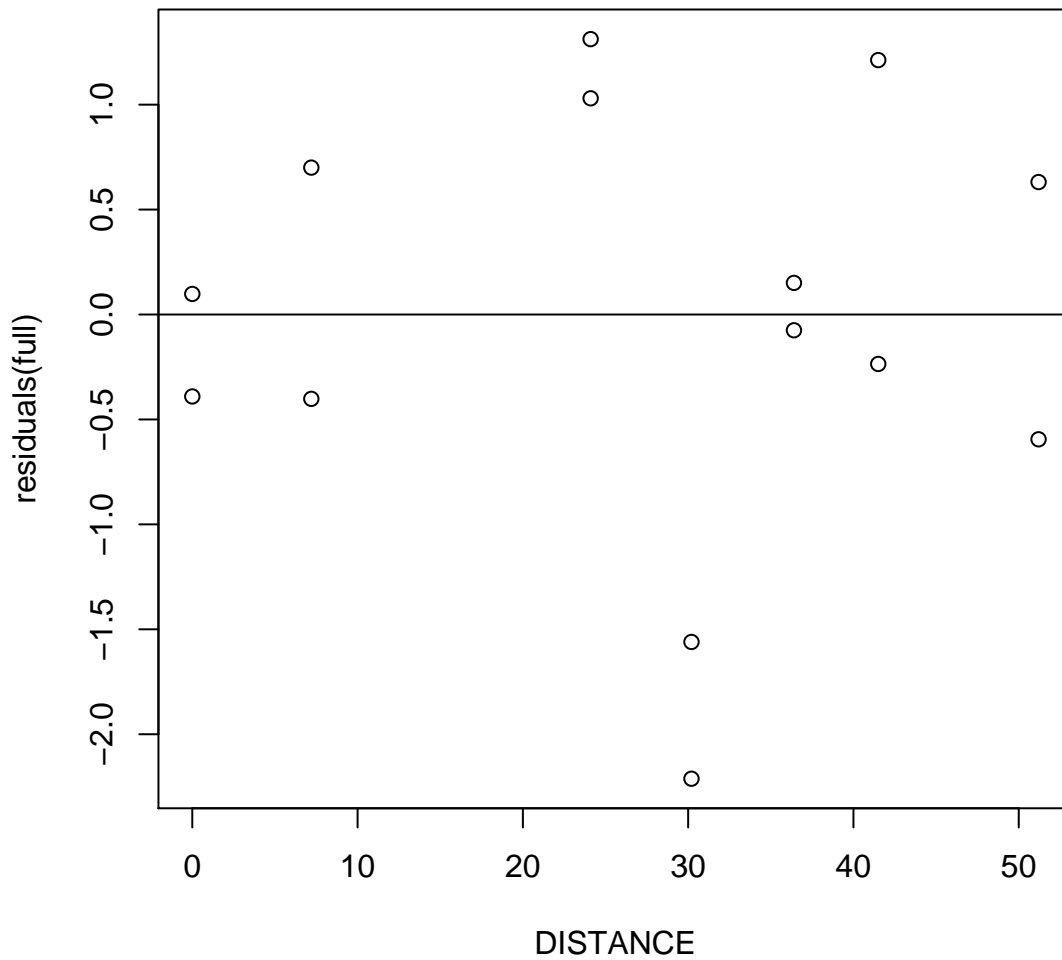
The odds of removal for light colored moths change by 0.991 per kilometer distance from Liverpool while the odds of removal for dark colored moths change by 1.019 per kilometer distance from Liverpool.

Look at a residual plot.

```
> plot(fitted(full), residuals(full))  
> abline(h = 0)
```



```
> plot(DISTANCE, residuals(full))  
> abline(h = 0)
```



```
> plot(DISTANCE, log(REMOVED/LEFT), type = "n")
> points(DISTANCE[MORPH == "light"], log(REMOVED/LEFT)[MORPH ==
+ "light"], pch = 1)
> points(DISTANCE[MORPH == "dark"], log(REMOVED/LEFT)[MORPH ==
+ "dark"], pch = 16)
```

