Discrete Response Model
Lecture 3

datascience@berkeley

Odds Ratio in the Context of Categorical Explanatory Variables

Confidence Interval

In order to compare the no control to the chemical control, the odds ratio is

$$OR = \frac{Odds_{C=0,N=1}}{Odds_{C=1,N=0}} = \frac{e^{\beta_0+\beta_3}}{e^{\beta_0+\beta_2}} = e^{\beta_3-\beta_2}$$

The estimated odds ratio is $e^{0.5152-(-0.7933)} = 3.70$

```
> K<-matrix(data = c(0, 0,  1,  0,
+                    0, 0,  0,  1),  nrow = 2, ncol = 4,
+       byrow = TRUE)
> linear.combo<-mcprofile(object = mod.fit, CM = K)
> ci.log.OR<-confint(object = linear.combo, level = 0.95, adjust = "none")
> ci.log.OR

   mcprofile - Confidence Intervals

level:          0.95
adjustment:     none

   Estimate  lower  upper
C1   -0.793 -1.054 -0.536
C2    0.515  0.258  0.773
```

```
> comparison<-c("C vs. B", "N vs. B")
> data.frame(comparison, OR = exp(ci.log.OR$confint))
  comparison  OR.lower  OR.upper
1    C vs. B 0.3486325 0.5848772
2    N vs. B 1.2945688 2.1665987
```

For example, the 95% profile LR confidence interval comparing level N to level B is 1.29 to 2.17. Thus,

With 95% confidence, the odds of plants showing symptoms are **between 1.29 and 2.17 times as large** when using no control methods rather than using a biological control (holding the infestation method constant)

Alternatively, we could also say

With 95% confidence, the odds of plants showing symptoms are between 0.46 and 0.77 times as large when using a biological control method rather than using no control methods (holding the infestation method constant). **Thus, using the spider mites (biological control) is estimated to reduce the odds of a plant showing symptoms by approximately 23% to 54%.**

Model with Interactions

The estimated model is

$$\operatorname{logit}(\hat{\pi}) = -1.0460 + 0.9258\text{Infest2} - 0.1623C + 1.1260N$$
$$-1.2114\text{Infest2} \times C - 1.1662\text{Infest2} \times N$$

To understand the effect of Control on the response, we will need to calculate odds ratios where the level of Infest2 is fixed at either 0 or 1. The odds ratio comparing level N to level B with Infest2 = 0 is

$$OR = \frac{Odds_{C=0,N=1,infest2=0}}{Odds_{C=0,N=0,infest2=0}} = \frac{e^{\beta_0+\beta_3}}{e^{\beta_0}} = e^{\beta_3}$$

1

The odds ratio comparing level N to level B with Infest2 = 1 is

$$OR = \frac{Odds_{C=0,N=1,infest2=1}}{Odds_{C=0,N=0,infest2=1}} = \frac{e^{\beta_0+\beta_1+\beta_3+\beta_5}}{e^{\beta_0+\beta_1}} = e^{\beta_3+\beta_5}$$

Model with Interactions

Other odds ratios can be calculated in a similar manner. Below are all of the estimated odds ratios and corresponding confidence intervals for Control holding Infest2 constant:

```
mcprofile - Confidence Intervals

level:              0.95
adjustment:         none

    Estimate  lower  upper
C1    1.1260  0.750  1.508
C2   -0.0402 -0.400  0.318
C3   -0.1623 -0.536  0.210
C4   -1.3738 -1.750 -1.009
C5    1.2884  0.905  1.678
C6    1.3336  0.934  1.742
```

Model with Interactions

```
data.frame(Infest2 = c(0, 1, 0, 1, 0, 1), comparison, OR
  = round(exp(ci.log.OR$estimate),2), OR.CI =
  round(exp(ci.log.OR$confint),2))
```

```
   Infest2 comparison Estimate OR.CI.lower OR.CI.upper
C1       0   N vs. B      3.08        2.12        4.52
C2       1   N vs. B      0.96        0.67        1.37
C3       0   C vs. B      0.85        0.58        1.23
C4       1   C vs. B      0.25        0.17        0.36
C5       0   N vs. C      3.63        2.47        5.36
C6       1   N vs. C      3.79        2.54        5.71
```

Model with Interactions

```
ci.logit.wald<-confint(object = save.wald, level = 0.95,
  adjust = "none")
data.frame(Infest2 = c(0, 1, 0, 1, 0, 1), comparison, OR
  = round(exp(ci.log.OR$estimate),2), lower =
  round(exp(ci.logit.wald$confint[,1]),2), upper =
  round(exp(ci.logit.wald$confint[,2]),2))
```

	Infest2	comparison	Estimate	lower	upper
C1	0	N vs. B	3.08	2.11	4.50
C2	1	N vs. B	0.96	0.67	1.38
C3	0	C vs. B	0.85	0.59	1.23
C4	1	C vs. B	0.25	0.17	0.37
C5	0	N vs. C	3.63	2.46	5.34
C6	1	N vs. C	3.79	2.53	5.68

Model with Interactions

The columns of K are ordered corresponding to the 6 parameters estimated by the model. For example, row 2 corresponds to estimating

$$OR = \frac{Odds_{C=0,N=1,infest2=1}}{Odds_{C=0,N=0,infest2=1}} = \frac{e^{\beta_0+\beta_1+\beta_3+\beta_5}}{e^{\beta_0+\beta_1}} = e^{\beta_3+\beta_5}$$

where the 4^{th} and 6^{th} columns of K have 1's for the 4^{th} and 6^{th} parameters. Remember the first parameter in the model is β_0 so this is why the column numbers are 1 higher than the indices for the β's.

The estimated odds ratio comparing level N to level B with Infest2 = 1 is $e^{1.1260 - 1.1662} = 0.96$. Thus,

> The estimated odds of plants showing symptoms are 0.96 times as large for using no control than using a biological control when infected thrips are released into the greenhouse.

We can also see why the interaction between Infest and Control was significant. The N vs. B and C vs. B odds ratio differ by a large amount over the two levels of Infest. However, there is not much of a difference for N vs. C over the levels of Infest.

Model with Interactions

The 95% profile likelihood ratio interval comparing level N to level B with Infest2 = 1 is 0.67 < OR < 1.38. Thus,

> With 95% confidence, the odds of plants showing symptoms are between 0.67 and 1.38 times as large for using no control methods than using a biological control when infected thrips are released in the greenhouse. Because 1 is within the interval, there is not sufficient evidence to conclude a biological control is effective in this setting.

Notice the interval for comparing level N to level B with Infest2 = 0 is 2.11 < OR < 4.50. Because the interval is above 1, there is sufficient evidence to conclude the biological control reduces the odds of plants showing symptoms when interspersing infected plants with uninfected thrips.

Berkeley SCHOOL OF INFORMATION