Discrete Response Model
Lecture 4

datascience@berkeley

Odds Ratio

Remarks

- The odds ratio interpretation specifically states "the odds of a category j vs. a category 1" comparison. In the past when Y was a binary response, we said something like "the odds of a success" only because it was assumed that a comparison was being made to the one other response category (failure).
- When there is more than one explanatory variable, we will need to include a statement like "holding the other variables in the model constant."
- Adjustments need to be made to an odds ratio interpretation when interactions or transformations are present in the model.
- Wald and LR-based inference methods for odds ratios are performed in the same ways as for likelihood procedures discussed in earlier weeks.

Example: Continue With the Wheat Example

Recall that we have these variables in the dataframe:

```
> summary(wheat)
 class        density             hardness             size             weight
 hrw:143   Min.   :0.7352   Min.   :-44.080   Min.   :0.5973   Min.   : 8.532
 srw:132   1st Qu.:1.1358   1st Qu.:  0.689   1st Qu.:1.8900   1st Qu.:21.982
           Median :1.2126   Median : 24.465   Median :2.2303   Median :27.610
           Mean   :1.1885   Mean   : 25.564   Mean   :2.2047   Mean   :27.501
           3rd Qu.:1.2687   3rd Qu.: 45.606   3rd Qu.:2.5125   3rd Qu.:32.882
           Max.   :1.6454   Max.   :111.934   Max.   :4.3100   Max.   :46.334
    moisture            type
 Min.   : 6.486   Healthy:96
 1st Qu.: 9.540   Scab   :83
 Median :11.909   Sprout :96
 Mean   :11.192
 3rd Qu.:12.538
 Max.   :14.514
```

Example: Continue With the Wheat Example

- Because most of the explanatory variables are "continuous," we use a value of **c** equal to 1 standard deviation.
- Ideally, it would be best to talk to the subject matter researcher about possible values for c.

```
> sd.wheat<-apply(X = wheat[,-c(1,7,8)], MARGIN = 2, FUN = sd)
> c.value<-c(1, sd.wheat)  # class = 1 is first value
> round(c.value,2)
         density hardness     size   weight moisture
    1.00    0.13    27.36     0.49     7.92     2.03
```

```
#beta.hat_jr for r = 1, ..., 6 and j = 2, 3
beta.hat2<-coefficients(mod.fit)[1,2:7]
beta.hat3<-coefficients(mod.fit)[2,2:7]
```

```
#OR for j = 2 (scab vs. healthy)
round(exp(c.value*beta.hat2), 2)
      density hardness     size   weight moisture
0.52     0.06     0.65     1.69     0.10     1.25
round(1/exp(c.value*beta.hat2), 2)
      density hardness     size   weight moisture
1.91    17.04     1.55     0.59     9.90     0.80
```

```
#OR for j = 3 (sprout vs. healthy)
round(exp(c.value*beta.hat3), 2)
      density hardness     size   weight moisture
0.80     0.14     0.56     1.54     0.69     0.92
round(1/exp(c.value*beta.hat3), 2)
      density hardness     size   weight moisture
1.25     7.28     1.78     0.65     1.45     1.09
```

Example: Continue With the Wheat Example

```
> sd.wheat<-apply(X = wheat[,-c(1,7,8)], MARGIN = 2, FUN = sd)
> c.value<-c(1, sd.wheat)  # class = 1 is first value
> round(c.value,2)
        density hardness      size  weight moisture
   1.00    0.13    27.36      0.49    7.92     2.03
```

3

```
#OR for j = 2 (scab vs. healthy)
round(exp(c.value*beta.hat2), 2)
        density hardness      size   weight moisture
0.52       0.06     0.65      1.69     0.10     1.25
round(1/exp(c.value*beta.hat2), 2)
        density hardness      size   weight moisture
1.91      17.04     1.55      0.59     9.90     0.80
```

```
#OR for j = 3 (sprout vs. healthy)
round(exp(c.value*beta.hat3), 2)
        density hardness      size   weight moisture
0.80       0.14     0.56      1.54     0.69     0.92
round(1/exp(c.value*beta.hat3), 2)
        density hardness      size   weight moisture
1.25       7.28     1.78      0.65     1.45     1.09
```

- The estimated odds of a sprout vs. a healthy response change by **7.28 times for a 0.13 decrease in the density**, holding the other variables constant.
- The estimated odds of a scab vs. healthy response change by **9.90 times for a 7.92 decrease in the weight**, holding the other variables constant.

- The estimated odds of a scab vs. a healthy response change by **0.06 times for a 0.13 increase in the density**, holding the other variables constant.
- Likewise, the estimated odds of a scab vs. a healthy response change by **17.04 times for a 0.13 decrease in the density**, holding the other variables constant.

Interpretation (cont.)

- The estimated odds of a sprout vs. healthy response change by 1.45 times for a 7.92 decrease in the weight, holding the other variables constant. Note that a Wald test of H_0: $\beta_{35} = 0$ vs. H_a: $\beta_{35} \neq 0$, which uses the parameter needed for this sprout vs. healthy odds ratio, has a p-value of 0.2, so this odds ratio may not be interpreted in actual applications.

Berkeley SCHOOL OF INFORMATION