Discrete Response Model
Lecture 4

datascience@berkeley

Odds Ratio

Odds Ratio

Odds ratios are easily formed because the proportional odds model equates log-odds to the linear predictor. The main difference now is the odds involve cumulative probabilities.

Consider the model again of

$$\text{logit}[P(Y \leq j)] = \beta_{j0} + \beta_1 x$$

The odds ratio is

$$\frac{\text{Odds}_{x+c}(Y \leq j)}{\text{Odds}_x(Y \leq j)} = \frac{e^{\beta_{j0}+\beta_1(x+c)}}{e^{\beta_{j0}+\beta_1 x}} = e^{c\beta_1}$$

where $\text{Odds}_x(Y \leq j)$ denotes the odds of observing category j or smaller for Y.

Odds Ratios: Interpretation

The formal interpretation of the odds ratio is

The odds of $Y \leq j$ vs. $Y > j$ change by $\exp(\beta_1)$ times for a c-unit increase in x

Interestingly, this odds ratio stays the same no matter what response category is used for j.

This is again due the absence of a j subscript on β_1 in the model.

Notes:

1. When there is more than one explanatory variable, we will need to include a statement like "holding the other variables in the model constant."
2. Adjustments need to be made to an odds ratio interpretation when interactions or transformations are present in the model.
3. Wald and LR-based inference methods for odds ratios are performed in the same ways as for likelihood procedures discussed in earlier weeks.

Example

The estimated odds ratios for each explanatory variable are calculated as $\exp(c\hat{\beta}_r)$ for r = 1, …, 6.

c is set to be equal to one standard deviation for each continuous explanatory variable and c = 1 for the SRW variable.

Below are the calculations (remember that $-\hat{\eta}_r$ is $\hat{\beta}_r$):

```
# Information about each variable to help with choosing c
summary(wheat)
sd.wheat<-apply(X = wheat[,-c(1,7,8)], MARGIN = 2, FUN = sd)
c.value<-c(1, sd.wheat)
round(c.value, 2)  # class = 1 is first value
```

Example

```
> head(wheat)
  class  density hardness    size  weight moisture    type type.order
1   hrw 1.349253 60.32952 2.30274 24.6480 12.01538 Healthy    Healthy
2   hrw 1.287440 56.08972 2.72573 33.2985 12.17396 Healthy    Healthy
3   hrw 1.233985 43.98743 2.51246 31.7580 11.87949 Healthy    Healthy
4   hrw 1.336534 53.81704 2.27164 32.7060 12.11407 Healthy    Healthy
5   hrw 1.259040 44.39327 2.35478 26.0700 12.06487 Healthy    Healthy
6   hrw 1.300258 48.12066 2.49132 33.2985 12.18577 Healthy    Healthy
> levels(wheat$type.order)
[1] "Scab"    "Sprout"  "Healthy"
```

```
> 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      type.order
 Min.   : 6.486   Healthy:96   Scab   :83
 1st Qu.: 9.540   Scab   :83   Sprout :96
 Median :11.909   Sprout :96   Healthy:96
 Mean   :11.192
 3rd Qu.:12.538
 Max.   :14.514
> sd.wheat<-apply(X = wheat[,-c(1,7,8)], MARGIN = 2, FUN = sd)
> c.value<-c(1, sd.wheat)
> round(c.value, 2)  # class = 1 is first value
         density hardness     size   weight moisture
    1.00    0.13    27.36     0.49     7.92     2.03
```

Example—Interpretation

```
> round(exp(c.value*(-mod.fit.ord$coefficients)),2)
        density hardness     size   weight moisture
   0.84    0.17     0.75     1.15     0.37     1.08
> round(1/exp(c.value*(-mod.fit.ord$coefficients)),2)
        density hardness     size   weight moisture
   1.19    5.89     1.33     0.87     2.74     0.92
```

Example interpretations include:

- The **estimated odds of a scab (Y ≤ 1) vs. sprout or healthy (Y > 1)** response are 0.84 times as large for soft rather than hard red winter wheat. Note that the corresponding 95% confidence interval for the class variable contains 1, as we will see shortly.
- The **estimated odds of a scab vs. sprout or healthy** response change by 5.89 times for a **0.13 decrease in the density**, holding the other variables constant.
- The **estimated odds of a scab vs. sprout or healthy** response change by 2.74 times for a **7.92 decrease in the weight**, holding the other variables constant.

Example—Interpretation

- Because of the proportional odds, each of the previous interpretations can start with "**The estimated odds of a scab or sprout vs. healthy response are ... ,**" and the same estimated odds ratios would be used in the interpretation.

- One could put the interpretation in the following form (due to the proportional odds):

The estimated odds of kernel quality being below a particular level change by ___ times for a ___ increase in ___, holding the other variables constant.

Overall, we see that the larger the density and weight, the more likely a kernel is healthy. We can again relate these results back to parallel coordinates plot to see why these interpretations make sense.

Profile Likelihood Ratio Confidence Intervals

```
> conf.beta<-confint(object = mod.fit.ord, level = 0.95)
Waiting for profiling to be done...

Re-fitting to get Hessian

> conf.beta
                2.5 %     97.5 %
classsrw -0.595305729  0.9435846
density  10.315429541 17.0363926
hardness -0.001207582  0.0221078
size     -1.103021561  0.5245184
weight    0.069318186  0.1872189
moisture -0.213254701  0.1339876
> c.value*(-conf.beta)
              2.5 %     97.5 %
classsrw  0.5953057 -0.9435846
density  -1.3544373 -2.2369136
hardness  0.0330348 -0.6047844
size      0.5411561 -0.2573352
weight   -0.5486839 -1.4819199
moisture  0.4335923 -0.2724253
> c.value[2]*(-conf.beta[2,])
    2.5 %    97.5 %
-1.354437 -2.236914
```

```
> ci<-exp(c.value*(-conf.beta))
> round(data.frame(low = ci[,2], up = ci[,1]), 2)
           low   up
classsrw  0.39 1.81
density   0.11 0.26
hardness  0.55 1.03
size      0.77 1.72
weight    0.23 0.58
moisture  0.76 1.54
> round(data.frame(low = 1/ci[,1], up = 1/ci[,2]), 2)
           low   up
classsrw  0.55 2.57
density   3.87 9.36
hardness  0.97 1.83
size      0.58 1.29
weight    1.73 4.40
moisture  0.65 1.31
```

The density odds ratio can be interpreted as: **With 95% confidence, the odds of a scab instead of a sprout or healthy response change by 3.87 to 9.36 times when density is decreased by 0.13, holding the other variables constant.**

Berkeley SCHOOL OF INFORMATION