BST 210 Lab Solution: Week 9
Multinomial and Ordinal Regression

Multinomial Regression: A Review

Logistic Regression

Let's start by taking another look at logistic regression. In logistic regression, we consider binary outcomes, where an individual either experienced an event (success, disease, death, etc.) or not. First, let's recall exactly how we specify our model:

$$\text{logit}(P(Y=1|X)) = \beta_0 + \beta_1 X_1 + \ldots \beta_p X_p$$

Equivalently,

$$P(Y=1|X) = \frac{\exp(\beta_0 + \beta_1 X_1 + \ldots \beta_p X_p)}{1 + \exp(\beta_0 + \beta_1 X_1 + \ldots \beta_p X_p)} = \pi_1$$

Since $1 = P(Y=1|X) + P(Y=0|X)$, we also get that:

$$P(Y=0|X) = 1 - \frac{\exp(\beta_0 + \beta_1 X_1 + \ldots \beta_p X_p)}{1 + \exp(\beta_0 + \beta_1 X_1 + \ldots \beta_p X_p)} = \frac{1}{1 + \exp(\beta_0 + \beta_1 X_1 + \ldots \beta_p X_p)} = \pi_0$$

We fit a logistic regression model and estimate parameters by using maximum likelihood. The likelihood can be seen as how likely we were to observe the data from our sample, given our chosen probability model and parameter values. For a binary outcome, the likelihood looks like:

$$\begin{aligned} L(\beta_0, \beta_1, \ldots, \beta_p) &= \prod_{i=1}^{n} \pi_{i1}^{I(Y_i=1)} \pi_{i0}^{I(Y_i=0)} \\ &= \prod_{i=1}^{n} \left(\frac{\exp(\beta_0 + \beta_1 X_{i1} + \ldots \beta_p X_{ip})}{1 + \exp(\beta_0 + \beta_1 X_1 + \ldots \beta_p X_p)} \right)^{y_i} \left(\frac{1}{1 + \exp(\beta_0 + \beta_1 X_{i1} + \ldots \beta_p X_{ip})} \right)^{1-y_i} \end{aligned}$$

This is solely a function of our beta coefficients and can be maximized using iterative computation techniques as our software packages do.

Multinomial Regression

Now let's consider extending what we know about logistic regression using binary outcomes to multinomial regression. Here, instead of Y only taking on two possible values (0 or 1), Y can take on c different values (from 1 up to c). Because each individual must have only one outcome, we can say:

$$\sum_{j=1}^{c} \pi_j = 1$$

where $\pi_k = P(Y = k|X)$. In multinomial regression, our model looks like (use category 1 as the baseline):

$$\log\left(\frac{\pi_k}{\pi_1}\right) = \beta_{k0} + \beta_{k1}X_1 + \ldots + \beta_{kp}X_p \qquad (k \in (2, \ldots, c))$$

Using our model and the fact that the π_k must sum to 1, we can show:

$$\pi_1 = \frac{1}{1 + \sum_{j=2}^{c} \exp(\beta_{j0} + \beta_{j1}X_1 + \ldots + \beta_{jp}X_p)}$$

(See Appendix 1 for detailed derivations.)
Once again, using our model, for $k = 2, ..., c$, we can write:

$$\begin{aligned}\pi_k &= \pi_1 \times \exp\left(\beta_{k0} + \beta_{k1}X_1 + \ldots + \beta_{kp}X_p\right) \\ &= \frac{\exp(\beta_{k0} + \beta_{k1}X_1 + \ldots + \beta_{kp}X_p)}{1 + \sum_{j=2}^{c} \exp(\beta_{j0} + \beta_{j1}X_1 + \ldots + \beta_{jp}X_p)}\end{aligned}$$

Using these probabilities, we can now extend our likelihood from logistic regression to a multinomial likelihood:

$$L(\beta_0, \beta_1, \ldots, \beta_p) = \prod_{i=1}^{n} \pi_{i1}^{I(Y_i=1)} \times \ldots \times \pi_{ic}^{I(Y_i=c)}$$

Replacing the π_ks with their formula involving βs, we can simply use iterative methods once again to find our maximum likelihood estimates.

Multinomial Example

A study examined the relative efficacy of penicillin and spectinomycin in treating gonorrhea. Three treatments were considered: (1) penicillin, (2) spectinomycin [Low Dose], and (3) spectinomycin [High Dose]. Three possible responses were recorded: (1) positive smear, (2) negative smear, positive culture, and (3) negative smear, negative culture. The results of the study are presented in the figure below:

Treatment	Response			Total
	+ Smear	- Smear + Culture	- Smear - Culture	
Penicillin	40	30	130	200
Spectinomycin, low dose	10	20	70	100
Spectinomycin, high dose	15	40	45	100
Total	65	90	245	400

Figure 1: Results from Gonorrhea Study

Using the following Stata outputs, answer the following questions:

1. Write out the full model.

$$\log\left(\frac{\pi_2}{\pi_1}\right) = \beta_{2,0} + \beta_{2,1} I(trt = LD) + \beta_{2,2} I(trt = HD)$$
$$\log\left(\frac{\pi_3}{\pi_1}\right) = \beta_{3,0} + \beta_{3,1} I(trt = LD) + \beta_{3,2} I(trt = HD)$$

Multinomial Model Fit

```
. **********************
. * BST 210 Week 9 Lab *
. **********************
. * Input data
. clear all
. use "data_lab9.dta"

. *Multinomial logistic regression
. mlogit outcome i.treatment [fweight = count], baseoutcome(1)

Iteration 0:   log likelihood = -372.45951
Iteration 1:   log likelihood = -358.97348
Iteration 2:   log likelihood = -358.51621
Iteration 3:   log likelihood = -358.51603
Iteration 4:   log likelihood = -358.51603

Multinomial logistic regression         Number of obs     =        400
                                        LR chi2(4)        =      27.89
                                        Prob > chi2       =     0.0000
Log likelihood = -358.51603             Pseudo R2         =     0.0374

------------------------------------------------------------------------------
outcome      |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
1            |  (base outcome)
-------------+----------------------------------------------------------------
2            |
treatment    |
Spect-Low    |   .9808293   .4564355     2.15   0.032     .0862322    1.875426
Spect-Hight  |   1.268511   .3872983     3.28   0.001     .5094205    2.027602
             |
_cons        |  -.2876821   .2415229    -1.19   0.234    -.7610583    .1856942
-------------+----------------------------------------------------------------
3            |
treatment    |
Spect-Low    |   .7672552   .3833771     2.00   0.045     .0158498    1.518661
Spect-Hight  |  -.0800427    .348685    -0.23   0.818    -.7634527    .6033672
             |
_cons        |   1.178655   .1808101     6.52   0.000     .8242736    1.533036
------------------------------------------------------------------------------
```

```
* We can directly get RRRs instead of beta coefficients with the 'rrr' option
. mlogit outcome i.treatment [fweight = count], baseoutcome(1) rrr

Iteration 0:   log likelihood = -372.45951
Iteration 1:   log likelihood = -358.97348
Iteration 2:   log likelihood = -358.51621
Iteration 3:   log likelihood = -358.51603
Iteration 4:   log likelihood = -358.51603

Multinomial logistic regression         Number of obs     =        400
                                        LR chi2(4)        =      27.89
                                        Prob > chi2       =     0.0000
Log likelihood = -358.51603             Pseudo R2         =     0.0374

------------------------------------------------------------------------------
outcome      |        RRR   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
1            |  (base outcome)
-------------+----------------------------------------------------------------
2            |
treatment    |
Spect-Low    |   2.666667   1.217161     2.15   0.032     1.090059      6.5236
Spect-Hight  |   3.555556   1.377061     3.28   0.001     1.664327    7.595851
             |
_cons        |        .75   .1811422    -1.19   0.234     .4671717    1.204054
-------------+----------------------------------------------------------------
3            |
treatment    |
Spect-Low    |   2.153846   .8257354     2.00   0.045     1.015976    4.566105
Spect-Hight  |   .9230769    .321863    -0.23   0.818     .4660545    1.828265
             |
_cons        |       3.25    .587633     6.52   0.000     2.280224    4.632221
------------------------------------------------------------------------------
```

2. Interpret $e^{\beta_{2,1}}$ based on our model and the outputs. (You can actually directly hand calculate this value from the data table! See Appendix 2 for details.)

   $e^{\beta_{2,1}} = e^{0.98} = 2.66$: we estimate that the risk ratio of having outcome 2 (Negative Smear/Positive Culture) to having outcome 1 (Positive Smear) for a population given Specitinomycin Low Dose is 2.667 times this risk ratio for a population given Penicillin.

3. In the above outputs, the baseline category is outcome 1 (Positive Smear). If you are interested in comparing outcome 2 and 3, what can you do? Only based on the above outputs, can you calculate the Relative Risk Ratios comparing Smear-negative/Culture-positive to Smear-negative/Culture-negative populations receiving Spectinomycin Low Dose compared to Penicillin?

$$RRR = \frac{P(Y=2|LD)}{P(Y=3|LD)} \Big/ \frac{P(Y=2|Pen)}{P(Y=3|Pen)}$$

   Now let's take the logarithm to make the math simpler and to match our model:

$$\begin{aligned}
\log(RRR) &= \log\left(\frac{P(Y=2|LD)}{P(Y=3|LD)} \Big/ \frac{P(Y=2|Pen)}{P(Y=3|Pen)}\right) \\
&= \log\left(\frac{P(Y=2|LD)}{P(Y=3|LD)}\right) - \log\left(\frac{P(Y=2|Pen)}{P(Y=3|Pen)}\right) \\
&= \log\left(\frac{P(Y=2|LD)/P(Y=1|LD)}{P(Y=3|LD)/P(Y=1|LD)}\right) - \log\left(\frac{P(Y=2|Pen)/P(Y=1|Pen)}{P(Y=3|Pen)/P(Y=1|Pen)}\right) \\
&= \log\left(\frac{P(Y=2|LD)}{P(Y=1|LD)}\right) - \log\left(\frac{P(Y=3|LD)}{P(Y=1|LD)}\right) - \log\left(\frac{P(Y=2|Pen)}{P(Y=1|Pen)}\right) + \log\left(\frac{P(Y=3|Pen)}{P(Y=1|Pen)}\right) \\
&= (\beta_{2,0} + \beta_{2,1}) - (\beta_{3,0} + \beta_{3,1}) - \beta_{2,0} - \beta_{3,0} \\
&= \beta_{2,1} - \beta_{3,1}
\end{aligned}$$

   Exponentiating both sides:

$$RRR = \exp(\beta_{2,1} - \beta_{3,1}) \approx 1.24$$

4. Using the model output, calculate the fitted probabilities of the outcome for a population having the Penicillin treatment $(P(Y=1|Pen), P(Y=2|Pen), P(Y=3|Pen))$.

$$P(Y=1|Pen) = \pi_1^{Pen} = \frac{1}{1 + \sum_{j=2}^{3} \exp(\beta_{j0} + \beta_{j1}X_1 + \beta_{j2}X_2)} = \frac{1}{1 + \exp(\beta_{2,0}) + \exp(\beta_{3,0})}$$

$$P(Y=2|Pen) = \pi_2^{Pen} = \pi_1^{Pen}\exp(\beta_{2,0} + \beta_{2,1}X_1 + \beta_{2,2}X_2) = \frac{\exp(\beta_{2,0})}{1 + \exp(\beta_{2,0}) + \exp(\beta_{3,0})}$$

$$P(Y=3|Pen) = \pi_3^{Pen} = \pi_1^{Pen}\exp(\beta_{3,0} + \beta_{3,1}X_1 + \beta_{3,2}X_2) = \frac{\exp(\beta_{3,0})}{1 + \exp(\beta_{2,0}) + \exp(\beta_{3,0})}$$

   To find the fitted probabilities, $\hat{\pi}_1^{Pen}, \hat{\pi}_2^{Pen}, \hat{\pi}_3^{Pen}$, use the estimates for the coefficients $(\hat{\beta}_{2,0}, \hat{\beta}_{3,0})$ in the above formulas.

Fitted Values

```
. predict fitm1 fitm2 fitm3
(option pr assumed; predicted probabilities)

. list outcome treatment fitm1 fitm2 fitm3

   +-----------------------------------------------+
   | outcome     treatment   fitm1   fitm2   fitm3 |
   |-----------------------------------------------|
1. |       1    Penicillin      .2     .15     .65 |
2. |       1     Spect-Low      .1      .2      .7 |
3. |       1   Spect-Hight     .15      .4     .45 |
4. |       2    Penicillin      .2     .15     .65 |
5. |       2     Spect-Low      .1      .2      .7 |
|-----------------------------------------------|
6. |       2   Spect-Hight     .15      .4     .45 |
7. |       3    Penicillin      .2     .15     .65 |
8. |       3     Spect-Low      .1      .2      .7 |
9. |       3   Spect-Hight     .15      .4     .45 |
   +-----------------------------------------------+
```

5. How do the fitted probabilities from the model compare to the observed proportions in the data table? Why is this the case?

   They match exactly. The model we fit is the saturated model. We have 3 fitted probabilities for each outcome, but each comes with the constraint that they must sum to 1. That means we only truly have 6 freely varying probabilities. We have a model with 6 coefficients, meaning we can perfectly fit those probabilities.

Some More Stata Results:

```
. *If we don't specifiy the reference group, Stata automatically uses
. *the most frequent outcome as the reference group.

. mlogit outcome i.treatment [fweight = count]

Iteration 0:   log likelihood = -372.45951
Iteration 1:   log likelihood = -358.97348
Iteration 2:   log likelihood = -358.51621
Iteration 3:   log likelihood = -358.51603
Iteration 4:   log likelihood = -358.51603

Multinomial logistic regression         Number of obs     =        400
                                        LR chi2(4)        =      27.89
                                        Prob > chi2       =     0.0000
Log likelihood = -358.51603             Pseudo R2         =     0.0374

------------------------------------------------------------------------------
outcome      |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
1            |
treatment    |
```

```
Spect-Low    |  -.7672552   .3833771    -2.00   0.045    -1.518661   -.0158498
Spect-Hight  |   .0800427    .348685     0.23   0.818    -.6033672    .7634527
             |
_cons        |  -1.178655   .1808101    -6.52   0.000    -1.533036   -.8242736
-------------+----------------------------------------------------------------
2            |
treatment    |
Spect-Low    |   .2135741   .3245171     0.66   0.510    -.4224677     .849616
Spect-Hight  |   1.348554   .2970654     4.54   0.000     .7663165    1.930792
             |
_cons        |  -1.466337   .2025479    -7.24   0.000    -1.863324   -1.069351
-------------+----------------------------------------------------------------
3            |  (base outcome)
------------------------------------------------------------------------------



. estat ic

Akaike's information criterion and Bayesian information criterion

-----------------------------------------------------------------------------
Model        |        Obs  ll(null)  ll(model)      df          AIC        BIC
-------------+---------------------------------------------------------------
.            |        400 -372.4595   -358.516       6    729.0321   752.9808
-----------------------------------------------------------------------------
Note: N=Obs used in calculating BIC; see [R] BIC note.
```

Ordinal Logistic Regression Review

For ordinal logistic regression, we are going to make the proportional odds assumption, giving us the following model specification:

$$\log\left(\frac{P(Y \geq j)}{P(Y < j)}\right) = \alpha_j + \beta_1 X_1 + \ldots + \beta_p X_p$$

With the proportional odds assumption, we assume that for each cutpoint j, we are only shifting our logistic curve by a constant, namely, the α_js. This in turn implies that the effect of a certain covariate β_1 is constant across different cutpoints in our outcome.

NOTE: Different packages will specify this model in different ways. For example, many will instead have:

$$\log\left(\frac{P(Y \leq j)}{P(Y > j)}\right) = \alpha_j + \beta_1 X_1 + \ldots + \beta_p X_p$$

Also, some packages will specify the model using negative intercepts. Please pay attention to how your package is specifying the model when it comes to interpretation.

Ordinal Logistic Regression Example

We'll be using the same data as the previous example, but instead, take into account the ordinal nature of our outcome [Outcome Smear-positive is worse than Outcome Smear-negative/Culture-positive is worse than outcome Smear-negative/Culture-negative].

Using the following Stata output, answer the questions below:

1. Write out the proportional odds ordinal logistic model. How is this different than the multinomial model? Why would you use one over the other?

$$\log\left(\frac{P(Y \geq j)}{P(Y < j)}\right) = \alpha_j + \beta_1 I(trt = LD) + \beta_2 I(trt = HD) \qquad (j \in \{2, 3\})$$

Or, writing it out fully:

$$\log\left(\frac{P(Y \geq 3)}{P(Y < 3)}\right) = \alpha_3 + \beta_1 I(trt = LD) + \beta_2 I(trt = HD)$$
$$\log\left(\frac{P(Y \geq 2)}{P(Y < 2)}\right) = \alpha_2 + \beta_1 I(trt = LD) + \beta_2 I(trt = HD)$$

Proportional odds model assumes the outcome levels are ordinal, and that the effect of a certain covariate is constant across different levels of outcomes. If the proportional odds assumption holds, fitting this model would be much more efficient than fitting the multinomial model, because proportional odds model has a much smaller number of parameters to estimate (especially when there are many covariates and outcome levels). Also, proportional odds model gives us simpler interpretations (odds ratios). However, by imposing this assumption, we do lose some information. And if the proportional odds assumption does not hold, our estimates would be invalid.

Proportional Odds

```
. *Ordinal logistic regression
. ologit outcome i.treatment [fweight = count]

Iteration 0:   log likelihood = -372.45951
Iteration 1:   log likelihood = -367.41299
Iteration 2:   log likelihood = -367.38955
Iteration 3:   log likelihood = -367.38955

Ordered logistic regression            Number of obs     =        400
                                       LR chi2(2)        =      10.14
                                       Prob > chi2       =     0.0063
Log likelihood = -367.38955            Pseudo R2         =     0.0136

------------------------------------------------------------------------------
outcome      |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
treatment    |
Spect-Low    |   .3410902    .260078     1.31   0.190    -.1686534    .8508338
Spect-Hight  |  -.5194078   .2325509    -2.23   0.026    -.9751992   -.0636163
-------------+----------------------------------------------------------------
/cut1        |   -1.73064    .175232                     -2.074088   -1.387192
/cut2        |  -.5251288   .1487098                     -.8165947   -.2336628
------------------------------------------------------------------------------
```

```
* We can directly get ORs intead of beta coefficients with the 'or' option
. ologit outcome i.treatment [fweight = count], or

Iteration 0:   log likelihood = -372.45951
Iteration 1:   log likelihood = -367.41299
Iteration 2:   log likelihood = -367.38955
Iteration 3:   log likelihood = -367.38955

Ordered logistic regression          Number of obs     =        400
LR chi2(2)        =      10.14
Prob > chi2       =     0.0063
Log likelihood = -367.38955          Pseudo R2         =     0.0136

------------------------------------------------------------------------------
outcome      | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
treatment    |
Spect-Low    |    1.40648   .3657946     1.31   0.190     .8448017    2.341598
Spect-Hight  |   .5948727   .1383382    -2.23   0.026     .3771172    .9383649
-------------+----------------------------------------------------------------
/cut1        |   -1.73064    .175232                     -2.074088   -1.387192
/cut2        |  -.5251288   .1487098                     -.8165947   -.2336628
------------------------------------------------------------------------------

. estat ic
```

```
Akaike's information criterion and Bayesian information criterion

------------------------------------------------------------------------------
Model        |        Obs  ll(null)  ll(model)      df         AIC        BIC
-------------+----------------------------------------------------------------
.            |        400 -372.4595  -367.3895       4    742.7791    758.745
------------------------------------------------------------------------------
Note: N=Obs used in calculating BIC; see [R] BIC note.
```

2. Provide brief interpretations β_1.
   $\beta_1 = 0.34$ is the log Odds Ratio for being in outcome 3 vs. combined outcome 2 and 1 for Low Dose compared to Penicillin populations. It is also the log Odds Ratio for being in combined outcome 3 and 2 vs. outcome 1 for Low Dose compared to Penicillin populations. This is because of our proportional odds assumption.

   Fitted Values

```
. predict fito1 fito2 fito3
(option pr assumed; predicted probabilities)

. list outcome treatment fito1 fito2 fito3

+-----------------------------------------------------------+
| outcome     treatment      fito1      fito2      fito3 |
|-----------------------------------------------------------|
1. |       1    Penicillin   .1505057    .221148   .6283463 |
2. |       1     Spect-Low    .111875   .1841667   .7039583 |
3. |       1   Spect-Hight   .2294831   .2690867   .5014302 |
4. |       2    Penicillin   .1505057    .221148   .6283463 |
5. |       2     Spect-Low    .111875   .1841667   .7039583 |
|-----------------------------------------------------------|
6. |       2   Spect-Hight   .2294831   .2690867   .5014302 |
7. |       3    Penicillin   .1505057    .221148   .6283463 |
8. |       3     Spect-Low    .111875   .1841667   .7039583 |
9. |       3   Spect-Hight   .2294831   .2690867   .5014302 |
+-----------------------------------------------------------+
```

3. How do the fitted probabilities from the ordinal model compare to the observed proportions in the data? Why?
   They are close to, but not the same as, the observed proportions. The model is not saturated (4 parameters), hence we don't expect a perfect fit.

4. How could you assess if the proportional odds assumption is violated? What are you conclusions in this case?
   We can use an approximate likelihood ratio test between an ordinal model with the proportional odds assumption and a generalized ordinal model without the proportional odds assumption.
   In this example, we reject the null hypothesis and conclude the proportional odds assumption is not an appropriate assumption for this data set.

Relaxing the Proportional Odds Assumption

```
. gologit2 outcome i.treatment [fweight = count], rrr

Generalized Ordered Logit Estimates     Number of obs     =        400
LR chi2(4)        =      27.89
Prob > chi2       =     0.0000
Log likelihood = -358.51603             Pseudo R2         =     0.0374

------------------------------------------------------------------------------
outcome      |        RRR   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
1            |
treatment    |
Spect-Low    |   2.250001   .8489427     2.15   0.032     1.074033    4.713544
Spect-Hight  |   1.416667   .4691744     1.05   0.293     .7402279    2.711254
|
_cons        |   3.999999   .7071064     7.84   0.000     2.828702    5.656299
-------------+----------------------------------------------------------------
2            |
treatment    |
Spect-Low    |    1.25641   .3314568     0.87   0.387     .7491586    2.107119
Spect-Hight  |   .4405594   .1100359    -3.28   0.001     .2700256    .7187932
|
_cons        |   1.857143   .2753212     4.18   0.000     1.388848    2.483338
------------------------------------------------------------------------------

. estat ic

Akaike's information criterion and Bayesian information criterion

------------------------------------------------------------------------------
Model        |        Obs  ll(null)  ll(model)      df         AIC        BIC
-------------+----------------------------------------------------------------
.            |        400 -372.4595   -358.516       6    729.0321   752.9808
------------------------------------------------------------------------------
Note: N=Obs used in calculating BIC; see [R] BIC note.
```

Testing Proportional Odds Assumption

```
. gen spectlow = (treatment == 2)

. gen specthigh = (treatment == 3)

. omodel logit outcome spectlow specthigh [fweight = count]

Iteration 0:   log likelihood = -372.45951
Iteration 1:   log likelihood = -367.41299
Iteration 2:   log likelihood = -367.38955
Iteration 3:   log likelihood = -367.38955

Ordered logit estimates                 Number of obs   =        400
```

```
LR chi2(2)      =      10.14
Prob > chi2     =     0.0063
Log likelihood = -367.38955            Pseudo R2       =     0.0136

------------------------------------------------------------------------------
outcome      |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
spectlow     |   .3410902    .260078     1.31   0.190    -.1686534    .8508338
specthigh    |  -.5194078   .2325509    -2.23   0.026    -.9751992   -.0636163
-------------+----------------------------------------------------------------
_cut1        |   -1.73064    .175232          (Ancillary parameters)
_cut2        |  -.5251288   .1487098
------------------------------------------------------------------------------

Approximate likelihood-ratio test of proportionality of odds
across response categories:
chi2(2) =     18.16
Prob > chi2 =    0.0001

. *use "ssc install omodel" if you don't have this package
```

Food for thought: can you think of other ways to assess the assumption (not necessarily a formal statistical test)? See Appendix 3 for one possible way.

5. Which model should we use for this data? Why?

   Clearly, we should not use the proportional odds ordinal model, as we rejected the test for proportional odds. Technically, the multinomial model and the generalized ordinal model (fit using gologit2) result from the same likelihood (they have the same log-likelihood, pseudo-R2, etc.). Since our outcomes do truly represent an ordinal response, it is more appropriate to use the results from the generalized ordinal model and report the separate Odds Ratios for each cut point.

Appendix

1. Derivation of π_1

Use our multinomial logistic regression model and the fact that the π_k must sum to 1, we can show:

$$\pi_1 = \frac{1}{1 + \sum_{j=2}^{c} \exp(\beta_{j0} + \beta_{j1}X_1 + \ldots + \beta_{jp}X_p)}$$

Let's denote $\boldsymbol{X\beta_k} = \beta_{k0} + \beta_{k1}X_1 + \ldots + \beta_{kp}X_p$ (note this is actually the matrix form), and our model becomes:

$$\begin{aligned} \log\left(\frac{\pi_k}{\pi_1}\right) &= \boldsymbol{X\beta_k} \\ \Leftrightarrow \frac{\pi_k}{\pi_1} &= \exp(\boldsymbol{X\beta_k}) \\ \Leftrightarrow \pi_k &= \pi_1 \exp(\boldsymbol{X\beta_k}) \end{aligned}$$

Now let's return our attention to the sum of the πs:

$$\begin{aligned} 1 &= \pi_1 + \pi_2 + \ldots + \pi_c \\ \Leftrightarrow 1 &= \pi_1 + \pi_1 \exp(\boldsymbol{X\beta_1}) + \ldots + \pi_1 \exp(\boldsymbol{X\beta_c}) \\ \Leftrightarrow 1 &= \pi_1 \left(1 + \sum_{j=2}^{c} \exp(\boldsymbol{X\beta_j})\right) \\ \Leftrightarrow \pi_1 &= \frac{1}{1 + \sum_{j=2}^{c} \exp(\boldsymbol{X\beta_j})} \end{aligned}$$

2. Hand calculation of RRR

Use the data table, compute (by hand), the relative risk ratios for having outcome 2 versus outcome 1 comparing people given Specitinomycin Low Dose to people given Penicillin, and compare them to the results found in the package output.

$$RRR = \frac{\frac{P(Y=2|LD)}{P(Y=1|LD)}}{\frac{P(Y=2|Pen)}{P(Y=1|Pen)}} = \frac{\frac{\frac{20}{100}}{\frac{10}{100}}}{\frac{\frac{30}{200}}{\frac{40}{200}}} = \frac{\frac{20}{10}}{\frac{30}{40}} = \frac{8}{3}$$

This is the same as what we got from the package.

3. Assess proportional odds assumption

We can compare two logistic regression models as follows:

```
. gen r1 = (outcome > 1)
```

```
. gen r2 = (outcome > 2)

. logistic r1 i.treatment [fweight = count]

Logistic regression                     Number of obs     =        400
                                        LR chi2(2)        =       5.31
                                        Prob > chi2       =     0.0701
Log likelihood = -174.85969             Pseudo R2         =     0.0150

------------------------------------------------------------------------------
r1           | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
treatment    |
Spect-Low    |       2.25   .8489424     2.15   0.032     1.074033    4.713542
Spect-Hight  |   1.416667   .4691742     1.05   0.293     .7402275    2.711253
             |
_cons        |          4   .7071068     7.84   0.000     2.828703    5.656302
------------------------------------------------------------------------------

. logistic r2 i.treatment [fweight = count]

Logistic regression                     Number of obs     =        400
                                        LR chi2(2)        =      15.31
                                        Prob > chi2       =     0.0005
Log likelihood = -259.38964             Pseudo R2         =     0.0287

------------------------------------------------------------------------------
r2           | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
treatment    |
Spect-Low    |    1.25641   .3314568     0.87   0.387     .7491587    2.107119
Spect-Hight  |   .4405594   .1100359    -3.28   0.001     .2700256    .7187933
             |
_cons        |   1.857143   .2753212     4.18   0.000     1.388848    2.483338
------------------------------------------------------------------------------
```

What do we notice about the coefficients in these two models? Do they tell you anything?

Notice that here we did two binomial logistic regressions rather than a single ordinal regression, and we get the same coefficients and standard errors as when we relaxed the proportional odds assumption. You can look at the package documentation for `gologit2` in Stata here for more information:

https://www.stata.com/meeting/4nasug/gologit2.pdf.