

# ORDINAL LOGISTIC REGRESSION

---

Dr. Aric LaBarr

Institute for Advanced Analytics

# INTRODUCTION

---

# Logistic Regression

- What if there are more than two categories?
  - Ordinal Logistic Regression
  - Multinomial Logistic Regression
- When the outcomes are **ordered** we can generalize the binary logistic regression model.
- Examples:
  - Disagree, Neutral, Agree
  - Tropical Depression, Tropical Storm, Category 1, 2, 3, 4, 5 Hurricanes

# Ordinal Logistic Regression

- Models are used when the response variable is ordinal.
- Models can also be used when the continuous response variable has a **restricted range** and need to be split into categories.

# Logistic Models

- Binary Logistic Regression (probability that observation  $i$  has the event):

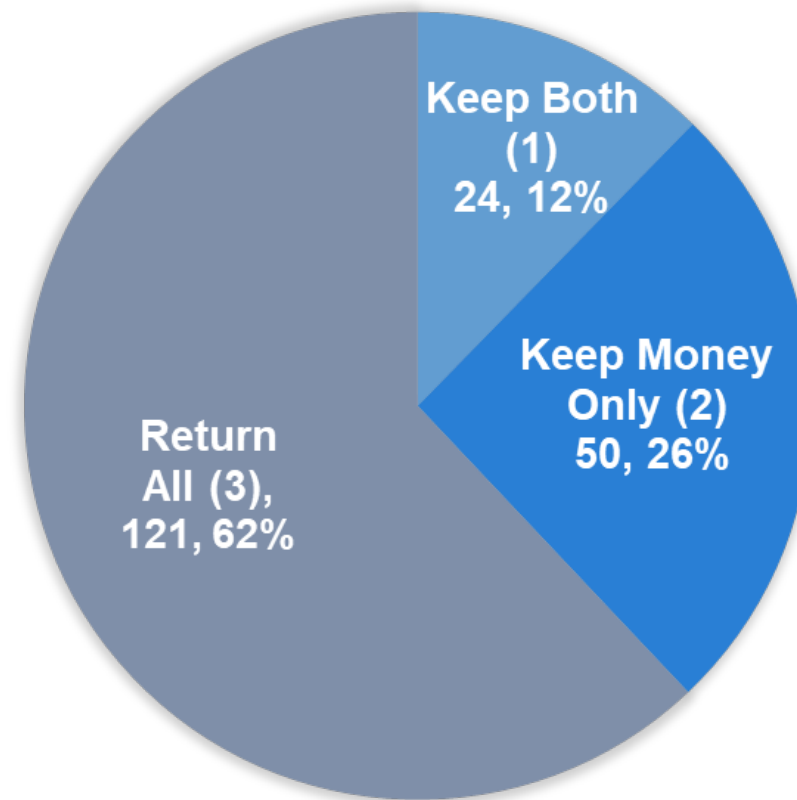
$$= \beta_0 + \beta_1 x_{1,i} + \cdots \beta_k x_{k,i}$$

- Ordinal Logistic Regression (probability that observation  $i$  has **at most** event  $j$ , and  $j = 1, \dots, m$ ):

$$= \beta_{0,j} + \beta_1 x_{1,i} + \cdots \beta_k x_{k,i}$$

# “Found a Wallet?” Data Set

- Model the association between various factors and different levels of ethical responses on finding a wallet.
- 195 observations in the data set.



# “Found a Wallet?” Data Set

- Model the association between various factors and different levels of ethical responses on finding a wallet.
- Students at UPenn.
- Predictors:
  - **male:** indicator for a male student
  - **business:** indicator for student enrolled in business school
  - **punish:** how often the student was punished as a child – low (1), moderate (2), high (3)
  - **explain:** indicator of whether explanation for punishment was given






# PROPORTIONAL ODDS MODEL

---

# Methods for Modeling

- There are three methods for modeling ordinal logistic regression models:
  1. Cumulative Logit Model
  2. Adjacent Categories Model
  3. Continuation Ratio Model

# Methods for Modeling

- There are three methods for modeling ordinal logistic regression models:
    1. Cumulative Logit Model
    2. Adjacent Categories Model
    3. Continuation Ratio Model
- 
- Easy to implement and interpret! Also, most common...

# Cumulative Logits

- Instead of modeling the typical logit, we will model the cumulative logits.
- If an ordinal variable has  $m$  levels with probabilities  $(p_1, p_2, \dots, p_m)$ , then the cumulative logits are:

$$\log\left(\frac{p_{i,1}}{p_{i,2} + p_{i,3} + \dots + p_{i,m}}\right), \log\left(\frac{p_{i,1} + p_{i,2}}{p_{i,3} + \dots + p_{i,m}}\right), \dots, \log\left(\frac{p_{i,1} + \dots + p_{i,m-1}}{p_m}\right)$$

# Cumulative Logits

- Instead of modeling the typical logit, we will model the cumulative logits.
- If an ordinal variable has  $m$  levels with probabilities  $(p_1, p_2, \dots, p_m)$ , then the cumulative logits are:

$$\log\left(\frac{p_{i,1}}{p_{i,2} + p_{i,3} + \dots + p_{i,m}}\right), \log\left(\frac{p_{i,1} + p_{i,2}}{p_{i,3} + \dots + p_{i,m}}\right), \dots, \log\left(\frac{p_{i,1} + \dots + p_{i,m-1}}{p_m}\right)$$

$$p_{i,1} + p_{i,2} + \dots + p_{i,m} = 1$$

# Cumulative Logits

- Instead of modeling the typical logit, we will model the cumulative logits.
- If an ordinal variable has  $m$  levels with probabilities  $(p_1, p_2, \dots, p_m)$ , then the cumulative logits are:

$$\log\left(\frac{p_{i,1}}{p_{i,2} + p_{i,3} + \dots + p_{i,m}}\right), \log\left(\frac{p_{i,1} + p_{i,2}}{p_{i,3} + \dots + p_{i,m}}\right), \dots, \log\left(\frac{p_{i,1} + \dots + p_{i,m-1}}{p_m}\right)$$

$$p_{i,1} + p_{i,2} + \dots + p_{i,m} = 1 \longrightarrow p_{i,2} + \dots + p_{i,m} = 1 - p_{i,1}$$

# Cumulative Logits

- Instead of modeling the typical logit, we will model the cumulative logits.
- If an ordinal variable has  $m$  levels with probabilities  $(p_1, p_2, \dots, p_m)$ , then the cumulative logits are:

$$\log\left(\frac{p_{i,1}}{1 - p_{i,1}}\right), \log\left(\frac{p_{i,1} + p_{i,2}}{p_{i,3} + \dots + p_{i,m}}\right), \dots, \log\left(\frac{p_{i,1} + \dots + p_{i,m-1}}{p_m}\right)$$

$$p_{i,1} + p_{i,2} + \dots + p_{i,m} = 1 \longrightarrow p_{i,2} + \dots + p_{i,m} = 1 - p_{i,1}$$

# Cumulative Logits

- Instead of modeling the typical logit, we will model the cumulative logits.
- If an ordinal variable has  $m$  levels with probabilities  $(p_1, p_2, \dots, p_m)$ , then the cumulative logits are:

$$\log\left(\frac{p_{i,1}}{1 - p_{i,1}}\right), \log\left(\frac{p_{i,1} + p_{i,2}}{p_{i,3} + \dots + p_{i,m}}\right), \dots, \log\left(\frac{p_{i,1} + \dots + p_{i,m-1}}{p_m}\right)$$



# Cumulative Logits

- Instead of modeling the typical logit, we will model the cumulative logits.
- If an ordinal variable has  $m$  levels with probabilities  $(p_1, p_2, \dots, p_m)$ , then the cumulative logits are:

$$\log\left(\frac{p_{i,1}}{1 - p_{i,1}}\right), \log\left(\frac{p_{i,1} + p_{i,2}}{p_{i,3} + \dots + p_{i,m}}\right), \dots, \log\left(\frac{p_{i,1} + \dots + p_{i,m-1}}{p_m}\right)$$

$$p_{i,1} + p_{i,2} + \dots + p_{i,m} = 1 \longrightarrow p_{i,3} + \dots + p_{i,m} = 1 - (p_{i,1} - p_{i,2})$$

# Cumulative Logits

- Instead of modeling the typical logit, we will model the cumulative logits.
- If an ordinal variable has  $m$  levels with probabilities  $(p_1, p_2, \dots, p_m)$ , then the cumulative logits are:

$$\log\left(\frac{p_{i,1}}{1 - p_{i,1}}\right), \log\left(\frac{p_{i,1} + p_{i,2}}{1 - (p_{i,1} + p_{i,2})}\right), \dots, \log\left(\frac{p_{i,1} + \dots + p_{i,m-1}}{p_m}\right)$$

$$p_{i,1} + p_{i,2} + \dots + p_{i,m} = 1 \longrightarrow p_{i,3} + \dots + p_{i,m} = 1 - (p_{i,1} - p_{i,2})$$

# Cumulative Logits

- Instead of modeling the typical logit, we will model the cumulative logits.
- If an ordinal variable has  $m$  levels with probabilities  $(p_1, p_2, \dots, p_m)$ , then the cumulative logits are:

$$\log\left(\frac{p_{i,1}}{1 - p_{i,1}}\right), \log\left(\frac{p_{i,1} + p_{i,2}}{1 - (p_{i,1} + p_{i,2})}\right), \dots, \log\left(\frac{p_{i,1} + \dots + p_{i,m-1}}{1 - (p_{i,1} + \dots + p_{i,m-1})}\right)$$

# Cumulative Logits

- Instead of modeling the typical logit, we will model the cumulative logits.
- If an ordinal variable has  $m$  levels with probabilities  $(p_1, p_2, \dots, p_m)$ , then the cumulative logits are:

$$\log\left(\frac{p_{i,1}}{1 - p_{i,1}}\right), \log\left(\frac{p_{i,1} + p_{i,2}}{1 - (p_{i,1} + p_{i,2})}\right), \dots, \log\left(\frac{p_{i,1} + \dots + p_{i,m-1}}{1 - (p_{i,1} + \dots + p_{i,m-1})}\right)$$

## ***m-1* Binary Logistic Regressions!**

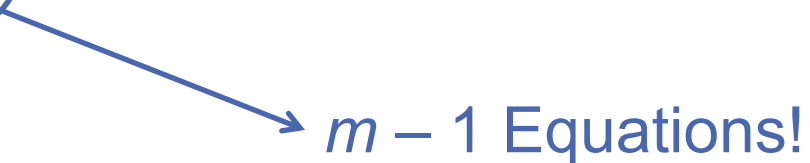
- Event now becomes outcome  $\leq j$  for categories  $j = 1, \dots, m$

# Logistic Models

- Binary Logistic Regression (probability that observation  $i$  has the event):

$$= \beta_0 + \beta_1 x_{1,i} + \cdots \beta_k x_{k,i}$$

- Ordinal Logistic Regression (probability that observation  $i$  has **at most** event  $m$ , and  $j = 1, \dots, m$ ):

$$= \beta_{0,j} + \beta_1 x_{1,i} + \cdots \beta_k x_{k,i}$$


$m - 1$  Equations!

# Logistic Models

- Binary Logistic Regression (probability that observation  $i$  has the event):

$$= \beta_0 + \beta_1 x_{1,i} + \cdots \beta_k x_{k,i}$$

- Ordinal Logistic Regression (probability that observation  $i$  has **at most** event  $m$ , and  $j = 1, \dots, m$ ):

$$= \beta_{0,j} + \beta_1 x_{1,i} + \cdots \beta_k x_{k,i}$$

- Intercept changes, but **slope parameters stays the same** (called **proportional odds** assumption)!

# “Found a Wallet?” Data Set

- Model the association between various factors and different levels of ethical responses on finding a wallet.

$$\log \left( \frac{p_{i,1}}{p_{i,2} + p_{i,3}} \right) = \beta_{0,1} + \beta_1 \text{male}_i + \beta_2 \text{business}_i \\ + \beta_3 \text{punishM}_i + \beta_4 \text{punishH}_i + \beta_5 \text{explain}_i$$

$$\log \left( \frac{p_{i,1} + p_{i,2}}{p_{i,3}} \right) = \beta_{0,2} + \beta_1 \text{male}_i + \beta_2 \text{business}_i \\ + \beta_3 \text{punishM}_i + \beta_4 \text{punishH}_i + \beta_5 \text{explain}_i$$

# Proportional Odds Model

```
train <- wallet
```

```
train$punish <- factor(train$punish)
```

```
clogit.model <- polr(factor(wallet) ~ male + business + punish + explain,  
                     method = "logistic", data = train)
```

```
summary(clogit.model)
```



# Proportional Odds Model

Call:

```
polr(formula = factor(wallet) ~ male + business + punish + explain,  
      data = train, method = "logistic")
```

Coefficients:

	Value	Std. Error	t value
male	-1.0598	0.3274	-3.237
business	-0.7389	0.3556	-2.078
punish2	-0.6276	0.4048	-1.551
punish3	-1.4031	0.4823	-2.909
explain	1.0519	0.3408	3.086

Intercepts:

	Value	Std. Error	t value
1 2	-2.5679	0.4190	-6.1287
2 3	-0.7890	0.3709	-2.1273

Residual Deviance: 307.3349

AIC: 321.3349

# Proportional Odds Model

Call:

```
polr(formula = factor(wallet) ~ male + business + punish + explain,  
      data = train, method = "logistic")
```

Coefficients:

	Value	Std. Error	t value
male	-1.0598	0.3274	-3.237
business	-0.7389	0.3556	-2.078
punish2	-0.6276	0.4048	-1.551
punish3	-1.4031	0.4823	-2.909
explain	1.0519	0.3408	3.086

Intercepts:

	Value	Std. Error	t value
1 2	-2.5679	0.4190	-6.1287
2 3	-0.7890	0.3709	-2.1273

Residual Deviance: 307.3349

AIC: 321.3349

# Proportional Odds Model

Call:

```
polr(formula = factor(wallet) ~ male + business + punish + explain,  
      data = train, method = "logistic")
```

Coefficients:

	Value	Std. Error	t value
male	-1.0598	0.3274	-3.237
business	-0.7389	0.3556	-2.078
punish2	-0.6276	0.4048	-1.551
punish3	-1.4031	0.4823	-2.909
explain	1.0519	0.3408	3.086

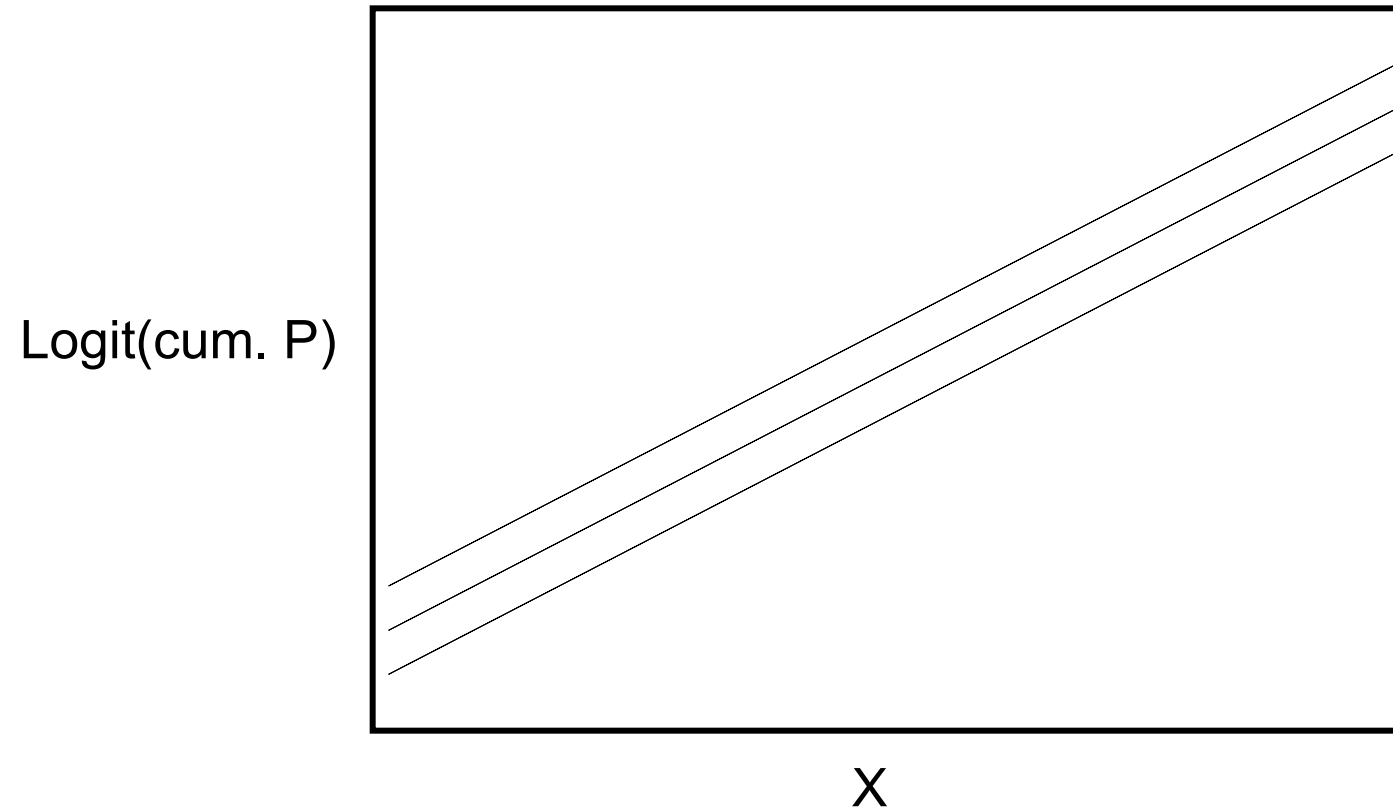
Intercepts:

	Value	Std. Error	t value
1 2	-2.5679	0.4190	-6.1287
2 3	-0.7890	0.3709	-2.1273

Residual Deviance: 307.3349

AIC: 321.3349

# Testing Assumptions



HOW DO WE TEST IF SLOPES ARE THE SAME?

# Score Test (Brant Test) for Proportional Odds

- Need to test to see if the slopes are statistically different from each other in the proportional odds model.
  - Null: Proportional Odds Correct (Slopes Equal Across Models)
  - Alternative: Proportional Odds Incorrect (Slopes NOT Equal Across Models)

# Brant Test

```
brant(clogit.model)
```

-----			
Test for	X2	df	probability
-----			
Omnibus	5.46	5	0.36
male	0.51	1	0.47
business	0.58	1	0.45
punish2	0.99	1	0.32
punish3	2.81	1	0.09
explain	0.25	1	0.62
-----			

H0: Parallel Regression Assumption holds

# What if Assumption Fails?

- The proportional odds assumption may not be met for all variables.
- 2 Approaches:
  1. Partial Proportional Odds Model
  2. Multinomial Logistic Regression

# What if Assumption Fails?

- The proportional odds assumption may not be met for all variables.
- 2 Approaches:

1. Partial Proportional Odds Model
2. Multinomial Logistic Regression



**Some** variables  
fail assumption

**All** variables  
fail assumption



# Partial Proportional Odds

```
plogit.model <- vglm(factor(wallet) ~ male + business + punish + explain,  
                     data = train,  
                     family = cumulative(parallel = F ~ business))  
  
summary(plogit.model)
```

# Partial Proportional Odds

Call:

```
vglm(formula = factor(wallet) ~ male + business + punish + explain,
      family = cumulative(parallel = F ~ business), data = train)
```

Coefficients:

	Estimate	Std. Error	z value	Pr(> z )	
(Intercept):1	-2.6695	0.4466	-5.978	2.26e-09	***
(Intercept):2	-0.7730	0.3678	-2.102	0.03557	*
male	1.0707	0.3258	3.287	0.00101	**
business:1	0.9722	0.4789	2.030	0.04236	*
business:2	0.6376	0.3810	1.674	0.09423	.
punish2	0.6300	0.4008	1.572	0.11594	
punish3	1.3956	0.4727	2.952	0.00316	**
explain	-1.0532	0.3413	-3.086	0.00203	**
---					

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



# INTERPRETATION

---

# Model Notation


- With cumulative logits, increasing the right-hand side of the equation leads to an increased log(odds) of **higher** outcome category:


$$\log \left( \frac{p_{i,3}}{p_{i,1} + p_{i,2}} \right) = \beta_{0,1} + \beta_1 \text{male}_i + \beta_2 \text{business}_i \\ + \beta_3 \text{punishM}_i + \beta_4 \text{punishH}_i + \beta_5 \text{explain}_i$$

$$\log \left( \frac{p_{i,3} + p_{i,2}}{p_{i,1}} \right) = \beta_{0,2} + \beta_1 \text{male}_i + \beta_2 \text{business}_i \\ + \beta_3 \text{punishM}_i + \beta_4 \text{punishH}_i + \beta_5 \text{explain}_i$$

# Model Notation

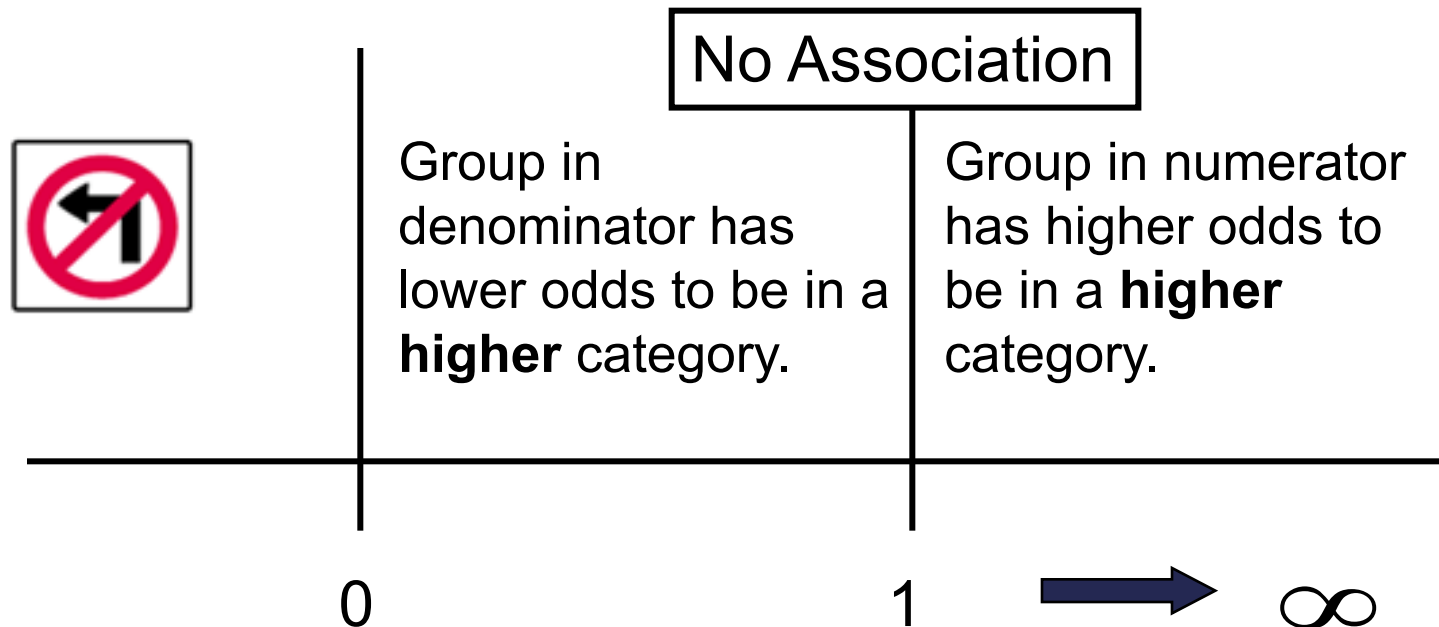
- With cumulative logits, increasing the right-hand side of the equation leads to an increased log(odds) of **higher** outcome category:

$$\log \left( \frac{p_{i,3}}{p_{i,1} + p_{i,2}} \right) = \beta_{0,1} + \beta_1 \text{male}_i + \beta_2 \text{business}_i + \beta_3 \text{punishM}_i + \beta_4 \text{punishH}_i + \beta_5 \text{explain}_i$$


$$\log \left( \frac{p_{i,3} + p_{i,2}}{p_{i,1}} \right) = \beta_{0,2} + \beta_1 \text{male}_i + \beta_2 \text{business}_i + \beta_3 \text{punishM}_i + \beta_4 \text{punishH}_i + \beta_5 \text{explain}_i$$


# Odds Ratio Interpretation – Descending

- Interpretation is still an odds ratio:  $100 * (e^{\hat{\beta}_j} - 1) \%$  **higher expected odds** of being in a higher category.



# Odds Ratio Interpretation – Descending

- Interpretation is still an odds ratio:  $100 * (e^{\hat{\beta}_j} - 1) \%$  **higher expected odds** of being in a higher category.
- Proportional odds model:
  - Same increase in odds across all singular jumps in category.
  - Wallet example: OR same comparing 3 to 1,2 and from 3,2 to 1.



# Odds Ratio Interpretation – Descending

- Interpretation is still an odds ratio:  $100 * (e^{\hat{\beta}_j} - 1) \%$  **higher expected odds** of being in a higher category.
- Proportional odds model:
  - Same increase in odds across all singular jumps in category.
  - Wallet example: OR same comparing 3 to 1,2 and from 3,2 to 1.
- Male variable (coefficient =  $-1.0598$ ,  $100 * (e^{-1.0598} - 1) = -65.35\%$  )
  - Males have **65.35% lower expected odds** of being in a **higher** ethical category as compared to females.

# Odds Ratio Interpretation – Descending

- Interpretation is still an odds ratio:  $100 * (e^{\hat{\beta}_j} - 1) \%$  **higher expected odds** of being in a higher category.
- Proportional odds model:
  - Same increase in odds across all singular jumps in category.
  - Wallet example: OR same comparing 3 to 1,2 and from 3,2 to 1.
- Business variable (coefficient =  $-0.7389$ ,  $100 * (e^{-0.7389} - 1) = -52.24 \%$  )
  - Business school students have **52.24% lower expected odds** of being in a **higher** ethical category as compared to students not in the business school.



# PREDICTIONS AND DIAGNOSTICS

---

# Similarities

- Ordinal logistic regression has a lot of the same aspects/issues as a binary logistic regression:
  - Multicollinearity still exists.
  - Non-convergence problems still exist.
  - Confidence intervals need profile likelihoods.
  - Concordance, Discordance, Tied pairs still exist – so the c statistic still exists.
  - Generalized  $R^2$  remains the same.

# Differences

- Ordinal logistic regression has a few aspects/issues that differ from a binary logistic regression:
  - A lot of the diagnostics for binary regression cannot be calculated easily since there are actually **multiple** models – ROC curves for each model?
  - Diagnostics / Influence plots are not available – residuals for each model?
  - Predicted probabilities are for **each** category.

# Predicted Probabilities

```
pred_probs <- predict(clogit.model, newdata = train, type = "probs")
```

```
head(pred_probs)
```

	1	2	3
1	0.12562481	0.3341195	0.5402557
2	0.04778463	0.1813420	0.7708734
3	0.02609095	0.1108549	0.8630542
4	0.12562481	0.3341195	0.5402557
5	0.07176375	0.2423258	0.6859105
6	0.02609095	0.1108549	0.8630542

# Confusion Matrix

- A confusion matrix is a matrix of all predicted responses compared to actual responses in terms of correct percentage.

	<b>Predicted</b>		
<b>Actual</b>	4	11	9
	3	9	38
	0	12	109



# Confusion Matrix

- A confusion matrix is a matrix of all predicted responses compared to actual responses in terms of correct percentage.

	Predicted		
Actual	16.7%	45.8%	37.5%
	6.0%	18.0%	76.0%
	0.0%	9.9%	90.1%

# Good Confusion Matrix

	Predicted		
Actual	1	2	3
	100%	0.0%	0.0%
	0.0%	100%	0.0%
	0.0%	0.0%	100%

# Confusion Matrix

- Weighted accuracy scores are common.

	Predicted		
Actual	1	0.5	0
	0.5	1	0.5
	0	0.5	1

# Notch Graph

- Some people use notch graphs to show the “accuracy” gains the further out the prediction is from the truth.

