

# Logistic Regression for Ordinal Responses

Edps/Psych/Soc 589

Carolyn J. Anderson

Department of Educational Psychology



©Board of Trustees, University of Illinois

Fall 2018

# I Outline

## Common models for ordinal responses:

- ▶ Cumulative logit model typically assuming “proportional odds”.
- ▶ Adjacent categories logit model typically assuming common slopes
- ▶ Continuation ratio logits.
- ▶ Baseline multinomial logistic regression but use the order to interpret and report odds ratios.

## They differ in terms of

- ▶ How logits are formed.
- ▶ Whether they summarize association with 1 parameter per predictor.
- ▶ Whether they allow for different models for different logits.

# I Logit Models for Ordinal Responses

The logit models for this situation

- ▶ Use the ordering of the categories in forming logits.
- ▶ Yield simpler models with simpler interpretations than (baseline) multinomial model.
- ▶ Are more powerful than nominal models.

# I Proportional Odds Model

## or Cumulative Logit Model

Form logits (dichotomize categories of  $Y$ ) incorporating the ordinal information.

Cumulative Probabilities:

- ▶  $Y = 1, 2, \dots, J$  and order is relevant.
- ▶  $\{\pi_1, \pi_2, \dots, \pi_J\}$ .
- ▶  $P(Y \leq j) = \pi_1 + \dots + \pi_j = \sum_{k=1}^j \pi_k$  for  $j = 1, \dots, J - 1$ .
- ▶ “Cumulative logits”

$$\begin{aligned} \log \left( \frac{P(Y \leq j)}{P(Y > j)} \right) &= \log \left( \frac{P(Y \leq j)}{1 - P(Y \leq j)} \right) \\ &= \log \left( \frac{\pi_1 + \dots + \pi_j}{\pi_{j+1} + \dots + \pi_J} \right) \quad \text{for } j = 1, \dots, J - 1 \end{aligned}$$

# I “Proportional Odds Model”

$$\text{logit}(P(Y \leq j)) = \log\left(\frac{P(Y \leq j)}{P(Y > j)}\right) = \alpha_j + \beta x \quad \text{for } j = 1, \dots, J-1$$

- ▶  $\alpha_j$  (intercepts) can differ.
- ▶  $\beta$  (slope) is constant.
  - ▶ The effect of  $x$  is the same for all  $J - 1$  ways to collapse  $Y$  into dichotomous outcomes (cumulatively).
  - ▶ A single parameter describes the effect of  $x$  on  $Y$  (versus  $J - 1$  slopes in the baseline model).
- ▶ Interpretation in terms of odds ratios.

# I Interpretation

For a given level of  $Y$  (say  $Y = j$ )

$$\begin{aligned} \frac{P(Y \leq j|X = x_2)/P(Y > j|X = x_2)}{P(Y \leq j|X = x_1)/P(Y > J|X = x_1)} &= \frac{P(Y \leq j|x_2)P(Y > j|x_1)}{P(Y \leq j|x_1)P(Y > j|x_2)} \\ &= \exp(\alpha_j + \beta x_2) / \exp(\alpha_j + \beta x_1) \\ &= \exp[\beta(x_2 - x_1)] \end{aligned}$$

or log odds ratio =  $\beta(x_2 - x_1)$ .

The log cumulative odds ratio is proportional to the difference (distance) between  $x_1$  and  $x_2$ .

Since the proportionality coefficient  $\beta$  is constant, this model is called the “**Proportional Odds Model**”.

# I Properties of Model

- ▶ Note that the cumulative probabilities are given by

$$P(Y \leq j) = \frac{\exp(\alpha_j + \beta x)}{1 + \exp(\alpha_j + \beta x)}$$

Since  $\beta$  is constant, curves of cumulative probabilities plotted against  $x$  are parallel.

- ▶ We can compute the probability of being in category  $j$  by taking differences between the cumulative probabilities.

$$P(Y = j) = P(Y \leq j) - P(Y \leq j - 1) \quad \text{for } j = 2, \dots, J$$

and

$$P(Y = 1) = P(Y \leq 1)$$

Since  $\beta$  is constant, these probabilities are guaranteed to be non-negative.

- ▶ In fitting this model to data, it must be simultaneous.

## I Example: HSB

Example: High School and Beyond

$X$  = mean of 5 achievement test scores.

$$\begin{aligned} Y &= \text{high school program type} \\ &= \begin{cases} 1 & \text{Academic} \\ 2 & \text{General} \\ 3 & \text{VoTech} \end{cases} \end{aligned}$$

So the logit model is

$$\text{Academic vs (Gen \& VoTech):} \quad \text{logit}(Y \leq 1) = \alpha_1 + \beta x$$

$$\text{(Academic \& Gen) vs VoTech:} \quad \text{logit}(Y \leq 2) = \alpha_2 + \beta x$$



# I Test of Proportional Odds Assumption

Score Test for the Proportional  
Odds Assumption

Chi-Square	DF	Pr > ChiSq
0.8194	1	0.3653

If this test is significant, then proportional odds model is not good one for the data. (Later we'll talk about what to do if it's significant.)

I will show R a bit later.

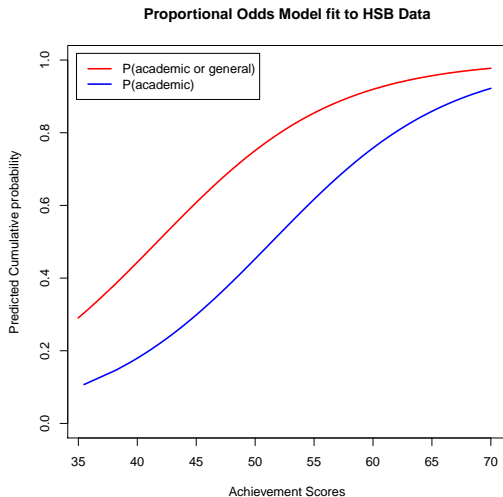
## I Example: Parameter Estimates

Parameter	Estimate	$e^{\beta}$	ASE	Wald	$p$
$\alpha_1$	-6.8408		.6118	125.04	< .001
$\alpha_2$	-5.5138		.5866	88.37	< .001
$\beta$	.1330	1.142	.0118	127.64	< .001

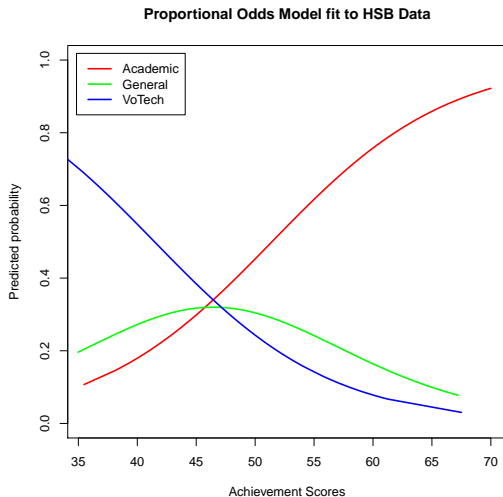
For a 10 point increase in mean achievement, the odds ratio (for either case) equals

$$\exp(10(.1330)) = 3.78$$

# I Fitted Cumulative Probabilities

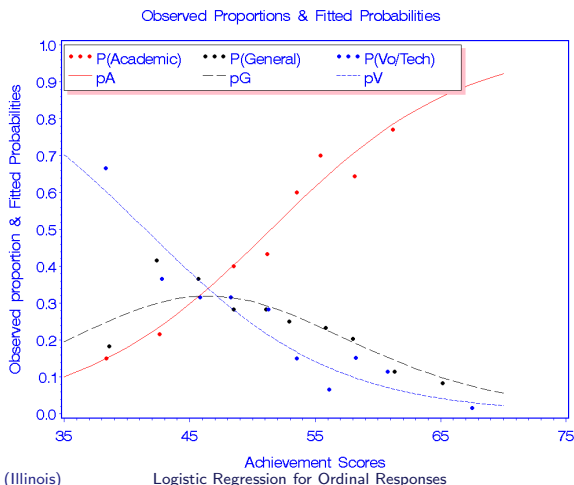


# I Fitted Category Probabilities



# I Observed Proportions and Fitted $\pi_j$ s

(Grouped only for plot)



# I Estimation in SAS and R

SAS:

- ▶ LOGISTIC (maximum likelihood).
- ▶ CATMOD (weighted least squares).
- ▶ GENMOD
- ▶ NLP or NLMIXED (maximum likelihood).
- ▶ Others?

For larger samples with categorical explanatory variables, results from MLE and WLS should be about same.

R:

- ▶ `plor` in the MASS package
- ▶ `vglm` in the VGAM package
- ▶ `lrm` in the rms package
- ▶ `ordinal` in the c1m package

# I SAS Logistic & GENMOD Code

```
proc logistic ;  
  model hsp = achieve;
```

In `proc logistic`, the cumulative logit model is the default if the response variable has more than 2 categories.

```
proc genmod;  
  model = achieve / dist=multinomial link=clogit type3;
```

“clogit” for Cumulative Logit, which is the default.

# I SAS PROC LOGISTIC: proportional odds assumption

Score Test for the Proportional Odds Assumption

Chi-Square	DF	Pr > ChiSq
0.8194	1	0.3653



# I SAS PROC LOGISTIC (edited) Output

## Response Profile

Ordered Value	program	Total Frequency
1	academic	308
2	general	145
3	vocation	147

Probabilities modeled are cumulated over the lower Ordered Values.

## I R: clm in ordinal package

```
# response has to be a factor
hsb$hsp ← as.factor(hsb$hsp)

po ← clm(hsp ~ achieve, data=hsb)
anova(po)

# test proportional odds assumption (before really looking at
model)
# — the test statistics is in column label "LRT"
nominal_test(po)

# Look at parameters
summary(po)

# Fitted probabilities for each type of hsp:
hsb$po.fit ← po$fitted
```

## I R: vglm in VGAM package

```
# Proportional odds model
# Note: response should be numeric (ordered)
po.vglm1 <- vglm(hsprog ~
  achieve,family=cumulative(parallel=TRUE),
                data=hsb)
summary(po.vglm1)
# Cumulative logits but allow slopes to differ
po.vglm2 <- vglm(hsprog ~
  achieve,family=cumulative(parallel=FALSE),
                data=hsb)
summary(po.vglm1)
# Difference in deviances of these two models:
lr <- 1082.413 - 1081.608
# p for testing proportional odds assumption
1 - pchisq(lr,1)
po.vglm1 <- vglm(hsprog ~ achieve,family=multinomial,
                data=hsb)
summary(po.vglm1)
```

## I R polr from MASS package

```
# response has to be a factor
hsb$hsp ← as.factor(hsb$hsp)
summary(po.polr ← polr(hsp ~ achieve,
data=hsb,Hess=TRUE )
# Check manual for definitions of these
names(po.polr)
# calculate and store p values
ctable ← coef(summary(po.polr))
p ← pnorm(abs(ctable[, "t value"]),lower.tail =
FALSE)*2
# combined table of coefficeints, se, t and pvalues
ctable ← cbind(ctable, "p value" = p)
```

## I R polr (continued)

```
# default method gives profiled CIs
ci ← confint(po.polr, level=.99)
#odds ratios
exp(coef(po.polr))
fit ← po.polr$fitted.values
```

## I Note on Example

We would get the exact same results regarding interpretation if we had used (i.e., put in **descending** option in proc LOGISTIC).

$$\begin{aligned} Y &= \text{high school program type} \\ &= \begin{cases} 1 & \text{VoTech} \\ 2 & \text{General} \\ 3 & \text{Academic} \end{cases} \end{aligned}$$

This reversal of the ordering of  $Y$  would

- ▶ Change the signs of the estimated parameters.
- ▶ Yield curves of cumulative probabilities that decrease (rather than increase).
- ▶ Essentially the same results.

## I Example 2: PIRLS

US 2006 Progress in International Reading Literacy Study (PIRLS) responses to item “How often to you use the Internet as a source of information for school-related work?” with responses

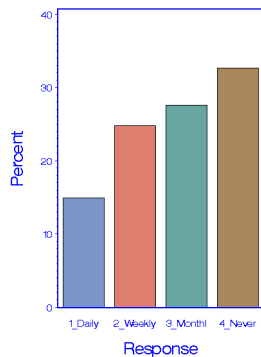
- ▶ Every day or almost every day ( $y_1 = 746, p_1 = .1494$ )
- ▶ Once or twice a week ( $y_2 = 1,240, p_2 = .2883$ )
- ▶ Once or twice a month ( $y_3 = 1,377, p_3 = .2757$ )
- ▶ Never or almost never ( $y_4 = 1,631, p_4 = .3266$ )

Predictors/Explanatory:

- ▶ Shortages at school.
- ▶ Time student spends in front of screen (electronic entertainment)
- ▶ Gender of student.

# I Graph of PIRLS Distribution

2006 US PIRLS on Internet Use for School





# I Problem with Model?

Score Test for the Proportional  
Odds Assumption

Chi-Square	DF	Pr > ChiSq
49.1500	6	<.0001

$H_o$ :  $\beta$ s are same over logits for all predictors.

$H_a$ : They are not all the same.

## I If Reject Proportional Odds Assumption

- ▶ If test is rejected this result could be due to large sample but not practical or substantively important. To investigate this fit separate logistic regressions to each logit.\*\*
- ▶ Add additional terms.
- ▶ Try non-symmetric link function.
- ▶ Use a different ordinal model.
- ▶ Add dispersion parameters.
- ▶ Permit separate effects for some variables (“partial proportional odds”)\*\*
- ▶ Use the baseline model but use order to interpret the results.\*\*

# I Is Problem Substantively Important?

Effect	Proportional Odds	Separate Binary Logistic Regression		
		$P(Y = 1)/P(Y > 1)$	$P(Y \leq 2)/P(Y > 2)$	$P(Y \leq 3)/P(Y = 4)$
shortages	-0.2055	-0.0685	-0.1986	-0.2571
girl	0.2225	0.1223	0.1213	0.3548
screenT	0.0599	0.1904	0.0725	-0.0028

- ▶ Shortages: Differ in terms of magnitude.
- ▶ Gender: Similar values.
- ▶ Screen Time: Different direction of effects.

Not just statistical but also substantive differences

# I Partial Proportional Odds

Relax assumption for shortages and allow different parameters for it.  
 Edited Output from PROC NL MIXED:

Parameter	Estimate	Standard Error	DF	<i>t</i> Value	Pr >   <i>t</i>	Gradient
Intercept 1	-1.9499	0.08314	4377	-23.45	< .0001	-0.00027
Intercept 2	-0.8769	0.06464	4377	-13.57	< .0001	0.000807
Intercept 3	0.6976	0.07237	4377	9.64	< .0001	-0.0008
Girl	0.1138	0.04423	4377	2.57	.0101	0.000621
ScreenT	0.0471	0.02001	4377	2.35	.0187	-0.00057
Shortage 1	-0.0603	0.08061	4377	-0.75	.4543	0.000023
Shortage 2	-0.1394	0.04256	4377	-3.27	.0011	0.000087
Shortage 3	-0.2560	0.05864	4377	-4.37	< .0001	-0.00045

## I Interpretation of Shortages

For fixed gender and screen time,

- ▶ The odds ratio daily versus more than daily usage for shortages  $x + 1$  equals  $\exp(-.0603) = 0.94$  the odds for shortage  $x \rightarrow$  equal odds.
- ▶ The odds ratio for daily or weekly use versus monthly or never for  $x + 1$  shortages equals  $\exp(-0.1394) = 0.87$  the odds for  $x$  shortages.
- ▶ The odds ratio for monthly or more usage versus never for shortages  $x + 1$  equals  $\exp(-.2560) = .77$

What does this mean:

- ▶ More shortages less frequently use computers?
- ▶ More shortage more frequently use computers?

## I Baseline Model but Use Order

All possible odds ratios: For 1 unit increase in shortages, the odds ratios for row versus column equal

	Daily	Weekly	Monthly	Never
Daily	—	1.11	.98	.80
Weekly	0.90	—	.89	.72
Monthly	1.02	1.12	—	.81
Never	1.25	1.39	1.23	—

- ▶ The odds of Daily versus Weekly are 1.11 the odds for 1 unit more on shortages.
- ▶ For greater shortages, daily use of computers is more likely than weekly.
- ▶ For fewer shortages, monthly or never using computers is more likely than daily use.

# I Interpretation of Shortages

For better and more proper analysis of data see Anderson, Kim & Keller (2010) and see results for multinomial model. . .

When take into account hierarchical structure, missing data and unequal probability sampling (particularly of the school), the impact of shortages of computer use quite different.

## I Final Comments on Cumulative Logit Models

- ▶ It takes into account the ordering of the categories of the response variable.
- ▶ One probability is monotonically increasing as a function of  $x$ . (see figure of estimated probabilities from HSB example).
- ▶ One probability is monotonically decreasing as a function of  $x$ . (see figure of estimated probabilities).
- ▶ Curves of probabilities for intermediate categories are uni-modal with the modes (maximum) corresponding to the order of the categories.
- ▶ The conclusions regarding the relationship between  $Y$  and  $x$  are not affected by the response category.



## I Final Comments on Cumulative Logit Models

- ▶ The specific combination of categories examined does not lead to substantially different conclusions regarding the relationship between responses and  $x$ .
- ▶ IRT connection: Samejima's (1969) graded response model for polytomous items is the same as the proportional odds model except that  $x$  is a latent continuous variable.

## I Adjacent-Categories Logit Models

Rather than using all categories in forming logits, we can just use  $J - 1$  pairs of them.

To incorporate the ordering of the response, we use adjacent categories:

$$\log \left( \frac{\pi_j}{\pi_{j+1}} \right) \quad j = 1, \dots, J - 1$$

The logit model for one (continuous) explanatory variable  $x$  is

$$\log \left( \frac{\pi_j}{\pi_{j+1}} \right) = \alpha_j + \beta x \quad j = 1, \dots, J - 1$$

## I Adjacent-Categories Logit Models

- ▶ This model is a special case of the baseline model. (shown below)
- ▶ It would not work for the PIRLs example

	Daily	Weekly	Monthly	Never
Daily	—	1.11		
Weekly	0.90	—	.89	
Monthly		1.12	—	.81
Never			1.23	—

- ▶ If we had a single  $\beta$ , these odds ratios would all be equal.

# I An Example for Adjacent Categories

GSS Happiness data from Agresti (2013):

- ▶ Response variable is happiness with categories 1= very happy, 2=pretty happy, and 3=not too happy.
- ▶ Predictors are
  - ▶ Race with categories 1=black and 0=white.
  - ▶ Number of traumatic events that happened to respondent or relatives in the last year. Values range from 0 to 5.
- ▶ Estimated model:

$$\log(P(Y_i = j)/P(Y_i = j+1)) = \hat{\alpha}_j - 0.357(\text{traumatic})_i - 1.84(\text{race})_i$$

note:  $\hat{\alpha}_1 = 2.532$  and  $\hat{\alpha}_2 = 3.028$

# I Interpretation

Estimated model:

$$\log(P(Y = j)/P(Y = j+1)) = \hat{\alpha}_j - 0.357(\text{traumatic})_i - 1.842(\text{race})_i$$

- ▶ Given number of traumatic events, the estimated odds of being **very happy versus pretty happy** for whites are  $\exp(1.842) = 6.31$  times the odds for blacks.
- ▶ Given number of traumatic events, the estimated odds of being **pretty happy versus not too happy** for whites are  $\exp(1.842) = 6.31$  times the odds for blacks — the same.
- ▶ Given race, the estimated odds of **very happy versus pretty happy** for  $x$  traumatic events are  $1/\exp(-.357) = 1.429$  times the odds for  $x + 1$  events.
- ▶ Odd ratio for **pretty happy versus not too happy** are the same as above.

# I Estimation

SAS:

- ▶ CATMOD: Weighted least squares, but if there are 0s, need to add small number to each cell.
- ▶ CATMOD: Maximum likelihood estimation involves design matrix that puts restrictions on parameters of the baseline model.
- ▶ NLMIXED: MLE for baseline but modify to correspond to adjacent categories.

R:

- ▶ vglm in VGAM
- ▶ others?

# I CATMOD and WLS

```
title 'Check for zeros';  
proc freq data=gss;  
  tables race*trauma*happy / nopercnt norow nocol sparse out=table;  
data fillin;  
  set table;  
  count2=count+ .01;  
title 'Adjacent Categories (WLS)';  
proc catmod data=fillin;  
  weight count2;  
  response alogits;  
  population race trauma;  
  direct trauma race ;  
  model happy = response race trauma;
```

WILL BE RUN IN LECTURE

## I To Use NLMIXED

We make use of the fact that the adjacent categories models is a special case of the baseline model.

Baseline odds = Product of adjacent categories odds ,and  
 logarithm of odds equals sum

$$\log \left( \frac{\pi_{ij}}{\pi_{iJ}} \right) = \log \left( \frac{\pi_{ij}}{\pi_{ij+1}} \right) + \log \left( \frac{\pi_{i(j+1)}}{\pi_{i(j+2)}} \right) + \dots \log \left( \frac{\pi_{i(j-1)}}{\pi_{iJ}} \right)$$

for  $j = 1, \dots, J - 1$ .

e.g., Taking a simple model for the adjacent categories,

$$\begin{aligned} \log \left( \frac{\pi_{ij}}{\pi_{iJ}} \right) &= (\alpha_j + \beta x_i) + (\alpha_{j+1} + \beta x_i) + \dots (\alpha_{J-1} + \beta x_i) \\ &= \underbrace{\sum_{k=j}^{J-1} \alpha_k}_{\alpha_j^*} + \underbrace{\beta(J-j)}_{\beta_j^*} x_i \end{aligned}$$



# I NLMIXED & MLE

```
title'Adjacent Categories (MLE)';
proc nlmixed data=gss;          * <— un-collapsed data;
  parms a1=0.1 a2=0.1 br=0.1 bt=0.1;
  /* Linear predictors */
  eta1 = a1 + br*(3-1)*race + bt*(3-1)*trauma;
  eta2 = a2 + br*(3-2)*race + bt*(3-2)*trauma;
  /* Define likelihood */
  if happy=1 then prob= exp(eta1)/(1 + exp(eta1) + exp(eta2));
  if happy=2 then prob= exp(eta2)/(1 + exp(eta1) + exp(eta2));
  if happy=3 then prob= 1 / (1 + exp(eta1) + exp(eta2));
  /* To make sure that probabilities are valid ones */
  p = (prob>0 and prob<=1)*prob + (prob<=0)*1e-8 + (prob>1);
  loglike = log(p);
  /* Specify distribution for response variable */
  model happy ~ general(loglike);
```

# I Comparison of WLS & MLE

## Weighted Least Squares Estimates from CATMOD

Parameter	Estimate	Standard Error	Chi-Square	Pr > ChiSq
Intercept	-1.1148	0.3698	9.09	0.0026
_RESPONSE_1	1.7341	0.3056	32.20	<.0001
race	1.7317	0.8031	4.65	0.0311
trauma	0.2053	0.1876	1.20	0.2740

## MLE from NLMIXED

Parameter	Estimate	Standard Error	df	<i>t</i>	<i>Pr</i> >   <i>t</i>
a1	2.5315	0.7464	23	3.39	0.0025
a2	3.0276	0.5740	23	5.27	<.0001
br	-1.8424	0.6419	23	-2.87	0.0087
bt	-0.3570	0.1640	23	-2.18	0.0400

# I R: vglm

```
summary(adj.cat1 ← vglm(happy ~ race + trauma, data=gss,  
                        family=acat(parallel=TRUE, reverse=TRUE)))  
# odds ratio for race (see notes for interpretation)  
exp(1.8423)  
# odds ratio for number of traumatic events  
exp(.3570)  
# Relax assumption on equality of slopes  
summary(adj.cat2 ←  
vglm(happy ~ race + trauma, data=gss,  
      family=acat(parallel=FALSE, reverse=TRUE)))  
# Difference in deviances  
lr ← 148.1996 - 146.8737  
df ← 190-188  
1-pchisq(lr,df)
```

# I Adjacent Categories or Proportional Odds Model?

(from Agresti, 2013)

- ▶ Both tend to fit (or not) for a particular data set.
- ▶ If prefer effects to refer to individual categories, use adjacent categories.
- ▶ If want to use entire scale for each logit or hypothesize underlying continuous latent variable, use proportional odds model.
- ▶ Effects for proportional odds tend to be larger because whole scale is used.
- ▶ Proportional odds models not effected by choice and number of response categories.
- ▶ Adjacent is more general than proportional odds model—if replace  $\beta$  by  $\beta_j$  in the adjacent model, cumulative probabilities will be in correct order—this isn't true for the partial proportional odds model.

# I Adjacent Categories for Ordered Grouped Data

- ▶ Recall... General Social Survey (1994) data from before.
  - ▶ Item 1: A working mother can establish just as warm and secure of a relationship with her children as a mother who does not work.
  - ▶ Item 2: Working women should have paid maternity leave.
- ▶ When using  $u_i = i$  and  $v_j = j$  as scores and fitting the independence log-linear model and the uniform association model

$$\log(\mu_{ij}) = \lambda + \lambda_i^I + \lambda_j^{II} + \beta_{ij}$$

- ▶ Results from model fitting

Model/Test	<i>df</i>	<i>G</i> <sup>2</sup>	<i>p</i>	Estimates
Independence	12	44.96	< .001	
Uniform Assoc	11	8.67	.65	$\hat{\beta} = .24, ASE = .0412$

## I Adjacent Categories for Ordered Grouped Data

Suppose that we consider item 2 as the response variable and model adjacent category logits with the restriction that  $\beta_j = \beta =$  a constant.

$$\begin{aligned}
 \log \left( \frac{\mu_{i(j+1)}}{\mu_{ij}} \right) &= \lambda + \lambda_i^I + \lambda_{j+1}^{II} + \beta i(j+1) \\
 &\quad - (\lambda + \lambda_i^I + \lambda_j^{II} + \beta ij) \\
 &= (\lambda_{j+1}^{II} - \lambda_j^{II}) + \beta(ij + i - ij) \\
 &= \alpha_j^* + \beta i
 \end{aligned}$$

So the estimated local odds ratio equals (and the effect of response on item 1 on item 2 for adjacent categories)

$$e^{\hat{\beta}} = e^{.24} = 1.28$$

# I Continuation-ratios Logit

In this approach, the order of the categories of the response variable is used to form  $(J - 1)$  logits as follows:

$$\log\left(\frac{\pi_1}{\pi_2}\right), \log\left(\frac{\pi_1 + \pi_2}{\pi_3}\right), \dots, \log\left(\frac{\pi_1 + \dots + \pi_{J-1}}{\pi_J}\right)$$

or

$$\log\left(\frac{\pi_1}{\pi_2 + \dots + \pi_J}\right), \log\left(\frac{\pi_2}{\pi_3 + \dots + \pi_J}\right), \dots, \log\left(\frac{\pi_{J-1}}{\pi_J}\right)$$

These are called “**continuation-ratio logits**”

# I Continuation-ratios Logit

- ▶ Just apply regular binary logistic regression to each one.
- ▶ The fitting is separate (no restrictions on parameters across the logits).
- ▶ The sum of the separate  $df$  and  $G^2$  provide an overall global goodness of fit test and measure.



## I Example

NYLS Example from Powers & Xie (2000) *Statistical Methods for Categorical Data Analysis* (1st edition). page 236=238.

$n = 978$  of 20-22 year old men from NYLS.

Race	Father's education	Employment Status		
		In school 1	Working 2	Inactive 3
White/other	$\leq 12$ yr	204	195	131
Black	$\leq 12$ yr	100	53	67
White/other	$> 12$ yr	78	90	28
Black	$> 12$ yr	18	5	9

Best baseling/multinomial model was (R,F).

# I Example

Logit Model	df	P(School) vs P(Working)		P(School or Working) vs P(Not Working)		
		$G^2$	p	df	$G^2$	p
null	3	19.1576	< .01	3	16.5575	< .01
(F)	2	17.8385	< .01	2	9.7941	< .01
(R)	2	2.6484	.27	2	6.0043	.05
(F,R)	1	2.3879	.12	1	1.3512	.25

- ▶ Test between (R) and (F,R),  
 $G^2((R)|(F,R)) = 6.0043 - 1.3512 = 4.6531$ ,  $df = 1$ ,  $p = .03$ .
- ▶ Total:  $G^2 = 2.3879 + 1.3512 = 3.7391$ ,  $df = 3$ ,  $p = .29$
- ▶ Only Race is needed for P(School)/P(Working).
- ▶ Father's education and race needed for P(School or Working)/P(Inactive).

## I Recommendation

The overriding determinate of which model you should reflect the goals of the analysis and that the model fits the data well.

e.g., Buki, Jamison, Anderson & Curdera (2007). Predictors of mammography and pap smear screening in Latina women. *Cancer*, 110, 1578-1585.

### Research Questions:

- ▶ What predicts whether a woman has even been screened?
- ▶ Among those who have ever been screened, what predicts whether screening is up to date?

What model should (did) we use?