The Analysis of Ordinal Data with Graphs and Odds Ratios

Robin High
Department of Biostatistics
University of Nebraska Medical Center
Omaha, NE

University of Nebraska Medical Center

Outline

Ordinal Data to Compute Scales

- **Data Processing**
- **Bar Charts**
- **Centered Bar Charts**
- **Diverging Stacked Bar Charts**

Ordinal Logistic Regression Models

- **Cumulative Logit**
- **Adjacent Logit**
- **Odds Ratios**
- **Predicted Probabilities**
- **Visual Display**

Power Analysis with Ordered Categories as Outcomes

SAS Software Procedures

Ordinal Data

Statistical Procedures with MODEL statement options

```
  PROC LOGISTIC: /                  link = clogit   OR  link=alogit
  PROC GENMOD   : / dist=multinomial link = cumlogit
  PROC GLIMMIX : / dist=multinomial link = cumlogit
  PROC NLMIXED :  write out equations
```

LOGISTIC and NLMIXED procedures:
   computation of ordinal response models with partial proportional odds

Visual Displays

PROC SGPLOT
PROC SGPANEL
PROC PLOT (rough graphs for diagnostic purposes)

Utility Procedures

PROC SUMMARY
PROC TABULATE
PROC FREQ
PROC FORMAT
PROC MODEL (ETS) for solving equations leading to power calculations

Ordinal Data
Categorical Data with an Inherent Order

Responses that reflect an ordered progression from the lowest (or highest) level to the next level without reversing trend

Evaluate responses difficult or impossible to quantify, qualitative or "subjective" endpoints, such as:

- **Pain**
- **Agreement / Disagreement**
- **Behavior (Frequency)**
- **Difficulty of a task**
- **Political viewpoints**

Types of Ordinal Data

Agreement (5 levels):

- Strongly Agree   Agree   [Uncertain / Indifferent]   Disagree   Strongly Disagree

Pain: Beck Anxiety Inventory (BAI) 21 items (4 increasing levels of severity)

- **Not at all   Mild   Moderate   Severe**

Behavior:
BPFAS – Behavior Pediatrics Feeding Assessment Scale (25+10 = 35 items)
CFSQ – Caregiver Feeding Style Questionnaire (19 Items)

- **Never   Rarely   Sometimes   Often   Always**

Frequency: how many times certain behaviors occurred in a given time period

- **0   1-5   6-10   11-20   21 or more**

Coding actual numbers into ordinal categories is generally NOT recommended for data analysis purposes

Coding Ordinal Data

Order matters: default settings of SAS procedures assume ordinal data values are sorted in increasing order

- Alphabetical: a b c d e ..
- Numerical: 1, 2, 3, ..k

Where to start:

- Response level of "greatest" interest coded as a 1

Assign verbal meaning of coded numbers with a format

```
PROC FORMAT;
VALUE rsp  1='Never' 2='Rarely' 3='Sometimes' 4='Often' 5='Always';
VALUE rspA 1='a Never' 2='b Rarely' 3='c Sometimes' 4='d Often' 5='e Always';
RUN;
```

Apply order= option when displaying data:

```
order = internal    ( with rsp  )
order = formatted   ( with rspA )
```

Recommendation: Do not assign formats in DATA step

Ordinal Data in the Computation of Scales

Reverse code items of the scale as directed
Multiple ordinal responses combined (added) to form scales
Compute Cronbach's alpha to assess internal consistency (reliability)

Caregiver Feeding Style Questionnaire (CFSQ)

Input data for 19 ordinal responses

- Each variable coded with an integer from 1 to 5
- Data stored in columns

```
INPUT y1 y3 y3 - y19;    * SAS INPUT statement reading external file;
```

```
LABEL
y1 = 'Physically struggle with child to get him/her to eat'
y2 = 'Promise child something other than food'
y3 = 'Encourage child to eat by arranging food'
y4 = 'Ask questions about food'
etc.
```

For graphs: convert variable labels into formats

Steps to Make a Diverging Stacked Bar Chart

Ordinal data coded as 1, 2, 3, 4, 5

READ data (variables stored in columns)

Add variable labels

Make a format of the variable labels

Convert individual data values to counts with PROC SUMMARY (macro loop)

PROC FREQ to compute row percents for each item

DATA step processing

PROC SGPLOT to make various graphs

Computing summary counts

Goal:

Summary counts for all values of each variable, including those not present in the response with count of 0

```
PROC FORMAT;
VALUE _rsp 1='1' 2='2' 3='3' 4='4' 5='5' ;
RUN;

PROC SUMMARY DATA=inpdat nway completetypes;
CLASS  y1  / preloadfmt;
VAR y1 ;
OUTPUT OUT=_cnts   n=count;
FORMAT y1 _rsp.;
RUN;
```

- Enter the PROC SUMMARY step into a macro and loop through the individual variables of the scale to collect counts for each variable
- Append summary data to a master file one variable at a time

Why not PROC FREQ?

```
ODS OUTPUT onewayfreqs=_onfr;

PROC FREQ DATA = inpdat;
TABLE y1 y2 y3   ;
FORMAT y1 y2 y3 _rsp. ;
RUN;

proc print data=_onfr; run;
```

Two reasons PROC FREQ ODS OUTPUT file not recommended

- Dataset needs considerable processing for next step
- Does not produce 0s for response levels not present

File with data in stacked layout with counts

- Append results from each ordinal variable with 5 levels into one file with frequency (count) often missing values of 4 and 5 in the data

PROC SUMMARY

item	response	count
y1	1	12
y1	2	11
y1	3	8
y1	4	0
y1	5	4
y2	1	16
y2	2	12
y2	3	4
y2	4	3
y2	5	0
y3	1	17
y3	2	11
y3	3	7
y3	4	0
y3	5	0

PROC FREQ

item	response	count
y1	1	12
y1	2	11
y1	3	8
y1	5	4
y2	1	16
y2	2	12
y2	3	4
y2	4	3
y3	1	17
y3	2	11
y3	3	7

Compute Individual Row and Cumulative Row Percents for Each Item

Obtain row percents with PROC FREQ, oneway counts, and cumulative percents for each ordered value 1, 2, 3, 4, 5

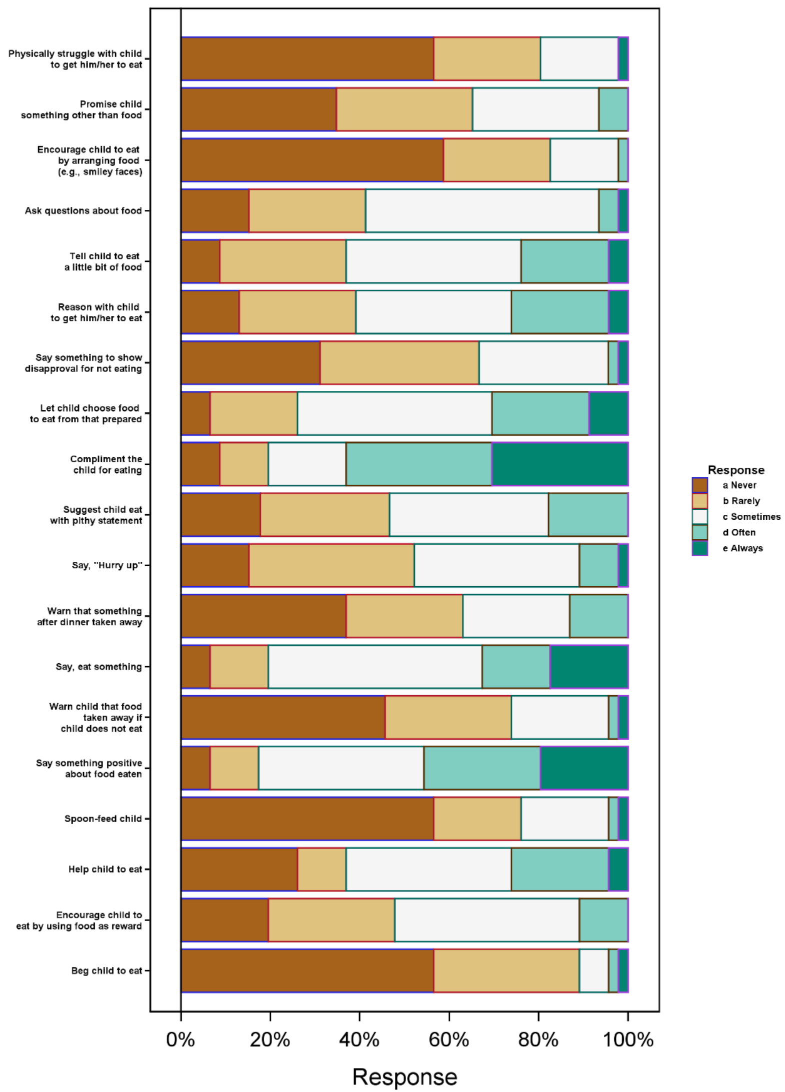
```
ODS OUTPUT onewayfreqs=_onewy
   (keep=va frequency rsp cumpercent rename=(cumpercent=rowpercent));

ODS LISTING close;

PROC FREQ DATA=vaplt order=internal;
BY va ;  * va is a number from 1 to n, index for items in the scale;
TABLE rsp ;
WEIGHT count / zero;  * count is from PROC SUMMARY;
TITLE 'Levels and Response Frequencies with Row Percents';
FORMAT va yC. ;  * Save the variable labels in the file;
RUN;

ODS listing;

proc print data=_onewy; run;
```

Two Examples of Bar Graphs

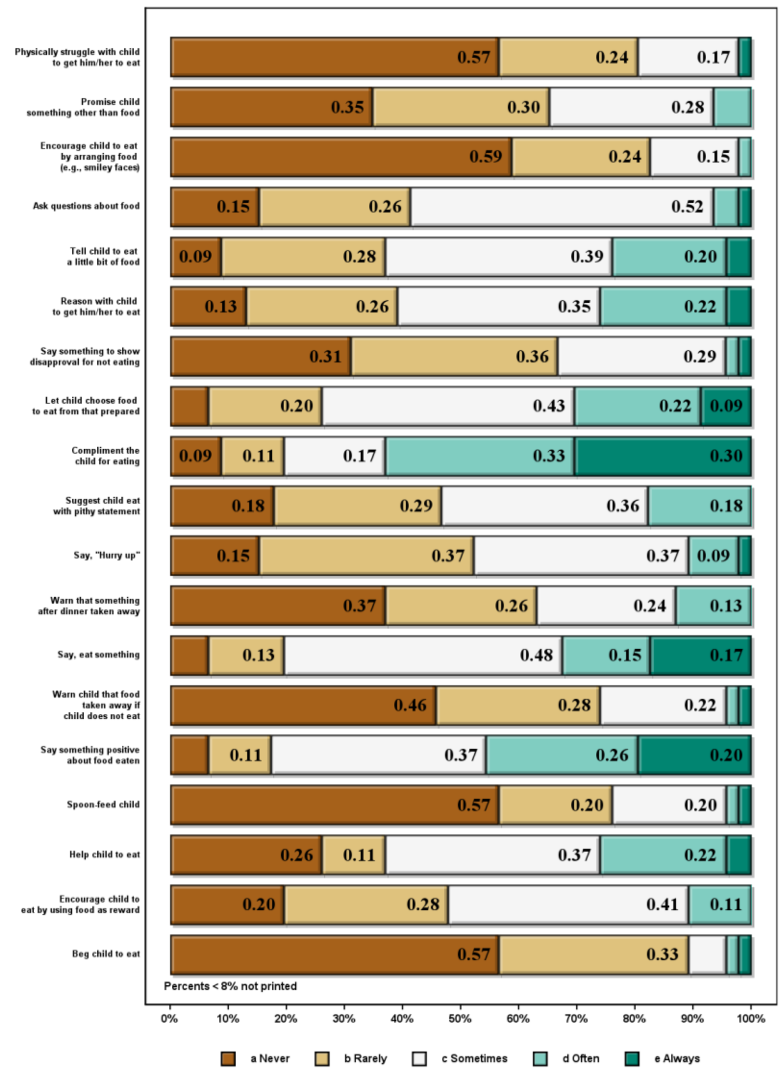
Make from PROC SUMMARY file with SGPLOT VBAR statement

Data processing and then SGPLOT with HIGHLOW statement (specify bar width) to display percents

SGPLOT: segmented horizontal bar graph



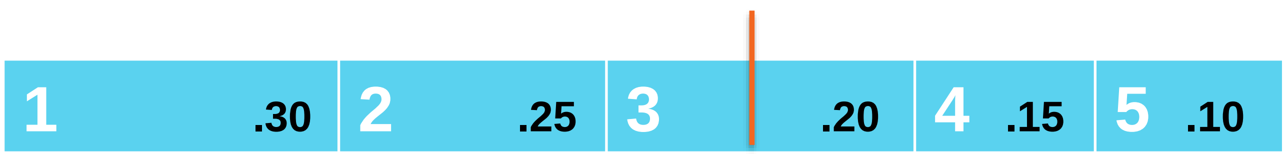
Five Response Likert Scale



Centered and Diverging plots
Segment boundaries

- Ordinal codes on left (white) | row proportions on right (black)
- Cumulative proportions printed below bar
- Orange line is midpoint of the ordinal codes
- Values to plot in bottom row: subtract the midpoint from each boundary value

0	0.35	0.60	0.80	1	cumulative
-.60	-.25	0	.2	.4	xaxis value

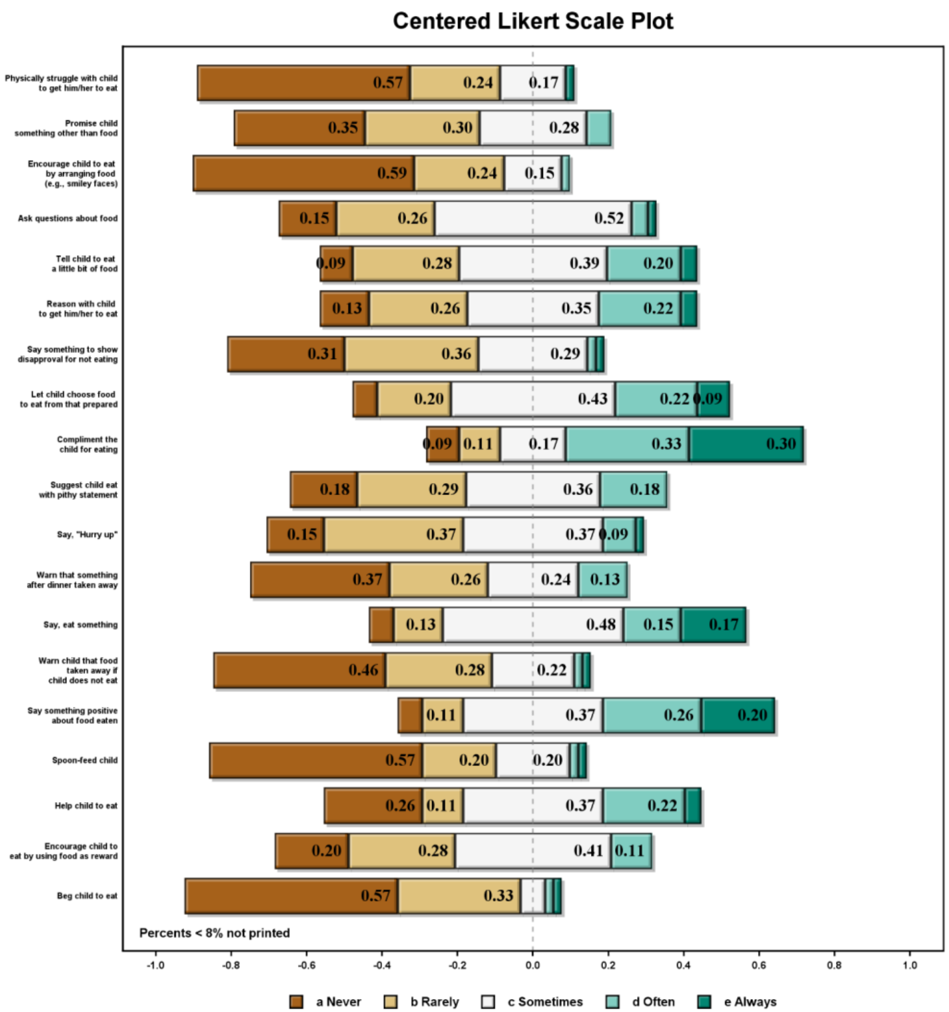


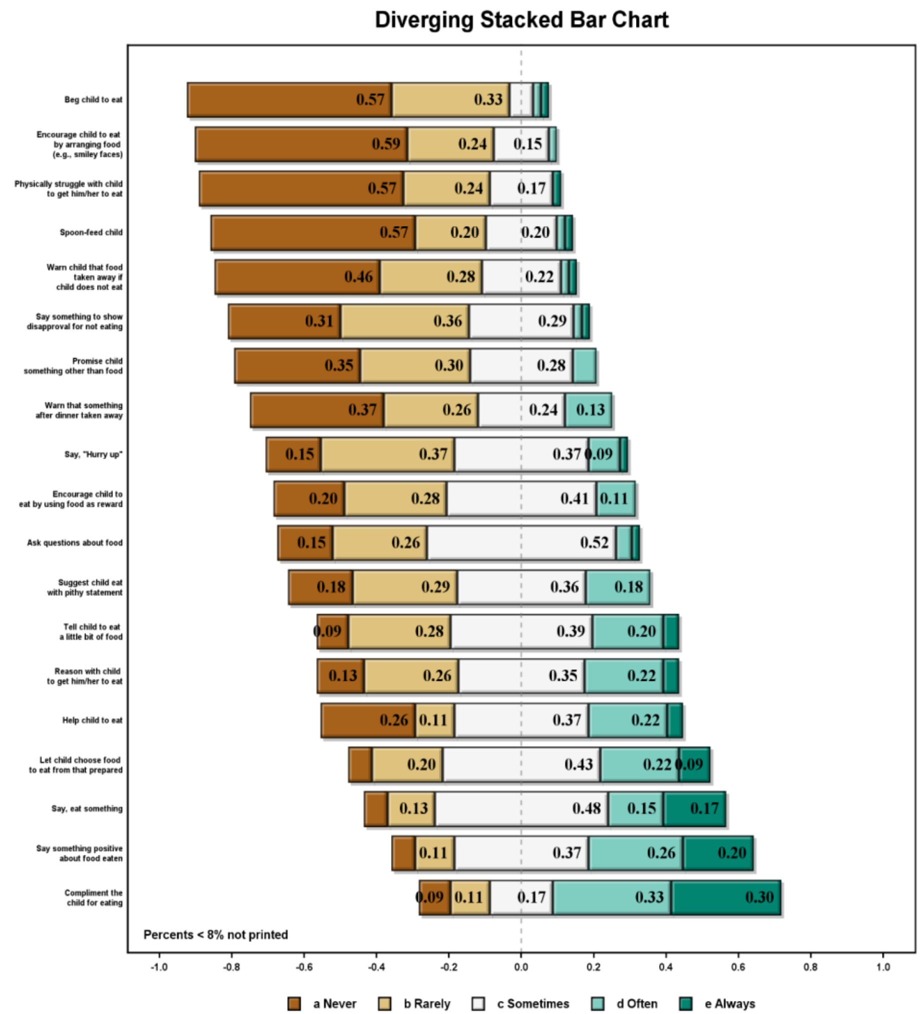
0	0.30	0.55	0.65	0.75	0.90	1	cumulative
-.65	-.35	-.10	0	.10	.25	.35	xaxis value

Adjust cumulative percents with cumulative midpoint

```
xL: Lower cumulative percent of each bar
xH: Upper cumulative percent of each bar

DATA crstb3;
MERGE crstb2 mdpnt;
BY va ;
xL = xL - midpnt;
xH = xH - midpnt;
run;
```

Bar Graphs Centered at midpoint (level 3 of 5 values)



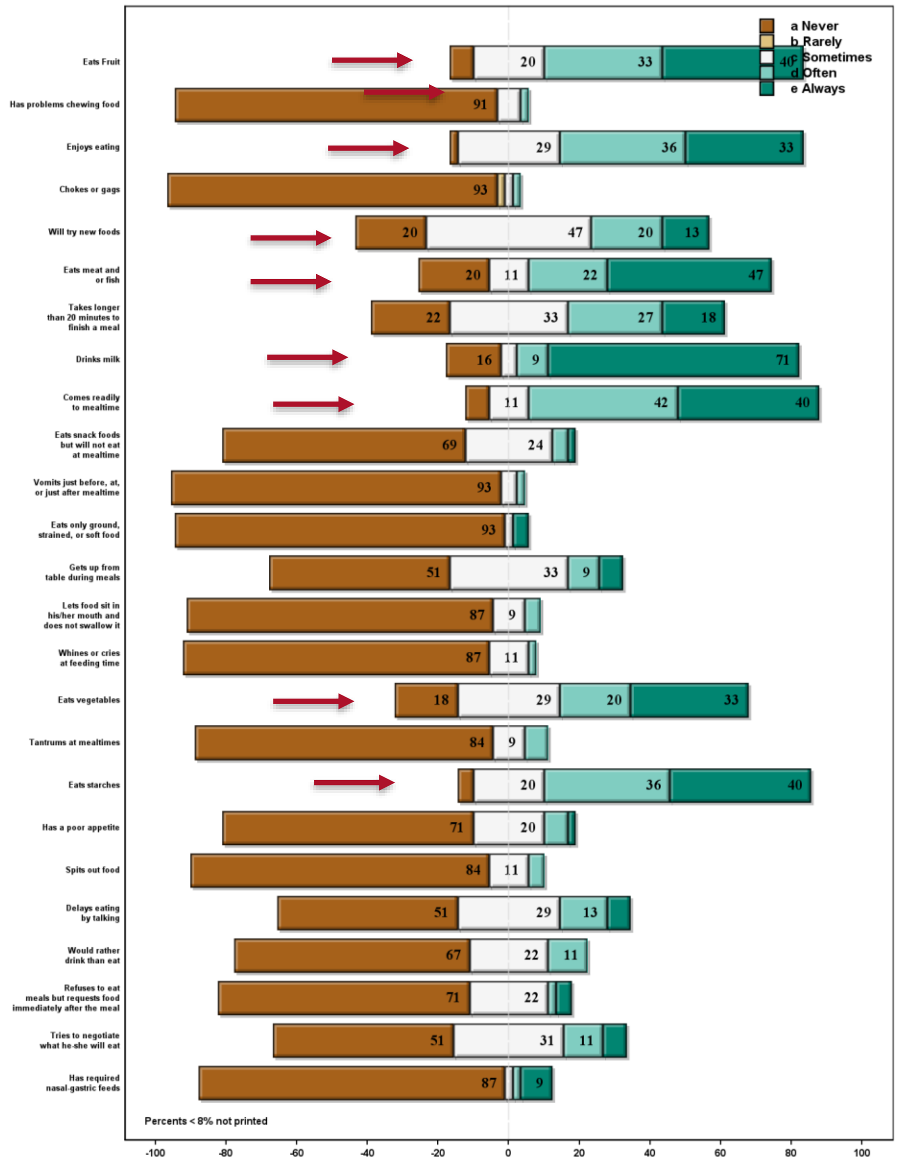


Choice of contrasting colors: colorbrewer2.org

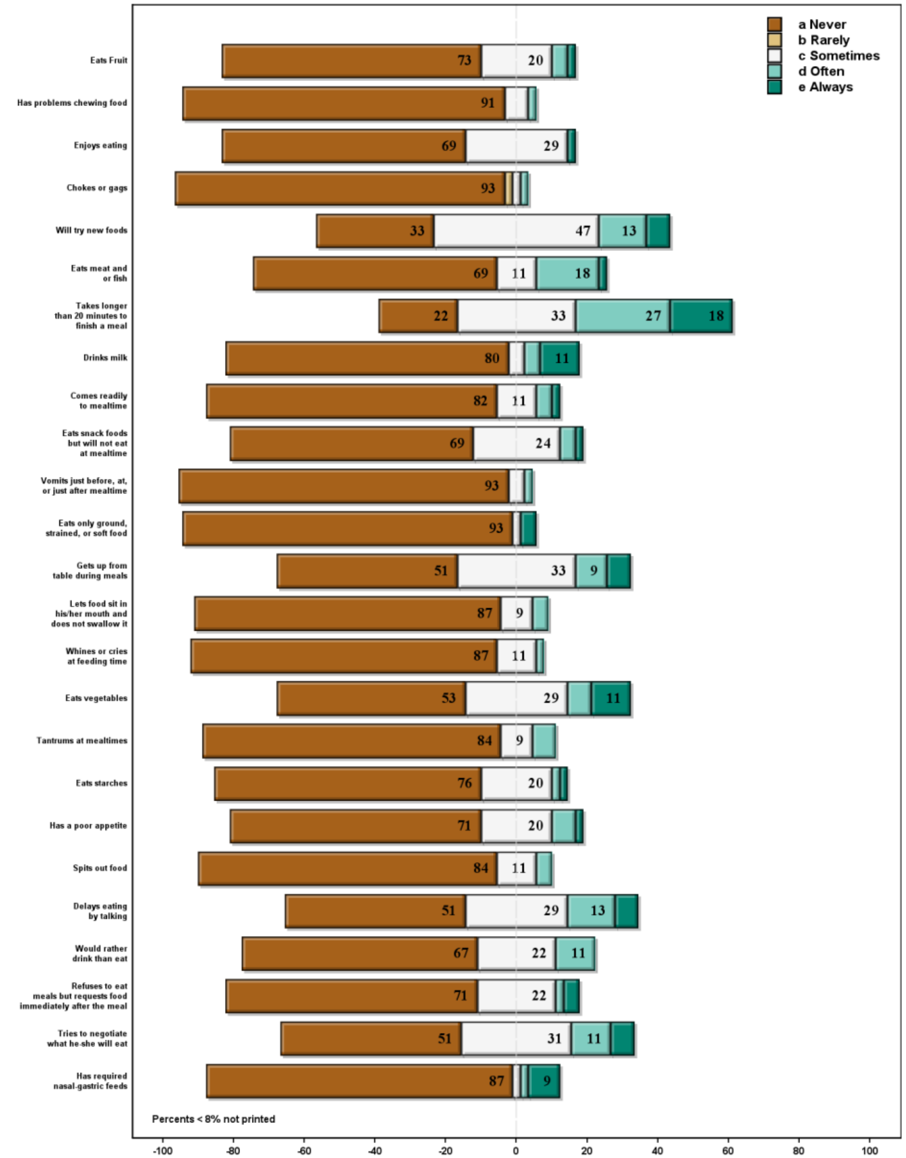
Original Data: alpha = 0.60
Need to reverse code

Items 1 3 5 6 8 9 16 18
reverse coded: alpha = 0.79

BPFAS Child Behavior: Likert Scale Centered at Midpoint



BPFAS Child Behavior: Likert Scale Centered at Midpoint



Diverging Stacked Bar Char

Reference:

Design of Diverging Stacked Bar Charts for Likert Scales
Richard M. Heiberger, Naomi B. Robbins
Journal of Statistical Software
March 2014, Volumne 57, Issue 5

Analysis of Ordinal Data

Rather than components of a scale, consider ordinal data as the subjective outcome or measure of an attribute

Ordinal nature of outcome data much too often ignored

Reduce k ordinal levels to two levels (recode as binary)

- May be necessary due to small cell counts

Analyze as continuous/interval data (treat as numbers)

- t-tests, ANOVA, regression methods

Preferred approach is to summarize with odds ratios and predicted probabilities summarized in tables and displayed with graphs

Binary Logistic Regression: PROC FREQ

	a Yes	b No	Odds	Odds Ratio
a Treatment	p11	p12	odds1 =p11 / p12	OR = odds1/odds2
b Control	p21	p22	odds2 = p21 / p22	

Data summarized as row percents in a 2x2 table

Columns: outcome

Primary level of interest placed in column 1

Rows: treatment

Reference category (Control) placed in row 2

Row probabilities sum to 1

p11 > p21 implies odds ratio will be > 1

In SAS enter:

```
PROC FREQ DATA= inpdat order=formatted;
TABLE x * y / nocol nopercent cmh2;
FORMAT x trt. y ysn. ;
RUN;
```

Ordinal Logistic Regression with k=4 levels

	1=Very Good	2=Good	3=Moderate	4=Poor
a Treatment	p11	p12	p13	p14
b Control	p21	p22	p23	p24

Response has k discrete categories (usually 3 to 5 levels)

Make a 2 x k table (k=4)

Columns: outcome
level of primary interest in column 1)

Rows: Treatment vs Control
reference category in row 2

Row i probabilities, p_{ij}, sum to 1 for i=1,2

Order of response defined as "ascending"

The objective is to compute the odds ratio(s) and predicted probabilities based on two methods of grouping the ordinal responses

Cumulative logit
Adjacent logit

Response Profile: Ascending

```
PROC LOGISTIC DATA=inpdat ;
MODEL y = x1 x2;
RUN;
```

```
Response Profile

 Ordered                       Total
   Value          y        Frequency

       1          1               60
       2          2               42
       3          3               21
       4          4               14

Probabilities modeled are cumulated over the lower Ordered Values.
```

```
Computed over lower order values in this setting means:

 [ Prob(y=1)                             ] / [ Prob(y=2) + Prob(y=3) + Prob(y=4) ]
 [ Prob(y=1) + Prob(y=2)                 ] / [             Prob(y=3) + Prob(y=4) ]
 [ Prob(y=1) + Prob(y=2) + Prob(y=3)     ] / [                         Prob(y=4) ]
```

Response Profile: Descending

```
PROC LOGISTIC DATA=inpdat descending;
MODEL y = x1 x2;
RUN;
```

Response Profile

Ordered Value	y	Total Frequency
1	4	14
2	3	21
3	2	42
4	1	60

Probabilities modeled are cumulated over the lower Ordered Values.

Computed over lower order values in this setting means:

```
[ Prob(y=4)                             ] / [ Prob(y=3) + Prob(y=2) + Prob(y=1) ]
[ Prob(y=4) + Prob(y=3)                 ] / [             Prob(y=2) + Prob(y=1) ]
[ Prob(y=4) + Prob(y=3) + Prob(y=2)     ] / [                         Prob(y=1) ]
```

Ordinal Logistic Regression: Cumulative Logit

	1 Very Good	2 Good	3 Moderate	4 Poor
Treatment	p11	p12	p13	p14
Control	p21	p22	p23	p24

```
Cumulative logit
Probabilities summed across all four response levels

Row T:   oddsT   =       A          /         B
Row C:   oddsC   =       C          /         D

T:  odds11  =  (p11              )   /  (p12 + p13 + p14)
C:  odds21  =  (p21              )   /  (p22 + p23 + p24)

T:  odds12  =  (p11 + p12        )   /   (p13 + p14)
C:  odds22  =  (p21 + p22        )   /   (p23 + p24)

T:  odds13  =  (p11 + p12 + p13)   /   ( p14)
C:  odds23  =  (p21 + p22 + p23)   /   ( p24)
```

Ordinal Logistic Regression: Adjacent Logits

	a Very Good	b Good	c Moderate	d Poor
a Treatment	p11	p12	p13	p14
b Control	p21	p22	p23	p24

```
Adjacent logit  (pairs of adjacent cells)

T: oddsT =  A   /   B
C: oddsC =  C   /   D

T: odds11 = p11 / p12
C: odds21 = p21 / p22

T: odds12 = p12 / p13
C: odds22 = p22 / p23

T: odds13 = p13 / p14
C: odds23 = p23 / p24
```

```
Both models produce one odds
ratio for the entire table

Odds Ratio refers to
    odds11 / odds21
    odds12 / odds22
    odds13 / odds23
```

Ordinal Logistic Regression
Checking the proportional odds condition

Score Test for the Proportional Odds Assumption

SAS Global Forum 2013
Paper 446-2013
Ordinal Response Modeling with the LOGISTIC Procedure
Bob Derr, SAS Institute Inc.

Graphical displays available with a set of macros found at:

http://support.sas.com/rnd/app/stat/papers/aastaffcode2013.html

Simple method:

- Divide the k levels of the response into k-1 sets of binary outcomes with format statements (as defined by odds computations on two previous slides)
- Graph the three beta coefficients vs the k-1 odds conditions
- Connect with a best fitting line
- Want to observe horizontal line with minimal deviations

Logistic Regr: Partial Proportional Odds

When diagnostics indicate non-proportional outcomes, apply the "unequalslopes" option in PROC LOGISTIC

```
PROC LOGISTIC DATA=indat;
CLASS x1 / param=ref;
MODEL y = x1 x2 / link=clogit unequalslopes;
RUN;
```

Or specify the specific variable name(s) to apply unequalslopes

```
MODEL y = x1 x2 / link=clogit unequalslopes=x1;
```

Ordinal Logistic Regression with NLMIXED
Odds Ratios for both Cumulative Logits

Proportional Odds: one odds ratio

```
proc nlmixed data=indat(rename=(outcome = y));
parms Int1 -2 Int2 0 Int3 1.5 bx .07 ;

eta1 = Int1 + bx * (trt=1) ; * trt=0 ref category;
eta2 = Int2 + bx * (trt=1) ;
eta3 = Int3 + bx * (trt=1) ;

cp1= 1 / (1 + exp(-eta1));
cp2= 1 / (1 + exp(-eta2));
cp3= 1 / (1 + exp(-eta3));

p1 = cp1;          * calculate single cell probs ;
p2 = cp2 - cp1;    * from the cumulative probs;
p3 = cp3 - cp2;
p4 =   1 - cp3;

lk =   (p1**(y EQ 1)) * (p2**(y EQ 2))
     * (p3**(y EQ 3)) * (p4**(y EQ 4)); * ascending;

  lk = max(min(lk,1-1E-9),1E-9);
lglk = log(lk);
ESTIMATE 'Odds Ratio' EXP(bx);

MODEL y ~ general(lglk);
TITLE "NLMIXED: ordinal Logr Regr with link = clogit";
run;
```

Partial Proportional odds: three odds ratios

```
parms Int1 -2 Int2 0 Int3 1.5 bx .07 da .1 db .1 ;

eta1 = Int1 + (bx   ) * (trt=1) ;
eta2 = Int2 + (bx+da) * (trt=1) ;
eta3 = Int3 + (bx+db) * (trt=1) ;

ESTIMATE 'Odds Ratio 1 vs 234' EXP(bx     );
ESTIMATE 'Odds Ratio 12 vs 34' EXP(bx + da);
ESTIMATE 'Odds Ratio 123 vs 4' EXP(bx + db);
```

Computations of Ordinal Logistic Regression Models in PROC NLMIXED

SAS Global Forum, 2013
Paper 445-2013
Models for Ordinal Response Data

Robin High
University of Nebraska Medical Center

Ordinal Logistic Regression
Display results: Effect plots

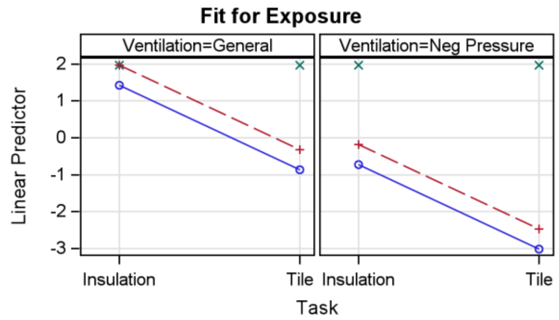
```
PROC LOGISTIC DATA=indat;
FREQ count;
CLASS x1 x2 / param=ref ;

MODEL response = x1 x2 / link=clogit aggregate scale=none;

EFFECTPLOT interaction(x=x1 plotby=x2 sliceby=response) / link;
EFFECTPLOT interaction(x=x1 plotby=x2 sliceby=response) / cluster;
EFFECTPLOT interaction(x=x1 plotby=x2 sliceby=response) / polybar;
EFFECTPLOT interaction(x=x1 plotby=x2 sliceby=response) / individual;
EFFECTPLOT interaction(x=x1 plotby=x2 sliceby=response) / individual Noconnect;

OUTPUT OUT=prd (keep=x1 x2 pred _level_ ) predicted=pred;
Run;
```

Output file prd

- Save the x data (x1 x2)
- Save pred, the cumulative probabilities (not cell specific)
- k predicted values for each combination of X values
- _level_ defines the outcome level

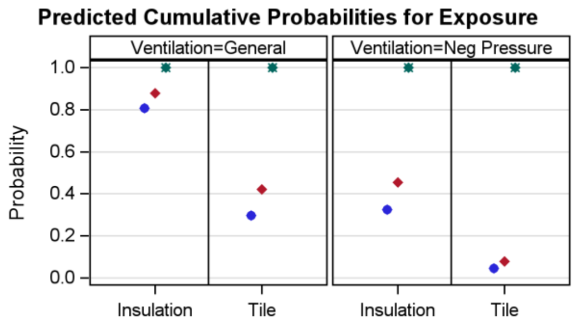
Types of Effect Plots

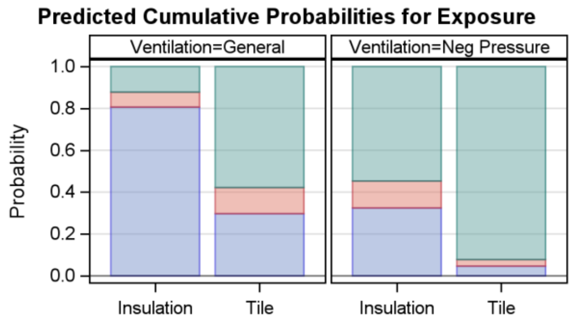
LINK



CLUSTER

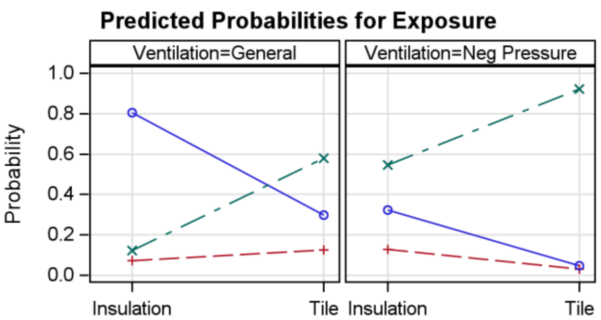
Cumulative Probabilities

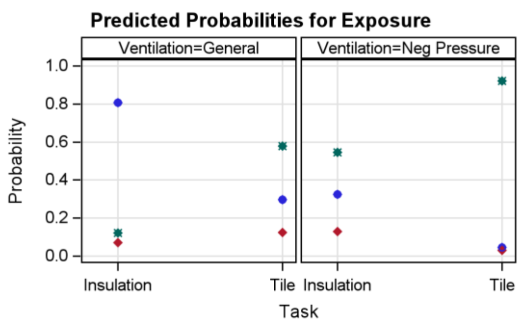




POLYBAR

Individual Probabilities





INDIVIDUAL NOCONNECT

Predicted Probabilities

```
PROC LOGISTIC;
< statements for cumulative logit >
OUTPUT OUT=prd(keep= x1 x2 _level_ pred) predicted=pred;
RUN;

* SORT by X variables and _level_ ; * _level_ is coded response 1,2,.. k;
PROC SORT DATA=prd;   BY x1 x2 _level_;

* Duplicate BY value records to one record;
DATA prd;
SET prd;
BY x1 x2 _level_;
IF first._level_ then OUTPUT;

* Cumulative Logit: transpose cumulative probabilities to one record;
PROC TRANSPOSE DATA=prd OUT=tprd(drop=_name_ _label_) prefix= _;
BY x1 x2; var pred; id _level_;

* Compute cell probabilities from predicted cumulative probs ;
DATA tprd2; SET tprd;
p1 = _1;
p2 = _2 - _1;
p3 = _3 - _2;
p4 = 1  - _3;
RUN;
```

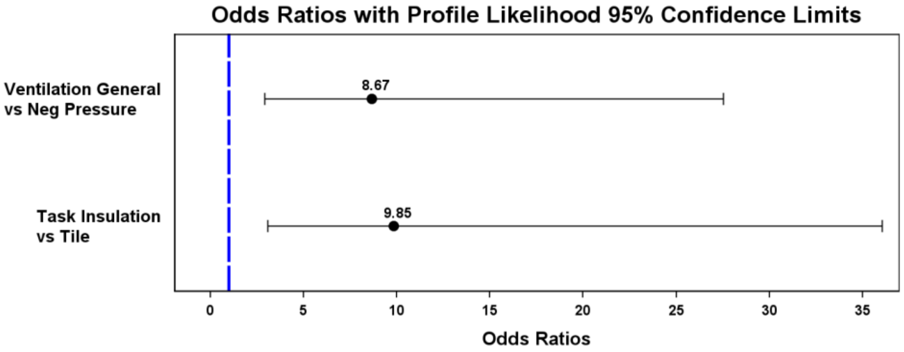
Predicted Probabilities

- Transpose to univariate layout;

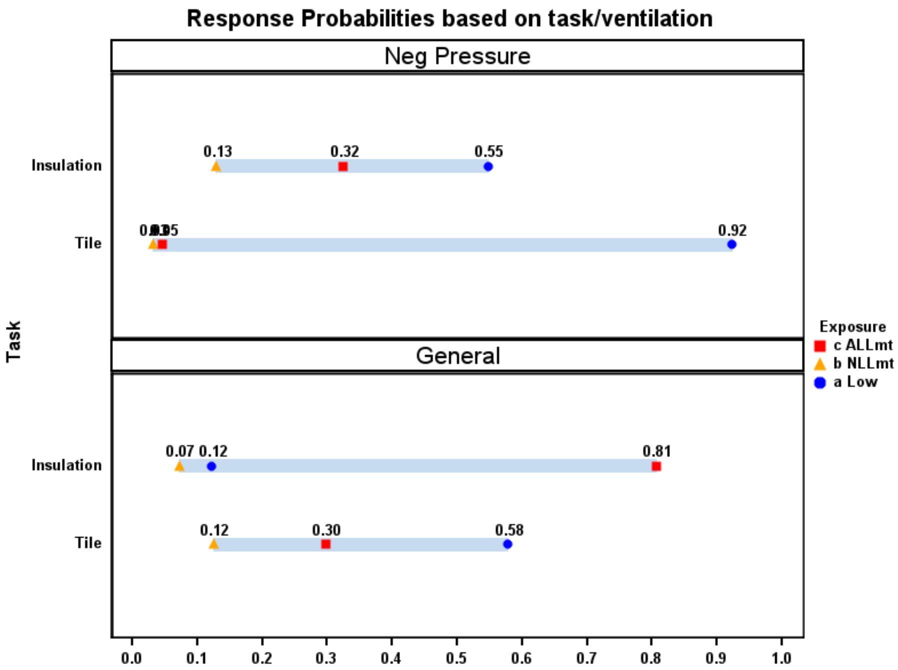
```
DATA prdplt; SET tprd2;
DROP p1 p2 p3 p4  ;
rsp=1; pct = p1;
rsp=2; pct = p2;
rsp=3; pct = p3;
rsp=4; pct = p4;
run;

PROC TABULATE DATA=prdplt;
CLASS rsp x1 x2;
VAR pct ;
TABLE x1 * x2, rsp='Response'*pct=' '*sum=' '*f=10.4
              / rts=15  ' Box="Predicted Probabilities";
RUN;
```

Ordinal LOGISTIC REGRESSION Graphical Displays

Odds Ratio (funnel plot when sorted by width of confidence interval)



Plot individual probabilities



Ordinal Logistic Regression
Adjacent Logits

```
PROC LOGISTIC DATA=indat;
CLASS x1 / param=ref;
MODEL y = x1 x2 / link = alogit ;
OUTPUT OUT=prd (keep= x1 x2 pred) predicted=pred;
RUN;
```

Save predicted values

Remove duplicates

Predicted values are the individual probabilities

Graphs of odds ratios and predicted probabilities

Example 48.4: PROC GENMOD Ice Cream Testing

Counts

	Very Good	Good	Middle	Bad	Very Bad
Brand 1	70	71	151	30	46
Brand 2	20	36	130	74	70
Brand 3	50	55	140	52	50

Row Percents

	Very Good	Good	Middle	Bad	Very Bad
Brand 1	19.0	19.3	41.0	8.2	12.5
Brand 2	6.1	10.9	39.4	22.4	21.2
Brand 3	14.4	15.9	40.3	15.0	14.4

Ice Cream Testing
Predicted Cell Probabilities

Cumulative logits

	Very Good	Good	Middle	Bad	Very Bad
Brand 1	0.186	0.196	0.405	0.112	0.100
Brand 2	0.076	0.105	0.388	0.193	0.238
Brand 3	0.135	0.161	0.419	0.143	0.141

```
Odds Ratios

1 vs 3: 1.47
2 vs 3: 0.52
```

Adjacent logits

	Very Good	Good	Middle	Bad	Very Bad
Brand 1	0.187	0.187	0.405	0.121	0.100
Brand 2	0.071	0.112	0.389	0.185	0.243
Brand 3	0.138	0.162	0.414	0.145	0.141

```
Odds Ratios

1 vs 3: 1.18
2 vs 3: 0.74
```

Non-proportional odds models (unequal slopes) works better for both models

Ordinal Logistic Regression

Which model to apply?

Both approaches fit well in similar situations and provide similar substantive results (with large cell counts)

Cumulative logit extends inference to underlying continuum (all k levels)

Adjacent logit gives effects in terms of the categories preferable to interpret given categories rather then the entire continuum

Power Calculations for Ordinal Data

Ordinal response with k levels is the experimental outcome, such as:

- Patient prognosis or outcome
- Behaviors or attitudes

Reference:

John Whitehead, (1993)
Sample Size for Ordered Categorical Data.
Statist. Med. 12: 2257 - 2271

Power Calculations for Ordinal Data Whitehead Example 1

Probabilities	Very Good	Good	Moderate	Poor	
Treatment	0.378	0.472	0.106	0.044	
Control	0.20	0.50	0.20	0.10	← known
Average	0.289	0.486	0.153	0.072	← Whitehead formula

Cumulative Probabilities	Very Good	Good	Moderate	Poor
Treatment	0.378	0.850	0.956	1.00
Control	0.20	0.70	0.90	1.00

Odds Ratio: [Very Good or Good vs Moderate or Poor]

$$= \frac{(0.850\ /\ 0.150)}{(0.70\ /\ 0.30\ )} = 2.43$$

John Whitehead formula (1993)

$N = 3*(A+1)^2 * (z_{\alpha/2} + z_{ß})^2 / [\ A * (\log(OR)^2 * (\ 1 - SUM\ p_i^3)\ ]$

$A = 1$: ratio of sample size of Treatment vs Control

$OR = 2.4286$ (odds ratio derived from cumulative probs)

$(1 - SUM\ p_i^{3\ )} = 0.857$

(p_i is the average of the probabilities in columns)

For $\alpha = 0.05$ $z_{\alpha/2} = 1.96$

power = 0.9 (ß=.1) $z_ß = 1.282$

N = 187 (or 94 in each group)

Make an Exemplary Data Set

Probabilities	Very Good	Good	Moderate	Poor
a Treatment	0.378	0.472	0.106	0.044
b Control	0.20	0.50	0.20	0.10

Counts	Very Good	Good	Moderate	Poor	Total
a Treatment	36	44	10	4	94
b Control	19	47	19	9	94

```
Cell counts derived from cell probabilities
given 94 subjects in each group
```

Analyze Data with PROC LOGISTIC: Compute Power

```
ODS SELECT nobs responseprofile oddsratios;
ODS OUTPUT globaltests=gblT(where=(substr(test,1,10)='Likelihood‘
                                or substr(test,1,5) ='Score' ));

PROC LOGISTIC DATA =counts ;
FREQ nn;
CLASS trt(ref=‘0’) / PARAM=ref; * Control level coded as 0;
MODEL rsp = trt / link=clogit ;
TITLE 'LOGISTIC: cumulative logit';
run;

* power computed from non-central chisquare ;

DATA pwr; SET gblT;
alpha=0.05 ;
c_crit = CINV(1-alpha,df);
powerC = 1-probchi(c_crit,df,ChiSq);
RUN;
```

Selected Output

Effect	Odds Ratio
Treatment vs Control	2.437

Test	ChiSq	total N	c_crit	power
Likelihood Ratio	10.25	188	3.84	0.893
Score	10.06	188	3.84	0.887

Cumulative Logit Model with k=4
Specify 4 control probs and 1 odds ratio

Four ordinal levels (k=4) provide three equations in the form

$$\text{Odds Ratio} = \text{OR} = \frac{(A / B)}{(C / D)}$$

Rearrange equation with components set equal to 0

$$((1/OR) * A * (D / C)) - B = 0$$

Solve for the unknown k=4 components of A and B under the restriction the sum of the 4 individual probabilities of A and B equals 1

Four equations, four unknowns

Ordinal Logistic Regression: Cumulative Logits

```
Row T:   oddsT   =   A   /   B
Row C:   oddsC   =   C   /   D
```

$$\text{Odds Ratio:} \quad ODRT = \frac{(A / B)}{(C / D)}$$

```
( (1/ODRT) * ( A ) *  ( (       D        ) / ( C ))) -  (       B         ) = 0
```

```
T: odds11 = p11 / (p12 + p13 + p14)  =  A / B
C: odds21 = p21 / (p22 + p23 + p24)  =  C / D
```

Odds ratio equation to solve (set equal to 0):

```
( (1/ODRT) * (p11) * ( (p22 + p23 + p24) / (p21))) - (p12 + p13 + p14)= 0
```

Repeat process for these odds ratio formulas

```
T: odds12 = (p11 + p12       )  /  (p13 + p14) = A / B
C: odds22 = (p21 + p22       )  /  (p23 + p24) = C / D

T: odds13 = (p11 + p12 + p13)  /  ( p14)   =  A / B
C: odds23 = (p21 + p22 + p23)  /  ( p24)   =  C / D
```

Verify Treatment Probabilities

```
Enter known values: p21 p22 p23 and odrt, the odds ratio

DATA cprbs;
p21=0.2; p22=0.5; p23=0.2; p24 = 1-(p21 + p22 + p23); * control group;
odrt = 2.4286;    * odds ratio ;
run;

* Four equations, four unknowns: p11 p12 p13 p14 ;
PROC MODEL DATA = cprbs;
(p11 + p12 + p13 + p14) - 1 = 0;
(( 1/odrt )*(p11            )*( (p22 + p23 + p24) / (p21            ))) - (p12 + p13 + p14)= 0;
(( 1/odrt )*(p11 + p12      )*( (      p23 + p24) / (p21 + p22      ))) - (      p13 + p14)= 0;
(( 1/odrt )*(p11 + p12 + p13)*( (            p24) / (p21 + p22 + p23))) - (            p14)= 0;
SOLVE p11 p12 p13 p14 / out=probs ;
RUN; QUIT;

proc print data=probs noobs; run;
```

p11	p12	p13	p14	p21	p22	p23	p24	odrt
0.378	0.472	0.106	0.044	0.2	0.5	0.2	0.1	2.4286

Ordinal Logistic Regression: Cumulative Logits

Given:

Odds Ratio = 1.75 and probabilities entered in row 2 for Control

Objective: Compute probabilities in row 1 for the outcomes of Treatment

Probabilities	Very Good	Good	Moderate	Poor
Treatment	0.4286	0.3747	0.1370	0.0597
Control	0.3	0.4	0.2	0.1

Make exemplary data set for various sample sizes and compute power for each

Power of OLR: cumulative logits

alpha=0.05 odds ratio = 1.75

Total N	Total N Whitehead	Power	Estimated Odds Ratio
120	117	0.378	1.74
160	148	0.456	1.72
200	195	0.565	1.74
300	284	0.727	1.73
350	356	0.820	1.76
400	404	0.865	1.76
450	448	0.897	1.75
500	493	0.922	1.75

Ordinal Logistic Regression: Adjacent Logits

```
Adjacent logit  (for respective pairs of adjacent cells)

Row T:  oddsT  =      A   /  B
Row C:  oddsC  =      C   /  D
Odds Ratio = ODRT = ( A / B ) / ( C / D )

columns
 1 and 2; ((1/odrt) * (p11 * p22) / p21 ) - p12 = 0
 2 and 3; ((1/odrt) * (p12 * p23) / p22 ) - p13 = 0
 3 and 4; ((1/odrt) * (p13 * p24) / p23 ) - p14 = 0
```

$$\boxed{p11 + p12 + p13 + p14 = 1}$$

Probabilities	Very Good	Good	Moderate	Poor
Treatment	0.438	0.389	0.130	0.043
Control	0.3	0.4	0.2	0.1

Compute Adjacent Logit Treatment Probabilities

```
Enter known control probabilities and desired odds ratio

DATA cntr_probs;
p21 = 0.3; p22 = 0.4; p23 = 0.2 ; p24 = 1 - (p21 + p22 + p23);
odrt = 1.5;
RUN;
* Solve for p11 p12 p13 p14 (four unknowns with four equations) ;

PROC MODEL DATA = cntr_probs;
(p11 + p12 + p13 + p14) - 1 = 0;
*p=3;  ((1/odrt) * (p11 * p22) / p21 ) - p12 = 0;
       ((1/odrt) * (p12 * p23) / p22 ) - p13 = 0;
*p=4;  ((1/odrt) * (p13 * p24) / p23 ) - p14 = 0;

SOLVE p11 p12 p13 p14 / out=root2 ;
RUN; QUIT;
```

Power of OLR: Adjacent Logits

Alpha=0.05, Odds Ratio = 1.5

total N	Power	Estimated Odds Ratio
100	0.484	1.55
150	0.619	1.52
200	0.743	1.52
250	0.842	1.53
300	0.900	1.53

Compute Treatment Probabilities with Prob(Y LE 2) Specified

```
Whitehead example:
 cumulative effect of treatment increases the probability
 of (Very Good or Good) from 0.7 to 0.85

* Five equations, five unknowns: estimate p11 p12 p13 p14 odrt (the odds ratio)

PROC MODEL DATA = cprbs;
(p11 + p12)             - 0.85 = 0;
(p11 + p12 + p13 + p14) - 1    = 0;

((1/odrt)*(p11          ) * ( (p22+p23+p24) /
          (p21          ) ) ) - (p12+p13+p14) = 0;

((1/odrt)*(p11+p12      ) * ( (      p23+p24) /
          (p21+p22      ) ) ) - (    p13+p14) = 0;

((1/odrt)*(p11+p12+p13) * ( (            p24) /
          (p21+p22+p23) ) ) - (           p14) = 0;

SOLVE p11 p12 p13 p14 odrt /out=root1a ;
RUN; QUIT;
```

Power with 3 Ordinal Categories two treatments, two odds ratios

```
DATA cprbs3;
p31 = .20 ; p32 = .33; p33 = 1 - (p31 + p32); * enter control probabilities ;
odrtA = 1.75;
odrtB = 2.25;    * enter desired odds ratios (non-proportional);
RUN;

PROC MODEL DATA = cprbs3;
(p11 + p12 + p13) - 1 = 0;
( (1/odrtA) * ((p11     )*(p32+p33)/(p31     )) ) - (p12+p13) = 0;
( (1/odrtB) * ((p11+p12)*(    p33)/(p31+p32)) ) - (    p13) = 0;

SOLVE p11 p12 p13 /  out=root3 ;  * solve for treatment probabilities;
RUN; QUIT;
```

odrtA	odrtB	p11	p12	p13	p21	p22	p23
1.75	2.25	0.30435	0.41295	0.28271	0.2	0.33	0.47

With PROC LOGISTIC power calculations

totalN	power	OR_1	OR_2
320	0.899	1.766	2.255

Power with 3 Ordinal Categories
Three Levels, Two Odds Ratios

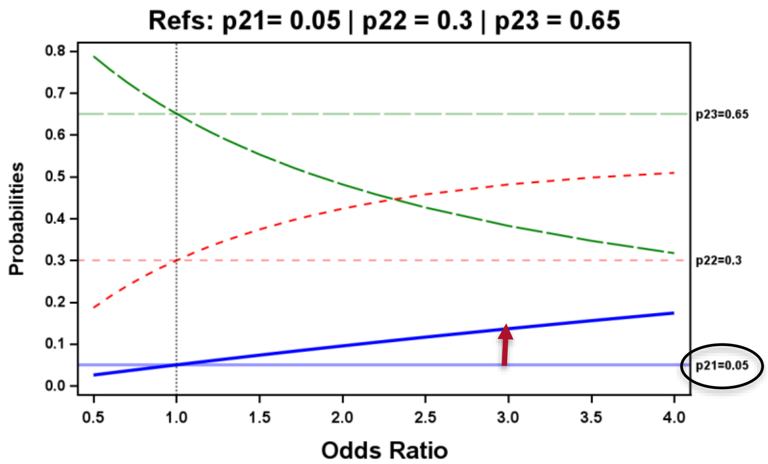
```
%LET odrtA = 1.6;
%LET odrtB = 1.3;

* x=1: treatment 1;      x=2: Treatment 2 ;     * x=3: Control;

DATA cprbs3;
p31 = .2 ; p32 = .33; p33 = 1 - (p31 + p32); * enter control probs ;
odrt1 = &odrtA. ;
odrt2 = &odrtB. ;   * enter desired odds ratios, treatments vs control;
run;

PROC MODEL DATA = cprbs3;
* Treatment 1 vs Control;
(p11+ p12+ p13) - 1 = 0;
( (1/odrt1) * ((p11    )*(p32+p33)/(p31    )) ) - (p12+p13)  = 0;
( (1/odrt1) * ((p11+p12)*(    p33)/(p31+p32)) ) - (    p13)  = 0;

* Treatment 2 vs Control;
(p21+ p22+ p23) - 1 = 0;
( (1/odrt2) * ((p21    )*(p32+p33)/(p31    )) ) - (p22+p23)  = 0;
( (1/odrt2) * ((p21+p22)*(    p33)/(p31+p32)) ) - (    p23)  = 0;

SOLVE p11 p12 p13
      p21 p22 p23 /  out=root3 ;

RUN; QUIT;
```

Results

	Y=1	Y=2	Y=3
Treatment 1	0.286	0.358	0.357
Treatment 2	0.245	0.349	0.406
Control	0.200	0.330	0.470

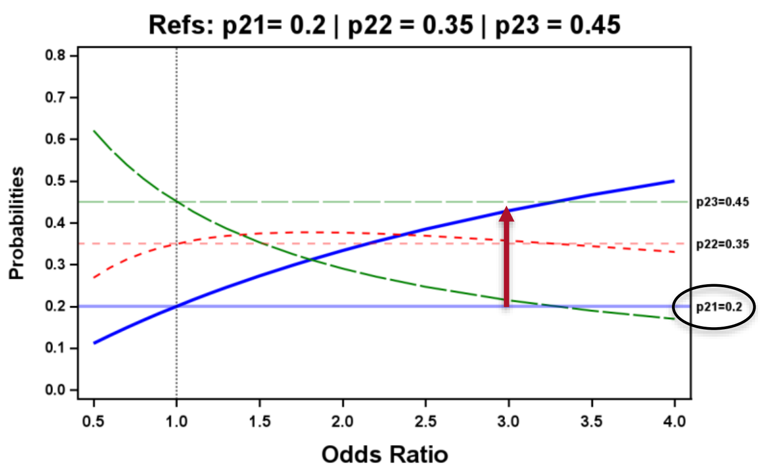
```
alpha=0.05
                   power_
total             likelihood_
  N                 ratio           OR_1        OR_2

 600                0.621           1.60        1.30
 675                0.673           1.60        1.31
 750                0.717           1.60        1.30
 900                0.803           1.60        1.30
```

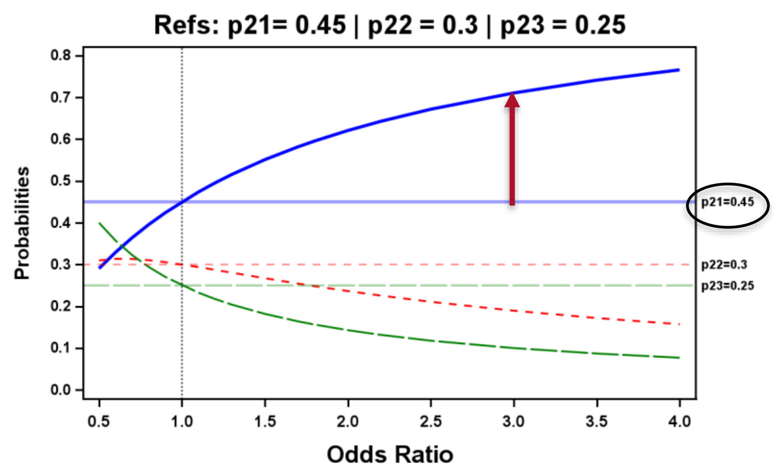
What does the odds ratio imply about the treatment effect?



Horizontal lines represent reference category probabilities

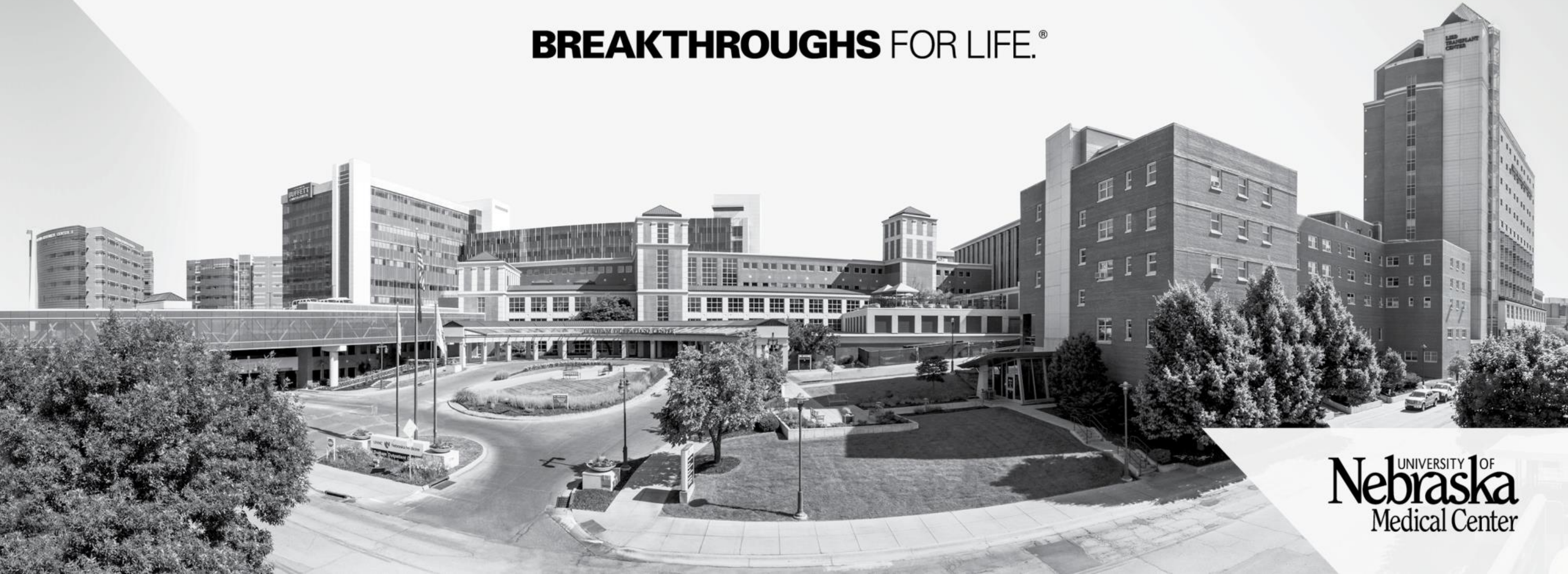
Notice length of arrow increases for p11 at an odds ratio = 3





University of Nebraska Medical Center℠

BREAKTHROUGHS FOR LIFE.®



University of Nebraska Medical Center