Moving beyond continuous outcomes: Multilevel modeling of binary, ordinal, and count outcomes

Scott C. Roesch, Ph.D.
Department of Psychology
San Diego State University
scroesch@sciences.sdsu.edu

Overview

- Brief review of multilevel regression modeling (MRM) with continuous outcomes
- Understanding Proportions (Probabilities), Odds, Odds Ratios
- MRM with
  - binary (multinomial) outcomes
  - ordinal outcomes
  - count outcomes
- Final Practical Issues
- Selected References

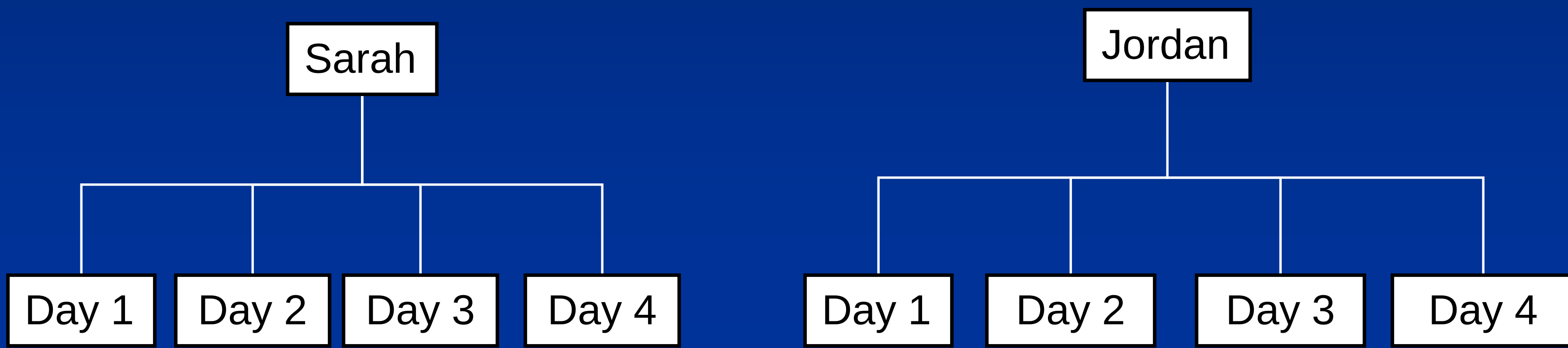
What is multilevel modeling?

- Many synonyms
  - Hierarchical linear modeling
  - Random-effects modeling
  - Mixed-effects modeling
  - Variance components modeling

- Statistical model that allows specifying and estimating relationships between variables
  - that have been observed at different levels of a hierarchical (or nested or clustered) data structure

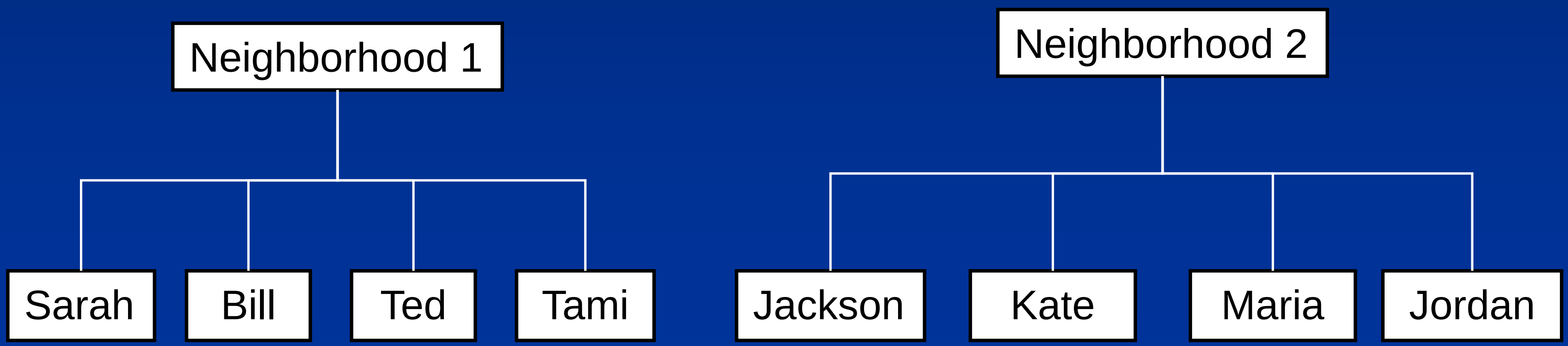
Why MRM?

- Nested data structures are everywhere
  - Time periods (or repeated observations) nested within individuals (2-level structure)



Why MRM?

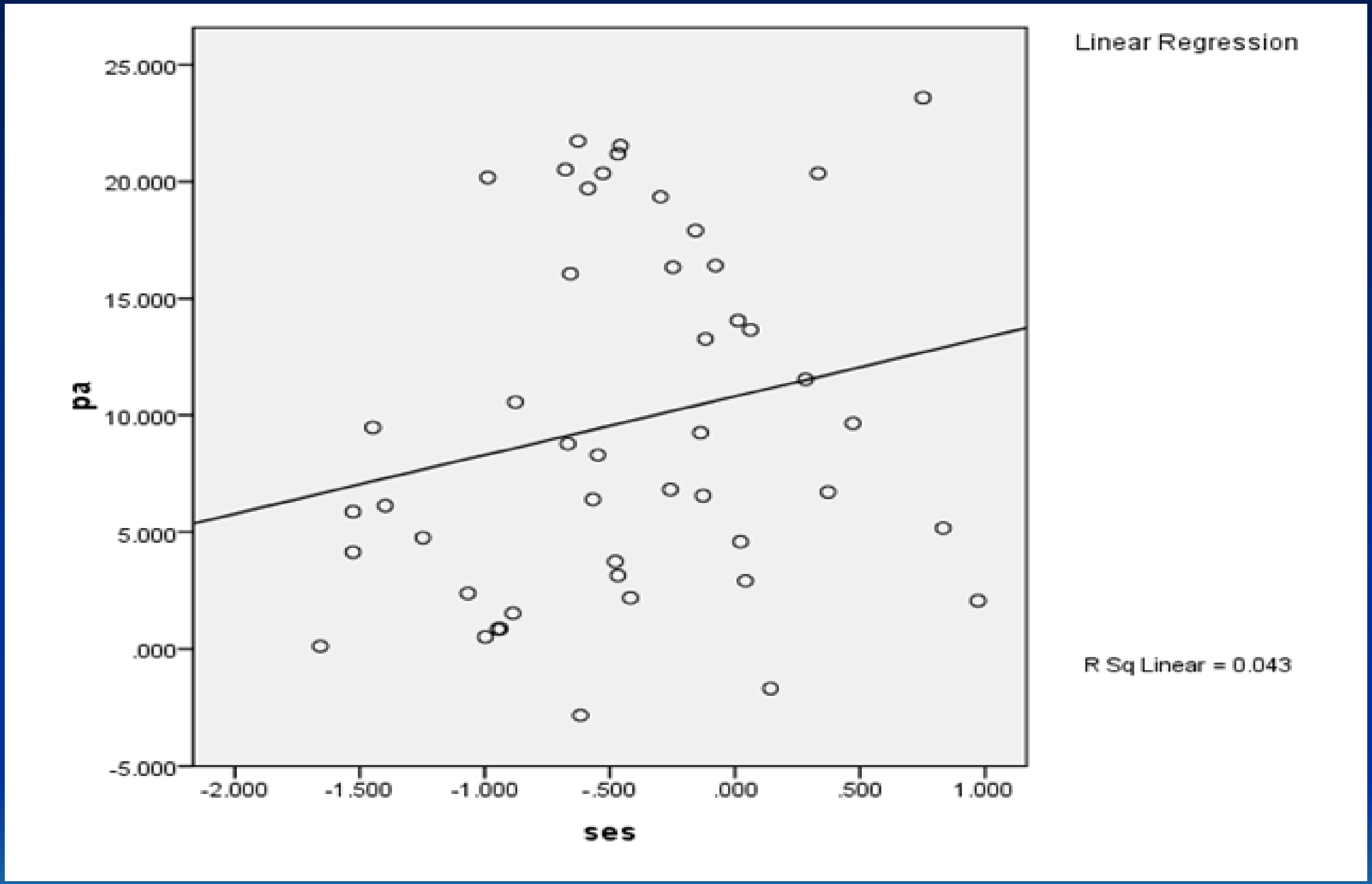
- Nested data structures are everywhere
  - Individuals nested within neighborhoods (2-level structure)



The logic of MRM extended to clusters

- Assume that we have individuals (level-1) nested in neighborhoods (level-2)
  - 1 level-1 continuous DV (physical activity [PA] Y_{ij})
  - 1 level-1 IV (SES; X_{ij})
    - $PA_{ij} = \beta_0 + \beta_1 SES_{ij} + r_{ij}$,
      - Assume that we have grand mean-centered the level-1 IV
      - β_0 is ………… β_1 is ……………………
      - $var(r_{ij}) = \sigma^2$, how different people are from their own neighborhood's regression line

Regression in a single neighborhood



$$PA' = 10.81 + 2.51(SES) + r_1$$

2 neighborhoods

- Neighborhood 1
  - $Y_i = \beta_{01} + \beta_{11}X_i + r_i$
- Neighborhood 2
  - $Y_i = \beta_{02} + \beta_{12}X_i + r_i$
- So each neighborhood has its own intercept and slope
  - this in effect serves as further "data"
  - distribution of intercepts and slopes can be summarized with
    - the mean
    - the variance relative to the mean

MRM Model

- Each level is represented by its own submodel
  - level-1 DV = PA (Y_{ij})
  - level-1 IV = individual's SES (X_{ij})
  - level-2 IV = neighborhood SES (Z_j)

- Equations for group structured data
  - Lowest *(individual)* level (level-1):
    - $Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + r_{ij}$
  - Upper *(group)* level (level-2):
    - $\beta_{0j} = \gamma_{00} + \gamma_{01}Z_j + u_{0j}$
    - $\beta_{1j} = \gamma_{10} + \gamma_{11}Z_j + u_{1j}$

The intercept-only (or empty) model

- Level-1 equation: $Y_{ij} = \beta_{0j} + r_{ij}$
  - β_{0j} = mean PA score for each neighborhood
  - r_{ij} = variance (σ^2) of each individual's PA score around the mean PA for their respective neighborhood

- level-2 equation: $\beta_{0j} = \gamma_{00} + u_{0j}$
  - γ_{00} = mean PA scores across neighborhood
    - i.e., grand mean
  - u_{0j} = variance (τ_{00}) of each neighborhood mean around the grand mean

The intercept-only model

- Intraclass correlation coefficient (variance partition coefficient) =
- $$\frac{\text{variance between groups}}{\text{variance between + variance within}}$$
- $\rho = \tau_{00} / (\tau_{00} + \sigma^2) =$
  - proportion of variance in PA between neighborhoods
    - e.g., ρ = .27 means that 27% of the variability in PA scores is between neighborhoods
    - Also refers to the intracluster correlation between two level-1 units in the same level-2 unit

Moving beyond continuous outcomes

- Categorical observed variables

Physical Activity and Gender

	n	No PA	Yes PA	Prob. PA
Female	1000	800	200	.20
Male	1000	500	500	.50
Total	2000	1300	700	

- Prob. or Risk (Yes PA) = 700/2000 = .35
  - Prob. differs by gender
- Risk Ratio or Relative Risk = .50/.20 = 2.50
  - Males are 2.50 are more likely to engage in PA than females

Moving beyond continuous outcomes

- Odds and Odds Ratios

Physical Activity and Gender

	n	PA	Prop.	Odds(Prop./1-Prop.)
Female	1000	200	.20	.25
Male	1000	500	.50	1.00

Odds Ratio (OR) = 1.00/.25 = 4

- The odds of engaging in PA (vs. not) is 4 times greater for Males (vs. Females)

McNutt et al. (2003). Estimating relative risk in cohort studies and clinical trials of common outcomes. American Journal of Epidemiology, 157, 940-943.

Osborne, J.W. (2006). Bringing balance and technical accuracy to reporting odds ratios and the results of logistic regression analyses. Practical Assessment Research & Evaluation, 11, 1-6.

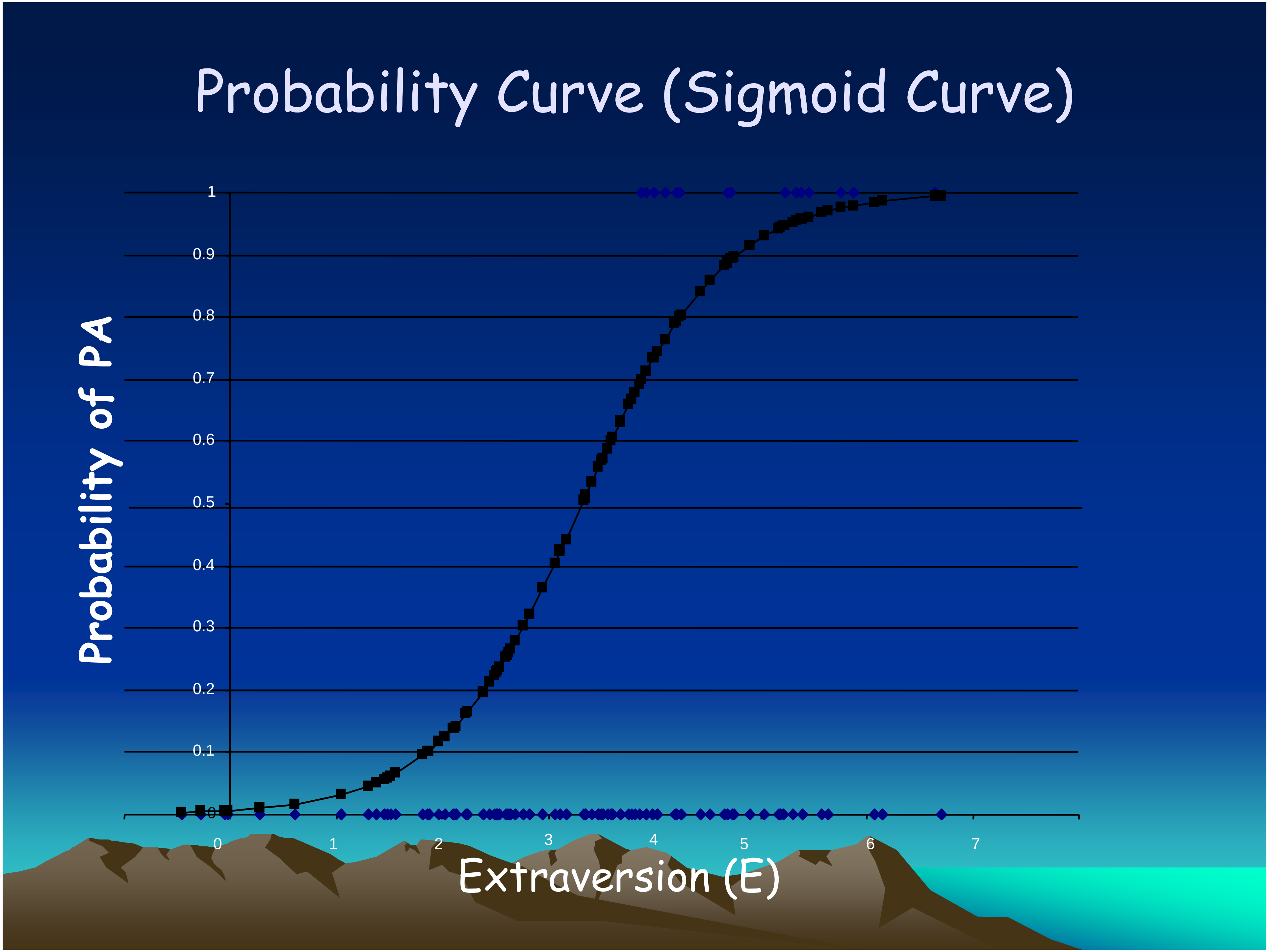
Kleinman, L.C., & Norton, E.C. (2009). What's the risk? A simple approach for estimating adjusted risk measures from nonlinear models including logistic regression. Health Services Research, 44, 288-302.

Moving beyond continuous outcomes

- Generalized linear (mixed) models ([hierarchical] generalized linear models)
  - Used when…
    - Outcomes violate OLS assumptions
      - Normality and homoscedasticity of residuals
    - Predicted outcome values will be "out of range"
    - Relationship of interest is nonlinear
- How to address these problems?
  - The link function: the log (natural)

The link function

- Binary case: Logistic regression model
  - Predicting the probability of group membership for an underlying variable (slide 16)
    - prediction not constant for full range of X
  - Log [P(y_i = 1) / 1- P(y_i = 1)] = B_0 + $B_i X_{i....}$
    - Logit [log odds] function
    - Model is linear for logits (not probabilities)
      - can convert back to probabilities by
        - Predicted Prob. = 1 / [1 + $e^{-(BO + B1X1)}$]



Probability Curve (Sigmoid Curve)
Probability of PA
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
1
2
3
4
5
6
7
Extraversion (E)

Interpreting Regression Coefficients

- Log [P(y_i = 1) / 1- P(y_i = 1)] = -0.60 + .39(E)
  - Regression coefficients interpreted as in OLS
    - Problem, the outcome is a logit value
    - Exponentiate B to get OR: exp(B) = e^B = $2.72^{.39}$
      - exp(.39) = 1.48 = OR
      - What does that mean? The case of PA
        - odds of engaging in PA (vs. not) are 1.48 greater for a 1-unit increase in E
        - <u>OR multiplier</u>: 2-unit increase E (1.48*1.48) = 2.19 odds of engaging in PA

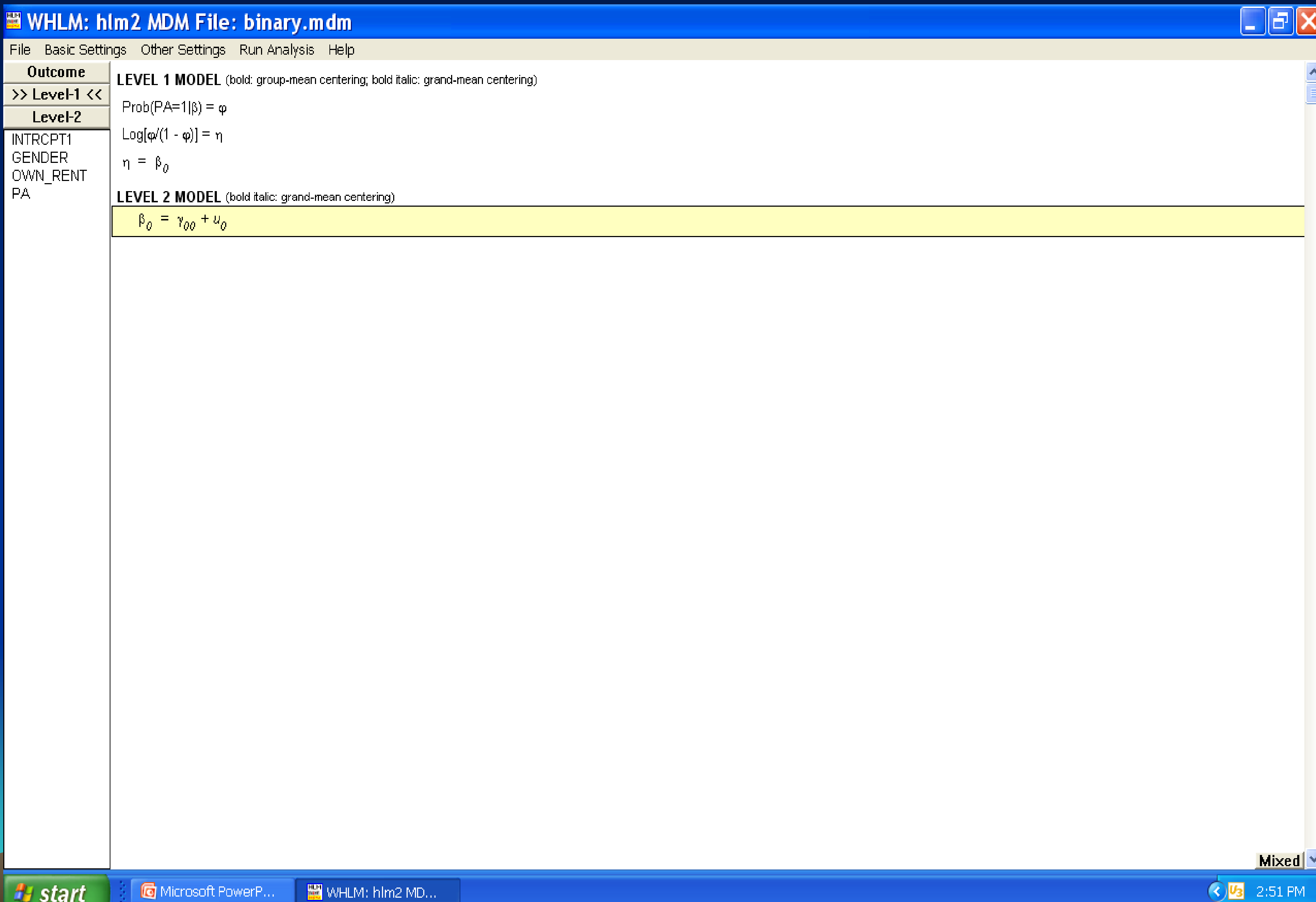
MRM: Logistic Regression Model

- Outcome of interest: PA (1=yes, 0=no)
- Assume individuals nested within neighborhoods
  - Regression equation: $Logit_{ij} = X_i\beta$
- Testing the intercept-only model
  - Allows us to gauge variation in PA across neighborhoods
    - Level-1: $Logit_{ij} = \beta_{0j}$, log-odds of PA in the jth neigh.
    - Level-2: $\beta_{0j} = \gamma_{00} + u_{0j}$
  - r_{ij} is missing from the level-1 equation

MRM: Logistic Regression Model

- Why r_{ij} is missing from the level-1 equation
  - Assume an underlying latent variable for PA
    - error structure must be fixed
  - variance of r_{ij} is assumed to have a standard logistic distribution (Mean = 0, variance = $\pi^2/3$)
    - see Snijders and Bosker (1999)
    - other methods: Goldstein et al. (2002). Partitioning variation in multilevel models. Understanding Statistics, 1, 223-231.
- Let's estimate the intercept-only model using HLM and identify parameters of interest

HLM



WHLM: hlm2 MDM File: binary.mdm
File Basic Settings Other Settings Run Analysis Help
Outcome
>> Level-1 <<
Level-2
INTRCPT1
GENDER
OWN_RENT
PA
LEVEL 1 MODEL (bold: group-mean centering; bold italic: grand-mean centering)
Prob(PA=1|β) = φ
Log[φ/(1 - φ)] = η
η = β_0
LEVEL 2 MODEL (bold italic: grand-mean centering)
$\beta_0 = \gamma_{00} + u_0$
Mixed
start
Microsoft PowerP...
WHLM: hlm2 MD...
2:51 PM

The intercept-only model

- Level-2 equation: $\beta_{0j} = \gamma_{00} + u_{0j}$
  - γ_{00} = -1.73, mean logit across neighborhoods
  - τ_{00} (variance of u_{0j}) = 1.32
    - Variance between neighborhoods around the grand mean logit (i.e., -1.73, $p < .001$)
  - Converting γ_{00} = -1.73 to a probability
    - Prob. (PA=1) = 1 / (1 + exp[-logit value])
    - Prob. (PA=*Yes*) = 1 / (1 + exp[1.73]) = .15
      - Neighborhood-wide PA rate (Prob.)

The intercept-only model

- Calculate a confidence interval (CI) to further probe neighborhood variability
  - 95% CI = mean logit ± (1.96 * √Var. logit)
  - 95% CI = -1.73 ± (1.96 * √1.32)
  - 95% CI = -3.98 to 0.52

  - Convert these to probabilities as previous
    - 95% CI = .02 to .63

The intercept-only model

- Calculate intraclass correlation coefficient

- $\rho = \tau_{00} / (\tau_{00} + \sigma^2) = 1.32 / (1.32 + \pi^2 / 3) = .29$

  - 29% of the variability in PA logit values is between neighborhoods

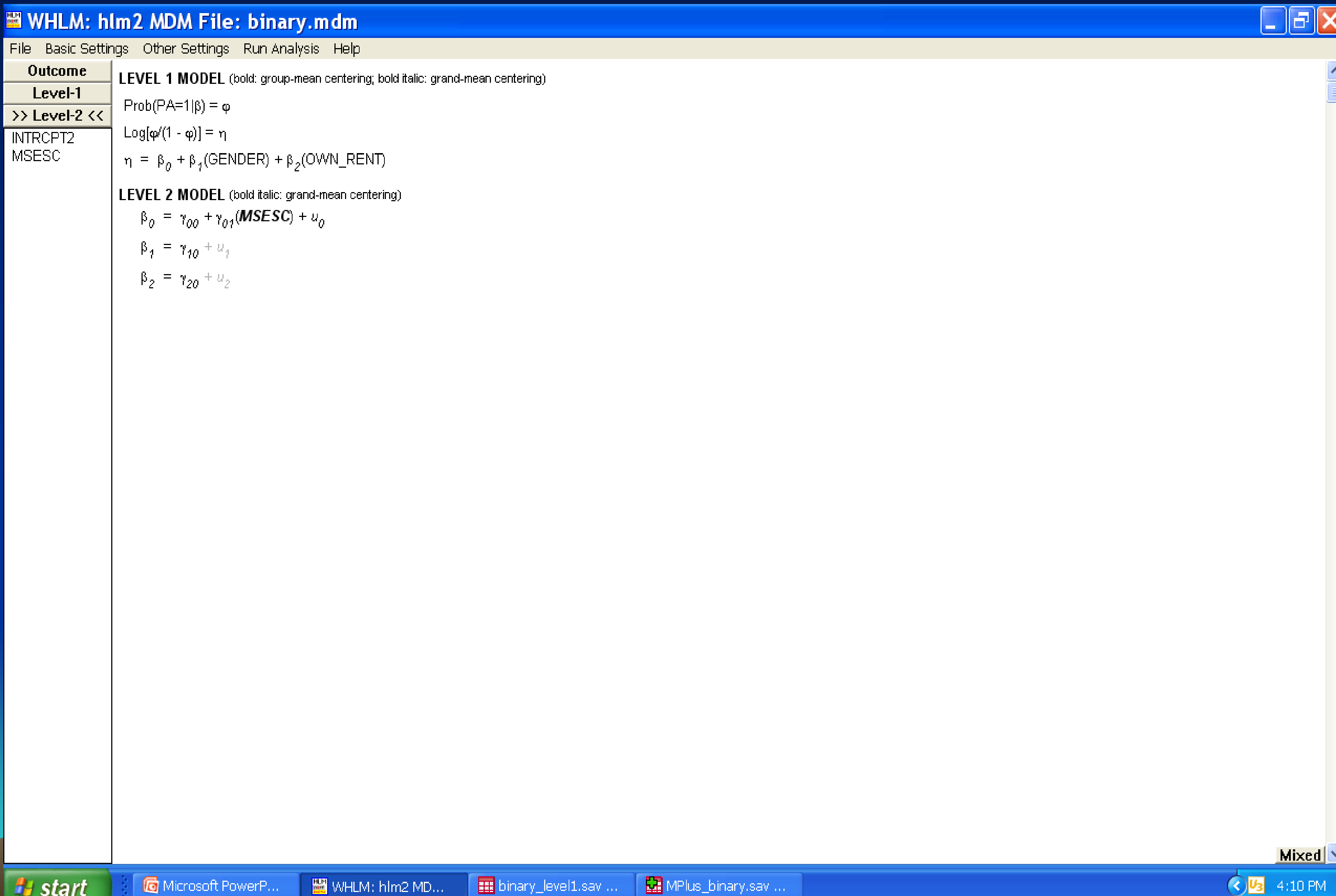
  - Other indices include median OR
    - See Merlo and Larsen in *Journal of Epidemiology and Community Health* (2003,2005,2006)

The conditional model

- Modeling variability
  - Level-1 predictors
    - Gender (0=female, 1=male)
    - Home Ownership (0=rent, 1=own)
  - Level-2 predictor
    - Neighborhood SES (grand-mean centered)

- Level-1 equation: $Logit_{ij} = \beta_{0j} + \beta_{1j} Gender_{ij} + \beta_{2j} Own_{ij}$
- Level-2 equations:
  - $\beta_{0j} = \gamma_{00} + \gamma_{01} SES_j + u_{0j}$
  - $\beta_{1j} = \gamma_{10}$ and $\beta_{2j} = \gamma_{20}$

HLM


WHLM: hlm2 MDM File: binary.mdm
File Basic Settings Other Settings Run Analysis Help
Outcome
Level-1
>> Level-2 <<
INTRCPT2
MSESC
LEVEL 1 MODEL (bold: group-mean centering; bold italic: grand-mean centering)
Prob(PA=1|β) = φ
Log[φ/(1 - φ)] = η
$\eta = \beta_0 + \beta_1(\text{GENDER}) + \beta_2(\text{OWN\_RENT})$
LEVEL 2 MODEL (bold italic: grand-mean centering)
$\beta_0 = \gamma_{00} + \gamma_{01}(\textbf{\textit{MSESC}}) + u_0$
$\beta_1 = \gamma_{10} + u_1$
$\beta_2 = \gamma_{20} + u_2$
Mixed
start
Microsoft PowerP...
WHLM: hlm2 MD...
binary_level1.sav ...
MPlus_binary.sav ...
4:10 PM

Model Results

Effect	Regression Coefficient	Adjusted OR	p-value
Intercept γ_{00}	-1.75		
SES γ_{01}	-0.28	0.75	.126
Gender γ_{10}	0.45	1.56	<.001
Own γ_{20}	-0.54	0.58	<.001

- Males and Renters more likely to engage in PA
- Interpreting ORs
  - Gender: The odds of engaging in PA (vs. not) is 1.56 times greater for males (vs. females)
  - Own: The odds of engaging in PA (vs. not) is 0.56 times less likely for homeowners (vs. renters)

MRM: Logistic Regression Model

- Converting to predicted probabilities to aid interpretation
  - We have a regression equation:

$Logit_{ij}$ = -1.75 + .45(Gender) - .54(Own) - .28(SES)…

  - Substitute predictor values in equation
    - For male, homeowner, average neighborhood SES
    - Prob. (PA=1|x) = 1 / (1 + exp[-logit value])
    - Prob. (PA=1|x) = 1 / (1 + exp[1.84]) = .14

MRM: Logistic Regression Model

- Converting to predicted probabilities to aid interpretation
  - We have a regression equation:

$Logit_{ij}$ = -1.75 + .45(Gender) - .54(Own) - .28(SES)…

  - Substitute predictor values in equation
    - For female, homeowner, average neighborhood SES
    - Prob. (PA=1|x) = 1 / (1 + exp[-logit value])
    - Prob. (PA=1|x) = 1 / (1 + exp[2.29]) = .09

Moving on: The Multinomial (Nominal) Case

- Similar to binary case in many ways
  - Multinomial logit is the link function, but now we have multiple equations
  - Assume 3 categories for the outcome
    - $\text{Log}[P(y_i = \text{category } 1) / P(y_i = \text{reference})] = X_i\beta$
    - $\text{Log}[P(y_i = \text{category } 2) / P(y_i = \text{reference})] = X_i\beta$
    - Thus, our outcome at level-1 will be the log-odds of falling into category 1 (relative to the reference category)
    - And similarly for category 2

MRM: Multinomial Logistic Regression Model

- Outcome of interest: intentions to engage in PA
  - (1=*yes*, 2=*not sure*, 3=*no*)
  - *no* category serves as the reference group
- Assume individuals nested within neighborhoods
- Testing the intercept only model
  - Level-1 Equations
    - $\text{Log}[P(y_i = \textit{yes}) / P(y_i = \textit{no})] = \beta_{0j(1)}$
    - $\text{Log}[P(y_i = \textit{not sure}) / P(y_i = \textit{no})] = \beta_{0j(2)}$

MRM: Multinomial Logistic Regression Model

- Level-2 Equations
  - $B_{0j(1)} = \gamma_{00(1)} + u_{0j(1)}$
  - $B_{0j(2)} = \gamma_{00(2)} + u_{0j(2)}$

- Estimate the model in HLM

HLM

WHLM: hlm2 MDM File: multinomial.mdm

File Basic Settings Other Settings Run Analysis Help

Outcome
>> Level-1 <<
Level-2

INTRCPT1
PA
SES_IND

LEVEL 1 MODEL (bold: group-mean centering; bold italic: grand-mean centering)

$\text{Prob}[PA(1)=1|\beta] = P(1)$

$\text{Prob}[PA(2)=1|\beta] = P(2)$

$\text{Prob}[PA(3)=1|\beta] = P(3) = 1 - P(1) - P(2)$

$\text{Log}[P(1)/P(3)] = \beta_{0(1)}$

$\text{Log}[P(2)/P(3)] = \beta_{0(2)}$

LEVEL 2 MODEL (bold italic: grand-mean centering)

For category 1

$\beta_{0(1)} = \gamma_{00(1)} + u_{0(1)}$

For category 2

$\beta_{0(2)} = \gamma_{00(2)} + u_{0(2)}$

Mixed

start | ALR2010_MLR | Calculator | WHLM: hlm2 MD... | 1:56 PM

The intercept-only model

- Level-2 equation 1: $B_{0j(1)} = \gamma_{00(1)} + u_{0j(1)}$
  - $\gamma_{00(1)} = 0.91$
    - Mean logit for saying *yes* to PA is greater than saying *no* to PA
  - $\tau_{00(1)}$ (variance of $u_{0j(1)}$) = 0.20 (p = .002)
    - Suggests statistically significant variation in logit values

The intercept-only model

- Level-2 equation 2: $B_{0j(2)} = \gamma_{00(2)} + u_{0j(2)}$
  - $\gamma_{00(2)} = -0.02$
    - Mean logit for saying *not sure* to PA is similar to saying *no* to PA
  - $\tau_{00(2)}$ (variance of $u_{0j(2)}$) = 0.04 (p = .302)
    - Suggests NO statistically significant variation in logit values between neighborhoods
      - Remove random effect $u_{0j(2)}$
      - Low likelihood that level-2 predictors will work

The intercept-only model

- Calculate ρ

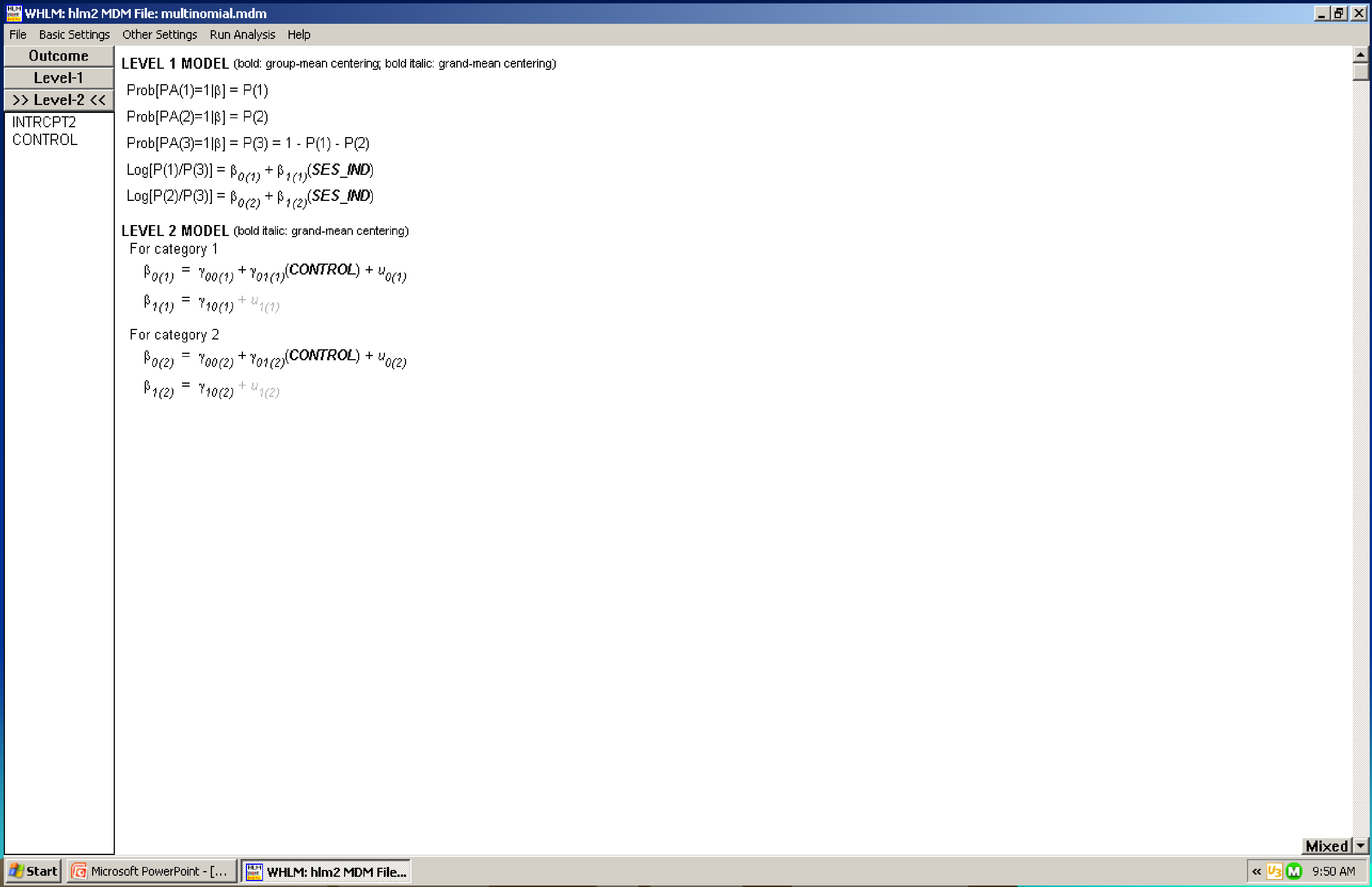
- $\rho_{(1)} = \tau_{00(1)} / (\tau_{00(1)} + \sigma^2) = 0.20 / (0.20 + \pi^2 / 3) = .06$
  - 6% of the variability in PA logit values is between neighborhoods

- $\rho_{(2)} = \tau_{00(2)} / (\tau_{00(2)} + \sigma^2) = 0.04 / (0.04 + \pi^2 / 3) = .01$
  - 1% of the variability in PA logit values is between neighborhoods

The conditional model

- Level-1 predictors
  - SES (grand mean centered)
- Level-2 predictor
  - Neighborhood control over crime (grand-mean centered)
- Same equations for both logit values
  - Level-1 equation: $Logit_{ij} = \beta_{0j} + \beta_{1j}SES_{ij}$
  - Level-2 equations:
    - $\beta_{0j} = \gamma_{00} + \gamma_{01}Control_j + u_{0j}$
    - $\beta_{1j} = \gamma_{10}$

HLM


WHLM: hlm2 MDM File: multinomial.mdm
File Basic Settings Other Settings Run Analysis Help
Outcome
Level-1
>> Level-2 <<
INTRCPT2
CONTROL
LEVEL 1 MODEL (bold: group-mean centering; bold italic: grand-mean centering)
Prob[PA(1)=1|β] = P(1)
Prob[PA(2)=1|β] = P(2)
Prob[PA(3)=1|β] = P(3) = 1 - P(1) - P(2)
Log[P(1)/P(3)] = $\beta_{0(1)} + \beta_{1(1)}$(*SES_IND*)
Log[P(2)/P(3)] = $\beta_{0(2)} + \beta_{1(2)}$(*SES_IND*)
LEVEL 2 MODEL (bold italic: grand-mean centering)
For category 1
$\beta_{0(1)} = \gamma_{00(1)} + \gamma_{01(1)}$(*CONTROL*) + $u_{0(1)}$
$\beta_{1(1)} = \gamma_{10(1)} + u_{1(1)}$
For category 2
$\beta_{0(2)} = \gamma_{00(2)} + \gamma_{01(2)}$(*CONTROL*) + $u_{0(2)}$
$\beta_{1(2)} = \gamma_{10(2)} + u_{1(2)}$
Mixed
Start
Microsoft PowerPoint - [...
WHLM: hlm2 MDM File...
9:50 AM

Model Results: Comparison 1

Effect	Regression Coefficient	Adjusted OR	p-value
Intercept $\gamma_{00(1)}$	1.08		
Control $\gamma_{01(1)}$	2.10	8.13	.001
SES $\gamma_{10(1)}$	0.40	1.49	.001

- High control over crime neighborhoods and individuals with higher SES more likely to say *yes* (relative to *no*) to engaging in PA

MRM: Multinomial Logistic Regression Model

- Converting to predicted probabilities to aid interpretation

$Logit_{ij}$ = 1.08 + 2.10(Control) + .40(SES)…

  - Substitute predictor values in equation
    - For individuals with average neighborhood control & individual SES
      - Prob. (PA=1|x) = 1 / (1 + exp[-1.08]) = .75
    - vs. 1 SD above the mean for control
      - Prob. (PA=1|x) = 1 / (1 + exp[-3.08]) = .95

Model Results: Comparison 2

Effect	Regression Coefficient	Adjusted OR	p-value
Intercept $\gamma_{00(2)}$	0.09		
Control $\gamma_{01(2)}$	0.04	1.04	.773
SES $\gamma_{10(2)}$	0.03	1.03	.830

- Predictors do not differentiate those who are *unsure* of intending to engage in PA and those who do *not* intend to engage in PA

Moving on: The Ordinal Case

- Similar to multinomial case in many ways…
  - …but preserves the *continuum* of the data
    - e.g., Likert items
    - e.g., (1=*never*, *2=sometimes*, 3=often)
  - Statistical model is the *cumulative* probability or logit model
    - Take our outcome from the previous analysis
      - Do you intend to engage in PA?
        - 1=*yes*, *2=not sure*, 3=*no*
      - Each of these outcomes can take on a probability and cumulative probability value

Moving on: The Ordinal Case

- To capture the ordered categorical nature of the data we consider cumulative logits (Clog)
- Assume our 3 ordered categories for the outcome
  - CLog[P(y_i = *yes*) / P(y_i = *not sure* & *no*)] = $X_i\beta$
  - CLog[P(y_i = *yes* & *not sure*)/ P(*no*)] = $X_i\beta$
  - Our outcomes at level-1 will be
    - the log-odds of falling into *yes* vs. the two higher categories
    - the log-odds of falling into *yes* & *not sure* vs. *no*

Moving on: The Ordinal Case

- Proportional Odds Model
  - Assumes that the effect of the predictors on the cumulative logits is identical
- Non-Proportional Odds Model relaxes this assumption
  - Similar to a multinomial case

- We will assume Proportional Odds in the current example
- Level-2 variance assessment the same as previous

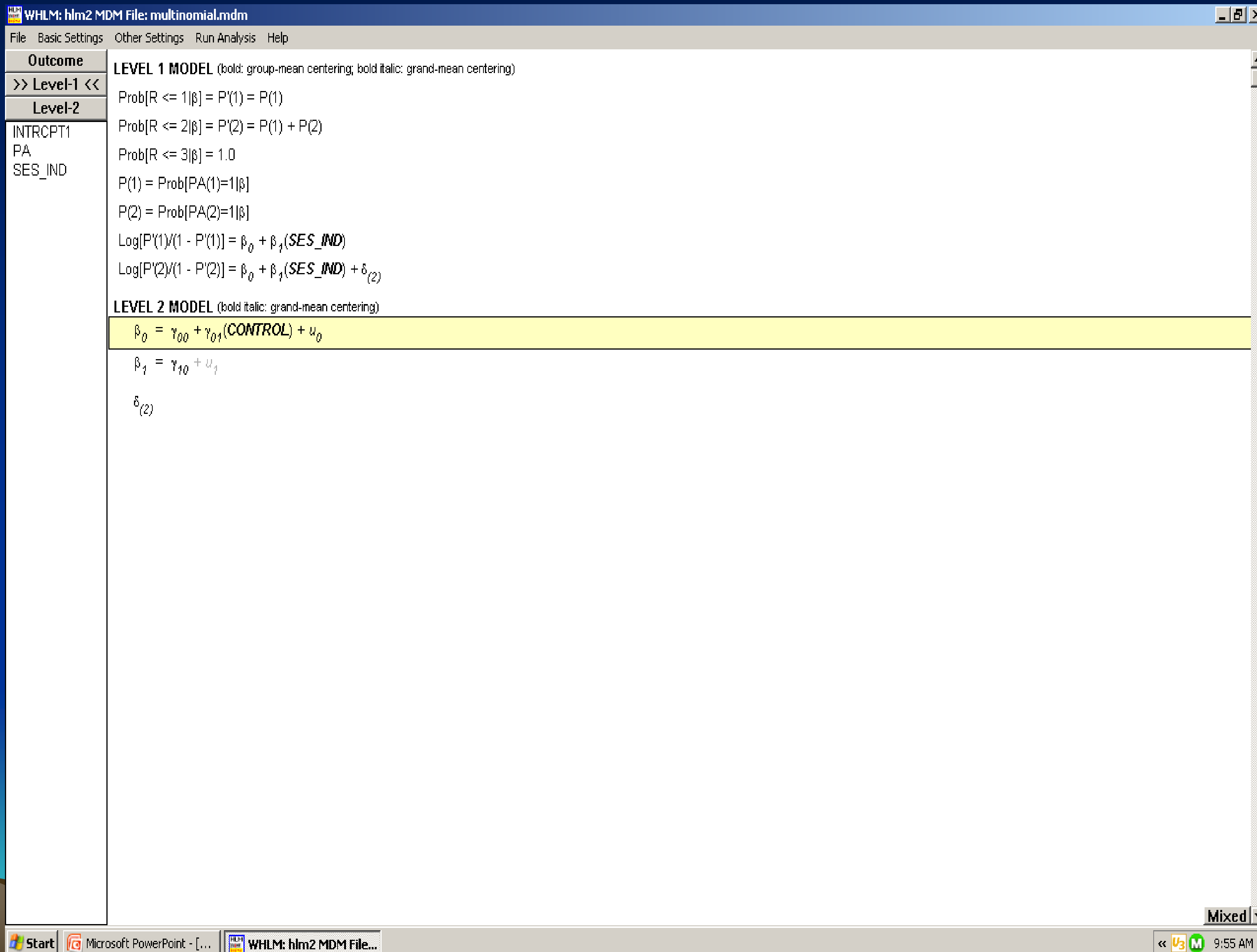
MRM: Ordinal Regression Model

- Outcome of interest: intentions to engage in PA
  - (1=*yes*, 2=*not sure*, 3=*no*)
- Assume individuals nested within neighborhoods
- Testing a conditional model
  - Level-1 predictors
    - SES (grand mean centered)
  - Level-2 predictor
    - Neighborhood control over crime (grand-mean centered)

The conditional model

- Same equations for both cumulative logit equations
  - Level-1 equation: $CLog_{ij} = \beta_{0j} + \beta_{1j}SES_{ij}$
  - Level-2 equation:
    - $\beta_{0j} = \gamma_{00} + \gamma_{01}Control_j + u_{0j}$
    - $\beta_{1j} = \gamma_{10}$
- Thresholds
  - Assumes a latent continuous variable underlies the outcome
  - These "cut-points" are intercept terms for each CLog equation

HLM


WHLM: hlm2 MDM File: multinomial.mdm
File Basic Settings Other Settings Run Analysis Help
Outcome
>> Level-1 <<
Level-2
INTRCPT1
PA
SES_IND
LEVEL 1 MODEL (bold: group-mean centering; bold italic: grand-mean centering)
Prob[R <= 1|β] = P'(1) = P(1)
Prob[R <= 2|β] = P'(2) = P(1) + P(2)
Prob[R <= 3|β] = 1.0
P(1) = Prob[PA(1)=1|β]
P(2) = Prob[PA(2)=1|β]
Log[P'(1)/(1 - P'(1))] = $\beta_0 + \beta_1$(*SES_IND*)
Log[P'(2)/(1 - P'(2))] = $\beta_0 + \beta_1$(*SES_IND*) + $\delta_{(2)}$
LEVEL 2 MODEL (bold italic: grand-mean centering)
$\beta_0 = \gamma_{00} + \gamma_{01}$(*CONTROL*) + u_0
$\beta_1 = \gamma_{10} + u_1$
$\delta_{(2)}$
Mixed
Start
Microsoft PowerPoint - [...
WHLM: hlm2 MDM File...
9:55 AM

Model Results

Effect	Regression Coefficient	Adjusted OR	p-value
Control γ_{01}	1.54	4.67	.001
SES γ_{10}	0.35	1.42	<.001

- High control over crime neighborhoods and individuals (within neighborhoods) with higher SES more likely to say *yes* (relative to *not sure* & *no*) to PA
  - And for *yes* and *not sure* relative to *no*

- Values from the regression equation can be converted to predicted probabilities as before

MRM: Ordinal Regression Model

- Other ordered logit models
  - Stage/Continuation Ratio Approach
    - "yes" vs. "not sure & no"
    - "not sure" vs. "no"
  - Adjacent Category Approach
    - "yes" vs. "not sure"
    - "not sure" vs. "no"

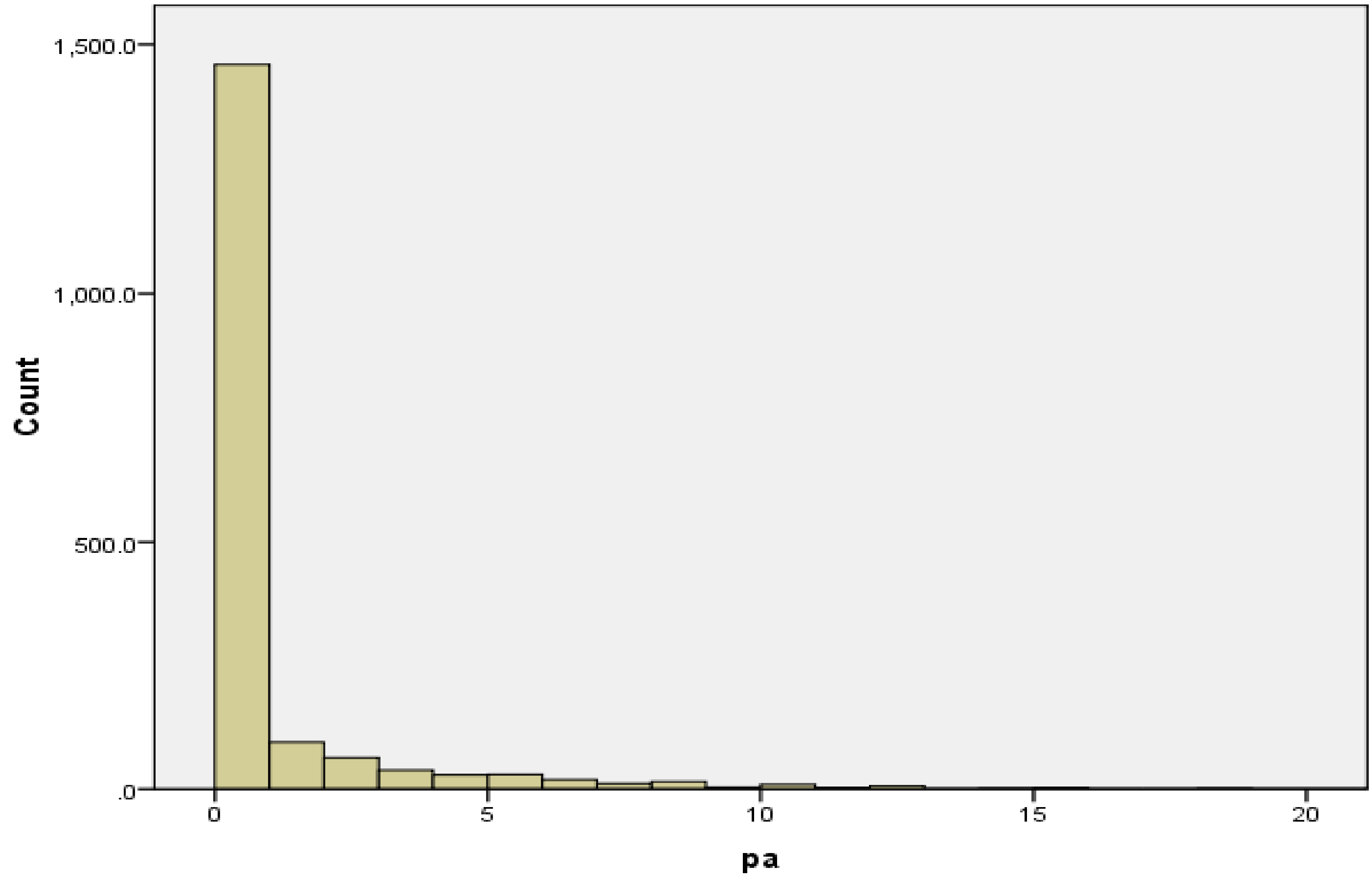
Fullerton, A.S. (2009). A conceptual framework for ordered logistic regression models. *Sociological Methods & Research, 38*, 306-347.

Liu, I., & Agresti, A. (2005). The analysis of ordered categorical data: An overview and a survey of recent developments. Test, 14, 1-73.

Moving on: Poisson MRM

- Seeks to model count variables
  - Nonnegative integers
  - Often have many zeros
    - Positively skewed distribution
  - # days engaged in PA over a 30 day period
    - Average number of days engaged in PA over this period (λ; rate parameter)
  - The natural log (ln) of the target "event" is the link function
    - $\ln(\lambda') = B_0 + B_1(\text{Extraversion}[E])\ldots$,
      - where λ'=predicted count variable

Poisson Distribution



Interpreting Regression Coefficients

- $\ln(PA') = 1.32 + .35(E)$
  - Like OLS regression interpreting $\ln(PA')$

  - Like logistic regression, exponentiating the equation/terms is helpful
    - $e^{\ln(PA')} = e^{(1.32 + .35[E])}$
    - $e^{\ln(PA')} = PA' \rightarrow PA' = e^{(1.32 + .35[E])}$
      - outcome is now in the original metric
    - Working on the right-side

$$e^{(1.32 + .35[E])} = e^{1.32}e^{.35(E)}$$

Interpreting Regression Coefficients

- $PA' = e^{1.32}e^{.35(E)}$
  - $\exp(B_0) = \exp(1.32) = 3.75$
    - Predicted days of PA for a person of average E (assume E was grand-mean centered)
  - $\exp(B_1) = \exp(.35) = 1.42$
    - Event (or Rate or Incidence) Ratio
      - Predicted <u>multiplicative</u> effect of a 1-unit change in E on days of PA
        - e.g., a 4 (relative to a 3) on E will engage in PA 1.42 times more

Interpreting Regression Coefficients

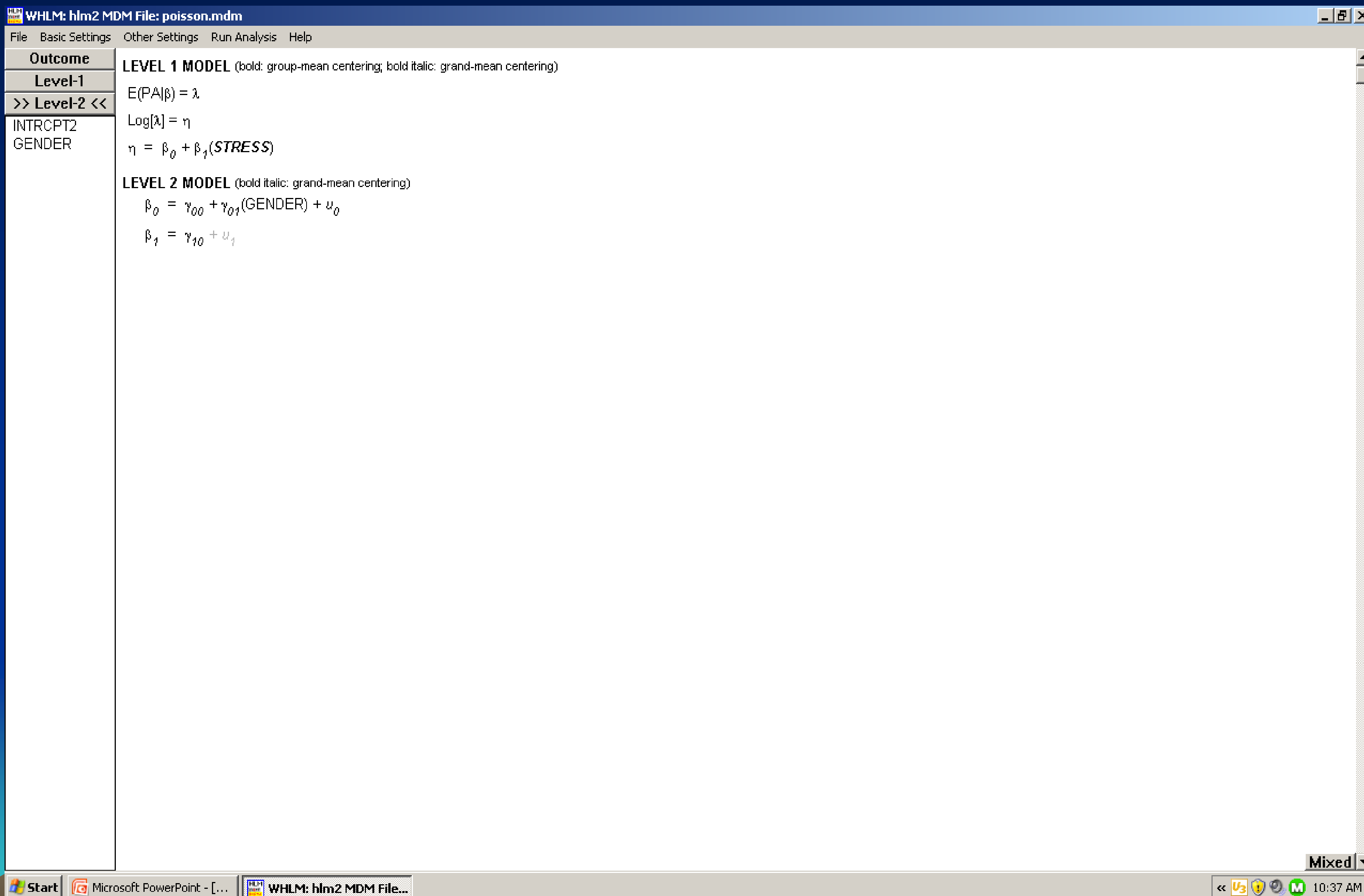
- Let's take a further look at this issue
  - $PA' = e^{B0 + B1(E)...}$
    - $PA' = e^{1.32 + .35(E)}$, E is grand-mean centered
    - Substitute values for E to get predicted number of events (days of PA)
      - e.g., $PA' = e^{1.32 + .35(4)} = 15.21$
      - e.g., $PA' = e^{1.32 + .35(3)} = 10.71$
        - 10.71 * 1.42 = 15.21

The conditional model

- Modeling variability
  - Level-1 predictors
    - Stress (continuous measure)
  - Level-2 predictor
    - Gender (0=female, 1=male)

- Level-1 equation: $\ln(PA)_{ij} = \beta_{0j} + \beta_{1j}Stress$
- Level-2 equations:
  - $\beta_{0j} = \gamma_{00} + \gamma_{01}Gender_j + u_{0j}$
  - $\beta_{1j} = \gamma_{10}$

HLM


WHLM: hlm2 MDM File: poisson.mdm
File Basic Settings Other Settings Run Analysis Help
Outcome
Level-1
>> Level-2 <<
INTRCPT2
GENDER
LEVEL 1 MODEL (bold: group-mean centering; bold italic: grand-mean centering)
$E(PA|\beta) = \lambda$
$Log[\lambda] = \eta$
$\eta = \beta_0 + \beta_1(STRESS)$
LEVEL 2 MODEL (bold italic: grand-mean centering)
$\beta_0 = \gamma_{00} + \gamma_{01}(GENDER) + u_0$
$\beta_1 = \gamma_{10} + u_1$
Mixed
Start
Microsoft PowerPoint - [...
WHLM: hlm2 MDM File...
10:37 AM

Model Results

Effect	Regression Coefficient	exp(B)	p-value
Intercept γ_{00}	0.46		
Gender γ_{01}	1.43	4.18	.013
Stress γ_{10}	-1.17	0.31	.001

- Males and individuals with lower stress are more likely to engage in PA

- Exponentiating regression coefficients gives us event (incident or rate) ratio

Model Results

- Further Interpretation
  - Gender: Males on average engage in PA 4.18 times more than females, holding stress constant
    - PA' (male) = $e^{0.46 + 1.43(1) - 1.17(0)}$ = 6.63
    - PA' (female) = $e^{0.46 + 1.43(0) - 1.17(0)}$ = 1.58
    - The Multiplier = 1.58 * 4.18 = 6.63
  - Stress: 1-unit change in stress associated with a .31 times change in PA, holding gender constant
    - PA' (stress=1) = $e^{0.46 + 1.43(0) - 1.17(1)}$ = 0.49
    - PA' (stress=0) = $e^{0.46 + 1.43(0) - 1.17(0)}$ = 1.58

Poisson MRM

- Assessing variability
  - Depends on the program used
    - Programs like HLM and Mplus do NOT estimate level-1 variance
      - Options
        - Assume a normal distribution for level-1 residuals
        - Use a simulation method
        - Assume a level-1 Poisson distribution with a specific mean/variance
        - Use statistical significance test and CI for level-2 variability

Poisson MRM

- Constant exposure vs. Variable exposure
  - Counts (Events) per unit time or population size is the outcome (offset)
    - becomes a rate
- Overdispersion
  - Variance > Mean
  - Largely influences standard errors
  - Use either
    - Overdispersed Poisson Model (Φ)
    - Negative Binomial (Φ plus other Poissons)

MRM: Zero-Inflated Poisson (ZIP) Regression Model

- Mixture of logistic and Poisson regression models
  - Used when
    - there are "excess 0s" for the Poisson
    - there are two ways that 0s can be generated
      - structural: those that will never engage in PA
      - regular Poisson 0s: those that will engage, but did not in the time-interval
    - Issue becomes finding predictors that differentiates these two groups

MRM: Zero-Inflated Poisson Regression Model

- ZIP models explore the prediction of latent groups ("always 0 group" vs. "not always 0 group")
  - Logit is used to model this latent binary outcome
    - "zero" class (coded 1)
- Part II: Poisson model
  - Regular Poisson without the excess 0s
- Can have different predictors for each equation

MPlus

```
TITLE: Multilevel ZIP regression

DATA:
       FILE is C:\Scott\Presentations\ALR2010\MPlus_poisson.dat;
       FORMAT is FREE;

VARIABLE:
       NAMES are id gender stress prior_pa pa;

       USEVARIABLES ARE pa gender stress prior_pa;

       MISSING ARE ALL (-999);

       COUNT IS pa (i);

       WITHIN =  stress prior_pa;

       BETWEEN = gender;

       CENTERING = GRANDMEAN (stress);

       CLUSTER is id;

ANALYSIS:
        TYPE = TWOLEVEL RANDOM;

MODEL:
        %WITHIN%
        pa ON stress;
        pa#1 ON stress prior_pa;

        %BETWEEN%
        pa ON gender;
        pa#1 ON gender;
```

Model Results: Logistic Model

Effect	Regression Coefficient	exp(B)	p-value
Intercept γ_{00}	1.28		
Gender γ_{01}	-0.28	0.76	.047
Prior PA γ_{10}	-1.23	0.29	.001
Stress γ_{20}	0.36	1.43	.023

Note. Prior PA (0=no, 1=yes)

- Females, individuals with no previous experience of formal PA, and individuals with higher stress are more likely in the "zero class"
- Interpret as in logistic MRM (Logit[PA])

Model Results: Poisson Model

Effect	Regression Coefficient	exp(B)	p-value
Intercept γ_{00}	1.42		
Gender γ_{01}	0.12	1.13	.074
Stress γ_{10}	-0.22	0.80	.014

- Higher stress, lower PA; but gender has no statistically significant effect

- Interpret like you would for Poisson MRM
  - ln(PA')

Practical Issues

- Model fit
  - -2 Log Likelihood (-2LL, deviance, likelihood ratio tests)
    - small values indicate better fit
    - Relative model fit for nested models

$$\text{“}R^2\text{”} = 1 - \frac{\text{deviance(fitted model)}}{\text{deviance(intercept-only model}}$$

  - Non-nested models (e.g., Poisson vs. ZIP)
    - Vuong (V) statistics
    - Akaike or Bayesian Information Criterion

Practical Issue

- Unit-specific vs. population-average models
  - All models with nonlinear link functions can be estimated via these two methods
  - When to use each…
    - Unit-specific (conditional models):
      - Estimates conditional on random effects
    - Population-average (marginal models):
      - Estimation based on average across random effects (in essence, ignoring them)

Practical Issues

- Centering (see Raudenbush & Bryk [2002], Enders [2007] Psychological Methods)
  - Uncentered, grand-mean, group-mean
- Software
  - HLM, Mplus, SuperMix, R, Stata, MLWin, SAS, etc.
  - See for a review:

Roberts, J.K., & McLeod, P. (2008). Software options for multilevel modeling. In A.A. O'Connell & D.B. Mc C oach (Eds.), Multilevel Modeling of Educational Data (pp. 427-467). Charlotte, NC : Information Age Publishing.

References

General references for nonlinear outcomes

Hedeker, D. & Gibbons, R.D. (2006). Longitudinal Data analysis. New York: Wiley.

Raudenbush, S.W. , & Bryk, A.S. (2002). Hierarchical linear models: Applications and data analysis methods (2nd Ed.). Thousand Oaks, CA: Sage.

Snijders, T.A.B., & Bosker, R.J. (1999). Multilevel analysis: An introduction to basic and advanced multilevel modeling London, England: Sage.

References

General Categorical and Ordinal Models

Hedeker, D. (2007). Multilevel models for ordinal and nominal variables. In J. de Leeuw & E. Meijer (Eds.), Handbook of multilevel analysis. New York: Springer.

Hedeker, D., et al. (1999). The thresholds of change model: An approach to analyzing stages of change data. Annals of Behavioral Medicine, 21, 61-70.

O'Connell, A.A. et al. (2008). Multilevel logistic models for dichotomous and ordinal data. In A.A. O'Connell & D.B. McCoach (Eds.), Multilevel Modeling of Educational Data (pp. 199-242). Charlotte, NC. Information Age Publishing.

References

Count Models

Coxe, S., et al. (2009). The analysis of count data: A gentle introduction to Poisson regression and its alternatives. Journal of Personality Assessment, 91, 121-136.

Lee, A.H., et al. (2006). Multi-level zero-inflated Poisson regression modelling of correlated count data with excess zeros. Statistical Methods in Medical Research, 15, 47-61.

Min, Y., & Agresti, A. (2005). Random effects models for repeated measures of zero-inflated count data. Statistical Modeling, 5, 1-19.

Poisson Distribution

- Probability distribution for the Poisson is:

$$\text{Prob.}(PA=y) = \frac{e^{-\lambda} * \lambda^{y}}{y!}$$

- Assume the rate parameter λ = 4

$$\text{Prob.}(Y=3) = \frac{e^{-4} * 4^{3}}{3!} = .19$$