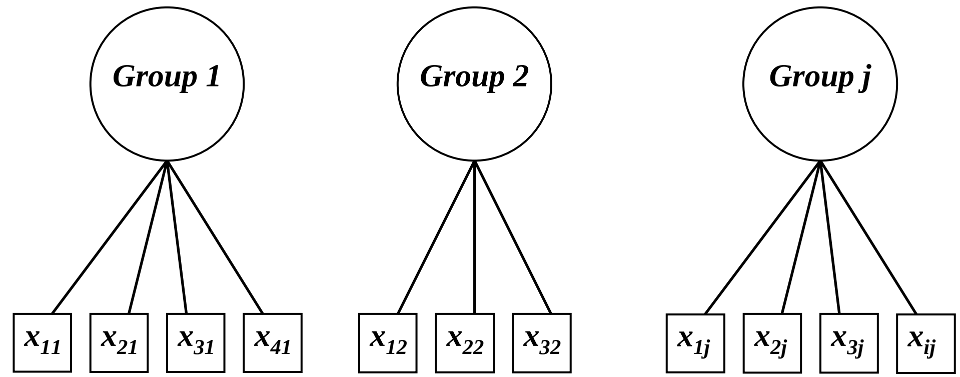
Longitudinal Data Analysis Using Multilevel Modeling

Presented by: Danny Huang

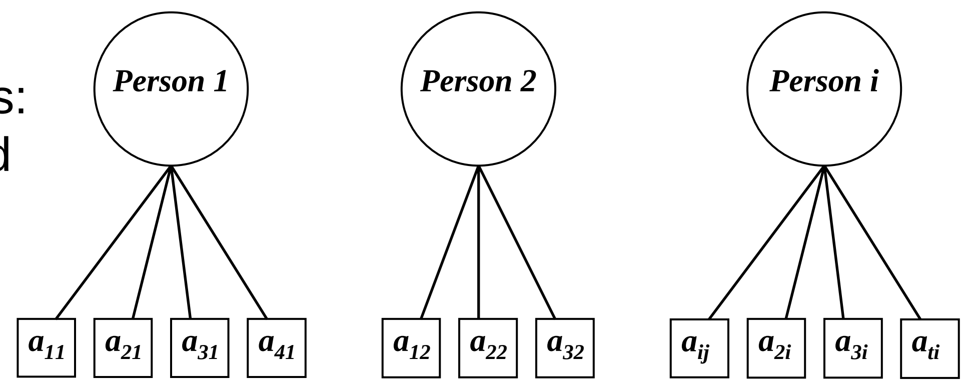
October 13, 2004

Nested Data Structure

Individuals nested within groups



Repeated Measures: observations nested within person



Repeated Measures ANOVA

Repeated Measures ANOVA:

- ❖ Tests whether the means are equal across occasions.
- ❖ Time is treated as ordered categories.
- ❖ Use contrasts to examine linear and polynomial trends.

Limitations:

- ➢ Variance and covariance of the repeated measures are assumed to be equal.
- ➢ Fixed effect of “time”.
- ➢ Missing data handling (listwise deletion).

Uni-level Regression

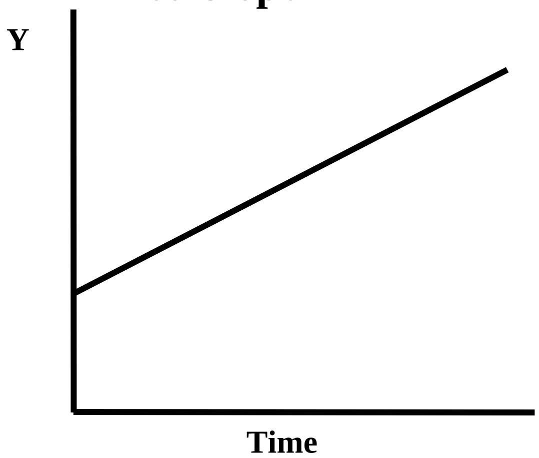
$$Y_i = \beta_0 + \beta_1 (Time) + e_i$$

Limitations:

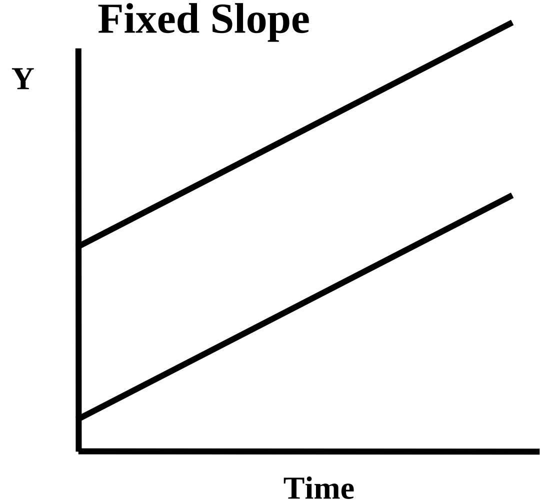
- Ignores nested data structure.
- Violation of data independence assumption.
- Assume equal (homogeneous) treatment effect: One slope for person.

Random vs. Non-Random

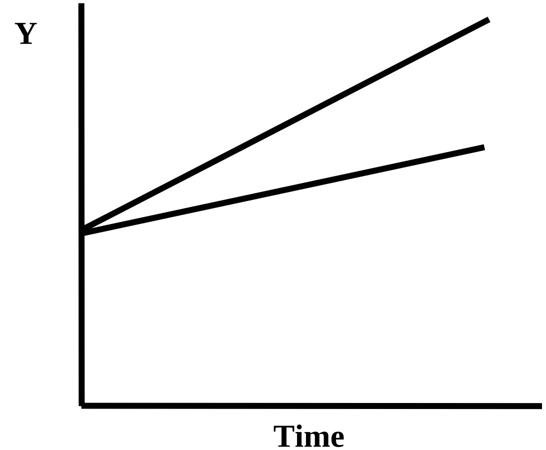
Fixed Intercept
Fixed Slope



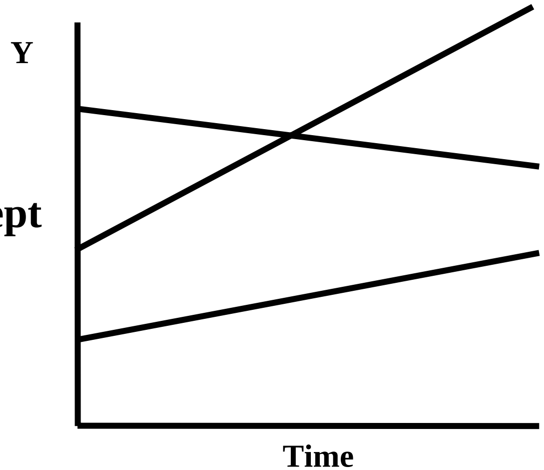
Random Intercept
Fixed Slope



Fixed Intercept
Random Slope



Random Intercept
Random Slope



Multilevel Model (Random Coefficient Model)

Level 1 (occasion level):

$$Y_{ti} = \pi_{0i} + \pi_{1i}(Time_{ti}) + e_{ti} \qquad e_{ti} \sim N(0, \sigma^2)$$

Y_{ti} = *outcome for person i at time t*
π_{0i} = *initial status for person i at time 0.*
π_{1i} = *growth rate for person i.*
$Time_{ti}$ = *time Variable*
e_{ti} = *residual for person i at time t*

Level 2 (person level):

$$\pi_{0i} = \beta_{00} + \mu_{0i} \qquad \mu_{0i} \sim N(0, \tau_{00})$$
$$\pi_{1i} = \beta_{10} + \mu_{1i} \qquad \mu_{1i} \sim N(0, \tau_{11})$$

β_{00} = *mean initial status (intercept)*
β_{10} = *mean growth rate (slope)*
μ_{0i} = *random component for initial status* (π_{0i})
μ_{0i} = *random component for growth rate* (π_{1i})

Multilevel Model (Intercept & Slope as Outcome)

Level 1 (occasion level):

$$Y_{ti} = \pi_{0i} + \pi_{1i}(Time_{ti}) + \pi_{1i}(X_{ti}) + e_{ti} \qquad e_{ti} \sim N(0, \sigma^2)$$

Level 2 (person level):

$$\pi_{0i} = \beta_{00} + \beta_{01}(Z_i) + \mu_{0i} \qquad \mu_{0i} \sim N(0, \tau_{00})$$

$$\pi_{1i} = \beta_{10} + \beta_{11}(Z_i) + \mu_{1i} \qquad \mu_{1i} \sim N(0, \tau_{11})$$

X_{ti} = *time varying covariate*
Z_i = *time invariant covariate*
β_{00} = *mean initial status (intercept)*
β_{01} = *effect of* Z_i *on initial status*
β_{10} = *mean growth rate (slope)*
β_{11} = *effect of* Z_i *on growth rate (slope)*
μ_{0i} = *random component for initial status* (π_{0i})
μ_{1i} = *random component for growth rate status* (π_{1i})

Variance Explained

1. Unconditional Model

$$Y_{ti} = \pi_{0t} + \pi_{1i}(Time_{ti}) + e_{ti} \qquad e_{ti} \sim N(0, \sigma^2)$$

$$\pi_{0i} = \beta_{00} + \mu_{0j} \qquad \mu_{0i} \sim N(0, \tau_{00})$$

$$\pi_{1i} = \beta_{10} + \mu_{1j} \qquad \mu_{1i} \sim N(0, \tau_{11})$$

2. Conditional Model

$$Y_{ti} = \pi_{0t} + \pi_{1i}(Time_{ti}) + e_{ti} \qquad e_{ti} \sim N(0, \sigma^2)$$

$$\pi_{0i} = \beta_{00} + \beta_{01}(Z_i) + \mu_{0j} \qquad \mu_{0i} \sim N(0, \tau_{00})$$

$$\pi_{1i} = \beta_{10} + \beta_{11}(Z_i) + \mu_{1j} \qquad \mu_{1i} \sim N(0, \tau_{11})$$

3. The proportion of variance explained in π_q by the model:

$$\frac{\hat{\tau}_{qq(unconditional)} - \hat{\tau}_{qq(conditional)}}{\hat{\tau}_{qq(unconditional)}}$$

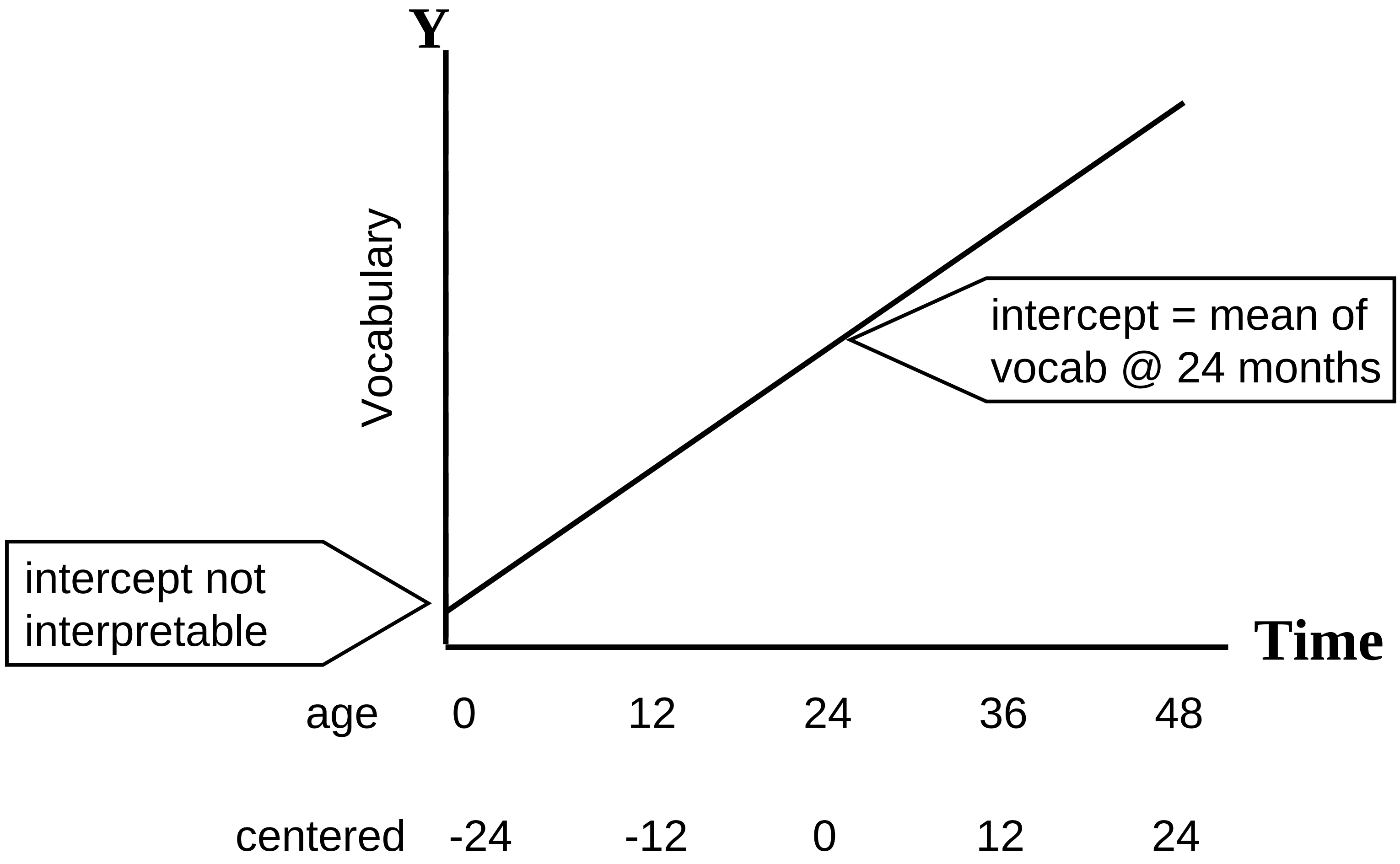
Advantages of Multilevel Modeling

- Different growth curve for each person.
- Number of repeated measures may vary across persons.
- Spacing between repeated measures may vary across persons.
- Allows for missing data at occasion level.
- Allow time varying and time invariant covariates.

Centering Time

- Fixed Occasions: all persons provide observations at the same set of occasions.
  - Example: Wave 1, Wave 2, ... Wave t
- Varying Occasions: Observations taken at different time points for different persons.
  - Example: Age at t months
    - t=0 may not have interpretable meaning.
    - set intercept at an interpretable time point (e.g., set t=0 as age at 24 months).

Centering (Illustration)



Piecewise Linear Growth Model

Time:	0	1	2	3	4	5	6	.
Period 1:	0	1	2	2	2	2	2	
Period 2:	0	0	0	1	2	2	2	
Period 3:	0	0	0	0	0	1	2	

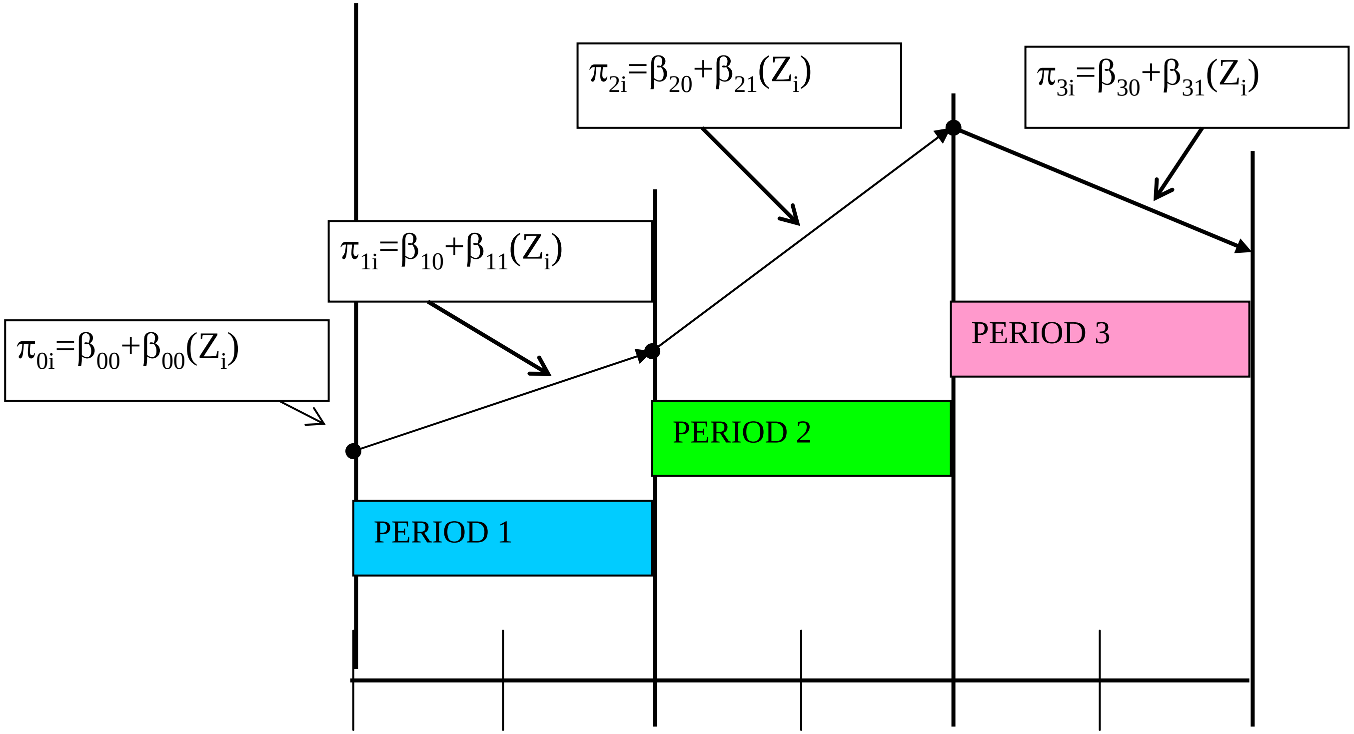
Level 1 (occasion level):

$$Y_{ti}=\pi_{0i}+\pi_{1i}\,(Period\ 1_{ti})+\pi_{2i}\,(Period\ 2_{ti})+\pi 2i\,(Period\ 3_{ti})+e_{ti}$$

Level 2 (person level):

$$\pi_{0i}=\beta_{00}+\beta_{01}\,(Z_i)+\mu_{0i}$$
$$\pi_{1i}=\beta_{10}+\beta_{11}\,(Z_i)+\mu_{1i}$$
$$\pi_{2i}=\beta_{20}+\beta_{21}\,(Zi)+\mu_{2i}$$
$$\pi_{3i}=\beta_{30}+\beta_{31}\,(Zi)+\mu_{3i}$$

Piecewise Linear Growth Model



Time	0	1	2	3	4	5	6
Period 1	0	1	2	2	2	2	2
Period 2	0	0	0	1	2	2	2
Period 3	0	0	0	0	0	1	2

Example

157 Schizophrenia patients

Repeated Measures (3 waves): Time (0, 1, 2)

Variables:

- DV: Social Non-Conformity
- Group = Control (0); Treatment (1)
- Impulsivity
- Perceptual Delusion
- Magic Ideation

Model Specification (Unconditional Model)

Level 1 (occasion level):

$$Non\text{-}Conformity_{ti} = \pi_{0i} + \pi_{1i}\,(Time_{ti}) + e_{ti}$$

$$e_{ti} \sim N(0, \sigma^2)$$

Level 2 (person level):

$$\pi_{0i} = \beta_{00} + \mu_{0i} \qquad \mu_{0i} \sim N(0, \tau_{00})$$

$$\pi_{1i} = \beta_{10} + \mu_{1i} \qquad \mu_{1i} \sim N(0, \tau_{11})$$

Model Specification (Unconditional Model)

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard Error	T-ratio	Approx. d.f.	P-value
For INTRCPT1, P0					
INTRCPT2, B00	15.261704	0.618308	24.683	156	0.000
For TIME slope, P1					
INTRCPT2, B10	-1.030751	0.284837	-3.619	156	0.001

Final estimation of variance components:

Random Effect		Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1,	U0	6.96844	48.55919	137	774.09715	0.000
TIME slope,	U1	1.87016	3.49749	137	194.42744	0.001
level-1,	R	3.31212	10.97015			

Model Specification (Conditional Model)

Level 1 (occasion level):

$$Non\text{-}Conformity_{ti} = \pi_{0i} + \pi_{1i}\,(Time_{ti}) + e_{ti}$$

$$e_{ti} \sim N(0, \sigma^2)$$

Level 2 (person level):

$$\pi_{0i} = \beta_{00} + \beta_{01}\,(Percept_i) + \beta 02\,(Magic_i) + \mu_{0i} \qquad \mu_{0i} \sim N(0, \tau_{00})$$

$$\pi_{1i} = \beta_{10} + \beta_{11}\,(Group_i) + \mu_{1i} \qquad \mu_{1i} \sim N(0, \tau_{11})$$

Model Specification (Conditional Model)

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard Error	T-ratio	Approx. d.f.	P-value
For INTRCPT1, P0					
INTRCPT2, B00	9.752993	0.740021	13.179	154	0.000
PERCEP, B01	0.334108	0.135700	2.462	154	0.014
MAGIC, B02	0.623003	0.127009	4.905	154	0.000
For TIME slope, P1					
INTRCPT2, B10	0.328386	0.353415	0.929	155	0.353
GROUP, B11	-2.362407	0.464579	-5.085	155	0.000

Final estimation of variance components:

Random Effect		Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1,	U0	5.19163	26.95297	135	489.92811	0.000
TIME slope,	U1	1.26078	1.58956	136	161.69813	0.066
level-1,	R	3.34419	11.18358			

Variance Explained

$$\frac{\hat{\tau}_{qq(unconditional)} - \hat{\tau}_{qq(conditional)}}{\hat{\tau}_{qq(unconditional)}}$$

Proportion of variance explained in the initial status by the model:

(48.56 – 26.95)/48.56 = 44.5%

Proportion of variance explained in the slope by the model:

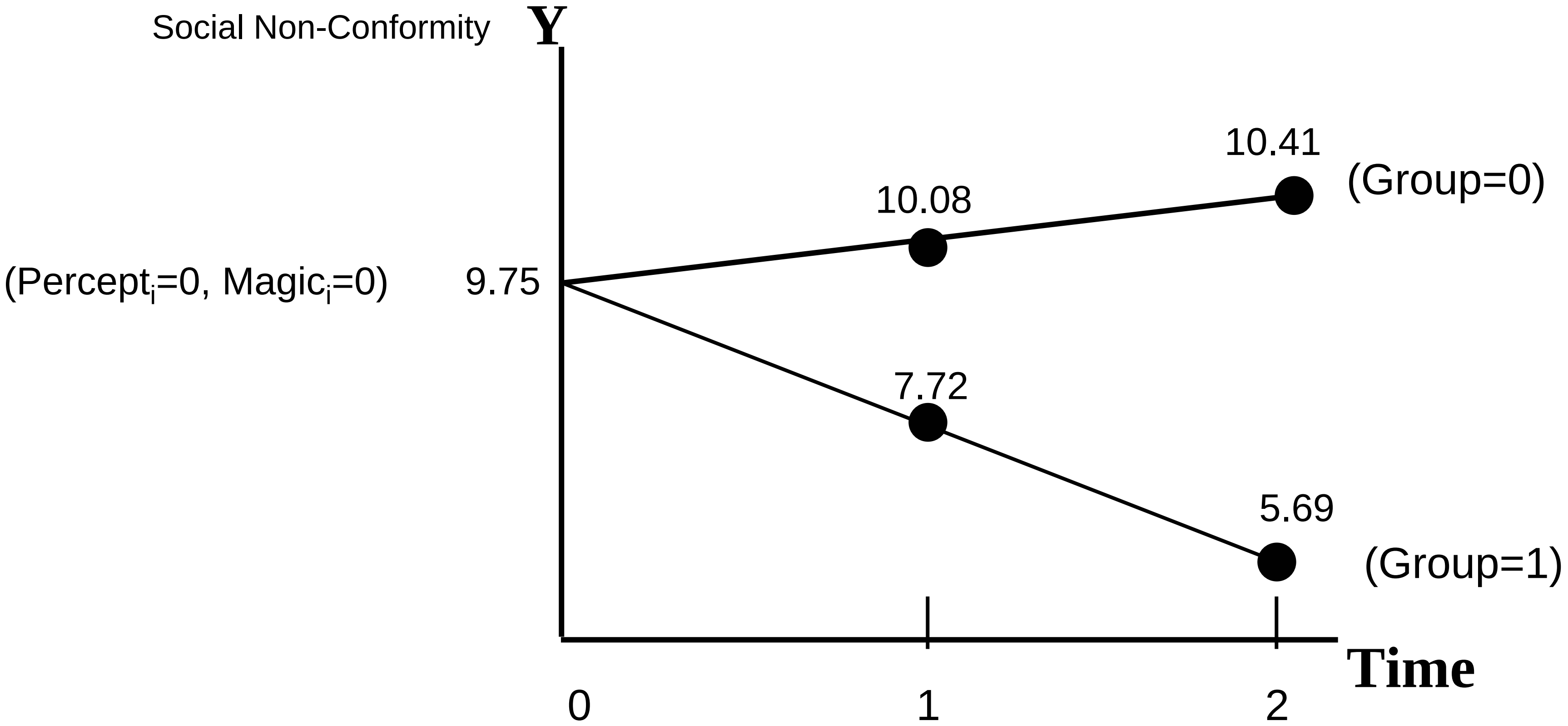
(3.50-1.59)/3.50 = 54.6%

Illustration

Solution:

$Y_{ti} = \pi_{0i} + \pi_{1i}(Time_{ti}) = \beta_{00}+\beta_{01}(Percept_i)+\beta_{02}(Magic_i) + ((\beta_{10}+ \beta_{11}(Group_i))(Time_{ti}))$

$Y_{ti} = (9.75+.33\ (Percept_i) +.62\ (Magic_i\ )) + (.33-2.36\ (Group_i)\ (Time))$



Assumptions & Limitations

Assumptions:

- Key variables in data must contain sufficient amounts of variances at both levels.
- The variance to be explained must have some theoretical meaning.
- Predictors and outcome must be normally distributed.
- Residuals are assumed to be normally distributed and independent of variables in equation.
- Linear relationship is assumed (unless otherwise specified).

Limitations:

- Cannot have missing level-2 data.
- Precision of the estimates is dependent on the within group sample size.
- The accuracy of the level-2 coefficient estimates depends on level-2 sample size.

Other Issues

- Non-Linear Models
  - Quadratic growth models
  - Logistic models

- 3 or more levels

- Sample Size & Power
  - Rules of thumb
  - A priori vs. post hoc power calculation
  - Costs

End of Presentation