Module I: Statistical Background on Multi-level Models

Francesca Dominici

Scott L. Zeger

Michael Griswold

The Johns Hopkins University

Bloomberg School of Public Health

Statistical Background on Multi-level Models

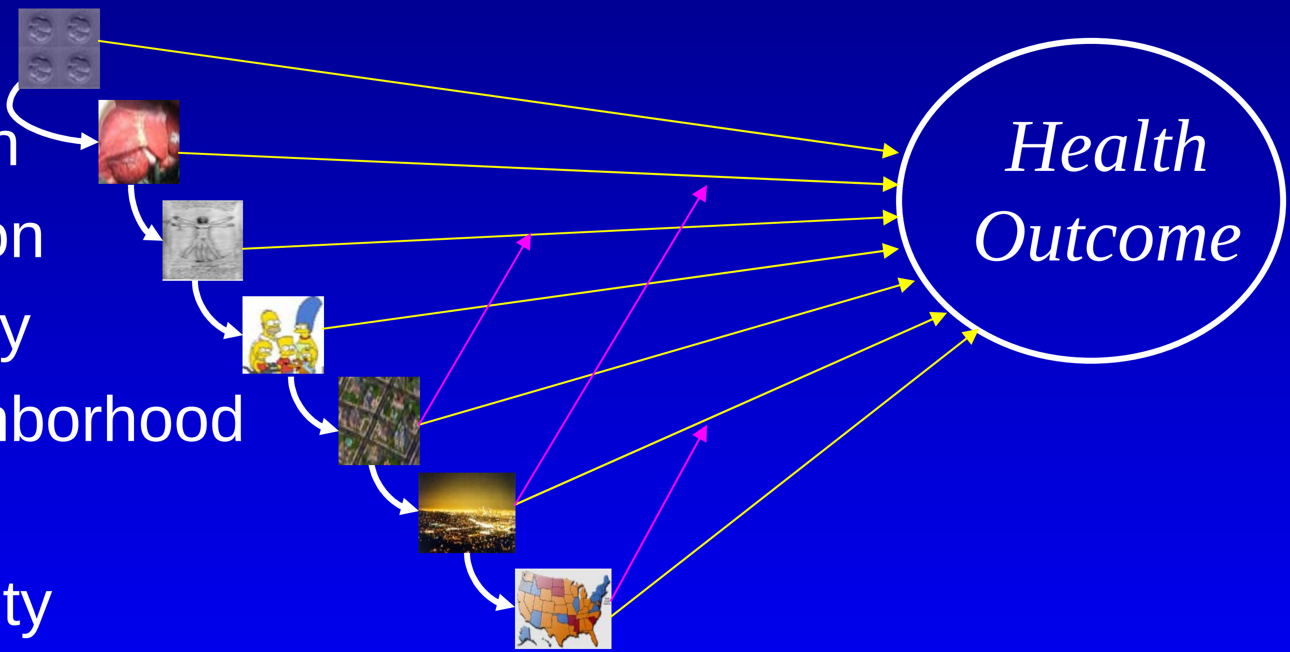
- Multi-level models
  - Main ideas
  - Conditional
  - Marginal
  - Contrasting Examples

A Rose is a Rose is a…

- Multi-level model
- Random effects model
- Mixed model
- Random coefficient model
- Hierarchical model

Multi-level Models – Main Idea

- Biological, psychological and social processes that influence health occur at many **levels**:
  - Cell
  - Organ
  - Person
  - Family
  - Neighborhood
  - City
  - Society



- An analysis of risk factors should consider:
  - Each of these levels
  - Their interactions

Example: Alcohol Abuse

Level:

1.	Cell:	Neurochemistry
2.	Organ:	Ability to metabolize ethanol
3.	Person:	Genetic susceptibility to addiction
4.	Family:	Alcohol abuse in the home
5.	Neighborhood:	Availability of bars
6.	Society:	Regulations; organizations; social norms

Example: Alcohol Abuse; Interactions among Levels

Level:

Level	
5	Availability of bars *and*
6	State laws about drunk driving
4	Alcohol abuse in the family *and*
2	Person's ability to metabolize ethanol
3	Genetic predisposition to addiction *and*
4	Household environment
6	State regulations about intoxication *and*
3	Job requirements

Notation:

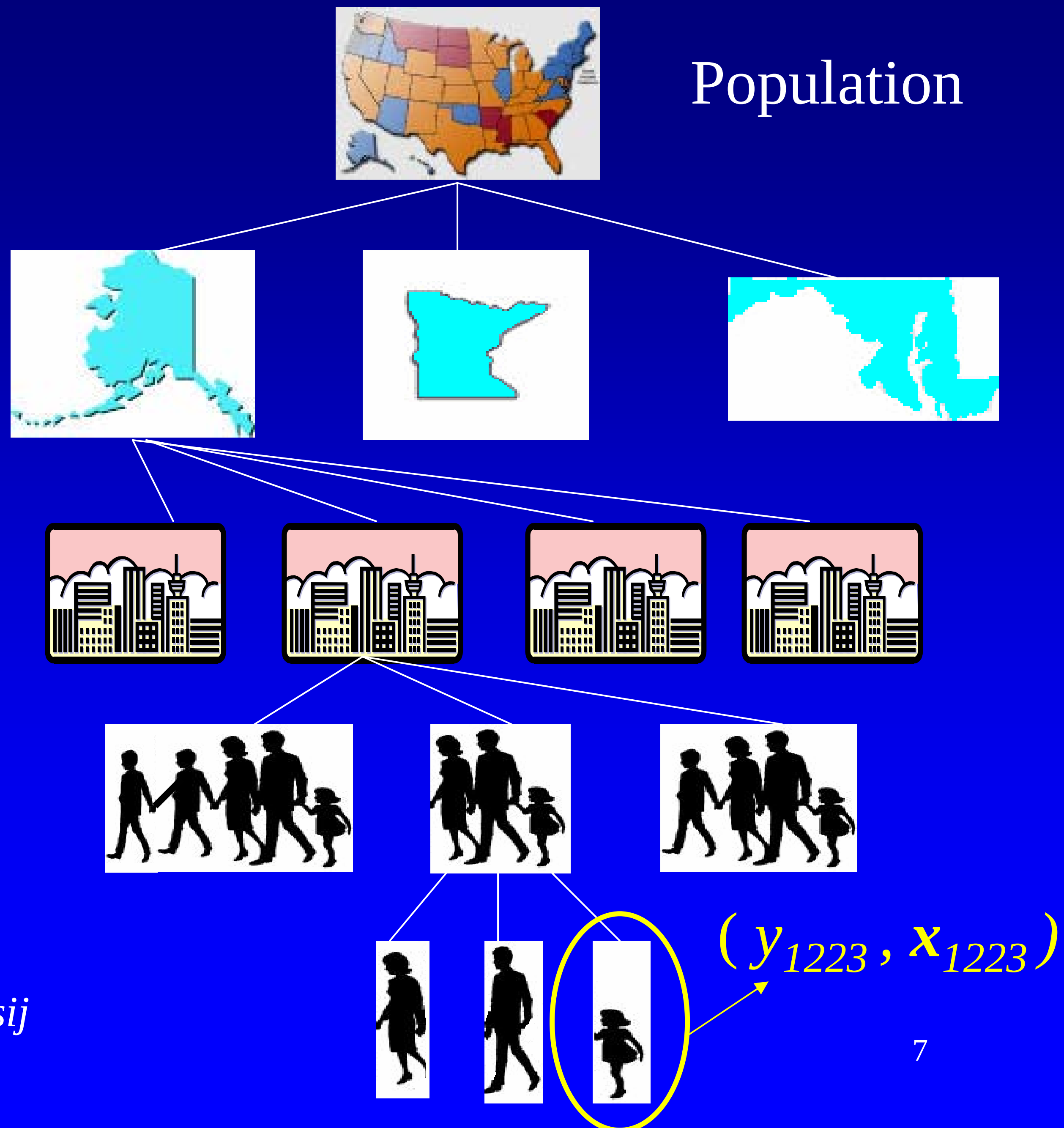
Person: $sijk$
Outcome: Y_{sijk}
Predictors: $\boldsymbol{X}_{sijk}$

State: $s=1,\ldots,S$

Neighborhood: $i=1,\ldots,I_s$

Family: $j=1,\ldots,J_{si}$

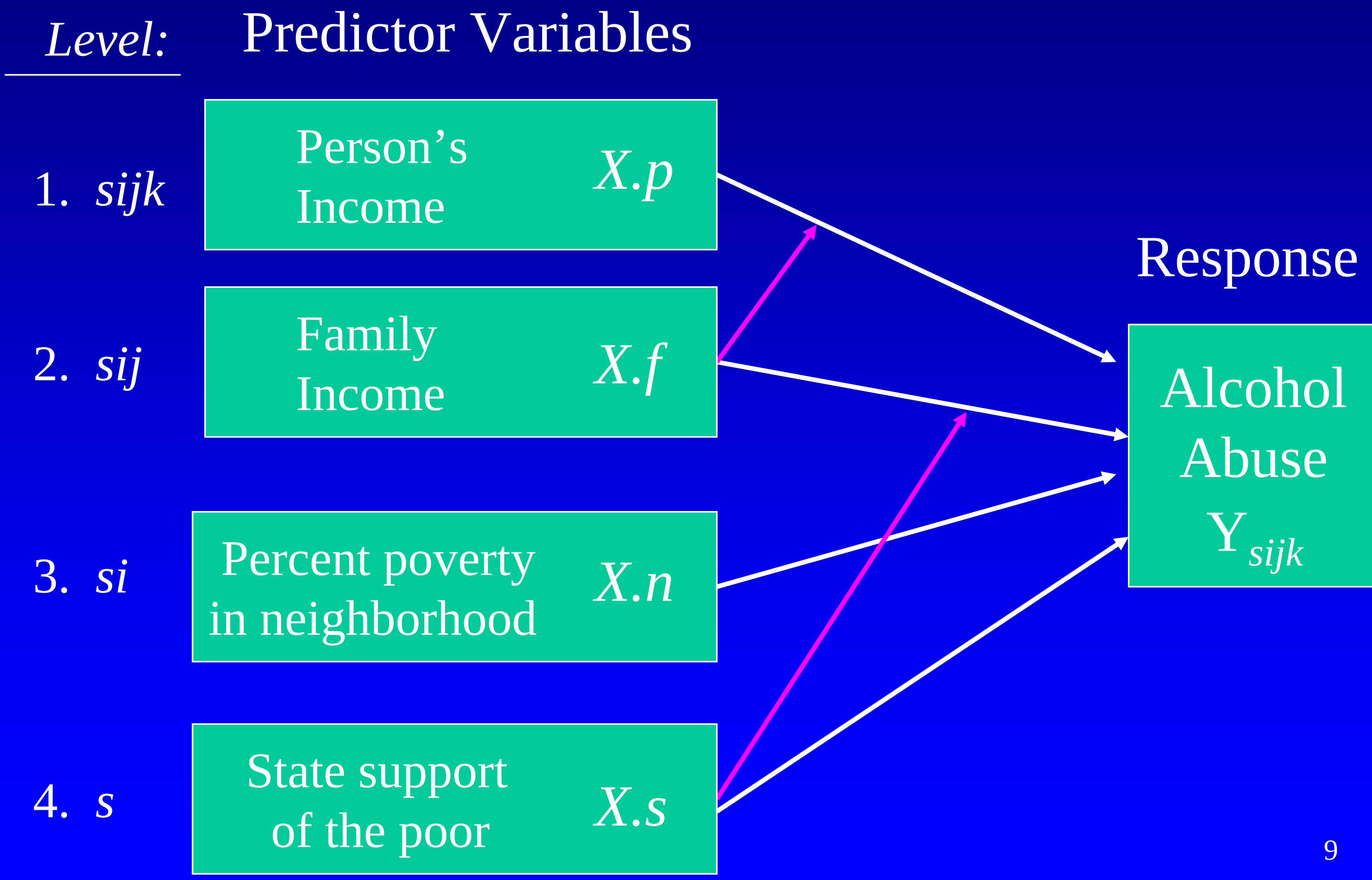
Person: $k=1,\ldots,K_{sij}$



Notation (cont.)

- (y_{sijk}, x_{sijk}) are (response, predictors) for
  - person $k = 1, ..., K_{sij}$ in
  - family $j = 1, ..., J_{si}$ in
  - neighborhood $i = 1, ..., I_s$ in
  - state $s = 1, ..., S$
- $\mu_{sijk} = \mathsf{E}(y_{sijk} | x_{sijk})$

Multi-level Models: Idea



Digression on Statistical Models

- A statistical model is an approximation to reality
- There is not a “correct” model;
  - (forget the holy grail)
- A model is a tool for asking a scientific question;
  - (screw-driver vs. sludge-hammer)
- A useful model combines the data with prior information to address the question of interest.
- Many models are better than one.

Generalized Linear Models (GLMs)

$$g(\mu) = \beta_0 + \beta_1 {*} X_1 + \ldots + \beta_p {*} X_p$$

where: $\mu = E(Y|X) = \text{mean}$

Model	Response	$g(\mu)$	Distribution	Coef Interp
Linear	Continuous (ounces)	μ	Gaussian	Change in avg(Y) per unit change in X
Logistic	Binary (disease)	$\log\left(\frac{\mu}{(1-\mu)}\right)$	Binomial	Log Odds Ratio
Log-linear	Count/Times to events	$\log(\mu)$	Poisson	Log Relative Risk

Generalized Linear Models (GLMs)

$$g(\mu) = \beta_0 + \beta_1 * X_1 + \ldots + \beta_p * X_p$$

Example: Age & Gender

Gaussian – Linear: $E(y) = \beta_0 + \beta_1 \text{Age} + \beta_2 \text{Gender}$

β_1 = Change in Average Response per 1 unit increase in Age, Comparing people of the SAME GENDER.

WHY?

Since: $E(y|\text{Age}+1, \text{Gender}) = \beta_0 + \beta_1(\text{Age}+1) + \beta_2 \text{Gender}$

And: $E(y|\text{Age}, \text{Gender}) = \beta_0 + \beta_1 \text{Age} + \beta_2 \text{Gender}$

$\Delta E(y) = \beta_1$

Generalized Linear Models (GLMs)

$$g(\mu) = \beta_0 + \beta_1{*}X_1 + \ldots + \beta_p{*}X_p$$

Example: Age & Gender

Binary – Logistic: $\log\{\text{odds}(Y)\} = \beta_0 + \beta_1\text{Age} + \beta_2\text{Gender}$

β_1 = log-OR of "+ Response" for a 1 unit increase in Age, Comparing people of the SAME GENDER.

WHY?

Since: $\log\{\text{odds}(y|\text{Age}+1,\text{Gender})\} = \beta_0 + \beta_1(\text{Age}+1) + \beta_2\text{Gender}$

And: $\log\{\text{odds}(y|\text{Age},\text{Gender})\} = \beta_0 + \beta_1\text{Age} + \beta_2\text{Gender}$

Δ log-Odds = β_1

⇒ log-OR = β_1

Generalized Linear Models (GLMs)

$$g(\mu) = \beta_0 + \beta_1 {*} X_1 + \ldots + \beta_p {*} X_p$$

Example: Age & Gender

Counts – Log-linear: $\log\{E(Y)\} = \beta_0 + \beta_1 \text{Age} + \beta_2 \text{Gender}$

β_1 = log-RR for a 1 unit increase in Age,
Comparing people of the SAME GENDER.

WHY?

Verify for Yourself Tonight

Most Important Assumptions of Regression Analysis?

A. Data follow normal distribution

B. All the key covariates are included in the model

C. Xs are fixed and known

D. Responses are independent

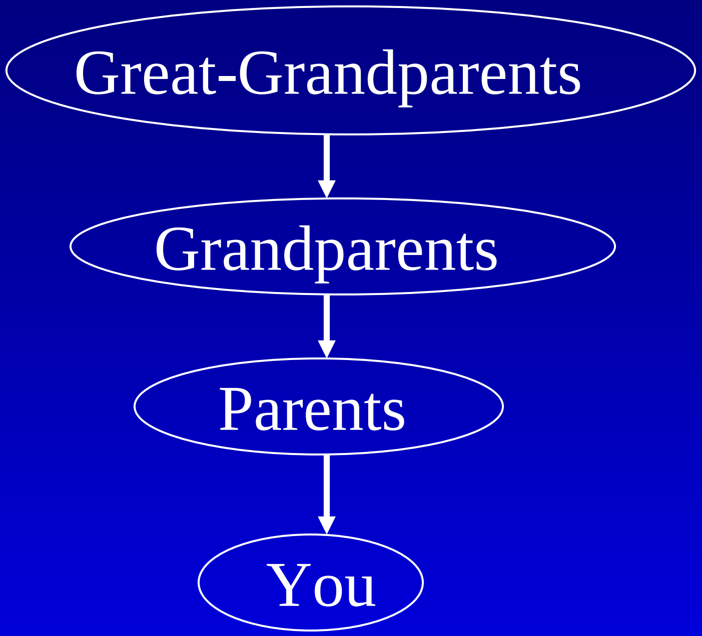
Within-Cluster Correlation

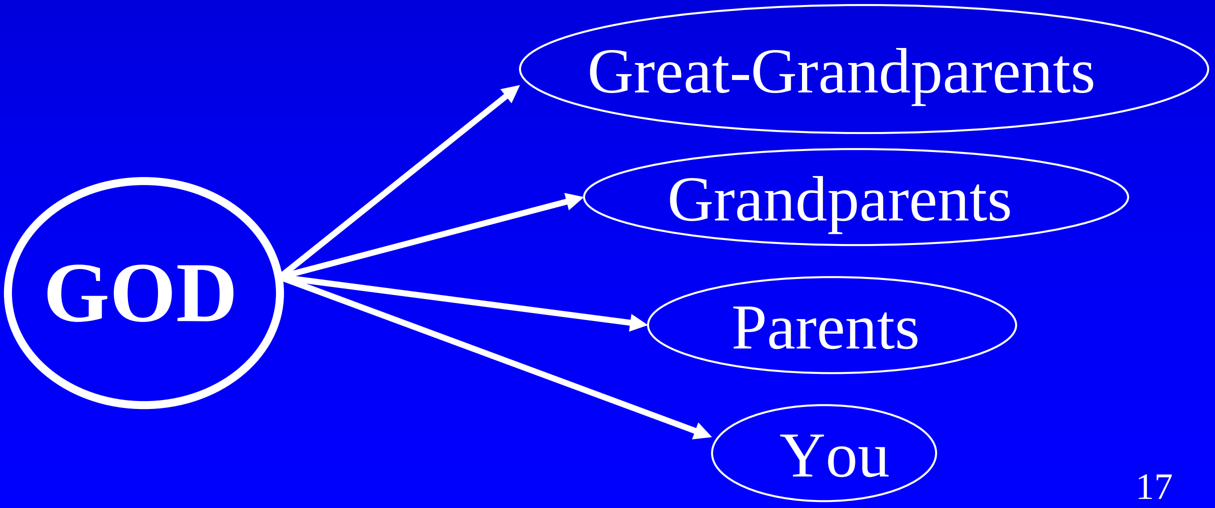
- Fact: two responses from the same family tend to be more like one another than two observations from different families

- Fact: two observations from the same neighborhood tend to be more like one another than two observations from different neighborhoods

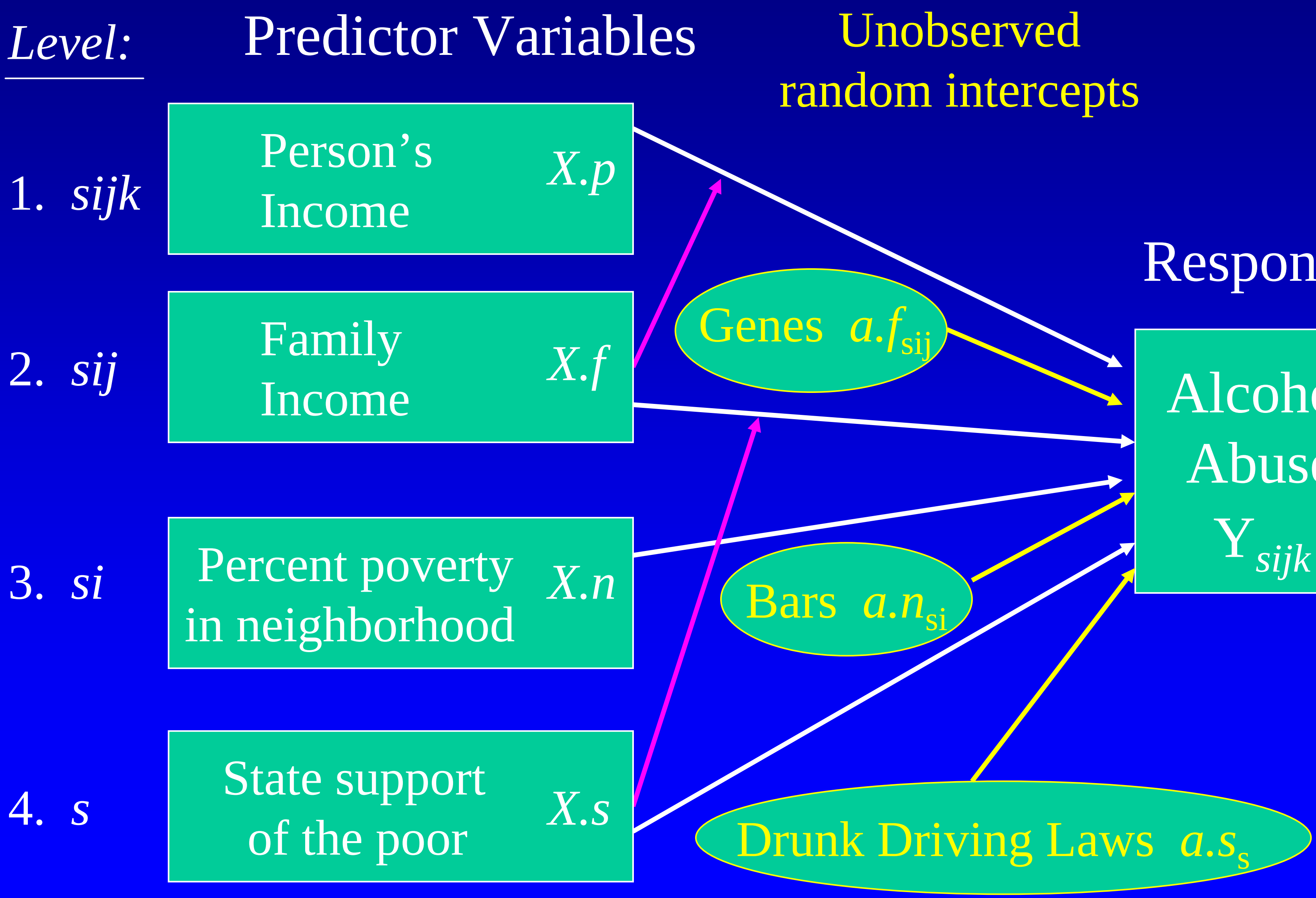
- Why?

Why? (Family Wealth Example)





Multi-level Models: Idea



Key Components of Multi-level Model

- Specification of predictor variables from multiple levels
  - Variables to include
  - Key interactions

- Specification of correlation among responses from same clusters

- Choices must be driven by the scientific question

Multi-level Shmulti-level

- Multi-level analysis of social/behavioral phenomena: an important idea
- Multi-level models involve predictors from multi-levels and their interactions
- They must account for correlation among observations within clusters (levels) to make efficient and valid inferences.

Key Idea for Regression with Correlated Data

Must take account of correlation to:

- Obtain valid inferences
  - standard errors
  - confidence intervals
  - posteriors
- Make efficient inferences

Logistic Regression Example: Cross-over trial

Ordinary logistic regression:

- Response: 1-normal; 0- alcohol dependence
- Predictors: period (x_1); treatment group (x_2)
- Two observations per person
- Parameter of interest: log odds ratio of dependence: treatment vs placebo

Mean Model: $\log\{\text{odds(AD)}\} = \beta_0 + \beta_1 \text{Period} + \beta_2 \text{Trt}$

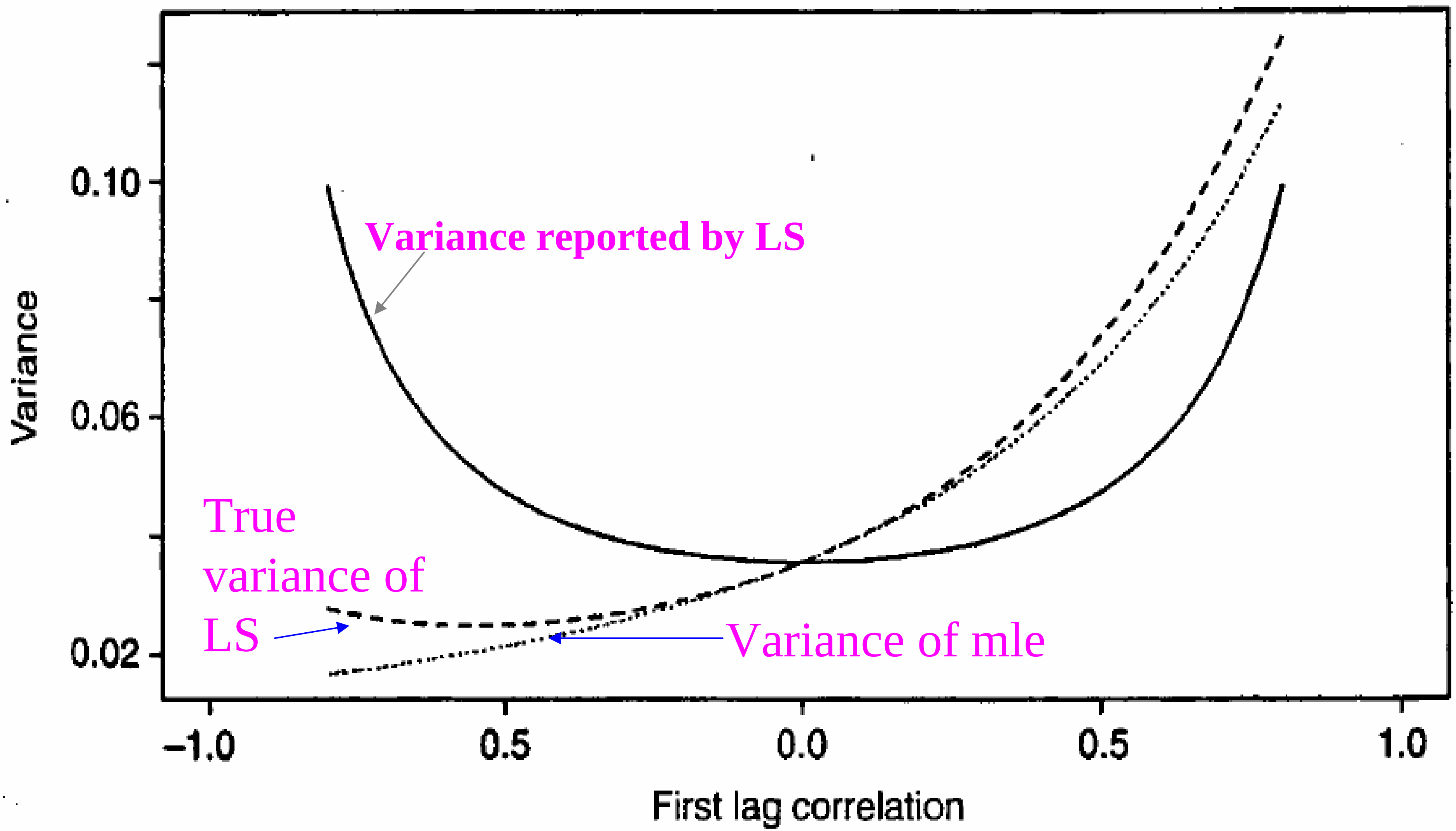
Results:
estimate, (standard error)

	Model	
Variable	Ordinary Logistic Regression	Account for correlation
Intercept (β_0)	0.66 (0.32)	0.67 (0.29)
Period (β_1)	-0.27 (0.38)	-0.30 (0.23)
Treatment (β_2)	0.56 (0.38)	0.57 (0.23)

Similar estimates,

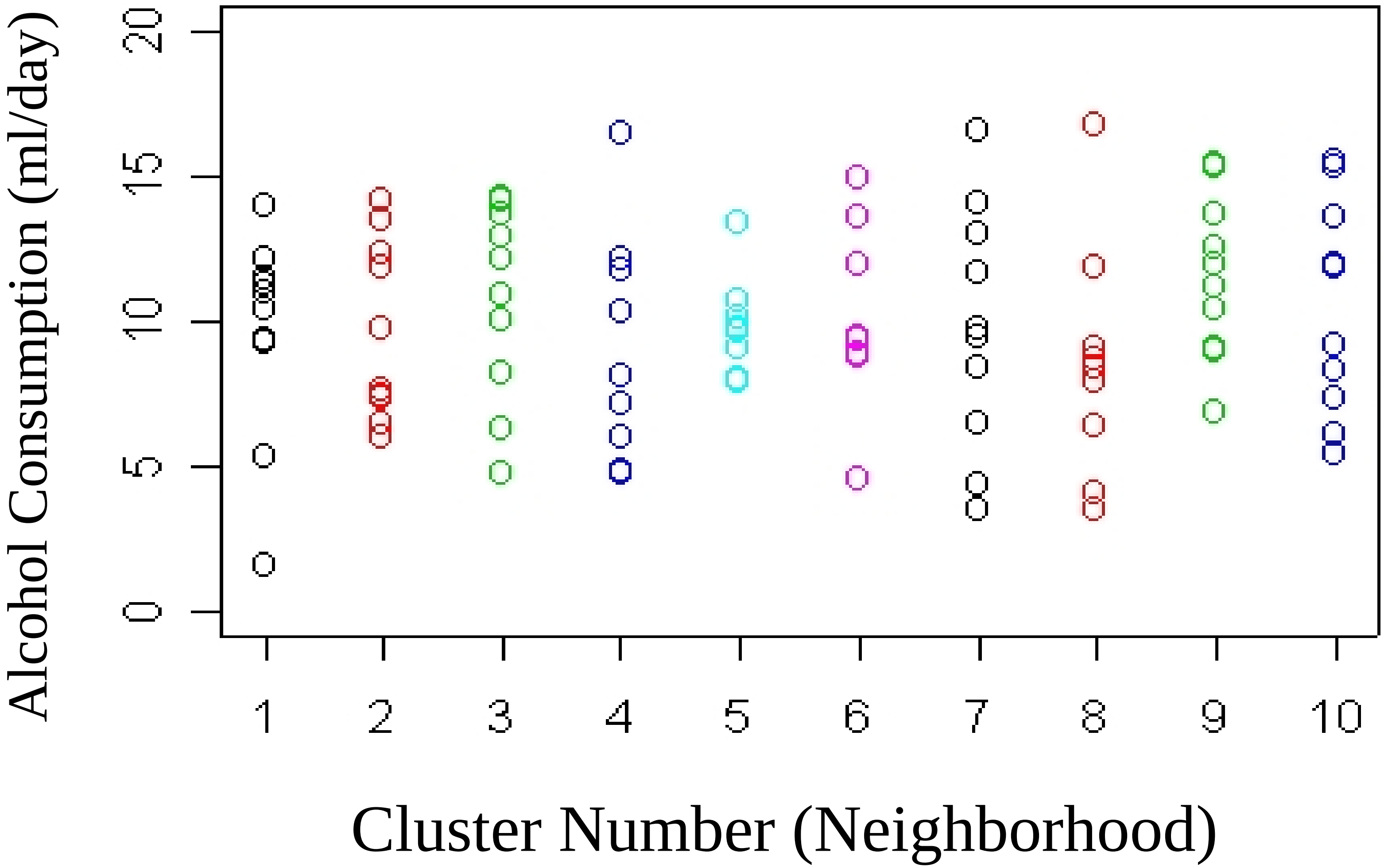
WRONG Standard Errors (& Inferences) for OLR

Variance of Least Squares and ML Estimators of Slope –vs- First Lag Correlation

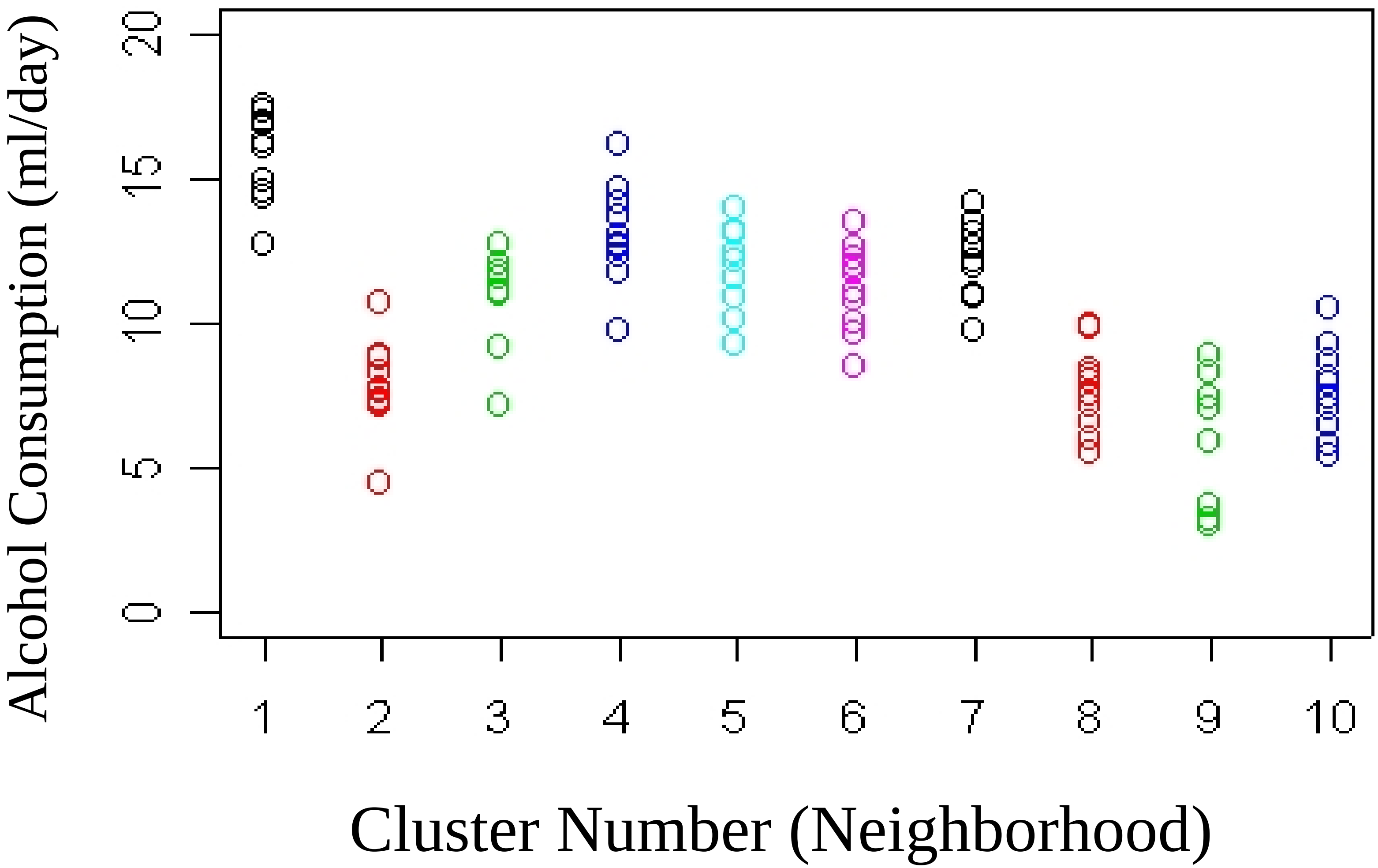


Source: DHLZ 2002 (pg 19)

Simulated Data: Non-Clustered



Simulated Data: Clustered



Within-Cluster Correlation

- Correlation of two observations from same cluster =

$$\boxed{\frac{\text{Total Var} - \text{Within Var}}{\text{Total Var}}}$$

- Non-Clustered = (9.8-9.8) / 9.8 = 0
- Clustered = (9.8-3.2) / 9.8 = 0.67

Models for Clustered Data

- Models are tools for inference
- Choice of model determined by scientific question
- Scientific Target for inference?
  - *Marginal* mean:
    - Average response across the population
  - *Conditional* mean:
    - Given other responses in the cluster(s)
    - Given unobserved random effects

Marginal Models

- Target – marginal mean or population-average response for different values of predictor variables
- Compare Groups
- Examples:
  - Mean alcohol consumption for Males vs Females
  - Rates of alcohol abuse for states with active addiction treatment programs vs inactive states

ex. mean model: $E(\text{AlcDep}) = \beta_0 + \beta_1 \text{Gender}$

- Public health (a.k.a. population) questions

Marginal GLMS for Multi-level Data: Generalized Estimating Equations (GEE)

- Mean Model: (Ordinary GLM - linear, logistic,..)
  - Population-average parameters
  - e.g. log{ odds(AlcDep$_{ij}$) } = $\beta_0 + \beta_1$Gender$_{ij}$

subject ***i*** in cluster ***j***

- Association Model: (for observations in clusters)
  - e.g. log{ Odds Ratio(Y_{ij},Y_{kj}) } = α_0

two different subjects (***i*** & ***k***) in cluster ***j***

- Solving GEE (DHLZ, 2002) gives nearly efficient and valid inferences about population-average parameters

OLR vs GEE
Cross-over Example

	Model	
Variable	Ordinary Logistic Regression	GEE Logistic Regression
Intercept	0.66 (0.32)	0.67 (0.29)
Period	-0.27 (0.38)	-0.30 (0.23)
Treatment	0.56 (0.38)	0.57 (0.23)
log(OR) (association)	0.0	3.56 (0.81)

Marginal Model Interpretations

- $\log\{ \text{odds(AlcDep)} \} = \beta_0 + \beta_1\text{Period} + \beta_2\text{trt}$

$$= 0.67 + (-0.30)\text{Period} + (0.57)\text{trt}$$

TRT Effect: (placebo vs. trt)

OR = exp(0.57) = 1.77,   95% CI (1.12, 2.80)

→ *Risk of Alcohol Dependence is almost twice as high on placebo, regardless of, (adjusting for), time period*

WHY?

Since: $\log\{\text{odds(AlcDep|Period, pl)}\} = \beta_0 + \beta_1\text{Period} + \beta_2$

And: $\log\{\text{odds(AlcDep|Period, trt)}\} = \beta_0 + \beta_1\text{Period}$

$$\Delta \text{ log-Odds} = \beta_2$$

$$\Rightarrow \text{OR} = \exp(\beta_2)$$

Conditional Models

- Conditional on other observations in cluster
  - Probability a person abuses alcohol as a function of the number of family members that do
  - A person's average alcohol consumption as a function on the average in the neighborhood

- Use other responses from the cluster as predictors in regressions like additional covariates

ex: $E(\text{AlcDep}_{ij}) = \beta_0 + \beta_1 \text{Gender}_{ij} + \beta_2 \overline{\text{AlcDep}_j}$

Conditional on Other Responses:
- Usually a Bad Idea -

- Definition of "other responses in cluster" depends on size/nature of cluster
  - e.g. "number of other family members who do"
    - 0 for a single person means something different that 0 in a family with 10 others

- The "risk factors" may affect the entire cluster; conditioning on the responses for the others will dilute the risk factor effect
  - Two eyes example

ex: $\log\{\text{odds}(\text{Blind}_{i,\text{Left}})\} = \beta_0 + \beta_1 \text{Sun} + \beta_2 \text{Blind}_{i,\text{Right}}$

≈ 0

Conditional Models

- Conditional on unobserved latent variables or “random effects”
  - Alcohol use within a family is related because family members share an unobserved “family effect”: common genes, diets, family culture and other unmeasured factors

  - Repeated observations within a neighborhood are correlated because neighbors share: common traditions, access to services, stress levels,…

Random Effects Models

- Latent (random) effects are unobserved
  - inferred from the correlation among residuals
- Random effects models describe the marginal mean and the source of correlation in one equation
- Assumptions about the latent variables determine the nature of the associations
  - ex: Random Intercept = Uniform Correlation

ex: $E(\text{AlcDep}_{ij} \mid b_j) = \beta_0 + \beta_1 \text{Gender}_{ij} + b_j$

where: $b_j \sim N(0,\sigma^2)$

Cluster specific random effect

OLR vs R.E.
Cross-over Example

	Model	
Variable	Ordinary Logistic Regression	Random Int. Logistic Regression
Intercept	0.66 (0.32)	2.2 (1.0)
Period	-0.27 (0.38)	-1.0 (0.84)
Treatment	0.56 (0.38)	1.8 (0.93)
log(σ) (association)	0.0	5.0 (2.3)

Conditional Model Interpretations

i^{th} subject's latent propensity for Alcohol Dependence

- $\log\{\, \text{odds}(\text{AlcDep}_i \mid b_i) \,\}$

$$= \beta_0 + \beta_1 \text{Period} + \beta_2 \text{trt} + b_i$$

$$= 2.2 + (-1.0)\text{Period} + (1.8)\text{trt} + b_i$$

where: $b_i \sim N(0,5^2)$

TRT Effect: (placebo vs. trt)

OR = exp(1.8) = 6.05, 95% CI (0.94, 38.9)

➡ *A Specific Subject's Risk of Alcohol Dependence is 6 TIMES higher on placebo, regardless of, (adjusting for), time period*

Conditional Model Interpretations

WHY?

Since: $\log\{\text{odds}(\text{AlcDep}_i|\text{Period}, \text{pl}, b_i)\} = \beta_0 + \beta_1\text{Period} + \beta_2 + b_i$

And: $\log\{\text{odds}(\text{AlcDep}|\text{Period}, \text{trt}, b_i)\} = \beta_0 + \beta_1\text{Period} + b_i$

$$\Delta \text{ log-Odds} = \beta_2$$

$$\Rightarrow \quad \text{OR} = \exp(\beta_2)$$

- In order to make comparisons we must keep the subject-specific latent effect (b_i) the same.
- In a Cross-Over trial we have outcome data for each subject on both placebo & treatment
- What about in a usual clinical trial / cohort study?

Marginal vs. Random Effects Models

- ***For linear models, regression coefficients in random effects models and marginal models are identical***:

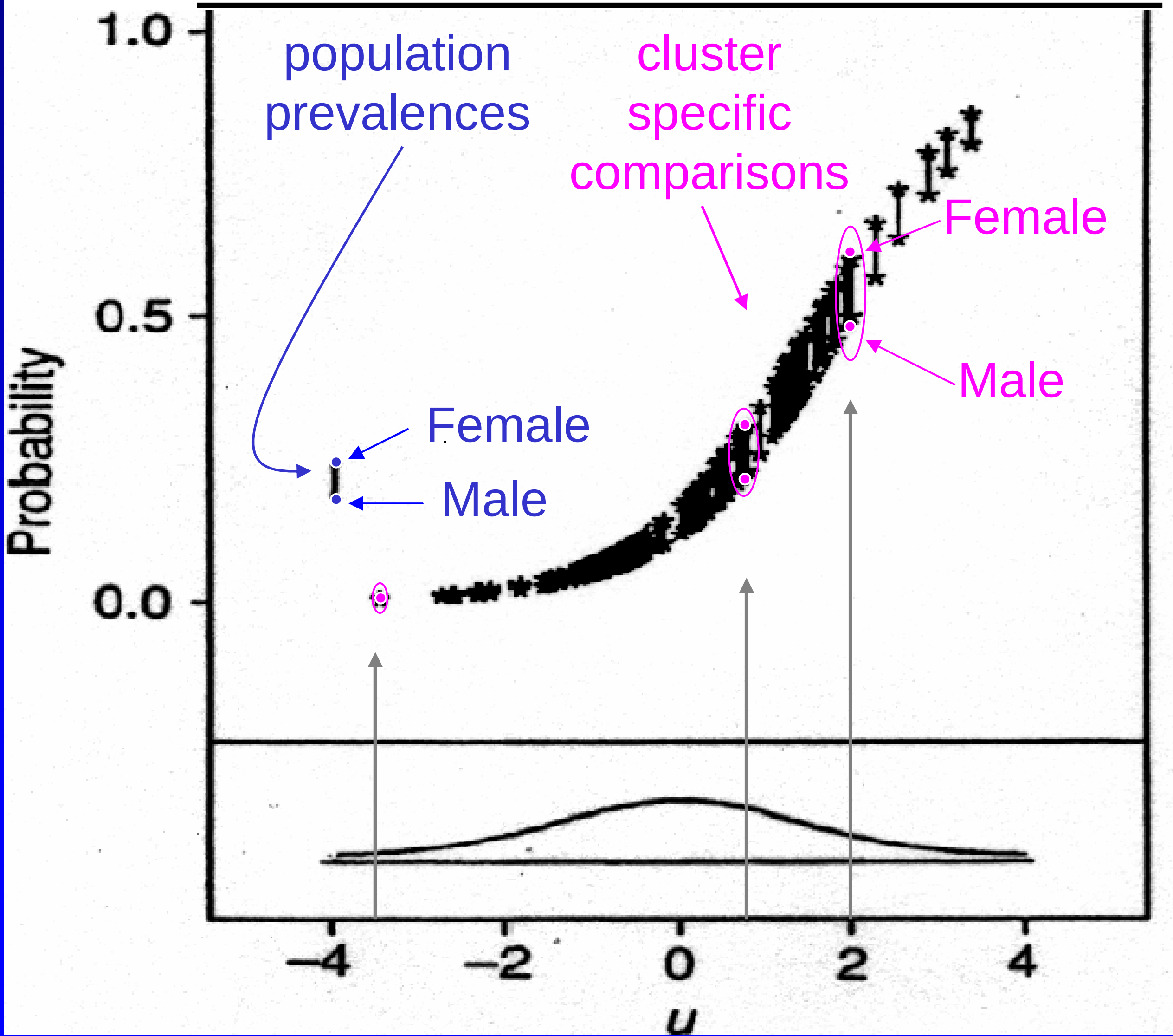
  average of linear function = linear function of average

- ***For non-linear models, (logistic, log-linear,…) coefficients have different meanings/values, and address different questions***
  - Marginal models -> *population-average* parameters
  - Random effects models -> cluster-*specific* parameters

Marginal –vs- Random Intercept Model

$\log\{\text{odds}(Y_i)\} = \beta_0 + \beta_1 * \text{Gender}$ VS.

$\log\{\text{odds}(Y_i \mid u_i)\} = \beta_0 + \beta_1 * \text{Gender} + u_i$



Source: DHLZ 2002 (pg 135)

Marginal -vs- Random Intercept Models; Cross-over Example

	Model		
Variable	Ordinary Logistic Regression	Marginal (GEE) Logistic Regression	Random-Effect Logistic Regression
Intercept	0.66 (0.32)	0.67 (0.29)	2.2 (1.0)
Period	-0.27 (0.38)	-0.30 (0.23)	-1.0 (0.84)
Treatment	0.56 (0.38)	0.57 (0.23)	1.8 (0.93)
Log OR (assoc.)	0.0	3.56 (0.81)	5.0 (2.3)

Comparison of Marginal and Random Effect Logistic Regressions

- Regression coefficients in the random effects model are roughly 3.3 times as large
  - Marginal: population odds (prevalence with/prevalence without) of AlcDep is exp(.57) = 1.8 greater for placebo than on active drug;

    population-average parameter
  - Random Effects: a person's odds of AlcDep is exp(1.8)= 6.0 times greater on placebo than on active drug;

    cluster-specific, here person-specific, parameter

Which model is better? They ask different questions.

Marginalized Multi-level Models

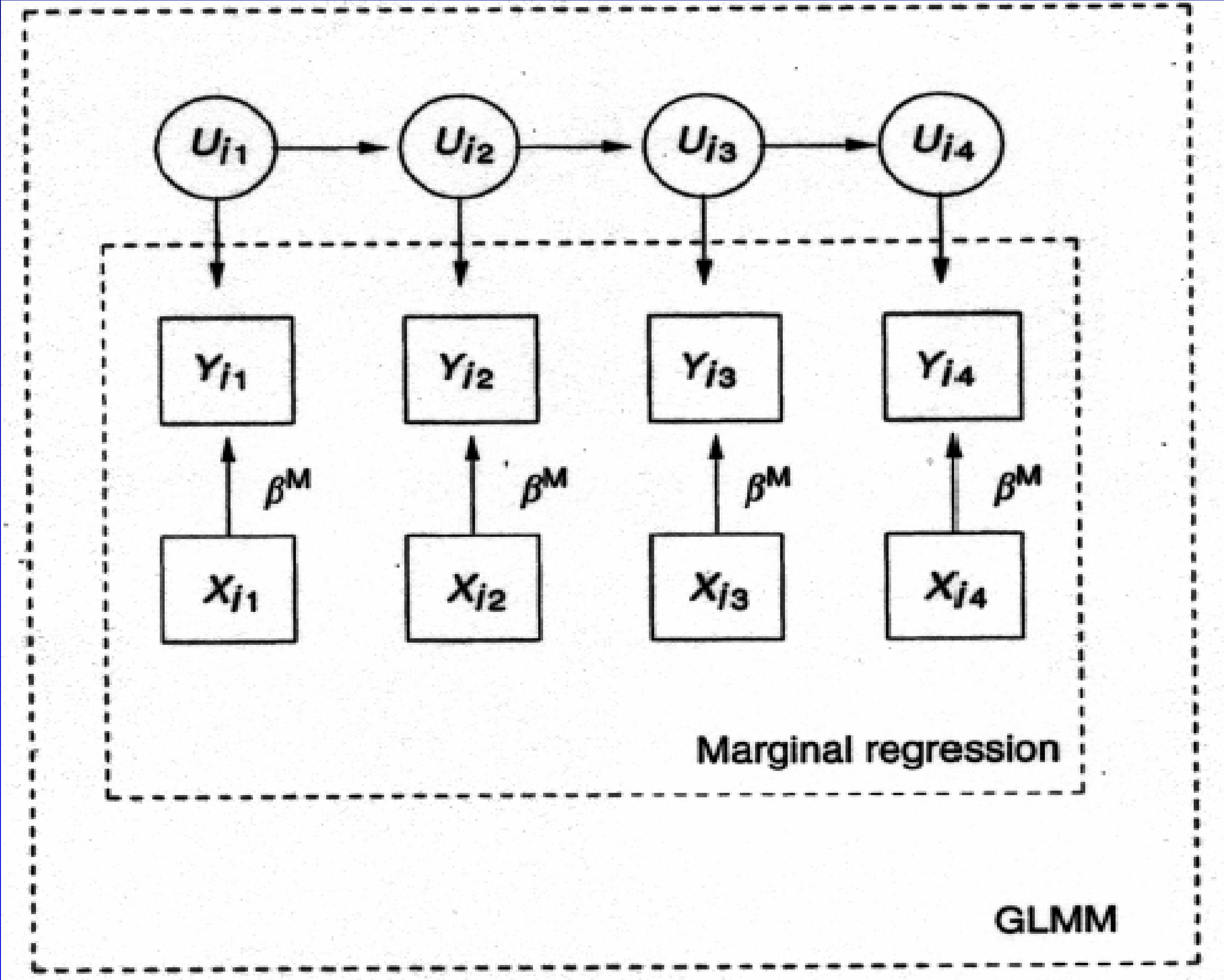
- Heagerty (1999, *Biometrics*); Heagerty and Zeger (2000, *Statistical Science*)

- Model:
  - marginal mean as a function of covariates
  - conditional mean given random effects as a function of marginal mean and cluster-specific random effects

- Random Effects allow flexible association models, but public health is usually concerned with population-averaged (marginal) questions.

$\Rightarrow$ MMM

Schematic of Marginal Random-effects Model



Marginal and Random Intercept Models
Cross-over Example

	Model			
Variable	Ordinary Logistic Regression	GEE Logistic Regression	MMM Logistic Regression	Random Int. Logistic Regression
Intercept	0.66 (0.32)	0.67 (0.29)	0.65 (0.28)	2.2 (1.0)
Period	-0.27 (0.38)	-0.30 (0.23)	-0.33 (0.22)	-1.0 (0.84)
Treatment	0.56 (0.38)	0.57 (0.23)	0.58 (0.23)	1.8 (0.93)
log(OR) (assoc.)	0.0	3.56 (0.81)	5.44 (3.72)	5.0 (2.3)

Refresher: Forests & Trees

Multi-Level Models:

- Explanatory variables from multiple levels
  - Family
  - Neighborhood
  - State
- Interactions

Must take account of correlation among responses from same clusters:

- Marginal: GEE, MMM
- Conditional: RE, GLMM

Illustration of Conditional Models and Marginal Multi-level Models; The British Social Attitudes Survey

- Binary Response: $Y_{ijk} = \begin{cases} 1 & \text{if favor abortion} \\ 0 & \text{if not} \end{cases}$
- Levels (notation)
  - Year: k=1,…,4 (1983-1986)
  - Subject: j=1,…,264
  - District: i=1,…54
  - Overall Sample: N = 1,056
- Levels (conception)
  - 1: time within person
  - 2: persons within districts
  - 3: districts

Covariates at Three Levels

- Level 1: time
  - Indicators of time

- Level 2: person
  - Class: upper working; lower working
  - Gender
  - Religion: protestant, catholic, other

- Level 3: district
  - Percentage protestant (derived)

Scientific Questions

- How does a person's religion influence her probability of favoring abortion?
- How does the predominant religion in a person's district influence her probability of favoring abortion?

Conditional model

- How does the rate of favoring abortion differ between protestants and otherwise similar catholics?
- How does the rate of favoring abortion differ between districts that are predominantly protestant versus other religions?

Marginal model

Conditional Multi-level Model

Levels:

1. Time: k
2. Person: j
3. District: i

$$\text{logit } E(Y_{ijk} \mid \mathbf{X}_{ijk}, b_{2,ij}, b_{3,i}) = \beta_0^C + \mathbf{X}_{1,ijk}\boldsymbol{\beta}_1^C + \mathbf{X}_{2,ij}\boldsymbol{\beta}_2^C + X_{3,i}\beta_3^C + b_{2,ij} + b_{3,i},$$

Person and district random effects

where we assume $b_{2,ij} = \sigma_2(\mathbf{X}_{2,ij})z_{2,ij}$, $z_{2,ij} \sim N(0,1)$, and $b_{3,i} = \sigma_3 z_{3,i}$, $z_{3,i} \sim N(0,1)$,

Conditional Multi-level Model Results

Coefficient		Model 1* Estimate	(s.e.)	Model 2* Estimate	(s.e.)	Model 3* Estimate	(s.e.)	Model 4* Estimate	(s.e.)
Conditional mean (β^C)									
Intercept		−1.388	(0.685)	−1.279	(0.693)	−1.001	(0.665)	−1.452	(0.679)
Year: 2	1	−0.761	(0.266)	−0.770	(0.267)	−0.758	(0.266)	−0.760	(0.267)
Year: 3		0.060	(0.252)	0.060	(0.252)	0.062	(0.252)	0.060	(0.252)
Year: 4		0.300	(0.251)	0.299	(0.250)	0.303	(0.250)	0.303	(0.250)
Class: upper working		−0.623	(0.378)	−0.587	(0.427)	−0.667	(0.374)	−0.708	(0.374)
Class: lower working		−0.499	(0.361)	−0.310	(0.405)	−0.513	(0.356)	−0.658	(0.371)
Gender	2	−0.600	(0.358)	−0.738	(0.354)	−0.876	(0.389)	−0.477	(0.356)
Religion: catholic		−0.609	(0.803)	−0.653	(0.730)	−0.725	(0.782)	−0.376	(0.950)
Religion: other		−1.049	(0.604)	−1348	(0.616)	−1.319	(0.615)	−0.487	(0.586)
Religion: none		1.263	(0.452)	0.861	(0.469)	1.019	(0.445)	1.384	(0.481)
% protestant	3	1.458	(0.837)	1.541	(0.778)	1.129	(0.809)	1.456	(0.821)
Level 2 heterogeneity (σ_2)									
Intercept		2.138	(0.236)	2.776	(0.592)	1.642	(0.294)	2.450	(0.408)
Class: upper working				−0.005	(0.769)				
Class: lower working				−1.135	(0.662)				
Gender						0.994	(0.494)		
Religion: catholic								−0.498	(1.055)
Religion: other								−1.222	(0.668)
Religion: none								−0.348	(0.608)
Level 3 heterogeneity (σ_3)									
Intercept		0.816	(0.295)	0.619	(0.320)	0.790	(0.285)	0.835	(0.282)
log L		−531.83		−529.42		−529.71		−530.35	

Conditional Scientific Answers

- How does a person's religion influence her probability of favoring abortion?

	Model 1*		Model 2*		Model 3*		Model 4*	
Coefficient	Estimate	(s.e.)	Estimate	(s.e.)	Estimate	(s.e.)	Estimate	(s.e.)
Religion: catholic	−0.609	(0.803)	−0.653	(0.730)	−0.725	(0.782)	−0.376	(0.950)
Religion: other	−1.049	(0.604)	−1348	(0.616)	−1.319	(0.615)	−0.487	(0.586)
Religion: none	1.263	(0.452)	0.861	(0.469)	1.019	(0.445)	1.384	(0.481)

- How does the predominant religion in a person's district influence her probability of favoring abortion?

	Model 1*		Model 2*		Model 3*		Model 4*	
Coefficient	Estimate	(s.e.)	Estimate	(s.e.)	Estimate	(s.e.)	Estimate	(s.e.)
% protestant	1.458	(0.837)	1.541	(0.778)	1.129	(0.809)	1.456	(0.821)

But Wait!…

Conditional Model Interpretations: Model 4

WHY?

$$\log\{\text{odds}(\text{Fav}|\text{Catholic}, X, b_{2,ij}, b_{3,ij})\}\} = \cancel{\beta_0} + \cancel{X\beta} + \beta_8 + \cancel{b_{2,0}} + b_C + \cancel{b_{3,0}}$$

$$\log\{\text{odds}(\text{Fav}|\text{Protestant}, X, b_{2,ij}, b_{3,ij})\}\} = \cancel{\beta_0} + \cancel{X\beta} + \cancel{b_{2,0}} + \cancel{b_{3,0}}$$

$$\Delta \text{ log-Odds} = \beta_8 + b_C$$

$$\Rightarrow \quad \text{OR} = \exp(\beta_8 + b_C)$$

$$\boxed{\Rightarrow \quad \text{OR} \neq \exp(\beta_8)}$$

Conditional Model Interpretations: Model 4

What happens if you simply report exp(β)??

$$\log\{odds(Fav|Catholic, X, b_{2,ij}, b_{3,ij}))\} = \beta_0 + X\beta + \beta_8 + b_{2,0} + b_C + b_{3,0}$$

$$\log\{odds(Fav|Prot/Cath, X, b_{2,ij}, b_{3,ij}))\} = \beta_0 + X\beta + b_{2,0} + b_C + b_{3,0}$$

$$\Delta \text{ log-Odds} = \beta_8$$

$$\Rightarrow \text{OR} = \exp(\beta_8)$$

But there were NO subjects in the study who were simultaneously BOTH Catholic AND Protestant

(Similar for % protestant!)

Marginal Multi-level Model

Levels:

1. Time: k
2. Person: j
3. District: i

Mean Model:

$$\text{logit}\, E(Y_{ijk}|\mathbf{X}_{ijk}) = \beta_0^M + \boldsymbol{\beta}_1^M \mathbf{X}_{1,ijk} + \boldsymbol{\beta}_2^M \mathbf{X}_{2,ij} + \beta_3^M X_{3,i}.$$

Association Model: (Separate)

$$\begin{aligned}\text{logit}\, & E(Y_{ijk}\,|\mathbf{X}_{ijk}, z_{2,ij}, z_{3,i}) \\ & = \Delta(\mathbf{X}_{ijk}) + \sigma_2(\mathbf{X}_{2,ij}) z_{2,ij} + \sigma_3 z_{3,i}\end{aligned}$$

Person and district random effects

Marginal Multi-level Model Results

Coefficient		Independence Estimate	Independence (s.e.)	Model 1 Estimate	Model 1 (s.e.)	Model 2 Estimate	Model 2 (s.e.)	Model 3 Estimate	Model 3 (s.e.)	Model 4 Estimate	Model 4 (s.e.)
Marginal mean (β^M)											
Intercept		−0.792	(0.287)	−0.763	(0.393)	−0.753	(0.393)	−0.751	(0.390)	−0.740	(0.395)
Year: 2	1	−0.433	(0.200)	−0.446	(0.153)	−0.417	(0.156)	−0.438	(0.154)	−0.453	(0.152)
Year: 3		0.038	(0.191)	0.025	(0.144)	0.031	(0.145)	0.040	(0.145)	0.014	(0.144)
Year: 4		0.181	(0.189)	0.165	(0.143)	0.172	(0.144)	0.147	(0.144)	0.155	(0.142)
Class: upper working		−0.328	(0.191)	−0.348	(0.216)	−0.335	(0.215)	−0.326	(0.216)	−0.370	(0.215)
Class: lower working		−0.431	(0.167)	−0.267	(0.206)	−0.269	(0.208)	−0.272	(0.206)	−0.300	(0.208)
Gender	2	−0.279	(0.140)	−0.349	(0.205)	−0.364	(0.206)	−0.315	(0.205)	−0.320	(0.205)
Religion: catholic		−0.421	(0.341)	−0.384	(0.480)	−0.416	(0.476)	−0.406	(0.477)	−0.389	(0.471)
Religion: other		−0.601	(0.250)	−0.634	(0.360)	−0.657	(0.366)	−0.700	(0.365)	−0.712	(0.343)
Religion: none		0.718	(0.179)	0.707	(0.256)	0.653	(0.260)	0.678	(0.253)	0.704	(0.258)
% protestant	3	0.858	(0.298)	0.799	(0.479)	0.806	(0.472)	0.768	(0.475)	0.796	(0.483)
Level 2 heterogeneity (σ_2)											
Intercept				2.140	(0.238)	2.274	(0.316)	1.689	(0.338)	2.433	(0.404)
Class: upper working						0.342	(0.513)				
Class: lower working						−0.460	(0.599)				
Gender								0.871	(0.464)		
Religion: catholic										−0.581	(0.946)
Religion: other										−1.143	(0.661)
Religion: none										−0.301	(0.589)
Level 3 heterogeneity (σ_3)											
Intercept				0.818	(0.295)	0.724	(0.308)	0.788	(0.287)	0.847	(0.281)
log L		−622.57		−531.92		−531.04		−530.36		−530.55	

Marginal Scientific Answers

- How does the rate of favoring abortion differ between protestants and otherwise similar catholics?

	Independence		Model 1		Model 2		Model 3		Model 4	
Coefficient	Estimate	(s.e.)	Estimate	(s.e.)	Estimate	(s.e.)	Estimate	(s.e.)	Estimate	(s.e.)
Religion: catholic	−0.421	(0.341)	−0.384	(0.480)	−0.416	(0.476)	−0.406	(0.477)	−0.389	(0.471)
Religion: other	−0.601	(0.250)	−0.634	(0.360)	−0.657	(0.366)	−0.700	(0.365)	−0.712	(0.343)
Religion: none	0.718	(0.179)	0.707	(0.256)	0.653	(0.260)	0.678	(0.253)	0.704	(0.258)

- How does the rate of favoring abortion differ between districts that are predominantly protestant versus other religions?

	Independence		Model 1		Model 2		Model 3		Model 4	
Coefficient	Estimate	(s.e.)	Estimate	(s.e.)	Estimate	(s.e.)	Estimate	(s.e.)	Estimate	(s.e.)
% protestant	0.858	(0.298)	0.799	(0.479)	0.806	(0.472)	0.768	(0.475)	0.796	(0.483)

Key Points

- "Multi-level" Models:
  - Have covariates from many levels and their interactions
  - Acknowledge correlation among observations from within a level (cluster)
- Conditional and Marginal Multi-level models have different targets; ask different questions
- When population-averaged parameters are the focus, use
  - GEE
  - Marginal Multi-level Models (Heagerty and Zeger, 2000)

Key Points (continued)

- When cluster-specific parameters are the focus, use random effects models that condition on unobserved latent variables that are assumed to be the source of correlation

- Warning: Model Carefully. Cluster-specific targets often involve extrapolations where there are no actual data for support
  - e.g. % protestant in neighborhood given a random neighborhood effect