$$Y_{ij} = \beta_0 + u_i + \beta_1 X_{ij} + \varepsilon_{ij}$$

u_i: variation between groups i

ε_{ij}: variation within groups

← assumes the effect of X is the same for every group

Mixed effects model w/ random slopes:

$$Y_{ij} = \beta_0 + u_i + \underbrace{(\beta_1 + v_i)}_{\text{effect of } X \text{ varies from group to group}} X_{ij} + \varepsilon_{ij}$$

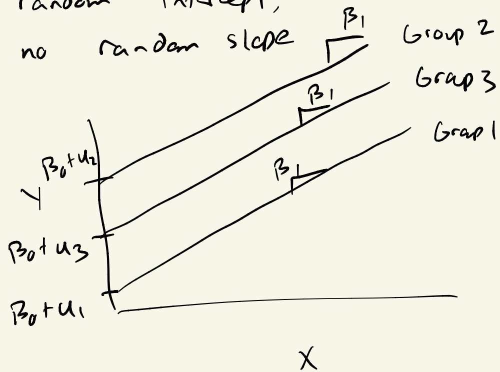
$$\begin{bmatrix} u_i \\ v_i \end{bmatrix} \overset{iid}{\sim} N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_u^2 & \rho_{uv}\sigma_u\sigma_v \\ \rho_{uv}\sigma_u\sigma_v & \sigma_v^2 \end{bmatrix} \right)$$

σ_u^2 = variance of u_i

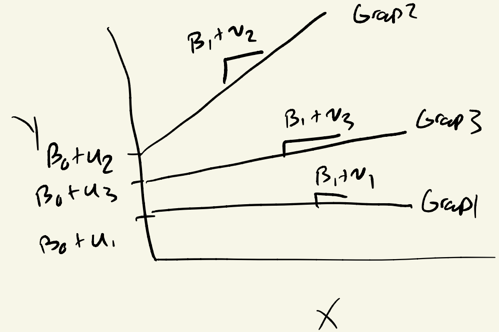
σ_v^2 = variance of v_i

ρ_{uv} = correlation between u_i and v_i

random intercept,
no random slope



random slopes + intercepts



(see., e.g., Q 7 on pg. 25
of review questions)

Logistic regression

- Wald test for individual βs — $H_0: \beta_1 = 0$, $H_A: \beta_1 \neq 0$; $H_0: \beta_1 = 1$, $H_A: \beta_1 > 1$
- LRT for multiple βs
  (G = deviance reduced − deviance full $\sim \chi^2_q$
  q = # parameters tested)

Multinomial regression

- Same as logistic regression
- for LRT, removing variables from all components simultaneously

---

Poisson regression
- quasi-Poisson
- NB
- ZIP

- χ^2 GOF test (residual deviance vs. residual df)
- dispersion test ($H_0: \phi = 1$ $H_A: \phi > 1$)
- Wald test for individual βs
  (inflate SEs by $\sqrt{\hat{\phi}}$ if $\phi \neq 1$)
- LRT for multiple βs
  - For Poisson: drop in deviance
  - For NB & ZIP: $2(\log L_{full} - \log L_{reduced})$

groups − p ≤ denominator df
≤ n − p
parameters in full model

Mixed effects models
approximate denominator df

- t tests for individual coefficients
- F tests for multiple coefficients (numerator df = # parameters tested)
- Bonus: Parametric bootstrap tests

Offsets in Poisson regression

- allow us to interpret βs in terms of rates
  (e.g. crimes per 1000 student, articles published per year, etc.)

- useful when observations come from groups of different sizes (e.g.. different #s of students on a campus, different lengths of time over which data is recorded, etc.)

- can use offsets for Poisson regression, NB regression, quasi-Poisson, or Poisson component of ZIP model

$Y_i \sim \text{Poisson}(\lambda_i)$

$\log(\lambda_i) = \beta_0 + \beta_1 X_i$

95% CI for β_1: $\hat{\beta}_1 \pm 1.96\, SE_{\hat{\beta}_1}$

"we are 95% confident that a unit increase in X is associated with a change in the log mean of Y by between .... and ...."

95% CI for e^{β_1}: $e^{\hat{\beta}_1 \pm 1.96\, SE_{\hat{\beta}_1}}$

"we are 95% confident that a unit increase in X is associated with a change in the mean of Y by a factor of between .... and ...."