Week 4 (I) Math Notes: Multilevel Data

Stat 209

Starting point $Y_{ij} = \bar{Y}_{i.} + (Y_{ij} - \bar{Y}_{i.})$ $\quad X_{ij} = \bar{X}_{i.} + (X_{ij} - \bar{X}_{i.})$ between/w/in

2 Levels, indivs w/ groups.

Basic Levels of Analysis Relations (DCD)

Regression $\quad \beta^{t}_{YX} = \eta^2_X \beta^{b}_{\bar{Y}\bar{X}} + (1 - \eta^2_X)\beta^{w-p}_{YX}$

Correlation: $\quad \rho^{t}_{YX} = \eta_X \eta_Y \rho^{b}_{\bar{Y}\bar{X}} + \sqrt{(1-\eta^2_X)(1-\eta^2_Y)}\,\rho^{w-p}_{YX}$

$\rho^{t}_{YX} = \rho^{b}_{\bar{Y}\bar{X}}$ iff $\rho^{w-p}_{YX} = \left(\dfrac{1 - \eta_X \eta_Y}{\sqrt{1-\eta^2_X}\sqrt{1-\eta^2_Y}}\right)\rho^{t}_{YX}$

where β^{t}_{YX} indiv ignore group $\quad \beta^{b}_{\bar{Y}\bar{X}}$ group mean "between" $\quad \beta^{w-p}_{YX}$ w/in group pooled, relative standing

ρ^{t}_{YX}
$\rho^{b}_{\bar{Y}\bar{X}}$
ρ^{w-p}_{YX}

grouping measures $(0,1)$ $\quad \eta^2_X = \dfrac{Var(\bar{X})}{Var(X)} \quad \eta^2_Y = \dfrac{Var(\bar{Y})}{Var(Y)}$

Derivation Y, X (unobserved) grouping var U

means $E(X|U) \cdot E(Y|U)$

Standard Decompositions: Stat 200

between $\qquad$ w/in

conditional variance $\quad Var(X) = Var(E(X|U)) + E(Var(X|U))$

$Var(Y) = Var(E(Y|U)) + E(Var(Y|U))$

conditional covariance $\quad Cov(Y,X) = Cov(E(X|U), E(Y|U)) + E(Cov(X,Y|U))$

Define $\eta^2_X = \dfrac{Var(E(X|U))}{Var(X)} \quad \beta^{t}_{YX} = \dfrac{Cov(X,Y)}{Var(X)}$

$\beta^{w-p}_{YX} = E(Cov(Y,X|U))/E(Var(X|U))$ $\quad \beta^{b}_{\bar{Y}\bar{X}} = Cov(E(X|U), E(Y|U))/Var(E(X|U))$

$$\frac{Cov(Y,X)}{Var(X)} = \frac{Cov(E(X|U), E(Y|U))}{Var(X)}\left(\frac{Var(E(X|U))}{Var(E(X|U))}\right)$$

$$+ \frac{E(Cov(X,Y|U))}{Var(X)}\left(\frac{E(Var(X|U))}{E(Var(X|U))}\right)$$

regroup to above

~~scale~~ repeat for ρ

Deep review: Anova, Nested Designs

Neter et al Ch 26, Training Example

$$Y_{ijk} = \mu + \alpha_i + \beta_{j(i)} + \varepsilon_{ijk}$$

α effect of school

$\beta_{j(i)}$ instructors nested w/in school

$$Y_{ijk} - \bar{Y}_{...} = \underbrace{(\bar{Y}_{i..} - \bar{Y}_{...})}_{\text{A main effect}} +$$

B w/in A $\quad (Y_{ij.} - \bar{Y}_{i..}) + (Y_{ijk} - \bar{Y}_{ij.})_{error}$

Anova Table

Schools	SSA
Instructors w/in School	SSB(A)
Error	SSE

Levels of Analysis, Aggregation Bias

Stab209 week 4

Robinson Table 3 2×2 Table review

```
> litnat = matrix(c(1304,11913,2614,81441), nr = 2,
          dimnames = list("Literacy" = c("No", "Yes"),
> litnat                                  "ForB" = c("Y", "N")))
         ForB
Literacy     Y     N
     No   1304  2614
     Yes 11913 81441
> chisq.test(litnat, correct = FALSE)

        Pearson's Chi-squared test

data:  litnat
X-squared = 1348.6, df = 1, p-value < 2.2e-16

> phicoeff = sqrt(1348.6/97272)
>  phicoeff                        # vs Robinson reported corr = .118
[1] 0.1177464
> # get conditional probabilities (cond'l on column)
> prop.table(litnat,2)
        ForBorn
Literacy          Y          N
     No  0.09866082 0.03109869
     Yes 0.90133918 0.96890131
> # so relative risk illit > 3
```

Ecological fallacy ex
Literacy, Foreign Born correlations

48 states −.53
9 regions −.62

sample $\phi_u = \sqrt{\frac{\chi^2}{n_{++}}}$
"phi coeff"

relative risk (n/f)

or odds ratio
vcd package

Multilevel, Contextual effects measured vars

$\beta^t_{YX} - \beta^b_{\bar{Y}\bar{X}}$ Aggregation bias

β^t_{YX}: indiv level regression ignoring groups

$\beta^b_{\bar{Y}\bar{X}}$: group level regression

$\beta_{Y\bar{X}\cdot X} = \beta^b_{\bar{Y}\bar{X}} - \beta^{w-p}_{YX}$

Contextual effect — effect of group on individ
mult regr interpretation
"increase in Y for increase $\bar{X}$ with X constant"
As if by experiment?

NELS data $\hat{\beta}^t = 3.6$ $\hat{\beta}^b = 7$
$\hat{\beta}^{wp} = 2.1$