Group Membership

(4ummy Variables)

Define $G_i = 1$ if unit $i$ is in group 1 (e.g. treatment)

$G_i = 0$ otherwise

(i.e. unit $i$ is in group 0, control)

$Y_i$ outcome for unit $i$.
$E(Y_i | G = 1) = \mu_1$ ; $E(Y_i | G = 0) = \mu_0$

$E(Y_i | G) = \beta_0 + \beta_1 G$

$\beta_0 = \mu_0$  \hspace{1cm} $\beta_1 = \mu_1 - \mu_0$
Nonexperimental Data:

Dichotomous and continuous predictors

$Y = \text{salary}$

$X = \text{experience (yrs on job)}$

Education level -- 3 groups

$(\text{HS}, \text{BS}, \text{Adv Deg})$

Management or Not

$E_1 = 1$ for HS  $E_1 = 0$ otherwise

$E_2 = 1$ for BS  $E_2 = 0$ otherwise

$M = 1$ for management pos

0 otherwise.
Analysis of Covariance

Week Ch.10 indicator vars: Chap 25 ancova
5 one-way 25.1-23.42

Using concurrent (pretest) information to improve precision of group comparisons.

Comparing conditional expectations $E(Y|x)$ more precise than comparing (unconditional) group means ($Y$).

Regression Approach: Null $p<.069$

Single classification: outcome $Y$, covariate $X$;

$I-1$ Group membership vars $G_1, \ldots, G_{I-1}$. 

Week 25.3 overview.
Comparing Conditional Expectations: Ancova
2 groups (1 and 0) outcome \( Y \), covariate \( X \)

Straight-line regressoin function within each group

\[
E(Y|X) = \alpha_1 + \beta_1 X \quad \text{(Group 1)}
\]

\[
E(Y|X) = \alpha_0 + \beta_0 X \quad \text{(Group 0)}
\]

Ancova restrictions: \( \beta_1 = \beta_0 = \beta \)
\[
\sigma_1^2 = \sigma_0^2 = \sigma^2 \quad \text{(equal residual variances)}
\]

Fit a straight-line within each group

For Group 1:
\[
\hat{Y} = \hat{\alpha}_1 + \hat{\beta}_1 X = \bar{Y}_1 + \hat{\beta}_1 (X - \bar{X}_1)
\]

For Group 0:
\[
\hat{Y} = \hat{\alpha}_0 + \hat{\beta}_0 X = \bar{Y}_0 + \hat{\beta}_0 (X - \bar{X}_0)
\]
Analysis of covariance presumes equal within group slopes (11 lines) so $\hat{\delta}_1, \hat{\delta}_0$ estimate common $\delta$

Replace within group slopes by \text{\footnotesize \textbf{NWU} 23.35}

\[
\hat{\delta}_p = \frac{\hat{\delta}_1 \text{ SS}x_1 + \hat{\delta}_0 \text{ SS}x_0}{\text{ SS}x_1 + \text{ SS}x_0}
\]

to obtain two parallel within group regressions

\[
\hat{y} = \bar{y}_1 + \hat{\delta}_p (x - \bar{x}_1)
\]

\[
\hat{y} = \bar{y}_0 + \hat{\delta}_p (x - \bar{x}_0)
\]

Adjusted mean difference is vertical distance between 11 lines

\[
\bar{y}_1 + \hat{\delta}_p (x - \bar{x}_1) - [\bar{y}_0 + \hat{\delta}_p (x - \bar{x}_0)]
\]

\[
= \bar{y}_1 - \bar{y}_0 - \hat{\delta}_p (\bar{x}_1 - \bar{x}_0)
\]

\text{\footnotesize aucoed adjustment}.
adjusted mean difference
ancova estimate of treatment effect (G1 vs GO membership)

corrected adj treatment means Nuck 23.40, Fig 23.97
precision of group comparison (95%)

Improved in ancova to the extent that Y/X more precise than Y
(i.e. magnitude of $\sqrt{1-r^2_{xy}}$)
Ancova via Regression with Group membership variables.

For each individual we have $Y, X, \text{ group membership } (G_1, \ldots, G_{i-1})$

Two groups $G = 1, G = 0$

General model:

$$E(Y|X, G) = \beta_0 + \beta_1 G + \beta_2 X + \beta_3 X_G$$

within groups

$$E(Y|X, G = 1) = (\beta_0 + \beta_1) + (\beta_2 + \beta_3)X$$

$$E(Y|X, G = 0) = \beta_0 + \beta_2 X$$
Thus difference between ungroup slopes $\beta_3 = (\delta_1, -\delta_2)$

Ancova assumes $\beta_3 = 0$

Thus ancova model is

$$E(Y|X, G) = \delta_0 + \delta_1 G + \delta_2 X$$

< Model eqs 10.4, 10.5 >

$\hat{\delta}_2 = \delta_0$ (common ungroup slope)

$\hat{\delta}_1$ = adjusted mean difference (ancova treatment effect)

Ancova $H_0: \delta_1 = 0$ $H_a: \delta_1 \neq 0$

Inferences for $\delta_1$ $\hat{\delta}_1 / \text{s.e.}(\hat{\delta}_1)$

Test statistic $\hat{\delta}_1 / \text{s.e.}(\hat{\delta}_1)$

Interval estimate for $\delta_1$. 
Comparing Nonparallel Regr Lines

Rogosa (1980)

2 group comparison single background variable \((X)\), outcome \((Y)\)

\[ G = 1, 0 \]

Fit separate within group regr's via

\[ E(Y|G,X) = \beta_0 + \beta_1 G + \beta_2 X + \beta_3 X \times G \]

Treatment effect as a function of \(X\):

\[ \Delta(X) = \beta_1 + \beta_3 X \]

ATI research (TTI)

(ancova assumes/requires \( \beta_3 = 0 \), why do it?)

Point estimate \( \Delta(X) = \hat{\beta}_1 + \hat{\beta}_3 X \)

<see cnr1lis, dat>

<see NML paper online>