Multilevel Factors and Contrasts

Salary Example: Let

x_1 = indicator of Asst
x_2 = indicator of Assoc
x_3 = indicator of Prof

Book uses u_1, u_2, u_3 and calls these "dummy" variables.

Parametrizing the Model

First try $E[Y|\underset{\sim}{x}] = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$

$$X = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix}$$

Estimation problem: X[,1] = X[,2] + X[,3] + X[,4]
X is not full rank
X'X is not invertible
NO unique OLS solution

Model problem:
$E[Y|\text{Asst}] = \beta_0 + \beta_1 = \mu_1$
$E[Y|\text{Assoc}] = \beta_0 + \beta_2 = \mu_2$
$E[Y|\text{Prof}] = \beta_0 + \beta_3 = \mu_3$

Too many parameters!
The model is overparametrized!
We need to drop a column of X.

Solution 1: "ANOVA parametrization" (drop the intercept)

$$X = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

Corresponds to $E[Y|\underset{\sim}{x}] = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$

$E(Y|\text{Asst}) = \beta_1 = \mu_1$
$E(Y|\text{Assoc}) = \beta_2 = \mu_2$
$E(Y|\text{Prof}) = \beta_3 = \mu_3$

Can show that
$$\begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{pmatrix} = \begin{pmatrix} \bar{y}_1 \\ \bar{y}_2 \\ \bar{y}_3 \end{pmatrix} = \begin{pmatrix} \text{mean}(y[x_1 == 1]) \\ \vdots \\ \vdots \end{pmatrix}$$

Solution 2: "Set to zero parametrization": eliminate x_1
default in R and other software

$$X = \begin{array}{cccc} \text{int} & \text{Assoc} & \text{Prof} \\ \left[\begin{array}{c} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{array}\right. & \begin{array}{c} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \end{array} & \left.\begin{array}{c} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{array}\right] \end{array}$$

β_1 "set to zero"

$$E[Y \mid \underline{x}] = \beta_0 + \beta_2 x_2 + \beta_3 x_3$$

$$\begin{aligned} E(Y \mid \text{Asst}) &= \beta_0 &&= \mu_1 \\ E(Y \mid \text{Assoc}) &= \beta_0 + \beta_2 &&= \mu_2, \quad \beta_2 = \mu_2 - \mu_1 \\ E(Y \mid \text{Prof}) &= \beta_0 + \beta_3 &&= \mu_3, \quad \beta_3 = \mu_3 - \mu_1 \end{aligned}$$

β_0 is the expect. for <u>baseline group</u>
β_2, β_3 represent <u>deviations from baseline</u>
"<u>contrasts</u>"

<u>Testing + Evaluating Contrasts</u>

Suppose we want to estimate / test / obtain CI for $l = \underset{\sim}{a}^T \underset{\sim}{\beta}$

<u>eg</u>: test/estimate $\mu_{Prof} - \mu_{Assoc}$

$$\left.\begin{aligned} \mu_{prof} &= \beta_0 + \beta_{prof} \\ \mu_{Assoc} &= \beta_0 + \beta_{Assoc} \end{aligned}\right\} \rightarrow l = \beta_{prof} - \beta_{assoc}$$

So if $\underset{\sim}{\beta} = \begin{pmatrix} \beta_0 \\ \beta_{Assoc} \\ \beta_{prof} \end{pmatrix}$, $l = (0\ -1,\ 1) \cdot \beta$

Estimation: Under A1, $E[\hat{\beta} \mid x] = \beta$

$$\begin{aligned} E[a^T \hat{\beta} \mid x] &= a^T E(\hat{\beta} \mid x) \\ &= a^T \beta \end{aligned}$$

So $\hat{l} = a^T \hat{\beta}$ is an unbiased estimator of $l = a^T \beta$.

(3)

Testing / CI: t-stat: $\hat{\ell} / SE(\hat{\ell})$, CI: $\hat{\ell} \pm 2\, \widehat{SE}(\hat{\ell})$

Need $SE(\hat{\ell})$: Recall, under A1+A2,

$$V(\hat{\beta}) = \sigma^2 (X^T X)^{-1} = E[(\hat{\beta} - \beta)(\hat{\beta} - \beta)^T]$$

$$V(a^T \hat{\beta}) = E[(a^T\hat{\beta} - a^T\beta)(a^T\hat{\beta} - a^T\beta)]$$

$a^T\hat{\beta} = \hat{\beta}^T a$

$$= E[a^T(\hat{\beta} - \beta)(\hat{\beta} - \beta)^T a]$$

$$= a^T E[(\hat{\beta} - \beta)(\hat{\beta} - \beta)^T] a$$

$$= a^T \sigma^2 (X^T X)^{-1} a$$

$$\Rightarrow \hat{V}(a^T \beta) = \hat{\sigma}^2 a^T (X^T X)^{-1} a$$

$$SE(a^T\beta) = \hat{\sigma} \sqrt{a^T (X^T X)^{-1} a}$$ (p. 103 in W)

Now, with A1 + A2 + A3, can do t-tests and form CIs.

(Return to numerical example)