Math Facts

Week 3, part I

Stat 209

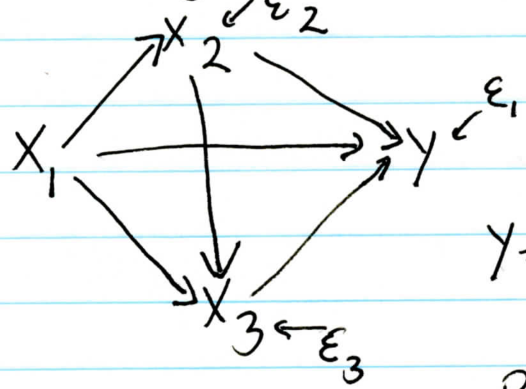
Path Analysis Basics

4-var, 3eq recursive

X_1 exogenous

standardized metric

ex Alcohol consumption on slides



$X_2 = \beta_{21} X_1 + \varepsilon_2$

$X_3 = \beta_{31} X_1 + \beta_{32} X_2 + \varepsilon_3$

$Y = \beta_{Y1} X_1 + \beta_{Y2} X_2 + \beta_{Y3} X_3 + \varepsilon_1$

paths from disturbance have coeff $\sqrt{1-R^2_{eq}}$ st dev residual

Solve Normal Equations to get association decomposition "Direct" and "Indirect" effects

(Sewell Wright)

correl's

total = direct + indirect

$$\rho_{12} = \beta_{21}$$

$$\rho_{13} = \underset{\text{[direct]}}{\beta_{31}} + \underset{\text{[indirect]}}{\beta_{32}\beta_{21}}$$

$$\rho_{1Y} = \beta_{Y1} + \underbrace{\beta_{Y2}\beta_{21} + \beta_{Y3}(\beta_{31} + \beta_{32}\beta_{21})}_{\text{indirect}}$$

$$\rho_{2Y} = \underset{\text{direct}}{\beta_{Y2}} + \underset{\text{indirect}}{\beta_{Y3}\beta_{32}} + \underbrace{\beta_{Y1}\beta_{21} + \beta_{Y3}\beta_{31}\beta_{21}}_{\text{common cause}}$$

R implementations on handout (our) and examples [Lab1 using corr, cov matrix]

s.e.m extension (sem handout)

latent vars LISREL, CALIS, R

Path Analysis Formulas (Corr matrix)

Items r_{xx} (cor Pred) r_{xy} (cor Pred Resp) n, p

Coeff $\hat{\beta} = r_{xx}^{-1} r_{xy}$

R^2 $\hat{\beta} r_{xy}$ residual st error $= \sqrt{1-R^2}$

$s.e.(\hat{\beta}_k) = \sqrt{(1-R^2) r_{xx}^{-1} / n-p}$

sqrt diagonal elements of $\frac{(1-R^2)}{n-p} r_{xx}^{-1}$ $\boxed{\text{cf Lab1 HW3}}$

Path Analysis Basics

path handout computations

C.f. Lab 1

```
> # Maruyama (1988) Basics of structural equation modeling ex p.57
> selfesteempredR = matrix(c(1, .39, -.33, .39, 1, -.33, -.33, -.33,1), nr=3)
> selfesteempredR
      [,1]  [,2]  [,3]
[1,]  1.00  0.39 -0.33
[2,]  0.39  1.00 -0.33
[3,] -0.33 -0.33  1.00
> selfesteemR = c(.19, .14, -.14)
> pathcoeff = selfesteemR%*%solve(selfesteempredR)
> pathcoeff     #coeffs for ability social class famsize respectively
          [,1]       [,2]        [,3]
[1,] 0.1423315 0.06036429 -0.07311039
> selfesteemR%*%t(pathcoeff)  #Rsq for eq
           [,1]
[1,] 0.04572944
```

$\sqrt{1-Rsq}$ is st dev residuals path coeff for disturbance see below

standard errors for path coeffs from $cov(\hat{\beta}) = \frac{(1-R^2)(r_{xx})^{-1}}{n-p}$ (or $n-p-1$) standardized

See Lab 1

DAF

```
> #Blau-Duncan, stratification US (DAF p.76)
> #Do the Y-eq (son occupation)
> bdpredR = matrix(c(1,.538,.417,.538,1,.438,.417,.438,1), nr = 3, byrow=T)
> bdpredR
      [,1]  [,2]  [,3]
[1,] 1.000 0.538 0.417
[2,] 0.538 1.000 0.438
[3,] 0.417 0.438 1.000
> bdYR = c(.541,.596,.405)
> bdYcoef = bdYR%*%solve(bdpredR)

> bdYcoef   #coefs for W U X respectively (cf DAF p.76)
          [,1]      [,2]      [,3]
[1,] 0.2807282 0.3945428 0.1151266

> bdYR%*%t(bdYcoef)  #Rsq for eq
          [,1]
[1,] 0.4336477
> sqrt( 1 - bdYR%*%t(bdYcoef))  #see p.76 Y eq disturbance term
          [,1]
[1,] 0.7525638

> #third example, Kline fitness
```

CAUSAL MODELS

STAT 209 ~~XX~~

1. Path Analysis

A

```
> # Maruyama (1988) Basics of structural equation modeling ex p.57
> selfesteempredR = matrix(c(1, .39, -.33, .39, 1, -.33, -.33, -.33,1), nr=3)
> selfesteempredR
      [,1]  [,2]  [,3]
[1,]  1.00  0.39 -0.33
[2,]  0.39  1.00 -0.33
[3,] -0.33 -0.33  1.00
> selfesteemR = c(.19, .14, -.14)
> pathcoeff = selfesteemR%*%solve(selfesteempredR)
> pathcoeff     #coeffs for ability social class famsize respectively
          [,1]       [,2]        [,3]
[1,] 0.1423315 0.06036429 -0.07311039
> selfesteemR%*%t(pathcoeff)  #Rsq for eq
           [,1]
[1,] 0.04572944
```

$\underset{n\times p}{X}$ transformed to (0,1) standardized metric

$r_{xx} = (X'X)$ $(p\times p)$

r_{yx}

$r_{xx} b = r_{yx}$

```
class    1.00 . . . .
famsize -.33 1.00  . . .
ability  .39 -.33 1.00 . .
esteem   .14 -.14  .19 1.00 .
achieve  .43 -.28  .67  .22 1.00
```

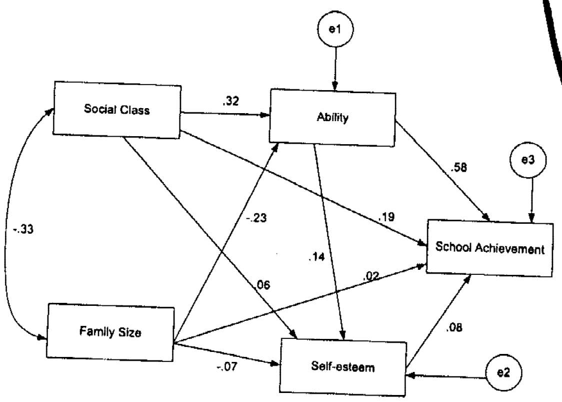
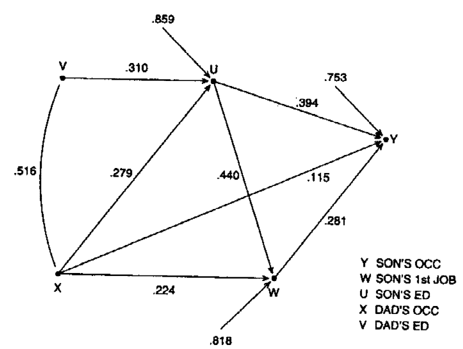


Table 1. Correlation matrix for variables in Blau and Duncan's path model.

		Y Son's occ	W Son's 1st job	U Son's ed	X Dad's occ	V Dad's ed
Y	Son's occ	1.000	.541	.596	.405	.322
W	Son's 1st job	.541	1.000	.538	.417	.332
U	Son's ed	.596	.538	1.000	.438	.453
X	Dad's occ	.405	.417	.438	1.000	.516
V	Dad's ed	.322	.332	.453	.516	1.000

Figure 1. Path model. Stratification, US, 1962.



B

```
> #Blau-Duncan, stratification US (DAF p.76)
> #Do the Y-eq (son occupation)
> bdpredR = matrix(c(1,.538,.417,.538,1,.438,.417,.438,1), nr = 3, byrow=T)
> bdpredR
      [,1]  [,2]  [,3]
[1,] 1.000 0.538 0.417
[2,] 0.538 1.000 0.438
[3,] 0.417 0.438 1.000
> bdYR = c(.541,.596,.405)
> bdYcoef = bdYR%*%solve(bdpredR)

> bdYcoef  #coefss for W U X respectively (cf DAF p.76)
          [,1]      [,2]      [,3]
[1,] 0.2807282 0.3945428 0.1151266

> bdYR%*%t(bdYcoef)  #Rsq for eq
          [,1]
[1,] 0.4336477
> sqrt( 1 - bdYR%*%t(bdYcoef))  #see p.76 Y eq disturbance term
          [,1]
[1,] 0.7525638
```

st errors for HW

Decomposing correlations (normal eq's)

C.

Path Analysis con'd

Roth, D. L., Wiebe, D. J., Fillingim, R. B., & Shay, K. A. (1989). Life events, fitness, hardiness, and health: A simultaneous analysis of proposed stress-resistance effects. *Journal of Personality and Social Psychology, 57,* 136–142.

Kline p. 117

TABLE 5.1. Analysis of a Recursive Path Model of Factor of Illness with Multiple Regression

Correlations, means, and standard deviations (Roth et al., 1989; $N = 373$ university students)

Variable	1	2	3	4	5
1. Exercise	—				
2. Hardiness	–.03	—			
3. Fitness	.39	.07	—		
4. Stress	–.05	–.23	–.13	—	
5. Illness	–.08	–.16	–.29	.34	—
M	40.90	0.00	67.10	4.80	716.70
SD	66.50	3.80	18.40	6.70	624.80

(b) Standardized Solution

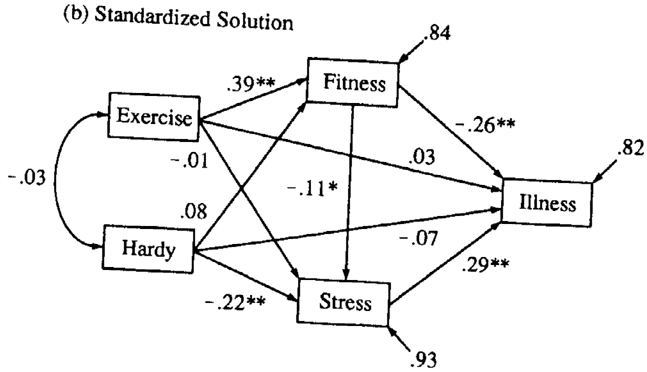


TABLE 5.2. Decomposition of Standardized Effects for a Model of Illness

Causal variable	Endogenous variable: Fitness	Stress	Illness
Exercise			
Direct effect	.39**	–.01	.03
Indirect via Fitness	—	–.04*	–.10**
Indirect via Stress	—	—	.00
Indirect via Fitness and Stress	—	—	–.01nt
Total effect	.39**	–.05	–.08
Hardiness			
Direct effect	.08	–.22**	–.07
Indirect via Fitness	—	–.01	–.02
Indirect via Stress	—	—	–.06*
Indirect via Fitness and Stress	—	—	.00nt
Total effect	.08	–.23**	–.15**
Fitness			
Direct effect	—	–.11*	–.26**
Indirect via Stress	—	—	–.03
Total effect	—	–.11*	–.29**
Stress			
Direct effect	—	—	.29**

Trace rule Kline p.121

path handout computations cf Lab 1 part II

```
> # Maruyama (1988) Basics of structural equation modeling ex p.57
> selfesteempredR = matrix(c(1, .39, -.33, .39, 1, -.33, -.33, -.33,1), nr=3)
> selfesteempredR
      [,1]  [,2]  [,3]
[1,]  1.00  0.39 -0.33
[2,]  0.39  1.00 -0.33
[3,] -0.33 -0.33  1.00
> selfesteemR = c(.19, .14, -.14)
> pathcoeff = selfesteemR%*%solve(selfesteempredR)
> pathcoeff     #coeffs for ability social class famsize respectively
          [,1]       [,2]        [,3]
[1,] 0.1423315 0.06036429 -0.07311039
> selfesteemR%*%t(pathcoeff)  #Rsq for eq
           [,1]
[1,] 0.04572944



> #Blau-Duncan, stratification US (DAF p.76)
> #Do the Y-eq (son occupation)
> bdpredR = matrix(c(1,.538,.417,.538,1,.438,.417,.438,1), nr = 3, byrow=T)
> bdpredR
      [,1]  [,2]  [,3]
[1,] 1.000 0.538 0.417
[2,] 0.538 1.000 0.438
[3,] 0.417 0.438 1.000
> bdYR = c(.541,.596,.405)
> bdYcoef = bdYR%*%solve(bdpredR)

> bdYcoef  #coefss for W U X respectively (cf DAF p.76)
          [,1]      [,2]      [,3]
[1,] 0.2807282 0.3945428 0.1151266

> bdYR%*%t(bdYcoef)  #Rsq for eq
          [,1]
[1,] 0.4336477
> sqrt( 1 - bdYR%*%t(bdYcoef))  #see p.76 Y eq disturbance term
          [,1]
[1,] 0.7525638

> #third example, Kline fitness
```