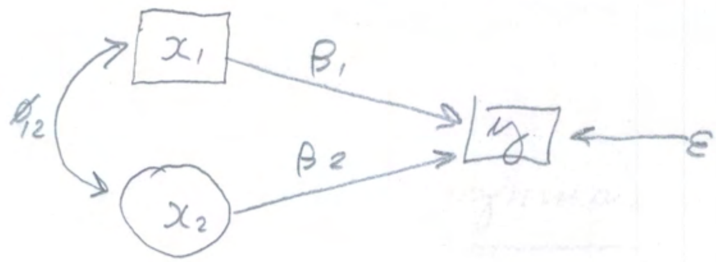
STA 431 Assignment 5

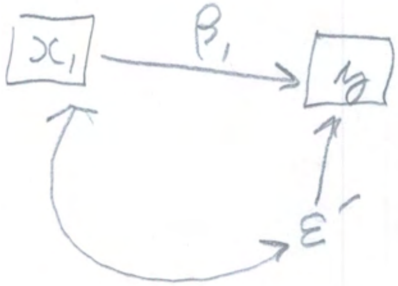
1. (a)



(b) $Cov(x_{i,1}, \varepsilon_i') = cov(x_{i,1}, \beta_2 x_{i2} + \varepsilon_i)$

$$= \beta_2 \phi_{12}$$

(c)



(d)

	x_1	y
x_1	ϕ_{11}	$\beta_1\phi_{11} + \beta_2\phi_{12}$
y		$\beta_1^2\phi_{11} + \beta_2^2\phi_{22} + 2\beta_1\beta_2\phi_{12} + \sigma^2$

$$cov(x_1, y) = cov(x_1, \beta_0' + \beta_1 x_1 + \varepsilon')$$
$$= \beta_1\phi_{11} + \beta_2\phi_{12}$$

(1e) $$\hat{\beta}_1 = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_{i1}-\bar{x}_1)(y_i-\bar{y})}{\frac{1}{n}\sum_{i=1}^{n}(x_{i1}-\bar{x}_1)^2} = \frac{\hat{\sigma}_{xy}}{\hat{\sigma}_x^2}$$

$$\xrightarrow{p} \frac{\sigma_{xy}}{\sigma_x^2} = \frac{\beta_1\phi_{11}+\beta_2\phi_{12}}{\phi_{11}} = \beta_1 + \frac{\beta_2\phi_{12}}{\phi_{11}}$$

(f) $\hat{\beta}_1 \xrightarrow{c}$ at all points where $\beta_2 = 0$, $\phi_{12} = 0$, or both

(2) $$f_{\varepsilon|x}(\varepsilon|x) = \frac{f_{\varepsilon,x}(\varepsilon,x)}{f_x(x)} = \underset{\uparrow \atop N(0,\sigma^2)}{f_\varepsilon(\varepsilon)}$$

$\Rightarrow f_{\varepsilon,x}(\varepsilon,x) = f_\varepsilon(\varepsilon)\,f_x(x)$ <u>Independence</u>

And so $Cov(X,\varepsilon) = 0$

(3) First, $E(\varepsilon) = E\{E(\varepsilon|X)\} = E\{0\} = 0$

Then $Cov(X,\varepsilon) = E(X\varepsilon) - E(X)\underbrace{E(\varepsilon)}_{0}$

$$= E(X\varepsilon) = E_x\{E_\varepsilon(X\varepsilon|X)\}$$

$$= E_x\{X E_\varepsilon(\varepsilon|X)\} = E\{X\cdot 0\} = E\{0\}$$

$$= 0$$

4 (a) $E(x_i y_i) = E(x_i, \beta x_i + \varepsilon_i) = \beta E(x_i^2) + E(x_i \varepsilon_i)$

Since $Cov(x_i, \varepsilon_i) = E(x_i \varepsilon_i) - E(x_i) E(\varepsilon_i)$ and $E(\varepsilon_i) = 0$, $E(x_i y_i) = \beta E(x_i^2) + c$, so

by Law of Large numbers and continuous mappings,

$$\hat{\beta}_n = \frac{\frac{1}{n}\sum_{i=1}^n x_i y_i}{\frac{1}{n}\sum_{i=1}^n x_i^2} \xrightarrow{p} \frac{E(x_i y_i)}{E(x_i^2)} = \frac{\beta E(x_i^2) + c}{E(x_i^2)}$$

$$= \beta + \frac{c}{E(x_i^2)} = \beta + \frac{c}{\sigma_x^2 + \mu_x^2} \neq \beta \text{ unless } c = 0$$

(b) $\bar{y}_n = \frac{1}{n}\sum_{i=1}^n y_i = \frac{1}{n}\sum_{i=1}^n (x_i \beta + \varepsilon_i) = \beta \bar{x}_n + \bar{\varepsilon}_n$,

so $\tilde{\beta}_n = \frac{\bar{y}_n}{\bar{x}_n} = \frac{\beta \bar{x}_n + \bar{\varepsilon}_n}{\bar{x}_n} = \beta + \frac{\bar{\varepsilon}_n}{\bar{x}_n}$

$$\xrightarrow{p} \beta + \frac{0}{\mu_x} = \beta$$

not zero

"Consistent," if the parameter space does not include points with $\mu_x = 0$.

5 (a) $$E\begin{pmatrix} x_i \\ y_i \\ z_i \end{pmatrix} = \begin{pmatrix} \mu_x \\ \hline \beta_0 + \beta_1 \mu_x \\ \hline \mu_z \end{pmatrix}$$

$$\mathrm{cov}\begin{pmatrix} x_i \\ y_i \\ z_i \end{pmatrix} = \begin{array}{c|c|c|c} & x & y & z \\ \hline x & \Phi_x & \Phi_x \beta_1^T + \varsigma & \kappa \\ \hline y & & \beta_1 \Phi_x \beta_1^T + \beta_1 \varsigma + \varsigma^T \beta_1^T + \psi & \beta_1 \kappa \\ \hline z & & & \Phi_z \end{array}$$

$$\mathrm{cov}(x, y) = \mathrm{cov}(x, \beta_1 x + \varepsilon) = \Phi_x \beta_1^T + \varsigma$$

$$\mathrm{cov}(y, y) = \mathrm{cov}(\beta_1 x + \varepsilon, \beta_1 x + \varepsilon) = \beta_1 \Phi_x \beta_1^T + \beta_1 \varsigma + \varsigma^T \beta_1^T + \psi$$

$$\mathrm{cov}(y, z) = \mathrm{cov}(\beta_1 x + \varepsilon, z) = \beta_1 \kappa$$

Or check the lecture slides

(5b) (i) $K = \Sigma_{13}$ and $\beta_1 K = \Sigma_{23}$, so

$$\beta_1 = \Sigma_{23} \Sigma_{13}^{-1}$$

(ii) $\hat{\beta}_1 = \hat{\Sigma}_{23} \hat{\Sigma}_{13}^{-1}$

(iii)
- $\Phi_x = \Sigma_{11}$
- $\Phi_z = \Sigma_{33}$
- $K = \Sigma_{13}$
- $\beta_1 = \Sigma_{23} \Sigma_{13}^{-1}$
- $\varsigma = \Sigma_{12} - \Phi_x \beta_1^T$
- $\Psi = \Sigma_{22} - \beta_1 \Phi_x \beta_1^T - \beta_1 \varsigma - \varsigma^T \beta_1^T$

6

6 (a) Independently for $i = 1, \ldots, n$,

$$y_{i1} = \alpha_0 + \alpha_1 x_i + \varepsilon_{i1}$$

$$y_{i2} = \beta_0 + \beta_1 x_i + \varepsilon_{i2} \quad \text{, where}$$

- $E(x_i) = \mu_x$, $E(z_i) = \mu_z$, $E(\varepsilon_{i1}) = E(\varepsilon_{i2}) = 0$
- $Var(x_i) = \phi_x$ $\quad Var(z_i) = \phi_z$
- $cov\begin{pmatrix} \varepsilon_{i1} \\ \varepsilon_{i2} \end{pmatrix} = \begin{pmatrix} \psi_{11} & \psi_{12} \\ \psi_{12} & \psi_{22} \end{pmatrix}$
- $cov(x_i, \varepsilon_{i1}) = c_1 \qquad cov(x_i, \varepsilon_{i2}) = c_2$
- $cov(x_i, z_i) = K$

(b) $\underset{\sim}{y}_i = \begin{pmatrix} y_{i1} \\ y_{i2} \end{pmatrix}$, $\quad \underset{\sim}{\varepsilon}_i = \begin{pmatrix} \varepsilon_{i1} \\ \varepsilon_{i2} \end{pmatrix}$

$$\underset{\sim}{\beta}_0 = \begin{pmatrix} \alpha_0 \\ \beta_0 \end{pmatrix}, \quad \underset{\sim}{\beta}_1 = \begin{pmatrix} \alpha_1 \\ \beta_1 \end{pmatrix}$$

$$\underset{\sim}{\Psi} = \begin{pmatrix} \psi_{11} & \psi_{12} \\ \psi_{12} & \psi_{22} \end{pmatrix} \qquad \underset{\sim}{C} = (c_1 \quad c_2)$$

$$\underset{\sim}{K} = (K) \quad 1 \times 1$$

(6c) $y_1 = \alpha_0 + \alpha_1 x + \varepsilon_1$
$y_2 = \beta_0 + \beta_1 x + \varepsilon_2$

	x	y_1	y_2	z
x	ϕ_x	$\alpha_1 \phi_x + c_1$	$\beta_1 \phi_x + c_2$	K
y_1		$\alpha_1^2 \phi_x + 2\alpha_1 c_1 + \psi_{11}$	$\alpha_1 \beta_1 \phi_x + \alpha_1 c_2 + \beta_1 c_1 + \psi_{12}$	$\alpha_1 K$
y_2			$\beta_1^2 \phi_x + 2\beta_1 c_2 + \psi_{22}$	$\beta_1 K$
z				ϕ_z

$Var(y_1) = cov(\alpha_1 x + \varepsilon_1, \alpha_1 x + \varepsilon_1) = \alpha_1^2 \phi_x + 2\alpha_1 c_1 + \psi_{11}$

$$cov(y_1, y_2) = cov(\alpha_1 x + \varepsilon_1, \beta_1 x + \varepsilon_2)$$
$$= \alpha_1 \beta_1 \phi_x + \alpha_1 c_2 + \beta_1 c_1 + \psi_{12}$$

$cov(y_1, z) = cov(\alpha_1 x + \varepsilon_1, z) = \alpha_1 K$

(d) $\tilde{\alpha}_1 = \dfrac{\hat{\sigma}_{24}}{\hat{\sigma}_{14}}$, $\tilde{\beta}_1 = \dfrac{\hat{\sigma}_{34}}{\hat{\sigma}_{14}}$

$$\begin{pmatrix} \hat{\alpha}_1 \\ \hat{\beta}_1 \end{pmatrix} = \hat{\Sigma}_{23} \hat{\Sigma}_{13}^{-1}$$

checks

$$\parallel$$

$$\begin{pmatrix} \hat{\sigma}_{24} \\ \hat{\sigma}_{34} \end{pmatrix} \cdot \frac{1}{\hat{\sigma}_{14}}$$

(6e) This is a special case of the general model given in Problem 5. We have already solved the equations in matrix form.

(f) $\theta = (\alpha_0, \alpha_1, \beta_0, \beta_1, \mu_x, \mu_z, \phi_x, \phi_z, \psi_{11}, \psi_{12}, \psi_{22}, c_1, c_2, \kappa_1)$ 14 parameters

(g) 4 expected values plus $4(4+1)/2 = 10$ unique variances and covariances for a total of 14 moments

(h) The moments are a function of the parameters (always true), and the parameters are a function of the moments. The function is 1-1, and invariance tells us that the MLE of the function of the $\hat{\sigma}_{ij}$ and $\hat{\mu}_j$ is the function of those MLEs.

A one word answer for full marks is "invariance."

Assignment 5, Question 7

```
> # Smoking and lung cancer.
>
> rm(list=ls())
> # install.packages("lavaan", dependencies = TRUE) # Only need to do this once
> library(lavaan)
This is lavaan 0.6-13
lavaan is FREE software! Please report any bugs.
>
> smoke = read.table("https://www.utstat.toronto.edu/brunner/openSEM/data/CancerTax2.data.txt")
> head(smoke)
       State CigaretteTax SmokingRate LungCancer BrainCancer
1    Alabama        0.675        20.2       59.5         6.3
2     Alaska        2.000        17.4       51.4         5.5
3    Arizona        2.000        14.9       39.2         5.8
4   Arkansas        1.150        20.2       72.9         6.1
5 California        2.870        10.0       37.1         5.6
6   Colorado        1.940        13.5       37.7         5.8
>
> # a) Fit the model
>
> mod1 = 'LungCancer ~ alpha0*1 + alpha1*SmokingRate # + epsilon1
+         BrainCancer ~ beta0*1 + beta1*SmokingRate  # + epsilon2
+           SmokingRate  ~ mux*1                     # E(SmokingRate)    = mux
+           CigaretteTax ~ muz*1                     # E(CigaretteTax)   = muz
+         LungCancer ~~ psi11*LungCancer             # Var(epsilon1) = psi11
+         BrainCancer ~~ psi22*BrainCancer           # Var(epsilon2) = psi22
+         LungCancer ~~ psi12*BrainCancer            # Cov(epsilon1,epsilon2) = psi12
+         SmokingRate ~~ phi11*SmokingRate           # Var(SmokingRate) = phi11
+         CigaretteTax ~~ phi22*CigaretteTax # Var(CigaretteTax) = phi22
+         SmokingRate ~~ kappa*CigaretteTax          # Cov(SmokingRate, CigaretteTax) = kappa
+         SmokingRate ~~ c1*LungCancer               # Cov(SmokingRate, epsilon1) = c1
+         SmokingRate ~~ c2*BrainCancer              # Cov(SmokingRate, epsilon2) = c2
+        ' # End of model string
>
> fit1 = lavaan(mod1, data=smoke)
> summary(fit1)
lavaan 0.6.13 ended normally after 57 iterations

  Estimator                                         ML
  Optimization method                           NLMINB
  Number of model parameters                        14

                                                  Used       Total
  Number of observations                            50          51

Model Test User Model:

  Test statistic                                 0.000
  Degrees of freedom                                 0

Parameter Estimates:

  Standard errors                             Standard
  Information                                 Expected
  Information saturated (h1) model          Structured

Regressions:
                   Estimate  Std.Err  z-value  P(>|z|)
  LungCancer ~
    SmkngRt (alp1)    1.383    0.576    2.401    0.016
  BrainCancer ~
    SmkngRt (bet1)    0.035    0.054    0.646    0.518
```

```
Covariances:
                   Estimate  Std.Err  z-value  P(>|z|)
 .LungCancer ~~
   .BrnCncr (ps12)    1.408    1.048    1.343    0.179
  SmokingRate ~~
    CgrttTx (kapp)   -2.369    0.652   -3.636    0.000
 .LungCancer ~~
    SmkngRt   (c1)   11.185    5.236    2.136    0.033
 .BrainCancer ~~
    SmkngRt   (c2)   -0.188    0.465   -0.405    0.685

Intercepts:
                   Estimate  Std.Err  z-value  P(>|z|)
   .LngCncr (alp0)   31.899    9.397    3.395    0.001
   .BrnCncr (bet0)    5.642    0.873    6.462    0.000
    SmkngRt  (mux)   16.198    0.465   34.804    0.000
    CgrttTx  (muz)    1.925    0.170   11.338    0.000

Variances:
                   Estimate  Std.Err  z-value  P(>|z|)
   .LngCncr (ps11)   64.570   18.241    3.540    0.000
   .BrnCncr (ps22)    0.557    0.113    4.920    0.000
    SmkngRt (ph11)   10.830    2.166    5.000    0.000
    CgrttTx (ph22)    1.442    0.288    5.000    0.000

>
> # b) MOM estimate of alpha1
>
> S = var(smoke[,2:5], na.rm=TRUE); S
             CigaretteTax SmokingRate LungCancer BrainCancer
CigaretteTax   1.47126004   -2.417818  -3.342772 -0.08357608
SmokingRate   -2.41781780   11.051220  26.692641  0.18980000
LungCancer    -3.34277176   26.692641 118.572004  2.09409388
BrainCancer   -0.08357608    0.189800   2.094094  0.56877143
> S[1,3]/S[1,2] # Should equal alpha1hat = 1.383
[1] 1.382557
> # Not surprising: Invariance. n-1 versus n does not matter because they
> # cancel in numerator and denominator.
>
> # (c) z = 2.401, p = 0.016. Higher smoking rate increases the rate of lung cancer.
>
> # (d) z = 0.646, p = 0.518. There is no evidence that smoking rate influences the
> # rate of brain cancer.
```