Name Jerry

Student Number

STA 431 Quiz 5

1. In Question 7 of this week's assignment, you tested $H_0 : \alpha_1 = 0$. *Do not answer this question if you do not have a printout.*

   (a) (2 points) Write the value of the test statistic and the p-value in the space below. These are numbers from your printout. On your printout, circle the numbers and write "Question 1" beside them.

   $Z = 2.401, \quad p = 0.016$

   (b) (2 points) In terms of the influence of smoking on cancer (which is the point of all this), what do you conclude from this test? If a conclusion is justified, draw a directional conclusion. *Be guided by the $\alpha = 0.05$ significance level, but do not mention it.* You have a lot mote room tan you need.

Smoking increases the chances of lung cancer.

one mark

(okay to say higher smoking rate increases the lung cancer rate, or anything like that.)

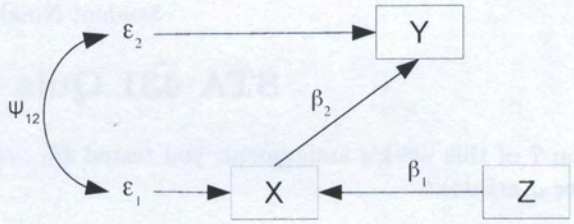
Even Smoking causes lung cancer.

But not "More smoking is associated with more lung cancer", because it's a causal model.

And not just "smoking causes cancer, because there was another kind of cancer (brain cancer) in the study.

2. In one version of the instrumental variables model, the instrumental variable z has a direct influence on the explanatory variable x. The model equations and path diagram are

$$\begin{aligned} X_i &= \alpha_1 + \beta_1 Z_i + \epsilon_{i1} \\ Y_i &= \alpha_2 + \beta_2 X_i + \epsilon_{i2} \end{aligned}$$



where $E(Z_i) = \mu_z$, $Var(Z_i) = \sigma_z^2$, $E(\epsilon_{i1}) = E(\epsilon_{i2}) = 0$, $Var(\epsilon_{i1}) = \psi_1$, $Var(\epsilon_{i2}) = \psi_2$ and $Cov(\epsilon_{i1}, \epsilon_{i2}) = \psi_{12}$.

(a) (2 points) Calculate $\sigma_{12} = Cov(Z_i, X_i)$. Show a little work and **circle your answer**.

$$\begin{aligned} Cov(Z_i, X_i) &= cov(Z_i, \beta_1 Z_i + \varepsilon_{i1}) \\ &= \beta_1 cov(Z_i, Z_i) + 0 \\ &= \beta_1 \sigma_z^2 \end{aligned}$$

(b) (2 points) Calculate $\sigma_{13} = Cov(Z_i, Y_i)$. Show a little work and **circle your answer**.

$$\begin{aligned} Cov(Z_i, Y_i) &= cov(Z_i, \beta_2 X_i + \varepsilon_{i2}) \\ &= cov(Z_i, \beta_2(\beta_1 Z_i + \varepsilon_{i1}) + \varepsilon_{i2}) \\ &= cov(Z_i, \beta_1\beta_2 Z_i + \beta_2 \varepsilon_{i1} + \varepsilon_{i2}) \\ &= \beta_1 \beta_2 \sigma_z^2 \end{aligned}$$

(c) (2 points) Give the formula for a Method of Moments estimate of the parameter β_2 in terms of $\widehat{\sigma}_{ij}$ values.

$$\hat{\beta}_2 = \hat{\sigma}_{13} / \hat{\sigma}_{12}$$

Please attach your printout to the quiz paper. The printout should show your *complete* R input and output. Make sure your name and student number appear on the printout.

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
```

Question 1

```
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.
```