Problem 1.-

Download dataset and run SAS to get the following descriptive statistics and run simple linear regression analysis. Dataset contains two columns: Y and X. Dependent variable is Y (first column) and independent variable is X (second column).

(1) Please show three digits

Sample size; n (from SAS/R)	
Sample mean of Y; $\bar{Y}$ (from SAS/R)	
Sample variance of Y; (from SAS/R) $$\boldsymbol{S_Y^2} = \frac{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}{(n-1)}$$	
Sample mean of X; $\bar{X}$ (from SAS) /R	
Sample variance of X; (from SAS/R) $$\boldsymbol{S_X^2} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}$$	
Sample correlation of X and Y; $\boldsymbol{r}$ (from SAS/R) $$r = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2 \sum_{i=1}^{n}(Y_i - \bar{Y})^2}}$$	
SSE; (from SAS/R) $$SSE = \sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$$	

Intercept; (from SAS/R) $\boldsymbol{\hat{\beta}_0}$	
Slope; (from SAS/R) $\boldsymbol{\hat{\beta}_1}$	
Standard error for slope estimate; (from SAS/R) $\boldsymbol{S_{\hat{\beta}_1}}$	

Coefficient of Determination; (from SAS/R) **R^2**	
95% confidence interval for $\boldsymbol{\beta_1}$	(,)
95% confidence interval for ρ	(,)
T statistics for ***H_0: β_1 = 0***	
F statistics for ***H_0: β_1 = 0***	

Use the above quantities to obtain the following estimates by hand calculation.

Show the equation first and use information in the above table to get the value.

(2) [10 points]

(a)

$\widehat{\boldsymbol{\beta}}_{\mathbf{1}} =$

(b)

$\widehat{\boldsymbol{\beta}}_{\mathbf{0}} =$

(c)

$\boldsymbol{S^2_{Y|X}} =$

(d)

R^2=

(3)

(a)

Interpretation: $\hat{\beta}_1$ and R^2.

For $\hat{\beta}_1$

For R^2

(b)

Obtain $\boldsymbol{S_{\hat{\beta}_1}} =$

(c)

Calculate T statistic for testing ***H_0: $\boldsymbol{\beta_1}$ = 0*** *vs.* ***H_A: $\boldsymbol{\beta_1}$ ≠ 0***

T= [ ] / [ ] =

(d)

Calculate F statistic for testing ***H_0: $\boldsymbol{\beta_1}$ = 0*** *vs.* ***H_A: $\boldsymbol{\beta_1}$ ≠ 0***

F= [ ] / [ ] =

(4)

(a)

Using either T statistic or F statistic, determine test result based on **5% significance level**. (Note: $t_{N-2,0.95}=1.66$; $t_{N-2,0.975}=1.98$; $t_{N-2,0.995}=2.63$;

or $F_{1,\ N-2,0.90}$=2.76; $F_{1,\ N-2,0.95}$=3.94; $F_{1,\ N-2,0.99}$=6.90)

(b)

Obtain **two-sided 95% confidence interval** of $\boldsymbol{\beta_1}$ (for slope);

Note $t_{N-2,0.95}$=1.66; $t_{N-2,0.975}$=1.98; $t_{N-2,0.995}$=2.63

[ ] $\pm$ [ ]$\times$[ ] = (,)

(5)

(a)

Calculate sample correlation coefficient *r* using the information in the above table.

r=

(b)

Testing ***H₀: ρ = 0*** *vs.* ***H_A: ρ ≠ 0*** *at* ***1% significance level (two-sided test)***

Note $t_{N-2,0.95}$=1.66; $t_{N-2,0.975}$=1.98; $t_{N-2,0.995}$=2.63

(c)

Testing $\boldsymbol{H_0: \rho = 0.8}$ *vs.* $\boldsymbol{H_A: \rho \neq 0.8}$ *(using table) at* ***5% significance level (two-sided test)*** Note $Z_{0.95}$=1.64; $Z_{0.975}$=1.96; and $Z_{0.995}$=2.58

Problem 2

(sample size=100) and reported the following results: **regression coefficient for Drug A=1.82** and **regression coefficient of Drug B=0.24**.

They want to test whether the joint effect of drug A and drug B is equal to 2 with regard to enhancing cognitive ability among patients who have brain surgery.

They also reported **variance of regression coefficient of drug A (0.31), variance of regression coefficient of drug B (0.23), and covariance of two regression coefficients of drug A and drug B (-0.015)**.

(1) Formulate the null hypothesis.

H_0:

(2) What is L (linear combination of regression coefficients)?

L=

(3) Test statistic=

T=

Note: Var(aX+bY)=a^2*Var(X)+b^2*Var(Y) + 2*a*b*cov(X,Y)

(4) What is your decision? (note: critical value=$t_{100-2-1,\ 0.975}$=1.98)

Then they want to test whether the effect of drug A is equal to that of drug B with regard to enhancing cognitive ability among patients who have brain surgery.

(5) Formulate the null hypothesis.

H_0:

(6) What is L (linear combination of regression coefficients)?

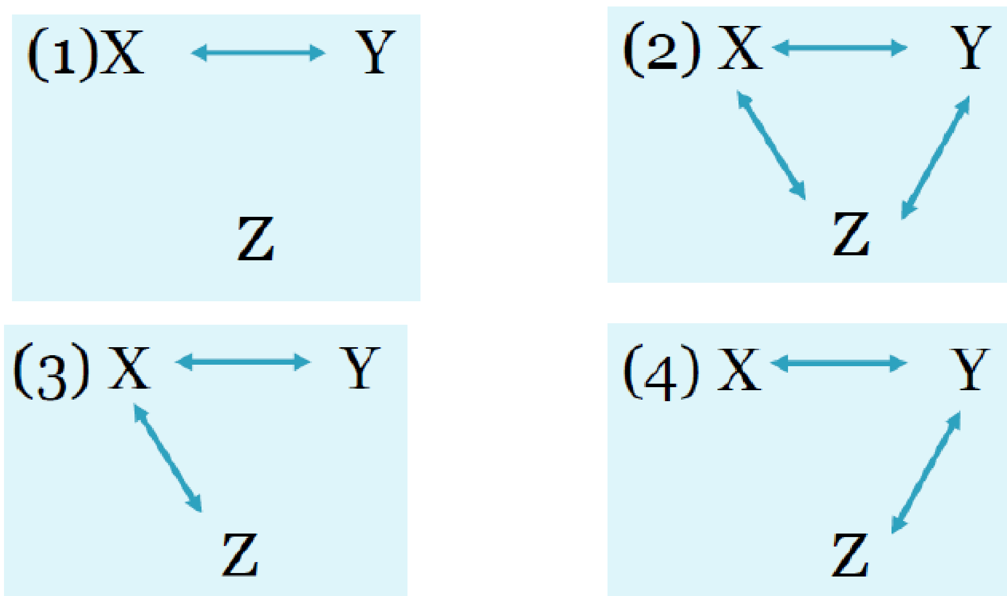
(7) Test statistic=

T=

(8) What is your decision? (note: critical value=$t_{100-2-1,\ 0.975}$=1.98)

Problem 3

(1) Which one has confounding variable in the following four cartoons?



(3) How can you identify confounding variable(s)?