

Quiz 2

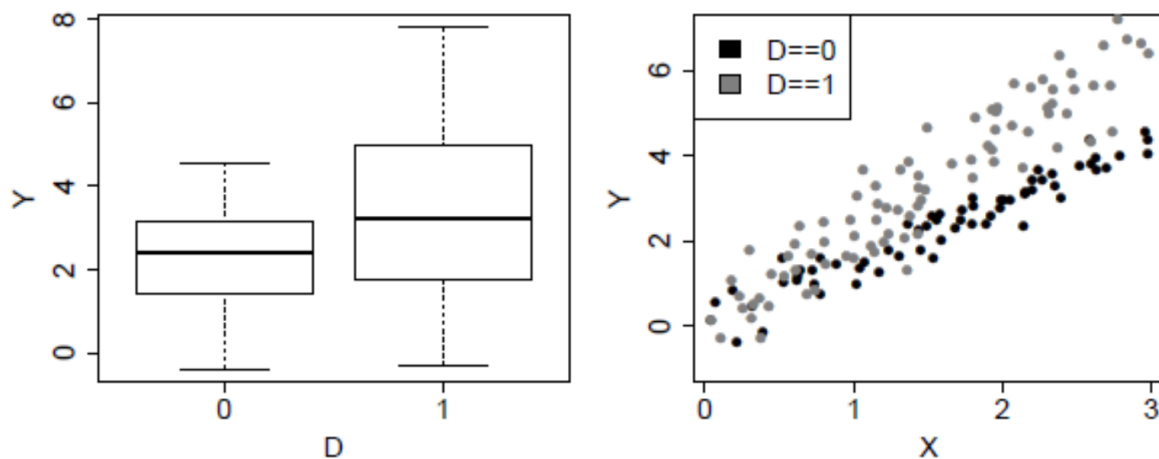
2023-06-06

Question 1

Prove: $SST=SSR+SSE$ (Please write out SST, SSR, SSE respectively and then prove the equality)

Question 2

Plotted below are data on 3 variables: a quantitative variable Y and X and a dummy variable D (either 0 or 1).



Based on these two plots, what model specification would you recommend where Y is the outcome and X and D are the independent variables? Write it down in term of $E(Y)$ as a function of the two independent variables.

Question 3

What is the level of stress (if any) that cows undergo prior to being slaughtered? To answer the question, the heart rate (beats per minute) of a cow was measured at four different pre-slaughter phases - (1) first phase of visual contact with pen mates, (2) initial isolation from pen mates for prepping, (3) restoration of visual contact with pen mates, and (4) first contact with human prior to slaughter. Eight cows were used for this experiment.

The researcher used the one-way ANOVA to determine whether the mean heart rate of cows differed in the four pre-slaughter phases.

```

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  108.000      3.615  29.873  <2e-16 ***
PHASE[T.2]   -11.000      5.113  -2.151  0.0402 *
PHASE[T.3]    -4.875      5.113  -0.953  0.3485
PHASE[T.4]    -2.875      5.113  -0.562  0.5784
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Multiple R-squared:  0.1511, Adjusted R-squared:  0.06015
F-statistic:      ?      on 3 and 28 DF,  p-value: 0.1979

```

(a) Write down H_0 and H_A in terms of μ .

(b) Write a conclusion based on the p-value approach with $\alpha = 0.05$ (clearly specify p-value)

(c) Compute each sample mean heart rate for four phases.(i.e. $\hat{Y}_1, \hat{Y}_2, \hat{Y}_3, \hat{Y}_4$)

(d) Construct an ANOVA summary table based on the above R output.

Source	df	SS	MS	F
PHASE	_____	_____	_____	_____
Error	_____	_____	_____	
Total	_____	_____		

(e) We re-analyzed this by the randomized block design (Encode every cow as a dummy variable). R output is as follows

```

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  120.688      4.056  29.757  < 2e-16 ***
PHASE[T.2]   -11.000      3.459  -3.180 0.004504 **
PHASE[T.3]    -4.875      3.459  -1.409 0.173332
PHASE[T.4]    -2.875      3.459  -0.831 0.415195
COW[T.2]     -17.250      4.891  -3.527 0.002003 **
COW[T.3]     -14.250      4.891  -2.913 0.008311 **
COW[T.4]     -21.500      4.891  -4.395 0.000252 ***
COW[T.5]      -1.000      4.891  -0.204 0.839977
COW[T.6]     -15.750      4.891  -3.220 0.004109 **
COW[T.7]     -21.000      4.891  -4.293 0.000322 ***
COW[T.8]     -10.750      4.891  -2.198 0.039322 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

```

Residual standard error: 6.917 on 21 degrees of freedom
Multiple R-squared:  0.7086, Adjusted R-squared:  0.5699
F-statistic: 5.108 on 10 and 21 DF,  p-value: 0.0008148

```

- i) Give the complete model appropriate for this design.

$E(Y) = \dots$

- ii) Give the reduced model appropriate for testing for differences in the mean heart beat rates of the four pre-slaughter phases.

$E(Y) = \dots$

- iii) Based on the following R output, test whether mean heart rate of cows differ in the four pre-slaughter phases.

Analysis of Variance Table

Model 1: BPM ~ COW

Model 2: BPM ~ PHASE + COW

Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
1	24	1526.0			
2	21	1004.9	3	521.12	(?) (?)

Analysis of Variance Table

Model 1: BPM ~ PHASE

Model 2: BPM ~ PHASE + COW

Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
1	28	2927.8			
2	21	1004.9	7	1922.9	(?) (?)

- First, write down hypothesis in terms of β : (Write H_0 and H_A)

$H_0 = \dots$

$H_A = \dots$

- Compute the missing F test statistic value in R output for the test:

$F = \dots$