- Web page coming!
- 1st hour: KKM 2, 6, 8, 10
- 2nd hour: questions

KKM #2 p.64:

32 white males over 40 years old from Angina

Table gives:

- SBP $\rightarrow$ systolic blood pressure
- QUET $\rightarrow$ body size
- AGE
- SMK $\rightarrow$ smoking history
  =1 if current or previous smoker
  =0 if nonsmoker

a. Draw scatter diagrams on graph paper for

(1) SBP (Y) vs. QUET (X)

(2) SBP (Y) vs. SMK (X)

(3) QUET (Y) vs. AGE (X)

(4) SBP (Y) vs. AGE (X)

b. For 1, 3, 4, sketch line by eye. Comment on relationships

c. (1) Determine l.s. estimates of slope (β_1) and intercept (β_0)

for straight line regression of SBP (Y) on SMK (X)

From printout, $\widehat{\beta}_1 = 7.024$, $\widehat{\beta}_0 = 140.800$

(2) Compare $\widehat{\beta}_0$ and mean SBP for nonsmokers.

$\widehat{\beta}_0 = 140.800$ and mean SBP for nonsmokers $= 140.800$

Compare $\widehat{\beta}_0 + \widehat{\beta}_1$ with the mean SBP for smokers.

$\widehat{\beta}_0 + \widehat{\beta}_1 = 147.824$ and mean SBP for smokers=147.824

(3) Test hypothesis $\beta_1 = 0$.

From page 24 of class notes, to test
H_0: $\beta_1 = 0$ vs. H_1: $\beta_1 \neq 0$

Use test statistic

$$t = \frac{\widehat{\beta}_1}{\text{s.e.}(\widehat{\beta}_1)} = \frac{7.024}{5.0236} = 1.398$$

And

$$\begin{aligned} p &= \Pr\{|t_{30}| > 1.398\} \\ &= 2 \times \Pr\{t_{30} > 1.398\} \\ &= 0.1723 \end{aligned}$$

Since $p = .1723 > .05$, we don't reject H_0 at the $\alpha = .05$ level.

(4) Is test in (3) equivalent to two-sample t test?

Yes!

2-sample t-test uses statistic:

d. (1) Determine l.s. estimates of slope, intercept for regression of SBP (Y) on QUET (X)

From printout, $\widehat{\beta}_1 = 21.492$, $\widehat{\beta}_0 = 70.576$.

(2) Sketch estimated regression line on scatter diagram (#2a1)

(3) Test the hypothesis of zero slope (i.e. $\beta_1 = 0$)

Use statistic

$$t = \frac{\widehat{\beta}_1}{\text{s.e.}(\widehat{\beta}_1)} = \frac{21.492}{3.545} = 6.063$$

So that

So we reject H_0: $\beta_1 = 0$ at the $\alpha = 0.05$ level.

(4) Find a 95% CI for $\mu_{Y|\overline{X}}$

From p60 of KKM,

$$\begin{aligned}\widehat{Y}_{\overline{X}} &= \widehat{\beta}_0 + \widehat{\beta}_1 \overline{X} \\ &= 70.576 + 21.492(3.441) \\ &= 144.552\end{aligned}$$

$$S_{\widehat{Y}_{\overline{X}}} = S_{Y|X}\sqrt{\frac{1}{n}} = (9.812)\left(\sqrt{\frac{1}{32}}\right) = 1.735$$

So a 95% CI for $\mu_{Y|\overline{X}}$ is

$$\begin{aligned}&\widehat{Y}_{\overline{X}} \pm t_{32-2,1-.95/2} S_{\widehat{Y}_{\overline{X}}} \\ &= 144.552 \pm (2.042)(1.735) \\ &= 144.552 \pm 3.543 \\ &= (141.009, 148.095)\end{aligned}$$

(5) Calculate 95% prediction bands

Use the following formula (5.17, p62, KKM):

$$144.531 + 21.492(X_0 - 3.441) \pm 20.035\sqrt{1.031 + \frac{(X_0 - 3.441)^2}{7.660}}$$

(6) Blood pressure increases as body size increases?

Yes, based on visual inspection (plot) and rejection of hypothesis in (3)

(7) Any regression assumptions clearly violated? No.

e., f. left to student

KKM #6 p.69:

13 children & adolescents in study to measure relationship between age and ATST

Table gives ATST and AGE for each subject.

(a) Draw scatter diagram for ATST (Y) on AGE (X)

(b) Calculate l.s. estimates of slope and intercept

$$\widehat{\beta}_0 = 646.483, \widehat{\beta}_1 = 13.152$$

(c) Draw the estimated regression line on the scattergram.

(d) Are any regression assumptions clearly violated?

(e) Obtain a 95% CI for β_1.

$$\begin{aligned}
&\widehat{\beta}_1 \pm t_{11,.975} \times \text{s.e.}(\widehat{\beta}_1) \\
&= 13.152 \pm (2.201)(1.368) \\
&= 13.152 \pm 3.011 \\
&= (10.141, 16.163)
\end{aligned}$$

(f) Reject H_0? Yes, since above CI doesn't included 0!

(g) Calculate and graph 95% conf. bands based on estimated line: see answer in back of book, found using equation (5.16) p60, KKM.

KKM #8 p.71:

(a) Draw scattergram.

(b) Calculate l.s. estimates of slope and intercept

(c) Draw regression line

(d) Obtain a 95% CI for β_1.

A 95% CI is

$$\widehat{\beta}_1 \pm t_{30-2,1-.05/2}\text{s.e.}(\widehat{\beta}_1)$$

$$= 3630.561 \pm (2.048)(465.769)$$

$$= 3630.561 \pm 953.895$$

$$= (2676.666, 4584.456)$$

(e) Reject H_0: $\beta_1 = 4000$ at α=0.05?

No, because 4000 is in the 95% CI above.

(f) <u>95% confidence bands</u>:

Use the formula:

$$\overline{Y} + \widehat{\beta}_1(X_0 - \overline{X}) \pm t_{30-2,1-.05/2} S_{Y|X} \sqrt{\frac{1}{30} + \frac{(X_0-\overline{X})^2}{(30-1)S_X^2}}$$

<u>95% prediction bands</u>:

Use (5.17) p62 KKM

(g) Graph the confidence and prediction bands above

(h) Reject H_0: $\mu_{Y|X_0} = 11,500$ at $X_0 = 2.75$?

$\widehat{Y}_{2.75} \pm t_{30-2,1-.05/2} S_{\widehat{Y}_{2.75}}$

Use (5.16) and (5.15)