Stat 615/415 (Regression) HW #2 – Hints

Regression Basics (chap. 1)

Main formulas obtained on Thursday:

Sample regression slope	$b_1 = \dfrac{S_{XY}}{S_{XY}} = \dfrac{s_{xy}}{s_x^2}$
Sample regression intercept	$b_0 = \overline{Y} - b_1 \cdot \overline{X}$
where	$S_{XX} = \sum_{i=1}^{n} (X_i - \overline{X})^2$
	$S_{XY} = \sum_{i=1}^{n} (X_i - \overline{X})(Y_i - \overline{Y})$
	$s_x^2 = \dfrac{S_{XX}}{n-1}$ is the sample variance of X
	$s_{xy} = \dfrac{S_{XY}}{n-1}$ is the sample covariance of X and Y
Predicted values	$\widehat{Y}_i = b_0 + b_1 \cdot X_i$
Residuals	$e_i = Y_i - \widehat{Y}_i$
Error sum of squares	$\sum_{i=1}^{n} e_i^2$
Sample variance	$s^2 = \dfrac{\sum e_i^2}{n-2}$
Sample standard deviation	$s = \sqrt{s^2}$

All this material is in Chap. 1 of our textbook.

##1-3. These exercises require the very basic understanding of linear regression, its meaning, and assumptions.

4-5. Instead of a sample covariance s_{xy}, these problems give us a correlation coefficient

$$r = \frac{s_{xy}}{s_x s_y}.$$

We can certainly use it to obtain the sample covariance $s_{xy} = r s_x s_y$ and then compute the slope and intercept from formulas above. But also, there is a shortcut, once we compare equations for the correlation coefficient r and the slope b_1,

$$r = \frac{s_{xy}}{s_x s_y} \qquad \Longrightarrow \qquad b_1 = \frac{s_{xy}}{s_x^2} = r\,\frac{s_y}{s_x}.$$

This is how the slope b_1 depends on the correlation coefficient r.

6. To show unbiasedness, as always, we need to find $\mathbf{E}(b_0)$ and demonstrate that it equals β_0.

7. All the needed R commands are in R lab #3 on Blackboard. SAS commands are also on Blackboard, in the folder "SAS help". We did a very similar example with Copiers data in class.