Practice Problems 6 - Solutions

1. upper left: -0.783
upper right: 0.847
lower left: 0.0556
lower right: -0.997

2. a. $\bar{x} = \dfrac{4+5+6+8+10}{5} = \dfrac{33}{5} = 6.6$

$\bar{y} = \dfrac{-3 + -2 + 1 + 3 + 2}{5} = \dfrac{1}{5} = 0.2$

$$var(X) = \frac{(4-6.6)^2 + (5-6.6)^2 + (6-6.6)^2 + (8-6.6)^2 + (10-6.6)^2}{5-1}$$

$$= \frac{(-2.6)^2 + (-1.6)^2 + (-.6)^2 + (1.4)^2 + (3.4)^2}{4}$$

$$= \frac{6.76 + 2.56 + .36 + 1.96 + 11.56}{4}$$

$$= \frac{23.20}{4} = 5.8$$

$$Var(y) = \frac{(-3-.2)^2 + (-2-.2)^2 + (1-.2)^2 + (3-.2)^2 + (2-.2)^2}{5-1}$$

$$= \frac{(-3.2)^2 + (-2.2)^2 + (.8)^2 + (2.8)^2 + (1.8)^2}{4}$$

$$= \frac{10.24 + 4.84 + .64 + 7.64 + 3.24}{4}$$

$$= \frac{26.70}{4} = 6.675$$

$\Rightarrow SD(X) = \sqrt{5.8} \approx 2.408$

$SD(Y) = \sqrt{6.675} \approx 2.584$

b. $r_{xy} = \frac{1}{n-1} \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{s_x s_y}$

$\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y}) = (-2.6)(-3.2) + (-1.6)(-2.2)$
$+ (-.6)(.8) + (1.4)(2.8) + (3.4)(1.8)$
$= 21.4$

$\Rightarrow r_{xy} = \frac{1}{4}\frac{(21.4)}{(2.408)(2.584)} = 0.8598$

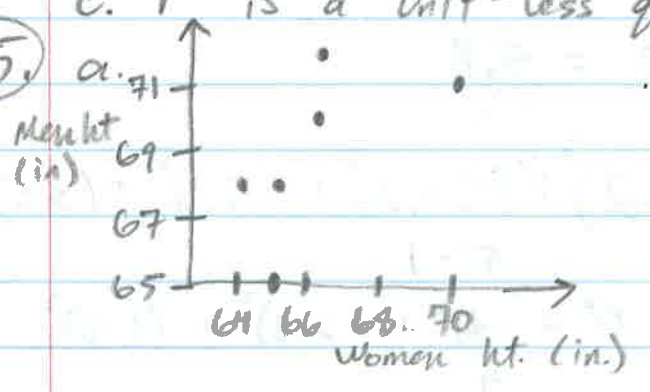
3. Shoe size and score are probably highly correlated through the variable age.

4. a. The correlation coefficient as we have defined it in class cannot be applied to the variables x and y if one of them is not numeric. Gender in this case is a categorical variable.

b. r must be between -1 and 1.

c. r is a unit-less quantity.

5. a.



* let x = women
y = men

$S_x = \sqrt{\frac{1}{5}\sum_{i=1}^{5}(x_i-66)^2}$

$S_x = \sqrt{\frac{1}{5}\sum_{i=1}^{5}(y_i-69)^2}$

$\Rightarrow S_x = 2.098$
$S_y = 2.530$

* r looks to be positive, probably not very close to 1.

b. $\bar{x}$ = avg. women ht = 66, $\bar{y}$ = avg male ht = 69

$r_{xy} = \left(\frac{1}{6-1}\right) \times \frac{\sum_{i=1}^{6}(x_i-66)(y_i-69)}{S_x S_y} = 0.565$

c. r would ~~not~~ change since the deviations between each male height and the average male height would remain the same. Hence, the correlation will not tell you if women date men taller than themselves.

d. r would equal 1 since there is a precise linear relationship, in this case male ht. = female ht. + 3 (in.)

e. r = 0.565, the same as b.

6. a. amt. of sugar

b. $r = b_1 \times \frac{S_x}{S_y} \Rightarrow b_1 = r \times \frac{S_y}{S_x} = (0.372)\left(\frac{.26}{16}\right)$

$\Rightarrow b_1 = 0.006045$

$b_0 = \bar{y} - b_1\bar{x} \Rightarrow b_0 = 1.93 - (.006045)(35.4)$

$= 1.716$

c. $\hat{y} = 1.716 + .006045x$

$\Rightarrow \hat{y} = 1.716 + .006045(7) = \$1.76$

7. a. y = time until child distracted

b. $r = \frac{1}{n-1}\left(\frac{\sum x_i y_i - n\bar{x}\bar{y}}{S_x S_y}\right)$

$= \frac{1}{242}\left(\frac{4762.8 - (243)(4.3)(6.92)}{(2.44)(3.77)}\right)$

$= -0.9525$

c. This means there is a strong, negative relationship between a child's TV watching and their time until distraction.

d. $b_1 = r \times \frac{s_x}{s_y} = (-.9525)\left(\frac{2.84}{3.77}\right) = -.7175$

$b_0 = \bar{y} - b_1\bar{x} = 6.92 - (-.7175)(4.3)$
$= 10.01$

$\Rightarrow \hat{y} = 10.01 - .7175x$

e. $\hat{y} = 10.01 - .7175(4.5)$
$= 6.78$ minutes