2.1 Consider the following five observations. You are to do all the parts of this exercise using only a calculator.

x	y	$x-\bar{x}$	$(x-\bar{x})^2$	$y-\bar{y}$	$(x-\bar{x})(y-\bar{y})$
3	4	2	4	2	4
2	2	1	1	0	0
1	3	0	0	1	0
−1	1	−2	4	−1	2
0	0	−1	1	−2	2
$\sum x_i =$ 5	$\sum y_i =$ 10	$\sum(x_i-\bar{x}) =$ 0	$\sum(x_i-\bar{x})^2 =$ 10	$\sum(y_i-\bar{y}) =$ 0	$\sum(x_i-\bar{x})(y_i-\bar{y}) =$ 8

a. Complete the entries in the table. Put the sums in the last row. What are the sample means $\bar{x}$ and $\bar{y}$?

b. Calculate b_1 and b_2 using (2.7) and (2.8) and state their interpretation.

c. Compute $\sum_{i=1}^{5} x_i^2$, $\sum_{i=1}^{5} x_i y_i$. Using these numerical values, show that $\sum(x_i-\bar{x})^2 = \sum x_i^2 - N\bar{x}^2$ and $\sum(x_i-\bar{x})(y_i-\bar{y}) = \sum x_i y_i - N\bar{x}\bar{y}$.

d. Use the least squares estimates from part (b) to compute the fitted values of y, and complete the remainder of the table below. Put the sums in the last row.

Calculate the sample variance of y, $s_y^2 = \sum_{i=1}^{N}(y_i-\bar{y})^2/(N-1)$, the sample variance of x, $s_x^2 = \sum_{i=1}^{N}(x_i-\bar{x})^2/(N-1)$, the sample covariance between x and y, $s_{xy} = \sum_{i=1}^{N}(y_i-\bar{y})(x_i-\bar{x})/(N-1)$, the sample correlation between x and y, $r_{xy} = s_{xy}/(s_x s_y)$ and the coefficient of variation of x, $CV_x = 100(s_x/\bar{x})$. What is the median, 50th percentile, of x?

2.1.a

$$\sum X_i = 3+2+1-1+0 = 5 \quad , \quad \bar{X} = 1$$

$$\sum y_i = 4+2+3+1+0 = 10 \quad , \quad \bar{y} = 2$$

2.1.b

$$b_2 = \frac{\sum(X_i-\bar{X})(y_i-\bar{y})}{\sum(X_i-\bar{X})^2} = \frac{4+0+0+2+2}{10} = 0.8$$

$b_2 = \dfrac{\partial E(y|x)}{\partial X} = 1$, One unit increase in X will bring 0.8 (b_2) unit increase of y on average.

$$b_1 = \bar{y} - b_2\bar{X} = 2 - 0.8\times 1 = 1.2$$

$E(y|x) = b_1 + b_2 X$ if $X=0$, y on average $= 1.2$ (b_1) unit

2.1.c

$$\sum X_i^2 = 3^2+2^2+1^2+(-1)^2+0^2 = 15$$

$$\sum X_i y_i = 3\times4+2\times2+1\times3+(-1)\times1+0\times0 = 18$$

$\Sigma(X_i-\bar{X}) = 10$

$\Rightarrow \Sigma(X_i-\bar{X}) = \Sigma X_i^2 - N\bar{X}^2$

$\Sigma X_i^2 - N\bar{X}^2 = 15 - 5 \times 1^2 = 10$

$\Sigma(X_i-\bar{X})(y_i-\bar{y}) = 8$

$\Rightarrow \Sigma(X_i-\bar{X})(y_i-\bar{y}) = \Sigma X_i y_i - N\bar{X}\bar{y}$

$\Sigma X_i y_i - N\bar{X}\bar{y} = 18 - 5 \times 1 \times 2 = 8$

x_i	y_i	$\hat{y}_i$	$\hat{e}_i$	$\hat{e}_i^2$	$x_i\hat{e}_i$
3	4	3.6	0.4	0.16	1.2
2	2	2.8	−0.8	0.64	−1.6
1	3	2	1	1	1
−1	1	0.4	0.6	0.36	−0.6
0	0	1.2	−1.2	1.44	0
$\sum x_i =$	$\sum y_i =$	$\sum \hat{y}_i =$ 10	$\sum \hat{e}_i =$ 0	$\sum \hat{e}_i^2 =$ 3.6	$\sum x_i\hat{e}_i =$ 0

$\hat{y}_i = 1.2 + 0.8X$

2.1.d

$$s_y^2 = \frac{\Sigma(y_i-\bar{y})^2}{N-1} = \frac{4+0+1+1+4}{4} = 2.5$$

$$s_x^2 = \frac{\Sigma(X_i-\bar{X})^2}{N-1} = \frac{10}{4} = 2.5$$

$$s_{xy} = \frac{\Sigma(X_i-\bar{X})(y_i-\bar{y})}{N-1} = \frac{8}{4} = 2$$

$$r_{xy} = \frac{S_{xy}}{S_x S_y} = \frac{2}{\sqrt{2.5}\sqrt{2.5}} = 0.8$$

$$CV_x = \left(\frac{S_x}{\bar{X}}\right)100 = \frac{\sqrt{2.5}}{1} \times 100\% = 158.1139\%$$

$Me(X) = 1$

e. On graph paper, plot the data points and sketch the fitted regression line $\hat{y}_i = b_1 + b_2 x_i$.

f. On the sketch in part (e), locate the point of the means $(\bar{x}, \bar{y})$. Does your fitted line pass through that point? If not, go back to the drawing board, literally.

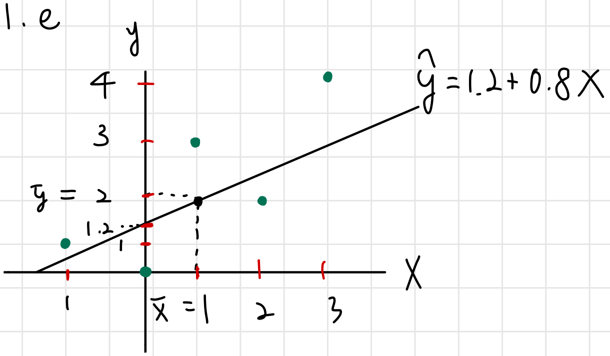
g. Show that for these numerical values $\bar{y} = b_1 + b_2\bar{x}$.

h. Show that for these numerical values $\bar{\hat{y}} = \bar{y}$, where $\bar{\hat{y}} = \sum \hat{y}_i / N$.

i. Compute $\hat{\sigma}^2$.

j. Compute $\widehat{\text{var}}(b_2|\mathbf{x})$ and $\text{se}(b_2)$.

2.1.e



2.1.f

$(\bar{x}, \bar{y})$ does pass through the fitted line of $\hat{y} = 1.2 + 0.8X$

2.1.g

$$b_1 + b_2\bar{X} = 1.2 + 0.8 \times 1 = 2 = \bar{y}$$

2.1.h

$$\bar{\hat{y}} = \frac{\Sigma \hat{y}}{N} = 2 = \bar{y}$$

2.1.i

$$\hat{\sigma}^2 = \frac{\Sigma \hat{e}^2}{N-2} = \frac{3.6}{3} = 1.2 \#$$

2.1.j

$$\text{Var}(\hat{b}_2 | x) = \frac{\hat{\sigma}^2}{\Sigma (X_i - \bar{X})^2} = \frac{1.2}{10} = 0.12$$

$$SE(b_2) = \sqrt{\text{Var}(\hat{b}_2 | x)} = \sqrt{0.12} = 0.3464$$

2.14 Consider the regression model $WAGE = \beta_1 + \beta_2 EDUC + e$, where $WAGE$ is hourly wage rate in U.S. 2013 dollars and $EDUC$ is years of education, or schooling. The regression model is estimated twice using the least squares estimator, once using individuals from an urban area, and again for individuals in a rural area.

$$\text{Urban} \quad \widehat{WAGE} = \underset{(2.27)}{-10.76} + \underset{(0.16)}{2.46}\, EDUC, \quad N = 986 \quad \text{(se)}$$

$$\text{Rural} \quad \widehat{WAGE} = \underset{(3.29)}{-4.88} + \underset{(0.24)}{1.80}\, EDUC, \quad N = 214 \quad \text{(se)}$$

a. Using the estimated rural regression, compute the elasticity of wages with respect to education at the "point of the means." The sample mean of $WAGE$ is \$19.74.

b. The sample mean of $EDUC$ in the urban area is 13.68 years. Using the estimated urban regression, compute the standard error of the elasticity of wages with respect to education at the "point of the means." Assume that the mean values are "givens" and not random.

c. What is the predicted wage for an individual with 12 years of education in each area? With 16 years of education?

2.14.a

$$\frac{\partial \widehat{WAGE}}{\partial EDUC} = 1.8$$

$$\frac{\partial \widehat{WAGE} / \widehat{WAGE}}{\partial EDUC / EDUC} = 1.8 \times \frac{EDUC}{\widehat{WAGE}}$$

$$\overline{\widehat{WAGE}} = 19.74 = -4.88 + 1.8\, \overline{EDUC}, \quad \overline{EDUC} = 13.68$$

$$\Rightarrow \frac{\partial \widehat{WAGE} / \widehat{WAGE}}{\partial EDUC / EDUC} = 1.8 \times \frac{13.68}{19.74} = 1.25 \#$$

2.14.b

$$SE\left(\frac{\partial \widehat{WAGE} / \widehat{WAGE}}{\partial EDUC / EDUC} \right) = SE\left(b_{urban} \times \frac{EDUC}{\widehat{WAGE}} \right)$$

$$= \frac{EDUC}{\widehat{WAGE}}\, SE(b_{urban}) = \frac{13.68}{19.74} \times 0.16 = 0.11 \#$$

2.14.c

	$EDUC = 12$	$EDUC = 16$
$\widehat{WAGE}_{urban}$	18.76	28.6
$\widehat{WAGE}_{rural}$	16.72	23.92

2.16 a.

The CAPM formula

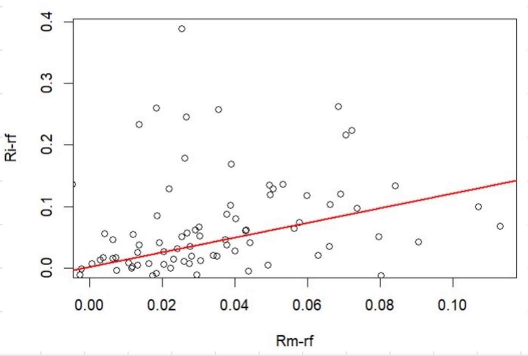
$E(r_i) = r_f + \beta(R_m - r_f)$ illustrates the linear relationship between risk premium of individual security and market portfolio. We can thus modify the structure by simple linear regression.

2.16.b

	Company	Alpha	Beta	Alpha_t	Beta_t
1	ge	-0.0026184496	1.1467864	-0.5932833	12.948606
2	ibm	0.0035908459	0.9778792	0.7446829	10.106073
3	ford	0.0045553259	$\boxed{1.6700002}$	0.4461734	8.151233
4	msft	0.0018303349	1.1927958	0.3040629	9.874648
5	dis	-0.0012334666	1.0233550	-0.2635431	10.896152
6	xom	0.0003840979	$\boxed{0.4650803}$	0.1089866	6.576301

According to the reported coefficients, Ford seem to be the most aggressive one and Exxon-mobil is the most defensive one.

2.16.c There's no significant result to support the six firms' stock has abnormal return α. The results of the data follow decently with the CAPM predicted.



2.16.d

無截距 有截距

	Company	Beta	Beta
1	ge	1.1485592	1.1467864
2	ibm	0.9754481	0.9778792
3	ford	1.6669162	1.6700002
4	msft	1.1915566	1.1927958
5	dis	1.0241901	1.0233550
6	xom	0.4648202	0.4650803

Betas seems to have a slight difference if we force $\alpha = 0$