HW0224_Pinyo 312712017

HW0224Q1

a. Complete the entries in the table. Put the sums in the last row. What are the sample means x and y?

	x	y	x – x_bar	(x-x_bar)^2	y – y_bar	(x-x_bar)(y-y_bar)
	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
∑	5	10	0	10	0	8

x_bar	1
y_bar	2

b. Calculate b1 and b2 using (2.7) and (2.8) and state their interpretation.

b.)

$$b_2 = \frac{\sum (x_i-\bar{x})(y_i-\bar{y})}{\sum (x_i-\bar{x})^2} = \frac{8}{10} = 0.8$$

$$b_1 = \bar{y} - b_2\bar{x} = 2 - 0.8(1) = 1.2$$

From b_1, b_2 result, fitted regression line is $\hat{y}_i = 0.8x_i + 1.2$

Intercept of this regression line is 1.2, meaning that at $x = 0$ estimated $y = 1.2$. Slope of regression equation = 0.8, meaning that change of x 1 unit affect estimated $y = 0.8$

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 $\sum(x_i - \bar{x})(y_i - \bar{y}) = \sum x_i y_i - N\bar{x}\bar{y}$.

	x	y	x – x_bar	(x-x_bar)^2	y – y_bar	(x-x_bar)(y-y_bar)	x^2	x * y
	3	4	2	4	2	4	9	12
	2	2	1	1	0	0	4	4
	1	3	0	0	1	0	1	3
	-1	1	-2	4	-1	2	1	-1
	0	0	-1	1	-2	2	0	0
∑	5	10	0	10	0	8	15	18

c.)

$\Sigma(x_i-\bar{x})^2 = \Sigma x_i^2 - N\bar{x}^2$ and $\Sigma(x_i-\bar{x})(y_i-\bar{y}) = \Sigma x_i y_i - N\bar{x}\bar{y}$

$10 = 15 - 5(1)$ | $8 = 18 - 5(1)(2)$

$10 = 10$. | $8 = 8$

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?

	x	y	ŷI ('=0.8xi + 1.2)	êi	êi^2	xi * êi
	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
∑	5	10	10	0	3.6	0

d.)

$$S_y^2 = \sum_{i=1}^{N} (y_i-\bar{y})^2/(N-1) = \frac{2^2+0+1^2+(-1)^2+(-2)^2}{5-1} = \frac{10}{4} = 2.5$$

$$S_x^2 = \sum_{i=1}^{N} (x_i-\bar{x})^2/(N-1) = \frac{2^2+1^2+0+(-2)^2+(-1)^2}{4} = \frac{10}{4} = 2.5$$

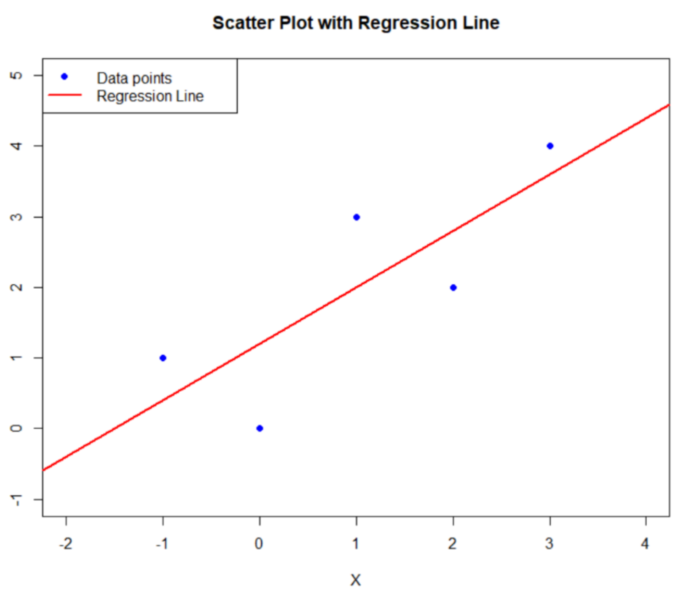
$$S_{xy} = \sum_{i=1}^{N} (y_i-\bar{y})(x_i-\bar{x})/(N-1) = \frac{8}{4} = 2$$

$$r_{xy} = S_{xy}/(S_x S_y) = \frac{2}{(\sqrt{2.5})^2} = 0.8$$

$$cv_x = 100\,(S_x/\bar{x}) = \frac{100\,(\sqrt{2.5})}{1} = 100\sqrt{2.5}$$

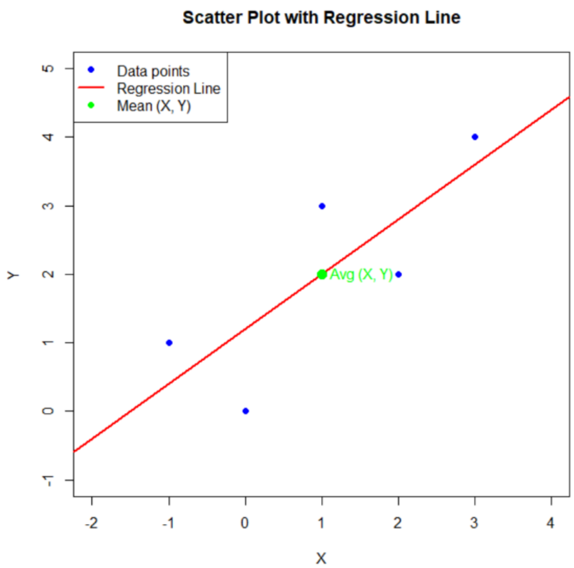
$$x_{med.} = 1$$

e. On graph paper, plot the data points and sketch the fitted regression line ŷi = b1 + b2xi



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.



Yes, the regression line pass the point (x mean, y mean)

g.

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

g.) $\bar{y} = b_1 + b_2\bar{x}$

$\bar{x} = 1, \ \bar{y} = 2, \ b_1 = 1.2, \ b_2 = 0.8$

so, $2 = 1.2 + 0.8(1)$

$2 = 2$

h.

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

h.) $\bar{\hat{y}} = \bar{y}$, $\bar{\hat{y}} = \Sigma\hat{y}_i/N$

$\bar{\hat{y}} = \Sigma\hat{y}_i/N = \frac{10}{5} = 2 \rightarrow \bar{\hat{y}} = \bar{y} = 2$

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

i.) $$\hat{\sigma}^2 = \frac{\sum \hat{e}_i^2}{N-2} = \frac{3.6}{4} = 0.9$$

j.

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

j.) $$\widehat{\text{var}}(b_2|x) = \frac{\hat{\sigma}^2}{\sum (x_i - \bar{x})^2} = \frac{0.9}{10} = 0.09$$

$$\text{se}(b_2) = \sqrt{\widehat{\text{var}}(b_2|x)} = \sqrt{0.09} = 0.3$$

HW0224Q14

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.

$$\begin{array}{lll} \text{Urban} & \widehat{WAGE} = -10.76 + 2.46\,EDUC, & N = 986 \\ & \text{(se)} \qquad (2.27) \quad (0.16) & \end{array}$$

$$\begin{array}{lll} \text{Rural} & \widehat{WAGE} = -4.88 + 1.80\,EDUC, & N = 214 \\ & \text{(se)} \qquad (3.29) \quad (0.24) & \end{array}$$

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.

a.) $\varepsilon = b_2 \dfrac{\overline{EDUC}}{\overline{WAGE}}$ —(1)

From rural estimated eq. $\widehat{WAGE} = -4.88 + 1.80\,EDUC$, $N = 214$

$\overline{EDUC} = (\overline{WAGE} + 4.88)/1.8 = (19.74 + 4.88)/1.8$

$= 13.7$

$\varepsilon = 1.8\left(\dfrac{13.7}{19.74}\right) = 1.25$

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.

b.)

$\overline{EDUC}_{urban} = 13.68$, $se(\hat{\varepsilon})_{wage} = ?$

From $\widehat{WAGE} = -10.76 + 2.46\,EDUC$

$\overline{WAGE} = -10.76 + 2.46(13.68) = 22.89$

$se(\hat{\varepsilon}) = \sqrt{Var(\hat{\varepsilon})} = \sqrt{var\left(b_2 \times \frac{\overline{EDUC}}{\overline{WAGE}}\right)}$

$= \sqrt{\frac{\overline{EDUC}}{\overline{WAGE}}} \times \sqrt{var(b_2)}$

$= \sqrt{\frac{13.68}{22.89}} \times se(b_2) = 0.77 \times 0.16$

$= 0.12$ #

c.

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

c.)

Urban $\widehat{WAGE} = -10.76 + 2.46(12) = 18.76$

$\widehat{WAGE} = -10.76 + 2.46(16) = 28.6$

Rural $\widehat{WAGE} = -4.88 + 1.8(12) = 16.72$

$\widehat{WAGE} = -4.88 + 1.8(16) = 23.92$ #

HW0224Q16

The capital asset pricing model (CAPM) is an important model in the field of finance. It explains variations in the rate of return on a security as a function of the rate of return on a portfolio consisting of all publicly traded stocks, which is called the *market* portfolio. Generally, the rate of return on any investment is measured relative to its opportunity cost, which is the return on a risk-free asset. The resulting difference is called the *risk premium*, since it is the reward or punishment for making a risky investment. The CAPM says that the risk premium on security *j* is *proportional* to the risk premium on the market portfolio. That is,

$$r_j - r_f = \beta_j(r_m - r_f)$$

where r_j and r_f are the returns to security *j* and the risk-free rate, respectively, r_m is the return on the market portfolio, and β_j is the *j*th security's "*beta*" value. A stock's *beta* is important to investors since it reveals the stock's volatility. It measures the sensitivity of security *j*'s return to variation in the whole stock market. As such, values of *beta* less than one indicate that the stock is "defensive" since its variation is less than the market's. A *beta* greater than one indicates an "aggressive stock." Investors usually want an estimate of a stock's *beta* before purchasing it. The CAPM model shown above is the "economic model" in this case. The "econometric model" is obtained by including an intercept in the model (even though theory says it should be zero) and an error term

$$r_j - r_f = \alpha_j + \beta_j(r_m - r_f) + e_j$$

a.

From the CAPM model, it is a simple linear regression model because:

1.) rj – rf can be considered as y which is only one dependent variable
2.) rm – rf can be considered as x which is only one independent variable
3.) x and y in this model has a linear relationship
4.) The model consist of error term which is a random value (not be controlled by market factors)

b.

```
CAPM Regression for ge :

Call:
lm(formula = formula, data = capm_data)

Residuals:
      Min        1Q    Median        3Q       Max 
-0.174157 -0.032861 -0.005409  0.040441  0.204256 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept) -0.003692   0.004435  -0.832    0.406    
mkt          1.153550   0.089665  12.865   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.05915 on 178 degrees of freedom
Multiple R-squared:  0.4818,	Adjusted R-squared:  0.4789 
F-statistic: 165.5 on 1 and 178 DF,  p-value: < 2.2e-16
```

```
CAPM Regression for ibm :

Call:
lm(formula = formula, data = capm_data)

Residuals:
      Min        1Q    Median        3Q       Max 
-0.257101 -0.035182 -0.006376  0.033002  0.275049 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 0.003725   0.004848   0.768    0.443    
mkt         0.982032   0.098019  10.019   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.06466 on 178 degrees of freedom
Multiple R-squared:  0.3606,	Adjusted R-squared:  0.357 
F-statistic: 100.4 on 1 and 178 DF,  p-value: < 2.2e-16
```

```
CAPM Regression for ford :

Call:
lm(formula = formula, data = capm_data)

Residuals:
     Min       1Q   Median       3Q      Max 
-0.27611 -0.07757 -0.01051  0.04632  1.09147 

Coefficients:
              Estimate Std. Error t value Pr(>|t|)    
(Intercept) -0.0001288  0.0102808  -0.013     0.99    
mkt          1.6609865  0.2078489   7.991 1.62e-13 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.1371 on 178 degrees of freedom
Multiple R-squared:  0.264,	Adjusted R-squared:  0.2599 
F-statistic: 63.86 on 1 and 178 DF,  p-value: 1.617e-13
```

```
CAPM Regression for msft :

Call:
lm(formula = formula, data = capm_data)

Residuals:
     Min       1Q   Median       3Q      Max 
-0.27671 -0.04764 -0.01104  0.03710  0.35487 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 0.000369   0.006045   0.061    0.951    
mkt         1.211274   0.122208   9.912   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.08062 on 178 degrees of freedom
Multiple R-squared:  0.3556,	Adjusted R-squared:  0.352 
F-statistic: 98.24 on 1 and 178 DF,  p-value: < 2.2e-16
```

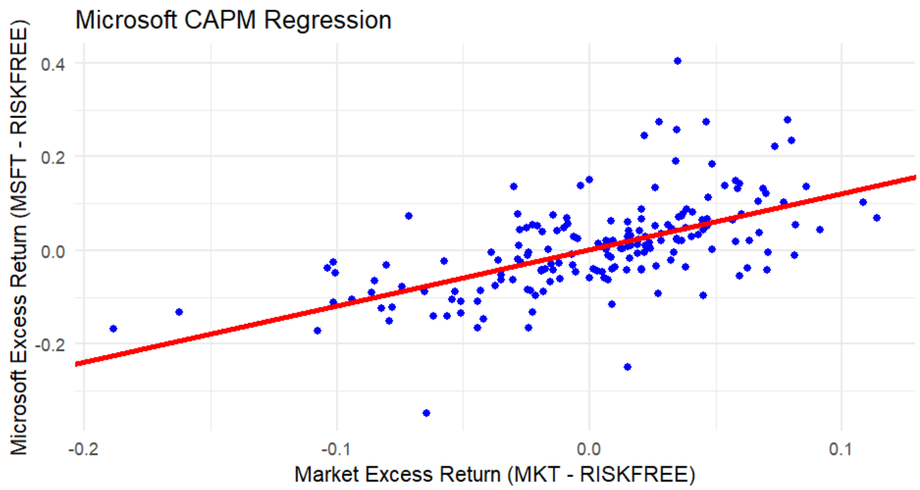
```
CAPM Regression for dis :

Call:
lm(formula = formula, data = capm_data)

Residuals:
      Min        1Q    Median        3Q       Max 
-0.176350 -0.029738 -0.004262  0.028411  0.278193 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept) -0.001332   0.004706  -0.283    0.778    
mkt          1.010968   0.095147  10.625   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.06277 on 178 degrees of freedom
Multiple R-squared:  0.3881,	Adjusted R-squared:  0.3847 
F-statistic: 112.9 on 1 and 178 DF,  p-value: < 2.2e-16
```

```
CAPM Regression for xom :

Call:
lm(formula = formula, data = capm_data)

Residuals:
      Min        1Q    Median        3Q       Max 
-0.115635 -0.030757 -0.001142  0.026255  0.215456 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 0.004191   0.003547   1.182    0.239    
mkt         0.459759   0.071704   6.412 1.25e-09 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.0473 on 178 degrees of freedom
Multiple R-squared:  0.1876,	Adjusted R-squared:  0.1831 
F-statistic: 41.11 on 1 and 178 DF,  p-value: 1.252e-09
```

```
Most Aggressive Stock: ford (Highest Beta = 1.660987 )
> cat("Most Defensive Stock:", most_defensive, "(Lowest Beta =", min(betas), ")\n")
Most Defensive Stock: xom (Lowest Beta = 0.4597585 )
```

c.

From the regression of CAPM model, alpha or intercept values from each stock are close to zero, this might be ignored from the model



d.

```
--------------------- GE ---------------------
Beta (With Intercept): 1.147952
Beta (No Intercept)  : 1.146763

--------------------- IBM ---------------------
Beta (With Intercept): 0.9768898
Beta (No Intercept)  : 0.9843954

--------------------- Ford ---------------------
Beta (With Intercept): 1.662031
Beta (No Intercept)  : 1.666717

--------------------- MSFT ---------------------
Beta (With Intercept): 1.20184
Beta (No Intercept)  : 1.205869

--------------------- Disney ---------------------
Beta (With Intercept): 1.011521
Beta (No Intercept)  : 1.012819

--------------------- ExxonMobil ---------------------
Beta (With Intercept):
Beta (No Intercept)  :
```

Estimated beta values of each stock by CAPM model from with and without interception (alpha) are not significantly different