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

$\bar{x}=1$ $\bar{y}=2$

x	y	$x-\bar{x}$	$(x-\bar{x})^2$	$y-\bar{y}$	$(x-\bar{x})(y-\bar{y})$
3	4	$3-1=2$	4	$4-2=2$	4
2	2	$2-1=1$	1	$2-2=0$	0
1	3	$1-1=0$	0	$3-2=1$	0
-1	1	$-1-1=-2$	4	$1-2=-1$	2
0	0	$0-1=-1$	1	$0-2=-2$	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?

$y = 1.2 + 0.8x$ $\hat{e}_i = y_i - \hat{y}_i$

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
$\sum x_i = 5$	$\sum y_i = 10$	$\sum \hat{y}_i = 10$	$\sum \hat{e}_i = 0$	$\sum \hat{e}_i^2 = 3.6$	$\sum x_i\hat{e}_i = 0$

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 $\text{se}(b_2)$.

(a)

$$\bar{x} = \frac{\sum x_i}{5} = 1 \qquad \bar{y} = \frac{\sum y_i}{5} = 2$$

(b)

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

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

$$\Rightarrow y = 1.2 + 0.8x$$

b1 b2

當$x=0$，$y=1.2$；當x增加1單位，y增加0.8單位

(C).

$$\sum_{i=1}^{5} x_i^2 = 3^2+2^2+1^2+(-1)^2+0 = 15$$

$$\sum_{i=1}^{5} x_i y_i = 3\times 4+2\times 2+1\times 3+(-1)\times 1+0 = 18$$

① $$\sum_{i=1}^{N}(x_i-\bar{x})^2 = \sum_{i=1}^{N}(x_i^2-2x_i\bar{x}+\bar{x}^2) = \sum_{i=1}^{N}x_i^2-2\bar{x}\sum_{i=1}^{N}\bar{x}+\sum_{i=1}^{N}\bar{x}^2$$

$$= \sum_{i=1}^{N}x_i^2-2\cdot\bar{x}\cdot N\bar{x}+N\cdot\bar{x}^2 = \sum_{i=1}^{N}x_i^2-N\bar{x}^2$$

得 $\sum_{i=1}^{N}(x_i-\bar{x})^2 = 10 = \sum_{i=1}^{N}x_i^2-N\bar{x}^2 = 15-5\cdot1^2 = 10$

② $$\sum_{i=1}^{N}(x_i-\bar{x})(y_i-\bar{y}) = \sum_{i=1}^{N}(x_iy_i-x_i\bar{y}-\bar{x}y_i+\bar{x}\bar{y})$$

$$= \sum_{i=1}^{N}x_iy_i-\bar{y}\sum_{i=1}^{N}x_i-\bar{x}\sum_{i=1}^{N}y_i+\sum_{i=1}^{N}\bar{x}\bar{y}$$

$$= \sum_{i=1}^{N}x_iy_i-\bar{y}\cdot N\cdot\bar{x}\cancel{-\bar{x}\cdot N\cdot\bar{y}}+\cancel{N\cdot\bar{x}\bar{y}} = \sum_{i=1}^{N}x_iy_i-N\cdot\bar{x}\bar{y}$$

得 $\sum_{i=1}^{5}(x_i-\bar{x})(y_i-\bar{y}) = 8 = \sum_{i=1}^{5}x_iy_i-N\cdot\bar{x}\cdot\bar{y} = 18-5\cdot1\cdot2 = 8$

(d)

$$S_y^2 = \frac{\sum_{i=1}^{N}(y_i-\bar{y})^2}{N-1} = \frac{(4-2)^2+(2-2)^2+(3-2)^2+(1-2)^2+(0-2)^2}{5-1} = \frac{10}{4} = 2.5$$

$$S_x^2 = \frac{\sum_{i=1}^{N}(x_i-\bar{x})^2}{N-1} = \frac{(3-1)^2+(2-1)^2+(1-1)^2+(-1-1)^2+(0-1)^2}{5-1} = \frac{10}{4} = 2.5$$

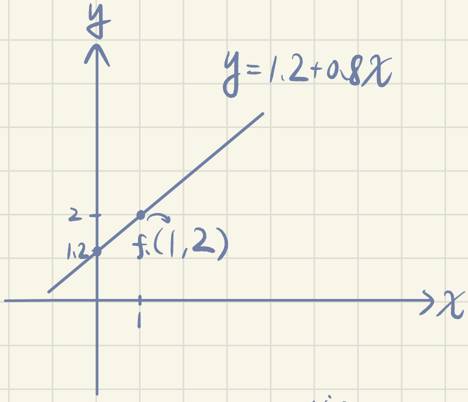
$$S_{xy} = \frac{\sum_{i=1}^{N}(y-\bar{y})(x_i-\bar{x})}{N-1} = \frac{8}{5-1} = 2$$

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

$$CV = 100 \cdot \frac{S_x}{\bar{x}} = 100 \times \frac{\sqrt{2.5}}{1} = 100\sqrt{2.5} \doteqdot 158.1139$$

median of $x = 1$ (排序 x: $-1, 0, 1, 2, 3$)

(e)



(f). $\bar{x}=1$, $\bar{y}=2$ $\overset{\text{代入}}{\Rightarrow}$ $y=2=1.2+0.8\times 1$ #

(g). $\bar{y} = b_1 + b_2 x = 2 = 1.2+0.8\times 1$ #

(h). $\bar{\hat{y}} = \dfrac{\sum \hat{y}_i}{N} = \dfrac{10}{5} = 2 = \bar{y}$ #

(i). $\hat{\sigma}^2 = \dfrac{\sum_{i=1}^{5} \hat{e}_i^2}{N-2} = \dfrac{\sum_{i=1}^{5} (y_i - \hat{y}_i)^2}{5-2}$

$= \dfrac{(0.4)^2 + (-0.8)^2 + (1)^2 + (0.6)^2 + (-1.2)^2}{5-2} = \dfrac{3.6}{3} = 1.2$ #

(j). $\widehat{var}(b_2|x) = \dfrac{\hat{\sigma}^2}{\sum_{i=1}^{5} (x_i - \bar{x})^2} = \dfrac{1.2}{10} = 0.12$ #

$se(b_2) = \sqrt{var(b_2|x)} = \sqrt{0.12} \approx 0.3464$ #

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.

$$\text{Urban} \quad \widehat{WAGE} = \underset{(2.27)}{\underset{b_1}{-10.76}} + \underset{(0.16)}{\underset{b_2}{2.46}}\, EDUC, \quad N = 986 \quad \text{(se)}$$

$$\text{Rural} \quad \widehat{WAGE} = \underset{(3.29)}{\underset{b_1}{-4.88}} + \underset{(0.24)}{\underset{b_2}{1.80}}\, EDUC, \quad N = 214 \quad \text{(se)}$$

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?

(a) $\boxed{\text{Rural}}$

$$\widehat{WAGE} = -4.88 + 1.80 EDUC = 19.74$$

$$\Rightarrow EDUC = \frac{19.74 + 4.88}{1.8} = 13.6778$$

$$\text{Elasticity of wage} = \frac{\Delta WAGE}{\Delta EDUC} \times \frac{\overline{EDUC}}{\overline{WAGE}} = 1.8 \times \frac{13.6778}{19.74}$$

$$\approx 1.2472 \#$$

(b). $\boxed{\text{Urban}}$

$$\widehat{WAGE} = -10.76 + 2.46 \times 13.68 = 22.8928$$

$$\text{the standard error of the elasticity of wages} = \beta_2 \times \frac{\overline{EDUC}}{\overline{WAGE}}$$

$$= 0.16 \times \frac{13.68}{22.8928} = 0.0956 \#$$

(C).

$\boxed{\text{Urban}}$

$$EDUC = 12, \ \widehat{WAGE} = -10.76 + 2.46 \times 12 = 18.76$$

$$EDUC = 16, \ \widehat{WAGE} = -10.76 + 2.46 \times 16 = 28.60$$

$\boxed{\text{Rural}}$

$$EDUC = 12, \ \widehat{WAGE} = -4.88 + 1.8 \times 12 = 16.72$$

$$EDUC = 16, \ \widehat{WAGE} = -4.88 + 1.8 \times 16 = 23.92$$

POEHW0224

Ch2.16

(a)

$rj - rf = \alpha j + \beta j(rm - rf) + ej$ 為簡單線性回歸模型

依變數$rj - rf$，自變數$rm - rf$。

根據 CAPM 理論，$\alpha j = 0$；ej是為誤差項。

該模型滿足$y = \beta 0 + \beta 1x + e$的形式，故為一個簡單回歸模型。

(b)

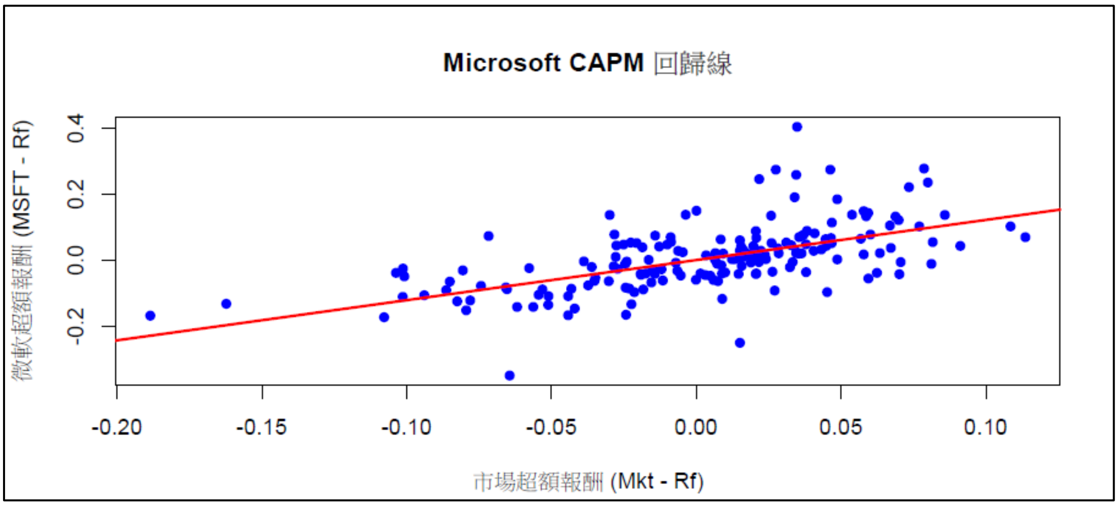
Ford β 值為 1.66＞1，是為 aggressive stock；

Exxon-Mobil 的 β 值為 0.46＜1，是為 defensive stock。

```
              Company         Alpha      Beta P_Value_Alpha P_Value_Beta
(Intercept)        ge -0.0036916942 1.1535496     0.4063037 3.316175e-27
(Intercept)1      ibm  0.0037247457 0.9820323     0.4433496 5.111456e-19
(Intercept)2     ford -0.0001287859 1.6609865     0.9900193 1.617334e-13
(Intercept)3     msft  0.0003690433 1.2112741     0.9513861 1.021569e-18
(Intercept)4      dis -0.0013317654 1.0109677     0.7775221 9.830474e-21
(Intercept)5      xom  0.0041912103 0.4597585     0.2388874 1.251880e-09
             T_Value_Alpha T_Value_Beta R_Squared
(Intercept)    -0.83239086    12.865157 0.4818235
(Intercept)1    0.76826190    10.018799 0.3605778
(Intercept)2   -0.01252689     7.991316 0.2640405
(Intercept)3    0.06105219     9.911595 0.3556321
(Intercept)4   -0.28297853    10.625278 0.3880986
(Intercept)5    1.18173455     6.411911 0.1876323
```

(c)

呈上圖，α 的 p 值和 t 值皆不顯著，故不能拒絕$\alpha j = 0$的假設。



(d)

比較 CAPM 原始模型(有截距 α)和假設$\alpha j = 0$時的模型， β 值的差異。

結果如下圖，可以發現$\alpha j = 0$時 β 值的變化不大。

	Company	Beta_Original	Beta_No_Intercept	Difference
mkt	ge	1.1535496	1.1454280	-0.0081215451
mkt1	ibm	0.9820323	0.9902265	0.0081942569
mkt2	ford	1.6609865	1.6607032	-0.0002833226
mkt3	msft	1.2112741	1.2120860	0.0008118771
mkt4	dis	1.0109677	1.0080379	-0.0029298182
mkt5	xom	0.4597585	0.4689790	0.0092204559

Ch2.16 code

(a)(b)

```
# 設定工作目錄
setwd("C:/Users/bella/Desktop/NYCU/計量經濟/POEdata")

# 確認檔案是否存在
list.files()

# 載入 capm5.rdata
load("capm5.rdata")

# 檢查數據格式
head(capm5)

# 公司名稱
companies <- c("ge", "ibm", "ford", "msft", "dis", "xom")

# 建立空的 data frame 來存放結果
results <- data.frame(
  Company = character(),
  Alpha = numeric(),
  Beta = numeric(),
  P_Value_Alpha = numeric(),
  P_Value_Beta = numeric(),
  T_Value_Alpha = numeric(),
  T_Value_Beta = numeric(),
  R_Squared = numeric(),
  stringsAsFactors = FALSE
```

```
)

# 估計 CAPM 模型
for (company in companies) {
  formula <- as.formula(paste0(company, " - riskfree ~ mkt - riskfree"))
  model <- lm(formula, data = capm5)
  summary_model <- summary(model)

  # 提取回歸結果
  alpha <- coef(model)[1]   # 截距
  beta <- coef(model)[2]    # β 值

  p_value_alpha <- coef(summary_model)[1, 4]   # 截距的 p 值
  p_value_beta <- coef(summary_model)[2, 4]    # β 的 p 值

  t_value_alpha <- coef(summary_model)[1, 3]   # 截距的 t 值
  t_value_beta <- coef(summary_model)[2, 3]    # β 的 t 值

  r_squared <- summary_model$r.squared   # R²

  # 將結果存入 data frame
  results <- rbind(results, data.frame(
    Company = company,
    Alpha = alpha,
    Beta = beta,
    P_Value_Alpha = p_value_alpha,
    P_Value_Beta = p_value_beta,
    T_Value_Alpha = t_value_alpha,
    T_Value_Beta = t_value_beta,
    R_Squared = r_squared
  ))
}

# 顯示結果
print(results)
```

(c)

```
# 微軟 (MSFT) 的回歸模型
```

```
msft_model <- capm_models[["msft"]]

# 顯示回歸結果
summary(msft_model)

# 繪製微軟的回歸線
plot(capm5$mkt - capm5$riskfree,
     capm5$msft - capm5$riskfree,
     main = "Microsoft CAPM 回歸線",
     xlab = "市場超額報酬 (Mkt - Rf)",
     ylab = "微軟超額報酬 (MSFT - Rf)",
     pch = 16, col = "blue")

# 加入回歸線
abline(msft_model, col = "red", lwd = 2)

# 檢查截距值 alpha
alpha_value <- coef(msft_model)[1]
cat("微軟的截距 (alpha):", alpha_value, "\n")

# 如果 alpha 接近 0，輸出結論
if (abs(alpha_value) < 0.05) {
  cat("截距接近 0，符合 CAPM 理論。\n")
} else {
  cat("截距不為 0，可能表示市場無效率或模型假設不完全。\n")
}
```

(d)

```
# 公司名稱
companies <- c("ge", "ibm", "ford", "msft", "dis", "xom")

# 建立空的 data frame 來存放結果
results_no_intercept <- data.frame(
  Company = character(),
  Beta_Original = numeric(),
  Beta_No_Intercept = numeric(),
  Difference = numeric(),
  stringsAsFactors = FALSE
```

```r
)

for (company in companies) {
  # 原本的 CAPM 迴歸 (有截距)
  formula1 <- as.formula(paste0(company, " - riskfree ~ mkt - riskfree"))
  model1 <- lm(formula1, data = capm5)
  beta_original <- coef(model1)[2]  # 原始 β

  # 假設 α = 0 的 CAPM 迴歸 (無截距)
  formula2 <- as.formula(paste0(company, " - riskfree ~ mkt - riskfree - 1"))
  model2 <- lm(formula2, data = capm5)
  beta_no_intercept <- coef(model2)[1]  # 無截距時的 β

  # 計算 β 的變化
  difference <- beta_no_intercept - beta_original

  # 存入 data frame
  results_no_intercept <- rbind(results_no_intercept, data.frame(
    Company = company,
    Beta_Original = beta_original,
    Beta_No_Intercept = beta_no_intercept,
    Difference = difference
  ))
}

# 顯示結果
print(results_no_intercept)
```