HW0224

Yung-Jung Cheng

2025-02-26

Q1

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

Q1 (a)

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

Ans

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

$\bar{x} = \frac{5}{5} = 1, \bar{y} = \frac{10}{5} = 2$

Q1 (b)

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

Ans

$$b_1 = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sum(x-\bar{x})^2}$$

$b_1 = \frac{8}{10} = 0.8$

$$b_0 = \bar{y} - b_1\bar{x}$$

$b_0 = 2 - 0.8 \times 1 = 1.2$

Q1 (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 $\sum(x_i - \bar{x})(y_i - \bar{y}) = \sum x_i y_i - N\bar{x}\bar{y}$.

Ans

$\sum_{i=1}^{5} x_i^2 = 9 + 4 + 1 + 1 + 0 = 15$, $\sum_{i=1}^{5} x_i y_i = 3 * 4 + 2 * 2 + 3 - 1 = 18$

$\sum(x_i - \bar{x})^2 = 10 = 15 - 5 * 1^2 = \sum x_i^2 - N\bar{x}^2$

$\sum(x_i - \bar{x})(y_i - \bar{y}) = 8 = 18 - 5 * 1 * 2 = \sum x_i y_i - N\bar{x}\bar{y}$

Q1 (d)

Ans

x_i	y_i	$\hat{y_i}$	e_i	e_i^2	$x_i e_i$
3	4	3.6	0.4	0.16	1.2
2	2	2.8	-0.8	0.64	-1.6
1	3	2	1.0	1	1
-1	1	0.4	0.6	0.36	-0.6
0	0	1.2	-1.2	1.44	0
$\sum x_i = 5$	$\sum y_i = 10$	$\sum \hat{y}_i = 10$	$\sum \hat{e}_i = 0$	$\sum \hat{e}_i^2 = 3.6$	$\sum x_i - \hat{e}_i = 0$

x={0,-1,1,2,3}, Thus the median is 1.

Q1 (e)

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

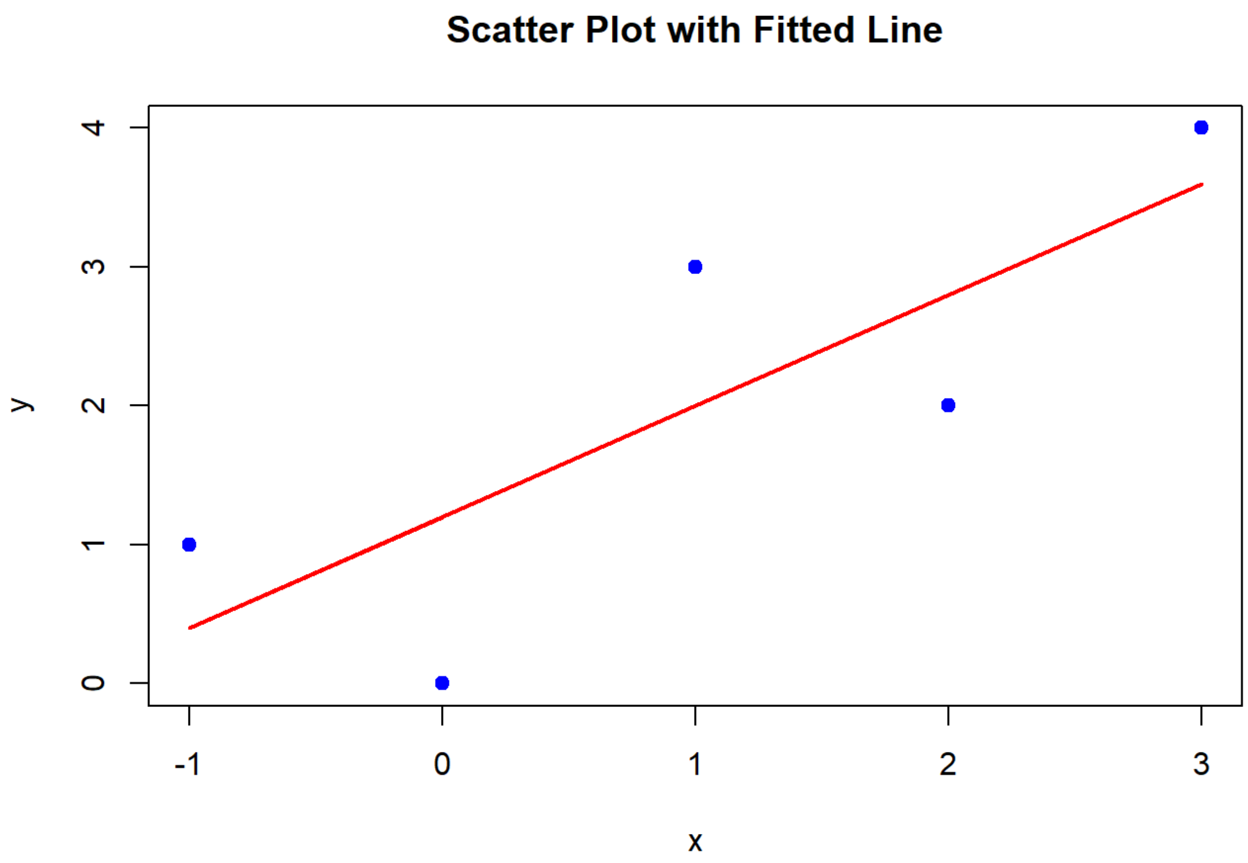
Ans

```
# 定義 x 和 y 的數據點
x <- c(3, 2, 1, -1, 0)
y <- c(4, 2, 3, 1, 0)

# 計算擬合線的 y 值
y_fit <- 1.2 + 0.8 * x

# 繪製散點圖
plot(x, y, main="Scatter Plot with Fitted Line", xlab="x", ylab="y", pch=19, col="blue", ylim=
c(min(y, y_fit), max(y, y_fit)))

# 添加擬合線
lines(x, y_fit, type="l", col="red", lwd=2)
```



Q1 (f)

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

Ans

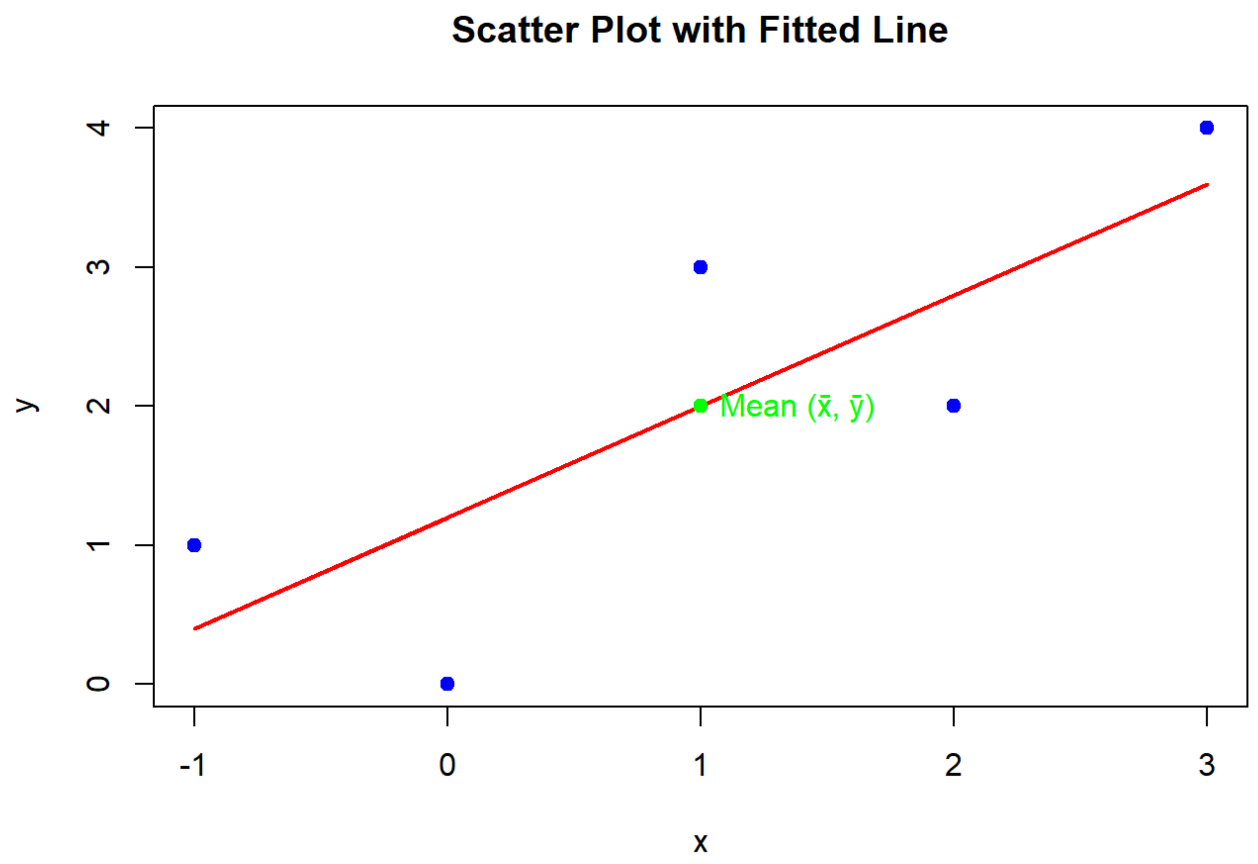
```
# 定義 x 和 y 的數據點
x <- c(3, 2, 1, -1, 0)
y <- c(4, 2, 3, 1, 0)

# 計算擬合線的 y 值
y_fit <- 1.2 + 0.8 * x

# 繪製散點圖
plot(x, y, main="Scatter Plot with Fitted Line", xlab="x", ylab="y", pch=19, col="blue", ylim=
c(min(y, y_fit), max(y, y_fit)))

# 添加擬合線
lines(x, y_fit, type="l", col="red", lwd=2)

# 標記平均值 (x̄, ȳ)
points(mean(x), mean(y), pch=19, col="green")
text(mean(x), mean(y), labels="Mean (x̄, ȳ)", pos=4, col="green")
```



Q1 g

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

Ans

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

Q1 h Show that for these numerical values $\bar{\hat{y}} = \bar{y}$, where $\bar{\hat{y}} = \sum \hat{y}_i/N$

Ans

$$\bar{\hat{y}} = \sum \hat{y}_i/N = \frac{3.6 + 2.8 + 2 + 0.4 + 1.2}{5} = \frac{10}{5} = 2 = \bar{y}$$

Q1 i

Compute $\hat{\sigma}^2$

Ans

$$\hat{\sigma}^2 = \frac{\sum e_i^2}{N-2} = \frac{3.6}{3} = 1.2$$

Q1 j

Compute $\hat{\mathrm{var}}(b_2|x)$ ans $\mathrm{se}(b_2)$

Ans

1. $\hat{\mathrm{var}}(b_2|x) = \frac{\hat{sigma}^2}{\sum(x_i-\bar{x})^2} = \frac{1.2}{10} = 0.12$
2. $\mathrm{se}(b_2) = \sqrt{\mathrm{var}(\hat{b}_1|x)} = \sqrt{0.12} \approx 0.3464$

Q14

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.

$$

$$\begin{aligned} Urban W\hat{AGE} &= -10.76 + 2.46 EDUC, N = 986 \\ (se) &= (2.27)\ (0.16) \\ Rural W\hat{AGE} &= -4.88 + 1.80 EDUC, N = 986 \\ (se) &= (3.29)\ (0.24) \end{aligned}$$

$$

Q14 (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.

Ans

$$ 19.74=-4.88+1.8 \=13.6778\ E==1.8=1.247

$$ ## Q14 (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.

Ans

$$SE(E) = SE(\beta_2) \times \frac{\overline{EDUC}}{\overline{WAGE}}$$

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

$\Rightarrow SE(E) = 0.16 \times \frac{13.68}{22.89} = 0.09562$

Q14 c

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

Ans

1. 12 years:

- Urban: $-10.76 + 2.46 * 12 = 18.76$
- Rural: $-4.88 + 1.8 * 12 = 16.72$

2. 16 years:

- Urban: $-10.76 + 2.46 * 16 = 28.6$
- Rural: $-4.88 + 1.8 * 16 = 23.92$

Q16

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 jth 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:

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

Q16 (a)

Explain why the econometric model above is a simple regression model like those discussed in this chapter.

Ans

1. Let the $r_j - r_f$ at the left side be the Dependent Variable
2. $r_m - r_f$ from the right side be tje Independent Variable
3. α_j is the Intercept term
4. β_j is Regression Coefficient
5. and e_j stand for the error term

Thus $r_j - r_f = \alpha_j + \beta_j(r_m - r_f) + e_j$ look just the same as simple linear regression $Y = \beta_0 + \beta_1 X + e$.

Q16 (b)

In the data file *capm5* are data on the monthly returns of six firms (GE, IBM, Ford, Microsoft, Disney, and Exxon-Mobil), the rate of return on the market portfolio (*MKT*), and the rate of return on the risk-free asset (*RISKFREE*). The 180 observations cover January 1998 to December 2012. Estimate the CAPM model for each firm, and comment on their estimated “beta” values. Which firm appears most aggressive? Which firm appears most defensive?

Ans

```
library(POE5Rdata)
data(capm5)
# 計算各公司超額報酬
capm5$ge_excess <- capm5$ge - capm5$riskfree
capm5$ibm_excess <- capm5$ibm - capm5$riskfree
capm5$ford_excess <- capm5$ford - capm5$riskfree
capm5$msft_excess <- capm5$msft - capm5$riskfree
capm5$dis_excess <- capm5$dis - capm5$riskfree
capm5$xom_excess <- capm5$xom - capm5$riskfree
capm5$mkt_excess <- capm5$mkt - capm5$riskfree


# 建立回歸模型（OLS）
ge_model <- lm(ge_excess ~ mkt_excess, data = capm5)
ibm_model <- lm(ibm_excess ~ mkt_excess, data = capm5)
ford_model <- lm(ford_excess ~ mkt_excess, data = capm5)
msft_model <- lm(msft_excess ~ mkt_excess, data = capm5)
dis_model <- lm(dis_excess ~ mkt_excess, data = capm5)
xom_model <- lm(xom_excess ~ mkt_excess, data = capm5)

# 顯示回歸結果
summary(ge_model)
```

```
##
## Call:
## lm(formula = ge_excess ~ mkt_excess, data = capm5)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.173801 -0.033907 -0.003789  0.038858  0.202748
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.0009587  0.0044244  -0.217    0.829
## mkt_excess   1.1479521  0.0895394  12.821   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.05925 on 178 degrees of freedom
## Multiple R-squared:  0.4801, Adjusted R-squared:  0.4772
## F-statistic: 164.4 on 1 and 178 DF,  p-value: < 2.2e-16
```

```
summary(ibm_model)
```

```
##
## Call:
## lm(formula = ibm_excess ~ mkt_excess, data = capm5)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.254880 -0.035266 -0.005322  0.031490  0.274520
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.006053   0.004834   1.252    0.212
## mkt_excess  0.976890   0.097839   9.985   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.06474 on 178 degrees of freedom
## Multiple R-squared:  0.359,  Adjusted R-squared:  0.3554
## F-statistic: 99.69 on 1 and 178 DF,  p-value: < 2.2e-16
```

```
summary(ford_model)
```

```
##
## Call:
## lm(formula = ford_excess ~ mkt_excess, data = capm5)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.27315 -0.07875 -0.01198  0.04720  1.08874
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.003779   0.010225   0.370    0.712
## mkt_excess  1.662031   0.206937   8.032 1.27e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1369 on 178 degrees of freedom
## Multiple R-squared:  0.266,  Adjusted R-squared:  0.2619
## F-statistic: 64.51 on 1 and 178 DF,  p-value: 1.271e-13
```

```
summary(msft_model)
```

```
##
## Call:
## lm(formula = msft_excess ~ mkt_excess, data = capm5)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.27424 -0.04744 -0.00820  0.03869  0.35801
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.003250   0.006036   0.538    0.591
## mkt_excess  1.201840   0.122152   9.839   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.08083 on 178 degrees of freedom
## Multiple R-squared:  0.3523, Adjusted R-squared:  0.3486
## F-statistic:  96.8 on 1 and 178 DF,  p-value: < 2.2e-16
```

```
summary(dis_model)
```

```
##
## Call:
## lm(formula = dis_excess ~ mkt_excess, data = capm5)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.176233 -0.030021 -0.004232  0.029179  0.280528
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.001047   0.004676   0.224    0.823
## mkt_excess  1.011521   0.094638  10.688   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.06262 on 178 degrees of freedom
## Multiple R-squared:  0.3909, Adjusted R-squared:  0.3875
## F-statistic: 114.2 on 1 and 178 DF,  p-value: < 2.2e-16
```

```
summary(xom_model)
```

```
##
## Call:
## lm(formula = xom_excess ~ mkt_excess, data = capm5)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.114474 -0.030596 -0.001884  0.026849  0.215396
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.005284   0.003535   1.494    0.137
## mkt_excess  0.456521   0.071550   6.380 1.48e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.04734 on 178 degrees of freedom
## Multiple R-squared:  0.1861, Adjusted R-squared:  0.1816
## F-statistic: 40.71 on 1 and 178 DF,  p-value: 1.48e-09
```

```
# 提取各公司的 Beta 值
beta_values <- data.frame(
  Company = c("GE", "IBM", "Ford", "Microsoft", "Disney", "ExxonMobil"),
  Beta = c(coef(ge_model)[2], coef(ibm_model)[2], coef(ford_model)[2],
           coef(msft_model)[2], coef(dis_model)[2], coef(xom_model)[2])
)

# 印出 Beta 值
print(beta_values)
```

```
##      Company      Beta
## 1         GE 1.1479521
## 2        IBM 0.9768898
## 3       Ford 1.6620307
## 4  Microsoft 1.2018398
## 5     Disney 1.0115207
## 6 ExxonMobil 0.4565208
```

Ford（1.662）appears most aggressive. ExxonMobil（0.457）appears most defensive.

Q16 c

Finance theory says that the intercept parameter α_j should be zero. Does this seem correct given your estimates? For the Microsoft stock, plot the fitted regression line along with the data scatter.

Ans

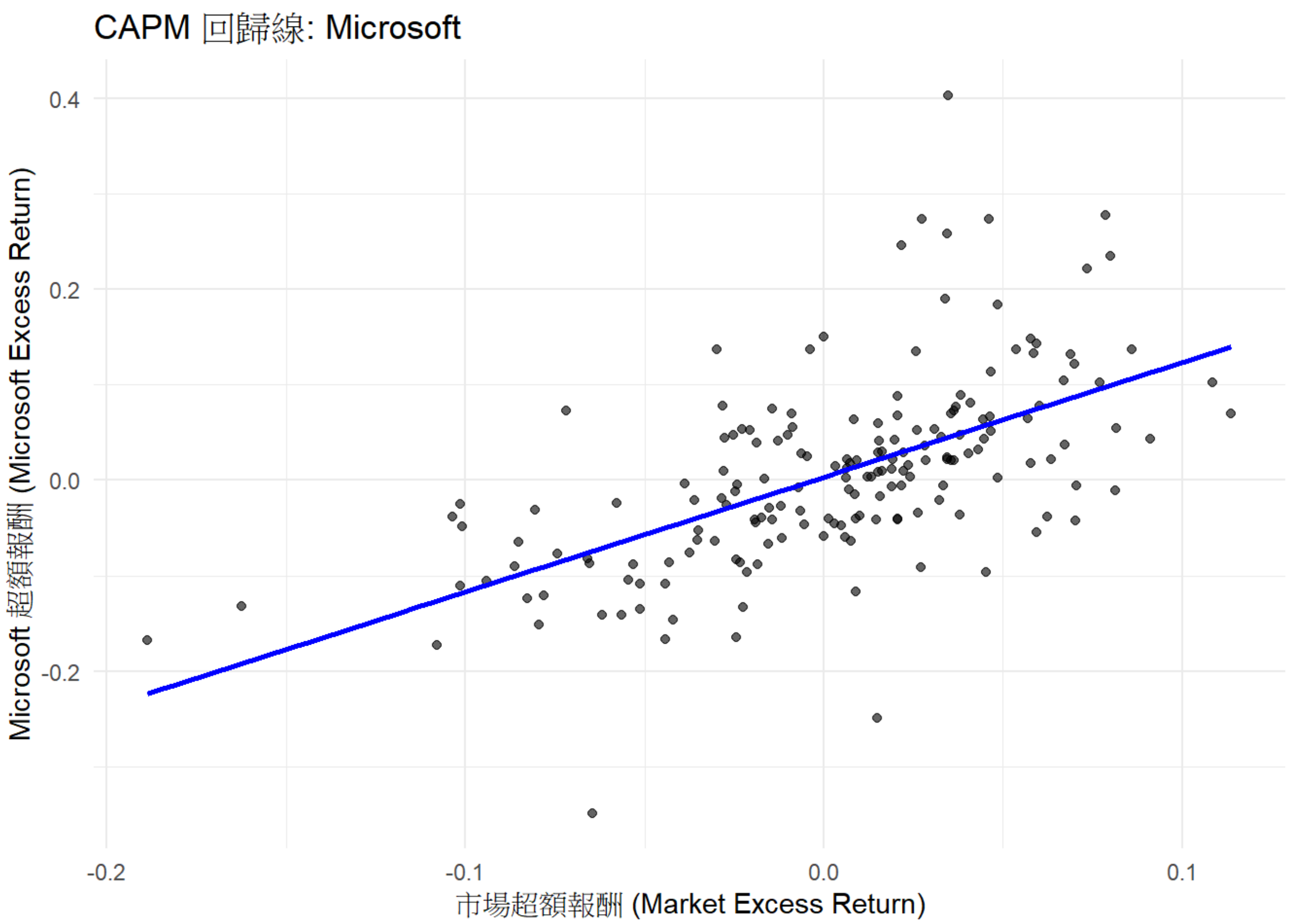
Forall companies' Intercept term's p-value is greater than 0.05. Thus we can not reject $\alpha_j = 0$, which make the assumption of CAPM holds.

```
# 建立回歸結果表格（使用 (b) 小題已經執行的回歸模型）
reg_results <- data.frame(
  Company = c("GE", "IBM", "Ford", "Microsoft", "Disney", "ExxonMobil"),
  Alpha = c(coef(ge_model)[1], coef(ibm_model)[1], coef(ford_model)[1],
            coef(msft_model)[1], coef(dis_model)[1], coef(xom_model)[1]),
  Alpha_p_value = c(summary(ge_model)$coefficients[1,4],
                    summary(ibm_model)$coefficients[1,4],
                    summary(ford_model)$coefficients[1,4],
                    summary(msft_model)$coefficients[1,4],
                    summary(dis_model)$coefficients[1,4],
                    summary(xom_model)$coefficients[1,4]),
  Beta = c(coef(ge_model)[2], coef(ibm_model)[2], coef(ford_model)[2],
           coef(msft_model)[2], coef(dis_model)[2], coef(xom_model)[2])
)

# 印出結果表格
print(reg_results)
```

```
##      Company         Alpha Alpha_p_value      Beta
## 1         GE -0.0009586682     0.8287072 1.1479521
## 2        IBM  0.0060525497     0.2122303 0.9768898
## 3       Ford  0.0037789112     0.7121467 1.6620307
## 4  Microsoft  0.0032496009     0.5909844 1.2018398
## 5     Disney  0.0010469237     0.8231091 1.0115207
## 6 ExxonMobil  0.0052835329     0.1368343 0.4565208
```

```r
# 繪製 Microsoft 的 CAPM 回歸圖（使用 (b) 小題已有的數據）
library(ggplot2)
ggplot(capm5, aes(x = mkt_excess, y = msft_excess)) +
  geom_point(alpha = 0.6) +  # 散點圖
  geom_smooth(method = "lm", formula = y ~ x, color = "blue", se = FALSE) +  # 回歸線
  labs(title = "CAPM 回歸線: Microsoft",
       x = "市場超額報酬 (Market Excess Return)",
       y = "Microsoft 超額報酬 (Microsoft Excess Return)") +
  theme_minimal()
```



Q16 (d)

Estimate the model for each firm under the assumption that $\alpha_j = 0$. Do the estimates of the "beta" values change much?

Ans

```
# 重新估計 CAPM 模型，假設 α_j = 0 (無截距模型)
ge_model_no_intercept <- lm(ge_excess ~ mkt_excess - 1, data = capm5)
ibm_model_no_intercept <- lm(ibm_excess ~ mkt_excess - 1, data = capm5)
ford_model_no_intercept <- lm(ford_excess ~ mkt_excess - 1, data = capm5)
msft_model_no_intercept <- lm(msft_excess ~ mkt_excess - 1, data = capm5)
dis_model_no_intercept <- lm(dis_excess ~ mkt_excess - 1, data = capm5)
xom_model_no_intercept <- lm(xom_excess ~ mkt_excess - 1, data = capm5)

# 提取新的 Beta 值
beta_no_intercept <- data.frame(
  Company = c("GE", "IBM", "Ford", "Microsoft", "Disney", "ExxonMobil"),
  Beta_Without_Alpha = c(coef(ge_model_no_intercept)[1], coef(ibm_model_no_intercept)[1],
                         coef(ford_model_no_intercept)[1], coef(msft_model_no_intercept)[1],
                         coef(dis_model_no_intercept)[1], coef(xom_model_no_intercept)[1])
)

# 合併原來的 Beta 值，做比較
beta_comparison <- merge(reg_results[, c("Company", "Beta")], beta_no_intercept, by = "Compan
y")
colnames(beta_comparison) <- c("Company", "Beta_With_Alpha", "Beta_Without_Alpha")

# 顯示結果
print(beta_comparison)
```

```
##      Company Beta_With_Alpha Beta_Without_Alpha
## 1     Disney       1.0115207          1.0128190
## 2 ExxonMobil       0.4565208          0.4630727
## 3       Ford       1.6620307          1.6667168
## 4         GE       1.1479521          1.1467633
## 5        IBM       0.9768898          0.9843954
## 6  Microsoft       1.2018398          1.2058695
```

Company	Beta (Intercept)	Beta (without Intercept)	Change
Disney	1.0115	1.0128	**+0.0013**
ExxonMobil	0.4565	0.4631	**+0.0066**
Ford	1.6620	1.6667	**+0.0047**
GE	1.1480	1.1468	**-0.0012**
IBM	0.9769	0.9844	**+0.0075**
Microsoft	1.2018	1.2059	**+0.0041**

- change the most is **IBM**（**+0.0075**） but almost approach to 0
- less change on **GE**（**-0.0012**），barely changed。

The estimates of the beta value change nearly nothing.