Exercise 10

0316213 Yu-Wen Pu

2018-06-10

11.1

a) 解釋如下：

- 在其他條件不變的情況下，`Temp` 每增加 1 單位，`Quality of Life` 就會降低 0.01 單位。
- 在其他條件不變的情況下，`Income` 每增加 1 單位，`Quality of Life` 就會增加 0.05 單位。
- 在其他條件不變的情況下，`SocSer` 每增加 1 單位，`Quality of Life` 就會增加 0.003 單位。
- 在其他條件不變的情況下，`Popul` 每增加 1 單位，`Quality of Life` 就會降低 0.01 單位。

b) Quality of Life =

```
5.37 - 0.01 * 55 + 0.05 * 12 + 0.003 * 500 - 0.01 * 200
```

```
## [1] 4.92
```

c) Quality of Life =

```
5.37 - 0.01 * 55 + 0.05 * 12 + 0.003 * 100 - 0.01 * 200
```

```
## [1] 3.72
```

11.2

a) No, not reliable. 因為 `Rel Inf`、`Rel Invol`、`Rel Hope` 這三個變數很可能高度相關。
b) Little relationship. 因為 adjusted r square 只有 0.099。
c) **a** 小題的答案不變，**b** 小題關係可能變強。

11.3

`Rel Inf` 與 `Rel Hope` 都有顯著貢獻。

11.6

```
d = data.frame(Y = c(5, 0, 5, 9, 4,
                     8, 3, 7, 0, 4,
                     7, 1, 4, 7, 9),
               X1 = c(3, 8, 1, 5, 8,
                      2, 4, 7, 9, 1,
                      3, 5, 6, 8, 9),
               X2 = c(7, 6, 4, 3, 1,
                      9, 7, 5, 3, 1,
                      8, 6, 0, 3, 7),
               X3 = c(1, 7, 4, 1, 8,
                      8, 6, 8, 3, 6,
                      1, 9, 7, 7, 7),
               X4 = c(3, 6, 0, 5, 1,
                      3, 5, 9, 1, 1,
                      7, 4, 2, 0, 9))
```

```
ds <- d[1:15, ]
summary(lm(ds$Y ~ ds$X1 + ds$X2 + ds$X3 + ds$X4))
```

```
##
## Call:
## lm(formula = ds$Y ~ ds$X1 + ds$X2 + ds$X3 + ds$X4)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.7772 -1.7964 -0.0455  1.8490  4.4261
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   5.8709     2.9505   1.990   0.0746 .
## ds$X1        -0.2621     0.3563  -0.736   0.4788
## ds$X2        -0.1076     0.4126  -0.261   0.7996
## ds$X3        -0.1253     0.3226  -0.388   0.7059
## ds$X4         0.4210     0.3790   1.111   0.2926
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.225 on 10 degrees of freedom
## Multiple R-squared:  0.1726, Adjusted R-squared:  -0.1583
## F-statistic: 0.5216 on 4 and 10 DF,  p-value: 0.7223
```

```
ds <- d[1:10, ]
summary(lm(ds$Y ~ ds$X1 + ds$X2 + ds$X3 + ds$X4))
```

```
##
## Call:
## lm(formula = ds$Y ~ ds$X1 + ds$X2 + ds$X3 + ds$X4)
##
## Residuals:
##       1       2       3       4       5       6       7       8       9
## -0.2963 -3.3886 -0.3027  2.9375  2.4465  3.1572 -2.3285  1.0603 -0.8870
##      10
## -2.3984
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  7.27451    3.34760   2.173   0.0818 .
## ds$X1       -0.66550    0.39968  -1.665   0.1568
## ds$X2       -0.23342    0.46818  -0.499   0.6392
## ds$X3       -0.09317    0.38812  -0.240   0.8198
## ds$X4        0.58179    0.46215   1.259   0.2636
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.141 on 5 degrees of freedom
## Multiple R-squared:  0.4021, Adjusted R-squared:  -0.07627
## F-statistic: 0.8406 on 4 and 5 DF,  p-value: 0.5543
```

```
ds <- d[1:6, ]
summary(lm(ds$Y ~ ds$X1 + ds$X2 + ds$X3 + ds$X4))
```

```
##
## Call:
## lm(formula = ds$Y ~ ds$X1 + ds$X2 + ds$X3 + ds$X4)
```

```
##
## Residuals:
##       1       2       3       4       5       6
##  0.4092 -2.8515 -2.7987  1.5974  1.5182  2.1254
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  11.8967     8.4488   1.408    0.393
## ds$X1        -1.4297     1.8578  -0.770    0.582
## ds$X2        -0.9047     1.6330  -0.554    0.678
## ds$X3         0.2377     1.0822   0.220    0.862
## ds$X4         1.0262     2.0631   0.497    0.706
##
## Residual standard error: 5.05 on 1 degrees of freedom
## Multiple R-squared:  0.4983, Adjusted R-squared:  -1.509
## F-statistic: 0.2483 on 4 and 1 DF,  p-value: 0.8848
```

```
ds <- d[1:5, ]
summary(lm(ds$Y ~ ds$X1 + ds$X2 + ds$X3 + ds$X4))
```

```
##
## Call:
## lm(formula = ds$Y ~ ds$X1 + ds$X2 + ds$X3 + ds$X4)
##
## Residuals:
## ALL 5 residuals are 0: no residual degrees of freedom!
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.300e+01         NA      NA       NA
## ds$X1        6.484e-16         NA      NA       NA
## ds$X2       -1.000e+00         NA      NA       NA
## ds$X3       -1.000e+00         NA      NA       NA
## ds$X4       -9.118e-16         NA      NA       NA
##
## Residual standard error: NaN on 0 degrees of freedom
## Multiple R-squared:      1,  Adjusted R-squared:    NaN
## F-statistic:   NaN on 4 and 0 DF,  p-value: NA
```

```
ds <- d[1:4, ]
summary(lm(ds$Y ~ ds$X1 + ds$X2 + ds$X3 + ds$X4))
```

```
##
## Call:
## lm(formula = ds$Y ~ ds$X1 + ds$X2 + ds$X3 + ds$X4)
##
## Residuals:
## ALL 4 residuals are 0: no residual degrees of freedom!
##
## Coefficients: (1 not defined because of singularities)
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.300e+01         NA      NA       NA
## ds$X1       -2.147e-17         NA      NA       NA
## ds$X2       -1.000e+00         NA      NA       NA
## ds$X3       -1.000e+00         NA      NA       NA
```

```
## ds$X4               NA         NA      NA       NA
##
## Residual standard error: NaN on 0 degrees of freedom
## Multiple R-squared:      1,  Adjusted R-squared:    NaN
## F-statistic:   NaN on 3 and 0 DF,  p-value: NA
```

當資料量太少時，沒有自由度，無法進行多重相關 / 迴歸分析。

11.24

當所有 predictor 都為 0 時，$\hat{Y}$ = intercept。因為在現實世界裡幾乎不可能所有 predictor 都為 0，所以我們說 intercept 沒有實際意義。可是如果沒有 intercept，當所有 predictor 都為 0 時，$\hat{Y} = 0$ 。整條迴歸線被平移了，將非 0 的值代入迴歸方程式得到的 $\hat{Y}$ 也會增加了 intercept 大小的誤差。