Homework 4

數據所 0656704 賴品儒

第一題 (Data：waste water)

- 廢水處理

- 三種去除 organic carbon 方法 (treatment)：

    a. Air flotation (AF)

    b. Foam separation (FS)

    c. Ferric-chloride coagulation (FCC)

(the organic carbon material measurements)

AF	FS	FCC
34.6	38.8	26.7
35.1	39.0	26.7
35.3	40.1	27.0
35.8	40.9	27.1
36.1	41.0	27.5
36.5	43.2	28.1
36.8	44.9	28.1
37.2	46.9	28.7
37.4	51.6	30.7
37.7	53.6	31.2

(a) Test $H_0 : \mu_1 = \mu_2 = \mu_3$ at the $\alpha = .10$ level

$n = 30$，$n_1 = n_2 = n_3 = 10$

$K=3$

```
            Df Sum Sq Mean Sq F value   Pr(>F)
treatment    2 1251.5   625.8   60.63 1.03e-10 ***
Residuals   27  278.7    10.3
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

$$F = \frac{\frac{SSTR}{K-1}}{\frac{SSE}{n-K}} \sim F(2,27)$$

$$F^*_{obs} = \frac{\frac{1251.5}{3-1}}{\frac{278.7}{30-3}} \approx 60.63$$

Reject H_0 if $F^*_{obs} > F_{\alpha=0.1}(2,27)$

$F_{\alpha=0.1}(2,27) \approx 2.510609$

結論：Reject H_0，我們有足夠證據顯示在 $\alpha = .10$ level 下至少有一個 μ_i 不相等

第二題 (Data：fern)

光線對蕨類生長的影響，因為不同 ages 有不同生長 rates，故 Variable 由 blocking 控制

- Block (age)
  - young plants (plants grown in the dark for 4 days)
  - older plants (plants grown in the dark for 12 days)
- Treatment (wavelength of light)

  720、460、600、720

Block (age)	Treatment (wavelength of light)			
	420 nm	460 nm	600 nm	720 nm
Young	1017.6	929.0	939.8	1081.5
Old	854.7	689.9	841.5	797.4

(a) Find the sample treatment, block, and grand totals and means

K=4 , b=2 ,n=K*b

- sample treatment：

Treatment	sample	total	mean
420 nm	1017.6, 854.7	1872.3	936.15
460nm	929.0, 689.9	1618.9	809.45
600nm	939.8, 841.5	1781.3	890.65
720nm	1081.5, 797.4	1878.9	939.45

- block：

block	sample	total	mean
Young	1017.6, 929.0, 939.8, 1081.5	3967.9	991.975
Old	854.7, 689.9, 841.5, 797.4	3183.5	795.875

- Grand total = 7151.4
- Grand mean = 893.925

(b) Test the null hypothesis of equal treatment means

$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$

```
              Df Sum Sq Mean Sq F value Pr(>F)
treatment      3  22004    7335   2.172 0.2702
block          1  76910   76910  22.777 0.0175 *
Residuals      3  10130    3377
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

在 H_0 下 $F = \dfrac{SSTR/(K-1)}{SSE/(b-1)(K-1)} = \dfrac{MSTR}{MSE} \sim F(3,3)$

$F^*_{obs} = \dfrac{\dfrac{22004}{4-1}}{\dfrac{10130}{3*1}} \approx 2.172$

Reject H_0 if $F^*_{obs} > F_{table}(p - value < 0.05)$

$F_{\alpha=0.05}(3,3) \approx 9.2766$

結論：Don't reject H_0，我們沒有足夠證據顯示在 $\alpha = .05$ level下 μ_i 不相等

第三題 (Data：cotinine)

Cotinine → indicator of tobacco smoke exposure
- Data：cotinine level mg/ml

  檢測不同種族的 cotinine level

	White	Black
Male	210	245
	300	347
	150	125
	325	250
(1085)	100	260 (1227)
Female	177	252
	300	152
	106	315
	150	267
(893)	160	275 (1261)

- Factor 1： racial (White, Black)
- Factor 2 ：gender (Male, Female)

(a) Construct a two-way ANOVA table for these data and use it to test the null hypothesis of no interaction

Race <fctr>	Gender <fctr>	Count <int>	Sum <int>	Mean <dbl>
Black	Female	5	1261	252.2
Black	Male	5	1227	245.4
White	Female	5	893	178.6
White	Male	5	1085	217.0

假設 model : $Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}$

Testing : $H_0 : (\alpha\beta)_{11} = \cdots = (\alpha\beta)_{k_1 k_2} = 0$ (no interaction)

- factor1 : Race

Factor2 : Gender

```
                  Df Sum Sq Mean Sq F value Pr(>F)
factor1            1  13005   13005   2.129  0.164
factor2            1   1248    1248   0.204  0.657
factor1:factor2    1   2554    2554   0.418  0.527
Residuals         16  97731    6108
```

Reject H_0 if p-value $< \alpha = 0.05$

p-value = 0.527 >0.05

結論：Don't reject H_0，我們沒有足夠證據顯示在$\alpha = 0.05$ level下有交互作用

(b) If no interaction is found, test for main effects
- 針對 factor1 : Race

```
            Df Sum Sq Mean Sq F value Pr(>F)
factor1      1  13005   13005   2.306  0.146
Residuals   18 101533    5641
```

$H_0 : \mu_{white} = \mu_{black}$

p-value = 0.145 > 0.05

結論：Don't reject H_0，我們沒有足夠證據顯示在$\alpha = 0.05$ level下μ_i不相等

- 針對 factor1 : Gender

```
            Df Sum Sq Mean Sq F value Pr(>F)
factor2      1   1248    1248   0.198  0.661
Residuals   18 113290    6294
```
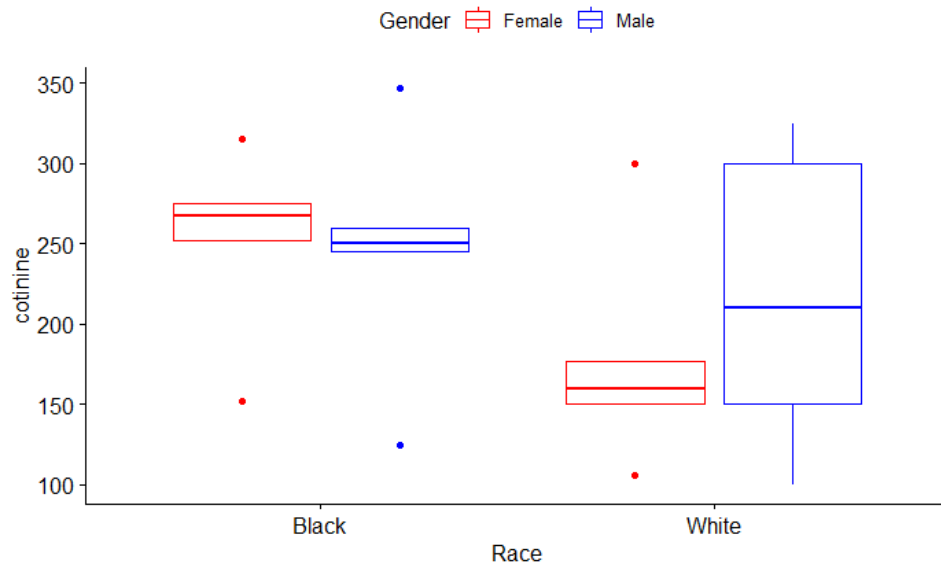
$H_0 : \mu_{female} = \mu_{male}$

p-value = 0.661 > 0.05

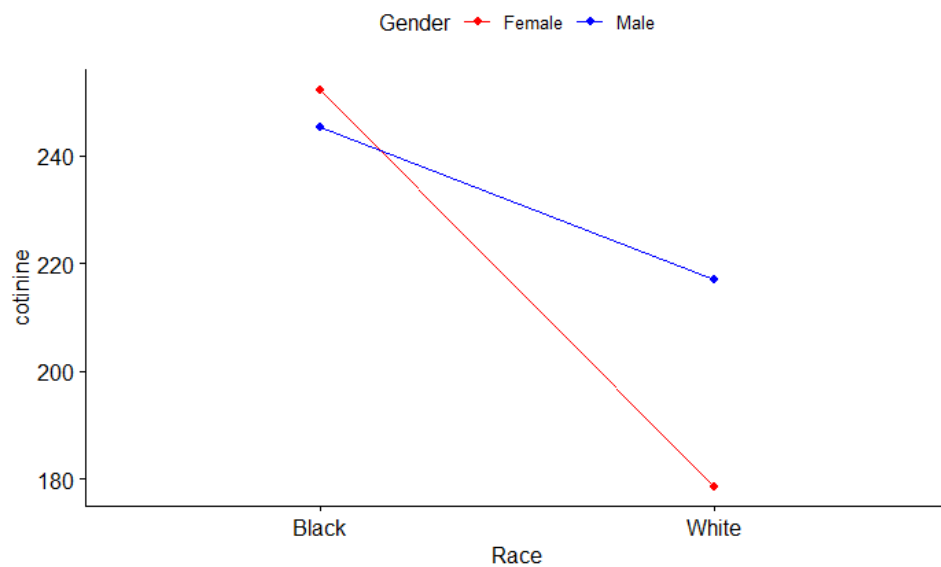結論：Don't reject H_0，我們沒有足夠證據顯示在$\alpha = 0.05$ level下μ_i不相等

(c) If interaction is detected, construct a diagram similar to that shown in Figure 10.6 to investigate the nature of the interaction

→ no interaction is detected

- 有 noise (outlier)，影響 mean

- 下圖只顯示平均，看不出個體差異，即便有相交，但 factor Race、Gender 交互作用不顯著

(d) If interaction is detected, compare the mean cotinine level between whites and blacks for females via a one-way ANOVA. Do the same for males

→ no interaction is detected

- For female：

```
            Df  Sum Sq  Mean Sq  F value  Pr(>F)
treatment    1   13542    13542    3.018   0.121
Residuals    8   35898     4487
```

$H_0 : \mu_{white} = \mu_{black}$

p-value = 0.121 > 0.05

結論：Don't reject H_0，我們沒有足夠證據顯示在$\alpha = 0.05$ level下μ_i不相等

- For male：

```
            Df  Sum Sq  Mean Sq  F value  Pr(>F)
treatment    1    2016     2016    0.261   0.623
Residuals    8   61833     7729
```

$H_0 : \mu_{white} = \mu_{black}$

p-value =0.623 > 0.05

結論：Don't reject H_0，我們沒有足夠證據顯示在$\alpha = 0.05$ level下μ_i不相等

原始碼：https://github.com/ayamisea/Applied-Methods-in-Statistics/blob/master/homework/HW04/HW04.Rmd