HW0414　丁振朝 311707021

8a.

檢驗男性與女性的誤差變異數是否相等。

$H_0$（虛無假設）：$\sigma_M^2 = \sigma_F^2$

$H_1$（對立假設）：$\sigma_M^2 \neq \sigma_F^2$

男性變異數 = $$\hat{\sigma}_M^2 = \frac{97161.9174}{577-4} = \frac{97161.9174}{573} \approx 169.54$$

$$F = \frac{\hat{\sigma}_M^2}{\hat{\sigma}_F^2} = \frac{169.54}{12.024} \approx 14.10$$

檢定統計量 F≈14.10

臨界區間：若 F<0.8 或 F>1.25 就拒絕 $H_0$

因為 14.10 > 1.25，所以 **拒絕虛無假設**

b.

$H_0$（虛無假設）：$\sigma^2_{\text{SINGLE}} = \sigma^2_{\text{MARRIED}}$

$H_1$（對立假設）：$\sigma^2_{\text{MARRIED}} > \sigma^2_{\text{SINGLE}}$（右尾檢定）

$$\hat{\sigma}_S^2 = \frac{SSE_S}{n_S - k} = \frac{56231.0382}{395} \approx 142.37$$

$$\hat{\sigma}_M^2 = \frac{SSE_M}{n_M - k} = \frac{100703.0471}{595} \approx 169.22$$

$$F = \frac{\hat{\sigma}_M^2}{\hat{\sigma}_S^2} = \frac{169.22}{142.37} \approx 1.1886$$

檢定統計量：F=1.1886

臨界值：F>1.21

因為 1.1886 < 1.21，所以**無法拒絕虛無假設**

c.

$$\chi^2_{(4),0.95} = 9.488$$

NR² = 59.03 > 9.488 → **落在拒絕域**

所以我們 **拒絕虛無假設**

代表：有統計證據顯示存在**異方差**

d.

degrees of freedom = 14

$$\chi^2_{(14),0.95} \approx 23.685$$

因為 78.82 > 23.685 → 落在拒絕區間

e.

| 變數 | 一般 SE | Robust SE | 寬窄變化 |
|---|---|---|---|
| 截距 | 2.36 | 2.50 | **變寬** |
| EDUC | 0.14 | 0.16 | **變寬** |
| EXPER | 0.031 | 0.029 | **變窄** ✅ |
| METRO | 1.05 | 0.84 | **變窄** ✅ |
| FEMALE | 0.81 | 0.80 | 幾乎不變 ✅ |

f.

在 (b)，我們測的是「**誤差變異**是否因 MARRIED 而異」，也就是 MARRIED 是否與 **異方差有關**。在這題，我們測的是「**MARRIED 是否顯著影響薪資水準（係數顯著性），不衝突**

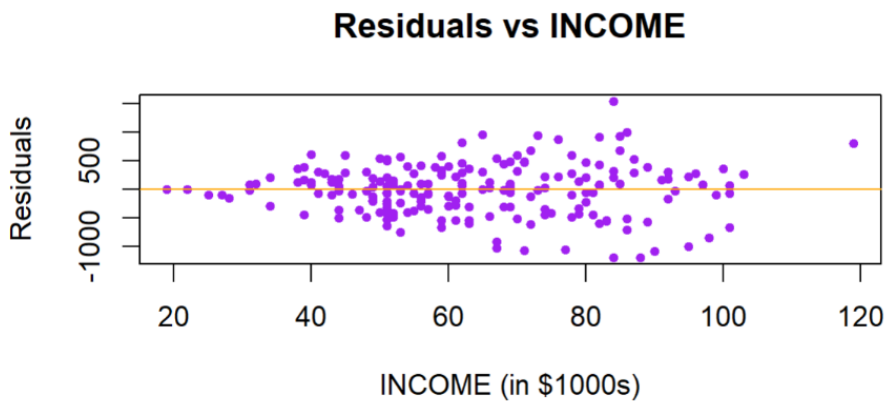
16

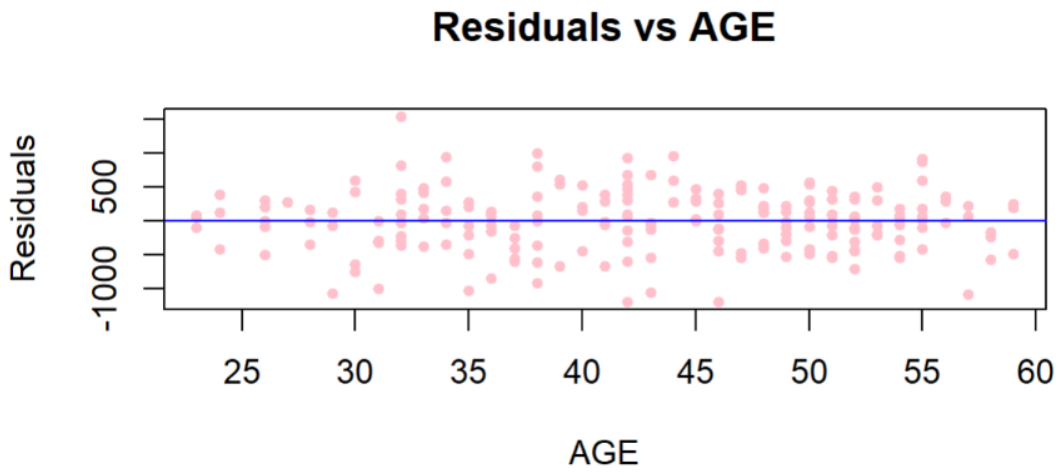
a.

```
Residuals:
     Min       1Q   Median       3Q      Max
-1198.14  -295.31    17.98   287.54  1549.41

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -391.548    169.775  -2.306   0.0221 *
income        14.201      1.800   7.889 2.10e-13 ***
age           15.741      3.757   4.189 4.23e-05 ***
kids         -81.826     27.130  -3.016   0.0029 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 452.3 on 196 degrees of freedom
Multiple R-squared:  0.3406,    Adjusted R-squared:  0.3305
F-statistic: 33.75 on 3 and 196 DF,  p-value: < 2.2e-16

> confint(model, "kids", level = 0.95)
          2.5 %    97.5 %
kids -135.3298 -28.32302
```

在控制收入（INCOME）與年齡（AGE）的情況下，每多一位小孩，平均每年旅行距離減少約 **81.826 英里**。此估計在 95% 信賴區間下為顯著的負向影響（因為信賴區間不包含 0）。

b.






圖形顯示 **殘差與 INCOME 之間存在變異擴大的現象**，這表明回歸模型中**可能存在異質變異數（heteroskedasticity）**。

c.

| 項目 | 數值 |
|---|---|
| SSE（低收入組） | 101744.65 |
| SSE（高收入組） | 315821.55 |
| F 統計量 | **3.104** |
| 臨界值（$\alpha = 0.05$，$df_1 = 86, df_2 = 86$） | **1.4286** |

- **虛無假設（H_0）**：$\sigma^2_1 = \sigma^2_2$ (同方差性) 高、低收入組變異數相等
- **對立假設（H_1）**：$\sigma^2_1 < \sigma^2_2$ (異方差性) 高收入組誤差變異較大

d.

```
t test of coefficients:

             Estimate Std. Error t value  Pr(>|t|)
(Intercept) -391.5480   142.6548 -2.7447 0.0066190 **
income        14.2013     1.9389  7.3246 6.083e-12 ***
age           15.7409     3.9657  3.9692 0.0001011 ***
kids         -81.8264    29.1544 -2.8067 0.0055112 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

> confint_robust <- coeftest(ols_model, vcov. = robust_se)
> beta_kids <- confint_robust["kids", 1]
> se_kids <- confint_robust["kids", 2]
> lower <- beta_kids - 1.96 * se_kids
> upper <- beta_kids + 1.96 * se_kids
> cat("Robust 95% CI for KIDS:", round(lower, 3), "to", round(upper, 3), "\n")
Robust 95% CI for KIDS: -138.969 to -24.684
```

- **估計值一致**：因為模型本身沒改變。
- **標準誤變大**：robust SE 考慮了異質變異數影響。
- **信賴區間變寬**：提高推論的保守性與真實性。

e.

```
Call:
lm(formula = miles ~ income + age + kids, data = vacation, weights = weights_gls)

Weighted Residuals:
     Min       1Q   Median       3Q      Max
-15.1907  -4.9555   0.2488   4.3832  18.5462

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -424.996    121.444  -3.500 0.000577 ***
income        13.947      1.481   9.420  < 2e-16 ***
age           16.717      3.025   5.527 1.03e-07 ***
kids         -76.806     21.848  -3.515 0.000545 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 6.765 on 196 degrees of freedom
Multiple R-squared:  0.4573,    Adjusted R-squared:  0.449
F-statistic: 55.06 on 3 and 196 DF,  p-value: < 2.2e-16

 t test of coefficients:

              Estimate Std. Error t value  Pr(>|t|)
 (Intercept) -424.9962    95.8035 -4.4361 1.526e-05 ***
 income        13.9473     1.3470 10.3545 < 2.2e-16 ***
 age           16.7175     2.7974  5.9761 1.061e-08 ***
 kids         -76.8063    22.6186 -3.3957 0.0008286 ***
 ---
 Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

解釋與比較：

- GLS 模型的估計值與 OLS 相比 **略小**（影響較弱），但其 **標準誤顯著縮**

小。

- 使用 robust GLS 後，標準誤略微上升，但仍**明顯小於 robust OLS**。
- Robust GLS 確實提供更窄的信賴區間與更有效的估計，顯示推論更精確

Q18

a.

由計算的到的 F 值 = 1.05076，沒有落在拒絕域，無法拒絕 H0

b.

H0 : a5 =a6 = a7 = 0, H1 : not all H0 =0

檢定值 = 23.55681 大於臨界值 = 11.3487，拒絕 H0

c.

H0 = homoskedasticity, H1 = hetroskedasticity

檢定統計 = 194.4447, 臨界值 = 60.48089，拒絕 H0

d.

OLS:

```
lm(formula = log(wage) ~ educ + exper + I(exper^2) + female +
    black + metro + south + midwest + west, data = data)

Residuals:
     Min       1Q   Median       3Q      Max
-2.31711 -0.30038 -0.00584  0.30238  3.00061

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.201e+00  3.211e-02  37.409  < 2e-16 ***
educ         1.012e-01  1.758e-03  57.574  < 2e-16 ***
exper        2.962e-02  1.300e-03  22.780  < 2e-16 ***
I(exper^2)  -4.458e-04  2.635e-05 -16.915  < 2e-16 ***
female      -1.655e-01  9.529e-03 -17.368  < 2e-16 ***
black       -1.115e-01  1.694e-02  -6.583 4.86e-11 ***
metro        1.190e-01  1.231e-02   9.671  < 2e-16 ***
south       -4.576e-02  1.356e-02  -3.374 0.000744 ***
midwest     -6.394e-02  1.410e-02  -4.534 5.86e-06 ***
west        -6.589e-03  1.440e-02  -0.458 0.647321
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.466 on 9789 degrees of freedom
Multiple R-squared:  0.3173,    Adjusted R-squared:  0.3167
F-statistic: 505.6 on 9 and 9789 DF,  p-value: < 2.2e-16
```

White:

```
t test of coefficients:

               Estimate  Std. Error  t value  Pr(>|t|)
(Intercept)  1.2014e+00  3.2777e-02  36.6527 < 2.2e-16 ***
educ         1.0123e-01  1.9048e-03  53.1431 < 2.2e-16 ***
exper        2.9622e-02  1.3142e-03  22.5391 < 2.2e-16 ***
I(exper^2)  -4.4578e-04  2.7583e-05 -16.1615 < 2.2e-16 ***
female      -1.6550e-01  9.4834e-03 -17.4517 < 2.2e-16 ***
black       -1.1153e-01  1.6085e-02  -6.9333 4.371e-12 ***
metro        1.1902e-01  1.1576e-02  10.2814 < 2.2e-16 ***
south       -4.5755e-02  1.3895e-02  -3.2931 0.0009946 ***
midwest     -6.3943e-02  1.3717e-02  -4.6615 3.180e-06 ***
west        -6.5891e-03  1.4549e-02  -0.4529 0.6506470
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

變數 educ, exper, exper$^2$, south, west. SE 變大 CI 變寬表示低估不確定性

變數 female, black, metro, Midwest. SE 變小 CI 變窄表示高估不確定性

e.

```
Call:
lm(formula = log(wage) ~ educ + exper + I(exper^2) + female +
    black + metro + south + midwest + west, data = data, weights = weights)

Weighted Residuals:
    Min      1Q  Median      3Q     Max
-4.7199 -0.6168 -0.0112  0.6182  6.1542

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.196e+00  3.184e-02  37.571  < 2e-16 ***
educ         1.015e-01  1.761e-03  57.604  < 2e-16 ***
exper        2.986e-02  1.299e-03  22.988  < 2e-16 ***
I(exper^2)  -4.510e-04  2.657e-05 -16.971  < 2e-16 ***
female      -1.658e-01  9.505e-03 -17.446  < 2e-16 ***
black       -1.112e-01  1.697e-02  -6.553 5.91e-11 ***
metro        1.184e-01  1.186e-02   9.979  < 2e-16 ***
south       -4.527e-02  1.354e-02  -3.343 0.000833 ***
midwest     -6.355e-02  1.405e-02  -4.524 6.13e-06 ***
west        -6.060e-03  1.439e-02  -0.421 0.673671
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.948 on 9789 degrees of freedom
Multiple R-squared:  0.3193,    Adjusted R-squared:  0.3187
F-statistic: 510.2 on 9 and 9789 DF,  p-value: < 2.2e-16
```

變數 educ, exper, exper$^2$, south, west. SE 變小 CI 變窄表示估計精準度上升

變數 female, black, metro, Midwest. SE 變大 CI 變寬表示估計不確定性增加

f.

| Variable | OLS_Robust_SE | FGLS_Robust_SE | SE_Change | OLS_Robust_Width | FGLS_Robust_Width | CI_Change |
|---|---|---|---|---|---|---|
| (Intercept) | 3.277743e-02 | 3.250910e-02 | 變小 | 0.1285010626 | 0.1274490902 | 變窄 |
| educ | 1.904848e-03 | 1.895323e-03 | 變小 | 0.0074677917 | 0.0074304492 | 變窄 |
| exper | 1.314237e-03 | 1.307055e-03 | 變小 | 0.0051523513 | 0.0051241957 | 變窄 |
| I(exper^2) | 2.758278e-05 | 2.744395e-05 | 變小 | 0.0001081359 | 0.0001075916 | 變窄 |
| female | 9.483417e-03 | 9.445177e-03 | 變小 | 0.0371789103 | 0.0370289927 | 變窄 |
| black | 1.608548e-02 | 1.595853e-02 | 變小 | 0.0630617199 | 0.0625640260 | 變窄 |
| metro | 1.157624e-02 | 1.155933e-02 | 變小 | 0.0453836492 | 0.0453173516 | 變窄 |
| south | 1.389454e-02 | 1.384176e-02 | 變小 | 0.0544723330 | 0.0542654167 | 變窄 |
| midwest | 1.371725e-02 | 1.369010e-02 | 變小 | 0.0537772884 | 0.0536708228 | 變窄 |
| west | 1.454941e-02 | 1.450663e-02 | 變小 | 0.0570397063 | 0.0568719873 | 變窄 |

| | FGLS_SE | FGLS_Robust_SE | SE_Change | FGLS_Width | FGLS_Robust_Width | CI_Change |
|---|---|---|---|---|---|---|
| (Intercept) | 3.184437e-02 | 3.250910e-02 | 變大 | 0.124843079 | 0.1274490902 | 變寬 |
| educ | 1.761461e-03 | 1.895323e-03 | 變大 | 0.006905656 | 0.0074304492 | 變寬 |
| exper | 1.298873e-03 | 1.307055e-03 | 變大 | 0.005092118 | 0.0051241957 | 變寬 |
| I(exper^2) | 2.657195e-05 | 2.744395e-05 | 變大 | 0.000104173 | 0.0001075916 | 變寬 |
| female | 9.505454e-03 | 9.445177e-03 | 變小 | 0.037265303 | 0.0370289927 | 變窄 |
| black | 1.696582e-02 | 1.595853e-02 | 變小 | 0.066513034 | 0.0625640260 | 變窄 |
| metro | 1.186360e-02 | 1.155933e-02 | 變小 | 0.046510222 | 0.0453173516 | 變窄 |
| south | 1.354227e-02 | 1.384176e-02 | 變大 | 0.053091297 | 0.0542654167 | 變寬 |
| midwest | 1.404549e-02 | 1.369010e-02 | 變小 | 0.055064111 | 0.0536708228 | 變窄 |
| west | 1.438967e-02 | 1.450663e-02 | 變大 | 0.056413445 | 0.0568719873 | 變寬 |

變數 educ, exper, exper², south, west：SE 變大、CI 變寬，表示 Robust FGLS 在這些變數上，估計更保守。

變數 female, black, metro, midwest：SE 變小、CI 變窄，表示 FGLS 標準誤可能高估變異數（權重模型不夠準確），Robust FGLS 修正了過高估計。

g.

8.18c 確認數據存在異質變異數（NR2=194.4447NR^2 = 194.4447NR2=194.4447），傳統 OLS（8.18d）標準誤不可靠。

Robust FGLS 提供比 Robust OLS 略高的效率（所有變數信賴區間變窄），同時保持穩健性，在效率與可靠性間取得最佳平衡，可以提供精確且可靠的結果。