10.20 R

```r
#10.20(a)
# 計算超額報酬
capm5$Rj_minus_Rf <- capm5$msft - capm5$riskfree
capm5$Rm_minus_Rf <- capm5$mkt - capm5$riskfree

# 執行 OLS 回歸 (CAPM 模型)
capm_model <- lm(Rj_minus_Rf ~ Rm_minus_Rf, data = capm5)

# 顯示回歸結果
summary(capm_model)

#(b)
# 對 Rm_minus_Rf 排序後產生 RANK 變數
capm5$RANK <- rank(capm5$Rm_minus_Rf, ties.method = "first")

#第一階段回歸: 用 RANK 解釋 Rm_minus_Rf
first_stage <- lm(Rm_minus_Rf ~ RANK, data = capm5)
summary(first_stage)

#(c)
# 取得第一階段殘差: v^
capm5$vhat <- resid(first_stage)
# 已計算過 Rj_minus_Rf = msft - riskfree
capm5$Rj_minus_Rf <- capm5$msft - capm5$riskfree

# 增廣回歸: 加入 v^
augmented_model <- lm(Rj_minus_Rf ~ Rm_minus_Rf + vhat, data = capm5)
summary(augmented_model)

#(d)
iv_model <- ivreg(Rj_minus_Rf ~ Rm_minus_Rf | RANK, data = capm5)
ols_model <- lm(Rj_minus_Rf ~ Rm_minus_Rf, data = capm5)

# 顯示 OLS 與 IV 模型的比較
summary(ols_model)
summary(iv_model)

#(e)
capm5$POS <- ifelse(capm5$Rm_minus_Rf > 0, 1, 0)
first_stage_multiple <- lm(Rm_minus_Rf ~ RANK + POS, data = capm5)
summary(first_stage_multiple)
linearHypothesis(first_stage_multiple, c("RANK = 0", "POS = 0"))

#(f)
capm5$vhat <- resid(first_stage_multiple)
augmented_model <- lm(Rj_minus_Rf ~ Rm_minus_Rf + vhat, data = capm5)
summary(augmented_model)

#(g)
# 二階段最小平方法估計 (使用 RANK 和 POS 作為工具變數)
iv_model_2iv <- ivreg(Rj_minus_Rf ~ Rm_minus_Rf | RANK + POS, data = capm5)

# 顯示結果
summary(iv_model_2iv)

#(h)
capm5$iv_resid <- resid(iv_model_2iv)
# 工具變數: RANK 和 POS
sargan_aux_model <- lm(iv_resid ~ RANK + POS, data = capm5)
summary(sargan_aux_model)
# 計算 Sargan 統計量
n <- nrow(capm5)
R2 <- summary(sargan_aux_model)$r.squared
sargan_stat <- n * R2

# p-value
p_value <- 1 - pchisq(sargan_stat, df = 1)

# 顯示結果
cat("Sargan Test Statistic =", round(sargan_stat, 4), "\n")
cat("p-value =", round(p_value, 4), "\n")
```

10,24R

```r
#10.24(a)
# 選擇參與勞動市場且無缺漏值的觀測值
iv_data4 <- subset(mroz, lfp == 1 & !is.na(wage) & !is.na(educ) &
                   !is.na(mothereduc) & !is.na(fathereduc) & !is.na(exper))

# 建立 exper² 變數
iv_data4$exper2 <- iv_data4$exper^2

# 2SLS 模型：使用 MOTHEREDUC 和 FATHEREDUC 為工具變數
iv_model4 <- ivreg(log(wage) ~ educ + exper + exper2 |
                   mothereduc + fathereduc + exper + exper2, data = iv_data4)

#  抽取 IV 模型的殘差
iv_resid <- resid(iv_model4)

ggplot(iv_data4, aes(x = exper, y = iv_resid)) +
  geom_point(color = "black", alpha = 0.6) +
  geom_smooth(method = "loess", color = "blue", se = FALSE) +
  labs(title = "Residuals vs. Experience (2SLS)",
       x = "EXPER", y = "IV residuals") +
  theme_minimal()

#(b)
# 取得殘差平方
iv_data4$iv_resid2 <- iv_resid^2  # iv_resid 是你前面算好的殘差

# 對殘差平方進行迴歸
bp_model <- lm(iv_resid2 ~ exper, data = iv_data4)

# 計算統計量 N * R^2
N <- nobs(bp_model)
R2 <- summary(bp_model)$r.squared
NR2 <- N * R2

# 計算 p-value (檢定 H0: 同變異性)
p_value <- 1 - pchisq(NR2, df = 1)

# 輸出結果
cat("NR² 統計量 =", round(NR2, 4), "\n")

cat("p-value =", round(p_value, 4), "\n")

# 結論
if (p_value < 0.05) {
  cat("結論：拒絕同變異性假設，存在異變異性。\n")
} else {
  cat("結論：無法拒絕同變異性假設，殘差可能為同變異。\n")
}

#(c)
# 使用異質變異穩健標準誤來重新估計模型
robust_summary <- summary(iv_model4, vcov. = vcovHC)
# 查看結果
print(robust_summary)

# 抽出 EDUC 係數與 robust 標準誤
educ_coef <- robust_summary$coefficients["educ", "Estimate"]
educ_se <- robust_summary$coefficients["educ", "Std. Error"]

# 計算 95% 信賴區間
lower_bound <- educ_coef - 1.96 * educ_se
upper_bound <- educ_coef + 1.96 * educ_se

cat("教育變數 EDUC 的 95% 信賴區間為：\n")
cat("[", round(lower_bound, 4), ",", round(upper_bound, 4), "]\n")

# Baseline model summary (未使用 robust SE)
baseline_summary <- summary(iv_model4)

# 比較兩個模型中的 EDUC 標準誤
baseline_se <- baseline_summary$coefficients["educ", "Std. Error"]
robust_se <- robust_summary$coefficients["educ", "Std. Error"]

cat("EDUC baseline SE:", round(baseline_se, 4), "\n")
cat("EDUC robust SE:", round(robust_se, 4), "\n")

#(d)
# 自訂 bootstrap 函數：每次重抽樣都估計一次 IV 模型，回傳 EDUC 係數
boot_iv <- function(data, indices) {
  d <- data[indices, ]  # 重抽樣資料
  model <- ivreg(log(wage) ~ educ + exper + exper2 |
                 mothereduc + fathereduc + exper + exper2, data = d)
  return(coef(model)["educ"])
}

# 執行 bootstrap：B = 200
set.seed(123)  # 為了可重現性
boot_results <- boot(data = iv_data4, statistic = boot_iv, R = 200)

# 查看結果
boot_se <- sd(boot_results$t)
boot_coef <- mean(boot_results$t)

# 95% 信賴區間 (使用正態近似)
boot_ci_lower <- boot_coef - 1.96 * boot_se
boot_ci_upper <- boot_coef + 1.96 * boot_se

cat("Bootstrap 標準誤 (EDUC):", round(boot_se, 4), "\n")
cat("Bootstrap 估計值 (EDUC):", round(boot_coef, 4), "\n")
cat("95% Bootstrap CI for EDUC: [", round(boot_ci_lower, 4), ",", round(boot_ci_upper, 4), "]\n")

baseline_se <- summary(iv_model4)$coefficients["educ", "Std. Error"]
robust_se   <- robust_summary$coefficients["educ", "Std. Error"]

cat("Baseline SE for EDUC:", round(baseline_se, 4), "\n")
cat("Robust SE for EDUC:", round(robust_se, 4), "\n")
cat("Bootstrap SE for EDUC:", round(boot_se, 4), "\n")
```