

202122852 이혜준 AI소프트웨어학과

문제:

p5 데이터 프레임에 대해 비선형회귀분석을 수행하고, SST, SSR, SSE, 결정계수, F-통계량, p-value를 구하시오.
코드 실행결과 캡처해서 붙이세요. pdf 파일로 제출하세요. 숙제 제출기한은 5월 7일 오후 6시까지 입니다.
숙제 미제출시 최종성적은 F 처리합니다.

코드:

```
getwd()
load("p5.RData")
plot(p5)

# 1. 일반 선형회귀
lin_model <- lm(y ~ x, data = p5)
abline(lin_model, lwd = 3, col = "blue")
summary(lin_model)

# 2. 비선형회귀
p5_filtered <- subset(p5, y > 0) # y > 0인 데이터만 사용
log_y <- log(p5_filtered$y)
exp_model <- lm(log_y ~ x, data = p5_filtered)

# 계수 추출 및 곡선 그리기
log_w0_hat <- coef(exp_model)[1]
w1_hat <- coef(exp_model)[2]
w0_hat <- exp(log_w0_hat)

curve(w0_hat * exp(w1_hat * x),
      lwd = 3, col = "darkgreen", add = TRUE)

summary(exp_model)

# y > 0인 데이터만 사용
p5_filtered <- subset(p5, y > 0)

# 로그 변환
log_y <- log(p5_filtered$y)

# 로그 변환된 y에 대해 선형 회귀
exp_model <- lm(log_y ~ x, data = p5_filtered)

# 예측값, 평균
log_y_hat <- fitted(exp_model)
log_y_bar <- mean(log_y)

# SST, SSR, SSE 계산 (모두 로그 스케일에서)
SST <- sum((log_y - log_y_bar)^2)
SSR <- sum((log_y_hat - log_y_bar)^2)
SSE <- sum((log_y - log_y_hat)^2)

# 결정계수
R_squared <- SSR / SST

# 자유도
df_reg <- 1
df_err <- length(log_y) - 2

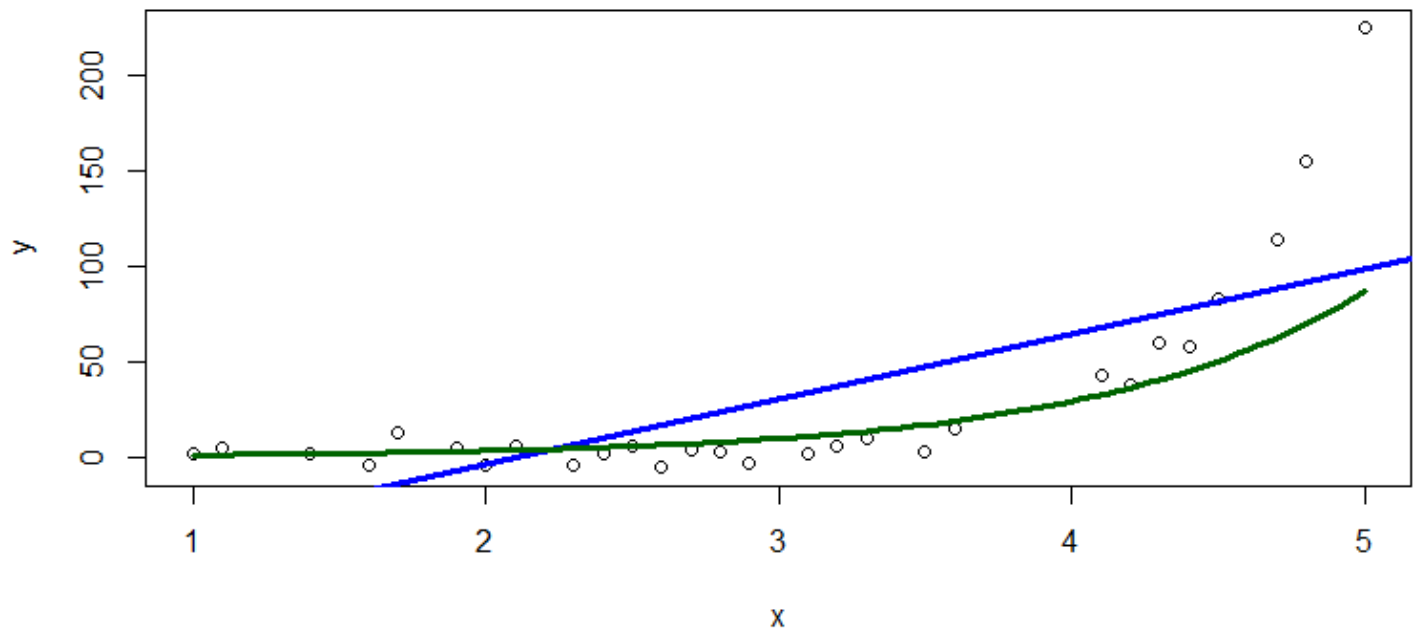
# 평균 제곱
MSR <- SSR / df_reg
MSE <- SSE / df_err

# F-통계량
F_stat <- MSR / MSE

# p-value
p_value <- 1 - pf(F_stat, df_reg, df_err)

# 결과 출력
cat("SST:", SST, "\n")
cat("SSR:", SSR, "\n")
cat("SSE:", SSE, "\n")
cat("R^2:", R_squared, "\n")
cat("F-statistic:", F_stat, "\n")
cat("p-value:", p_value, "\n")
```

실행결과:



비선형으로 변환 후 summary() 결과와 SST, SSR, SSE, 결정계수, F-통계량, p-value 값

```
call:
lm(formula = log_y ~ x, data = p5_filtered)

Residuals:
    Min       1Q   Median       3Q      Max
-1.8905 -0.4507  0.2464  0.5041  1.7101

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -1.0012     0.5453  -1.836   0.0806 .
x              1.0924     0.1619   6.746 1.13e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9383 on 21 degrees of freedom
Multiple R-squared:  0.6843,    Adjusted R-squared:  0.6692
F-statistic: 45.52 on 1 and 21 DF,  p-value: 1.128e-06
```

```
> # 결과 출력
> cat("SST:", SST, "\n")
SST: 58.55877
> cat("SSR:", SSR, "\n")
SSR: 40.07074
> cat("SSE:", SSE, "\n")
SSE: 18.48803
> cat("R^2:", R_squared, "\n")
R^2: 0.6842825
> cat("F-statistic:", F_stat, "\n")
F-statistic: 45.51516
> cat("p-value:", p_value, "\n")
p-value: 1.128065e-06
```