R 入門: OLS

Table of contents

1 Quick example

- Robust standard error を用いて信頼区間を計算する

```
library(tidyverse)
library(lmtest)
library(sandwich)

Data = read_csv("CPS1985.csv")

Fit = lm(log(wage) ~ poly(experience,2) + education,
         Data)

coefci(Fit, vcov = vcovHC, type = "HC3")
```

```
                          2.5 %     97.5 %
(Intercept)           0.6798416  1.1015158
```

```
poly(experience, 2)1  2.2320255  4.2146065
poly(experience, 2)2 -3.0556868 -0.9939588
education             0.0735447  0.1059675
```

2 marginal

- 平均的な education のもとで、モデル上の wage-experience の関係性を可視化

```
library(ggeffects)

Pred = predict_response(
  Fit,
  "experience",
  vcov_fun = "vcovHC",
  vcov_type = "HC3")

plot(Pred)
```

3 発展: pipe 演算子

- 以下は同じ出力をもたらす

```
Data = read_csv("CPS1985.csv")

summary(Data)

read_csv("CPS1985.csv") |>
  summary()
```

- functionA |> functionB: functionA の出力を functionB に入力する
  - Data として保存する必要がない

4 bins plot

- $y \sim x$ の推定結果と bin plot

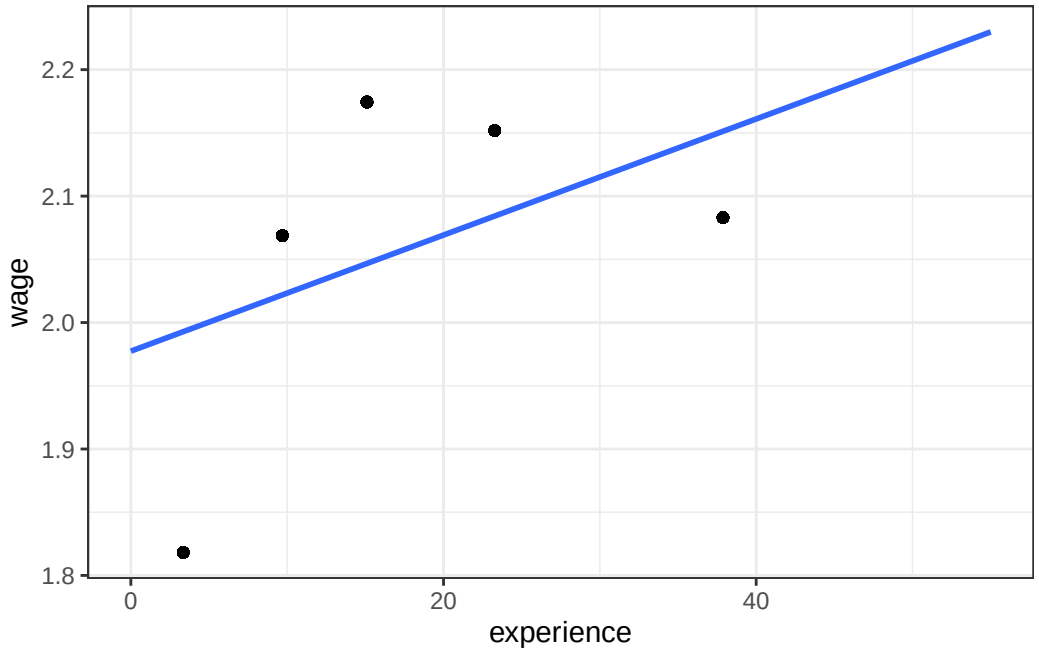
```
Fig = Data |>
  mutate(
    BinExp = ntile(experience,5),
    wage = log(wage)) |>
  mutate(
    AveWage = mean(wage),
    AveExp = mean(experience),
    .by = c(BinExp)
  ) |>
  ggplot(
    aes(
      x = experience,
      y = wage
    )
  ) +
  theme_bw() +
  geom_smooth(
    method = "lm",
    se = FALSE
  ) +
  geom_point(
    aes(
      x = AveExp,
      y = AveWage
    )
```

```
  )
```

5 bins plot

- $y \sim x$ の推定結果と bin plot

```
Fig
```



6 bins plot

- $y \sim poly(x, 2)$ の推定結果と bin plot

```
Fig = Data |>
  mutate(
    BinExp = ntile(experience,5),
    wage = log(wage)) |>
  mutate(
    AveWage = mean(wage),
    AveExp = mean(experience),
    .by = c(BinExp)
  ) |>
  ggplot(
    aes(
```

```
      x = experience,
      y = wage
    )
  ) +
  theme_bw() +
  geom_smooth(
    method = "lm",
    se = FALSE,
    formula = y ~ poly(x,2)
  ) +
  geom_point(
    aes(
      x = AveExp,
      y = AveWage
    )
  )
```

7 bins plot

- $y \sim poly(x, 2)$ の推定結果と bin plot

```
Fig
```