

MODELE Z INTERAKCJAMI I WIELOMIANAMI ZM. OBJAŚNIAJĄCEJ.

ZAJĘCIA NR 9.

Wszystkie ćwiczenia pochodzą z podręczników Wooldridge lub Stock & Watson. Pod tym [linkiem](#) znajduje się dokument z opisem zmiennych zbiorów danych, potrzebnych do rozwiązywania zadań.

I. Modele z wielomianami.

1. Skorzystaj z danych WAGE1 w tym ćwiczeniu. (W, C2 s. 202)

(i) Oszacuj metodą MNK poniższe równanie:

$$\log(wage) = \beta_0 + \beta_1 educ + \beta_2 exper + \beta_3 exper^2 + u$$

Zapisz wynik w formie równania.

- (ii) Używając przybliżenia, znajdź (przybliżony) zwrot z piątego roku doświadczenia. Jaki jest przybliżony zwrot z dwudziestego roku doświadczenia?
- (iii) Przy jakiej wartości *exper*, zwiększenie doświadczenia obniża przewidywany logarytm płacy? Ile osób ma więcej (niż ta wartość) doświadczenia w próbie?

2. Użyj danych BWGHT2 w tym ćwiczeniu (W, C10 p. 223).

(i) Oszacuj równanie MNK:

$$\log(bwght) = \beta_0 + \beta_1 npvis + \beta_2 npvis^2 + u$$

zapisz wyniki w formie równania.

- (ii) Pokaż korzystając z oszacowania z (i), że liczba wizyt prenatalnych która maksymalizuje $\log(bwght)$ wynosi około 22. Jak wiele kobiet miało przynajmniej 22 wizyty w próbie?
- (iii) Czy ma to sens, że waga urodzenia spada po przekroczeniu 22 wizyt prenatalnych? Wyjaśnij.
- (iv) Dodaj wiek matki do równania, używając funkcji kwadratowej. Przy *npvis* ustalonym, czy jakim wieku matki waga przy narodzinach dziecka jest największa? Ile jest w próbie kobiet starszych niż obliczony optymalny wiek?
- (v) Czy wiek matki oraz liczba wizyt prenatalnych wyjaśnia dużo zróżnicowania w $\log(bwght)$?
- (vi) Używając funkcji kwadratowej dla *npvis* oraz *age*, zdecyduj czy logarytm naturalny czy poziom *bwght* jest lepsze dla przewidywania *bwght*.

3. Użyj danych GPA2 do tego zadania (W, C4 p. 221).

(i) Oszacuj model

$$sat = \beta_0 + \beta_1 hsize + \beta_2 hsize^2 + u$$

gdzie *hsize* to rozmiar klasy absolwentów (w setkach), oraz zapisz model w postaci równania.

- (ii) Używając oszacowań z (i) powiedz jaki jest optymalny rozmiar klasy? Uzasadnij.
- (iii) Czy ta analiza jest reprezentatywnie przedstawia poziom osiągnięć akademickich wśród starszych uczniów liceum? Wyjaśnij.
- (iv) Znajdź optymalny poziom klasy, używając $\log(sat)$ jako zmienną zależną. Czy jest różny od tego otrzymanego w (iii)?

II. Modele z interakcjami

1. Użyj danych GPA2 (W, 3 p. 259).

(i) Oszacuj równanie

$$sat = \beta_0 + \beta_1 hsize + \beta_2 hsize^2 + \beta_3 female + \beta_4 black + \beta_5 female \times black + u$$

Zmienna *sat* to łączny wynik testu SAT, *hsize* to wielkość klasy absolwentów (w setkach), *female* to zmienna zero-jedynkowa oraz *black* to zmienna zero-jedynkowa oznaczająca rasę.

Korzystając z równania wyznacz optymalną wielkość klasy.

- (ii) Przy niezmiennym *hsize*, jaka jest szacowana różnica w wyniku SAT między nieczarnymi kobietami i czarnymi mężczyznami?
- (iii) Jaka jest szacowana zmiana w wyniku SAT między czarnymi mężczyznami a pozostałymi mężczyznami?
- (iv) Jaka jest szacowana różnica między czarnymi kobietami a pozostałymi kobietami?

2. Użyj danych TWOYEAR (W, 9 p. 261)

(i) Oszacuj równanie

$$\log(wage) = \beta_0 + \beta_1 female + \beta_2 totcoll + \beta_3 female \times totcoll$$

Używając oszacowań, znajdź wartości *totcoll* takie że przewidywane wartości $\log(wage)$ są takie same dla kobiet i mężczyzn.

- (ii) Korzystając z oszacowania z (i) powiedz, czy kobiety naprawdę mogą uzyskać tyle lat edukacji, aby ich zarobki dogoniły zarobki mężczyzn? Wyjaśnij.

3. Poniższy model pozwala aby zwrot z edukacji zależał od edukacji rodziców (*pareduc*) (W, 4 p. 218):

$$\log(wage) = \beta_0 + \beta_1 educ + \beta_2 educ \times pareduc + \beta_3 exper + \beta_4 tenure + u.$$

(i) Pokaż że w przybliżeniu zwrot z dodatkowego roku edukacji można opisać formułą:

$$\frac{\Delta \log(wage)}{\Delta educ} = \beta_1 + \beta_2 pareduc$$

Jakiego znaku spodziewasz się przy β_2

- (ii) Używając danych WAGE2 oszacuj model. Zinterpretuj współczynnik przy interakcji. Pomocnym może być wybranie dwóch wartości zmiennej *pareduc* np.: *pareduc* = 32 (obaj rodzice mają wykształcenie wyższe), lub *pareduc* = 24 (obaj rodzice mają wykształcenie średnie) i porównaj oszacowany zwrot z edukacji.
- (iii) Dodaj *pareduc* jako oddzielną zmienną i oszacuj model. Czy teraz zależność między zwrotem z edukacji a edukacją rodziców jest pozytywna?