

IMIĘ I NAZWISKO (NR GRUPY):

## KOŁOKWIUM Z EKONOMETRII 2

LUTY 2023

Czas pracy wynosi 90 min. Podpisz kartę z zadaniami oraz kartki z odpowiedziami. W zadaniach przyjmij że  $\alpha = 0.05$  oraz  $t^* = 1.96$ . Jeśli masz pytanie, podnieś rękę. Jeśli coś jest niejasne, pytaj. Maksymalna liczba punktów wynosi 36. Możesz używać kalkulatora. Gdy przeprowadzasz test statystyczny, zawsze podaj hipotezę zerową.

1. Daniel Hammermesh i Amy Parker w 2004 roku analizowali wpływ wyglądu fizycznego i innych cech nauczycieli oraz cech nauczanych przedmiotów na oceny jakości nauczania, wystawiane przez uczniów. Analizowane dane zostały zebrane na Uniwersytecie Texasu w Austin z 463 przedmiotów, i zawierały poniższe zmienne:

- Course\_eval - ocena jakości nauczania danego przedmiotu, w skali od 1 (bardzo niesatysfakcjonująca) do 5 (doskonała).
- Beauty - ocena wyglądu nauczyciela, średnia z ocen 6 uczniów, ma średnią równą zero.
- Female - przyjmuje wartość 1 jeśli nauczyciel jest kobietą, zero wpp.
- Minority - przyjmuje wartość 1 jeśli nauczyciel nie jest rasy białej, 0 wpp.
- NNenglish - przyjmuje wartość 1 jeśli angielski nie jest ojczystym językiem nauczyciela, 0 wpp.
- intro - przyjmuje wartość 1 jeśli dany przedmiot jest nauczany na pierwszym lub drugim roku.
- onecredit - przyjmuje wartość 1 jeśli przedmiot jest nieobowiązkowy (joga, aerobik, taniec etc.), 0 wpp.
- age - wiek nauczyciela

Na podstawie tych danych otrzymano oszacowania, zawarte w Tabeli 1:

**Tabela 1:** Oszacowania modelu objaśniającego oceny przedmiotów.

	Zmienna zależna: course_eval		
	(1)	(2)	(3)
beauty	0.166 (0.0307)	0.160 (0.0320)	0.231 (0.0422)
intro	0.0113 (0.0545)	0.00244 (0.0550)	-0.00123 (0.0545)
onecredit	0.635 (0.111)	0.620 (0.112)	0.657 (0.111)
female	-0.173 (0.0493)	-0.188 (0.0513)	-0.173 (0.0491)
minority	-0.167 (0.0763)	-0.180 (0.0772)	-0.135 (0.0773)
nnenglish	-0.244 (0.107)	-0.243 (0.107)	-0.268 (0.107)
age		0.0195 (0.0231)	
age <sup>2</sup>		-0.000222 (0.000237)	
female × beauty			-0.141 (0.0627)
stała	4.068 (0.0375)	3.677 (0.543)	4.075 (0.0375)
N	463	463	463
R <sup>2</sup>	0.155	0.157	0.164

W nawiasach () podano odchylenia standardowe.

- (a) Korzystając z oszacowań z kolumny (1) wyjaśnij ile wariancji zmiennej *Course\_eval* zostało objaśnione przez ten model. (1p.)
- (b) Korzystając z oszacowań z kolumny (1) oblicz skorygowany  $R^2$ . Wyjaśnij różnicę między zwykłym i skorygowanym  $R^2$ . (2p.)
- (c) Korzystając z oszacowań z kolumny (1), zinterpretuj współczynnik przy zmiennej *Beauty*. (1p.)
- (d) Korzystając z oszacowań z kolumny (1), oceń istotność statystyczną zmiennej *Beauty*. (2p.)
- (e) Korzystając z oszacowań z kolumny (1), zbuduj przedział ufności dla zmiennej *Minority*. Na jego podstawie oceń istotność statystyczną zmiennej *Minority*. (2p.)
- (f) Korzystając z oszacowań z kolumny (2) zweryfikuj, czy zmienne *age* oraz  $age^2$  są łącznie istotne statystycznie. (Jako wartość krytyczną wykorzystaj 0.3859) (2p.)
- (g) Skorzystaj z oszacowań z kolumny (2). Jeśli wiek nauczyciela wzrasta z 40 do 41 przy innych czynnikach niezmiennych, jak zmieni się ocena jakości nauczania? Użyj odpowiedniego przybliżenia. (2p.)
- (h) Korzystając z oszacowań z kolumny (3), zapisz równanie regresji dla kobiet i dla mężczyzn. Dla której grupy wpływ oceny wyglądu fizycznego wpływa silniej na ocenę jakości nauczania? (2p.)
2. Korzystając z danych CPS (*Current Population Survey*), otrzymano poniższe oszacowanie funkcji płac (w nawiasach podano odchylenia standardowe, liczebność próby wynosiła 4733):

$$\widehat{\log(wage)} = 0.41434 + 0.10925 \text{ educ} + 0.03925 \text{ exper} - 0.00066 \text{ exper}^2 - 0.24973 \text{ female} \quad (1)$$

(0.041618) (0.002813) (0.001966) (0.000046) (0.012968)

- wage - płaca na godzinę, w dolarach.
- educ - liczba lat edukacji.
- female - przyjmuje wartość 1 jeśli kobieta, zero wpp.
- exper - doświadczenie, w latach.

Dla tego modelu przeprowadzono zestaw testów diagnostycznych, których wyniki zawarto w Tabeli 2:

**Tabela 2:** Wyniki testów diagnostycznych modelu (1).

Test	Wartość statystyki testowej	p-value
Breuscha-Pagana	81.008	$2.2 \times 10^{-16}$
White'a	80.711	$3.641 \times 10^{-13}$
Ramsey-RESET	38.16	$7.064 \times 10^{-10}$
Jarque-Bera	187.82	$2.2 \times 10^{-16}$

- (a) Zinterpretuj oszacowanie przy zmiennej *female*. (1p.)
- (b) Czy w modelu występuje heteroskedastyczność? Uzasadnij, wykorzystując odpowiedni test. (2p.)
- (c) Czy zastosowana forma funkcyjna jest poprawna? Uzasadnij, wykorzystując odpowiedni test. (2p.)
- (d) Czy składnik losowy ma rozkład normalny w rozważanym modelu? Uzasadnij, wykorzystując odpowiedni test. (2p.)
- (e) Na podstawie przeprowadzonych testów, wskaż założenia Klasycznego Modelu Regresji Linowej które są niespełnione przez model (1). (3p.)
- (f) Jakie są konsekwencje niespełnionych założeń dla interpretacji oraz wnioskowania statystycznego z modelu? (2p.)
- (g) W jaki sposób można rozwiązać problemy zasygnalizowane przez wyniki testów? (2p.)

3. To zadanie jest inspirowane studium płciowej luki płacowej (*gender wage gap*) w korporacjach [Bertrand and Hallock (2001)]. W badaniu są porównywane całkowite zarobki managerów (zmienna *Earnings*) w dużym zbiorze danych korporacji z USA w latach 90-tych. (Każda z tych korporacji musi raportować zarobki dla 5-ciu najważniejszych pracowników).

- (a) Niech *Female* będzie zmienną przyjmującą wartość 1 dla kobiet i 0 dla mężczyzn. Oszacowano regresję logarytmu zarobków i zmiennej *Female*:

$$\log(\widehat{Earnings}) = 6.48 - 0.44Female$$

(0.01) (0.05)

$$R^2 = 0.305$$

- (i) Zinterpretuj oszacowanie przy zmiennej *Female*. (1p.)
- (ii) Czy oszacowanie przy zmiennej *Female* jest statystycznie istotne? Uzasadnij. (2p.)
- (iii) Zinterpretuj  $R^2$ . (1p.)
- (b) Dodano dwie nowe zmienne do regresji: wartość rynkową firmy (*MarketValue* - wartość rynkowa firmy, miliony dolarów) oraz zwrot z akcji (*Return* - stopa zwrotu, miara rezultatów firmy). Nowe oszacowanie:

$$\log(\widehat{Earnings}) = 3.86 - 0.28Female + 0.37\log(MarketValue) + 0.004Return$$

(0.03) (0.04) (0.004) (0.003)

$$n = 46,670, R^2 = 0.345.$$

- (i) Zinterpretuj współczynnik przy  $\log(MarketValue)$ . (1p.)
- (ii) Zinterpretuj współczynnik przy *Return*. (1p.)
- (iii) Czy dodane zmienne (czyli *MarketValue* oraz *Return*) są łącznie statystycznie istotne? Opisz jak byś to przetestował/a. (2p.)