

Zaawansowana Ekonometria

Paweł Struski

Uniwersytet Warszawski

23 lutego 2026

Sprostowanie odnośnie obecności na kolokwium i oceniania

- Ocena z ćwiczeń jest średnią ważoną z trzech elementów (wagi):
 - ▶ short-testy (30%)
 - ▶ case studies (30%)
 - ▶ kolokwium (40%)
- **Każdy** z tych elementów **trzeba zaliczyć** aby zaliczyć ćwiczenia tzn. niezaliczenie któregośkolwiek z nich skutkuje niezaliczeniem ćwiczeń.
- W związku z tym **kolokwium jest obowiązkowe**.
- Termin kolokwium nie jest jeszcze ustalony.

Plan na dzisiaj

- Test t vs test F
- Metodologia konstrukcji modeli "od ogółu do szczegółu"
- Porównywanie i selekcja modeli

Zadanie 1A

Stworzono model, w którym wyjaśniana była cena sprzedaży domu za pomocą jego spodziewanej ceny (określonej przed momentem sprzedaży) - *assess*, powierzchni w metrach kwadratowych - *sqrft*, powierzchni działki na której znajduje się dom - *lotsize*, oraz liczby sypialni - *bdrms*.

$$price = \beta_0 + \beta_1 assess + \beta_2 lotsize + \beta_3 sqrft + \beta_4 bdrms + u$$

W pierwszej kolejności oszacowano model:

$$\hat{price} = -\underset{16.27}{14.47} + \underset{.049}{.976} assess$$

$$n = 88, RSS = 165,644.51, R^2 = .820$$

1A. Przetestuj dwustronne hipotezy $H_0 : \beta_0 = 0$ oraz $H_0 : \beta_1 = 1$.
Skomentuj wyniki.

Zadanie 1B

$$\hat{price} = -14.47 + .976 assess$$

16.27 .049

$$n = 88, RSS = 165,644.51, R^2 = .820$$

1B. Przetestuj hipotezę łączną $H_0 : \beta_0 = 0, \beta_1 = 1$. SSR w modelu z ograniczeniami jest równe 209,448.99 (w jaki sposób się je liczy?).

Zadanie 1C-D

$$price = \beta_0 + \beta_1 assess + \beta_2 lotsize + \beta_3 sqrft + \beta_4 bdrms + u$$

1C. W powyższym modelu przetestuj hipotezę łączną

$H_0 : \beta_2 = \beta_3 = \beta_4 = 0$. Model oszacowano, wykorzystując te same dane i uzyskano $R_2 = 0.829$.

1D. Jeżeli wariancja zmiennej *price* zależy od zmiennych *assess*, *lotsize*, *sqrft*, i *bdrms* to co możemy powiedzieć o teście z podpunktu 1C?

Zadanie 2

Omów w jaki sposób należy przeprowadzić upraszczanie modelu zgodnie z metodą general-to-specific.

```
reg wage IQ educ exper tenure age married black south urban sibs meduc feduc
```

Source	SS	df	MS	Number of obs	=	722
Model	28733750	12	2394479.17	F(12, 709)	=	18.59
Residual	91341044	709	128830.81	Prob > F	=	0.0000
				R-squared	=	0.2393
				Adj R-squared	=	0.2264
Total	120074794	721	166539.243	Root MSE	=	358.93

wage	Coefficient	Std. err.	t	P> t	[95% conf. interval]	
IQ	3.870139	1.155821	3.35	0.001	1.600898	6.139381
educ	41.75525	8.497271	4.91	0.000	25.07243	58.43808
exper	11.00258	4.337995	2.54	0.011	2.485729	19.51944
tenure	3.693257	2.848831	1.30	0.195	-1.899897	9.286412
age	10.89169	5.337833	2.04	0.042	.4118388	21.37154
married	181.1174	42.94556	4.22	0.000	96.80172	265.4331
black	-79.60274	52.46551	-1.52	0.130	-182.6091	23.40361
south	-35.14277	29.9045	-1.18	0.240	-93.85474	23.56919
urban	169.9953	30.4923	5.58	0.000	110.1293	229.8613
sibs	2.177125	6.510146	0.33	0.738	-10.60435	14.9586
meduc	7.214735	6.008851	1.20	0.230	-4.582535	19.012
feduc	8.669905	5.260394	1.65	0.100	-1.657908	18.99772
_cons	-936.3304	194.8673	-4.80	0.000	-1318.916	-553.7443

Zadanie 2 c.d.

- ❶ $H_0 : \beta_{sibs} = 0, H_1 : \text{model ogólny}$
- ❷ $H_0 : \beta_{sibs} = \beta_{south} = 0, H_1 : \text{model ogólny}$
- ❸ $H_0 : \beta_{sibs} = \beta_{south} = \beta_{meduc} = 0, H_1 : \text{model ogólny}$
- ❹ itd. ...

Zadanie 3

Omów konsekwencje:

- 1 Uwzględniania w modelu zmiennej nieistotnej.
- 2 Pominięcia w modelu zmiennej istotnej.

Zadanie 4

6A. Wskaż, który z modeli zaprezentowanych poniżej wybierzesz. Uzasadnij.

Model	AIC	BIC
Model 1	10557.07	10616.63
Model 2	10554.62	10605.02
Model 3	10620.28	10652.36
Model 4	10826.83	10863.69

6B. Wskaż, który z modeli zaprezentowanych poniżej wybierzesz. Uzasadnij.

Model	AIC	BIC
Model 1	10557.07	10605.63
Model 2	10521.61	10655.02

AIC vs BIC

$$\text{AIC} = 2k - 2 \log(\hat{L}) \quad (1)$$

$$\text{BIC} = k \log(n) - 2 \log(\hat{L}) \quad (2)$$

gdzie:

- k – liczba parametrów modelu,
- n – liczba obserwacji,
- \hat{L} – zmaksymalizowana wartość funkcji wiarygodności modelu
- \log – logarytm **naturalny**

Uwaga: czasami podaje się wersję, gdzie AIC i BIC są podzielone przez liczbę obserwacji n . Przy porównywaniu modeli na tych samych danych nie ma to znaczenia.

AIC vs BIC

$$\text{AIC} = 2k - 2 \log(\hat{L}) \quad (1)$$

$$\text{BIC} = k \log(n) - 2 \log(\hat{L}) \quad (2)$$

gdzie:

- k – liczba parametrów modelu,
- n – liczba obserwacji,
- \hat{L} – zmaksymalizowana wartość funkcji wiarygodności modelu
- \log – logarytm **naturalny**

Uwaga: czasami podaje się wersję, gdzie AIC i BIC są podzielone przez liczbę obserwacji n . Przy porównywaniu modeli na tych samych danych nie ma to znaczenia.

Różnica między AIC a BIC:

$\log(n) > 2$ dla każdego $n \geq 2$. To znaczy, że w praktyce BIC penalizuje liczbę parametrów bardziej niż AIC.

Zadanie 5

Oszacowano za pomocą MNK trzy modele na zbiorze 1534 obserwacji. Który z modeli wybierzesz i dlaczego?

$$\hat{pr\acute{a}te} = \underset{.78}{80.29} + \underset{.52}{5.44}mr\acute{a}te + \underset{.045}{.269}age - \underset{.00004}{.00013}totemp \quad (3)$$

$$R^2 = .100, \bar{R}^2 = .098$$

$$\hat{pr\acute{a}te} = \underset{1.95}{97.32} + \underset{.51}{5.02}mr\acute{a}te + \underset{.044}{.314}age - \underset{.28}{2.66}\log(totemp) \quad (4)$$

$$R^2 = .144, \bar{R}^2 = .142$$

$$\hat{pr\acute{a}te} = \underset{.78}{80.62} + \underset{.52}{5.34}mr\acute{a}te + \underset{.045}{.290}age - \underset{.00009}{.00043}totemp + \underset{.00000000.10}{.00000000039}totemp^2 \quad (5)$$

$$R^2 = .108, \bar{R}^2 = .106$$

Podsumowanie

- Dziękuję i do zobaczenia za tydzień!