

Ekonometria dla IiE oraz MSEMAt— Kolokwium

20-01-2023

Zadanie 1 (30%)

Państwa ćwiczeniowiec znalazł na blogu ¹ post o filmach, w których wystąpił Tom Cruise². Zaintrygowany postanowił oszacować model regresji przychodów z filmów mierzonych w milionach dolarów (zmienna `Box office mln USD`) na jakości filmów mierzonych w skali 0-100% w serwisie Rotten Tomatoes³ zmienna `Rotten tomatoes`. Miara ta odzwierciedla odsetek krytyków, którym film się podobał. Hipoteza, którą ćwiczeniowiec chciał zweryfikować brzmiała " H_0 : wyższe środki przeznaczone na produkcję przekładają się na jakość filmu mierzoną przez serwis Rotten Tomatoes".

Call:

```
lm(formula = 'Box office mln USD' ~ 'Rotten tomatoes', data = Tom)
```

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	84.84	94.22	0.900	0.3753
'Rotten tomatoes'	241.18	137.33	1.756	0.0896 .

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Residual standard error: 158.1 on 29 degrees of freedom

Multiple R-squared: 0.09613, Adjusted R-squared: 0.06496

F-statistic: 3.084 on 1 and 29 DF, p-value: 0.08961

RESET test

data: model

RESET = 0.054783, df1 = 2, df2 = 27, p-value = 0.9468

> bptest(model, ~Tom\$'Rotten tomatoes')

studentized Breusch-Pagan test

data: model

BP = 1.6292, df = 1, p-value = 0.2018

Jarque Bera Test

data: model\$residuals

X-squared = 1.7441, df = 2, p-value = 0.4181

Hipotezy weryfikuj na poziomie istotności $\alpha = 10\%$.

- (5%) Zinterpretuj wartość współczynnika determinacji.
- (5%) Zinterpretuj parametr przy zmiennej `Box office mln USD` (niezależnie od jakości modelu).
- (5%) Czy zależność między zmiennymi jest pozorna?
- (10%) Czy model spełnia założenia KMRL? Określić które spełnia/niespełnia.
- (5%) Jak w świetle wyników modelu i testów statystycznych należałoby udzielić odpowiedzi na pytanie o weryfikowaną hipotezę?

¹<https://fivethirtyeight.com/>²<https://fivethirtyeight.com/features/the-four-types-of-tom-cruise-movies/>³<https://www.rottentomatoes.com/>

Zadanie 2 (40%)

Dane do tego zadania pochodzą z podręcznika R. Carter Hill, William E. Griffiths and Guay C. Lim, *Principles of Econometrics*, 3rd edition, John Wiley & Sons.

Dla zbioru 2610 obserwacji o cenach domów w miejscowości Stockton w stanie Kalifornia oszacowanych zostało 6 modeli.

- a) (10%) Przetestować formalnie hipotezę o tym, że wartość parametru przy powierzchni domu równa jest 0.06 w modelu (2).
- b) (5%) Który z przedstawionych modeli należy uznać za najlepszy? Uzasadnij.
- c) (5%) Czy powinniśmy usunąć z modelu (2) kolor oczu sprzedającego i płeć jednocześnie? Uzasadnij.
- d) (10%) Zweryfikuj hipotezę o łącznej nieistotności płci i liczby łazienek w modelu (2).
- e) (10%) Zweryfikować formalnie hipotezę o łącznej nieistotności zmiennych w modelu (1)

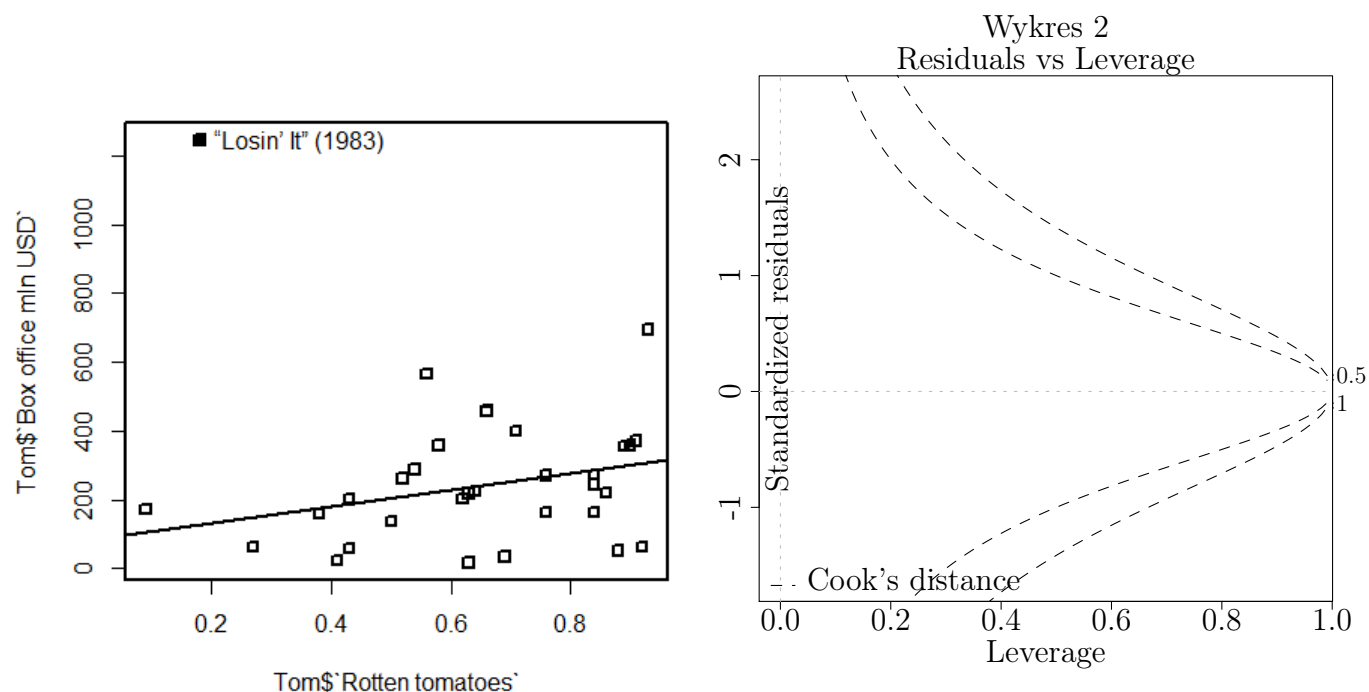
	1	2	3	4	5	6
	b/se	b/se	b/se	b/se	b/se	b/se
living area, 100 sq feet		0.057*** (0.00)	0.057*** (0.00)	0.057*** (0.00)	0.058*** (0.00)	0.058*** (0.00)
number of beds		-0.052*** (0.01)	-0.052*** (0.01)	-0.052*** (0.01)	-0.048*** (0.01)	-0.048*** (0.01)
number of baths		0.022 (0.01)	0.022 (0.01)	0.022 (0.01)		
=1 if lot size> 0.5 acres		0.209*** (0.02)	0.209*** (0.02)	0.209*** (0.02)	0.208*** (0.02)	0.208*** (0.02)
age of home, years		-0.003*** (0.00)	-0.003*** (0.00)	-0.003*** (0.00)	-0.004*** (0.00)	-0.004*** (0.00)
=1 if home has pool		0.071** (0.02)	0.071** (0.02)	0.071** (0.02)	0.071** (0.02)	0.071** (0.02)
brown_eyes		-0.006 (0.01)	-0.006 (0.01)		-0.006 (0.01)	
female		0.001 (0.01)				
Constant	11.602*** (0.01)	10.892*** (0.03)	10.892*** (0.03)	10.890*** (0.03)	10.908*** (0.03)	10.906*** (0.03)
r2	0.000	0.698	0.698	0.698	0.698	0.698
r2_a	0.000	0.697	0.697	0.697	0.697	0.697
N	2610.000	2610.000	2610.000	2610.000	2610.000	2610.000
F	0.000	751.233	858.874	1002.217	1000.844	1201.214
rss	501.264	151.412	151.413	151.433	151.578	151.600

* p<0.05, ** p<0.01, *** p<0.001

Zadanie 3 (30%)

Zadanie to jest dalszym ciągiem zadania 1.

Państwa ćwiczeniowiec postanowił sporządzić wykres zależności między zmiennymi **Box office** mln USD oraz **Rotten tomatoes** (lewy panel poniżej). Na tym wykresie został zaznaczony film "Losin' it", którego ćwiczeniowiec nie uwzględnił w modelu przedstawionym w zadaniu 1.



Odpowiedz i uzasadnij.

1. (10%) Zaznacz na wykresie (prawy panel) pozycję gdzie najprawdopodobniej powinna znaleźć się obserwacja dla filmu "Losin' it". Uzasadnij.
2. (10%) Czy ćwiczeniowiec postąpił właściwie nieuwzględniając filmu "Losin' it"? Uzasadnij.
3. (10%) Czy zmienna **Rotten tomatoes** jest zmienną egzogeniczną? Uzasadnij oraz odpowiedz jakie ma to konsekwencje dla modelu.

