
Egzamin z ekonometrii - wersja ogólna

06-02-2019

Regulamin egzaminu

1. Egzamin trwa 90 min.
2. Rozwiązywanie zadań należy rozpocząć po ogłoszeniu początku egzaminu a skończyć wraz z ogłoszeniem końca egzaminu. Złamanie tej zasady skutkuje usunięciem z sali i unieważnieniem pracy.
3. Każdy ze zdających musi podpisać się na regulaminie i liście obecności. Praca osoby, która nie złożyła tych podpisów, nie zostanie oceniona.
4. Rozwiązanie każdego zadania należy zapisać na kartce z tymże zadaniem lub na dodatkowych kartkach uzyskanych od prowadzących egzamin.
5. Każda kartka z rozwiązaniem musi być opatrzona imieniem, nazwiskiem i numerem zadania.
6. W razie braku podpisu lub numeru zadania na kartce, kartka nie zostanie oceniona. Nie będą też oceniane rozwiązania wpisane na kartkach innych niż te rozdawane przez prowadzących.
7. Na jednej kartce może znajdować się rozwiązanie tylko jednego zadania. Oceniane jest rozwiązanie tylko tego zadania, którego numer widnieje na kartce.
8. W trakcie egzaminu:
 - (a) wolno używać jedynie długopisu i kalkulatora,
 - (b) posiadanie jakichkolwiek innych "pomocy naukowych" jest traktowane jak ściąganie,
 - (c) rozmowy z innymi zdającymi są traktowane jak ściąganie.
9. Wszystkie pytania należy kierować bezpośrednio do jednego z pilnujących.
10. Minimalnym progiem zaliczenia jest uzyskanie 50% z pytań teoretycznych i 50% z zadań.

Regulamin zrozumiałam(em)

Warszawa, 06-02-2019 r.,

Podpis

Brudnopis

Pytania teoretyczne

1

1. Wymienić założenia Klasycznego Modelu Regresji Liniowej (KMRL) (1 pkt)

Pytania teoretyczne c.d.

2

2. Dlaczego zmienną dyskretną rozkodowywujemy na zmienne zerojedynkowe? (1 pkt)

Pytania teoretyczne c.d.

3

3. Wyprowadzić estymator MNK dla modelu z wieloma zmiennymi objaśniającymi (1 pkt)

Pytania teoretyczne c.d.

4

4. Jaki skutek może mieć pominięcie istotnej zmiennej w modelu? (1 pkt)

Zadanie 1

1	2	Σ

1. Mamy następujący model ze stałą i jedną zmienną objaśniającą: $y_i = \beta_1 + \beta_2 x_i + \varepsilon_i$, $\varepsilon \sim N(0, \sigma^2 I)$, gdzie X jest nielosowe. Wyznaczyć $cov(b_1, b_2)$, gdzie b_1 i b_2 są estymatorami parametrów uzyskanymi Metodą Najmniejszych Kwadratów. (3 pkt)

(Wskazówka: Należy wyznaczyć macierz wariancji – kowariancji dla wektora oszacowanych parametrów, czyli: $var(b) = \sigma^2(X'X)^{-1}$).

2. Oszacowano regresję $y_i = \beta x_i + \varepsilon_i$ i uzyskano estymator b . Utworzono nowe zmienne: $x_i^* = x_i a$ oraz $y_i^* = c y_i$, gdzie $c, a \in \mathbb{R}$, $c \neq 0$, $a \neq 0$. Estymowano regresję: $y_i^* = \gamma x_i^* + \varepsilon_i$ i otrzymano estymator $\hat{\gamma}$.

Jaki warunek muszą spełniać stałe c i a aby $\hat{\gamma} = b$? (3 pkt)

Zadanie 1

Zadanie 2

1	2	3	4a	4b	4c	Σ

Oszacowano regresję logarytmu ilości sprzedawanych nowych samochodów (zmienna ln_samoch) na logarytmie PKB per capita (zmienna ln_pkb) i logarytmie ceny nowych samochodów (zmienna ln_cena). Regresję oszacowano na danych kwartalnych dla Stanów Zjednoczonych z okresu 1976 – 1990. Wprowadzono także do modelu zmienne sezonowe (zmienne: $dq2, dq3, dq4$) przyjmujące wartość 1 dla odpowiedniego kwartału, a 0 dla pozostałych (poziom bazowy to I kwartał). Oszacowania parametrów znajdują się poniżej. Hipotezy testować na poziomie istotności 0,05. Odpowiedzi uzasadnić podając p -value.

Source	SS	df	MS	Number of obs =	60
Model	.910843346	5	.182168669	F(5, 54) =	14.07
Residual	.699371266	54	.01295132	Prob > F	= 0.0000
Total	1.61021461	59	.027291773	R-squared	= 0.5657
				Adj R-squared	= 0.5254
				Root MSE	= .1138

ln_samoch	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
ln_pkb	2.531049	.4403752	5.75	0.000	1.64815 3.413948
ln_cena	-1.314402	.1822032	-7.21	0.000	-1.679697 -.9491063
dq2	.1119025	.0415588	2.69	0.009	.0285822 .1952228
dq3	-.0019095	.0416029	-0.05	0.964	-.0853183 .0814994
dq4	-.0354907	.0416428	-0.85	0.398	-.1189795 .047998
_cons	7.576466	.4808595	15.76	0.000	6.6124 8.540531

Ramsey RESET test statistic: F(3, 51) = 3.45 p-value = 0.0231
 Jarque-Bera test statistic: chi2(2) = 0.12 p-value = 0.9438
 White's general test statistic: chi2(14) = 10.99 p-value = 0.6865
 Breusch-Godfrey test statistic: chi(4) = 24.42 p-value = 0.0001
 Durbin-Watson (6,60) = 0.97

Wartości krytyczne d_l i d_u dla testu DW przy 60 obserwacjach i 5 zmiennych i stałej na poziomie istotności 0,05 wynoszą odpowiednio 1.41 i 1.77.

1. Zbadać, czy w modelu występuje heteroscedastyczność. (1 pkt)
2. Zbadać, czy w modelu występuje autokorelacja. (1 pkt)
3. Sprawdzić, czy forma funkcyjna modelu jest prawidłowa.(1 pkt)
4. Jeżeli model nie spełnia założeń KMRL określić:
 - (a) Które założenia nie są spełnione? (1 pkt)
 - (b) Jakie to ma konsekwencje dla interpretacji modelu i wnioskowania statystycznego? (1 pkt)
 - (c) W jaki sposób można rozwiązać problemy zasygnalizowane przez wyniki testów? (1 pkt)

Zadanie 2

Zadanie 2

Zadanie 2

Zadanie 3

1	2	3	4	5a	5b	6	Σ

Na podstawie danych dla Stanów Zjednoczonych z roku 1988 dotyczących kobiet oszacowano wysokość wynagrodzenia (\ln_placa - logarytm wynagrodzenia). Zmiennymi objaśniającymi są wiek ($wiek$), wiek do kwadratu ($wiek_2$), rasa ($rasa$: 0 - inny, 1 - biały), płeć ($plec$: 0 - mężczyzna, 1 - kobieta), wykształcenie średnie ($srednie$: 1 – osoba ma PRZYNAJMNIEJ wykształcenie średnie (czyli średnie lub wyższe), 0 – w pozostałych przypadkach), wykształcenie wyższe ($wyzsze$: 1 – osoba ma wykształcenie wyższe, 0 – w pozostałych przypadkach), $srednie_rasa$ – interakcja między zmienną $srednie$ a $rasa$; $wyzsze_rasa$ – interakcja między zmienną $wyzsze$ a $rasa$. Oszacowania parametrów znajdują się poniżej. Hipotezy testować na poziomie istotności 0,05. Odpowiedzi uzasadnić podając p -value.

Source	SS	df	MS	Number of obs = 2258		
Model	108.54612	7	15.5065886	F(7, 2250)	=	50.15
Residual	695.775585	2250	.309233593	Prob > F	=	0.0000
-----				R-squared	=	0.1350
-----				Adj R-squared	=	0.1323
Total	804.321705	2257	.356367614	Root MSE	=	0.5561

\ln_placa	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
$wiek$.0711354	.1045639	0.68	0.496	-.1339163	.2761871
$wiek_2$	-.0009436	.0013242	-0.71	0.476	-.0035404	.0016532
$srednie$.3953436	.0931082	4.25	0.000	.2127567	.5779306
$wyzsze$.5662218	.0553494	10.23	0.000	.4576807	.674763
$rasa$.227602	.1316609	1.73	0.084	-.0305875	.4857914
$srednie_rasa$	-.0705119	.135399	-0.52	0.603	-.3360318	.1950081
$wyzsze_rasa$	-.1910754	.0631099	-3.03	0.002	-.3148351	-.0673158
$_cons$	-.0599334	2.055809	-0.03	0.977	-4.091414	3.971547

1. Dlaczego wprowadzono do modelu interakcje między wykształceniem a rasa? Odpowiedź należy dokładnie uzasadnić. (1 pkt)
2. W jakim celu umieszczono w modelu zmienną $wiek$ podniesioną do kwadratu? W jaki sposób należałoby przetestować hipotezę, że wiek nie wpływa na poziom dochodu? Odpowiedź należy dokładnie uzasadnić. (1 pkt)
3. Postanowiono wprowadzić do wyżej zaprezentowanej regresji zmienną $wyksztalcenie$, która oznacza liczbę lat nauki w latach. Jaki problem może wystąpić w „nowej” regresji (w modelu z dołączoną zmienną $wyksztalcenie$)? Odpowiedź należy dokładnie uzasadnić. (1 pkt)
4. Zdefiniowano nową zmienną $srednie_1$, która przyjmuje wartość równą 1, jeżeli najwyższym ukończonym poziomem wykształcenia jest wykształcenie średnie oraz 0 w pozostałych przypadkach. Jeżeli zdecydujemy się umieścić tę zmienną w modelu, to jaki problem ekonometryczny wystąpi? Odpowiedź należy dokładnie uzasadnić. (1 pkt)
5. Dokonać interpretacji parametrów przy następujących zmiennych:
 - (a) $wyzsze_rasa$ – interakcja między zmienną $wyzsze$ a $rasa$, (1 pkt)
 - (b) wykształcenie średnie ($srednie$: 1 – osoba ma PRZYNAJMNIEJ wykształcenie średnie (czyli średnie lub wyższe), 0 – w pozostałych przypadkach), (1 pkt)

6. Należy przetestować hipotezę, że kobiety z wykształceniem wyższym rasy białej i rasy innej niż biała zarabiają tyle samo. Zapisać hipotezę zerową za pomocą macierzy: $H\beta = h$ oraz wyjaśnić jak można ją testować (podać wzór na statystykę testową, podjęcie decyzji). (1 pkt)

Zadanie 3

Zadanie 3