

imię, nazwisko, nr indeksu: _____

Ekonometria — egzamin
07/03/2018

1. Egzamin trwa 90 minut.
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. W trakcie egzaminu wolno używać jedynie długopisu o innym kolorze atramentu niż czerwony oraz kalkulatora.
4. Przed przystąpieniem do pisania egzaminu należy podpisać **wszystkie kartki** pracy (na dole w przewidzianym miejscu). Złożenie podpisu pod regulaminem oznacza jego akceptację. Do egzaminu mogą przystąpić osoby, które akceptują regulamin.
5. 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.
6. Rozwiązanie każdego zadania należy zapisać na kartce z tymże zadaniem, ewentualnie na czystych kartkach znajdujących się na końcu arkusza egzaminacyjnego lub na dodatkowych kartkach uzyskanych od prowadzących egzamin.
7. Na jednej kartce może znajdować się rozwiązanie tylko jednego zadania. Oceniane jest rozwiązanie tylko tego zadania, którego numer widnieje na kartce.
8. Egzamin składa się z czterech pytań teoretycznych i 3 zadań.
9. Posiadanie przy sobie wszelkich materiałów drukowanych (w tym książek) oraz innych np. wykonanych własnoręcznie materiałów zostanie uznane za ściąganie.
10. Rozmowy z innymi zdającymi będą traktowane identycznie jak ściąganie.
11. Każda zauważona próba ściągania skutkuje usunięciem z egzaminu.
12. Wszystkie pytania należy kierować bezpośrednio do osób pilnujących.
13. Warunkiem uzyskania oceny pozytywnej jest zdobycie conajmniej 50 % punktów z części teoretycznej egzaminu oraz min. 40 % punktów z części zadaniowej.

Warszawa, 07/03/2018, _____

podpis

Powodzenia :-)

Wartości krytyczne testu Durбина - Watsona.

n	d_L	d_U
50	1,50	1,59
100	1,65	1,69
200	1,76	1,78

Wartości krytyczne rozkładu χ^2 .

liczba stopni swobody	$\alpha = 0,95$	$\alpha = 0,99$
1	3,84	6,64
2	5,99	9,21
3	7,81	11,35
4	9,49	13,28
5	11,07	15,09
6	12,59	16,81
7	14,07	18,48
8	15,51	20,09
9	16,91	21,67
10	18,31	23,21

Wartości krytyczne rozkładu F .

liczba stopni swobody	$\alpha = 0,95$
F(1,20)	4,35
F(2,20)	3,49
F(3,20)	3,10
F(20,1)	248,02
F(20,2)	19,45
F(20,3)	8,66
F(1,500)	3,86
F(2,500)	3,01
F(3,500)	2,62
F(1, ∞)	3,84
F(5, ∞)	2,21

1	2	Σ

Pytanie 1.

Pokaż, że w modelu ze stałą $TSS = ESS + RSS$.

UWAGA: Część osób po prawidłowym rozpisaniu sumy kwadratów zapomniała napisać dlaczego określone wyrazy są równe zero.

Pytanie 2.

Na czym polega współliniowość. Jakie są jej konsekwencje dla oszacowań parametrów modeli?

UWAGA: Oczekiwałem, że rozpoczną Państwo odpowiedź od zdania "Współliniowość to korelacja między zmiennymi objaśniającymi modelu", po czym rozpatrzą przypadek dokładnej i niedokładnej współliniowości. Nie uznawałem odpowiedzi: współliniowość to korelacja między zmiennymi.

Pytanie 3.

Zapisz założenia niezbędne do uzyskania zgodności estymatora MNK, pokaż, że przy tych założeniach estymator MNK jest zgodny.

UWAGA: Częstkowe punkty są za wypisanie założeń.

Pytanie 4.

Dlaczego statystyki R^2 nie można wykorzystywać do porównań dopasowania modeli.

UWAGA: Argumentów można wymienić kilka, jeżeli odpowiadający wymienił prawidłowe dwa otrzymał 100% punktów.

1	2	3	Σ

ZADANIE 1

Badacz dysponuje danymi dotyczącymi dochodów gospodarstw domowych (*doch*) oraz wydatków na żywność gospodarstw (*wyd*) zlokalizowanych w mieście i na wsi

Lokalizacja	Liczba obserwacji	Odchylenie standardowe
Razem	100	2,2
Miejskie	40	2,0
Wiejskie	60	1,5

Oszacowano parametry modelu dla całej próby

$$wyd_i = 3,23 + 0,69doch_i \quad R^2 = 0,70$$

następnie dla gospodarstw miejskich

$$wyd_i = 4,15 + 0,57doch_i \quad R^2 = 0,90$$

oraz gospodarstw wiejskich

$$wyd_i = 2,96 + 0,82doch_i \quad R^2 = 0,95$$

1. Jakie będą wartości oszacowanych współczynników w modelu

$$wyd_i = \beta_1 + \beta_2 D_i + \beta_3 doch_i + \beta_4 doch_i D_i + \varepsilon_i$$

gdzie D_i jest równe 1 dla gospodarstw miejskich i 0 dla wiejskich.

2. Oblicz sumę kwadratów reszt i wartość statystyki R^2 dla modelu z punktu 1.
3. Oblicz wartość statystyki testu weryfikującego hipotezę o identycznych wartościach współczynników regresji dla gospodarstw miejskich i wiejskich (test Chowa). Zinterpretuj uzyskany wynik.

1	2	3	Σ

ZADANIE 2

Badacz szacował parametry KMRL wykorzystując MNK. Przed uzyskaniem oszacowań przypadkowo pomnożył jeden z wierszy macierzy obserwacji X przez stałą c , różną od zera. Niech i_n będzie wektorem wybierającym n -tą obserwację. Wówczas macierz obserwacji można zapisać jako

$$X + i'_n X$$

Zatem badacz wykorzystał estymator $\tilde{b} = ((X + i'_n X)'(X + i'_n X))^{-1} X' y$

1. Pokaż, że estymator \tilde{b} jest obciążony.
2. Pokaż, że estymator \tilde{b} jest asymptotycznie nieobciążony.
3. Oblicz obciążenie estymatora macierzy wariancji-kowariancji.

1	2	3	4	5	Σ

ZADANIE 3 Na podstawie danych pochodzących z badania struktury wynagrodzeń według zawodów badacze postanowili sprawdzić czy płace przedstawicieli władz publicznych są ustalane zgodnie z teorią kapitału ludzkiego. W tym celu oszacowano równanie płacy typu Mincera. W tym równaniu logarytm miesięcznego wynagrodzenia w złotych (l_wyn) tłumaczono stażem pracy pracownika i jego kwadratem oraz poziomem wykształcenia (zmienna wskazująca *wyzsze* równa 1 oznacza pracownika o wykształceniu wyższym) i zmienną wskazującą czy pracuje w wymiarze pełnego etatu ($pelny=1$). Uzyskano następujące wyniki:

Source	SS	df	MS	Number of obs	=	335
Model	1.78074212	4	.445185529	F(4, 330)	=	10.69
Residual	13.7402531	330	.04163713	Prob > F	=	0.0000
				R-squared	=	0.1147
				Adj R-squared	=	0.1040
Total	15.5209952	334	.046470045	Root MSE	=	.20405

l_wyn	Coef.	Std. Err.	t	P> t
staz	0,0187	0,0043	4,33	0,000
staz2	-0,0003	0,0001	-3,31	0,001
wyzsze	0,0530	0,0389	1,36	0,174
pelny	0,1498	0,0825	1,82	0,070
_cons	8,8068	0,0922	95,56	0,000

Jarque- Bera	chi2(2)	= 18,208	P=0,001
RESET	F(5, 326)	= 1,69	P=0,136
Breusch-Pagan, wartosc dopasowana	chi2(1)	= 19,70	P=0,000
Breusch-Pagan, potegi zmiennych	chi2(4)	= 43,58	P=0,000
Mean VIF		= 9,40	

gdzie P oznacza wartość P.

Odpowiedz na poniższe pytania, uzasadniając swoje odpowiedzi stosownymi obliczeniami lub wartościami statystyk. Przyjmij poziom istotności 5%.

1. Udziel odpowiedzi na pytanie: Czy płace przedstawicieli władz publicznych są ustalane zgodnie z teorią kapitału ludzkiego?
2. Zinterpretuj wpływ długości stażu pracy pracownika na wysokość uzyskiwanego przez niego wynagrodzenia, jeżeli przeciętny staż pracownika wynosi 28,6 lat.
3. Czy forme funkcyjną modelu można uznać za poprawną?
4. Czy spełnione jest założenie o (hiper)sferyczności składnika losowego?
5. Jakie konsekwencje niesie niespełnienie założeń KMRL w tym modelu i w jaki sposób można rozwiązać ten problem.

Rozwiązania zadań

Zadanie 1

1.

$$\begin{aligned} \text{wyd}_i &= 2,96 + (4,15 - 2,96)D_i + 0,82\text{doch}_i + (0,57 - 0,82)\text{doch}_i D_i + \varepsilon_i \\ \text{wyd}_i &= 2,96 + 1,19D_i + 0,82\text{doch}_i - 0,25\text{doch}_i D_i + \varepsilon_i \end{aligned}$$

2.

$$\begin{aligned} TSS &= n \times SD^2; \quad TSS_m = 40 \times (2,0)^2 = 160; \quad RSS_m = 0,1TSS_m = 16. \\ TSS_w &= 60 \times (1,5)^2 = 135; \quad RSS_m = 0,05TSS_w = 6,75. \\ TSS_{m+w} &= 160 + 135 = 295; \quad RSS_{m+w} = 16 + 6,75 = 22,75. \\ R^2 &= 1 - \frac{16,75}{295} \approx 0,923. \end{aligned}$$

3.

$$\begin{aligned} F &= \frac{((RSS - RSS_{m+w})/J)}{RSS_{m+w}/(n - 2k)} \\ RSS &= 0,3TSS = 0,3(100 \times (2,2)^2) = 0,3 \times 484 = 145,2. \\ F &= \frac{((145,2 - 22,75)/2)}{22,75/(100 - 4)} = \frac{61,225}{0,237} \approx 258 > F^*(2, 96) \approx 3,01. \end{aligned}$$

Zatem współczynniki równania regresji są różne dla gospodarstw miejskich i wiejskich.

Zadanie 2

1.

$$\begin{aligned} \mathbb{E}(\tilde{b}) &= \mathbb{E} \left((X + i'_n X)'(X + i'_n X) \right)^{-1} X' y = \left((X + i'_n X)'(X + i'_n X) \right)^{-1} X' \mathbb{E}(X\beta + \varepsilon) \\ &= \left((X + i'_n X)'(X + i'_n X) \right)^{-1} X' (X\beta \neq \beta). \end{aligned}$$

2.

$$\begin{aligned} \lim_{n \rightarrow \infty} \left[\frac{1}{n} \left((X + i'_n X)'(X + i'_n X) \right) \right]^{-1} \left[\frac{X'X}{n} \right] &= \\ \lim_{n \rightarrow \infty} \left[\frac{X'X}{n} + \frac{X'i'_n X}{n \rightarrow 0} + \frac{X'i_n X}{n \rightarrow 0} + \frac{X'i_n i'_n X}{n \rightarrow 0} \right]^{-1} \left[\frac{X'X}{n} \right] &= \mathbb{E}(X'X)^{-1} X'X. \end{aligned}$$

3.

$$\begin{aligned} \text{var}(\tilde{b}) - \text{var}(b) &= \sigma^2 \left[(X + i'_n X)'(X + i'_n X) \right]^{-1} - \sigma^2 (X'X)^{-1} = \\ &= \sigma^2 \left[X'X + X'i'_n X + X'i_n X + X'i_n i'_n X - X'X \right]^{-1} = \\ &= \sigma^2 \left[X'i'_n X + X'i_n X + X'i_n i'_n X \right]^{-1} = \end{aligned}$$

Zatem estymator wariancji jest dodatnio obciążony.

Zadanie 3

1. Wartość statystyki t-Studenta dla zmiennej *wyższe* jest mniejsza od 1,96 zatem, ocena parametru w sensie statystycznym nie różni się od zera. Wobec tego należy stwierdzić, że wysokość płac władz publicznych nie jest ustanawiana zgodnie z teorią kapitału ludzkiego.

2. W tym celu należy obliczyć wartość pochodną cząstkowej względem stażu

$$\frac{\partial \ln \text{wyn}}{\partial \text{staz}} = \beta_{\text{staz}} + 2\beta_{\text{staz}^2} \bar{\text{staz}} = 0,0187 + 2 \times -0,0003 \times 28,6 = 0,0187 - 0,0172 = 0,0015$$

Dla osoby o przeciętnym w próbie wynagrodzeniu i stażu pracy, kolejny rok stażu pracy przyczynia się do wzrostu oczekiwanego wynagrodzenia o 0,15% przy innych czynnikach stałych.

3. Wynik testu RESET: $F(5, 326) = 1,69$; $P = 0,14 < \alpha$; wskazuje, że liniowa forma funkcyjna może zostać uznana za poprawną.
4. Testami pomagającymi weryfikować hipersferyczność składnika losowego są testy Breusch-Pagana. Wyniki testów sugerują konieczność odrzucenia założenia o stałej wariancji $\chi^2(1) = 19,70$; $P = 0,00 < \alpha$; $\chi^2(4) = 43,58$; $P = 0,00 < \alpha$. Zatem składnik losowy modelu nie jest hipersferą.

UWAGA: Test normalności rozkładu składnika losowego nie jest testem sferyczności.

5. Konsekwencją heteroscedastyczności są nieprawidłowe wartości obliczonych w standardowy sposób statystyk t oraz F . Rozwiązaniem problemu jest wykorzystanie odpornej macierzy White'a w celu oszacowania wariancji estymatora wektora parametrów.