



KURS EKONOMETRIA

LEKCJA 2

Dobór zmiennych objaśniających.
Metoda Hellwiga.

ZADANIE DOMOWE



Część 1: TEST

Zaznacz poprawną odpowiedź (tylko jedna jest prawdziwa).

Pytanie 1

Która z własności określa zmienne objaśniające w modelach ekonometrycznych?

- a) Odpowiednio mała zmienność.
- b) Słaba korelacja ze zmienną objaśnianą.
- c) Słaba korelacja pomiędzy zmiennymi objaśniającymi.
- d) Umiarkowana korelacja między zmiennymi endogenicznymi.

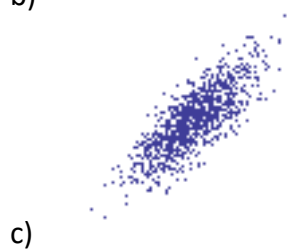
Pytanie 2

Jakie zmienne uznaje się za qasi-stałe?

- a) Nie wnoszące istotnych informacji do badań, czyli takie, które dla obranej wcześniej wartości krytycznej współczynnika zmienności V^* spełniają nierówność: $|V_j| \leq V^*$.
- b) Nie wnoszące istotnych informacji do badań, czyli takie, które dla obranej wcześniej wartości krytycznej współczynnika zmienności V^* spełniają nierówność: $|V_j| \geq V^*$.
- c) Wnoszące istotne informacje do badań, czyli takie, które dla obranej wcześniej wartości krytycznej współczynnika zmienności V^* spełniają nierówność: $|V_j| \leq V^*$.
- d) Wnoszące istotne informacje do badań, czyli takie, które dla obranej wcześniej wartości krytycznej współczynnika zmienności V^* spełniają nierówność: $|V_j| \geq V^*$.

Pytanie 3

Na którym z wykresów przedstawiono dane (X, Y) z odpowiadającą im niezerową wartością współczynnika korelacji liniowej Pearsona?



Pytanie 4

Które z własności charakteryzują współczynnik korelacji liniowej Pearsona?

- a) $r_{X,Y} \in (-1,1)$; $r_{X,Y} = r_{Y,X}$; $r_{X,X} = -1$
- b) $r_{X,Y} \in [0,1]$; $r_{X,Y} = r_{X,X}$; $r_{X,X} = 1$
- c) $r_{X,Y} \in [-1,1]$; $r_{X,Z} = r_{Y,X}$; $r_{Y,Y} = 1$
- d) $r_{X,Y} \in [-1,1]$; $r_{X,Y} = r_{Y,X}$; $r_{X,X} = 1$

Pytanie 5

Które z poniższych macierzy wskazują poprawne wyznaczenie wektora R_0 oraz macierzy R dla następującej macierzy korelacji r pomiędzy zmiennymi X, Y, Z , gdzie zmienną objaśnianą jest zmienna Y ?

$$r = \begin{bmatrix} 1 & 0,65 & -0,37 \\ 0,65 & 1 & 0,22 \\ -0,37 & 0,22 & 1 \end{bmatrix}$$

$$a) \quad R_0 = \begin{bmatrix} 0,22 \\ 0,65 \end{bmatrix}, \quad R = \begin{bmatrix} -0,37 & 1 \\ 1 & -0,37 \end{bmatrix}$$

$$b) \quad R_0 = \begin{bmatrix} 0,65 \\ 0,22 \end{bmatrix}, \quad R = \begin{bmatrix} 1 & -0,37 \\ -0,37 & 1 \end{bmatrix}$$

$$c) \quad R_0 = \begin{bmatrix} 0,65 \\ -0,37 \end{bmatrix}, \quad R = \begin{bmatrix} 1 & 0,22 \\ 0,22 & 1 \end{bmatrix}$$

$$d) \quad R_0 = \begin{bmatrix} 0,65 \\ 1 \end{bmatrix}, \quad R = \begin{bmatrix} 0,22 & -0,37 \\ -0,37 & 0,22 \end{bmatrix}$$

Pytanie 6

Który z poniższych zbiorów reprezentują zbiór wszystkich kombinacji zmiennych objaśniających ze zbioru $X = \{X, W, Z\}$.

$$a) \quad C_1 = \{X\}, C_2 = \{W\}, C_3 = \{Z\}, C_4 = \{X, X\}, C_5 = \{W, W\}, C_6 = \{Z, Z\}, \\ C_7 = \{X, W, Z\}.$$

$$b) \quad C_1 = \{X\}, C_2 = \{W\}, C_3 = \{Z\}, C_4 = \{X, W\}, C_5 = \{X, Z\}, C_6 = \{W, Z\}, \\ C_7 = \{X, W, Z\}.$$

$$c) \quad C_1 = \{X\}, C_2 = \{W\}, C_3 = \{Z\}, C_4 = \{X, W\}, C_5 = \{X, Z\}, C_6 = \{W, Z\}, \\ C_7 = \{X, W, Z\}, C_8 = \{XW, WZ, XZ\}.$$

$$d) \quad C_1 = \{X, W, Z\}, C_2 = \{X, W\}, C_3 = \{X, Z\}, C_4 = \{W, X\}, C_5 = \{X\}, C_6 = \{W\}, \\ C_7 = \{Z\}.$$

Pytanie 7

Które ze stwierdzeń jest prawdziwe?

- a) Indywidualnych wskaźników pojemności informacyjnej liczy się w ramach jednej kombinacji tyle, ile jest w niej zmiennych objaśniających.
- b) Integralnych wskaźników pojemności informacyjnej liczy się w ramach jednej kombinacji tyle, ile jest w niej zmiennych objaśniających.
- c) Indywidualny wskaźnik pojemności informacyjnej jest sumą integralnych wskaźników pojemności informacyjnych w ramach danej kombinacji.
- d) Obliczana jest zawsze taka sama ilość indywidualnych, jak i integralnych wskaźników pojemności informacyjnej.

Pytanie 8

Które ze stwierdzeń nie dotyczy metody Hellwiga?

- a) Każdą z „kandydatek” na zmienne objaśniające traktuje się jako nośnik informacji o zmiennej objaśnianej, można więc uważać ją za źródło wiedzy o zmiennej Y .
- b) Jest to metoda optymalnego doboru predyktant.
- c) W ramach tej metody obliczane są indywidualne h_{sj} oraz integralne H_s wskaźniki pojemności informacyjnej.
- d) Za najlepszą (optymalną) kombinację nośników informacji uznaje się ten podzbiór na zmienne objaśniające, dla którego wskaźnik indywidualnej pojemności informacyjnej jest największy, czyli: $C_{opt} : h_{opt} = \max \{h_s : s = 1, 2, \dots, 2^m - 1\}$.

Pytanie 9

Które z indywidualnych wskaźników pojemności informacyjnej są poprawnie rozpisane dla kombinacji $C_9 = \{X_2, X_4, X_5\}$?

- a) $h_{92} = \frac{r_2^2}{1 + r_{24} + r_{25}}$, $h_{94} = \frac{r_4^2}{r_{42} + 1 + r_{45}}$, $h_{95} = \frac{r_5^2}{r_{52} + r_{54} + 1}$
- b) $h_{92} = \frac{r_2}{|r_{22}| + |r_{42}| + |r_{52}|}$, $h_{94} = \frac{r_4}{|r_{24}| + |r_{44}| + |r_{54}|}$, $h_{95} = \frac{r_5}{|r_{25}| + |r_{45}| + |r_{55}|}$
- c) $h_{92} = \frac{r_2^2}{|r_{22}| + |r_{24}| + |r_{25}|}$, $h_{94} = \frac{r_4^2}{|r_{42}| + |r_{44}| + |r_{45}|}$, $h_{95} = \frac{r_5^2}{|r_{52}| + |r_{54}| + |r_{55}|}$
- d) $h_{29} = \frac{r_2^2}{1 + |r_{42}| + |r_{52}|}$, $h_{49} = \frac{r_4^2}{|r_{24}| + 1 + |r_{54}|}$, $h_{59} = \frac{r_5^2}{|r_{25}| + |r_{45}| + 1}$

Pytanie 10

Przy zastosowaniu metody Hellwiga wskazano kombinację $C_9 = \{X_2, X_4, X_5\}$ za najlepszą do opisu zmiennej objaśnianej Y . Która z poniższych form stanowi najbardziej odpowiedni zapis ogólnego liniowego modelu ekonometrycznego?

- a) $Y = \alpha_0 + \alpha_2 X_2 + \alpha_4 X_4 + \alpha_5 X_5 + \varepsilon$
- b) $Y = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \alpha_3 X_3 + \varepsilon$
- c) $Y = \alpha_0 + \alpha_1 X_2 + \alpha_2 X_4 + \alpha_3 X_5 + \varepsilon$
- d) $Y = \alpha_1 X_2 + \alpha_2 X_4 + \alpha_3 X_5$

Część 2: ZADANIA

Zad. 1

Czy podana zmienna może zostać „kandydatką” na zmienną objaśniającą we względu na odpowiednią zmienność? (Przyjmując wartość krytyczną $V^* = 0,15$).

- a) X_1 : 10, 15, 12, 8, 11, 11, 16, 10.
- b) X_2 : 10, 11, 10, 12, 9, 11, 11, 10.
- c) Z: -1; 0,5; -1; -1,5; -1,5; 0,5; 0,5; 0; -0,5; 0,5.
- d) W: -3, -5, -7, -4, -5, -12, -1.
- e) Q: 15; 15,6; 16; 17,2; 16,6; 14; 16,3; 18; 15,9; 16,7; 16.
- f) R: 1,2; 1,3; 5; 1,8 1,9; 2,5.
- g) T: 56, 85, 45, 72, 66, 47, 73, 68, 74, 77, 80, 82.

Zad. 2

W 2015 roku zebrano informacje w siedmiu krakowskich uczelniach o liczbie studentów [w tys. osób] (Y) oraz powierzchni [w m²] sal dydaktycznych tych uczelni (X):

Uczelnia	AR	AE	WSP	UJ	PK	ASP	AWF
Liczba studiujących w tys. (y_i)	4	12	6	21	9	1	3
Powierzchnia sal dydaktycznych w tys. (x_i)	4,5	8,8	4,2	17	6	1,8	2,5

Czy istnieje obustronna zależność między zmiennymi X oraz Y? Celem odpowiedzi na to pytanie należy wykorzystać współczynnik korelacji liniowej Pearsona.

Zad. 3

Zbadaj zależność między czasem przeznaczonym na reklamę telewizyjną telewizora marki A [w min/miesiąc] a miesięczną ich sprzedaż [w tys. sztuk]. W tym celu zebrano dane za 7 miesięcy 2012 roku:

Miesiąc (i)	1	2	3	4	5	6	7
Czas przeznaczony na reklamę w min. (x_i)	11	17	13	13	21	17	8
Liczba sprzedanych telewizorów w tys. (y_i)	2,2	4,3	5,3	4,0	5,7	3,1	1,2

Zad. 4

Dane są:

$$\sum_{i=1}^n x_i y_i = 480, \quad s^2(x) = 9, \quad s^2(y) = 49, \quad \bar{x} = 5, \quad \bar{y} = 8, \quad n = 10.$$

Znajdź wartość współczynnika korelacji między tymi zmiennymi.

Zad. 5

Dane są:

$$\sum_{i=1}^n x_i y_i = 5600, \quad \sum_{i=1}^n x_i w_i = 7300, \quad \sum_{i=1}^n w_i y_i = 4300, \quad s^2(x) = 121, \quad s^2(y) = 49, \quad s^2(w) = 216, \\ \bar{x} = 20, \quad \bar{y} = 15, \quad \bar{w} = 22, \quad n = 15.$$

Znajdź wartości współczynników korelacji pomiędzy zmiennymi X , Y oraz W .

Zad. 6

Oblicz współczynniki korelacji liniowej Pearsona pomiędzy podanymi zmiennymi. Wyznacz macierz R oraz wektor R_0 .

a)

Y:	10	15	12	8	11	11	16	10
X:	10	11	10	12	9	11	11	10

b)

Y:	8	10	16	22	30	37
X:	14	15	16	27	25	30

c)

Y:	19	21	37	14	43
X:	5	8	7	3	16
Z:	62	92	73	103	84

d)

Y:	60	80	90	110	110	130	195	220	250	300
X ₁ :	11	12	19	18	22	37	21	10	29	30
X ₂ :	62	92	73	103	84	85	90	95	100	137
X ₃ :	103	100	95	91	91	96	83	73	52	91

e)

Y:	17	5	14	18	10	19	21	7	11	18	22	20	12	8	10	25
X ₁ :	122	100	169	185	163	198	200	117	147	188	222	180	107	99	130	205
X ₂ :	5,5	1,5	4	4,7	2,5	6,5	5	2	3,3	7	7,3	9,1	5,2	3,6	2,7	6
X ₃ :	1002	507	689	742	420	950	800	436	445	775	200	710	114	876	963	1100
X ₄ :	56	98	44	50	61	68	70	83	44	52	57	66	98	51	75	49

Zad. 7

Dla zmiennej objaśnianej Y wybrano pewien zbiór „kandydatek” na zmienne objaśniające. Poniżej podano macierze współczynników korelacji R oraz R_0 .

1. Określ kombinacje zmiennych objaśniających.
2. Dobierz zmienne objaśniające do modelu ekonometrycznego korzystając z metody Hellwiga (wskaźników pojemności informacyjnej).
3. Zapisz ogólny wzór modelu w postaci liniowej.

a) Zbiór kandydatek: $X = \{X_1, X_2\}$

$$R = [r_{ij}] = \begin{bmatrix} 1 & 0,73 \\ 0,73 & 1 \end{bmatrix}, \quad R_0 = [r_j] = \begin{bmatrix} 0,22 \\ 0,84 \end{bmatrix}$$

b) Zbiór kandydatek: $X = \{X_1, X_2, X_3\}$

$$R = [r_{ij}] = \begin{bmatrix} 1 & -0,2 & 0,8 \\ -0,2 & 1 & 0,3 \\ 0,8 & 0,3 & 1 \end{bmatrix}, \quad R_0 = [r_j] = \begin{bmatrix} 0,3 \\ 0,8 \\ 0,5 \end{bmatrix}$$

c) Zbiór kandydatek: $\{X, U, Z\}$

$$R = [r_{ij}] = \begin{bmatrix} 1 & 0,44 & 0,63 \\ 0,44 & 1 & -0,11 \\ 0,63 & -0,11 & 1 \end{bmatrix}, \quad R_0 = [r_j] = \begin{bmatrix} 0,23 \\ 0,52 \\ -0,73 \end{bmatrix}$$

d) Zbiór kandydatek: $X = \{X_1, X_2, X_3, X_4\}$

$$R = \begin{bmatrix} 1 & & & \\ 0,2 & 1 & & \\ 0,5 & 0,1 & 1 & \\ 0,6 & 0,3 & 0,9 & 1 \end{bmatrix}, \quad R_0 = \begin{bmatrix} 0,7 \\ 0,1 \\ 0,4 \\ 0,8 \end{bmatrix}$$

Zad. 8

Dla zmiennej objaśnianej Y oraz dla potencjalnych zmiennych objaśniających X_1, X_2, X_3, X_4 wyznaczono podane współczynniki korelacji liniowej Pearsona (zapisane w postaci macierzowej lub za pomocą oznaczeń r_j, r_{ij}). Wykorzystując metodę wskaźników pojemności informacyjnej zbuduj liniowy model ekonometryczny opisujący zależność zmiennej Y :

a) od trzech zmiennych objaśniających, gdzie:

$$R = \begin{bmatrix} 1 & & & \\ 0,2 & 1 & & \\ 0,5 & 0,1 & 1 & \\ 0,6 & 0,3 & 0,9 & 1 \end{bmatrix}, \quad R_0 = \begin{bmatrix} 0,7 \\ 0,1 \\ 0,4 \\ 0,8 \end{bmatrix}$$

b) od dwóch zmiennych objaśniających, gdzie:

$$R = \begin{bmatrix} 1 & & & \\ 0,25 & 1 & & \\ 0,55 & 0,35 & 1 & \\ 0,95 & 0,2 & 0,7 & 1 \end{bmatrix}, \quad R_0 = \begin{bmatrix} 0,6 \\ 0,3 \\ 0,4 \\ 0,75 \end{bmatrix}$$

c) od zbioru trzech zmiennych objaśniających zawierających zmienną X_4 , gdzie:

$$R = \begin{bmatrix} & & & \\ 0,25 & & & \\ 0,55 & 0,35 & & \\ 0,95 & 0,2 & 0,7 & \end{bmatrix}, \quad R_0 = \begin{bmatrix} 0,6 \\ 0,3 \\ 0,4 \\ 0,75 \end{bmatrix}$$

d) od zbioru trzech zmiennych objaśniających zawierających zmienną X_2 , gdzie:

$$r_2 = r_{13} = r_{24} = 0,45, \quad r_1 = r_{23} = 0,6, \quad r_4 = r_{14} = 0,8, \quad r_{12} = -0,2, \quad r_3 = r_{34} = 0,73$$

e) od zbioru dwóch zmiennych objaśniających zawierających zmienną X_3 , gdzie:

$$r_2 = r_{13} = r_{24} = 0,45, \quad r_1 = r_{23} = 0,6, \quad r_4 = r_{14} = 0,8, \quad r_{12} = -0,2, \quad r_3 = r_{34} = -0,73$$

f) od zbioru trzech zmiennych objaśniających zawierających zmienną X_3 , gdzie:

$$r_3 = r_{23} = 0,65, \quad r_{13} = r_{34} = -0,26, \quad r_1 = r_4 = r_{14} = 0,84, \quad r_2 = 0,37, \quad r_{12} = r_{24} = -0,43$$

Zad. 9

Współczynnik korelacji między zmiennymi X_1 i X_2 wynosi -0,5. Indywidualne wskaźniki pojemności informacyjnej kombinacji $C_3 = \{X_1, X_2\}$ wynoszą: $h_{31} = 0,54$, $h_{32} = 0,375$.

Ile wynoszą odpowiednio współczynniki korelacji zmiennej Y ze zmiennymi X_1 oraz X_2 ?

**Zad. 10**

Mamy trzy potencjalne zmienne objaśniające: $\{X_1, X_2, X_3\}$ do liniowego jednorównaniowego modelu ekonometrycznego. Dane są niektóre integralne wskaźniki pojemności informacyjnej oraz wybrane wartości współczynników korelacji pomiędzy zmiennymi: $H_1 = h_{11} = 0,9216$, $H_2 = h_{22} = 0,3025$, $H_5 = h_{51} + h_{53} = 0,943$, $r_3 = 0,87$, $r_{12} = -0,273$, $r_{32} = 0,732$. Wyznacz brakujące współczynniki. Dobierz zmienne objaśniające do modelu ekonometrycznego korzystając z metody Hellwiga oraz napisz wzór ogólny tego modelu.

KONIEC