

Ekonometria 1

Lista nr 3

- (1) W próbie obejmującej 20 losowo wybranych osób w pewnym województwie (dane przekrojowe w tabeli poniżej) do opisu rocznego spożycia mięsa wieprzowego y w kg zaproponowano trzy potencjalne zmienne objaśniające: x_1 - roczny dochód w tyz. zł. na osobę, x_2 - roczne spożycie ryb w kg na osobę, x_3 - spożycie ziemniaków w kg na osobę. Dokonaj wyboru zmiennych w modelu stosując metodę Hellwiga.

Nr	y	x_1	x_2	x_3
1	18	40	12	165
2	18.3	28	14	155
3	22	20	10	111
4	25	22	8	107
5	26.6	18	4	140
6	27	21	7	110
7	27	20	4.5	170
8	30	19	5	109
9	31	20	6.5	138
10	32.5	23	6	142
11	34	27	6	149
12	34	28	5	128
13	35	25	3.5	139
14	36.6	31	3	115
15	38	33	3	141
16	40	37	2.5	121
17	42.5	38	3.5	150
18	43	26	2.5	170
19	43	38	1	149
20	44	41	2	168

- (2) W zadaniu (1) dokonaj wyboru zmiennych metodą regresji krokowej “wstecz” (z wykorzystaniem testu t -Studenta).
- (3) W zadaniu (1) dokonaj wyboru zmiennych stosując metodę regresji “w przód” (z wykorzystaniem testu t -Studenta).
- (4) W zadaniu 91) dokonaj wyboru zmiennych metodą regresji krokowej “w przód” w wersji wykorzystującej cząstkowe test F i cząstkowe współczynniki korelacji. Wskazówka (prawie rozwiązanie): obliczamy r_1, r_2, r_3 . Wybieramy x_2 . Empiryczna wartość cząstkowej statystyki F dla x_2 jest 67.71, a wartość krytyczna F dla $\alpha = 0.05, m_1 = 1, m_2 = 18$ jest $F_{in} = 4.41$ - zmienna x_2 zostaje w modelu. Obliczamy $r_{y1.2}, r_{y3.2}$ i na tej podstawie wybieramy x_1 . F_{in} określamy jako wartość krytyczną F dla $m_1 = 1, m_2 = 17$. Ponieważ wartość cząstkowej statystyki F dla x_1 jest większa od F_{in} , więc x_1 zostaje w modelu. F_{ou} określamy jako wartość krytyczną F dla $\alpha = 0.05, m_1 = 1, m_2 = 18$. Wartość cząstkowej statystyki F dla x_2 jest większa od F_{ou} , więc nie usuwamy x_2 . Potem wybieramy x_3 i obliczamy dla tej zmiennej F . Dostajemy $F < F_{in}$, więc zmiennej nie wprowadzamy do modelu. Wybraliśmy zatem x_1 i x_2 .
- (5) Na podstawie danych z 12 lat oszacowano dwa modele określające zależność zmiennej objaśnianej od trzech zmiennych objaśniających. Model (I) został opisany funkcją liniową, a model (II) potęgową. Dla modeli obliczono reszty:

Rok	1	2	3	4	5	6	7	8	9	10	11	12
Reszty w (I)	2	-3	1	2	4	3	-1	-4	-2	1	2	3
Reszty w (II)	-7	-5	-8	-4	1	2	4	3	2	-1	-4	-5

Który model lepiej oddaje zależność.

- (6) Na podstawie danych przekrojowych oszacowano model $\hat{y}_i = 2 + 3x_i, i = 1, \dots, 10$. Reszty modelu, po uporządkowaniu ze względu na zmienną objaśniającą, wynoszą $-10, -8, -9, -6, -1, -2, 1, 2, 5, 8$. Czy rozkład odchyłeń jest losowy?
- (7) Dla pewnego liniowego modelu tendencji rozwojowej otrzymano następujący ciąg reszt odpowiadający kolejnym latom badanego okresu: $-6, -4, -2, -1, 0, 1, 3, 2, 4, -1, -3, -4$. Czy rozkład odchyłeń jest losowy?
- (8) W pewnym modelu, którego parametry zostały oszacowane na podstawie 14 obserwacji, wartość statystyki Shapiro-Wilka dla reszt wynosi $W = 0.923$. Czy na poziomie istotności $\alpha = 0.05$ należy odrzucić hipotezę o normalności składnika losowego?

- (9) Dany jest ciąg reszt pewnego modelu: 12, -2, 0, -1, 1, 0, 10, -4. Na podstawie testu Shapiro-Wilka zweryfikować hipotezę o normalności rozkładu składnika losowego na poziomie istotności $\alpha = 0,05$.
- (10) Na podstawie próby 15-elementowej oszacowano metodą NK model $\hat{y}_i = 32 + 0.35x_{1i} - 0.46x_{2i}$, w którym wartość statystyki Durбина-Watsona wynosi $d = 3.34$. Na poziomie istotności $\alpha = 0.05$ dokonać weryfikacji hipotezy o autokorelacji składnika losowego.
- (11) Czy na podstawie testu Durбина-Watsona, na poziomie istotności $\alpha = 0.05$, można sądzić, że w modelu $\hat{y}_i = 2 + 0.1x_{1i} + 0.3x_{2i} = 4x_{3i}$, w którym $n = 40$, $d = 2.5$, nie występuje autokorelacja składnika losowego?
- (12) Dla modelu liniowego opisującego zależność zmiennej objaśnianej od jednej zmiennej objaśniającej otrzymano ciąg reszt:

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
e_t	5	0	-1	2	-3	8	1	-2	-4	-1	0	1	2	2	-1

Na poziomie istotności $\alpha = 0.05$ sprawdzić czy występuje autokorelacja składnika losowego.

- (13) Model $\hat{y}_i = 2 + 3x_{1i} + 4x_{2i} + 100x_{3i}$, został oszacowany na podstawie 50-elementowej próby. Po uporządkowaniu próby ze względu na zmienną x_1 dokonano estymacji modelu dla dwóch skrajnych podprób 20-elementowych i na tej podstawie otrzymano, że sumy kwadratów reszt wynoszą $S_1^2 = 4$, $S_2^2 = 14$. Czy można przypuszczać, że składniki losowe modelu, którego parametry oszacowano na podstawie całej próby są homoskedastyczne.
- (14) Kształtowanie się wielkości przychodów i wartości sprzedaży w 14 firmach przedstawia się następująco:

Firma	przychody (tys. zł.)	Sprzedaż (tys. zł.)
1	10	85
2	12	82
3	8	50
4	20	100
5	26	132
6	30	138
7	18	105
8	22	90
9	30	110
10	9	70
11	15	120
12	21	125
13	30	150
14	14	90

Oszacowano model opisujący wielkość przychodów w zależności od wartości sprzedaży. Czy założenie o homoskedastyczności składnika losowego zostało w tym modelu spełnione?