

6. Elementy prognozowania ekonometrycznego. Wyznaczanie prognoz na podstawie modelu ekonometrycznego, szacowanie średnich błędów prognozy, wyznaczenie przedziałów ufności dla zmiennej prognozowanej.

Zadanie 1. Oszacowano model opisujący ilość urodzeń (żywych) w zależności od ilości zawartych małżeństw i rozwodów w Polsce.

Model 1: Estymacja KMNK, wykorzystane obserwacje 2000-2013 ($N = 14$)
Zmienna zależna: U

	Współczynnik	Błąd stand.	t-Studenta	wartość p	
const	217,595	35,4462	6,1387	0,00007	***
M	0,689656	0,179001	3,8528	0,00269	***
R	0,254891	0,449597	0,5669	0,58214	
Średn. aryt. zm. zależnej	380,2643	Odch. stand. zm. zależnej		22,36882	
Suma kwadratów reszt	2186,072	Błąd standardowy reszt		14,09730	
Wsp. determ. R-kwadrat	0,663926	Skorygowany R-kwadrat		0,602822	
F(2, 11)	10,86544	Wartość p dla testu F		0,002485	
Logarytm wiarygodności	-55,22077	Kryt. inform. Akaike'a		116,4415	
Kryt. bayes. Schwarz	118,3587	Kryt. Hannana-Quinna		116,2641	
Autokorel. reszt - rho1	0,778231	Stat. Durbina-Watsona		0,465074	

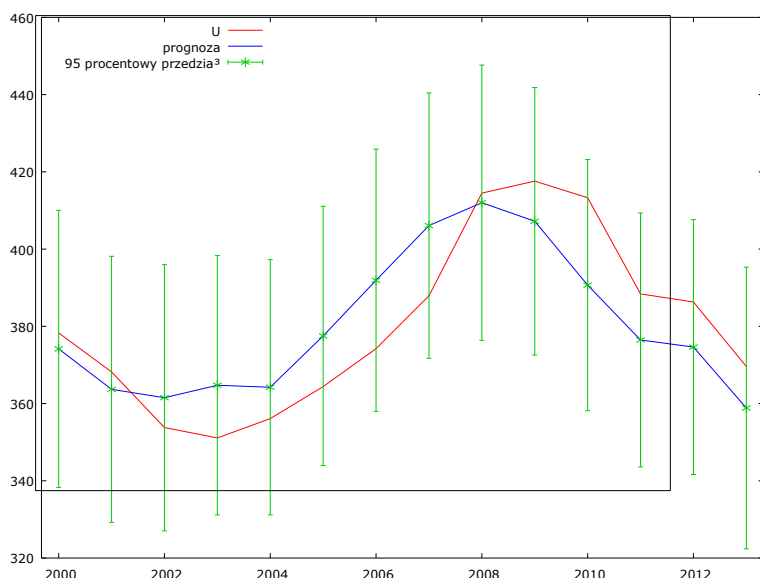
Dokonano prognozy. Zinterpretuj oszacowania i miary dokładności prognoz ex post

Dla 95% przedziału ufności, $t(11, 0,025) = 2,201$

Obs	U	prognoza	błąd standardowy	95% przedział
2000	378,300	374,159	16,2969	(338,290, 410,029)
2001	368,200	363,693	15,6500	(329,248, 398,138)
2002	353,800	361,512	15,6644	(327,035, 395,989)
2003	351,100	364,741	15,2566	(331,162, 398,321)
2004	356,100	364,221	15,0161	(331,171, 397,271)
2005	364,400	377,515	15,2418	(343,968, 411,062)
2006	374,200	391,922	15,4270	(357,967, 425,876)
2007	387,900	406,088	15,6110	(371,728, 440,447)
2008	414,500	412,014	16,1921	(376,376, 447,653)
2009	417,600	407,205	15,7325	(372,578, 441,832)
2010	413,300	390,668	14,7742	(358,150, 423,186)
2011	388,400	376,475	14,9326	(343,608, 409,341)
2012	386,300	374,631	14,9956	(341,625, 407,636)
2013	369,600	358,857	16,5686	(322,390, 395,324)

Miary dokładności prognoz ex post

Średni błąd predykcji	ME = 4,8723e-014
Błąd średniokwadratowy	MSE = 156,15
Pierwiastek błędu średniokwadr.	RMSE = 12,496
Średni błąd absolutny	MAE = 11,214
Średni błąd procentowy	MPE = -0,10628
Średni absolutny błąd procentowy	MAPE = 2,9465
Współczynnik Theila (w procentach)	I = 0,97714
Udział obciążoności predykc.	$I1^2/I^2 = 0$
Udział niedost. elastyczności	$I2^2/I^2 = 0$
Udział niezgodności kierunku	$I3^2/I^2 = 1$



Miary dokładności prognoz ex post

Nazwa	Wzór	Interpretacja
Średni błąd predykcji	$ME = \frac{1}{m} \sum_{\tau=1}^m (y_{\tau} - y_{\tau}^P)$	<ul style="list-style-type: none"> — wartość ME powinna być równa zero lub bliska zero; — średnie obciążenie predykcji przyjmuje wartość zero w przypadku predykcji nieobciążonej; — odchylenia wartości miernika ME od zera świadczą, że zasada predykcji nieobciążonej nie została zachowana; — gdy zaobserwowane odchylenie od zera jest dodatnie, wnioskujemy, że prognozy wygasłe są niedoszacowane; — gdy zaobserwowane odchylenie od zera jest ujemne, wnioskujemy, że prognozy wygasłe są przeszacowane
Błąd średniokwadratowy	MSE	
Pierwiastek błędu średniokwadratowego	$RMSE = \sqrt{\frac{1}{m} \sum_{\tau=1}^m (y_{\tau} - y_{\tau}^P)^2}$	<ul style="list-style-type: none"> — RMSE mierzy, o ile średnio odchylają się realizacje zmiennej prognozowanej od obliczonych prognoz — znacząca różnica wartości między MAE i RMSE wskazuje na występowanie w okresie prognozy błędów o bardzo dużych wartościach.
Średni błąd absolutny	$MAE = \frac{1}{m} \sum_{\tau=1}^m y_{\tau} - y_{\tau}^P $	informuje o ile średnio - w okresie predykcji - rzeczywiste realizacje zmiennej prognozowanej będą się odchyłać – co do bezwzględnej wartości – od prognoz
Średni błąd procentowy	$MPE = \frac{\sum_{\tau=1}^m PE_{\tau}}{m}$	informuje, jaki procent rzeczywistych realizacji zmiennej prognozowanej stanowią błędy prognozy w okresie predykcji
Średni absolutny błąd procentowy	$MAPE = \frac{1}{m} \sum_{\tau=1}^m \left \frac{y_{\tau} - y_{\tau}^P}{y_{\tau}} \right \cdot 100$	informuje o średniej wielkości błędów prognoz, wyrażonych w procentach rzeczywistych wartości zmiennej prognozowanej. Wartości MAPE pozwalają porównać dokładność prognoz otrzymywanych różnych modeli.
Współczynnik Theila (w procentach)	$I^2 = \frac{\sum_{\tau=1}^m (y_{\tau} - y_{\tau}^P)^2}{\sum_{\tau=1}^m y_{\tau}^2}$ <p>I – przeciętny względny błąd prognozy</p> $I^2 = I_1^2 + I_2^2 + I_3^2$	<p>Pierwiastek współczynnika Theila informuje o całkowitym względnym błędzie prognozy w okresie testowania</p> <p>I_1^2 odzwierciedla obciążenie predykcji, czyli to, w jakim stopniu nie udało się odgadnąć średniej wartości zmiennej prognozowanej.</p> <p>I_2^2 określa, w jakim stopniu zmienność prognozy i zmiennej prognozowanej są do siebie zbliżone.</p> <p>I_3^2 określa, czy wystąpiła zgodność kierunków zmian prognoz z rzeczywistym kierunkiem zmian zmiennej prognozowanej.</p>
Udział obciążoności predykcji	I_1^2 / I^2	miernik całkowitego względnego błędu prognozy, wynikającego z obciążenia predykcji (małej zgodności średnich wartości zmiennej objaśnianej i jej prognozy w

		przedziale weryfikacji prognoz)
Udział niedost. elastyczności	I_2^2 / I^2	miernik całkowitego względnego błędu prognozy, wynikającego z niedostatecznej elastyczności predykcji (małej zgodności zróżnicowania wartości zmiennej objaśnianej, a wartości prognozy w przedziale weryfikacji prognoz)
Udział niezgodności kierunku	I_3^2 / I^2	miernik całkowitego względnego błędu prognozy, wynikającego z niedostatecznej predykcji tak zwanych punktów zwrotnych (małej zgodności kierunków zmian wartości zmiennej objaśnianej i jej prognozy w przedziale weryfikacji prognoz)

Zadanie 2. Oszacowano model trendu z sezonowością opisujący ilość mieszkań oddanych do użytku w danym kwartale w okresie 1998:4-2012:3:

$$M_t = \alpha_0 + \alpha_1 t + \delta'_1 Q_1 + \delta'_2 Q_2 + \delta'_3 Q_3 + \varepsilon_t$$

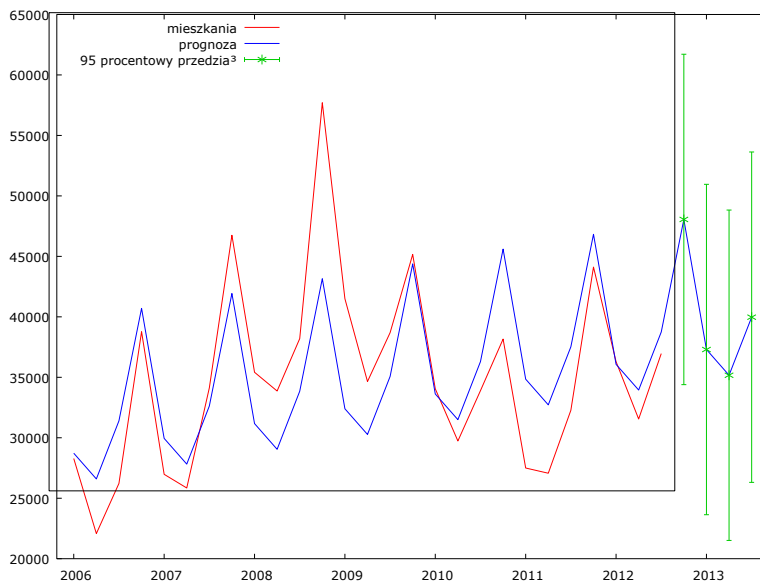
Model 1: Estymacja KMNK, wykorzystane obserwacje 1998:4-2012:3 (N = 56)
Zmienna zależna: mieszkania

	Współczynnik	Błąd stand.	t-Studenta	wartość p	
const	30600,4	2227,8	13,7357	<0,00001	***
time	306,171	52,9679	5,7803	<0,00001	***
dq1	-11061,7	2416,29	-4,5780	0,00003	***
dq2	-13493,7	2418,03	-5,5804	<0,00001	***
dq3	-8999,01	2420,93	-3,7172	0,00050	***
Średn. aryt. zm. zależnej	30937,66	Odch. stand. zm. zależnej		9326,297	
Suma kwadratów reszt	2,08e+09	Błąd standardowy reszt		6391,367	
Wsp. determ. R-kwadrat	0,564512	Skorygowany R-kwadrat		0,530356	
F(4, 51)	16,52748	Wartość p dla testu F		9,58e-09	
Logarytm wiarygodności	-567,5532	Kryt. inform. Akaike'a		1145,106	
Kryt. bayes. Schwarz	1155,233	Kryt. Hannana-Quinna		1149,033	
Autokorel. reszt - rho1	0,576612	Stat. Durbina-Watsona		0,844171	

Dokonano prognozy ex ante

Dla 95% przedziału ufności, $t(51, 0,025) = 2,008$

Obs	mieszkania	prognoza	błąd ex ante	95% przedział ufności
2011:3	32271,00	37522,26		
2011:4	44115,00	46827,45		
2012:1	36292,00	36071,95		
2012:2	31551,00	33946,09		
2012:3	36962,00	38746,95		
2012:4		48052,13	6803,855	34392,82 - 61711,44
2013:1		37296,63	6803,855	23637,32 - 50955,94
2013:2		35170,77	6803,855	21511,47 - 48830,08
2013:3		39971,63	6803,855	26312,32 - 53630,94



Okres, którego dotyczy sporządzana prognoza nazywa się okresem prognozy (T).

Liczba jednostek czasu, jaka upływa od terażniejszości do okresu prognozowania nazywa się wyprzedzeniem prognozy (τ).

Dla 95% przedziału ufności, $t(51, 0,025) = 2,008$

Obs	mieszkania	prognoza	błąd ex ante	95% przedział ufności
2012:2	31551,00	33946,09		
2012:3	36962,00	38746,95		
2012:4		48052,13	6803,855	34392,82 - 61711,44
2013:1		37296,63	6803,855	23637,32 - 50955,94
2013:2		35170,77	6803,855	21511,47 - 48830,08
2013:3		39971,63	6803,855	26312,32 - 53630,94

Interpretacja wyników prognozowania dla okresu 2012q4 jest następująca:

- w warunkach stabilności prognostycznej modelu oczekuję, że zmienna prognozowana w okresie 2012q4 przyjmie wartość równą $M_{2012q3|2012q4}^p = 48052,13$ ze średnim błędem $\pm \hat{\sigma}(\xi_{2012q3|2012q4}^p) = 6803,86$,
- z prawdopodobieństwem 0,95, w warunkach stabilności prognostycznej modelu przedział $P\{34392,82 \leq M_{2012q4} \leq 61711,44\} = 0,95$ pokryje nieznaną realizację zmiennej M_t w okresie 2012q4.