

Forecasting is the art of saying what will happen, and then explaining why it didn't.

Ch. Chatfield (1986)

PROGNOZY I SYMULACJE

Katarzyna Chudy – Laskowska

WYKŁAD VI

Szeregi czasowe I

1. Modele średniej ruchomej
2. Metoda średniej ruchomej ważonej
3. Model podwójnej średniej ruchomej
4. Wygładzanie wykładnicze - model prosty Browna
5. Wygładzanie wykładnicze - liniowy model Holta
6. Model trendu pełzającego z wagami harmonocznymi

1. MODELE ŚREDNIEJ RUCHOMEJ

Jeśli w szeregu czasowym występują oprócz innych składowych, stosunkowo duże wahania przypadkowe, to prognozy otrzymane przy użyciu metod naiwnych będą obarczone dużymi błędami. W celu poprawienia dokładności prognoz można zastosować metodę opartą na modelu średniej ruchomej. Modele te mogą być wykorzystane zarówno do wygładzania szeregu jaki do prognozowania.

Idea wyrównywania szeregu czasowego za pomocą średnich ruchomych (tzw. metoda mechaniczna) polega na zastąpieniu pierwotnych wartości zmiennej prognozowanej średnimi arytmetycznymi, obliczanymi sekwencyjnie dla wybranej liczby obserwacji. Wyznaczone obserwacje przyporządkowuje się środkowym obserwacjom, na których podstawie były obliczane średnie.

Używając tych modeli zakłada się, że wartość zmiennej prognozowanej w następnym momencie/okresie będzie równa średniej arytmetycznej z k ostatnich wartości tej zmiennej. Formuła obliczeniowa wygląda następująco:

$$y_t^* = \frac{1}{k} \sum_{i=t-k}^{t-1} y_i$$

y_t^* prognoza w wyznaczonym okresie t
 y_i wielkość w okresie t
 k stała wygładzania (określana przez prognozę)

Wraz ze wzrostem wartości stałej wygładzania rośnie efekt wyrównywania. Im większa liczba obserwacji użytych do wygładzania tym większy efekt wygładzania.

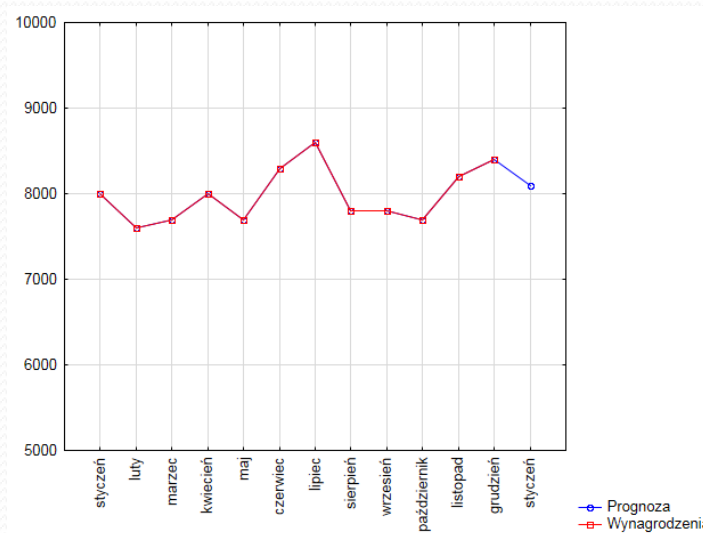
1. MODELE ŚREDNIEJ RUCHOMEJ - przykład 1

Wynagrodzenia w pewnym sektorze przedstawia poniższy szereg:

8000; 7000,6000; 7000,7000; 8000; 7700; 8300; 8600; 7800; 7800; 7700; 8200; 8400.

- 1) Określić składowe szeregu
- 2) Wyznaczyć prognozę 3 elementową średnią ruchomą prostą.

t	Y
1	8000
2	7600
3	7700
4	8000
5	7700
6	8300
7	8600
8	7800
9	7800
10	7700
11	8200
12	8400



$$s = 310,5$$

$$\bar{x} = 7983$$

$$: \frac{s}{\bar{x}} \cdot 100\% = \frac{310,5}{7983} \cdot 100\% = 3,89\%$$

Trzyelementowa średnia ruchoma prosta

$$y_t^* = \frac{1}{k} \sum_{i=t-k}^{t-1} y_i$$

$$y_{13}^* = \frac{1}{3} (7700 + 8200 + 8400) = 8100 \text{ [zł]}$$

1. METODA ŚREDNIEJ RUCHOMEJ - przykład 2

Liczba lokat założonych w pewnym oddziale w banku kolejnych dniach była następująca: 41; 38; 39; 40; 39; 42; 38; 39; 41; 42; 40; 39. Wyznaczyć prognozę ilości założonych lokat stosując 3-elementową średnią ruchomą prostą oraz 5-elementową średnią ruchomą prostą.

$$s = 1,4$$

$$\bar{x} = 39,93$$

$$V_z = \frac{s}{\bar{x}} \cdot 100\% = \frac{1,4}{39,93} \cdot 100\% = 3,52\%$$

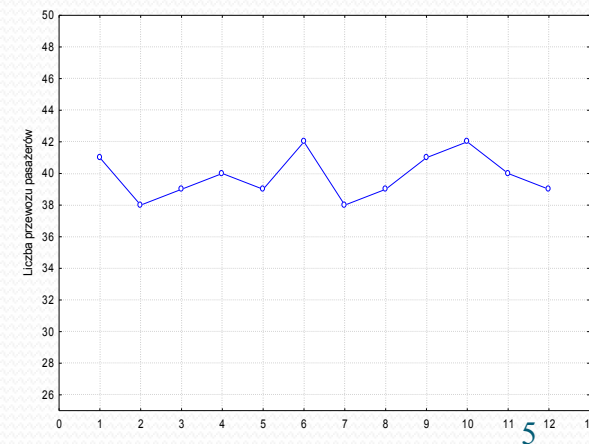
Średni względny błąd prognoz (ψ)

$$\Psi = \frac{1}{n-k} \sum_{t=n+1}^n \frac{|y_t - y_t^*|}{y_t} \cdot 100$$

$$\Psi^{(3-el)} = \frac{1}{12-3} \cdot 32,26 = 3,58\%$$

$$\Psi^{(5-el)} = \frac{1}{12-5} \cdot 24,16 = 3,45\%$$

t	Liczba pasażerów Y_t	Prognoza 3-elementowa	Błąd $\psi_t = \frac{ y_t - y_t^* }{y_t} \cdot 100$	Prognoza 5-elementowa	Błąd $\psi_t = \frac{ y_t - y_t^* }{y_t} \cdot 100$
1	41				
2	38				
3	39				
4	40	$(41+38+39)/3=39,3$	1,75		
5	39	$(38+39+40)/3=39,00$	0,00		
6	42	$(39+40+39)/3=39,33$	6,35	$(41+38+39+40+39)/5=39,4$	6,19
7	38	$(40+39+42)/3=40,33$	6,14	$(38+39+40+39+42)/5=39,6$	4,21
8	39	$(39+42+38)/3=39,67$	1,71	$(39+40+39+42+38)/5=39,6$	1,54
9	41	$(42+38+39)/3=39,67$	3,25	$(40+39+42+38+39)/5=39,6$	3,41
10	42	$(38+39+41)/3=39,33$	6,35	$(39+42+38+39+41)/5=39,8$	5,24
11	40	$(39+41+42)/3=40,67$	1,67	$(42+38+39+41+42)/5=40,4$	1,00
12	39	$(41+42+40)/3=41,00$	5,13	$(38+39+41+42+40)/5=40,0$	2,56
13		$(42+40+39)/3=40,33$		$(39+41+42+40+39)/5=40,2$	
	Σ		32,26		24,16



1. MODELE ŚREDNIEJ RUCHOMEJ - przykład 2

do rozwiązania

W pewnym banku podano liczbę obsłużonych klientów w ciągu 12 miesięcy w postaci szeregu: 720,710,704,699,732,715,713,724,696,699,703,705.
Dokonać prognozy metodą średniej ruchomej 3 elementowej oraz 5 elementowej.

$$V_z = ?$$

$$Y_{13}^* = ?$$

2. METODA ŚREDNIEJ RUCHOMEJ WAŻONEJ

W razie budowy prognoz na podstawie modelu średniej ruchomej ważonej należy określić liczbę wyrazów średniej oraz wagi nadawane poszczególnym wyrazom. Najmniejsze błędy prognoz daje średnia z co najmniej 12 obserwacji.

$$y_t^* = \sum_{i=t-k}^{t-1} y_i w_{i-t+k+1}$$

y_t^* prognoza w wyznaczonym okresie t

y_i wielkość w okresie t

k stała wygładzania (określana przez prognozę)

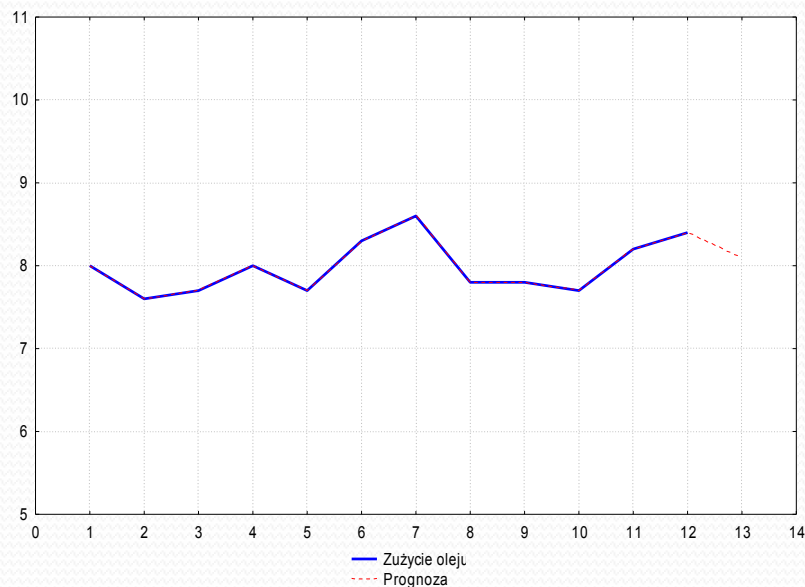
$w_{i-t+k+1}$ waga nadawana przez prognozę wartości zmiennej prognozowanej w okresie t

2. METODA ŚREDNIEJ RUCHOMEJ WAŻONEJ - przykład 1

Zużycie oleju w pewnym punkcie gastronomicznym przedstawia poniższy szereg:

8; 7,6; 7,7; 8; 7,7; 8,3; 8,6; 7,8; 7,8; 7,7; 8,2; 8,4.

- 1) Określić składowe szeregu
- 2) Wyznaczyć prognozę 3 elementową średnią ruchomą ważoną.



Nadano wagi poszczególnym obserwacjom:

I	0,2	0,3	0,5
II	0,15	0,25	0,6
III	0,1	0,2	0,7

- ! Wagi powinny być z przedziału $(0,1]$
- !! Suma wag powinna być równa 1

2. METODA ŚREDNIEJ RUCHOMEJ WAŻONEJ - przykład 1

$$y_t^* = \sum_{i=t-k}^{t-1} y_i w_{i-t+k+1}$$

Przyjęte wagi

I	0,2	0,3	0,5
II	0,15	0,25	0,6
III	0,1	0,2	0,7

Dla każdego z przyjętych modeli wyznaczono średni kwadratowy błąd ex post.
Procedura wygląda następująco: (dla III przypadku)

t	Zużycie Y _t	Prognoza ŚW3 (0,1;0,2;0,7) y_t^*	BBP ex post $q_t = y_t - y_t^*$	q_t^2
1	8		bezwzględny	
2	7,6		błąd	
3	7,7		prognozy	
4	8	$(8,0 \cdot 0,1 + 7,6 \cdot 0,2 + 7,7 \cdot 0,7) = 7,71$	0,29	0,084
5	7,7	$(7,6 \cdot 0,1 + 7,7 \cdot 0,2 + 8,0 \cdot 0,7) = 7,90$	-0,20	0,040
6	8,3	$(7,7 \cdot 0,1 + 8,0 \cdot 0,2 + 7,7 \cdot 0,7) = 7,76$	0,54	0,292
7	8,6	$(8,0 \cdot 0,1 + 7,7 \cdot 0,2 + 8,3 \cdot 0,7) = 8,15$	0,45	0,202
8	7,8	$(7,7 \cdot 0,1 + 8,3 \cdot 0,2 + 8,6 \cdot 0,7) = 8,45$	-0,65	0,422
9	7,8	$(8,3 \cdot 0,1 + 8,6 \cdot 0,2 + 7,8 \cdot 0,7) = 8,01$	-0,21	0,044
10	7,7	$(8,6 \cdot 0,1 + 7,8 \cdot 0,2 + 7,8 \cdot 0,7) = 7,88$	-0,18	0,032
11	8,2	$(7,8 \cdot 0,1 + 7,8 \cdot 0,2 + 7,7 \cdot 0,7) = 7,73$	0,47	0,221
12	8,4	$(7,8 \cdot 0,1 + 7,7 \cdot 0,2 + 8,2 \cdot 0,7) = 8,06$	0,34	0,116
13		$(7,7 \cdot 0,1 + 8,2 \cdot 0,2 + 8,4 \cdot 0,7) = 8,29$		
			Σ	1,454

$$s^* = \left[\frac{1}{n-k} \sum_{t=k+1}^n (y_t - y_t^*)^2 \right]^{\frac{1}{2}}$$

$$s^* = \left[\frac{1}{12-3} \cdot 1,454 \right]^{\frac{1}{2}} = 0,402$$

Średni kwadratowy błąd ex post

I = 0,407

II = 0,403

III = 0,402

$$y_{13}^* = (7,7 \cdot 0,1 + 8,2 \cdot 0,2 + 8,4 \cdot 0,7) = 8,29$$

Obliczyć prognozę dla wag: 0,7;0,2;0,1

3. MODEL PODWÓJNEJ ŚREDNIEJ RUCHOMEJ

W razie występowania w szeregu czasowym trendu liniowego można do konstrukcji prognozy zastosować model podwójnej średniej ruchomej, określanego w literaturze także mianem modelu liniowej średniej ruchomej.

Wygładzony wówczas szereg wartości zmiennej prognozowanej poddaje się powtórnemu wygładzaniu metodą średniej ruchomej.

4. METODA WYGŁADZANIA WYKŁADNICZEGO

Istota wygładzania wykładniczego polega na tym, że szereg czasowy zmiennej prognozowanej wygładza się za pomocą średniej ruchomej ważonej, przy czym wagi są określane według prawa wykładniczego.

Modeli wygładzania wykładniczego jest wiele Statistica dysponuje 12 różnymi przypadkami wygładzania wykładniczego w zależności od znajdujących się w szeregu czasowym składowych.

4 PROSTY MODEL WYGŁADZANIA WYKŁADNICZEGO BROWNA

Może być stosowany do prognozowania zjawisk, które charakteryzują się prawie stałym poziomem oraz niewielkimi wahaniami przypadkowymi.

$$y_t^* = \alpha y_{t-1} + (1 - \alpha) y_{t-1}^*, \quad \alpha \in (0, 1]$$

y_t^*, y_{t-1}^* prognoza wyznaczona na okres $t, t-1$

y_{t-1} wielkość zmiennej w okresie $t-1$

α waga nadana ostatniej (najnowszej) obserwacji zmiennej prognozowanej

Konstruując prognozę na okres t , przyjmuje się, że będzie ona równa prognozie wyznaczonej na okres poprzedni $t-1$, skorygowanej o pewną część (α) jej bezwzględnego błędu *ex post*.

Parametr α -zwany jest parametrem wygładzania i przyjmuje wartości z przedziału 0-1. Wartość tego parametru wyznacza się eksperymentalnie. Natomiast jako wartość prognozy Y_1^* niezbędną do konstrukcji modelu, przyjmuje się najczęściej wartość początkową Zmiennej prognozowanej (y_1) lub średnią arytmetyczną z kilku pierwszych obserwacji.

4. PROSTY MODEL WYGŁADZANIA WYKŁADNICZEGO - przykład

Liczba klientów w pewnej firmie ubezpieczeniowej w kolejnych kwartałach była następująca: 41; 38; 39; 40; 39; 42; 38; 39; 41; 42; 40; 39. (dane w tys. osób)

Wyznaczyć prognozę liczby klientów korzystając z metody prostego wygładzania wykładniczego dla parametru $\alpha=0,05$ oraz $\alpha=0,6$. Wybrać model, który charakteryzuje się mniejszymi błędami prognozy.

$$y_t^* = \alpha y_{t-1} + (1 - \alpha) y_{t-1}^*, \quad \alpha \in (0,1]$$

t	Liczba klientów y_t	Prognoza $\alpha=0,05$	Błąd	Prognoza $\alpha=0,6$	Błąd
1	41	$(41+38+39)/3=39,33$	4,07	$(41+38+39)/3=39,33$	4,07
2	38	$[(0,05)41+(1-0,05)39,33]=39,42$	3,73	$[(0,6)41+(1-0,6)39,33]=40,33$	6,14
3	39	$[(0,05)38+(1-0,05)39,42]=39,35$	0,89	$[(0,6)38+(1-0,6)40,33]=38,93$	0,17
4	40	$[(0,05)39+(1-0,05)39,35]=39,33$	1,68	$[(0,6)39+(1-0,6)38,93]=38,97$	2,57
5	39	$[(0,05)40+(1-0,05)39,33]=39,36$	0,93	$[(0,6)40+(1-0,6)38,97]=39,59$	1,51
6	42	$[(0,05)39+(1-0,05)39,36]=39,34$	6,32	$[(0,6)39+(1-0,6)39,59]=39,24$	6,58
7	38	$[(0,05)42+(1-0,05)39,34]=39,48$	3,98	$[(0,6)42+(1-0,6)39,24]=40,89$	7,62
8	39	$[(0,05)38+(1-0,05)39,48]=39,40$	1,03	$[(0,6)38+(1-0,6)40,89]=39,16$	0,4
9	41	$[(0,05)39+(1-0,05)39,40]=39,38$	3,94	$[(0,6)39+(1-0,6)39,16]=39,06$	4,72
10	42	$[(0,05)41+(1-0,05)39,38]=39,46$	6,04	$[(0,6)41+(1-0,6)39,06]=40,23$	4,23
11	40	$[(0,05)42+(1-0,05)39,46]=39,59$	1,02	$[(0,6)42+(1-0,6)40,23]=41,29$	3,23
12	39	$[(0,05)40+(1-0,05)39,59]=39,61$	1,57	$[(0,6)40+(1-0,6)41,29]=40,52$	3,89
13		$[(0,05)39+(1-0,05)39,61]=39,58$	-	$[(0,6)39+(1-0,6)40,52]=39,61$	
	Σ		35,10		45,12

$$\Psi_t = \frac{1}{n-k} \sum_{t=k+1}^n \frac{|y_t - y_t^*|}{y_t} \cdot 100$$

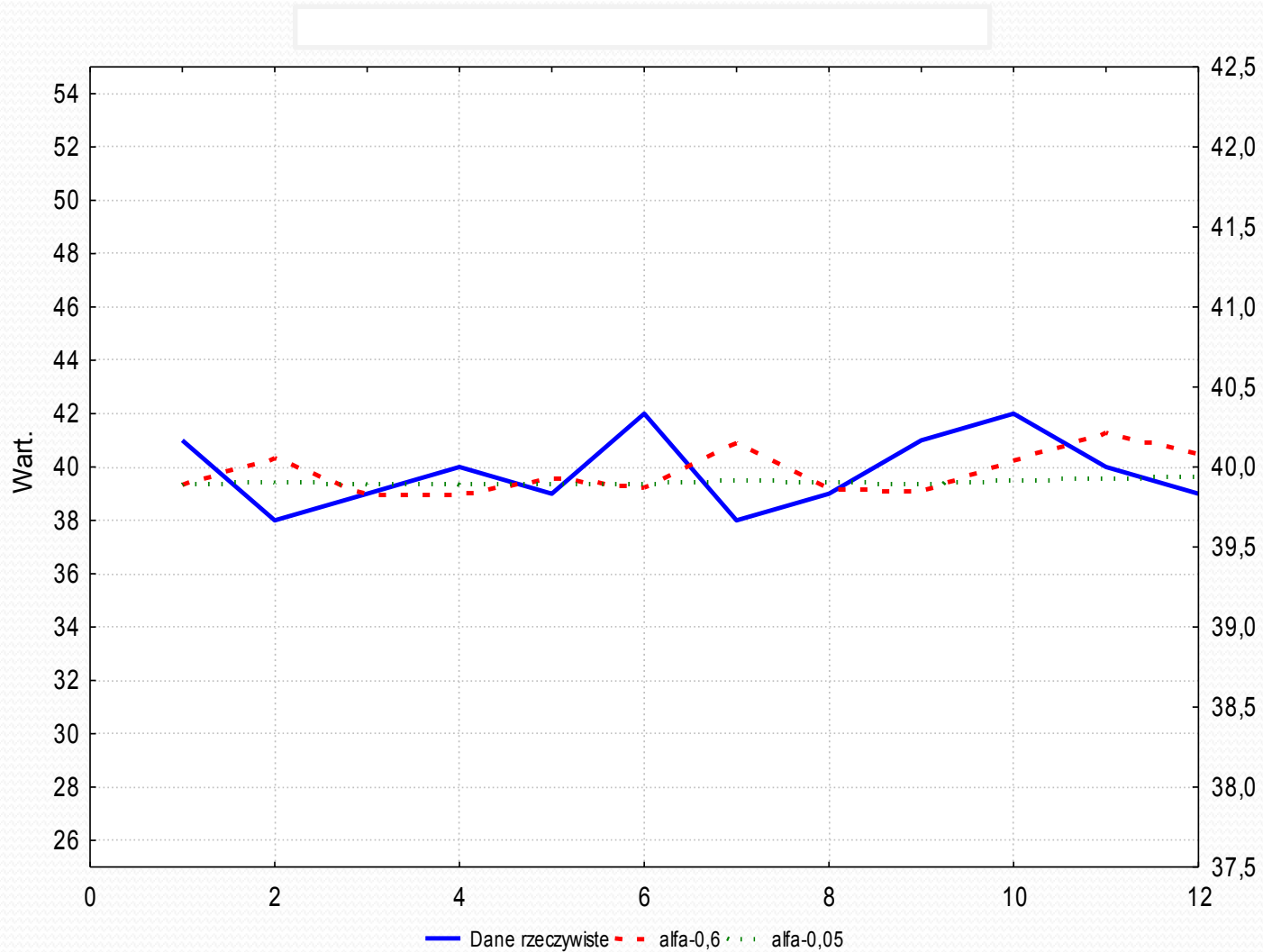
$$\Psi_t = \frac{1}{11} \cdot 35,1 = 3,19\%$$

$$\Psi_t = \frac{1}{11} \cdot 45,12 = 4,10\%$$

$$y_{13}^* = 0,05 \cdot 39 + (1 - 0,05) \cdot 39,61 = 39,58 \quad [tys.os.]$$

Obliczyć prognozę dla $\alpha=0,2$ wraz z błędami.

4. PROSTY MODEL WYGŁADZANIA WYKŁADNICZEGO - przykład



5. LINIOWY MODEL WYGŁADZANIA WYKŁADNICZEGO HOLTA

Model ten znajduje zastosowanie w przypadku gdy w szeregu czasowym występują trend oraz wahania przypadkowe. Jest to model dwurównaniowy, w którym do opisu tendencji używa się wielomianu stopnia pierwszego. Model ten jest bardziej elastyczny ze względu na występowanie dwóch parametrów wygładzania: α oraz β .

$F_{t-1} = \alpha y_{t-1} + (1-\alpha)(F_{t-2} + S_{t-2})$ Służy do wyznaczenia wygładzonych wartości szeregu czasowego

$S_{t-1} = \beta(F_{t-1} - F_{t-2}) + (1-\beta)S_{t-2}$ Służy do wyznaczenia wygładzonych wartości przyrostu trendu

F_{t-1} Ocena wartości średniej w okresie t-1

S_{t-1} Przyrostu trendu w okresie t-1

α, β Parametry modelu o wartościach z przedziału [0;1]

Równanie prognozy na okres $t > n$ ma postać: $y_t^* = F_n + (t - n)S_n$

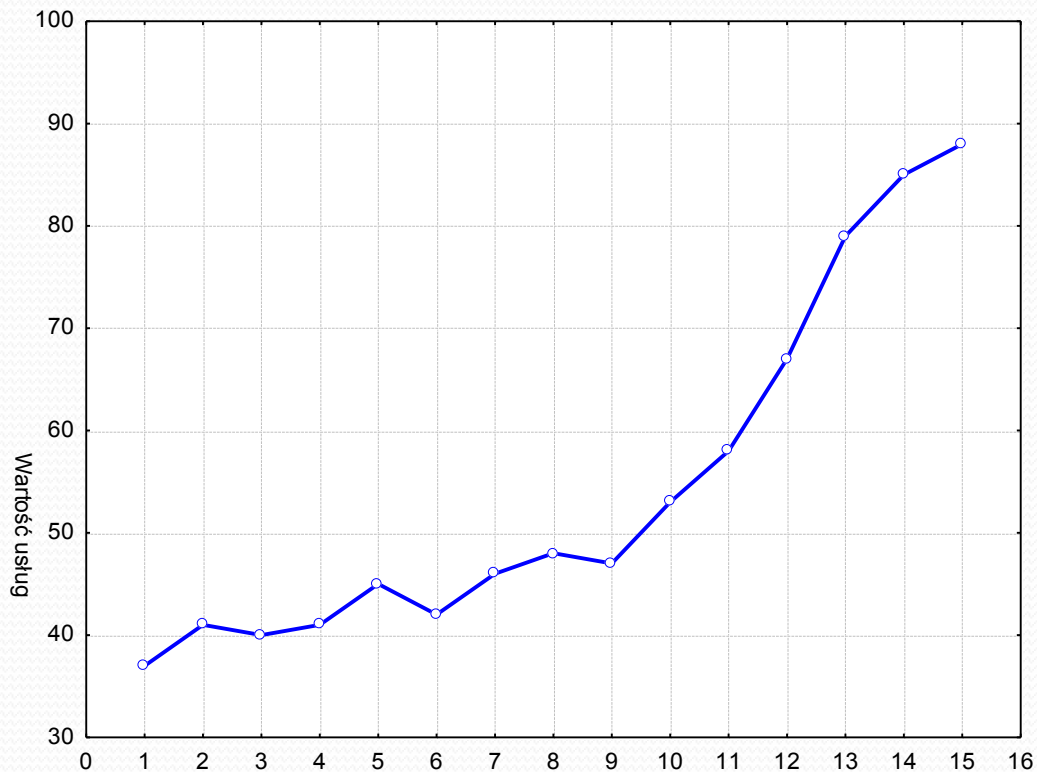
Do budowy modelu Holta potrzebne są początkowe wartości F i S, (F_1 i S_1).

Propozycja	F1	S1
1	Y1	0
2	Y1	Y2-Y1
3	Wyraz wolny liniowej funkcji trendu Oszacowanej na podstawie próbki wstępnej	Współczynnik kierunkowy liniowej funkcji trendu oszacowany na podstawie próbki wstępnej

5. LINIOWY MODEL WYGŁADZANIA WYKŁADNICZEGO HOLTA - przykład

Wartość usług w pewnym przedsiębiorstwie przedstawia szereg:
37, 41, 40, 41, 45, 42, 46, 48, 47, 53, 58, 67, 79, 85, 88.

1) Wyznaczyć prognozę wartości usług w kolejnym okresie.



Za wartości początkowe przyjęto:

$$F_1 = y_1 = 37$$

$$S_1 = y_2 - y_1 = 41 - 37 = 4$$

$$\alpha = 0,95$$

$$\beta = 0,45$$

5. LINIOWY MODEL WYGŁADZANIA WYKŁADNICZEGO HOLTA - przykład

$$F_{t-1} = \alpha y_{t-1} + (1 - \alpha)(F_{t-2} + S_{t-2})$$

$$S_{t-1} = \beta(F_{t-1} - F_{t-2}) + (1 - \beta)S_{t-2}$$

t	Yt	Ft $\alpha = 0,95$	S _t $\beta = 0,45$	$y_t^* = F_{t-1} + S_{t-1}$	$(y_t - y_t^*)^2$
1	37	37	4	-	-
2	41	$0,95 * 41 + (1 - 0,95) * (37,0 + 4,0) = 41,0$	$0,45 (41 - 37) + (1 - 0,45) * 4 = 4$	$37 + 4 = 41$	-
3	40	$0,95 * 40 + (1 - 0,95) * (41,0 + 4,0) = 40,3$	$0,45 (40,3 - 41) + (1 - 0,45) * 4 = 1,9$	$41 + 4 = 45$	25
4	41	$0,95 * 41 + (1 - 0,95) * (40,3 + 1,9) = 41,1$	$0,45 (41,1 - 40,3) + (1 - 0,45) * 1,9 = 1,4$	$40,3 + 1,9 = 42,2$	1,2
5	45	$0,95 * 45 + (1 - 0,95) * (41,1 + 1,4) = 44,9$	$0,45 (44,9 - 41,1) + (1 - 0,45) * 1,4 = 2,5$	$41,1 + 1,4 = 42,5$	6,5
6	42	$0,95 * 42 + (1 - 0,95) * (44,9 + 2,5) = 42,3$	$0,45 (42,3 - 44,9) + (1 - 0,45) * 2,5 = 0,2$	$44,9 + 2,5 = 47,4$	28,6
7	46	$0,95 * 46 + (1 - 0,95) * (42,3 + 0,2) = 45,8$	$0,45 (45,8 - 42,3) + (1 - 0,45) * 0,2 = 1,7$	$42,3 + 0,2 = 42,5$	12,5
8	48	$0,95 * 48 + (1 - 0,95) * (45,8 + 1,7) = 48,0$	$0,45 (48 - 45,8) + (1 - 0,45) * 1,7 = 1,9$	$45,8 + 1,7 = 47,5$	0,2
9	47	$0,95 * 47 + (1 - 0,95) * (48,0 + 1,9) = 47,1$	$0,45 (47,1 - 48) + (1 - 0,45) * 1,9 = 0,7$	$48 + 1,9 = 49,9$	8,3
10	53	$0,95 * 53 + (1 - 0,95) * (47,1 + 0,7) = 52,7$	$0,45 (52,7 - 47,1) + (1 - 0,45) * 0,7 = 2,9$	$47,1 + 0,7 = 47,8$	26,8
11	58	$0,95 * 58 + (1 - 0,95) * (52,7 + 2,9) = 57,9$	$0,45 (57,9 - 52,7) + (1 - 0,45) * 2,9 = 3,9$	$52,7 + 2,9 = 55,6$	5,6
12	67	$0,95 * 67 + (1 - 0,95) * (57,9 + 3,9) = 66,7$	$0,45 (66,7 - 57,9) + (1 - 0,45) * 3,9 = 6,1$	$57,9 + 3,9 = 61,8$	27,2
13	79	$0,95 * 79 + (1 - 0,95) * (66,7 + 6,1) = 78,7$	$0,45 (78,7 - 66,7) + (1 - 0,45) * 6,1 = 8,8$	$66,7 + 6,1 = 72,8$	37,6
14	85	$0,95 * 85 + (1 - 0,95) * (78,7 + 8,8) = 85,1$	$0,45 (85,1 - 78,7) + (1 - 0,45) * 8,8 = 7,7$	$78,7 + 8,8 = 87,5$	6
15	88	$0,95 * 88 + (1 - 0,95) * (85,1 + 7,7) = 88,2$	$0,45 (88,2 - 85,1) + (1 - 0,45) * 7,7 = 5,6$	$85,1 + 7,7 = 92,8$	23,3
16	-			$88,2 + 5,6 = 93,8$	
					209

Wykonać obliczenia dla pierwszych 4 przypadków gdzie alfa=0,45 a beta 0,95.

5. LINIOWY MODEL WYGŁADZANIA WYKŁADNICZEGO HOLTA - przykład

$$F_{t-1} = \alpha y_{t-1} + (1-\alpha)(F_{t-2} + S_{t-2}) \quad y_t^* = F_n + (t-n)S_n$$
$$S_{t-1} = \beta(F_{t-1} - F_{t-2}) + (1-\beta)S_{t-2}$$

$$F_n = F_{15} = 0,95 \cdot 88 + (1-0,95) \cdot (85,1 + 7,7) = 88,2$$

$$S_n = S_{15} = 0,45 \cdot (88,2 - 85,1) + (1-0,45) \cdot 7,7 = 5,6$$

$$y_{16}^* = 88,2 + (16-15) \cdot 5,6 = 93,8$$

Do oceny dopuszczalności prognozy można posłużyć się błędami prognoz wygasłych. Obliczono średni kwadratowy błąd ex post prognoz wygasłych.

$$s^* = \left[\frac{1}{n-k} \sum_{t=k+1}^n (y_t - y_t^*)^2 \right]^{\frac{1}{2}} \quad \text{Błąd względny } \frac{s^*}{y_{16}^*} \cdot 100\% = \frac{4}{39,8} \cdot 100\% = 4,3\%$$

$$s^* = \left[\frac{1}{13} \sum_{t=3}^{15} (y_t - y_t^*)^2 \right]^{\frac{1}{2}} = \left(\frac{209}{13} \right)^{0,5} = 4$$