Metody prognozowania:

Szeregi czasowe

Dr inż. Sebastian Skoczypiec

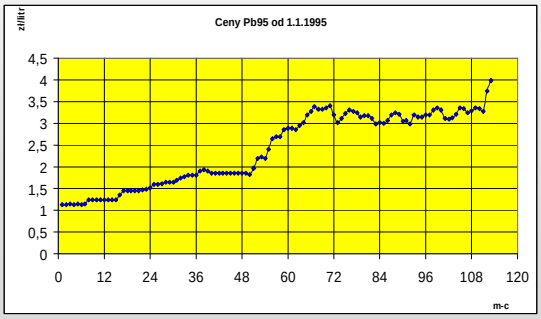
ver. 11.20.2009

Co to jest szereg czasowy?

Szereg czasowy: uporządkowany zbiór wartości badanej cechy lub wartości określonego zjawiska, zaobserwowanych w różnych momentach czasu:

$$\{y_t,\ t=1,\ 2,\ 3,....T\},$$

gdzie:
y_t - wyniki obserwacji,
t - momenty czasu , w których obserwowano wartości zmiennej.



Składowe szeregów czasowych (1):

- ✓ Składowa systematyczna: efekt oddziaływań stałego zestawu czynników na zmienną prognozowaną
- ✓ Składowa przypadkowa (składnik losowy, wahania przypadkowe):

Składowe szeregów czasowych (3):

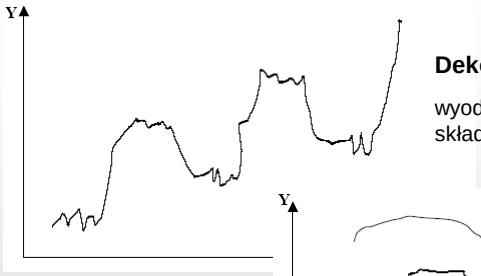
Tendencja rozwojowa (trend)
- długookresowa skłonność do jednokierunkowych zmian (wzrostu lub spadku) wartości badanej zmiennej.

Stały (przeciętny) poziom
- występuje, gdy w szeregu czasowym nie ma tendencji rozwojowej, wartości zmiennej prognozowanej oscylują zaś wokół pewnego (stałego poziomu).

Wahania cykliczne
- Długookresowe rytmiczne wahania wartości wokół tendencji rozwojowej lub stałego (przeciętnego) poziomu.

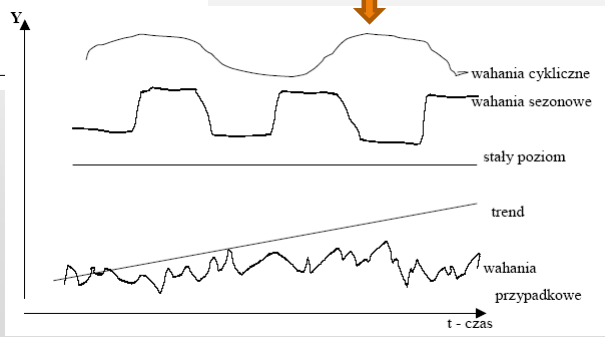
Wahania sezonowe
- wahaniami wartości obserwowanej zmiennej wokół tendencji rozwojowej lub stałego (przeciętnego) poziomu tej zmiennej. Mają skłonność do powtarzania się w określonym czasie

Składowe szeregów czasowych (2):



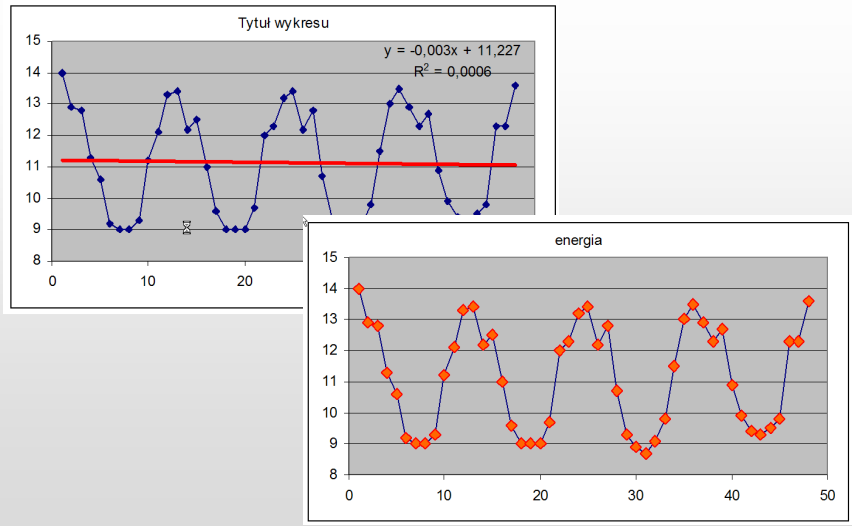
Dekompozycja szeregu czasowego:

wyodrębnianie poszczególnych składowych danego szeregu czasowego



Składowe szeregów czasowych (3):

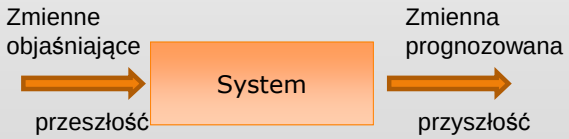
- ✓ Dekompozycja często na podstawie oceny wzrokowej
- ✓ Nie zawsze występują wszystkie składowe
- ✓ Obserwacje nietypowe: Wahania losowe czy może tylko błąd rejestracji danych ? A może gwałtowna zmiana rozpatrywanego zjawiska ? Często tego typu dane eliminujemy z analizowanego szeregu czasowego.
- ✓ Punkty zwrotne: zmiana kierunku tendencji rozwojowej -> wymaga użycia specjalnych metod prognozowania lub może nawet udaremnić prognozowanie.



Model szeregu czasowego (1)

W przypadku prognozowania na podstawie szeregu czasowego przetwarzanie informacji o przeszłości następuje przez **budowę odpowiedniego <u>modelu formalnego</u>**, przejście zaś od informacji przetworzonej do prognozy przez wybór <u>reguły prognozowania</u>.

Modelem szeregu czasowego służącym do określenia przyszłej wartości zmiennej prognozowanej Y w momencie (okresie) prognozowania t tj. y_t^* jest model formalny, którego zmiennymi objaśniającymi mogą być tylko zmienne czasowe oraz przyszłe wartości lub prognozy zmiennej Y .



Model szeregu czasowego (2)

Prognoza zmiennej Y jest wartością funkcji f zależnej od czasu, przeszłych wartości i (lub) prognoz tej zmiennej:

$$p_t = f(t, y_{t-1}, \ldots\ldots, y_{t-d}, p_{t-1}, \ldots\ldots, p_{t-d}, \varepsilon_t)$$

gdzie:

$p_t, p_{t-1}, \ldots\ldots, p_{t-d}$ *- prognoza zmiennej Y wyznaczona na okres t, t-1,....., t-d*

$y_t, y_{t-1}, \ldots\ldots, y_{t-d}$ *- zaobserwowane wartości zmiennej Y w okresie t, t-1,....., t-d*

t – zmienna czasowa, d – opóźnienie ε *- składnik losowy*

W zależności od przyjętych założeń odnośnie wpływy poszczególnych składowych szeregu czasowego za prognozowaną zmienną oraz wzajemnych relacji tych składowych modele szeregów czasowych mogą mieć postać:

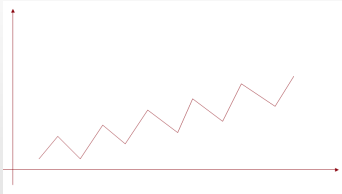
- ✓ addytywną
- ✓ multiplikatywną

Model szeregu czasowego (3)

✓ <u>Model addytywny</u>: zakłada się że obserwowane wartości zmiennej prognozowanej są sumą (wszystkich lub niektórych) składowych szeregu czasowego. Dla zmiennej czasowej, postać takiego modelu może być następująca:

$$y_t = f(t) + g(t) + h(t) + \varepsilon_t$$

$$y_t = const + g(t) + h(t) + \varepsilon_t$$



f(t) – funkcja czasu, charakteryzująca tendencję rozwojową (funkcją trendu);

g(t) – funkcja czasu, charakteryzująca wahania sezonowe;

h(t) – funkcja czasu, charakteryzująca wahania cykliczne

ε- zmienna losowa (składnik losowy)

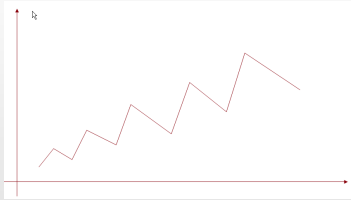
const – stały (średni) poziom prognozowanej zmiennej

Model szeregu czasowego (4)

✓ Model multiplikatywny: przyjmuje się że obserwowane wartości zmiennej prognozowanej są iloczynem składowych szeregu czasowego. Dla zmiennej czasowej, postać takiego modelu może być następująca:

$$y_t = f(t) \cdot g(t) \cdot h(t) \cdot \varepsilon_t$$

$$y_t = const \cdot g(t) \cdot h(t) \cdot \varepsilon_t$$



f(t) – funkcja czasu, charakteryzująca tendencję rozwojową (funkcją trendu);

g(t) – funkcja czasu, charakteryzująca wahania sezonowe;

h(t) – funkcja czasu, charakteryzująca wahania cykliczne

ε- zmienna losowa (składnik losowy)

const – stały (średni) poziom prognozowanej zmiennej

Medel szeregu czasowego (5)

Możliwe jest takie stosowanie różnego rodzaju postaci modeli mieszanych **„**addytywnomultiplikatywnych**”,** jednak są bardziej skomplikowane i trudniejsze zarówno do estymacji jak też interpretacji

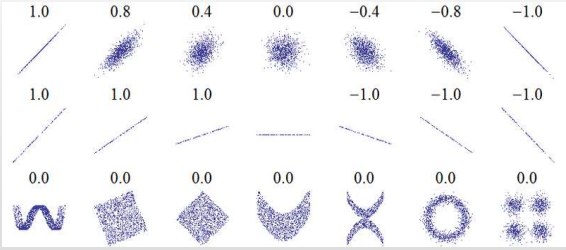
Prognoza na podstawie modelu szeregu czasowego to najczęściej ekstrapolacja funkcji trendu, a następnie jej korekcie uwzględniające oddziaływanie wahań okresowych na prognozowaną zmienną.

Identyfikacja składowych szeregu

Trend: istotność współczynnika korelacji r Pearsona

$$r_{yx} = \frac{\mathrm{cov}_{yx}}{S_x S_y}$$

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \overline{y})^2}},$$



Współczynnik korelacji – liczba określająca w jakim stopniu zmienne są współzależne.

Identyfikacja składowych szeregu

Sezonowość:
jednoczynnikowa analiza wariancji (ANOVA) – hipoteza o równości wielu wartości przeciętnych

Przy analizie wahań sezonowych taka hipoteza głosi, że odchylenia od trendu są we wszystkich miesiącach, średnio rzecz biorąc, takie same, a hipoteza alternatywna, że przynajmniej jeden miesiąc różni się istotnie od innych. Przy stosowaniu ANOVA potrzebna jest zmienna wskazująca numer miesiąca, pozwalająca programowi znaleźć odchylenia od trendu we wszystkich styczniach, wszystkich lutych itd. Wprowadziliśmy zmienną do zbioru danych, nazywając ją NR_MIES.

Test analizy wariancji wymaga spełnienia dwóch założeń: badana zmienna powinna m w każdej grupie rozkład normalny, a wariancje w grupach powinny być takie same. Założenie normalności sprawdzimy przy pomocy testu Shapiro-Wilka. Jest on dostępn module Podstawowe statystyki. W części Statystyki opisowe zaznaczamy opcję Test W Shapiro-Wilka i po naciśnięciu klawisza Tabele liczebności oprócz szeregu rozdzielcze otrzymujemy też wartość prawdopodobieństwa testowego w teście normalności. Wcześniej trzeba zdefiniować testowaną zmienną (ODCH_ADD). Przy pomocy klawisz Select cases ("wybierz przypadki") wybieramy kolejno poszczególne miesiące, wpisują v8=1 dla stycznia, potem v8=2 dla lutego, v8=3 dla marca itd. Wyniki, które otrzymujem w ten sposób, przedstawia poniższa tabela:

Rodzaje modeli szeregów czasowych

Typ szeregu czasowego	Modele szeregów czasowych
Szeregi stacjonarne	Naiwne Średniej ruchomej Prosty model wygładzania wykładniczego Modele ARMA i ARIMA
Szeregi wykazujące tendencję rozwojową (trend)	Wygładzanie wykładnicze drugiego rzędu Model liniowy Holta Modele analityczne Modele adaptacyjne
Szeregi wykazujące sezonowość	Model Wintersa Metoda wskaźników Analiza harmoniczna

Modele szeregów stacjonarnych (1):

między poszczególnymi wartościami zmiennej występują niewielkie wahania, których łączne odchylenia od pewnego stałego poziomu są niewielkie i niwelują się.

W przypadku stwierdzenia stacjonarności danego zjawiska można zastosować takie metody prognozowania jak:

- ✓ metody naiwne,
- ✓ metoda średniej ruchomej,
- ✓ wygładzanie wykładnicze pierwszego rzędu,
- ✓ modele ARMA i ARIMA.

Metody naiwne

Mogą być stosowane w przypadku stwierdzenia niewielkich wahań w szeregu zmiennej prognozowanej i wykorzystane w prognozowaniu krótkookresowym, obejmującym jeden okres naprzód.

Zakładają, iż nie nastąpią zmiany w sposobie oddziaływania czynników kształtujących wartości zmiennej prognozowanej:

$$p_t = y_{t-1}$$

Metody naiwne są szybkie i tanie w zastosowaniu, jednak jakość prognoz wyznaczonych z ich użyciem jest zwykle niska.

Przykłady:

✓ *sprzedaż przedsiębiorstwa w następnym kwartale będzie na dotychczasowym poziomie,*

✓*zysk wzrośnie w tym samym stopniu co ubiegłym miesiącu.*

Metody naiwne

Metoda naiwna

Algorytm prognozowania:

$$y_t^* = y_{t-1}$$

gdzie: y_t^* - prognoza zmiennej Y dla momentu t, y_{t-1} - obserwacja rzeczywistej wartości zmiennej *Y* dla chwili *t-1*.

Tab. Charakterystyka metody naiwnej prognozowania

Grupa metod	Prognozowanie na podstawie szeregu czasowego
Składowe szeregu czasowego	Stały (przeciętny) poziom i wahania przypadkowe
Przesłanki metody	Nie nastąpią zmiany w sposobie oddziaływania czynników określających zmienną prognozowaną, niewielkie wahania przypadkowe
Postawa, zasada, reguły prognostyczna	Postawa pasywna, zasada *status quo*, reguła podstawowa prognozowania
Horyzont prognozy	Jeden okres (moment)
Ocena prognozy	Błędy ex post
Zalety metody	Prosty algorytm, łatwość zrozumienia, szybkie i tanie prognozowanie
Wady metody	Niska jakość prognozy, ocena jedynie za pomocą błędów *ex post*

Metoda średniej ruchomej (1)

- Polega na zastępowaniu danych empirycznych dla kolejnych okresów **średnimi poziomami z okresu badanego i kilku okresów sąsiednich.**
- Średnie ruchome mogą być obliczane z parzystej bądź nieparzystej liczby kolejnych wyrazów szeregu. W pierwszym przypadku mówimy o średniej ruchomej scentrowanej, w drugim o zwykłej średniej ruchomej.
- Metoda średniej ruchomej (zwanej też łańcuchową) może być stosowana zarówno do wygładzania szeregu czasowego, jak i do prognozowania.

Metoda średniej ruchomej (2)

Metoda średniej ruchomej prostej

Algorytm prognozowania:

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

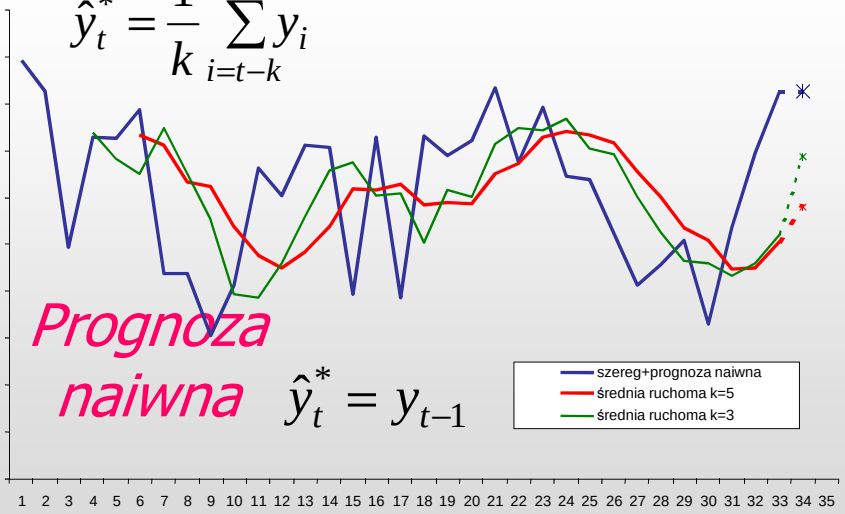
gdzie: y_t^* - prognoza zmiennej *Y* dla momentu *t*, y_i - obserwacja rzeczywistej wartości zmiennej *Y* dla chwili *i*, *k* – liczba ruchomych składników szeregu czasowego.

Tab. Charakterystyka metody średniej ruchomej prostej prognozowania

Grupa metod	Prognozowanie na podstawie szeregu czasowego
Składowe szeregu czasowego	Stały (przeciętny) poziom i wahania przypadkowe
Przesłanki metody	Nie nastąpią zmiany w sposobie oddziaływania czynników określających zmienną prognozowaną, mogą wystąpić duże wahania przypadkowe
Postawa, zasada, reguły prognostyczna	Postawa pasywna, zasada *status quo*, reguła podstawowa prognozowania
Horyzont prognozy	Prognoza krótkookresowa
Ocena prognozy	Ocena dopuszczalności prognozy za pomocą średniego kwadratowego błędu s^* prognozy lub średniego błędu ψ^* *ex post*
Zalety metody	Względnie prosty algorytm, łatwość zrozumienia, względnie szybkie i tanie prognozowanie
Wady metody	Konieczność doboru stałej *k* (minimalizacja błędów), konieczność przechowywania dużej ilości danych dla dużego *k*

Średnia ruchoma prosta

$$\hat{y}_t^* = \frac{1}{k} \sum_{i=t-k}^{t-1} y_i$$



Prognoza naiwna $\hat{y}_t^* = y_{t-1}$

Metoda średniej ruchomej (3)

Średnia ruchoma szeregu chronologicznego: $y_1, y_2, \ldots, y_n$ jest wielkością, którą obliczamy według ogólnej formuły:

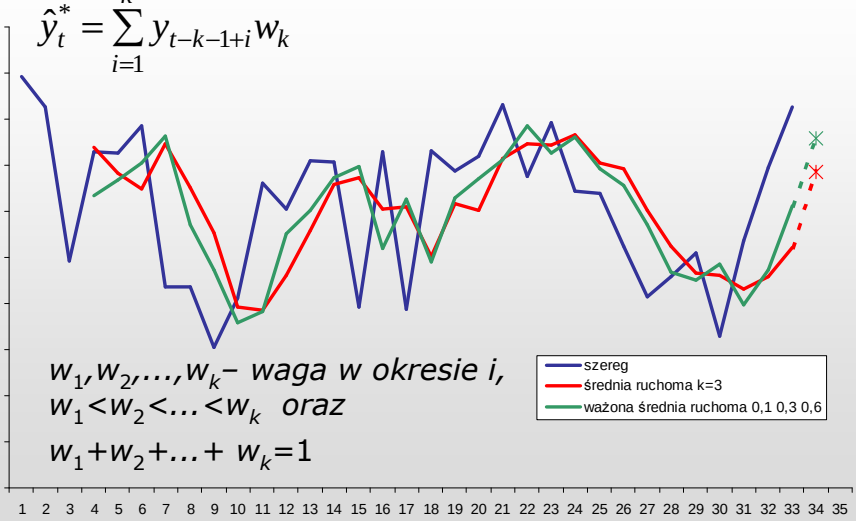
$$\overline{y}_i = \sum_{j=1}^{m} w_j y_{i-j} \qquad (i=m, m+1, \ldots, n),$$

gdzie: w_j – ustalony zbiór wag spełniający warunek $\sum_{j=1}^{m} w_j = 1$

m – liczba odpowiadająca podzbiorowi ($m < n$) jednostek czasu.

Średnia ruchoma ważona liniowo

$$\hat{y}_t^* = \sum_{i=1}^{k} y_{t-k-1+i} w_k$$



$w_1, w_2, \ldots, w_k$ – *waga w okresie i,*

$w_1 < w_2 < \ldots < w_k$ *oraz*

$w_1 + w_2 + \ldots + w_k = 1$

Metoda średniej ruchomej (4)

✓ Prognozowanie:
przyjmuje się, że wartość zmiennej prognozowanej w następnym momencie lub okresie będzie równa średniej arytmetycznej z *k* ostatnich wartości tej zmiennej:

- przy zastosowaniu **średniej ruchomej prostej**: $$p_t = \frac{1}{k}\sum_{i=t-k}^{t-1} y_i ,$$

- przy zastosowaniu **średniej ruchomej ważonej**: $$p_t = \sum_{i=t-k}^{t-1} y_i w_{i-t+k+1},$$

$w_{i-t+k+1}$ – waga nadana przez prognostę wartości zmiennej prognozowanej w okresie *i*,

k – stała wygładzania

Metoda średniej ruchomej (5)

Metoda średniej ruchomej ważonej

Algorytm prognozowania:

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

gdzie: y_t^* - prognoza zmiennej *Y* dla momentu *t*, y_i - obserwacja rzeczywistej wartości zmiennej *Y* dla chwili *i*, *k* – liczba ruchomych składników szeregu czasowego, w_i - współczynniki wagowe z przedziału [0, 1].

Tab. Charakterystyka metody średniej ruchomej ważonej prognozowania

Grupa metod	Prognozowanie na podstawie szeregu czasowego
Składowe szeregu czasowego	Stały (przeciętny) poziom i wahania przypadkowe
Przesłanki metody	Nie nastąpią zmiany w sposobie oddziaływania czynników określających zmienną prognozowaną, mogą wystąpić duże wahania przypadkowe
Postawa, zasada, reguły prognostyczna	Postawa pasywna, zasada *status quo*, reguła podstawowa prognozowania
Horyzont prognozy	Prognoza krótkookresowa
Ocena prognozy	Ocena dopuszczalności prognozy za pomocą średniego kwadratowego błędu *s** prognozy *ex post*
Zalety metody	Względnie prosty algorytm, łatwość zrozumienia, względnie szybkie i tanie prognozowanie
Wady metody	Konieczność doboru stałej *k* i współczynników wagowych *w* (minimalizacja błędów), konieczność przechowywania dużej ilości danych danych dla dużego *k*

Prosty model wygładzania wykładniczego (1) tzw. Metoda Browna

✓ Jest rozszerzeniem idei średniej ruchomej i polega na:

1. wygładzeniu oryginalnego szeregu tak, jak robi to średnia ruchoma

1. użycia otrzymanego szeregu do uzyskania przyszłych wartości zmiennej.

✓ Istotne jest zwiększenie wpływu ostatnich wartości szeregu na prognozę, w stosunku do wpływu bardziej odległych obserwacji. Największa waga jest nadana bieżącej obserwacji, mniejsza waga poprzedniej obserwacji i tak dalej. Wagi zmniejszają się geometrycznie w miarę cofania się w czasie

✓ Przydatna do prognozowania szeregów nie mających wyraźnego trendu i wahań sezonowych.

Prosty model wygładzania wykładniczego (2)

✓ Prognozowanie polega na wyznaczeniu prognozy na okres *t*, która jest korektą prognozy na okres *t-1* po porównaniu jej z wynikiem rzeczywistym y_{t-1}.

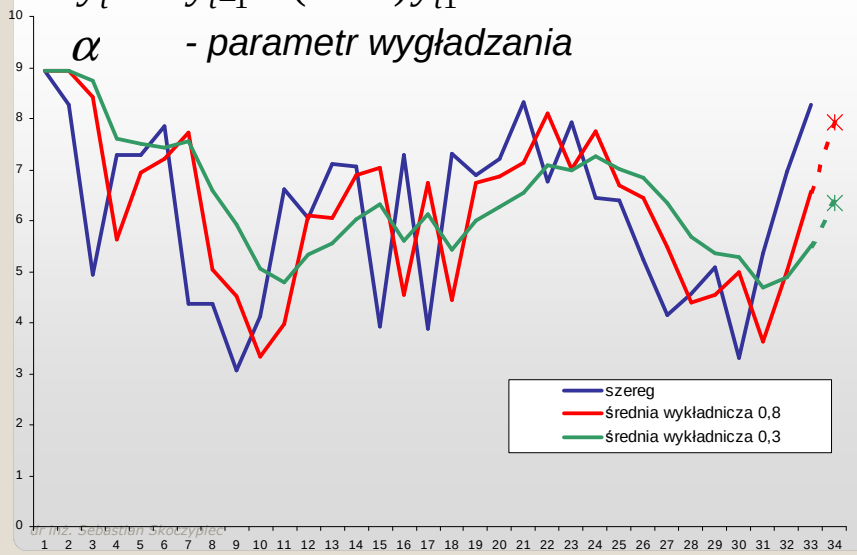
✓ Zgodnie z rekurencyjną procedurą każda nowa wygładzona wartość jest obliczana jako średnia ważona bieżącej obserwacji i poprzedniej obserwacji wygładzonej (prognozy); poprzednia wygładzona obserwacja była obliczona znów z poprzedniej wartości obserwowanej i wygładzonej wartości przed poprzednią obserwacją itd.

$p_1 = y_1$
$p_2 = \alpha\, y_1 + (1-\alpha)p_1$
$p_3 = \alpha\, y_2 + (1-\alpha)p_2$

Wygładzanie wykładnicze

$$\hat{y}_t^* = \alpha y_{t-1} + (1-\alpha)\hat{y}_{t1}$$

α - parametr wygładzania



Prosty model wygładzania wykładniczego (4)

- α = 0,2 (obserwacje teraźniejsze) 1-α = 0,8 (obserwacje przeszłe) Większą wagę przywiązuje się do obserwacji przeszłych niż obecnych.

- α = 0,7 (obserwacje teraźniejsze) 1- α = 0,3 (obserwacje przeszłe) Większą wagę przywiązuje się do obserwacji obecnych niż przeszłych.

- Jedynie dla wartości α = 0,5 takie same wagi są zarówno dla obserwacji historycznych jak i teraźniejszych.

- Im α bliższe **1**, tym większa wagę przykładamy do najnowszych obserwacji, a budowana prognoza będzie uwzględniała w bardzo wysokim stopniu błędy ex post prognoz poprzednich.

Metoda prostego uśredniania wykładniczego

Algorytm prognozowania:

$$y_t^* = \alpha \cdot y_{t-1} + (1-\alpha) \cdot y_{t-1}^* = y_{t-1}^* + \alpha \cdot \left(y_{t-1} - y_{t-1}^*\right)$$

gdzie: y_t^* - prognoza zmiennej *Y* dla momentu *t*, y_{t-1} - obserwacja rzeczywistej wartości zmiennej *Y* dla chwili *t-1*, *α* – parametr uśredniania (wygładzania) z przedziału [0, 1].

Grupa metod	Prognozowanie na podstawie szeregu czasowego
Składowe szeregu czasowego	Stały (przeciętny) poziom i wahania przypadkowe
Przesłanki metody	Nie nastąpią zmiany w sposobie oddziaływania czynników określających zmienną prognozowaną, mogą wystąpić duże wahania przypadkowe
Postawa, zasada, reguły prognostyczna	Postawa pasywna, zasada *status quo*, reguła podstawowa prognozowania
Horyzont prognozy	Prognoza krótkookresowa
Ocena prognozy	Ocena dopuszczalności prognozy za pomocą średniego kwadratowego błędu *s** prognozy *ex post*
Zalety metody	Względnie prosty algorytm, łatwość zrozumienia, względnie szybkie i tanie prognozowanie, uwzględnienie w prognozie wartości ostatniego błędu *ex post*
Wady metody	Konieczność i trudność doboru parametru *α* wygładzania (minimalizacja błędów)

Prosty model wygładzania wykładniczego (5)

W praktyce parametr wygładzania jest często wybierany przez poszukiwanie sieciowe przestrzeni parametrów; to znaczy wypróbowuje się różne rozwiązania zaczynając, na przykład, od wartości **α = 0.1 do α = 0.9, przy przyroście 0.1.**

Następnie wybiera się a tak, aby otrzymać najmniejszą sumę kwadratów (lub najmniejszy średni kwadrat) różnicy pomiędzy wartościami empirycznymi a wartościami prognozowanymi na jeden okres naprzód; jest to tzw. **średniokwadratowy błąd ex post, (ex post MSE).**

Prosty model wygładzania wykładniczego (6)

W oparciu o wygładzony szereg p_t wyznaczamy **prognozy dla danych historycznych** y_t:

$$Y_{t+j}^{P} = p_t + \Delta p_t \times j = p_t + (p_t - p_{t-1}) \times j = 2p_t j - p_{t-1} j$$

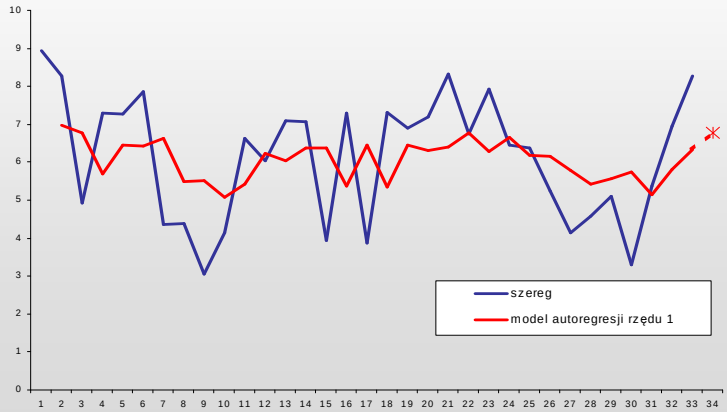
t – okres z którego pochodzi obserwacja,
j – wyprzedzenie prognozy tzn. na ile okresów wcześniej chcemy prognozować niż jesteśmy obecnie.

Prognozowanie w przyszłość:
Jeżeli jesteśmy w okresie bieżącym i dokonamy prognozy z wyprzedzeniem j-okresowym, to uzyskamy prognozę wartości przyszłej na podstawie danych historycznych.

Model autoregresji

$$\hat{y}_t^* = a_0 + a_1 y_{t-1} + \ldots + a_p y_{t-p}$$

$a_0, a_1, \ldots a_p$ *- oceny parametrów wyznaczone MNK*



Modele szeregów czasowych z trendem (1)

✓Trend (tendencja rozwojowa):
w statystyce funkcja matematyczną opisującą zmiany wielkości w czasie.

✓Wyznaczenie trendu:
znalezienie odpowiedniej krzywej, możliwie dokładnie pasującej do wartości, które badany proces przyjmował w minionych okresach.

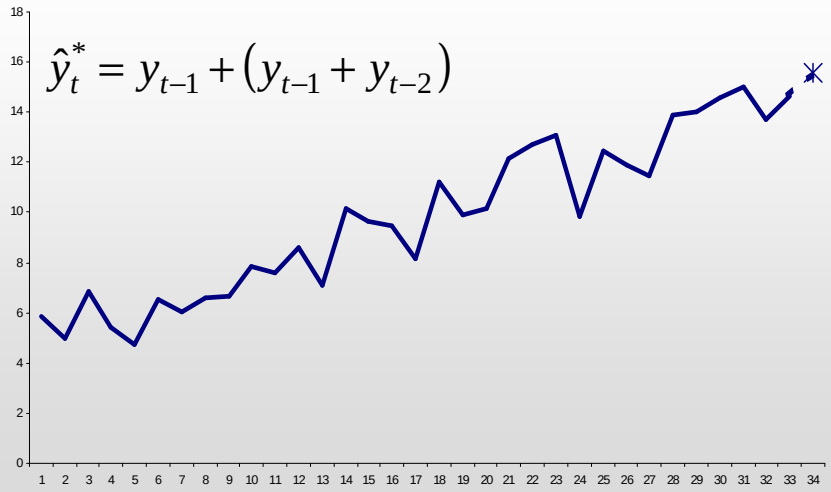
⇓

Konstrukcja jedno-równaniowego modelu matematycznego, w którym rolę jedynej zmiennej objaśniającej pełni czas.

1. Wybór postaci analitycznej funkcji trendu.
2. Szacowanie parametrów funkcji trendu.
3. Ocena jakości otrzymanego modelu.

Metoda naiwna

$$\hat{y}_t^* = y_{t-1} + (y_{t-1} + y_{t-2})$$



Modele szeregów czasowych z trendem (2)

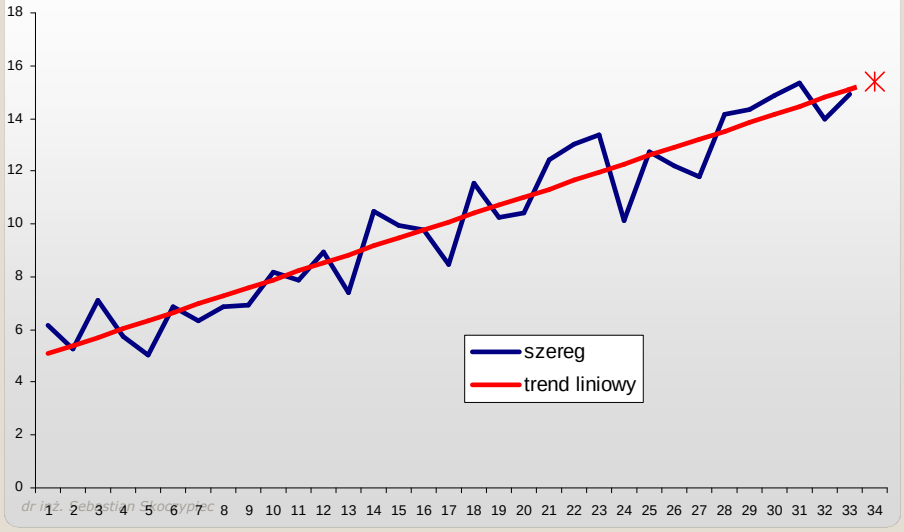
Wybór postaci analitycznej funkcji może być oparty na teoretycznych przesłankach dotyczących mechanizmu rozwojowego zmiennej prognozowanej.

Najczęstsze postacie funkcji:

- ✓ *funkcja liniowa* $y_t = \alpha + \beta t$
- ✓ *funkcję potęgową,* $y_t = \alpha t^{\beta}, \quad \beta > 1$
- ✓ *wykładniczą,* $y_t = e^{\alpha + \beta t} \qquad \beta > 0$
- ✓ *wielomian drugiego stopnia* $y_t = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \ldots + \alpha_k t^k$
- ✓ *funkcję logistyczną (0 - 1).*

Model trendu liniowego

$$\hat{y}_t = a_0 + a_1 t$$



Modele szeregów czasowych z trendem (3)

Pomocne przy wyborze postaci funkcji może być również obliczenie przyrostów zmiennej prognozowanej (objaśnianej) Y:

1. $(y_i - y_{i-1}) \approx \text{const}$ → **funkcją liniowa**.

1. $\Delta y_i - \Delta y_{i-1} \approx \text{const}$ → **funkcja wielomianowa drugiego stopnia.**

1. Przyrosty względne $\frac{\Delta y_i}{y_{i-1}} \approx const$ **wykładnicza funkcja trendu.**

Modele szeregów czasowych z trendem (4)

✓ Oszacowanie parametrów modelu:

Najczęściej wykorzystywana w tym celu jest klasyczna metoda najmniejszych kwadratów (najpopularniejsza, choć nie najlepsza, z metod wyznaczania linii trendu na podstawie zbioru danych w postaci par liczb).

Żądamy minimalizacji funkcji, która mierzy odchylenie zadanej zależności funkcyjnej od punktów doświadczalnych.

$$\hat{y}_i = a_0 + a_1 \cdot t_i$$

$$\begin{cases} \sum_{i=1}^{n} y_i = na_0 + a_1 \sum_{i=1}^{n} t_i, \\ \sum_{i=1}^{n} y_i t_i = a_0 \sum_{i=1}^{n} t_i + a_1 \sum_{i=1}^{n} t_i^2, \end{cases}$$

Modele szeregów czasowych z trendem (5)

Miara dopasowania	Interpretacja
Odchylenie standardowe resztowe $$s_\varepsilon = \sqrt{s_\varepsilon^2} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$	Pierwiastek kwadratowy z wariancji resztowej. Informuje o ile średnio różnią się wartości zmiennej objaśnianej od wartości oszacowanych na podstawie funkcji trendu
Współczynnik zbieżności $$\varphi^2 = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$	Wskazuje jaka część zmienności w wartościach zmiennej zależnej nie jest związana ze zmiennością zmiennej objaśniającej; im bliższy 0, tym lepiej dopasowana funkcja
Współczynnik determinacji $$R^2 = 1 - \varphi^2$$	Informuje jaka część zmienności zmiennej zależnej związana jest ze zmiennością zmiennej objaśniającej; im bliższy 1 tym lepiej dopasowana funkcja

Podwójne wygładzanie wykładnicze – Metoda HOLTA

- ✓ Przystosowane do szeregów czasowych ze składnikami sezonowymi i trendami.
- ✓ Ogólna idea polega na tym, że prognozy są obliczane nie tylko na podstawie kolejnych poprzednich obserwacji (jak w prostym wyrównywaniu wykładniczym), ale można także dodać niezależny trend i składnik sezonowy.
- ✓ Wiele empirycznych szeregów czasowych zawiera wahania sezonowe. (np. roczna sprzedaż komputerów osiąga prawdopodobnie szczyt w listopadzie i grudniu i być może latem) Wzorzec ten będzie się prawdopodobnie powtarzał co roku, chociaż względna wielkość wzrostu sprzedaży w grudniu może się powoli zmieniać z roku na rok.

Wygładzanie wykładnicze (5)

Trend liniowy, wykładniczy i gasnący:

Sprzedaż komputerów:

✓może przejawiać rosnący trend liniowy (np. co roku sprzedaż rośnie średnio o 1 milion €),

✓wzrost wykładniczy (np. co roku sprzedaż rośnie średnio o 30%)

✓lub trend gasnący (w pierwszym roku sprzedaż rośnie o 1 milion €; w drugim roku wzrost wynosi tylko 80%
✓Zeszłorocznego itd.).

Wygładzanie wykładnicze (5)

Każdy typ trendu ujawnia się w szeregu w charakterystyczny sposób.

Trend może się powoli zmieniać w czasie i znowu może być sensowne wygładzenie składnika trendu przy pomocy oddzielnego parametru tzw. parametru wyrównania trendu.

$$F_t = \alpha \cdot y_t + (1-\alpha)\cdot(F_{t-1} + T_{t-1})$$
$$T_t = \beta \cdot (F_t - F_{t-1}) + (1-\beta)\cdot T_{t-1}$$

$$y^*_\tau = F_n + (\tau - n)\cdot T_n$$
$$\tau > n$$

Ft - wygładzona wartość zmiennej prognozowanej

Tt (St) - wygładzona wartość przyrostu trendu

alfa - stała wygładzania poziomu zmiennej F

beta - stała wygładzania współczunnika trendu T

n - numer ostatniej obserwacji

Wygładzanie wykładnicze (6)

Jeśli do procesu wyrównywania wykładniczego zostanie włączony składnik trendu, to dla każdego momentu czasu zostaje obliczony niezależny **składnik trendu T_t.**

Zostaje on zmodyfikowany jako funkcja błędu prognozy i odpowiedniego parametru.

T_t = f(błąd prognozy, γ)

Jeśli **γ = 0, to składnik trendu jest stały dla** wszystkich wartości szeregu czasowego (i dla wszystkich prognoz).

Jeśli **γ = 1, to składnik trendu zostaje "maksymalnie" zmodyfikowany z obserwacji na** obserwację przez odpowiedni błąd prognozy.

Wartości parametrów, które znajdują się pośrodku reprezentują mieszanki tych dwóch wartości skrajnych.

Model Holta

$$\hat{y}_t = S_t + F_t$$

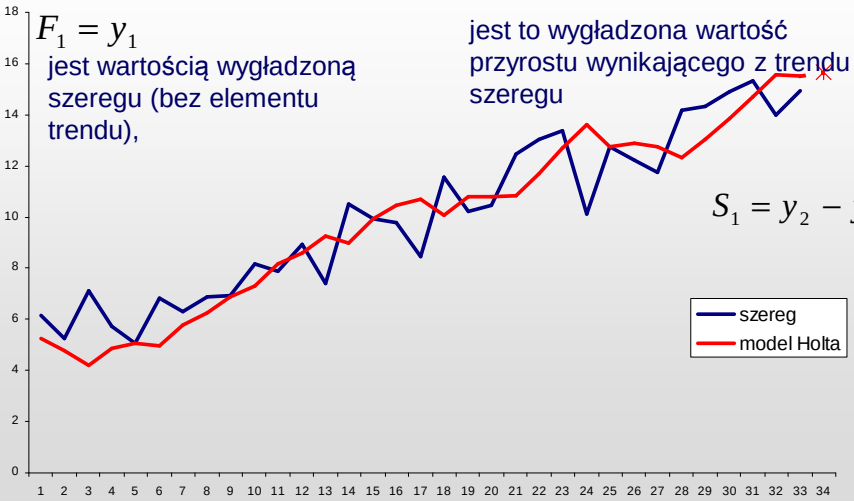
$$F_t = \alpha y_{t-1} + (1-\alpha)(F_{t-1} + S_{t-1}) \qquad S_t = \beta(F_t - F_{t-1}) + (1-\beta)S_{t-1}$$

$F_1 = y_1$

jest wartością wygładzoną szeregu (bez elementu trendu),

jest to wygładzona wartość przyrostu wynikającego z trendu szeregu

$S_1 = y_2 - y_1$



Metoda tendencji rozwojowej liniowego modelu Holta (wygładzania wykładniczego)

Algorytm prognozowania:

Równanie I (uśrednienie szeregu czasowego): $F_{t-1} = \alpha \cdot y_{t-1} + (1-\alpha) \cdot (F_{t-2} + S_{t-2})$

Równanie II (uśrednienie trendu): $S_{t-1} = \beta \cdot (F_{t-1} - F_{t-2}) + (1-\beta) \cdot S_{t-2}$

Równanie III (wyznaczenie prognozy): $y_T^* = F_n + (T-n) \cdot S_n$

gdzie: F_{t-1} - uśredniona prognoza zmiennej *Y* dla chwili *t-1*, S_{t-1} - uśredniony trend zmiany wartości zmiennej *Y* dla chwili *t-1*, y_{t-1} - rzeczywista wartość obserwacji zmiennej *Y* dla chwili *t-1*, *α*, *β* – parametry uśredniania z przedziału [0, 1], y_T^* - prognoza zmiennej *Y* dla chwili *T*, *n* – liczba elementów szeregu czasowego (obserwacji).

Tab. Charakterystyka metody Holta liniowego wygładzania wykładniczego prognozowania

Grupa metod	Prognozowanie na podstawie szeregu czasowego
Składowe szeregu czasowego	Trend i wahania przypadkowe
Przesłanki metody	Nie nastąpią zmiany w sposobie oddziaływania czynników określających zmienną prognozowaną, niezmienność trendu i regularność zmian
Postawa, zasada, reguły prognostyczna	Postawa pasywna, zasada *status quo*, reguła podstawowa prognozowania
Horyzont prognozy	Prognoza krótkookresowa
Ocena prognozy	Ocena dopuszczalności prognozy za pomocą średniego kwadratowego błędu *s** prognozy *ex post*
Zalety metody	Elastyczność metody (model adaptacyjny)
Wady metody	Konieczność i trudność doboru parametrów wygładzania *α* i *β* (minimalizacja błędów), konieczność przechowywania dużej ilości danych

Metoda tendencji rozwojowej trendu pełzającego z wagami harmonicznymi

Algorytm prognozowania:

Trend pełzający jest modelem adaptacyjnym służącym do budowy prognoz krótkookresowych. Procedura konstrukcji i ekstrapolacji trendu pełzającego jest następująca:

Krok 1: Ustalenie wartości stałej wygładzania $k < n$.

Krok 2: Oszacowanie na podstawie kolejnych fragmentów szeregu o długości *k* liniowych funkcji trendu.

Krok 3: Obliczenie wygładzonych wartości zmiennej $\hat{y}_{t(i)}$, tzn. wartości teoretycznych wynikających z *i*-tej funkcji trendu.

Krok 4: Obliczenie średniej wartości wygładzonej $\bar{y}_t$ dla każdego okresu *t* jako średniej arytmetycznej wartości wygładzonych obliczonych dla tego okresu w kroku 3. Po połączeniu odcinkami liniowymi kolejnych punktów (t, $\bar{y}_t$) otrzymuje się wykres wygładzonych wartości szeregu czasowego w postaci funkcji segmentowej, zwanej trendem pełzającym.

Krok 5: Ekstrapolacja modelu trendu pełzającego. Obliczenie przyrostów funkcji trendu dla wartości wygładzonych:

$$w_{t+1} = \bar{y}_{t+1} - \bar{y}_t, t = 1, \ldots, n-1$$

Krok 6: Nadanie wag poszczególnym przyrostom. Są to tzw. wagi harmoniczne.

$$C_{t+1}^n = \frac{1}{n-1} \cdot \sum_{i=1}^{t} \frac{1}{n-i}, t = 1, \ldots, n-1$$

Krok 7: Określenie średniego przyrostu trendu jako średniej ważonej (wagami harmonicznymi) wszystkich obliczonych w kroku 5. przyrostów.

$$\overline{w} = \sum_{t=1}^{n-1} C_{t+1}^{n} \cdot w_{t+1}$$

Krok 8: Wyznaczenie prognozy punktowej na moment/okres T.

$$y_T^* = \overline{y}_n + (T-n) \cdot \overline{w}$$

Tab. Charakterystyka metody trendu pełzającego prognozowania

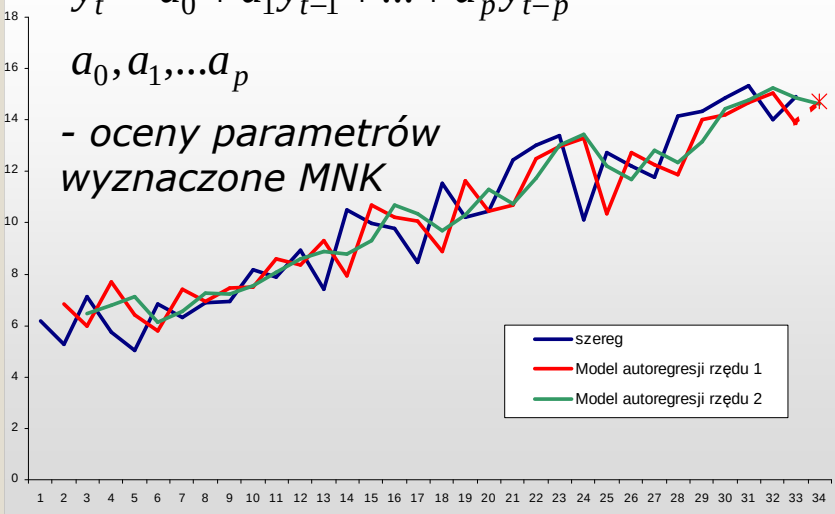
Grupa metod	Prognozowanie na podstawie szeregu czasowego
Składowe szeregu czasowego	Trend i wahania przypadkowe
Przesłanki metody	Nie nastąpią zmiany w sposobie oddziaływania czynników określających zmienną prognozowaną, utrzymanie trendu, zmiany mechanizmu rozwojowego badanego zjawiska
Postawa, zasada, reguły prognostyczna	Postawa pasywna, zasada *status quo*, reguła podstawowa prognozowania, model addytywny (wahania bezwzględnie stałe), model multiplikatywny (wahania względnie stałe)
Horyzont prognozy	Prognoza krótkookresowa i średniookresowa
Ocena prognozy	Ocena dopuszczalności prognozy za pomocą błędu prognozy *ex ante*
Zalety metody	Pozwala na adaptacyjność do tendencji zmian trendu, umożliwia wyznaczenie błędu *ex ante*
Wady metody	Złożoność obliczeniowa i trudność doboru stałej wygładzania *k* (minimalizacja błędów), konieczność posiadania dużej ilości danych i ich przechowywania

Model autoregresji

$$\hat{y}_t^* = a_0 + a_1 y_{t-1} + \ldots + a_p y_{t-p}$$

$$a_0, a_1, \ldots a_p$$

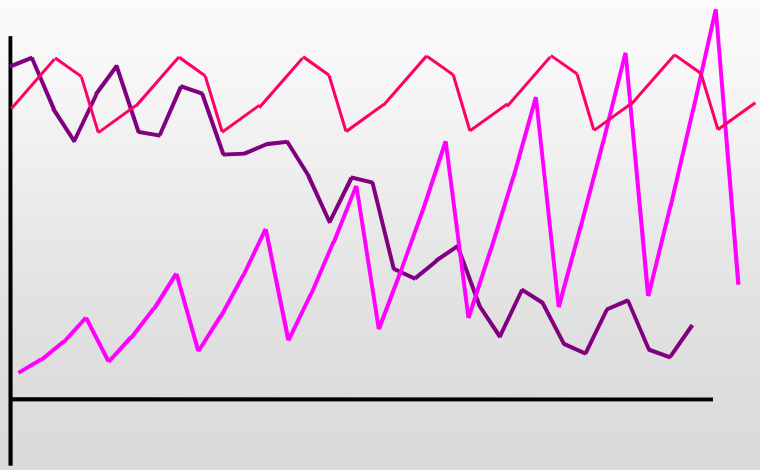
- oceny parametrów wyznaczone MNK



Metody oceny dopuszczalności prognoz

Metoda oceny	Zakres zastosowań
średni względny błąd dopasowania modelu $$MAPE = \frac{1}{n}\sum_{t=1}^{n}\frac{\lvert y_t - \hat{y}_t \rvert}{y_t}$$	•metoda naiwna • średnia ruchoma prosta • średnia ruchoma ważona • wygładzanie wykładnicze, model Holta, Wintersa •metoda wskaźników
względny błąd ex ante $$\eta_t = \frac{S(\hat{y}_t^*)}{\hat{y}_t^*}$$	• model trendu, m. trendu ze zmiennymi sezonowymi •model autoregresji

Prognozowanie na podstawie sezonowych szeregów czasowych



Identyfikacja składowych szeregu – trend i sezonowość

Dwuczynnikowa analiza wariancji (ANOVA)

Hipotezy:

- o równości wartości przeciętnych między grupami dla różnych wariantów czynnika A – sezonu (występuje sezonowość),
- o równości wartości przeciętnych między grupami dla różnych wariantów czynnika B – cyklu – roku (występuje trend),

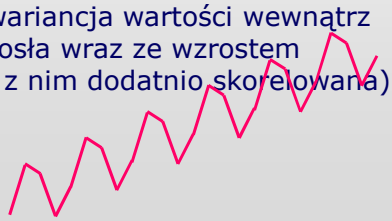
(zakłada się addytywny charakter zmian)

Identyfikacja składowych szeregu – charakter sezonowości

Oczekujemy, że:

- dla sezonowości multiplikatywnej: wariancja wartości wewnątrz kolejnych okresów będzie rosła wraz ze wzrostem średniego poziomu (będzie z nim dodatnio skorelowana)

- dla sezonowości addytywnej: wariancja wartości wewnątrz kolejnych okresów nie będzie rosła wraz ze wzrostem średniego poziomu (nie będzie z nim dodatnio skorelowana)



Szereg z sezonowością (bez trendu)

Metody prognozowania:

metoda naiwna

metoda wskaźników

model autoregresji

model regresji ze zmiennymi sezonowymi

analiza harmoniczna

Szereg z trendem i sezonowością

Charakter sezonowości:

- *addytywny*
- *multiplikatywny*

Metody prognozowania:

- *metoda naiwna*
- *metoda wskaźników dla wygładzonego szeregu*
- *model regresji ze zmiennymi czasową i sezonowymi*
- *model Wintersa*
- *model autoregresji*
- *model trendów jednoimiennych okresów*

Metoda naiwna (bez trendu)

$$y_t^* = y_{t-k}$$

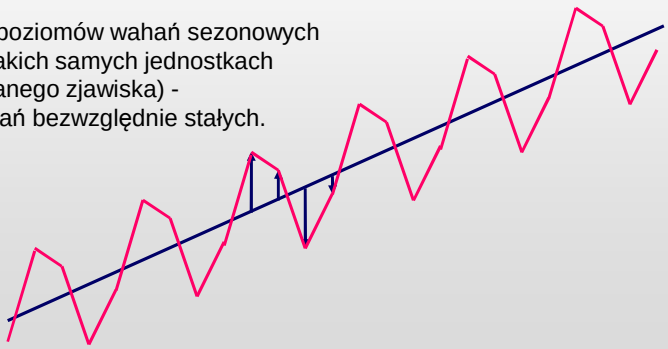
Metoda naiwna (z trendem, dla wahań addytywnych)

$$y_t^* = y_{t-k} + \left(y_{t-1} - y_{t-k-1}\right)$$

Addytywne wskaźniki sezonowości (także w szeregu bez trendu)

Wahania okresowe wyodrębnia się za pomocą:

- wskaźników sezonowości (wyrażonych w %) - w przypadku wahań okresowych względnie stałych;
- absolutnych poziomów wahań sezonowych (wyrażonych w takich samych jednostkach jak obserwacje danego zjawiska) - w przypadku wahań bezwzględnie stałych.



$$S_i = \frac{\sum_{j=1}^{r} y_{ij}}{\sum_{j=1}^{r} \hat{y}_{ij}} * 100;$$

gdzie:

i - numery podokresów cyklu ($i = 1,2,3,....,r$);

j - numer kolejnego cyklu ($j = 1,2,3,...,k$);

S_i - surowy wskaźnik sezonowości dla i-tego podokresu cyklu;

y_{ij} - wartości empiryczne zjawiska w i-tym podokresie cyklu w kolejnych latach;

$\hat{y}_{ij}$ - analogiczne wartości teoretyczne wyznaczone za pomocą funkcji trendu.

Suma S_i powinna być równa $100 \cdot r$ czyli (1200 w przypadku sezonowości miesięcznej, 400 w przypadku sezonowości kwartalnej, 200 w przypadku sezonowości półrocznej).

Jeżeli warunek powyższy jest spełniony to surowe wskaźniki sezonowości są jednocześnie czystymi wskaźnikami sezonowości. Jeżeli jednak jest inaczej należy obliczyć czyste wskaźniki sezonowości posługując się następującym wzorem:

$\hat{S}_i = k * S_i$ gdzie: k - współczynnik korygujący

$$k = \frac{100 * r}{\sum_{i=1}^{r} S_i}.$$

Absolutne poziomy wahań sezonowych obliczamy z następującego wzoru:

$$q_i = \left(\frac{\hat{S}_i}{100} - 1 \right) * \bar{y}$$

gdzie:

q_i - absolutne poziomy wahań sezonowych;

$\bar{y}$ - średnia arytmetyczna z wartości empirycznych y_t dla $t = 1,2,3,....,n$

Wyznaczanie prognozy:

$$\hat{y} = a_0 + a_1 * T + q_i$$

Metoda wskaźników sezonowych

Algorytm prognozowania:

$y_{ti}^{*} = y_t^{*(w)} \cdot c_i$ (model multiplikatywny) **lub** $y_{ti}^{*} = y_t^{*(w)} + c_i$ (model addytywny)

gdzie: y_{ti}^{*} - prognoza zmiennej *Y* dla chwili *t* oraz *i*-tego cyklu, $y_t^{*(w)}$ - wstępna prognoza w oparciu o trend liniowy, c_i - czysty wskaźnik sezonowości dla *i*-tego cyklu.

W metodzie wskaźników sezonowych prognozę wyznaczamy jako ekstrapolację dotychczasowej tendencji korygowanej wskaźnikiem sezonowości.

Procedura konstrukcji prognozy jest następująca:

Krok 1: Identyfikacja cykli i faz występujących w cyklach.

Krok 2: Opracowanie prognozy wstępnej jako ekstrapolacji zaobserwowanej tendencji rozwojowej z oszacowaniem parametrów za pomocą KMNK.

$$\hat{y}_t^{(w)} = a + b \cdot t$$

Krok 3: Wyznaczenie wartości wskaźników sezonowości z_{ti} jako ilorazu lub różnicy wartości rzeczywistych i wartości teoretycznych:

$$z_{ti} = \frac{y_{ti}}{\hat{y}_t} \text{ (model multiplikatywny)} \qquad z_{ti} = y_{ti} - \hat{y}_t \text{ (model addytywny)}$$

Krok 4: Wyznaczenie surowych wskaźników sezonowości z_i w celu wyeliminowania wahań przypadkowych we wskaźnikach z_{ti} przez wyznaczenie średniej arytmetycznej tych wartości z_{ti}, które odpowiadają jednoimiennym okresom.

Krok 5: Obliczenie średniej arytmetycznej *q* surowych wskaźników sezonowości z_i.

Krok 6: Wyznaczenie czystych wskaźników sezonowości jako ilorazów lub różnic surowych wskaźników sezonowosci z_i i wielkości *q*:

$$c_i = \frac{z_i}{q} \text{ (model multiplikatywny)} \qquad c_i = z_i - q \text{ (model addytywny)}$$

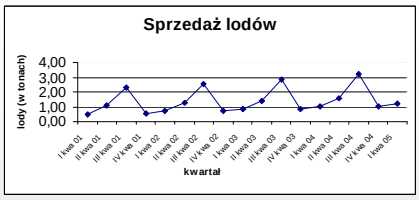
Krok 7: Wyznaczenie prognozy jako korektę prognozy wstępnej za pomocą czystego wskaźnika sezonowości dla danej fazy cyklu:

$$y_{ti}^{*} = y_t^{*(w)} \cdot c_i \text{ (model multiplikatywny)} \qquad y_{ti}^{*} = y_t^{*(w)} + c_i \text{ (model addytywny)}$$

Tab. Charakterystyka metody wskaźników sezonowych prognozowania

Grupa metod	Prognozowanie na podstawie szeregu czasowego
Składowe szeregu czasowego	Trend lub stały poziom i wahania sezonowe
Przesłanki metody	Nie nastąpią zmiany w sposobie oddziaływania czynników określających zmienną prognozowaną, utrzymanie trendu, niezmienność siły i rodzaju wahań sezonowych
Postawa, zasada, reguły prognostyczna	Postawa pasywna, zasada *status quo*, reguła podstawowa prognozowania, model addytywny (wahania bezwzględnie stałe), model multiplikatywny (wahania względnie stałe)
Horyzont prognozy	Prognoza krótkookresowa lub średniookresowa
Ocena prognozy	Ocena dopuszczalności prognozy za pomocą błędu prognozy *ex ante*
Zalety metody	Pozwala na obserwację tendencji zmian wahań sezonowych, umożliwia wyznaczenie błędu *ex ante*
Wady metody	Złożoność obliczeń

Potrójne wygładzanie wykładnicze
Metoda Wintersa



$$F_t = \alpha \cdot (y_t - S_{t-r}) + (1-\alpha) \cdot (F_{t-1} + T_{t-1})$$
$$T_t = \beta \cdot (F_t - F_{t-1}) + (1-\beta) \cdot T_{t-1}$$
$$S_t = \gamma \cdot (y_t - F_t) + (1-\gamma) \cdot S_{t-r}$$
$$y^*_\tau = F_n + (\tau - n) \cdot T_n + S_{\tau-r}$$
$$\tau > n$$

Metodę Wintersa można stosować wówczas, gdy zjawisko charakteryzuje się wahaniami o coraz większej bądź mniejszej amplitudzie (wersja multiplikatywna)

Ft - wygładzona wartość zmiennej prognozowanej

Tt = St - wygładzona wartość przyrostu trendu

St = Ct - ocena wskaźnika sezonowości na moment t

alfa - stała wygładzania poziomu zmiennej F

beta - stała wygładzania współczunnika trendu T

gama - stała wygładzania efektu sezonowego S

r - długość cykli sezonowości

n - numer ostatniej obserwacji

Metoda Wintersa

Algorytm prognozowania:

Gdy szereg czasowy zmiennej prognozowanej zawiera tendencję rozwojową, wahania sezonowe i wahania przypadkowe, wtedy stosuje się model Wintersa należący do klasy modeli wygładzania wykładniczego. Prognozę wyznacza się w sposób sekwencyjny, korzystając z 3 równań zawierających 3 parametry wygładzania. Model prognozy w metodzie Wintersa może mieć postać: addytywną lub multiplikatywną. Wygładzona wartość zmiennej po eliminacji wahań sezonowych na moment *t-1* dana jest wzorem:

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

Wygładzona wartość przyrostu trendu na moment *t-1* dana jest wzorem:

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

Ocena wskaźników sezonowości na moment *t-1* dana jest wzorem:

$$C_{t-1} = \gamma \frac{y_{t-1}}{F_{t-1-r}} + (1-\gamma) \cdot C_{t-1-r}$$

Gdzie *r* – długość cyklu sezonowego (liczba faz w cyklu)

Równanie prognozy na moment *t>n*:

$$y_t^* = (F_n + (t-n) \cdot S_n) + C_{t-r} \quad \text{(postać addytywna modelu)}$$

$$y_t^* = (F_n + (t-n) \cdot S_n) \cdot C_{t-r} \quad \text{(postać multiplikatywna modelu)}$$

Wyboru parametrów wygładzania dokonuje prognosta.

Wartości początkowe:

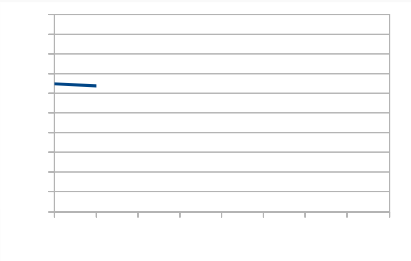
- dla komponenty *F* przyjąć wartość rzeczywistą zmiennej z szeregu czasowego odpowiadającą 1. fazie drugiego cyklu lub średnią wartość z 1. cyklu,
- dla komponenty *S* przyjąć różnicę średnich wartości z 2. i 1. cyklu bądź przyjąć 0,
- dla komponenty *C* (w poszczególnych fazach 1. cyklu) przyjąć ilorazy wartości rzeczywistej zmiennej z 1. cyklu w odniesieniu do średniej wartości w 1. cyklu bądź przyjąć 1.

Grupa metod	Prognozowanie na podstawie szeregu czasowego
Składowe szeregu czasowego	Tendencja rozwojowa, wahania sezonowe, wahania przypadkowe
Przesłanki metody	Nie nastąpią zmiany w sposobie oddziaływania czynników określających zmienną prognozowaną, utrzymanie trendu, niezmienność siły i rodzaju wahań sezonowych
Postawa, zasada, reguły prognostyczna	Postawa pasywna, zasada *status quo*, reguła podstawowa prognozowania, model addytywny (wahania bezwzględnie stałe), model multiplikatywny (wahania względnie stałe)
Horyzont prognozy	Prognoza krótkookresowa
Ocena prognozy	Ocena dopuszczalności prognozy za pomocą błędu prognozy *ex ante*
Zalety metody	Pozwala na obserwację tendencji zmian wahań sezonowych, umożliwia wyznaczenie błędu *ex ante*
Wady metody	Trudności obliczeniowe (obliczenia sekwencyjne), konieczność i trudność doboru parametrów α, β, γ (minimalizacja błędów)

Przykład 1:

Dochody ze sprzedaży (w tys. zł) odzieży rowerowej przedsiębiorstwa ”MTB Hard Dress” w latach 1999 - 2007 w Polsce kształtowały się następująco: 56,5; 56,4; 56,7; 55,9; 56,0; 55,9; 56,2; 56,6; 56,5.

Ocena wzrokowa: w badanym szeregu czasowym występują składowa systematyczna w postaci stałego (przeciętnego) poziomu oraz wahania przypadkowe.



Do oceny siły wahań przypadkowych zastosowano współczynnik zmienności (miara zróżnicowania rozkładu cechy. W odróżnieniu od odchylenia standardowego, które określa bezwzględne zróżnicowanie cechy, współczynnik zmienności jest miarą względną, czyli zależną od wielkości średniej arytmetycznej).

$$V = \frac{s}{\overline{x}}, \quad \overline{x} \neq 0$$

V = 0.55%

Przykład 1:

Niska wartość współczynnika zmienności badanej zmiennej dopuszcza zastosowanie metody naiwnej do prognozowania na następny okres.

Metoda prognozowania dochodów ze sprzedaży odzieży rowerowej:

Ponieważ w badanym szeregu czasowym występuje składowa systematyczna stałego poziomu oraz niewielkie wahania przypadkowe, w prognozowaniu można wykorzystać metody prognozowania na podstawie szeregu czasowego tj:

- metoda naiwna,
- wygładzanie wykładniczego.

Wyznaczenie prognozy:

- metoda naiwna: $y10^* = y9 = 56,5$
- wygładzanie wykładnicze: $y10^* = \alpha \cdot y9 + (1 - \alpha) \cdot y9^* = ???$

Przykład 1:

			naiwna	wygladzanie	
t	**rok**	**Dochody (y)**	y*	y* (alfa = 0,2)	y* (alfa = 0,8)
1	1999	56,5		56,5000	56,5000
2	2000	56,4		56,5000	56,5000
3	2001	56,7		56,4800	56,4200
4	2002	55,9		56,5240	56,6440
5	2003	56		56,3992	56,0488
6	2004	55,9		56,3194	56,0098
7	2005	56,2		56,2355	55,9220
8	2006	56,6		56,2284	56,1444
9	2007	56,5		56,3027	56,5089
10	**2008**		**56,5**	**56,3422**	**56,5018**

Przykład 2:

Liczba pasażerów przewożonych środkami komunikacji miejskiej w latach 1999-2008 (w mln osób) w pewnym mieście była następująca: 12,8; 12,6; 12,9; 13,1; 12,8; 12,7; 13,0; 13,1; 13,2; 13,0.

a) Sformułować przesłanki do wyznaczenia prognozy na 2006 r.

b) Wybrać metodę prognozowania. Wybór uzasadnić.

c) Wyznaczyć prognozę liczby przewożonych pasażerów w 2006 r.

Przykład 2:

Rozwiązanie:

Przesłanki: hipotezy badawcze określające wstępnie mechanizm rozwojowy (wskazanie zjawisk i kierunki wpływów). Mechanizm: trend liniowy + wahania przypadkowe

Wybór metody:

- prognozowanie na podstawie szeregu czasowego z trendem liniowym.

$$b = \frac{\sum y \cdot (t - \bar{t})}{\sum (t - \bar{t})^2} = \frac{3{,}3}{82{,}5} = \mathbf{0{,}04}$$

$$a = \bar{y} - b \cdot \bar{t} = \mathbf{12{,}7}$$

$$y_t^* = 0{,}04 \cdot t + 12{,}7$$

Przykład 2:

Metoda tendencji rozwojowej liniowego modelu Holta (wygładzania wykładniczego)

Algorytm prognozowania:

Równanie I (uśrednienie szeregu czasowego): $F_{t-1} = \alpha \cdot y_{t-1} + (1-\alpha) \cdot (F_{t-2} + S_{t-2})$

Równanie II (uśrednienie trendu): $S_{t-1} = \beta \cdot (F_{t-1} - F_{t-2}) + (1-\beta) \cdot S_{t-2}$

Równanie III (wyznaczenie prognozy): $y_T^* = F_n + (T-n) \cdot S_n$

gdzie: F_{t-1} - uśredniona prognoza zmiennej *Y* dla chwili *t-1*, S_{t-1} - uśredniony trend zmiany wartości zmiennej *Y* dla chwili *t-1*, y_{t-1} - rzeczywista wartość obserwacji zmiennej *Y* dla chwili *t-1*, *α*, *β* – parametry uśredniania z przedziału [0, 1], y_T^* - prognoza zmiennej *Y* dla chwili *T*, *n* – liczba elementów szeregu czasowego (obserwacji).

Przykład 2:

wygładzanie (alfa: 0.2, beta 0.4)

t	rok	y [mln]	y* (at+b)	St	Ft	
1	1999	12,8	12,74	0,04	12,74	12,74
2	2000	12,6	12,78	0,04	12,78	12,46
3	2001	12,9	12,82	0,04	12,82	12,54
4	2002	13,1	12,86	0,04	12,86	12,62
5	2003	12,8	12,9	0,04	12,9	12,7
6	2004	12,7	12,94	0,04	12,94	12,78
7	2005	12,7	12,98	0,04	12,98	12,86
8	2006	13	13,02	0,04	13,02	12,94
9	2007	13,1	13,06	0,04	13,06	13,02
10	2008	13,2	13,1	0,04	13,1	13,1
11	**2009**		**13,14**	0,04	13,14	**13,18**

Przykład 3:

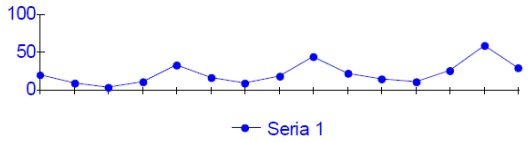
Liczba zawartych umów leasingowych w firmie finansowo-leasingowej w poszczególnych kwartałach lat 2002-2005 kształtowała się następująco:

20 10 4 11 33 17 9 18 45 23 14 11 25 60 30 13.

Sporządzić prognozy liczby zawartych umów leasingowych na kolejny rok (2006), stosując model Wintersa.

Rozwiązanie:

Sprawdzenie zmienności szeregu czasowego:



Gdy szereg czasowy zmiennej prognozowanej zawiera tendencję rozwojową, wahania sezonowe i wahania przypadkowe, wtedy stosuje się model Wintersa należący do klasy modeli wygładzania wykładniczego. Prognozę wyznacza się w sposób sekwencyjny, korzystając z 3 równań zawierających 3 parametry wygładzania.

Przykład 3:

Postawa prognosty: pasywna. Horyzont prognozy: prognoza krótkookresowa. Postaci modelu Wintersa: wybieramy postać multiplikatywną. Liczba umów jest najwyższa w 1. kwartałach i 4 kwartałach, najniższa w 3. kwartałach. Zauważyć można tendencję wzrostową liczby zawieranych umów.

Postać multiplikatywna:

Wygładzona wartość zmiennej po eliminacji wahań sezonowych na moment t-1 dana jest wzorem:

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

Wygładzona wartość przyrostu trendu na moment t-1 dana jest wzorem:

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

Ocena wskaźników sezonowości na moment t-1 dana jest wzorem:

$$C_{t-1} = \gamma \frac{y_{t-1}}{F_{t-1-r}} + (1-\gamma) \cdot C_{t-1-r}$$

Gdzie r – długość cyklu sezonowego (liczba faz w cyklu)

Równanie prognozy na moment t>n:

$$y_t^* = (F_n + (t-n) \cdot S_n) \cdot C_{t-r}$$

t	y	$y-y_{sr}$	$(y-y_{sr})^2$
1	20	-1,4375	2,066406
2	10	-11,4375	130,8164
3	4	-17,4375	304,0664
4	11	-10,4375	108,9414
5	33	11,5625	133,6914
6	17	-4,4375	19,69141
7	9	-12,4375	154,6914
8	18	-3,4375	11,81641
9	45	23,5625	555,1914
10	23	1,5625	2,441406
11	14	-7,4375	55,31641
12	11	-10,4375	108,9414
13	25	3,5625	12,69141
14	60	38,5625	1487,066
15	30	8,5625	73,31641
16	13	-8,4375	71,19141
Suma	343		3231,938
y_{sr}	**21,4375**		
s	**14,67864**		
V_z [%]	**68,47179**		

Przykład 3:

Wartości początkowe:

- dla komponenty F przyjąć wartość rzeczywistą zmiennej z szeregu czasowego odpowiadającą 1. fazie drugiego cyklu lub średnią wartość z 1. cyklu,
- dla komponenty S przyjąć różnicę średnich wartości z 2. i 1. cyklu bądź przyjąć 0,
- dla komponenty C (w poszczególnych fazach 1. cyklu) przyjąć ilorazy wartości rzeczywistej zmiennej z 1. cyklu w odniesieniu do średniej wartości w 1. cyklu bądź przyjąć 1.

W tabeli podano obliczenia prognoz na dla kolejnych kwartałów 2006 r. (t=17, 18, 19, 20) przyjmując założenie, że wartości parametrów są równe: α = 0.5, β = 0.95, γ = 0.2 (obliczenie 1), α = 0.5, β = 0.95, γ = 0.5 (obliczenie 2).

Obliczenie 1.

t	y_t	F_t	S_t	C_t	y_t^*	$(y_t-y_t^*)^2$
1	20			1,777778		
2	10			0,888889		
3	4			0,355556		
4	11			0,977778		
5	33	33	32	1,622222		
6	17	42,0625	10,20938	0,791943		
7	9	38,79219	-2,59633	0,330846		
8	18	27,30248	-11,045	0,914078		
9	45	21,99858	-5,59095	1,706895		
10	23	22,72506	0,410606	0,835974		
11	14	32,72574	9,52118	0,350236		
12	11	27,14045	-4,82997	0,812323		
13	25	18,47848	-8,47037	1,636101		
14	60	40,89033	20,86773	0,962247		
15	30	73,7073	32,21951	0,361592		
16	13	60,96515	-10,4941	0,692505		
17					**82,5758**	
18					**38,46779**	
19					**10,6608**	
20					**13,14991**	
α	**0,5**					
β	**0,95**					
γ	**0,2**					

Obliczenie 2.

t	y_t	F_t	S_t	C_t	y_t^*	$(y_t-y_t^*)^2$
1	20			1,777778		
2	10			0,888889		
3	4			0,355556		
4	11			0,977778		
5	33	33	32	1,388889		
6	17	42,0625	10,20938	0,646525		
7	9	38,79219	-2,59633	0,293781		
8	18	27,30248	-11,045	0,818529		
9	45	24,32872	-3,37732	1,619277		
10	23	28,26311	3,568805	0,730153		
11	14	39,74327	11,08459	0,323021		
12	11	32,1333	-6,67524	0,580427		
13	25	20,44852	-11,4343	1,42093		
14	60	45,59438	23,31685	1,023052		
15	30	80,89226	34,69883	0,346942		
16	13	68,99421	-9,56821	0,384424		
17					**84,44017**	
18					**51,00713**	
19					**13,97815**	
20					**11,81003**	
α	**0,5**					
β	**0,95**					
γ	**0,5**					

Przykład 3:

t	rok	kwartal	Obliczenie 1	Obliczenie 2
17	2006	1	**82,58**	**84,44**
18		2	**38,47**	**51,01**
19		3	**10,66**	**13,98**
20		4	**13,15**	**11,81**

Opracowano m.in. na podstawie:

Roman Pietroń

PROGNOZOWANIE I SYMULACJE

(Wybrane zagadnienia i materiały wykładu dla 4 roku ZiM)

Politechnika Wrocławska