

Statystyka Opisowa z Demografią oraz Biostatystyka
Analiza Zmian w czasie

Aleksander Denisiuk

denisjuk@euh-e.edu.pl

Elbląska Uczelnia Humanistyczno-Ekonomiczna

ul. Lotnicza 2

82-300 Elbląg

Analiza Zmian w czasie

Najnowsza wersja tego dokumentu dostępna jest pod adresem

<http://denisjuk.euh-e.edu.pl/>

Szereg czasowy

- prezentuje poziom badanej zbiorowości lub zjawiska w kolejnych okresach lub momentach czasu
 - stan na określony moment: zasób
 - liczba ludności
 - liczba budynków
 - stan zatrudnienia
 - przedział czasu: strumień
 - liczba urodzin
 - liczba oddanych mieszkań
 - wpłaty na rachunki
- szereg momentów, szereg okresów

Szereg momentów. Przykład

Tabela 1: Liczba ludności w woj. gdańskim w latach 1990–1997. Źródło: W. Makać, *Podstawy statystyki i demografii dla studentów demografii*

Rok	Liczba ludności w tys., stan na 31 XII
1990	1 431,6
1991	1 438,9
1992	1 437,7
1993	1 444,8
1994	1 450,1
1995	1 455,9
1996	1 460,7
1997	1 464,8

Szereg okresów. Przykład

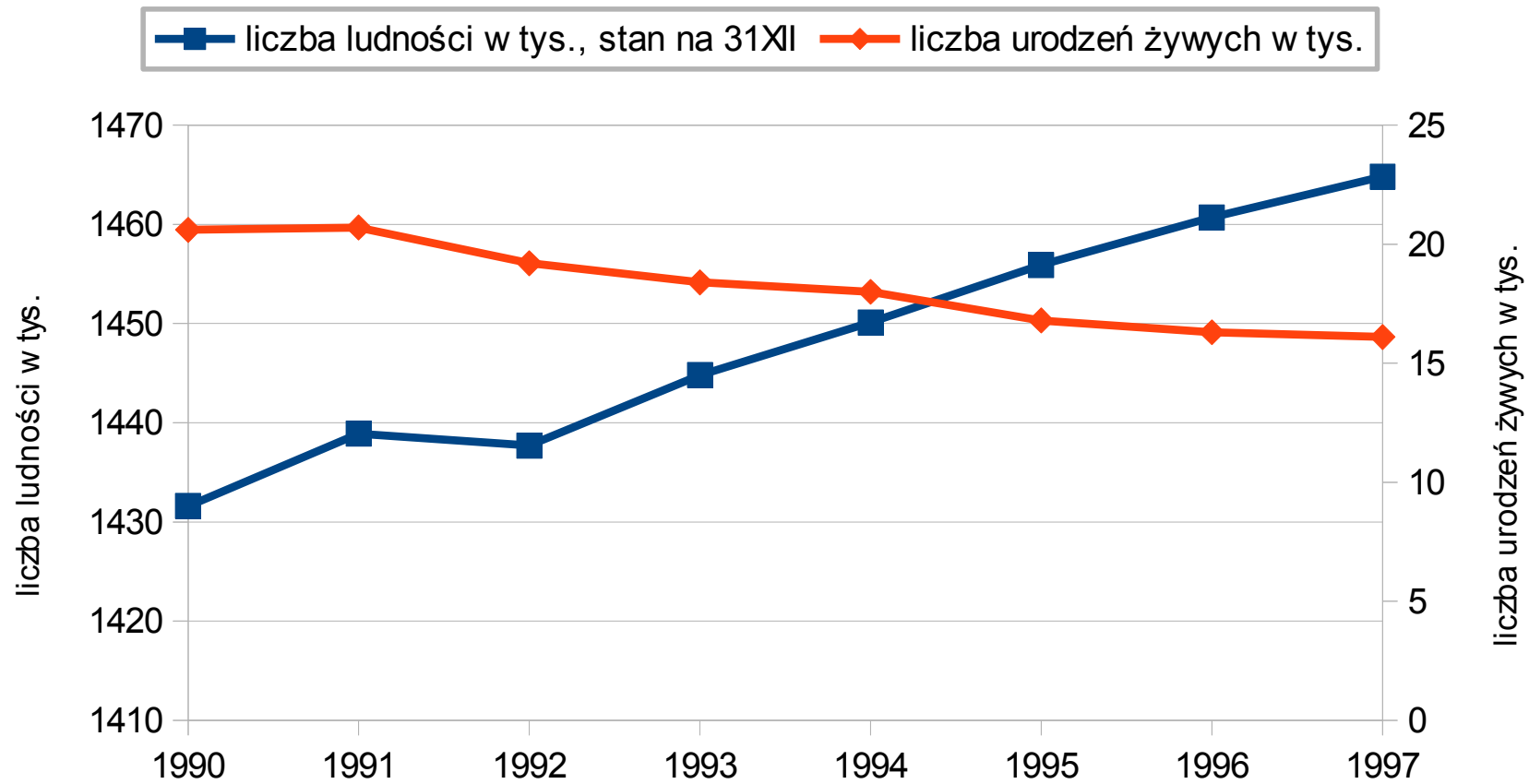
Tabela 2: Liczba urodzin w woj. gdańskim w latach 1990–1997. Źródło: W. Makać, *Podstawy statystyki i demografii dla studentów demografii*

Rok	Liczba urodzin żywych w tys., stan na 31 XII
1990	20,6
1991	20,7
1992	19,2
1993	18,4
1994	18,0
1995	16,8
1996	16,3
1997	16,1

Graficzne przedstawienie szeregów

- wykresy liniowe
- poziomo: jednostka czasu
- pionowo: poziom zjawiska
- dwa zjawiska na tym samym wykresie (różne jednostki lub różne skale): dwie osie pionowe

Graficzne przedstawienie szeregów. Przykład



Prawidłowości statystyczne

- wskaźniki dynamiki
 - natężenie zmian w poszczególnych okresach (momentach) w odniesieniu do innych okresów (momentów)
- tendencja rozwojowa (trend)
 - prawidłowości rozwojowe w dłuższym czasie
- wahania sezonowe
 - dane są w kwartałach (miesiącach) i zjawiska zmieniają się regularnie

Średni poziom

- w szeregu okresów — średnia arytmetyczna:

$$\bar{y} = \frac{y_1 + y_2 + \dots + y_{n-1} + y_n}{n} = \frac{\sum_{i=1}^n y_i}{n}$$

- w szeregu momentów — średnia chronologiczna:

$$\bar{y} = \frac{\frac{1}{2}y_1 + y_2 + \dots + y_{n-1} + \frac{1}{2}y_n}{n-1}$$

- Przykład: liczba pracujących w budownictwie w 1993 roku:

31 III	30 VI	30 IX	31 XII
724 tys.	712 tys.	696 tys.	666 tys.

- *średni poziom zatrudnienia* w 1993 roku: $\bar{y} = 701$ tys.
- *średni stan zatrudnienia* w ostatnich dniach kwartałów: 699,5 tys.

Analiza dynamiki

- porównujemy poziom zjawiska w wybranym momencie (okresie) y_t z poziomem w momencie (okresie), przyjętym za podstawę porównań, y_0 .
 - o ile jednostek poziom jest wyższy (niższy)? Przyrost absolutny: $y_t - y_0$
 - czy zaobserwowany przyrost (spadek) jest duży? Przyrost względny: $\frac{y_t - y_0}{y_0} \cdot 100\%$
 - ile razy poziom jest wyższy (niższy)? Wskaźnik dynamiki: $\frac{y_t}{y_0} \cdot 100\%$

Analiza dynamiki. Przykład

- liczba pracujących w sektorze prywatnym według stanu na 31 XII 1990 r. wynosiła 7 902 tys. osób, a liczba na koniec 1997 r. — 10 614 tys. osób. Ocenić dynamikę
 - przyrost absolutny: 2 712 tys. osób
 - przyrost względny: 34,3%
 - wskaźnik dynamiki: 134,3%

Indeksy dynamiki

- mierzą relatywne zmiany w szeregach czasowych
- indeksem zawywa się iloraz poziomu w okresie badanym, y_t do poziomu w okresie podstawowym, y_0
 - indeks jednopodstawowy $i_{t/0} = \frac{y_t}{y_0} \cdot 100\%$
 - indeksy łańcuchowe $i_{t/t-1} = \frac{y_t}{y_{t-1}} \cdot 100\%$
 - tempo $T = \frac{y_t - y_{t-1}}{y_{t-1}} \cdot 100\% = \left(\frac{y_t}{y_{t-1}} - 1 \right) \cdot 100\%$
 - $T = i_{t/t-1} - 100\%$
- przykład:
Liczba urodzin w woj. gdańskim w latach 1990–1997

Związki matematyczne między indeksami

- operujemy indeksami wyrażonymi ułamekami dziesiętnymi (nie odsetkami)
 - znane są wskaźniki dynamiki o podstawie łańcuchowej
⇒ wskaźniki o podstawie stałej

- $\frac{y_2}{y_0} = \frac{y_1}{y_0} \cdot \frac{y_2}{y_1}$

- $\frac{y_3}{y_0} = \frac{y_2}{y_0} \cdot \frac{y_3}{y_2}$

- $\frac{y_4}{y_0} = \frac{y_3}{y_0} \cdot \frac{y_4}{y_3}$

- znane są wskaźniki dynamiki o podstawie stałej
⇒ wskaźniki o podstawie łańcuchowej

- $\frac{y_2}{y_1} = \frac{y_2}{y_0} \cdot \frac{y_1}{y_0}$

- $\frac{y_3}{y_2} = \frac{y_3}{y_0} \cdot \frac{y_2}{y_0}$

- $\frac{y_4}{y_3} = \frac{y_4}{y_0} \cdot \frac{y_3}{y_0}$

Zmiana podstawy porównań

- $\frac{y_1}{y_k} = \frac{y_1}{y_0} \cdot \frac{y_k}{y_0}$
- $\frac{y_2}{y_k} = \frac{y_2}{y_0} \cdot \frac{y_k}{y_0}$
- $\frac{y_3}{y_k} = \frac{y_3}{y_0} \cdot \frac{y_k}{y_0}$

Średnie względne zmiany w pewnym czasie

- średnia geometryczna indeksów łańcuchowych:

$$\bar{i} = \sqrt[n-1]{i_{n/n-1} \cdots i_{1/0}} = \sqrt[n-1]{\frac{y_1}{y_0} \frac{y_2}{y_1} \frac{y_3}{y_2} \cdots \frac{y_n}{y_{n-1}}} = \sqrt[n-1]{\frac{y_n}{y_0}}$$

- średnie tempo (średnia względna zmiana:

$$\bar{T} = (\bar{i} - 1) \cdot 100\% = \left(\sqrt[n-1]{\frac{y_n}{y_0}} - 1 \right) \cdot 100\%$$

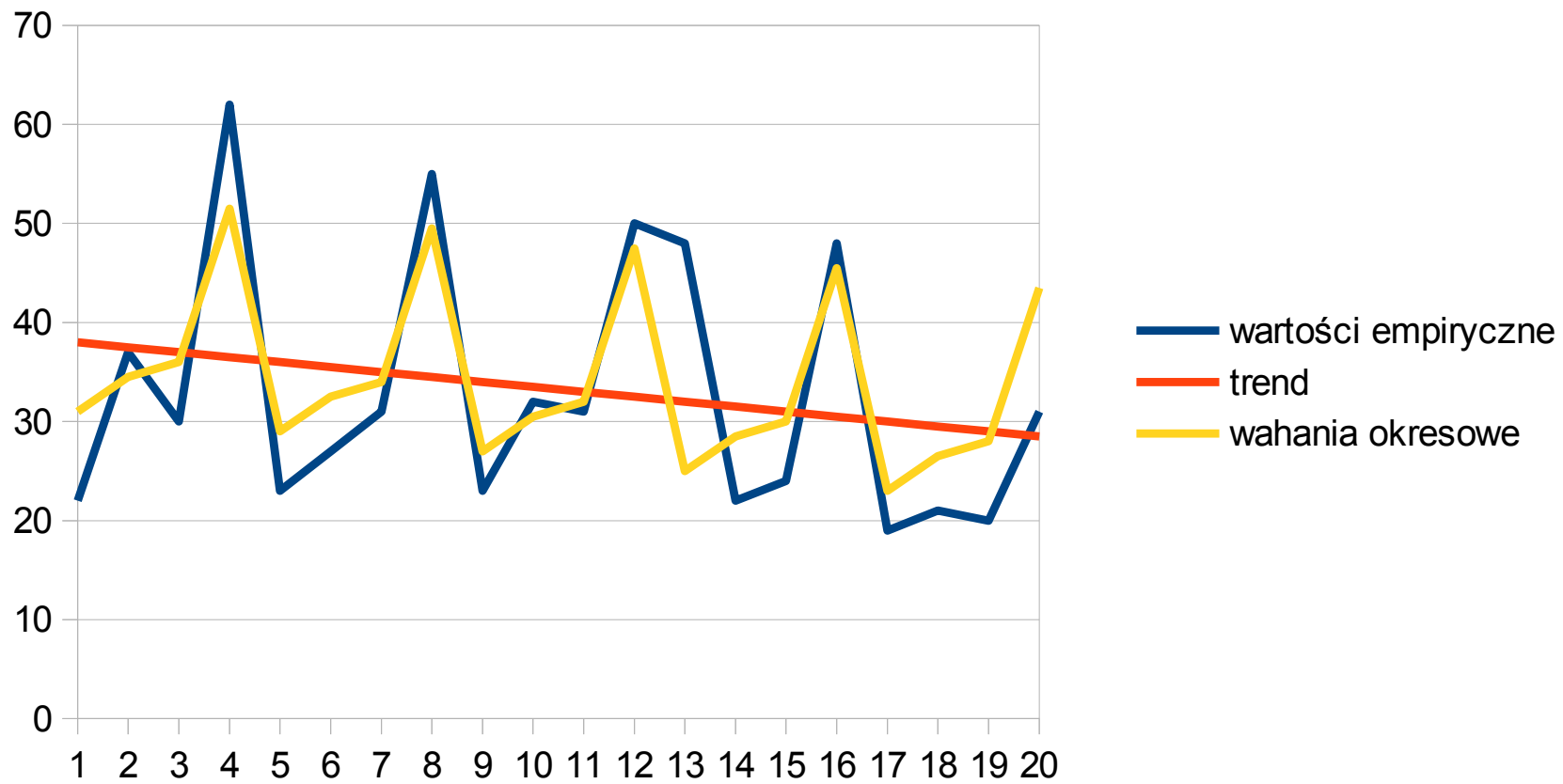
- przykład:

Liczba urodzin w woj. gdańskim w latach 1990–1997

- $\bar{i} = 0,965$
- $\bar{T} = -3,5\%$

Wyznaczenie tendencji rozwojowej

- główne składniki kształtujące poziom zjawiska:
 - tendencja rozwojowa (trend) — ogólny kierunek zmian, ujawniający się w długich okresach
 - wachania okresowe — powtarzające się regularnie
 - wachania przypadkowe — nie związane z istotą zjawiska



- wyodrębnienie tendencji poprzez metodę
 - średnich ruchomych
 - najmniejszych kwadratów

Metoda średnich ruchomych

- średnia ruchoma jest średnią arytmetyczną określonej liczby k kolejnych wartości szeregu
- k zazwyczaj jest nieparzystą liczbą (3-letnia, 5-letnia, itd)
- średnie zapisuje się przy środkowym okresie
- każdy kolejny szereg jest bardziej wygładzony, niż poprzedni
- stosuje się zazwyczaj do analizy szeregów z dużymi różnokierunkowymi zmianami
- przykład

Funkcja trendu

- znalezienie funkcji analitycznej, która najlepiej odzwierciadła rozwój zjawiska w czasie
- metoda najmniejszych kwadratów $\sum_{t=1}^n (y_t - \hat{y}_t)^2 \rightarrow \min$,
 $\hat{y}_t = f(t)$
- najczęściej $t = 1, 2, 3, \dots, n$ (lub $t = 0, 1, 2, \dots, n - 1$)
- najczęściej $f(t) = a + b \cdot t$

Liniowa funkcja trendu

- $$\begin{cases} \sum_t y_t = n \cdot s + b \sum_t t \\ \sum_t t_t \cdot t = a \sum_t t + b \sum_t t^2 \end{cases}$$

- $$\begin{cases} b = \frac{n \sum y_t t - \sum y_t \cdot \sum t}{n \sum t^2 - (\sum t)^2} \\ a = \frac{\sum y_t - b \sum t}{n} \end{cases}$$

- $b > 0$ interpretuje się jako średni przyrost, $b < 0$ jako spadek
- a pokazuje teoretyczny poziom zjawiska w okresie $t = 0$

Liczba pracujących (mln osób) w sektorze prywatnym

Rok	1990	1991	1992	1993	1994	1995	1996	1997
y_t	7,9	8,4	8,4	8,7	9,0	9,4	10,1	10,6

- $b = 0,365$ mln osób
- $a = 7,42$ mln osób

Dopasowanie funkcji do rozkładu

- $y_t = \hat{y}_t + e_t$,
 - \hat{y}_t — tendencja rozwojowa
 - e_t — wahania przypadkowe (reszty)
- odchylenie standardowe składnika resztowego:

$$Se(Y) = \sqrt{\frac{\sum_{t=1}^n (y_t - \hat{y}_t)^2}{n-2}},$$

- współczynnik zmienności przypadkowej: $Ve = \frac{Se(Y)}{\bar{y}} \cdot 100\%$
- w poprzednim przykładzie:
 - $Se(Y) = 0,216$ mln osób
 - $Ve = 2,4\%$

Ekstrapolacja

- $\hat{y}_{n+p} - Se(Y) < y_{n+p} < \hat{y}_{n+p} + Se(Y)$
- w przykładzie $11,219 < y_{2000} < 11,651$
- uwagi:
 - czy nie zmieniły się przyczyny główne, określające tendencje rozwoju?
 - czy nie uległy zmianom definicje i sposoby pomiaru?
 - czy zmiany badanego zjawiska nie są wywołane głównie zmianami cen (inflacja)?

Analiza wahań sezonowych

- wahania sezonowe — zmiany poziomu zjawiska, powtarzające się regularnie w kolejnych cyklach rocznych
- sezonowość:
 - miesięczna ($d = 12$)
 - kwartalna ($d = 4$)
 - półroczna ($d = 2$)
- obserwacja kilku cykli rocznych (najmniej 3–4)

Wskaźniki sezonowości

- liczby względne, mierzą zanężenie wahań sezonowych w wyróżnionych podokresach cyklu rocznego
 - odchylenie od średniej
 - odchylenie od trendu

Odchylenie od średniej

- brak wyraźnej tendencji wzrastania lub spadku
- oblicza się *średnie jednoimiennych okresów*, \bar{y}_i
- po podzieleniu średnich jednoimiennych przez średnią ogólną (\bar{y}), otrzymujemy *względne wskaźniki sezonowości* dla okresów jednoimiennych ${}_wS_i^* = \frac{\bar{y}_i}{\bar{y}} \cdot 100\%$
- pokazują stosunek poziomu zjawiska w i -tym podokresie do wielkości średniej

Liczba zawartych małżeństw

miesiąc		I	II	III	IV	V	VI
liczba zawartych małżeństw (tys.)	1995	10,8	12,9	4,7	28,1	6,1	23,1
	1996	10,2	10,9	4,9	27,3	6,0	27,6
	1997	9,8	7,9	8,1	25,7	6,3	25,2
miesiąc		VII	VIII	IX	X	XI	XII
liczba zawartych małżeństw (tys.)	1995	20,2	23,7	30,4	24,6	7,7	14,8
	1996	16,7	28,3	26,7	24,8	8,2	12,8
	1997	18,8	29,6	26,2	26,1	8,3	12,9

Odchylenie od trendu

- określamy funkcję trendu
- określamy wartości teoretyczne, \hat{y}_t
- ilorazy $\frac{y_t}{\hat{y}_t}$
- dla okresów jednoimiennych obliczamy *wskaźniki sezonowości* $wS_i = \frac{1}{c} \sum_c \frac{y_{it}}{\hat{y}_{it}}$
- korygujemy wskaźniki: $wS_{i(\text{kor})} = wS_i \cdot \frac{d}{\sum_{i=1}^d wS_i}$

Liczba mieszkań oddanych to użytku (tys.)

Kwartał	1989	1990	1991	1992	1993
I	23,5	23,3	23,7	47,3	18,7
II	35,8	28,1	29,9	19,6	19,6
III	29,4	31,0	28,5	22,4	19,1
IV	61,5	51,9	54,4	43,7	30,6