

# **SZEREGI CZASOWE**

## **Procesy stochastyczne** **Wprowadzenie do szeregów czasowych**

**Marzena Nowakowska**

**Wydział Zarządzania i Modelowania Komputerowego**  
**Politechnika Świętokrzyska**

**Budynek C, p. 3.21**

**[spimn@tu.kielce.pl](mailto:spimn@tu.kielce.pl)**

# Informacje organizacyjne

1. Budynek C, piętro III, p. 3.21.
2. Zaliczenie przedmiotu  
Zdobycie co najmniej 50% punktów możliwych do uzyskania ze sprawdzianu końcowego przeprowadzonego w formie testu.
3. Udział w wykładach jest wskazany ze względu na transfer wiedzy na ćwiczenia laboratoryjne z przedmiotu **SELECTED ASPECTS OF STOCHASTIC PROCESSES**
4. Materiały do przedmiotu w Internecie (uzupełniane na bieżąco):  
  
[staff.tu.kielce.pl/spimn](http://staff.tu.kielce.pl/spimn)  
[staff.tu.kielce.pl/spimn/sc/](http://staff.tu.kielce.pl/spimn/sc/)  
[staff.tu.kielce.pl/spimn/sasp/](http://staff.tu.kielce.pl/spimn/sasp/)
5. Po co jest ten przedmiot?

# Po co jest ten przedmiot

W zakresie wiedzy i umiejętności

- Pojęcia: proces stochastyczny i szereg czasowy
- Dekompozycja szeregu czasowego i analiza jego składnika losowego
- Wygładzanie szeregu czasowego
- Metody średniej ruchomej
- Autoregresja
- Problem stacjonarności szeregu czasowego
- Wykorzystanie narzędzi języka R do analiz i prognozowania szeregów czasowych
- Opracowanie sprawozdania z zadań analizy szeregów czasowych w języku angielskim

# Przestrzeń probabilistyczna

**Przestrzeń probabilistyczna** jest trójką  $(\Omega, \mathcal{F}, P)$ , w której poszczególne elementy są określone następująco:

- $\Omega$  to pewien niepusty zbiór, zwany zbiorem zdarzeń elementarnych;
- $\mathcal{F}$  to pewna rodzina podzbiorów zbioru  $\Omega$  o własnościach:
  - $\emptyset \in \mathcal{F}$ ,
  - jeżeli  $A \in \mathcal{F}$ , to dopełnienie zbioru  $A$  również należy do  $\mathcal{F}$ ;  $A^c = \Omega \setminus A \in \mathcal{F}$ ,
  - jeżeli  $A_1, A_2, \dots \in \mathcal{F}$ , to  $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$
- $P$  to funkcja,  $P: \mathcal{F} \rightarrow [0, 1]$ , o własnościach:
  - $P(\Omega) = 1$  (unormowanie),
  - dla  $A_1, A_2, \dots \in \mathcal{F}$ , parami rozłącznych (tzn.  $A_i \cap A_j = \emptyset$  dla  $i \neq j$ ) zachodzi przeliczalna addytywność: 
$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n)$$

Zdarzenie elementarne – to pojedynczy, elementarny wynik doświadczenia losowego.

# Proces stochastyczny

Niech będą dane:  $(\Omega, \mathcal{F}, P)$  – ustalona przestrzeń probabilistyczna oraz  $T \subset \mathbb{R}$  – zbiór liczb rzeczywistych lub podzbiór tego zbioru (przedział skończony lub nieskończony lub podzbiór dyskretny). Zakłada się, że dla każdego  $t \in T$  zdefiniowana na przestrzeni  $(\Omega, \mathcal{F}, P)$  zmienna losowa  $X_t: \Omega \rightarrow \mathbb{R}$  lub  $X: \Omega \times T \rightarrow \mathbb{R}$  określana jako

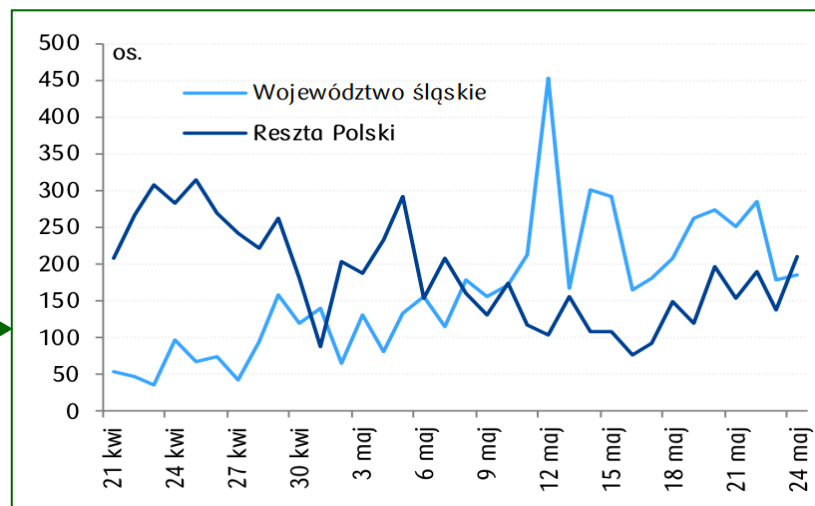
$X_t(\omega) = X(\omega, t)$  jest nazywana **procesem stochastycznym** ze zbiorem indeksującym  $T$ . Zazwyczaj  $T$  interpretuje się jako czas.

Ww. oznacza, że dla ustalonego  $t$  funkcja  $X$  rozważana jako funkcja argumentu  $\omega$  jest zmienną losową.

Proces stochastyczny w czasie dyskretnym  $X = \{X_n, n = 0, 1, 2, \dots\}$  jest policzalnym zbiorem zmiennych losowych indeksowanych przez nieujemne liczby całkowite.

Procesem stochastyczny w czasie ciągłym  $X = \{X_t, 0 \leq t < \infty\}$  jest nieprzeliczalnym zbiorem zmiennych losowych indeksowanych przez nieujemne liczby rzeczywiste.

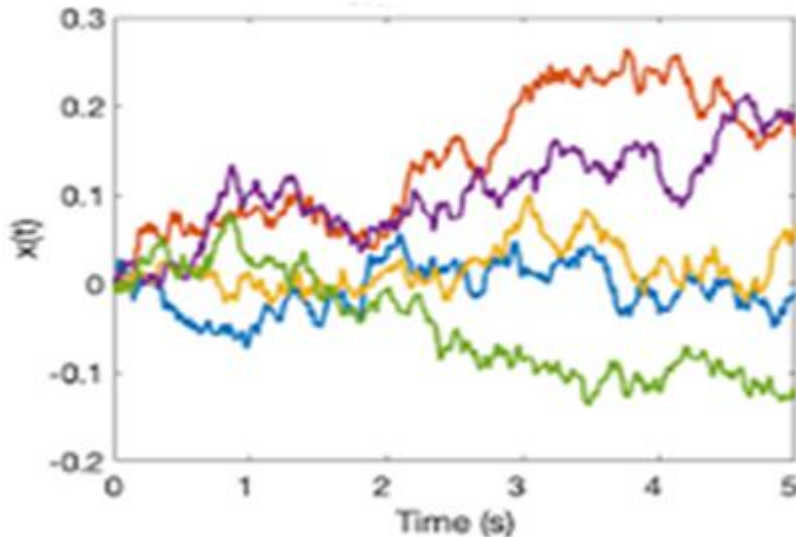
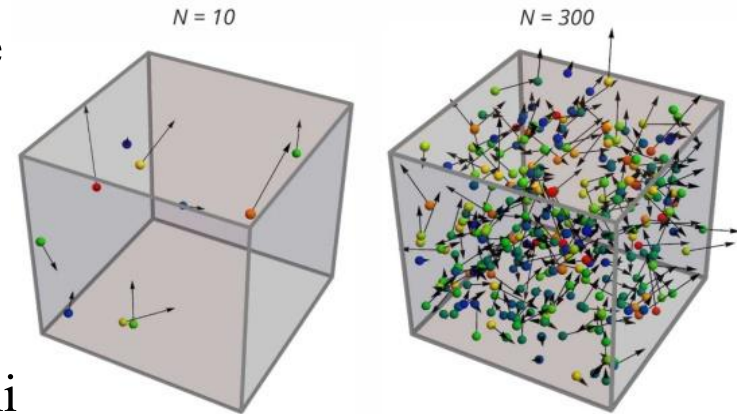
Jako proces stochastyczny można rozpatrywać np. ruch cząstki gazu, poziom wody w zbiorniku, wahania skrzydła samolotu, zmiany ceny akcji, liczbę zakażonych koronawirusem itp.



# Ruch Browna

Cząstka Browna wykonuje ruch losowy (błądzenie losowe).

- Ruch cząstki w jakimś kierunku jest doświadczeniem
- Można opisać położenie względem jednej współrzędnej, np. osi X (rzut ruchu w przestrzeni trójwymiarowej na jedną oś) –wynik doświadczenia.
- $x(t)$  jest położeniem cząstki w chwili  $t$ .



Możliwe położenie cząstki Browna dla pięciu doświadczeń.

- Wielokrotne powtarzanie doświadczenia, gdzie początkowe położenie cząstki (w chwili  $t = 0$ ) nie ma większego znaczenia.
- W kolejnych doświadczeniach otrzymuje się inną trajektorię cząstki.
- Gdyby przeprowadzać coraz więcej doświadczeń, trajektorie pokrywałyby coraz większą część płaszczyzny.

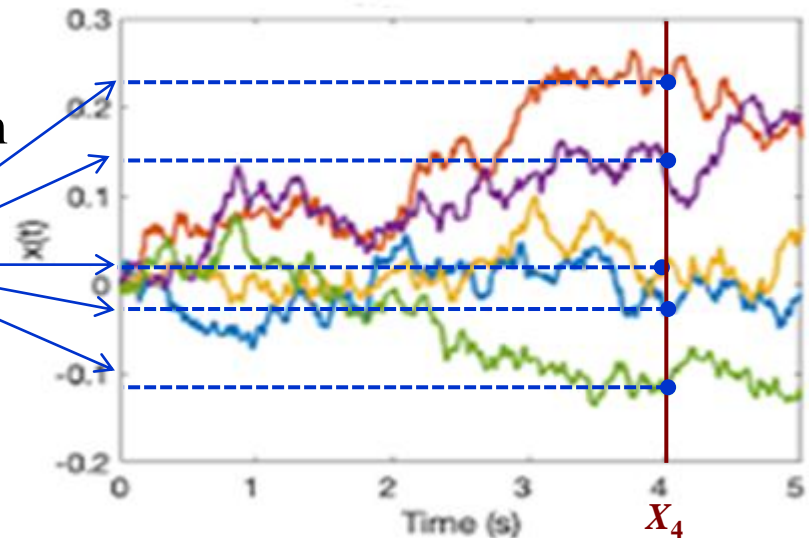
# Trajektorie ruchów Browna; realizacje procesu stochastycznego

Dla chwili czasu  $t = 4$  linia pionowa  $t = 4$  na rysunku przecina trajektorie w pięciu punktach. Te punkty  $x(t = 4)$  można potraktować jako możliwe wartości (realizacje) zmiennej losowej, oznaczonej np. jako  $X_4$ .

Zbór możliwych położenia cząstki Browna w różnych chwilach czasu indeksowanego przez wskaźnik czasu  $t$  można opisać za pomocą formuły procesu stochastycznego:  $X = \{X(\omega, t), t \in (0, T]\}$

- Dla ustalonego  $t$ :  $X(\omega, t) = X_t(\omega)$  jest zmienną losową opisującą możliwe położenia cząstki w tej samej chwili czasu ale dla wielu doświadczeń, czyli różnych  $\omega \in \Omega$  (np.  $X(\omega_1, 4)$ ,  $X(\omega_2, 4)$ ,  $X(\omega_3, 4)$ ,  $X(\omega_4, 4)$ ,  $X(\omega_5, 4)$  to wartości procesu w chwili  $t=4$  dla 5 doświadczeń).
- Jedna trajektoria cząstki to jedna realizacja procesu stochastycznego. Jest to funkcja czasu dla ustalonego  $\omega \in \Omega$ , czyli  $X(\omega_1, t)$  są to możliwe położenia cząstki dla pierwszego doświadczenia ( $\omega = \omega_1$ ) ale różnych chwil czasu,  $X(\omega_2, t)$  są to możliwe położenia cząstki dla drugiego doświadczenia ( $\omega = \omega_2$ ) ale różnych chwil czasu itd.

Różne doświadczenia są indeksowane przez  $\omega \in \Omega$ , różne chwile czasu są indeksowane przez  $t \in (0, T]$ .



Możliwe realizacje tego samego procesu.

# Specyfika procesu stochastycznego

Proces stochastyczny  $X = X(\omega, t)$  - funkcja dwóch zmiennych:  $\omega \in \Omega$  oraz  $t \in T$ .

Możliwe podejścia:

- rodzina zmiennych losowych indeksowanych "czasem"  $t \in T$ ,  
dla każdego ustalonego  $t$  funkcja  $X = X(\omega, t)$  argumentu  $\omega$  jest zmienną losową
- rodzina realizacji zależna od parametru  $\omega \in \Omega$   
dla ustalonego zdarzenia elementarnego  $\omega$  funkcję argumentu  $t \in T$  nie jest zmienną losową i nazywa się realizacją procesu stochastycznego lub trajektorią procesu stochastycznego (wyraża ewolucję w czasie wybranego zdarzenia losowego).

Oznaczenia procesu:  $X(\omega, t)$ ,  $X_t(\omega)$ , notacja uproszczona:  $X(t)$ ,  $\{X(t), t \in T\}$ .

Zbiór wartości, jakie może przyjmować  $X(\omega, t)$  często nazywa się przestrzenią stanów. Jeżeli  $X(\omega, t) = x$  to mówi się, że proces jest w stanie  $x \in R$ ; wartości procesu są stanami, zbiór wszystkich stanów jest przestrzenią stanów.



# Przykładowe rodzaje procesów stochastycznych

Stany	Czas	Nazwa procesu	Przykład
C	C	CC	Czas uzyskania połączenia z określoną stroną internetową, jeśli polecenie połączenia zostało wydane na przeglądarce w chwili $t$ .
C	D	CD	Czas efektywnej pracy modemu danego komputera w poszczególne dni konkretnego tygodnia.
D	C	DC	Liczba uczestników forum dyskusyjnego na określonej stronie internetowej, zalogowanych w chwili $t$ .
D	D	DD	Liczba zalogowań komputerów do danego serwera w poszczególne dni konkretnego roku

# Szereg czasowy jako jedna z kategorii danych statystycznych

**Dane** to zbiory wartości opisujące ilość, jakość, fakt, statystyki pewnych cech (obiektów ożywionych lub nieożywionych) reprezentowane przez sekwencje symboli (znaków), które mogą być dalej interpretowane i przetwarzane.

**Dane statystyczne** powstają jako wynik badań eksperymentalnych lub obserwacji.

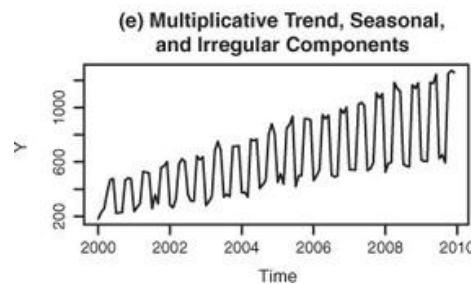
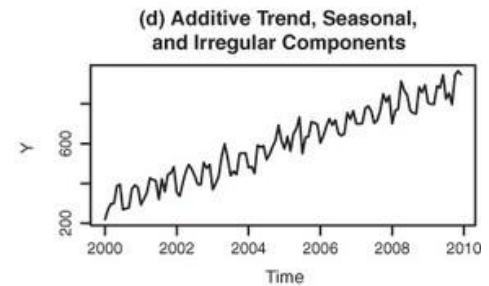
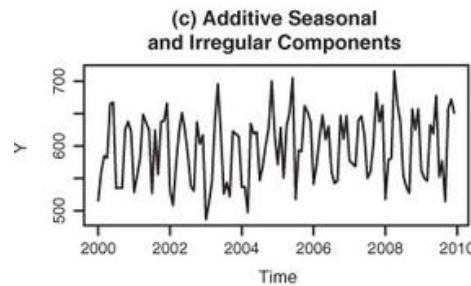
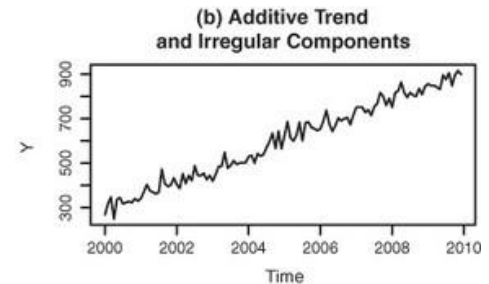
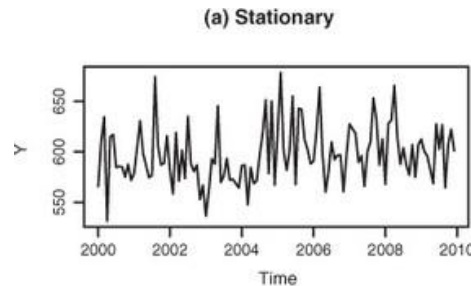
Ulokowanie szeregu statystycznego w kategorii danych statystycznych wg procesu rejestracji:

- **dane przekrojowe** (*cross sectional data*) wartości cechy zarejestrowane dla wielu jednostek w jednej jednostce czasu (np. wydatki powiatów woj. świętokrzyskiego w roku 2022 – stan na 31/12/2022)
- **szereg czasowy** (*time series data*) zarejestrowane wartości cechy w wielu jednostkach czasu (np. wydatki powiatu jędrzejowskiego w kolejnych miesiącach 2022 roku – stan na koniec każdego miesiąca)
- **dane panelowe** (*panel data, crosssectional time series data*) wartości cechy zarejestrowane dla wielu jednostek w wielu jednostkach czasu (np. wydatki powiatów woj. świętokrzyskiego w kolejnych miesiącach 2022 roku – stan na koniec każdego miesiąca)

# Szereg czasowy - definicja

**Szereg czasowy** (*time series data*) jest realizacją procesu stochastycznego, którego dziedziną jest czas (zazwyczaj z regularnymi odstępami). Jest to więc ciąg wartości (danych) uporządkowanych chronologicznie od wartości najstarszej historycznie do najbardziej aktualnej, przy czym pomiar jest zazwyczaj wykonywany z dokładnym i stałym krokiem czasowym, np. w kolejnych dniach, miesiącach czy kwartałach.

Częste oznakowania: szeregu:  $X_t$ ,  $Y_t$ ,  $Z_t$ , w których  $t$  oznacza czas i zależy od ustalonych punktów (chwili) czasowych. Często punkty czasowe indeksuje się kolejnymi liczbami naturalnymi.

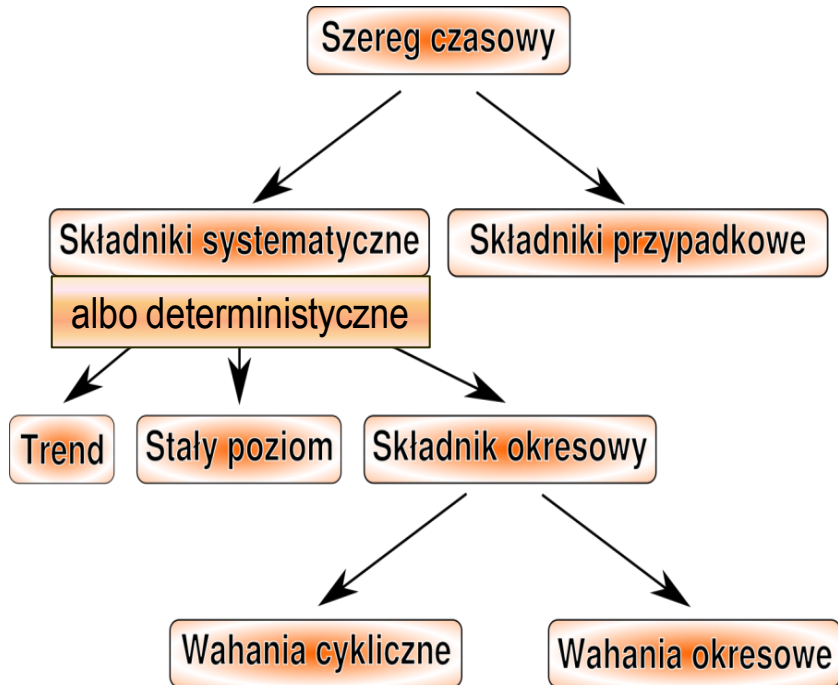


Źródło:

[https://drek453711klr.cloudfront.net/kabacoff2/Figure s/15fig04\\_alt.jpg](https://drek453711klr.cloudfront.net/kabacoff2/Figure%2Fs/15fig04_alt.jpg)

# Składowe szeregu czasowego

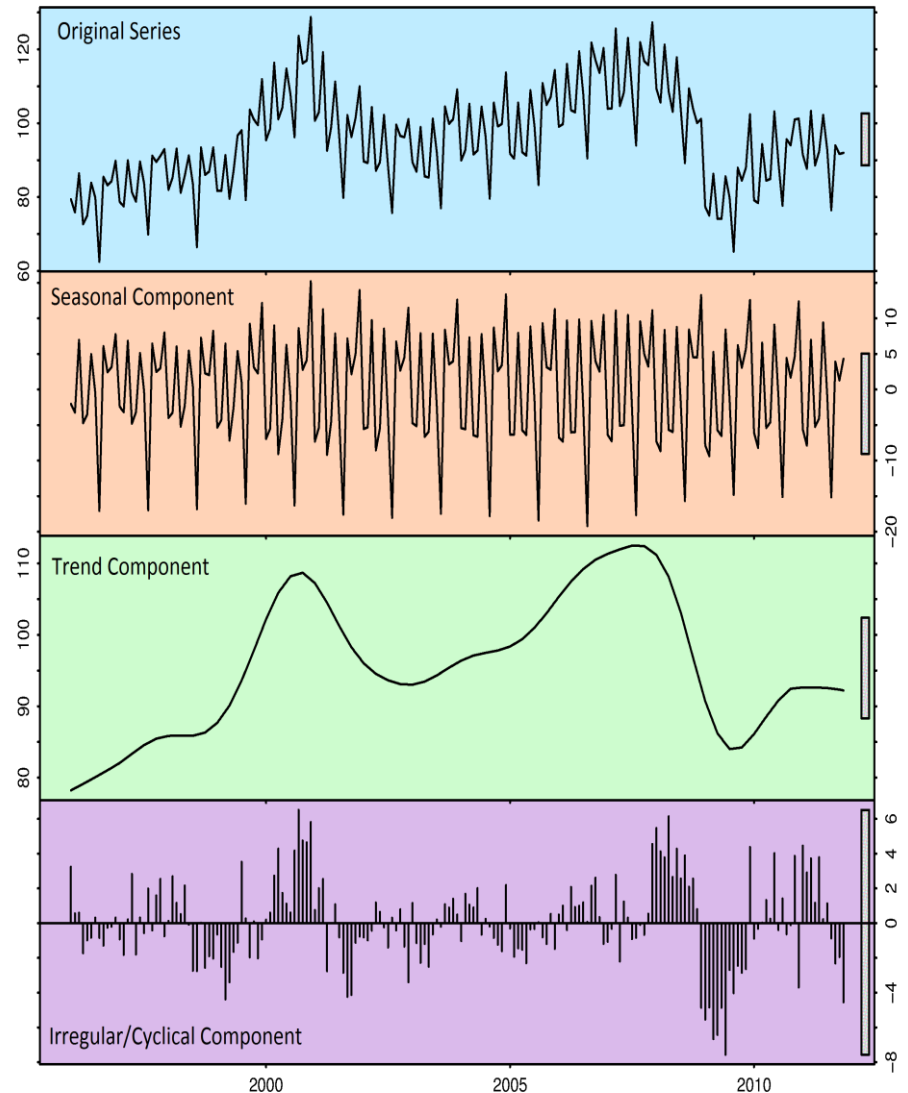
- trend (tendencja rozwojowa),  $T_t$
  - stały poziom,
  - wahania cykliczne  $C_r$
  - wahania sezonowe  $S_t$
- wahania przypadkowe)  $I_t$



Źródło:

[https://el.us.edu.pl/ekonofizyka/images/d/d4/Sk%C5%82adni\\_iki\\_szeregu\\_czasowego.png](https://el.us.edu.pl/ekonofizyka/images/d/d4/Sk%C5%82adni_iki_szeregu_czasowego.png)

## Dekompozycja szeregu czasowego



Źródło: <https://quantdare.com/wp-content/uploads/2014/09/decomp-example.png>

# Biały szum

**Biały szum** WN (*white noise*) jest ciągiem zmiennych losowych nieskorelowanych o jednakowym rozkładzie o średniej 0 (zero) i wariancji  $\sigma^2$ . Biały szum jest najprostszym modelem procesu stochastycznego. Opisuje zachowanie się składnika losowego, który charakteryzuje się zerową wartością oczekiwaną, stałą wariancją i nieskorelowaniem w czasie, co dla procesu  $\{\varepsilon_t\}$  jest zapisywane następująco:

$$\forall_t \varepsilon_t \sim (0, \sigma^2), \text{cov}(\varepsilon_t, \varepsilon_s) = 0$$

Jeżeli dodatkowo biały szum ma rozkład normalny, to jest nazywany **gaussowskim białym szumem** (*Gaussian white noise*). Gaussowski biały szum jest pochodną procesu Wienera. Najczęstsze oznaczenia:

$$\{\varepsilon_t\} \sim \text{BS}(0, \sigma^2) \sim \text{WN}(0, \sigma^2)$$

Jeżeli zmienne losowe  $Z_t$  mają takie same, niezależne od siebie rozkłady ze średnią zero i wariancją  $\sigma^2$  to wtedy oznacza się je:  $Z_t = \text{IID}(0, \sigma^2)$ .

# Błądzenie losowe

Błądzenie losowe to proces opisujący ścieżkę składającą się z kolejnych losowych kroków w pewnej przestrzeni matematycznej. **Błądzenie losowe (przypadkowe) RW** (*random walk*) jest zdefiniowany jako losowy proces określony rekurencyjną zależnością:

$$X_t = X_{t-1} + \varepsilon_t$$

$t$        $t > 0$ , chwila czasowa (numer kolejnego kroku)

$X_t$       Zmienna losowa  $X_t$  (realizacja procesu w chwili  $t$ , wyrażona wartością zmiennej  $X_t$ )

$X_{t-1}$       Zmienna losowa  $X_{t-1}$  (realizacja procesu w chwili  $t-1$ , wyrażona wartością zmiennej  $X_{t-1}$ )

$\varepsilon_t$        $\sim \text{WN}(0, \sigma^2)$

Przyjmując  $X_0 = z$  i rozpisując ww. zależność można przedstawić proces błądzenia przypadkowego w postaci (błądzenie losowe z dryfem):

$$X_t = z + \sum_{i=1}^t \varepsilon_i$$

Z własności białego szumu  $\varepsilon_t$  wynika, że wartość oczekiwana błądzenia losowego jest stała i równa  $z$ . Wariancja procesu jest równa sumie wariancji poszczególnych zmiennych  $\varepsilon_t$ .  $\text{Var}(X_t) = t\sigma^2$ , rośnie więc z czasem.

Często przyjmuje się, że  $z = 0$ , stąd w wielu opracowaniach wartość oczekiwana procesu RW jest określana jako 0.

# Częstotliwość próbkowania szeregu czasowego

**Częstotliwość próbkowania** określa, ile próbek jest pobieranych w naturalnej jednostce czasu, która dla R jest zdefiniowana jako rok. Ta częstość próbkowania jest właściwością potrzebną do określenia przy definiowaniu szeregu w środowisku programistycznym R.

W systemie R naturalną jednostką czasu jest 1 rok. Częstość danych określa się więc względem jego okresu:

- 52 - tygodniowe,
- 12 - miesięczne,
- 4 - kwartalne,
- 2 - półroczne,
- 1 - roczne.

Jeśli mierzy się wartość jakiejś zmiennej raz w miesiącu i posiada się dane z wielu lat, można użyć wartości 12 jako częstości.

# Generowanie liczb pseudolosowych

Obecnie w R w pakiecie **base** zaimplementowanych jest sześć algorytmów generowania liczb losowych - generatorów.

Generator to funkcja deterministyczna. Do losowania kolejnych liczb wykorzystuje tzw. ziarno (ang. *seed*), całkowicie determinujące wartości kolejnych liczb pseudolosowych. **Ziarno** to wartość (wartości) na podstawie której konstruowane będą kolejne liczby losowe. Dla ustalonego generatora i ziarna generowane będą identyczne liczby losowe. Sterując wyborem ziarna umożliwia się otrzymywanie identycznych ciągów liczb. W ten sposób można powtarzać wyniki symulacji, odtwarzać te wyniki na innych komputerach lub kontynuować obliczenia przerwane w wyniku wystąpienia jakiegoś błędu.

Zmienić generator liczb losowych w R można wywołując funkcję **RNGkind()** oraz podając jako jej argumenty nazwy dwóch wybranych generatorów, odpowiednio dla rozkładu jednostajnego i dla rozkładu normalnego.

Informacje o ziarnie i generatorze liczb losowych przechowywana jest w wektorze **.Random.seed**. Pierwszym elementem tego wektora jest informacja, który generator jest wykorzystywany, pozostałe elementy przechowują ziarno generatora. Ziarno nosi też nazwę **wartości początkowej** do generatora liczb pseudolosowych



# Symulacja białego szumu

Funkcja generująca liczby pseudolosowe o rozkładzie normalnym ma składnię:

**`rnorm(n, mean=0, sd=1)`**

gdzie  $n$  jest liczbą obserwacji do wygenerowania.

Pozostałe parametry funkcji definiują parametry rozkładu. Powstaje wektor  $n$ -elementowy reprezentujący realizację zmiennej losowej o zadanym rozkładzie.

```
WN_SN.50 <- rnorm(50, mean = 0, sd=1)
```

```
# standaryzowany rozkład normalny
```

```
WN_ND50 <- rnorm(50, mean =25, sd=15)
```

```
# rozkład normalny N(25, 15)
```

```
WN_GD <- rgeom(50,
```

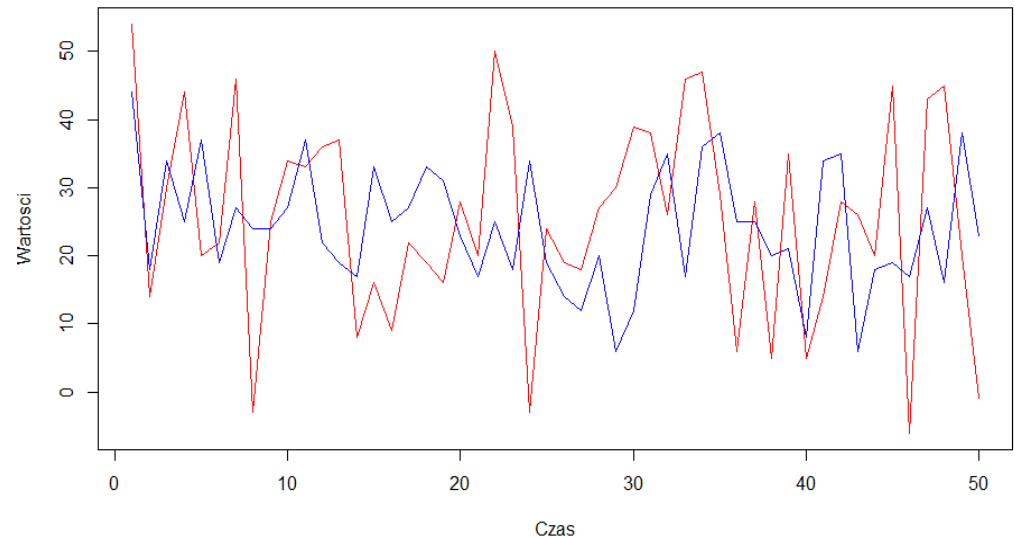
```
head(WN_50)
```

```
WN_50s <- as.ts(WN_GD)
```

## Dane a szereg czasowy w R

**`as.ts(x, ...)`**      **`is.ts(x)`**

Funkcje *as.ts* i *is.ts* odpowiednio przekształcają obiekt do szeregu czasowego (*as.ts*) i sprawdzają, czy obiekt jest szeregiem czasowym (*is.ts*).



# Symulacja błędzenia losowego

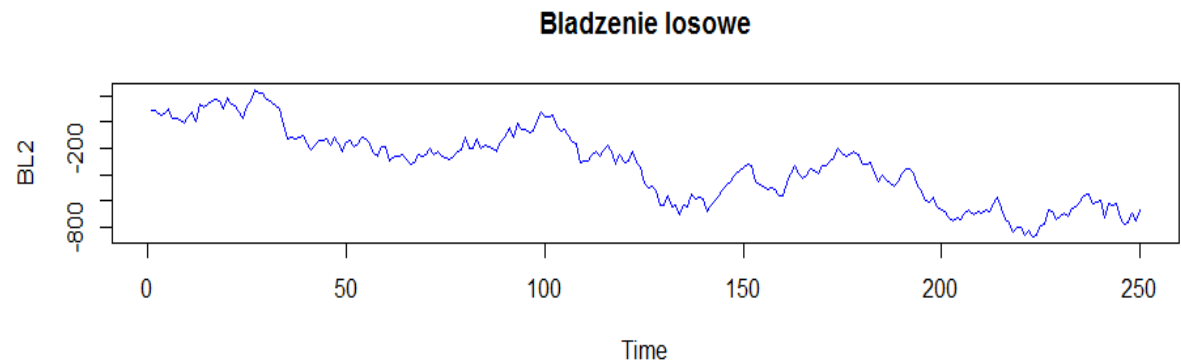
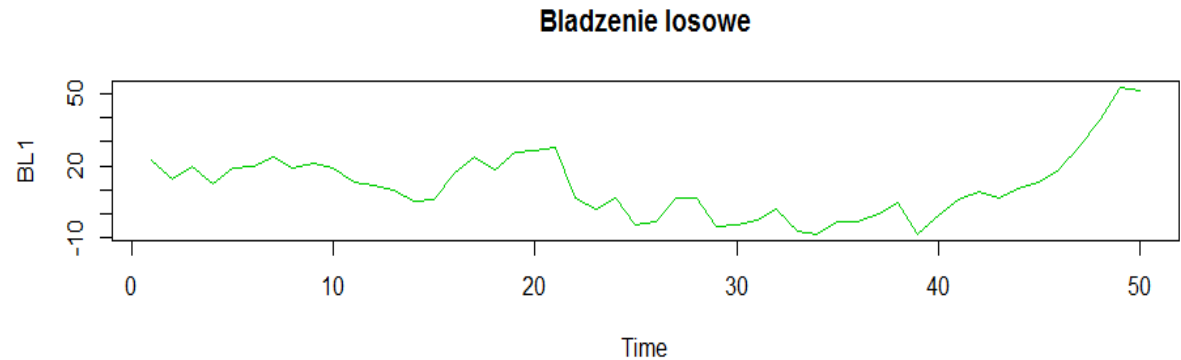
Zależność na błędzenie losowe przedstawiona jako suma zmiennych reprezentujących biały szum pozwalają na symulację tego błędzenia poprzez wyznaczenia skumulowanych sum na bazie białego szumu.

Do tego celu można wykorzystać funkcję *cum sum*.

```
sumySkum <- data.frame(1:50, cumsum(1:50))
```

```
View(sumySkum)
```

	X1.50	cumsum.1.50.
1	1	1
2	2	3
3	3	6
4	4	10
5	5	15
6	6	21
7	7	28
8	8	36
9	9	45
10	10	55
11	11	66
12	12	78



# Definiowanie szeregu czasowego w R

Funkcja *ts* służy do tworzenia obiektów szeregów czasowych.

```
ts(data = NA, start = 1, end = numeric(), frequency = 1, deltat = 1, ts.eps = getOption("ts.eps"), class = , names = )
```

Pierwszy parametr określa zbiór liczb reprezentujących wartości szeregu. Następne parametry informują, jak te liczby są ułożone w czasie.

data	Wektor lub macierz zaobserwowanych wartości szeregu czasowego. Ramka danych jest przekształcana do macierzy numerycznej poprzez <i>data.matrix</i> .
start	Punkt czasowy zarejestrowania pierwszej obserwacji. Może to być pojedyncza liczba lub wektor dwóch wartości. W przypadku dwuelementowego wektora, parametr ma postać: <i>c(rok_pierwszy, nr_okresu_w_pierwszym_roku)</i> . Parametr nie ma znaczenia bez dostarczenia wartości parametru <i>frequency</i> .
end	Punkt czasowy zarejestrowania ostatniej obserwacji. Jest wyspecyfikowany analogicznie, jak <i>start</i> . Nie ma znaczenia bez dostarczenie wartości parametru <i>frequency</i> .
frequency	Jeden z ważniejszych parametrów obiektu <i>ts</i> , wpływający na interpretację pozostałych parametrów. Określa liczbę obserwacji na jednostkę czasu (roczna częstotliwość danych)

# Drukowanie szeregu czasowego w R

`print(x, calendar, ...)`

`.preformat.ts(x, calendar, ...)`

Ww. funkcje są wykorzystywane do drukowania wartości szeregu czasowego wraz z informacją o okresach czasowych.

- x            Obiekt reprezentujący wartości szeregu czasowy
- calendar    Włącza/wyłącza wyświetlanie informacji o nazwach miesięcy, kwartałach lub roku podczas drukowania. Wartością domyślną jest TRUE dla częstotliwości (*frequency*) 4 lub 12, w przeciwnym razie FALSE. Jeżeli parametr ma wartość TRUE a częstotliwość jest różna od 4 i 12, to punkty czasowe są sygnowane z wykorzystaniem symboli: p1, p2 itd. (p to period).
- ...            Dodatkowe parametry funkcji *print*

# Typy szeregów czasowych R

**ts** - time series otrzymane z:  
data(AirPassengers)

	V1
1	112
2	118
3	132
4	129
5	121
6	135
7	148
8	148
9	136
10	119
11	104
12	118
13	115
14	126
15	141
16	135
17	125
18	149
19	170

**mts** - multivariate time series  
otrzymane z: data(EuStockMarkets)

	DAX	SMI	CAC	FTSE
1	1628.75	1678.1	1772.8	2443.6
2	1613.63	1688.5	1750.5	2460.2
3	1606.51	1678.6	1718.0	2448.2
4	1621.04	1684.1	1708.1	2470.4
5	1618.16	1686.6	1723.1	2484.7
6	1610.61	1671.6	1714.3	2466.8
7	1630.75	1682.9	1734.5	2487.9
8	1640.17	1703.6	1757.4	2508.4
9	1635.47	1697.5	1754.0	2510.5
10	1645.89	1716.3	1754.3	2497.4
11	1647.84	1723.8	1759.8	2532.5
12	1638.35	1730.5	1755.5	2556.8
13	1629.93	1727.4	1758.1	2561.0
14	1621.49	1733.3	1757.5	2547.3
15	1624.74	1734.0	1763.5	2541.5
16	1627.63	1728.3	1762.8	2558.5
17	1631.99	1737.1	1768.9	2587.9
18	1621.18	1723.1	1778.1	2580.5
19	1613.42	1723.6	1780.1	2579.6

# Wybrane operacje na szeregu czasowym 1

- **start()** - zwraca punkt czasowy (datę) pierwszej obserwacji
- **end()** - zwraca punkt czasowy (datę) ostatniej obserwacji
- **frequency()** - zwraca częstotliwość próbkowania szeregu; liczba obserwacji w jednostce cyklu (czasu), np. 12 dla danych miesięcznych (cykl miesięczny)
- **cycle()** - zwraca szereg czasowy zawierający numery obserwacji w kolejnych cyklach (jednostkach czasu), np. 1 dla stycznia itd.
- **time()** - zwraca szereg czasowy punktów czasowych obserwacji (współrzędna x używana na wykresie), dla których były notowane wartości szeregu (szereg był próbkowany).

## *data(EuStockMarkets)*

Z okna *Global environment*:

Name	Type	Length	Size	Value
EuStockMarkets	mts	7440	59.3KB	TimeSeries [1:1860, 1:4] from 1991 to 1999: 1629 1614

Jedna kolumna danych obiektu EuStockMarkets zawiera 1860 obserwacji, są 4 kolumny, co czyni wielkość (długość) całego zasobu równą  $1860 * 4 = 7440$ .

Liczba rekordów (obserwacji):  $(260 - 130 + 1)$  (liczba obserwacji w roku 1991) +  $6 * 260$  (liczba pełnych lat:  $1997 - 1992 + 1 = 6 * 260$ ) +  $169$  (liczba obserwacji w roku 1998) =  $131 + 1560 + 169 = 1860$

`start(EuStockMarkets)` → 1991 130

`end(EuStockMarkets)` → **1998** 169

`frequency(EuStockMarkets)` → 260 (552 tygodnie \* 5 dni rejestrowania kursu = 260)

# Wybrane operacje na szeregu czasowym 2

cycle(EuStockMarkets)



```
Time Series:
Start = c(1991, 130)
End = c(1998, 169)
Frequency = 260
 [1] 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148
 [20] 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167
 [39] 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186
 [58] 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205
 [77] 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224
 [96] 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243
[115] 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 1 2
[134] 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
[153] 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
[172] 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59
[191] 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78
[210] 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97
[229] 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116
[248] 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135
[267] 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154
[286] 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173
[305] 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192
[324] 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211
[343] 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230
[362] 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249
[381] 250 251 252 253 254 255 256 257 258 259 260 1 2 3 4 5 6 7 8
[400] 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27
```

```
Time Series:
Start = c(1991, 130)
End = c(1998, 169)
Frequency = 260
 [1] 1991.496 1991.500 1991.504 1991.508 1991.512 1991.515 1991.519 1991.523
 [9] 1991.527 1991.531 1991.535 1991.538 1991.542 1991.546 1991.550 1991.554
[17] 1991.558 1991.562 1991.565 1991.569 1991.573 1991.577 1991.581 1991.585
[25] 1991.588 1991.592 1991.596 1991.600 1991.604 1991.608 1991.612 1991.615
[33] 1991.619 1991.623 1991.627 1991.631 1991.635 1991.638 1991.642 1991.646
[41] 1991.650 1991.654 1991.658 1991.662 1991.665 1991.669 1991.673 1991.677
[49] 1991.681 1991.685 1991.688 1991.692 1991.696 1991.700 1991.704 1991.708
[57] 1991.712 1991.715 1991.719 1991.723 1991.727 1991.731 1991.735 1991.738
[65] 1991.742 1991.746 1991.750 1991.754 1991.758 1991.762 1991.765 1991.769
[73] 1991.773 1991.777 1991.781 1991.785 1991.788 1991.792 1991.796 1991.800
[81] 1991.804 1991.808 1991.812 1991.815 1991.819 1991.823 1991.827 1991.831
[89] 1991.835 1991.838 1991.842 1991.846 1991.850 1991.854 1991.858 1991.862
```

← time(EuStockMarkets)

# Wybrane operacje na szeregu czasowym 3

## Wycinek szeregu czasowego

`window(x, start = NULL, end = NULL, frequency = NULL, deltat = NULL, extend = FALSE, ts.eps = getOption("ts.eps"), ...) <- value` **przerwane**

## Parametry

x	Szereg czasowy, źródłowy
start	Startowy punkt czasowy definiujący początek wycinka szeregu źródłowego
end	Końcowy punkt czasowy definiujący koniec wycinka szeregu źródłowego
frequency/ deltat	Określenie nowej częstotliwości próbkowania dla wycinka szeregu czasowego
extend	Wartość logiczna określającą, czy wycinek może być poszerzony
ts.eps	Tolerancja porównania szeregu czasowego, jak w funkcji <i>ts</i> .
value	Wartości zastępcze

```
Time series:
Start = c(1995, 1)
End = c(1998, 1)
Frequency = 260
      DAX      SMI      CAC      FTSE
1995.000 2110.77 2673.5 1956.0 3083.4
1995.004 2097.34 2656.2 1927.8 3095.8
1995.008 2074.68 2628.8 1894.2 3065.6
1995.012 2097.51 2628.8 1881.2 3065.5
1995.015 2079.19 2628.8 1881.2 3065.5
1995.019 2068.92 2612.3 1885.9 3065.7
1995.023 2072.90 2632.4 1901.8 3051.6
1995.027 2051.46 2613.1 1871.5 3032.3
1995.031 2058.20 2622.3 1886.4 3065.0
1995.035 2053.41 2617.3 1864.2 3055.8
1995.038 2062.08 2600.4 1859.2 3060.4
1995.042 2061.76 2597.0 1849.1 3049.4
1995.046 2059.68 2600.5 1844.1 3033.2
1995.050 2064.14 2600.2 1854.0 3048.3
1995.054 2088.25 2591.6 1872.8 3076.7
1995.058 2081.39 2582.8 1856.9 3054.0
1995.062 2085.62 2593.6 1860.3 3054.9
1995.065 2079.60 2595.5 1837.1 3028.6
1995.069 2050.86 2574.6 1813.3 2995.9
```

`EUCut<-window(EuStockMarkets, start=1995, end=1998)` →

`EUCut`

Przy dużej sumarycznie liczbie wartości wyświetlanych jednorazowo, przekraczających opcję *max.print*, jest wyświetlany komunikat. Należy zmienić ustawienia tej opcji, np. `options(max.print=4000)`

```
1995.958 2277.70 3261.6 1849.6 3652.1
[ reached getoption("max.print") -- omitted 531 rows ]
```



# Wybrane operacje na szeregu czasowym 4

## Łączenie szeregów czasowych

`ts.union(..., dframe = FALSE)`

`ts.intersect(..., dframe = FALSE)`

Łączenie (scalenie) szeregów czasowych o takiej samej częstotliwości.

Za pomocą metody *union* łączenie rozszerza dziedzinę szeregu wynikowego (odpowiednik sumy mnogościowej zbiorów).

Za pomocą metody *intersect* łączenie zawęża dziedzinę szeregu wynikowego (odpowiednik iloczynu mnogościowego zbiorów).

Dane o liczbie zgonów miesięcznie z powodu chorób płuc są zapisane w zbiorze

*UKLungDeaths*. W tym zasobie są zawarte obiekty reprezentujące szeregi czasowe:

*ldeaths* (dane o zgonach osób obojga płci), *fdeaths* (dane o zgonach kobiet), *mdeaths*

(dane o zgonach osób mężczyzn). Aby uzyskać do nich dostęp należy załadować

zbiór *UKLungDeaths* (ww. zbiory są obiektami *promise*), po czym korzystać już z poszczególnych zasobów.

```
data(UKLungDeaths)
```

```
print(ldeaths, calendar = TRUE)
```

```
print(fdeaths, calendar = TRUE)
```

Struktury tych zasobów są takie same. Można więc przetestować na nich lub ich podzbiorach ww. operacje