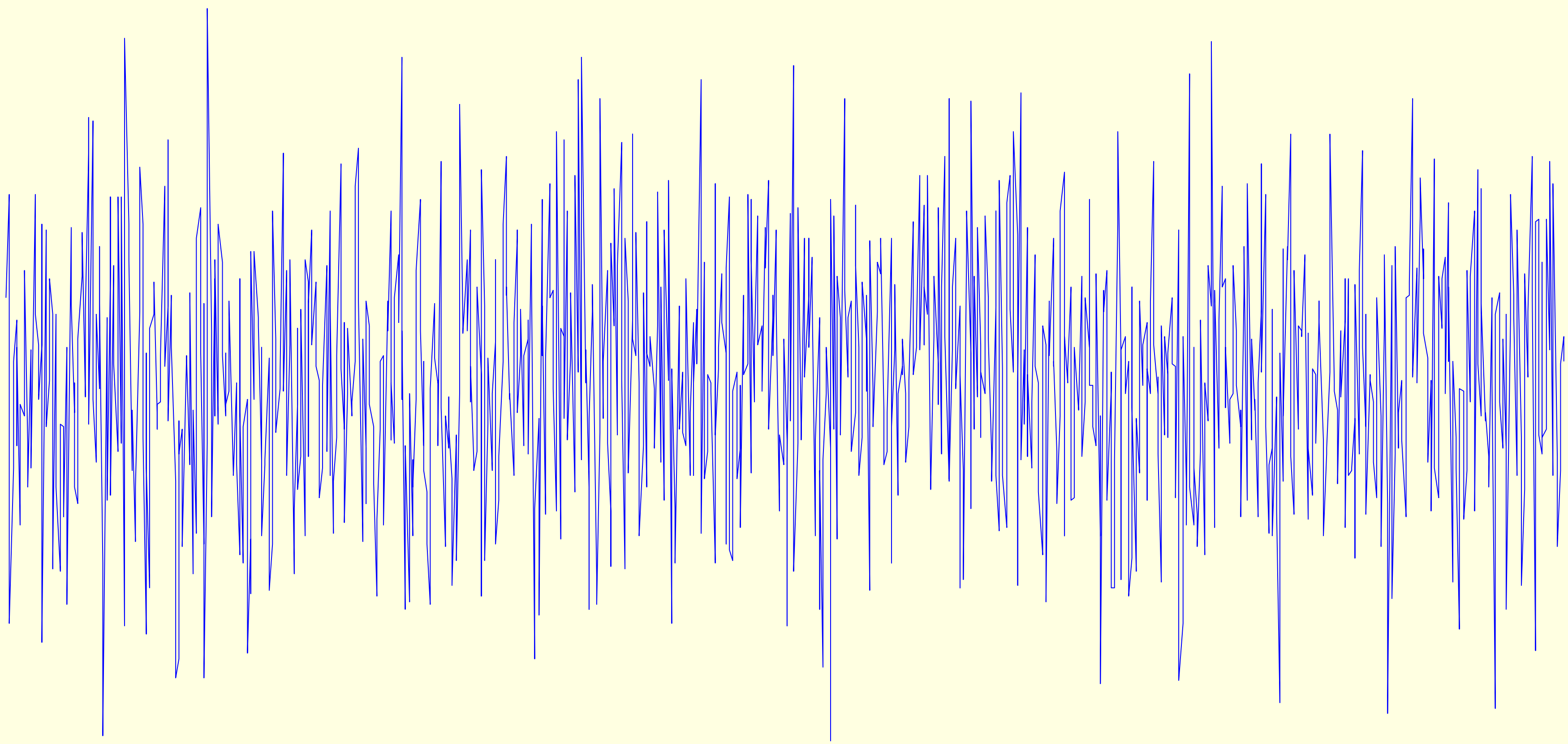
Sygnały losowe i ich analiza



Paweł Strumiłło, Instytut Elektroniki Politechniki Łódzkiej

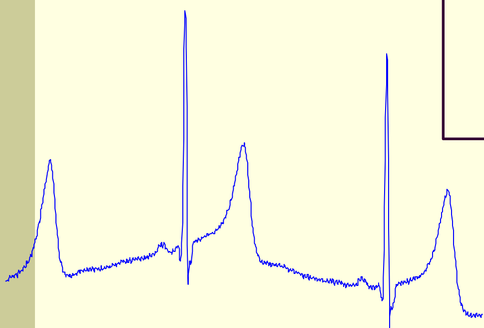
Sygnały biologiczne

Modele deterministyczne

Modele stochastyczne

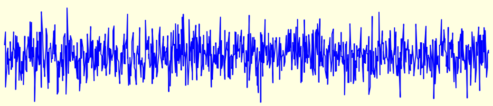
$s(t)$

$n(t)$



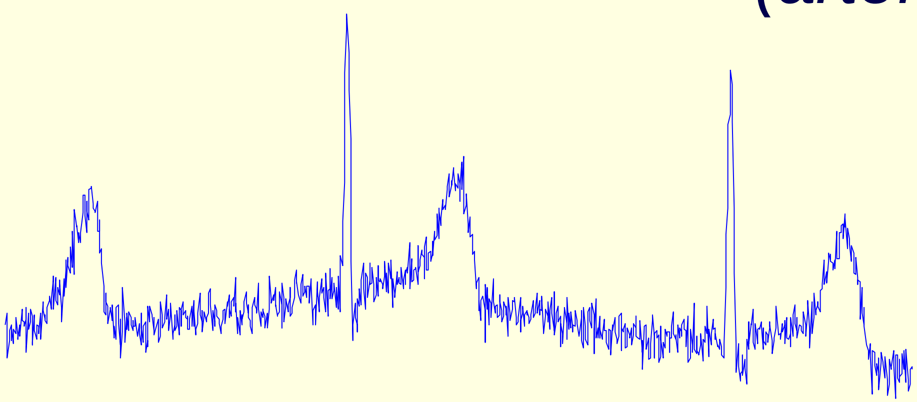
EKG

Sygnały biologiczne



zakłócenia EMG (artefakty)

$$x(t) = s(t) + n(t)$$



Modele sygnałów

Przykłady!

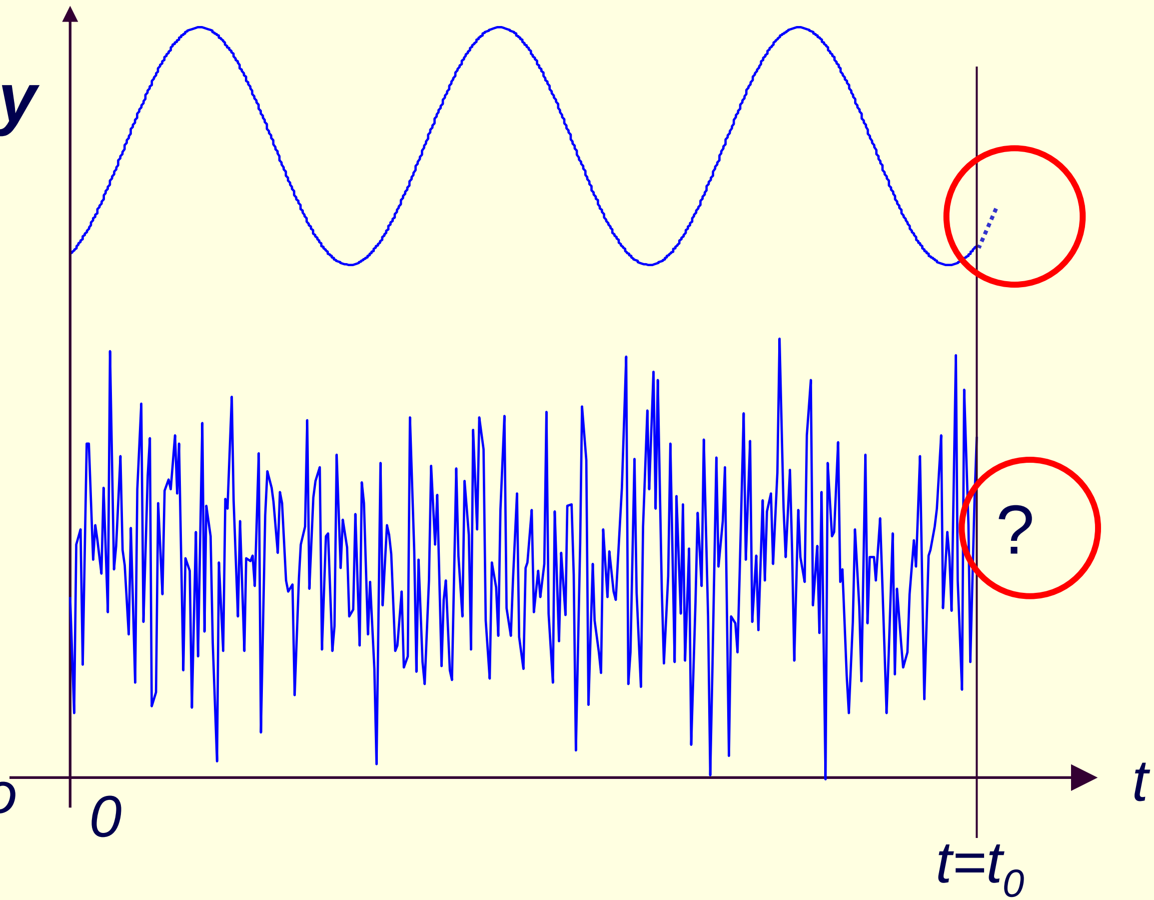
- **Sygnały**
  - Modele deterministyczne
    - Okresowe
      $x(t) = x(t + nT)$
      - Harmoniczne
      - Poliharmoniczne
      - Złożone
    - Nieokresowe
      - Quasi-okresowe
      - Przejściowe
      - Inne
  - Modele stochastyczne
    - **Stacjonarne**
      - Ergodyczne
      - Nieergodyczne
    - Niestacjonarne

Modele chaosu deterministycznego

Sygnał losowy a sygnał stochastyczny

sygnał deterministyczny
(próbki można przewidywać z dużą dokładnością)

sygnał losowy
(niemożliwe przewidywanie wartości próbek sygnału, można tylko określić zadane parametry sygnału)



Paweł Strumiłło, Instytut Elektroniki Politechniki Łódzkiej

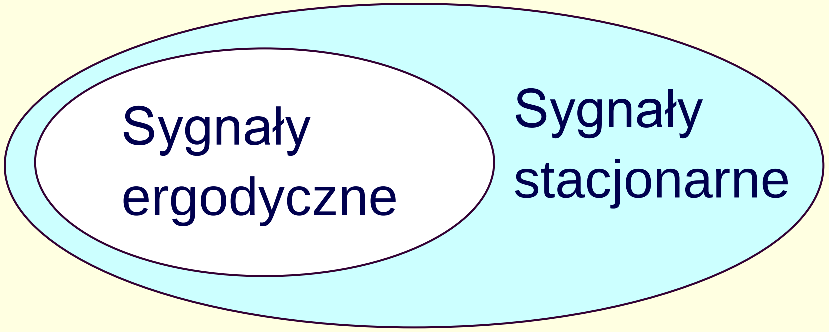
Sygnały losowe stacjonarne i ergodyczne

Sygnały stacjonarne

Momenty statystyczne wyznaczane dla danej realizacji w czasie (np. średnia, wariancja) nie podlegają zmianom w czasie.

Sygnały ergodyczne

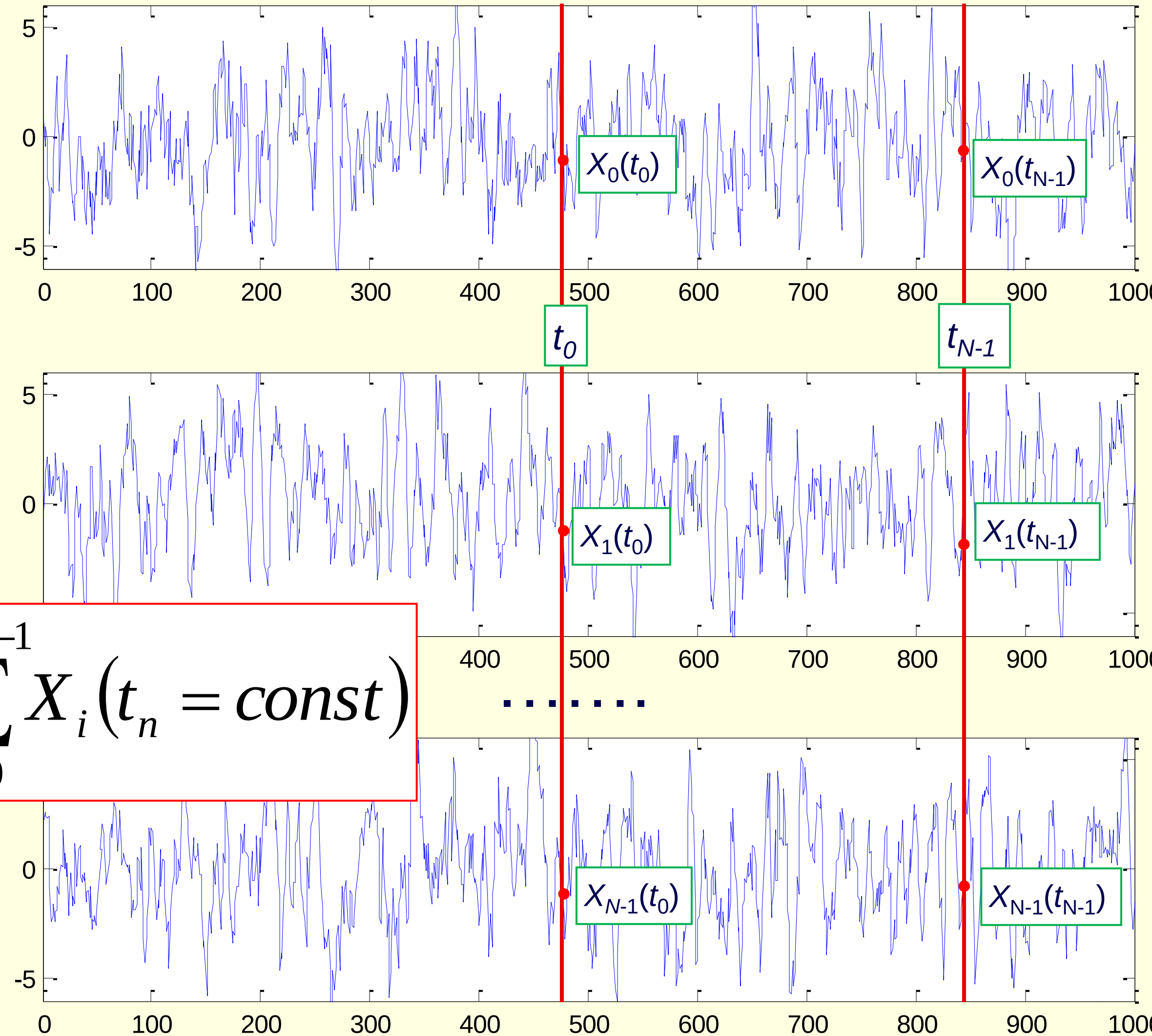
Momenty statystyczne (np. średnia, wariancja) wyznaczane w czasie i dla różnych zbiorów realizacji są sobie równe.



Sygnały losowe stacjonarne i ergodyczne

Sygnał losowy jest ergodyczny jeżeli:

$$\frac{1}{N}\sum_{t=0}^{t=N-1} X_{i=const}(t) = \frac{1}{N}\sum_{i=0}^{i=N-1} X_i(t_n = const)$$



Sygnał losowy (parametry)

Dokładny przebieg **sygnału losowego**, nie jest możliwy do wyznaczenia - można charakteryzować tylko jego odpowiednio zdefiniowane parametry statystyczne.

Jako zapis (obserwacja) pewnego procesu losowego,

sygnał losowy można interpretować jako zmienną losową i opisywać ją za pomocą wielkości i zależności probabilistycznych tj.

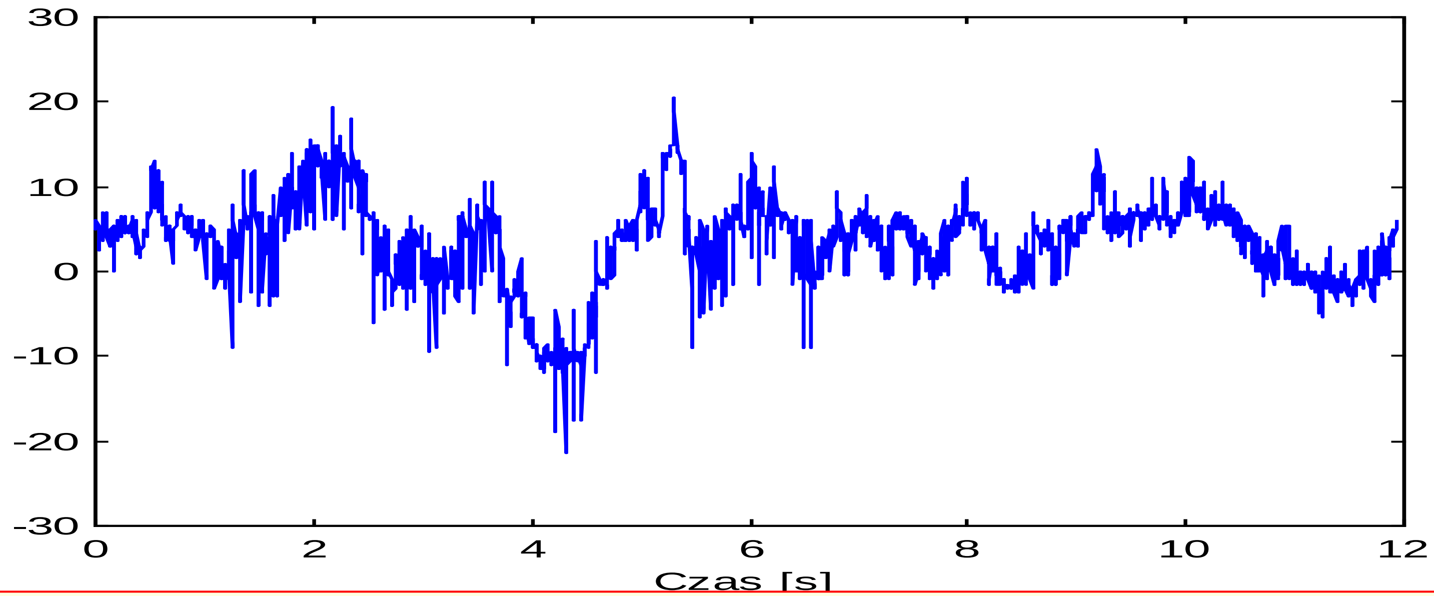
- **wartość oczekiwana,**
- **odchylenie standardowe,**
- **funkcja autokorelacji.**

Dla **stacjonarnych** sygnałów losowych parametry te nie zmieniają się w funkcji czasu.

Przykład - sygnał EEG

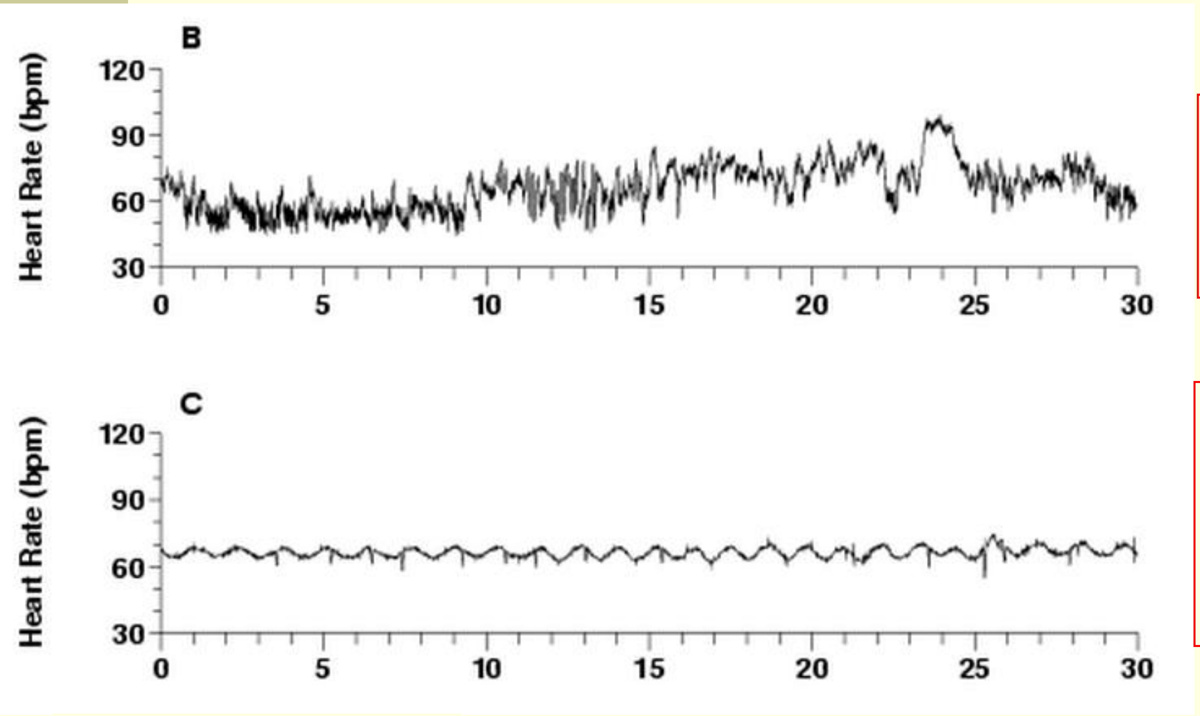
Dla sygnału elektroencefalograficznego można zastosować model stochastyczny.

Sygnał EEG jest wynikiem aktywności elektrycznej wielkiej liczby* komórek neuronowych.



*) W niektórych pracach dowodzi się, że procesy percepcji neuronowej polegają na synchronizacji czasowej specjalizowanych populacji komórek neuronowych

Rytm serca

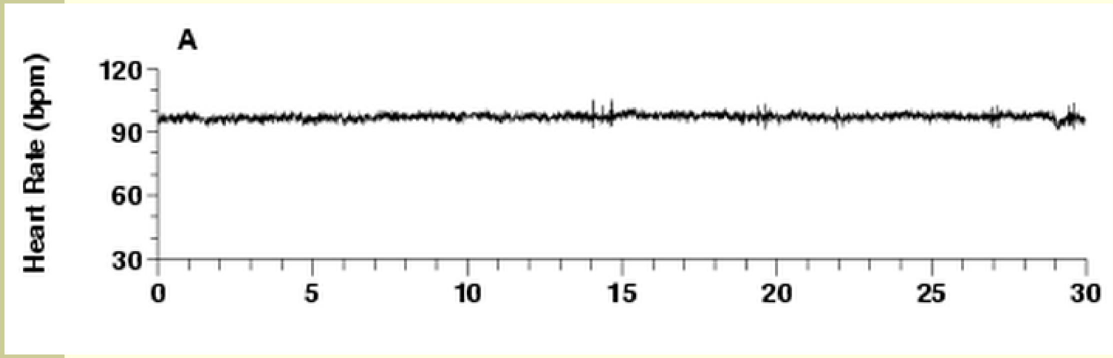


Rytm prawidłowy – charakter zmienności o naturze „fraktalnej”

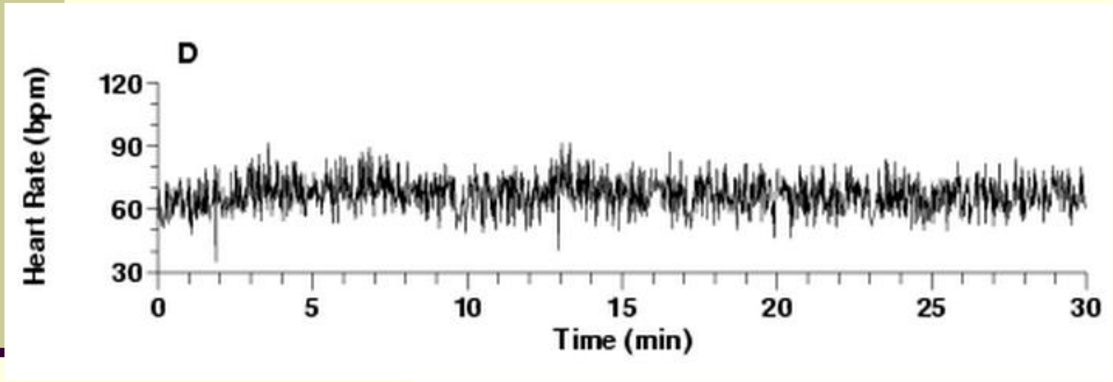
Rytm nieprawidłowy (niewydolność serca) – widoczne oscylacje szybkości rytmu (~1/min)

www.physionet.org

Rytm serca



Rytm serca w poważnej niewydolności serca – rytm zbyt regularny



Arytmia serca (fibrylacja przedsionków) – sygnał ma charakter losowy

www.physionet.org

Sygnał losowy – wartość oczekiwana

Wartość oczekiwana - sygnał ciągły:

$$E\{x(t)\} = \int_{-\infty}^{+\infty} x p(x) dx = x$$

```
#Python
#see also random?
x=random.randn(1000)
expected=x.mean()
```

Wartość oczekiwana (estymata)- sygnał dyskretny:

$$E\{x(n)\} = \sum_{n=-N}^{n=N} x(n)p(n) = \frac{1}{2N+1}\sum_{n=-N}^{n=N} x(n) = x$$

kolejne próbki sygnału są jednakowo prawdopodobne

Sygnał losowy - wariancja

Wariacja (*średnia kwadratów odchyleń od wartości oczekiwanej***)- sygnał ciągły:**

$$\sigma^2 = E\left\{\left[x(t) - x\right]^2\right\} = \int_{-\infty}^{+\infty} \left[x(t) - x\right]^2 p(x)dx$$

```
#Python
x=random.randn(1000)
std_dev=x.std()
variance=x.var()
```

Wariancja - sygnał dyskretny:

$$\sigma^2 = \frac{1}{2N+1}\sum_{n=-N}^{n=N}\left[x(n) - x\right]^2 = \frac{1}{2N+1}\sum_{n=-N}^{n=N} x^2(n) - x^2$$

Sygnał losowy – funkcja autokorelacji

Sygnał ciągły:

$$R(\tau) = \int_{-\infty}^{\infty} x(t)x(t+\tau)dt$$

Sygnał dyskretny:

$$R(m) = \sum_{n=-\infty}^{n=\infty} x(n)x(n+m)$$

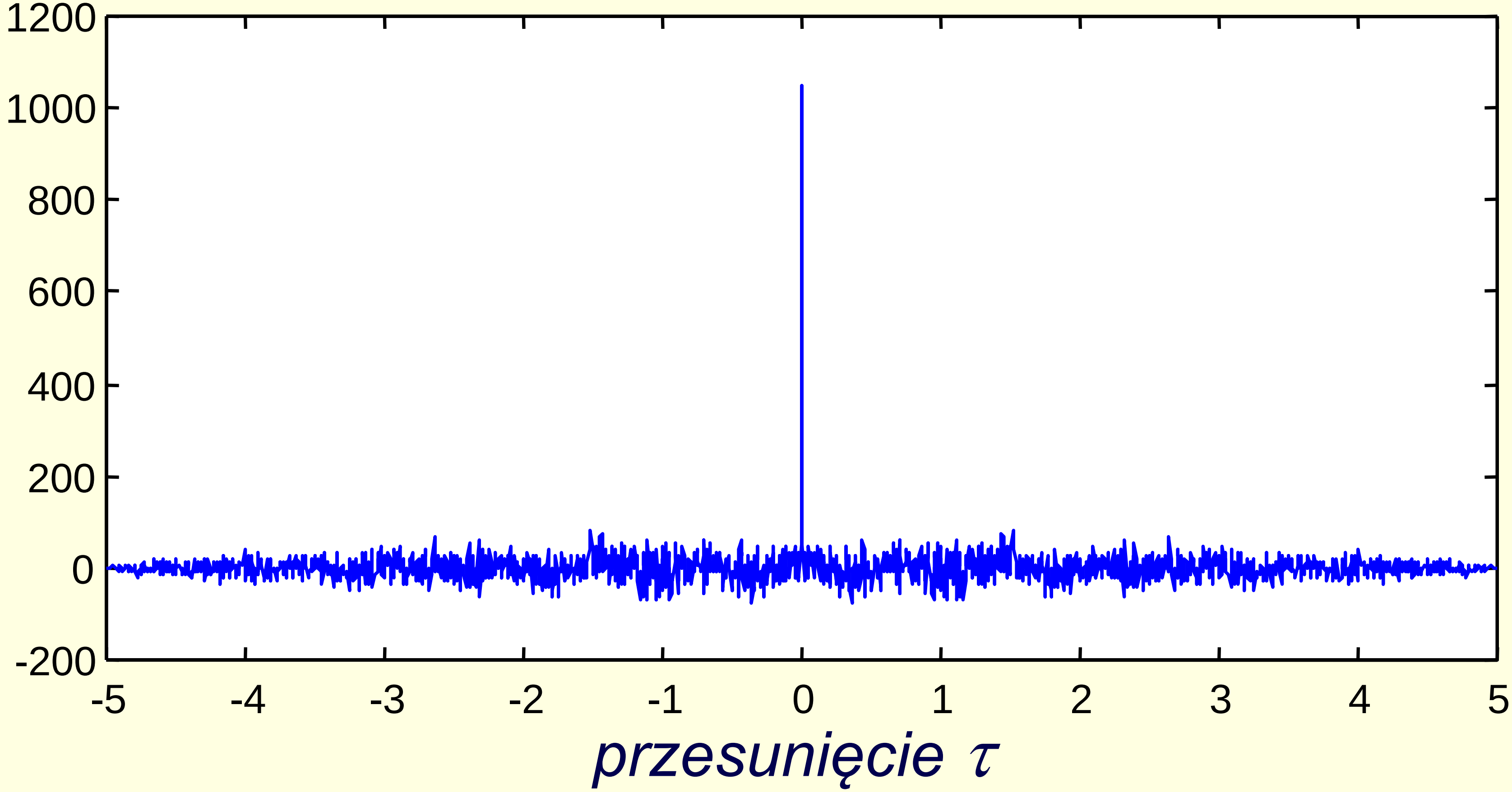
```
#Python
x=random.randn(1000)
R=correlate(x,x,mode='full')
plot(R)

#usuń składową DC
x=x-x.mean()
R=correlate(x,x,mode='full')
plot(R)
```

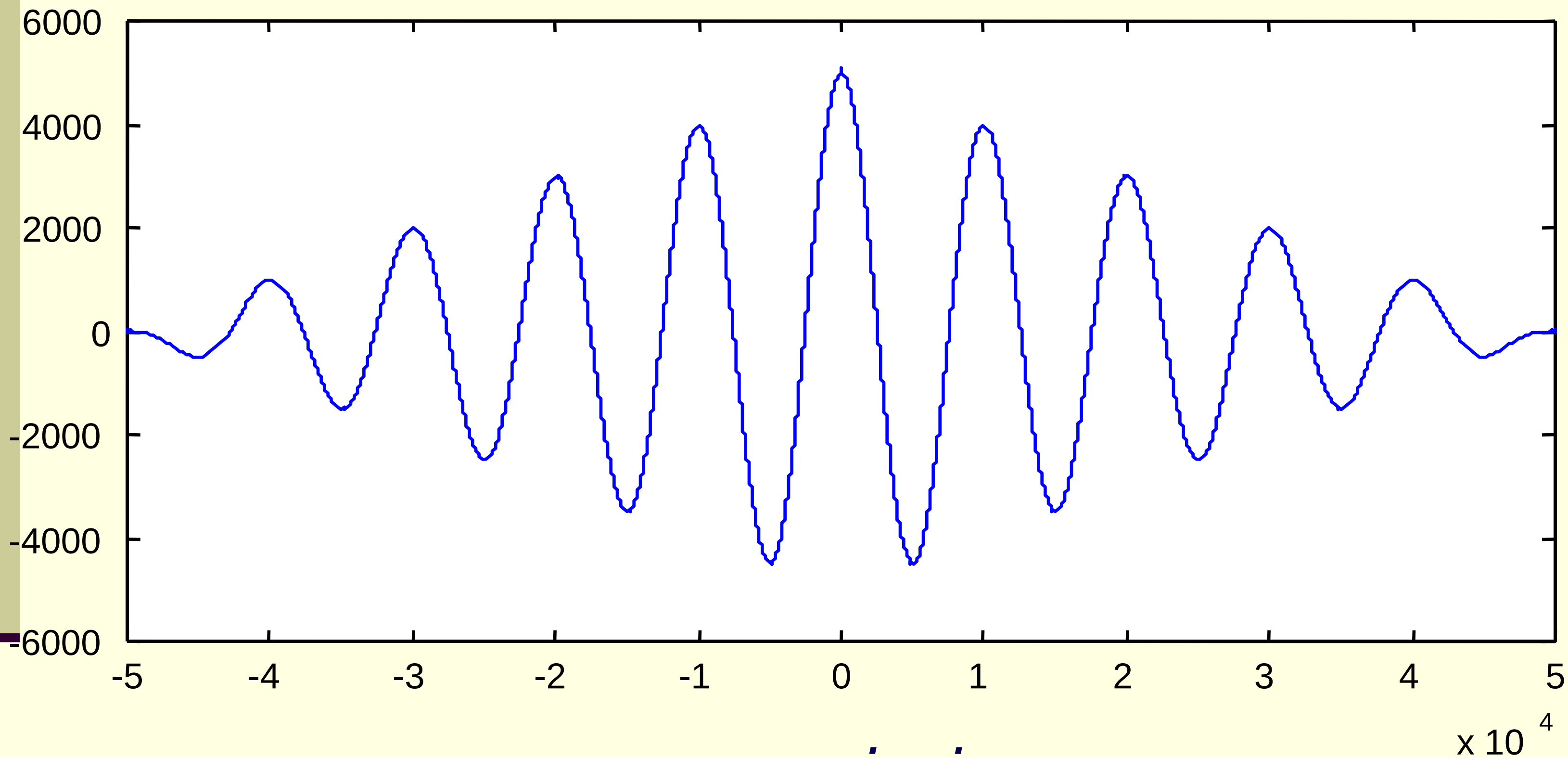
Pamiętaj o usunięciu składowej stałej sygnału!

Paweł Strumiłło, Instytut Elektroniki Politechniki Łódzkiej

Funkcja autokorelacji sygnału losowego
(*rozkład Gaussa* $N(\mu, \sigma)=N(0,1)$)



Funkcja autokorelacji sygnału harmonicznego



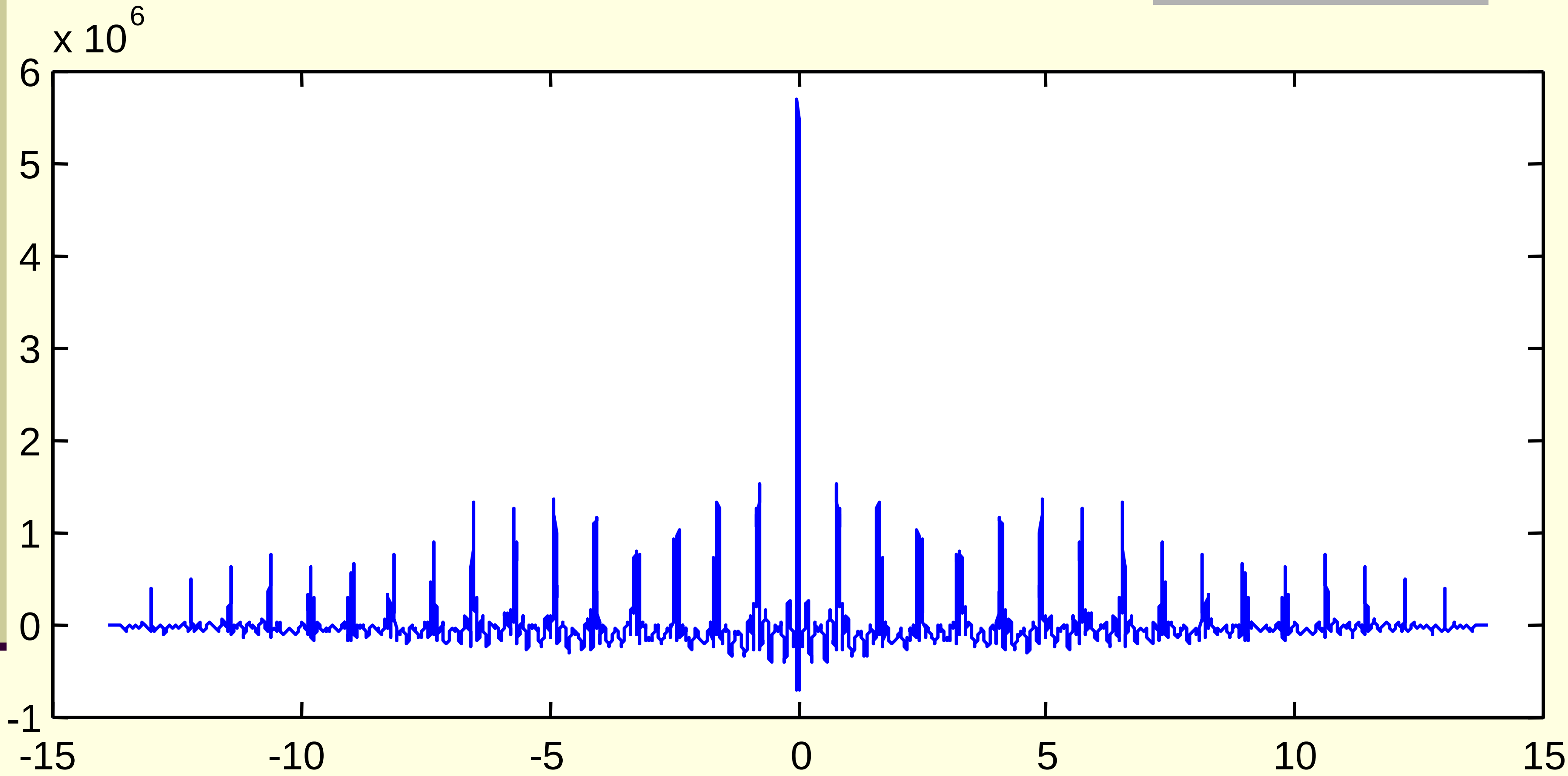
przesunięcie τ

Funkcja autokorelacji - ćwiczenie

Ćwiczenie komputerowe: funkcja autokorelacji

1. Załaduj zadany sygnał EKG (‘`ecg_s.mat`’).
2. Napisz procedurę do wyznaczania i wyświetlania funkcji autokorelacji sygnału dyskretnego o zadanej długości *N*.
3. Porównaj długość otrzymanej funkcji autokorelacji z długością sygnału.
4. Jakie własności funkcji autokorelacji zauważasz?
5. Zastanów się jakie zastosowanie może mieć funkcja autokorelacji w przetwarzaniu sygnałów?

Funkcja autokorelacji sygnału EKG



przesunięcie τ [s]

Funkcja korelacji wzajemnej

Sygnały ciągłe:

$$R_{xy}(\tau) = \int_{-\infty}^{\infty} x(t) y(t+\tau) dt$$

```
#Python
R=correlate(x,y,mode='full')
```

Sygnały dyskretne:

$$R_{xy}(m) = \sum_{n=-\infty}^{n=\infty} x(n) y(n+m)$$

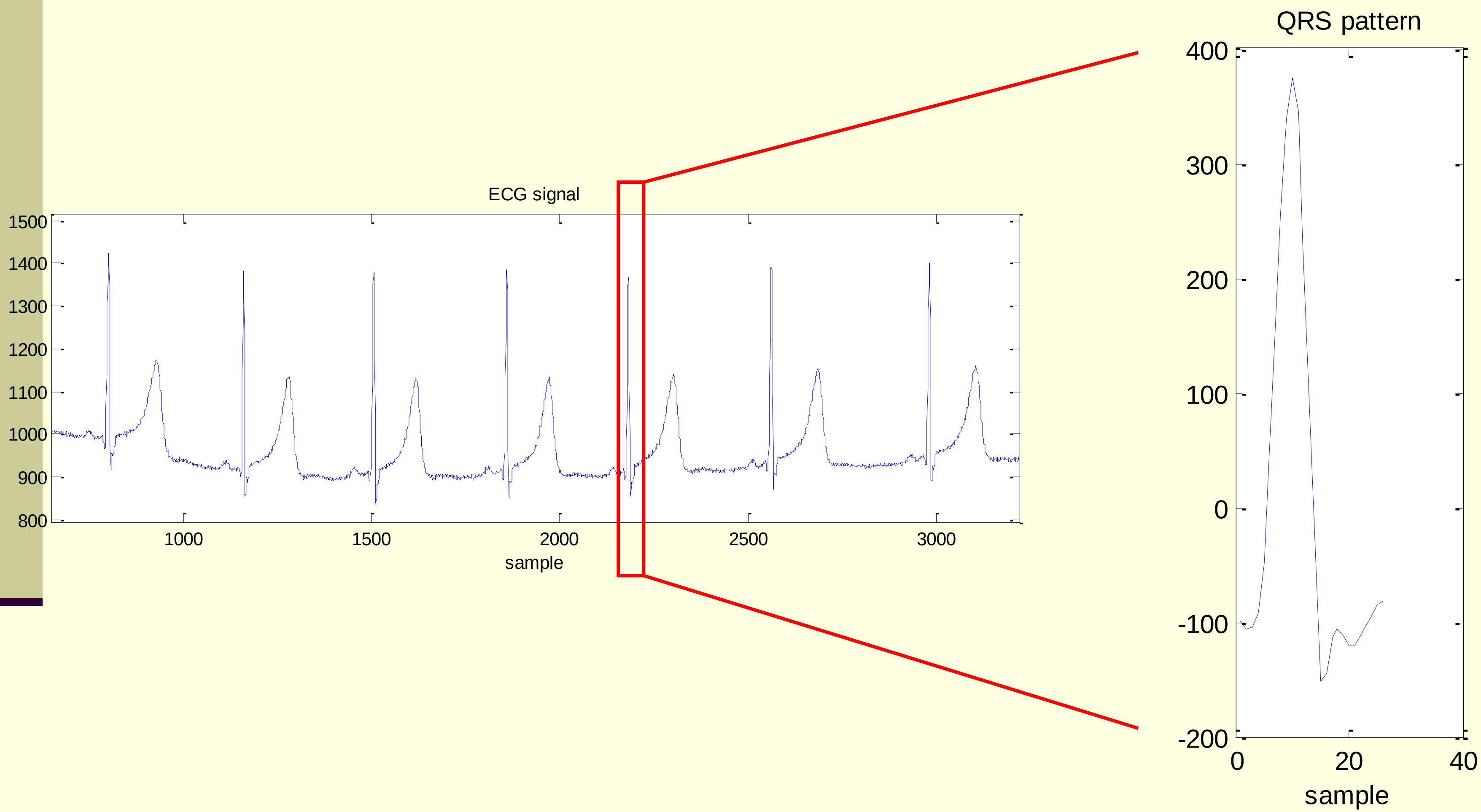
Pamiętaj o usunięciu składowej stałej sygnałow!

Funkcja korelacji wzajemnej - ćwiczenie

Ćwiczenie komputerowe: Detekcja zespołu QRS za pomocą wzorca QRS

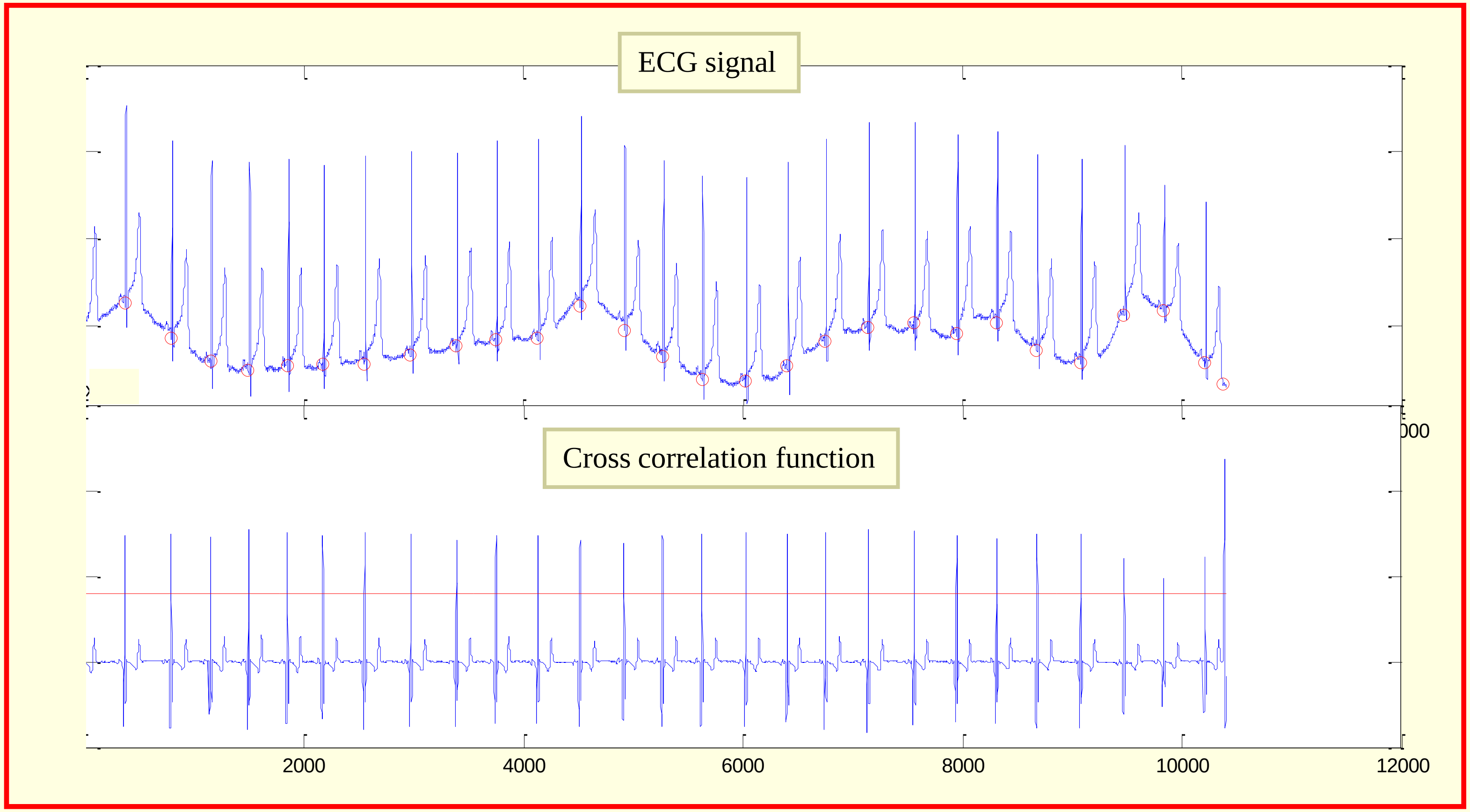
1. Załaduj zadany sygnał EKG (‘`ecg_mit.mat`’).
2. Wyznacz wzorzec zespołu QRS. Zastosuj ten wzorzec do detekcji zespołów QRS w sygnale EKG.
3. Zachowaj otrzymane wyniki do porównania z innymi metodami detekcji zespołu QRS w sygnale EKG.

Funkcja korelacji wzajemnej



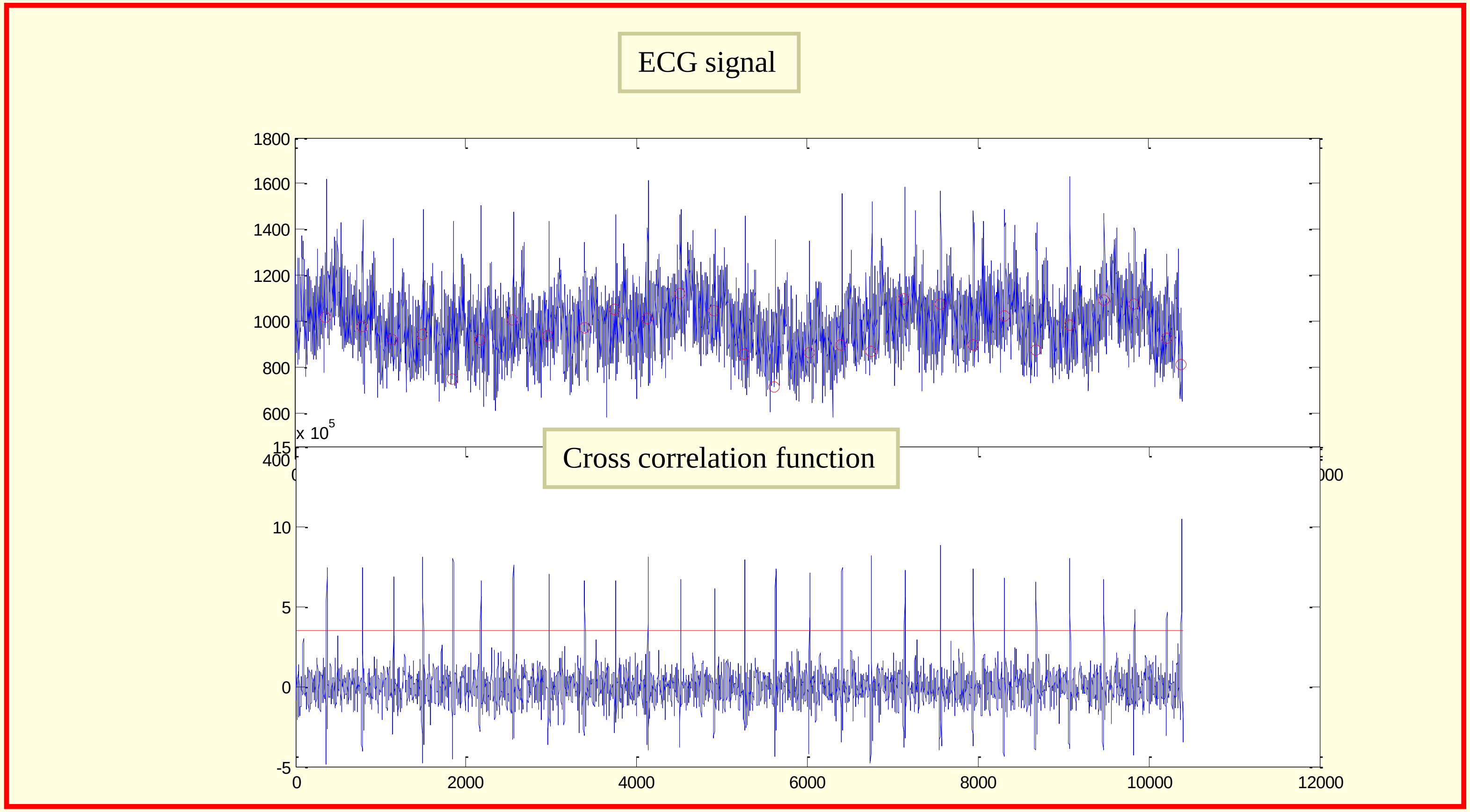
Funkcja korelacji wzajemnej

```
kor=xcorr(ECGsignal,QRSpattern)
```



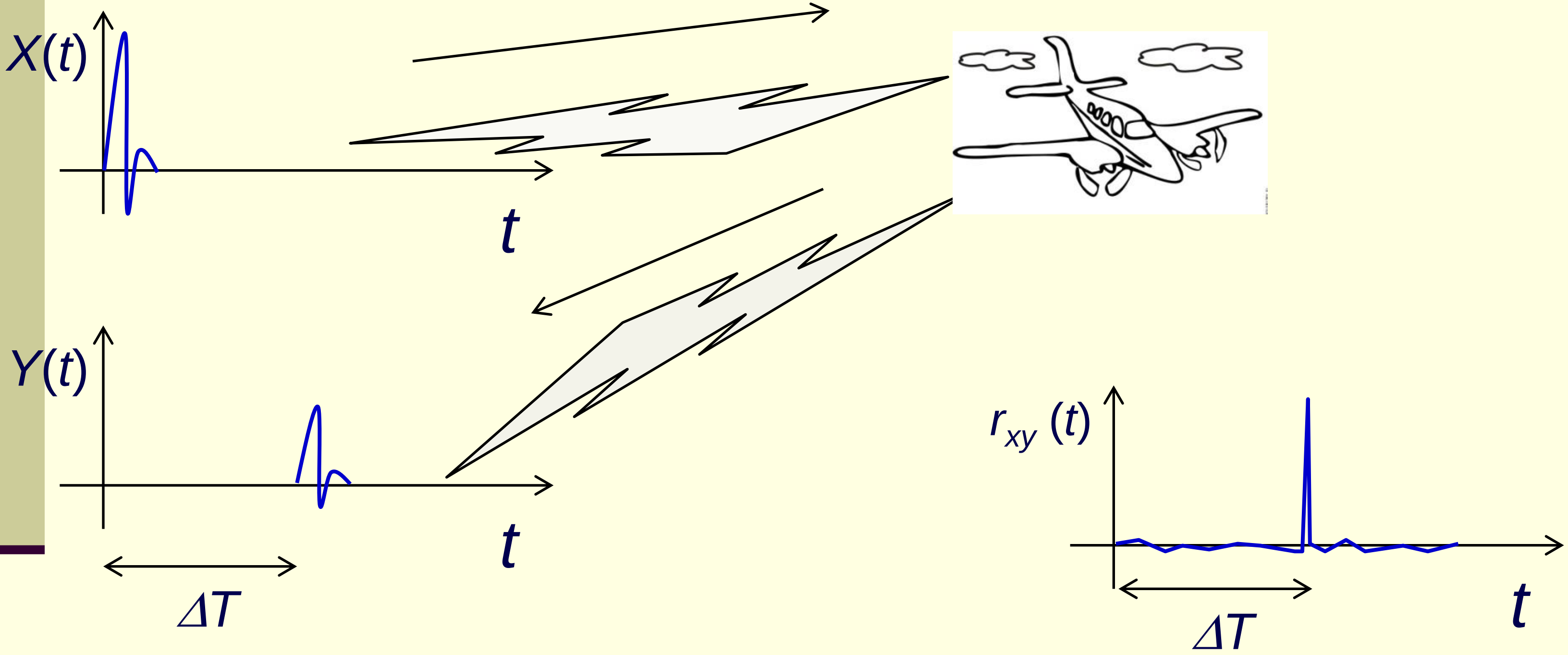
Funkcja korelacji wzajemnej

kor=xcorr(ECGsignal,QRSpattern);



Funkcja korelacji wzajemnej

Wyznaczanie opóźnienia w sygnale radarowym:



Gęstość widmowa mocy sygnału losowego (ang. *power spectral density - PSD*)

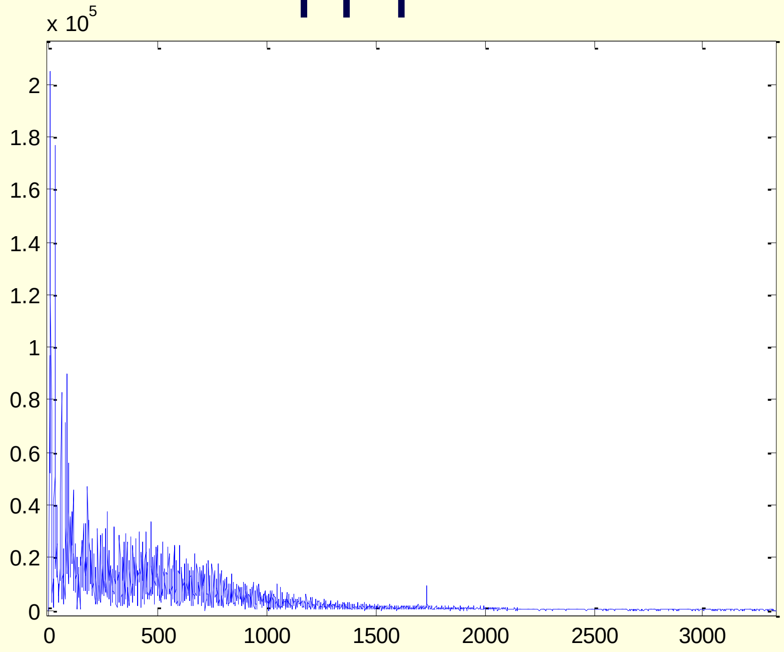
- **Funkcja autokorelacji** sygnału losowego jest przebiegiem **deterministycznym** zatem, można wyznaczyć widmo Fouriera tego przebiegu i badać rozkład mocy sygnału w dziedzinie częstotliwości

$$\boxed{FT\left\{R_{xx}(\tau)\right\} = X(j\omega)X^*(j\omega) = \left|X(j\omega)\right|^2}$$

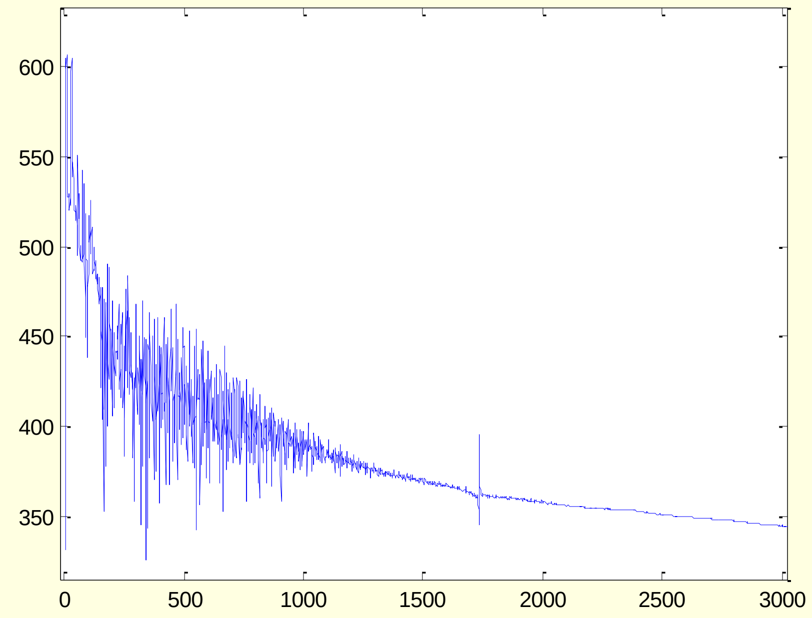
Ćwiczenie komputerowe: Wyznacz PSD sygnału EKG przez wyznaczenie kwadratu widma a następnie zastosuj funkcję ‘`psd`’. Porównaj otrzymane wyniki.

Power Spectral Density - PSD

FFT



PSD



Współczynnik korelacji

Podobieństwo dwóch sygnałów można badać za pomocą współczynnika korelacji:

kowariancja

$$r_{xy} = \frac{\sigma_{xy}}{\sqrt{\sigma_x \sigma_y}} = \frac{\sum_{n=1}^{N} [x(n) - x][y(n) - y]}{\sqrt{\sum_{n=1}^{N} [x(n) - x]^2 \sum_{n=1}^{N} [y(n) - y]^2}}$$

$$-1 \le r_{xy} \le 1$$

```
#Python

corrcoef?
```

Współczynnik korelacji - przykład

Wyznacz współczynnik korelacji pomiędzy danymi *a* i *b*:

$$a = [\ \ 1, 2, 3]$$
$$b = [-3, 4, 2]$$

$$\Longrightarrow \quad r_{xy} \cong 0.6934$$

$$r_{xy} = \frac{\sum_{n=1}^{N}[x(n)-\bar{x}][y(n)-\bar{y}]}{\sqrt{\sum_{n=1}^{N}[x(n)-\bar{x}]^2 \sum_{n=1}^{N}[y(n)-\bar{y}]^2}}$$

Współczynnik korelacji

Ćwiczenie: Utwórz trzy wektory:

1. Wektor x_1 którego elementami jest wzrost osób w grupie wyrażony w metrach
2. Wektor x_2 którego elementami jest waga osób wg porządku z wektora x_1 (nie oszukiwać!)
3. Wektor x_3 którego elementami jest rozmiar buta osób w grupie

Wyznacz: współczynniki korelacji pomiędzy tymi wektorami i wskaż która para wektorów jest „najbardziej podobna"

Rozkłady zmiennych losowych ciągłych

Rozkład równomierny (jednostajny):

$$p(x) = \begin{cases} \dfrac{1}{x_1 - x_0} & dla \quad x_0 \le x \le x_1 \\ 0 & dla \quad innych \quad x \end{cases}$$

Rozkład Laplace'a:

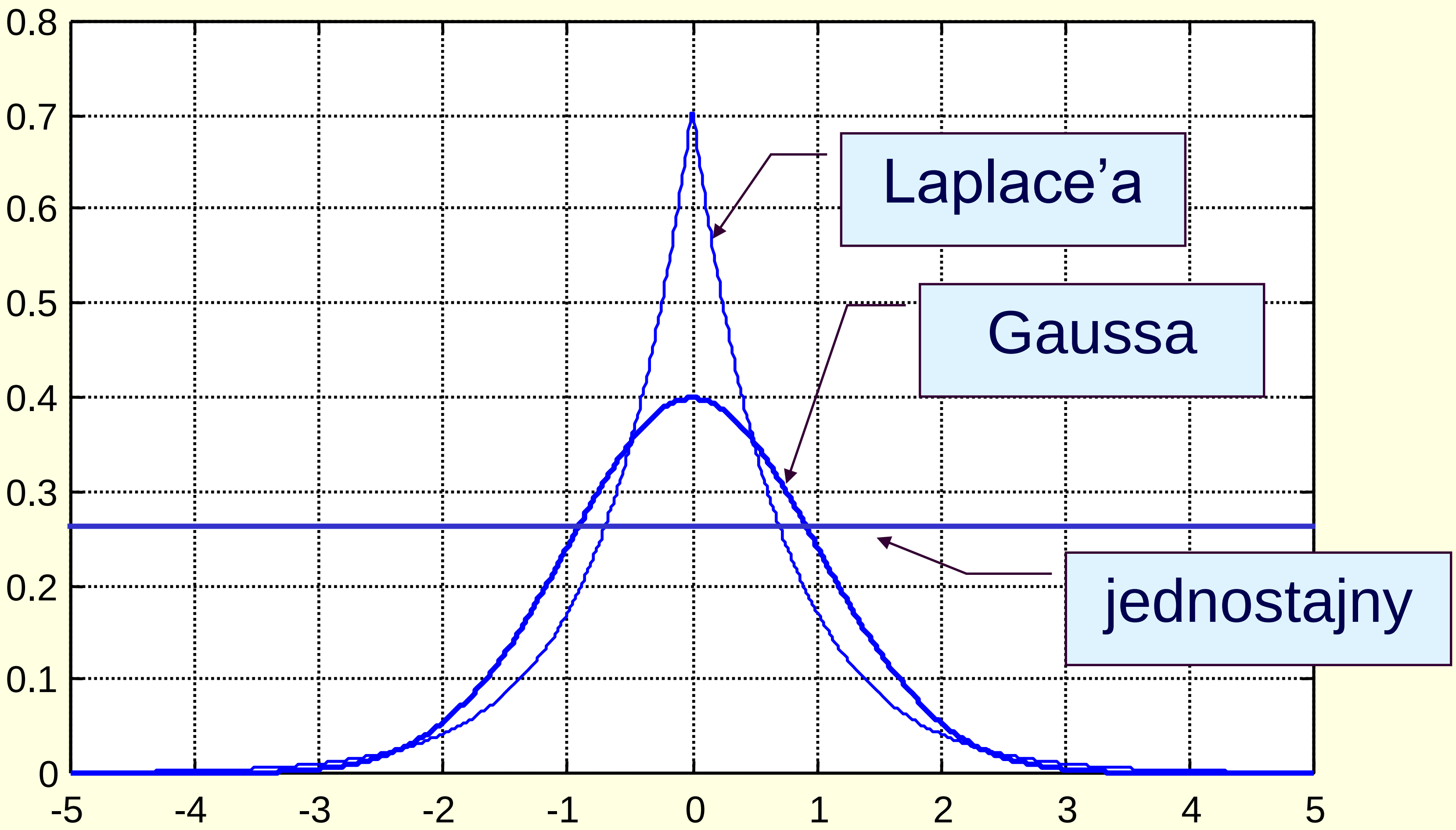
$$p(x) = \frac{1}{\sigma\sqrt{2}} e^{-\frac{\sqrt{2}\left|x-\mu\right|}{\sigma}}$$

Rozkład normalny (Gaussa):

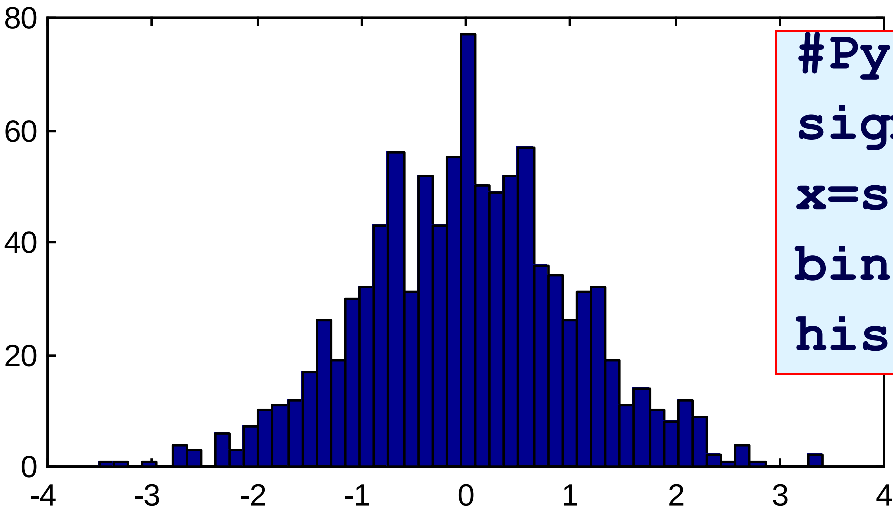
$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Paweł Strumiłło, Instytut Elektroniki Politechniki Łódzkiej

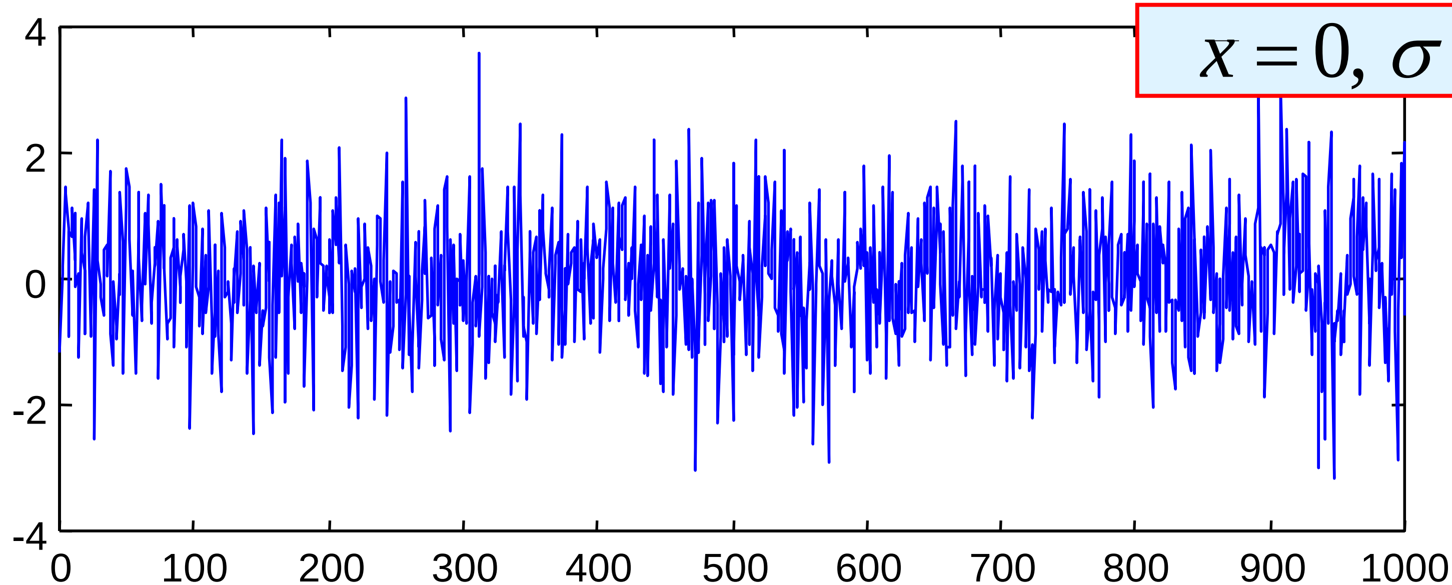
Rozkłady zmiennych losowych



Szum o rozkładzie Gaussa (biały)

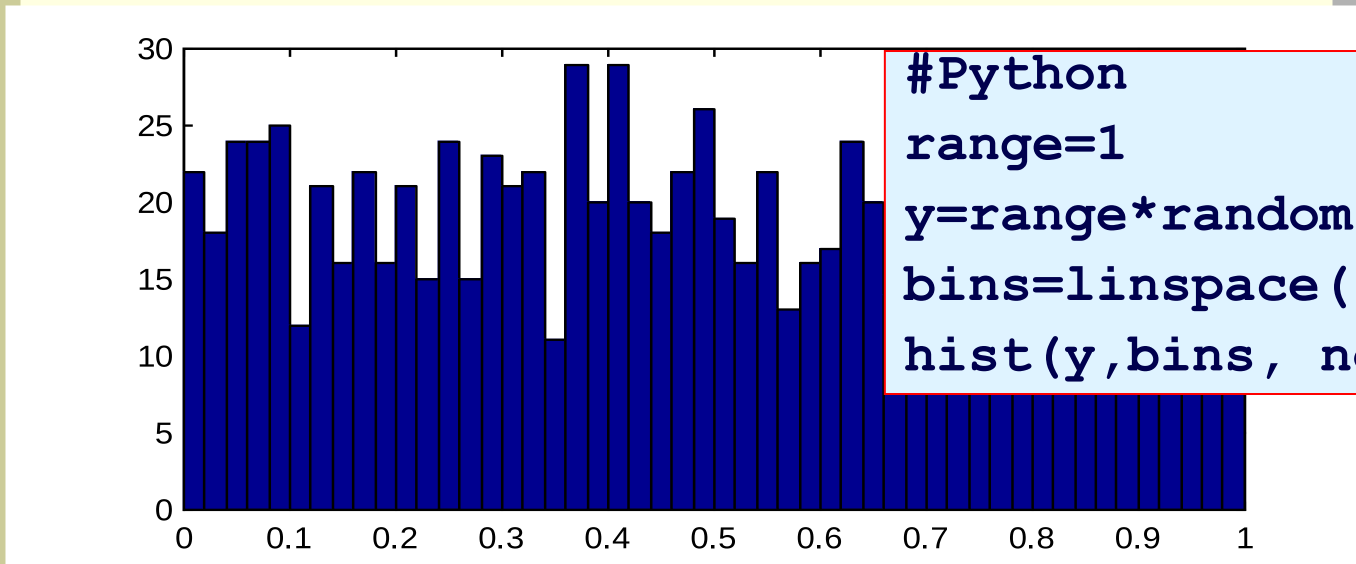



```
#Python
sigma=1
x=sigma*random.randn(1000)
bins=linspace(-5,5,50)
hist(x,bins, normed='True')
```

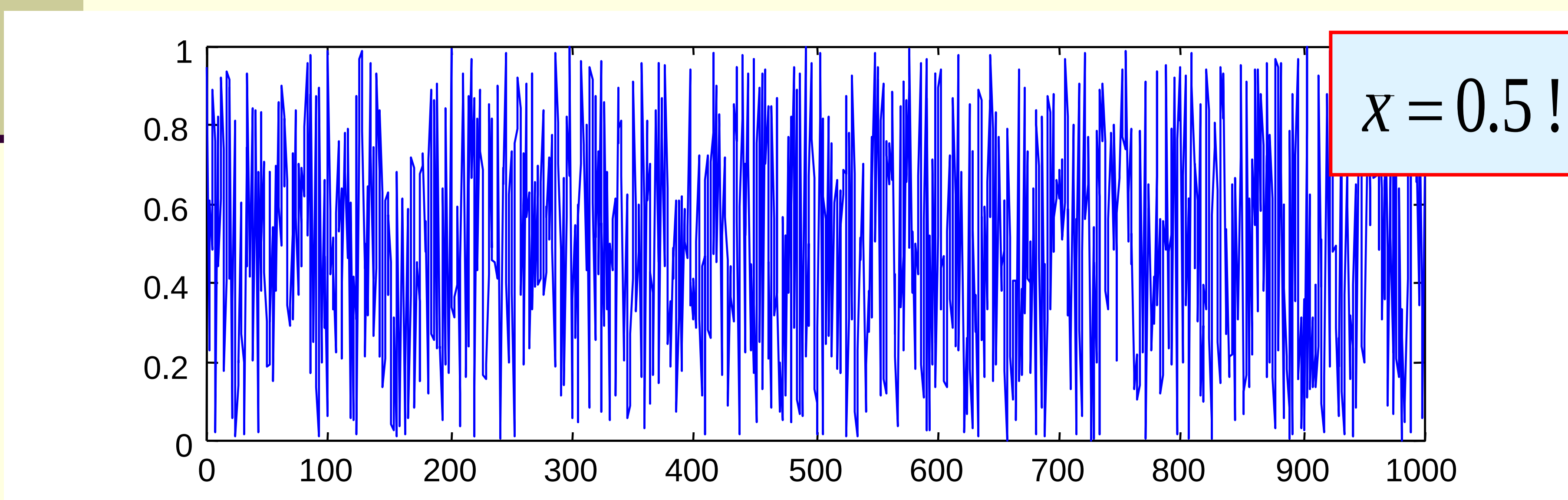


$$x = 0,\ \sigma = 1$$

Szum o rozkładzie równomiernym



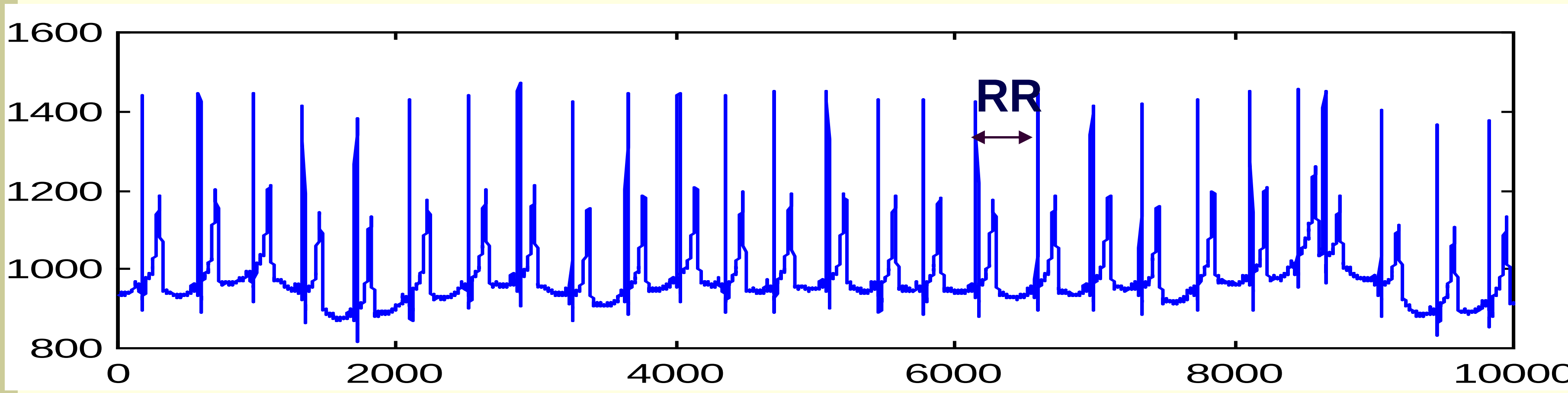

```
#Python
range=1
y=range*random.rand(1000)
bins=linspace(-1,1.5,50)
hist(y,bins, normed='True')
```



$$x = 0.5\,!$$

Proces Markowa

$$p[x(n)/\,x(n-1),x(n-2),\ldots,x(1)]=p[x(n)/\,x(n-1)]$$



Czy ciąg $\{RR_1, RR_2, \ldots, RR_n, \ldots\}$ jest procesem Markowa?

Rozkłady punktowe - rozkład Poissona

Rozkład Poissona ma zastosowanie w obserwacji niezależnych zjawisk o małym prawdopodobieństwie sukcesu (np. zjawiska rozpadu promieniotwórczego).

Prawdopodobieństwo zaobserwowania *j* zdarzeń w zadanym przedziale czasu Δt:

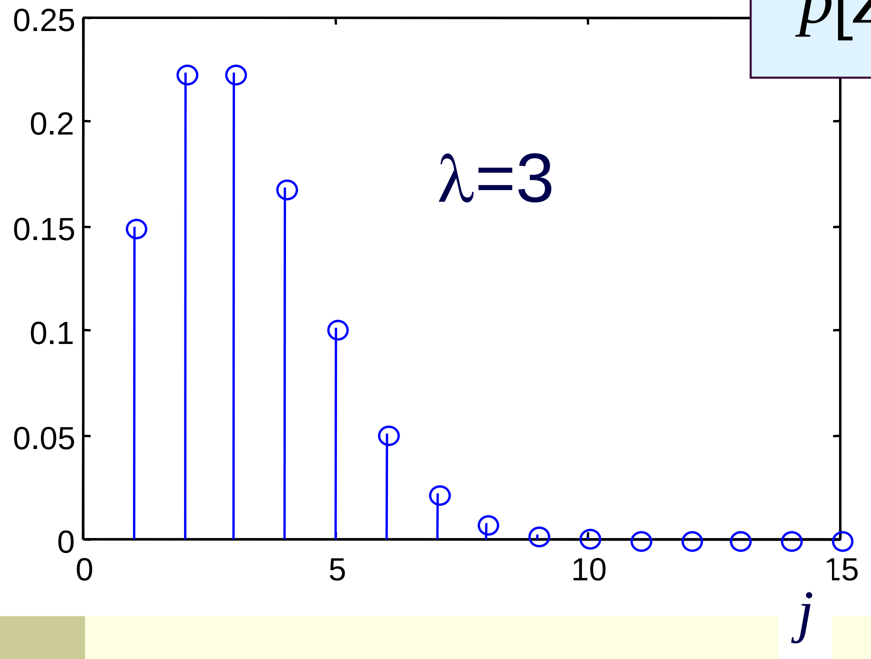
$$p\left[Z(\Delta t) = j\right] = \frac{(\lambda \Delta t)^j}{j!} e^{-\lambda \Delta t}, \ \lambda > 0 \quad \Rightarrow \quad \begin{aligned} E(Z) &= \lambda \\ D^2(Z) &= \lambda \end{aligned}$$

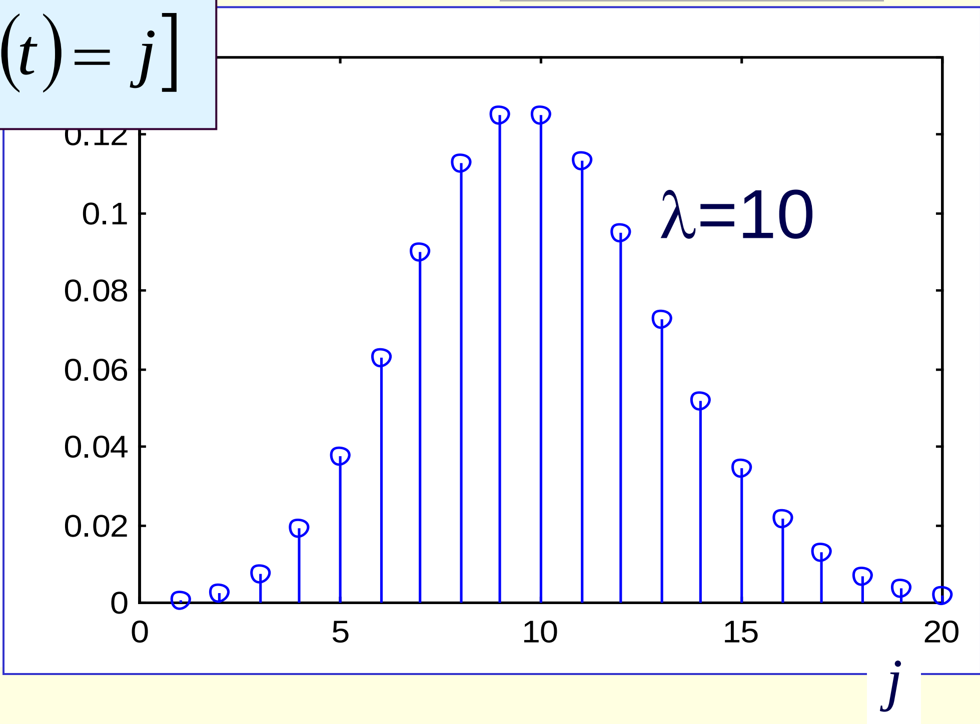
gdzie λ oznacza średnią liczbę zdarzeń występujących w czasie Δt.

Zastosowanie: w modelowaniu interwałów R-R sygnału EKG**,** sygnału EMG, serii impulsów wytwarzanych w komórkach neuronowych, kontroli jakości, ….

Rozkład Poissona

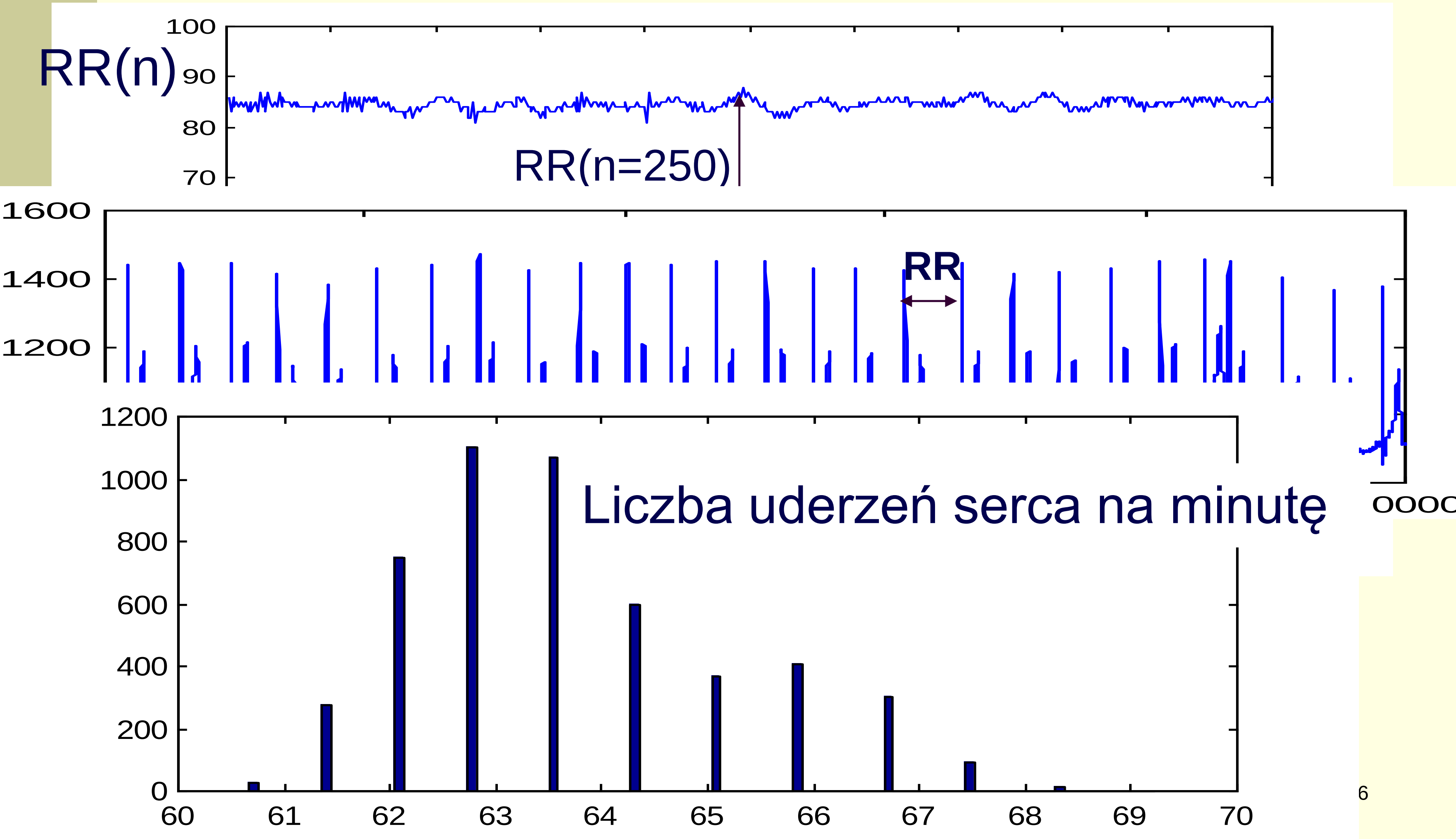
$$p[Z(t) = j]$$





Proces Poissona opisuje tzw. procesy bez pamięci, w których liczba bieżących zdarzeń w jednostce czasu nie zależy od ich liczby w przeszłości

Przykład procesu Poissona (Heart Rate Variability – HRV)



Zagadnienie regresji liniowej

Rozpatrzmy przypadek regresji liniowej:

$$\boxed{y = \alpha_0 + \alpha_1 x + \varepsilon}$$

gdzie:

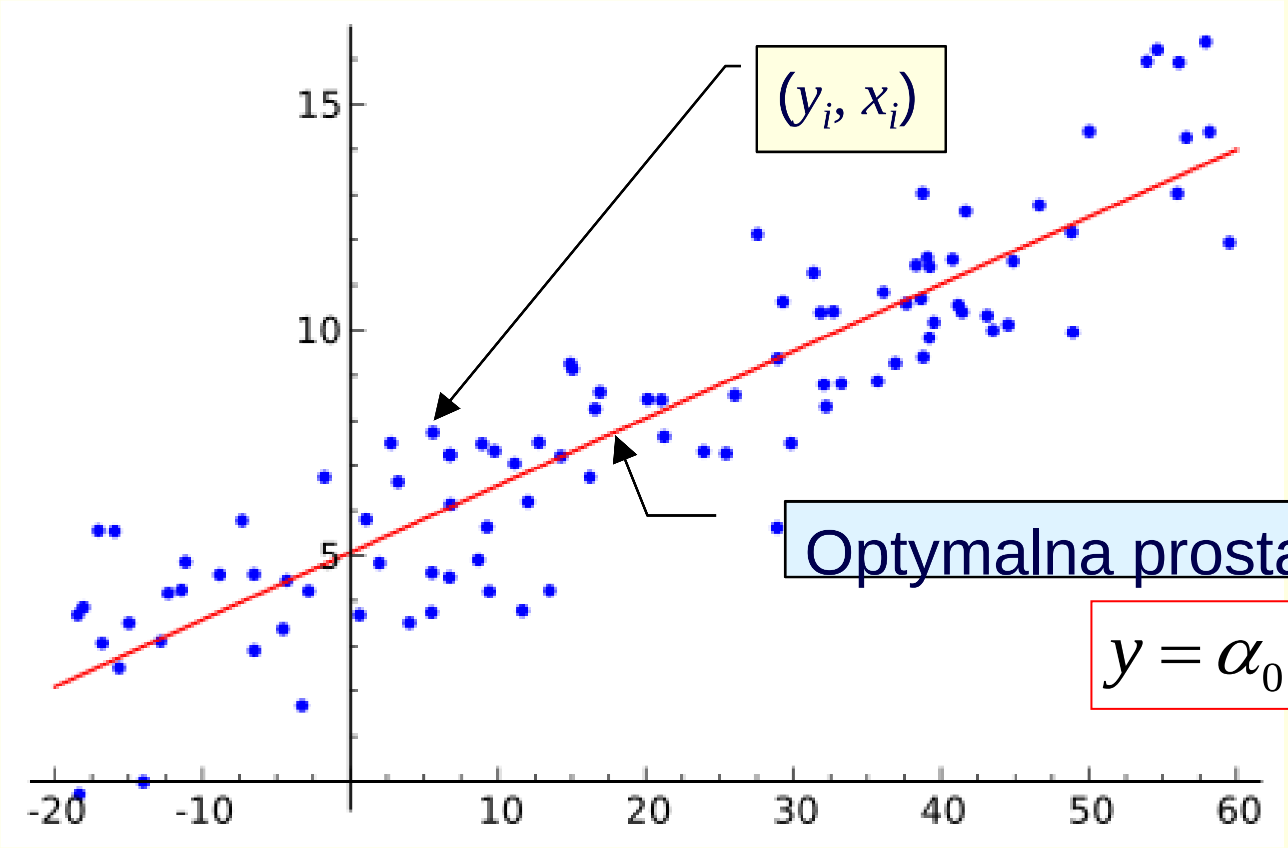
x - zmienna niezależna,

y - zmienna zależna,

ε - zakłócenia (np. błąd pomiaru),

α_0, α_1 - współczynniki regresji.

Regresja liniowa



$$y = \alpha_0 + \alpha_1 x + \varepsilon$$

Estymacja liniowej funkcji regresji

Zadanie regresji:

Załóżmy, że mamy *P* obserwacji $(x_1, y_1), (x_2, y_2), \ldots, (x_P, y_P)$, poszukujemy współczynników a_0, a_1 , dla których uzyskamy najdokładniejszy opis zależności pomiędzy zmiennymi *x* i *y*.

Rozwiązanie:

Estymaty współczynników można wyznaczyć na drodze poszukiwania minimum sumy kwadratów błędów:

$$\varepsilon_i = \underbrace{y_i}_{\text{obserwacja}} - \underbrace{(\alpha_0 + \alpha_1 x)}_{\text{model regresji liniowej}} = y_i - \hat{y}_i$$

Estymacja liniowej funkcji regresji

Suma kwadratów błędów *(ang. summed squared error - SSE)*

$$SSE = \sum_{i=1}^{P} \varepsilon_i^2 = \sum_{i=1}^{P} \left(y_i - (\hat{\alpha}_0 + \hat{\alpha}_1 x_i) \right)^2$$

zgodnie z ideą metody sumy najmniejszych kwadratów błędów zaproponowanej przez Gaussa w XIX wieku; poszukujemy minimum tego wyrażenia.

Estymacja liniowej funkcji regresji

Minimum znajdujemy wyznaczając pochodne cząstkowe względem obu współczynników i przyrównujemy je do zera:

$$\begin{cases} \dfrac{\partial(SEE)}{\partial\hat{\alpha}_0} = \dfrac{\partial}{\partial\hat{\alpha}_0}\left(\sum_{i=1}^{P}\left(y_i - \hat{\alpha}_0 - \hat{\alpha}_1 x_i\right)^2\right) = 0 \\ \dfrac{\partial(SEE)}{\partial\hat{\alpha}_1} = \dfrac{\partial}{\partial\hat{\alpha}_1}\left(\sum_{i=1}^{P}\left(y_i - \hat{\alpha}_0 - \hat{\alpha}_1 x_i\right)^2\right) = 0 \end{cases}$$

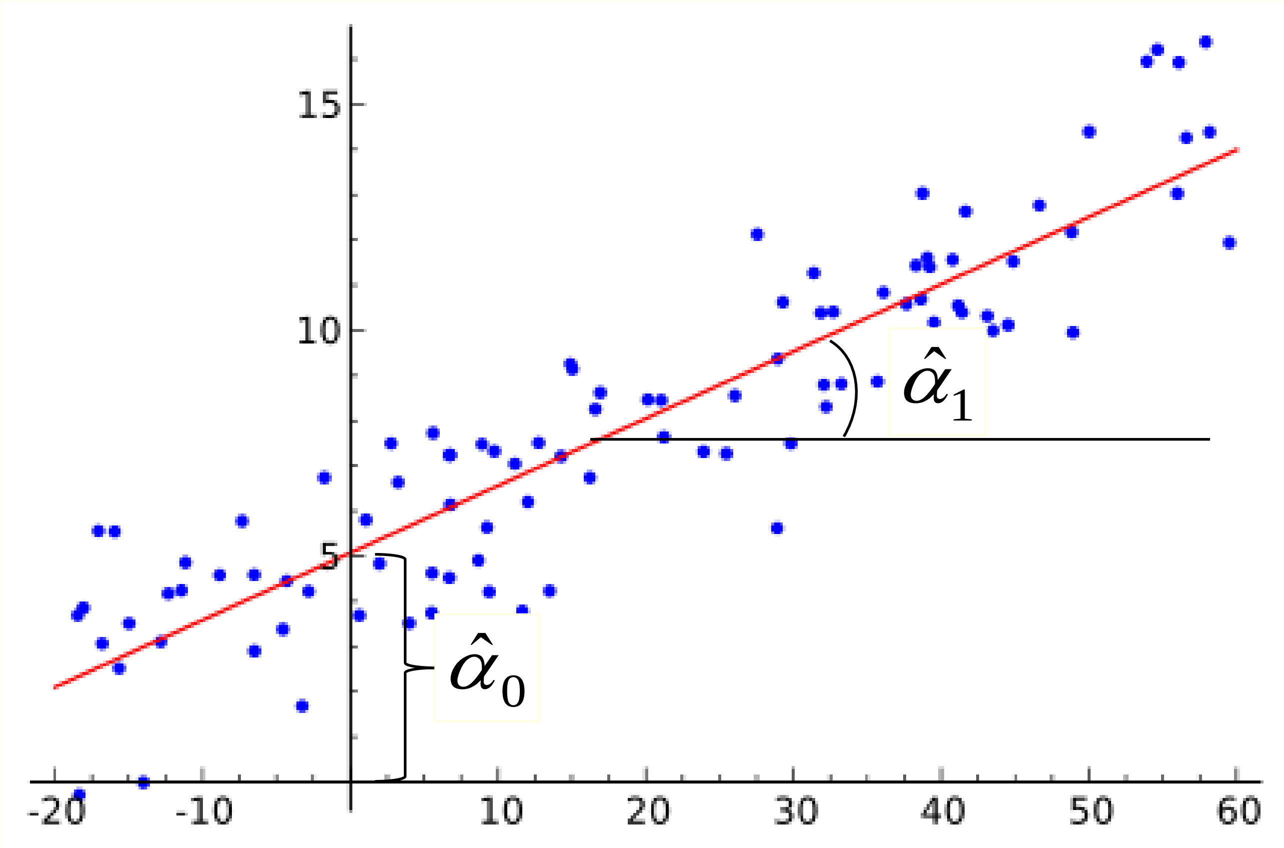
tzw. układ normalny równań

Estymacja liniowej funkcji regresji

Układ normalny równań posiada rozwiązanie:

$$
\begin{cases}
\boxed{\hat{\alpha}_0 = y_{śr} - \hat{\alpha}_1 x_{śr}} \\
\boxed{\hat{\alpha}_1 = \dfrac{\sum\limits_{i=1}^{P} \left(x_i - x_{śr}\right)\left(y_i - y_{śr}\right)}{\sum\limits_{i=1}^{P} \left(x_i - x_{śr}\right)^2}}
\end{cases}
$$

Estymacja liniowej funkcji regresji



Source: Wikipedia

Estymacja liniowej funkcji regresji

Regresja wyższych rzędów:

$$y_i = \alpha_0 + \alpha_1 x_{i1} + \alpha_1 x_{i1} + \ldots + \alpha_N x_{iN}, \; i = 1, \ldots, P$$

tj.

$$y_i = \alpha_0 + \sum_{j=1}^{N} \alpha_j x_{ij}, \; i = 1, \ldots, P$$

Podobnie jak dla regresji pierwszego rzędu, poszukujemy współczynników przez przyrównanie do zera pochodnych cząstkowych:

$$\frac{\partial (SEE)}{\partial \hat{\alpha}_j} = 0, \;\; j = 0, \ldots N$$

i otrzymujemy N+1 równań z N+1 niewiadomymi.

Inne modele regresji

Regresja wielomianowa:

$$y_i = \alpha_0 + \alpha_1 x_{i1}^1 + \alpha_1 x_{i2}^2 + \ldots + \alpha_N x_{iN}^N, \;\; i = 1, \ldots P$$

Przykłady regresji nieliniowej:

$$y_i = \alpha_0 e^{\alpha_1 x_i}$$

$$y_i = \alpha_0 \alpha_1^{x_i}$$

$$y_i = \frac{\alpha_0}{1 + \alpha_1 e^{x_i}}$$

Rekonstrukcja trajektorii układu dynamicznego

Takens (1981)udowodnił, że można odtwarzać własności trajektorii układu dynamicznego na podstawie próbek jednowymiarowego zapisu aktywności badanego układu:

$$\boldsymbol{y}(k)=\left[y(k), y(k-\Delta t),\ldots, y(k-(D-1)\Delta t)\right]^T$$

dla odpowiednio dużego wymiaru *D* wektora $\boldsymbol{y}(k)$ (ang. *time delay embedding*)

Rekonstrukcja trajektorii układu dynamicznego

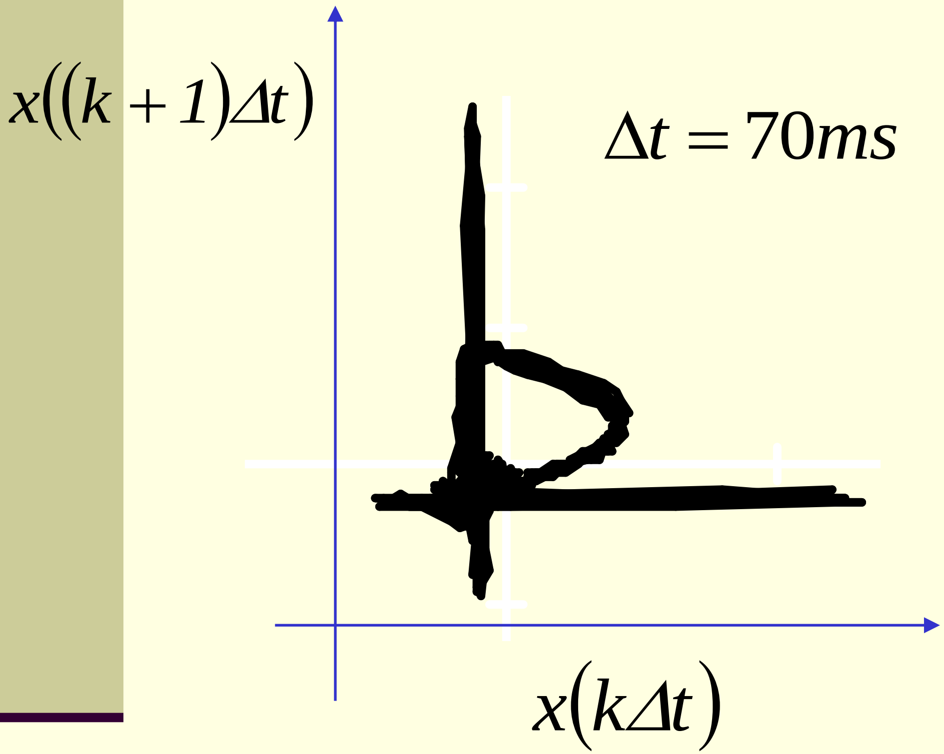
Praktyczne znaczenie tw. Takensa dla badania układów dynamicznych to dowód o istnieniu zależności:

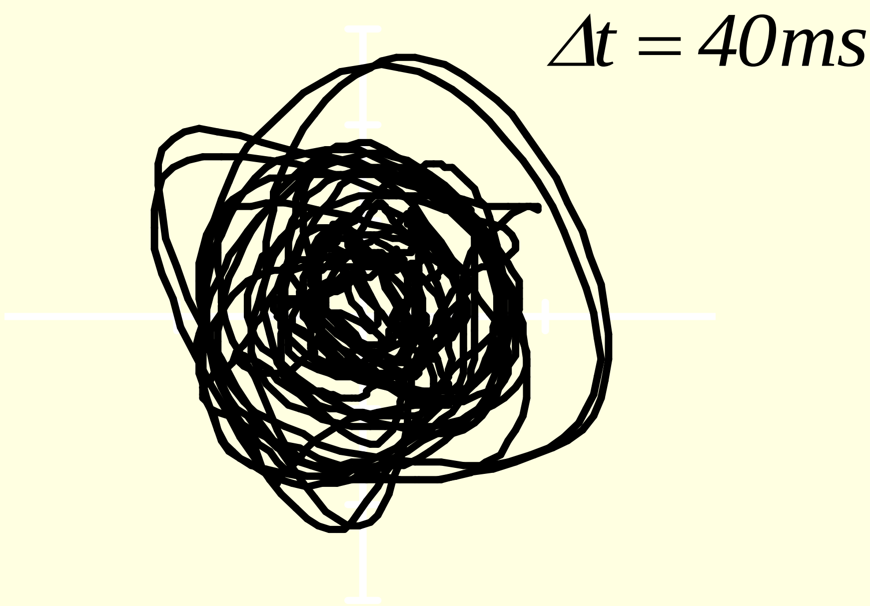
?

$$y((k+1)\Delta t) = F(\mathbf{y}(k))$$

na podstawie, której można wyznaczać prognozę próbki $y(k+1)\Delta t$ sygnału odwzorowującego zachowanie badanego układu.

Rekonstrukcja trajektorii dla rytmu serca





Normalny rytm serca

Fibrylacja komór

Δt – pierwsze zero funkcji autokorelacji ciągu $y(k)$

Chaos deterministyczny

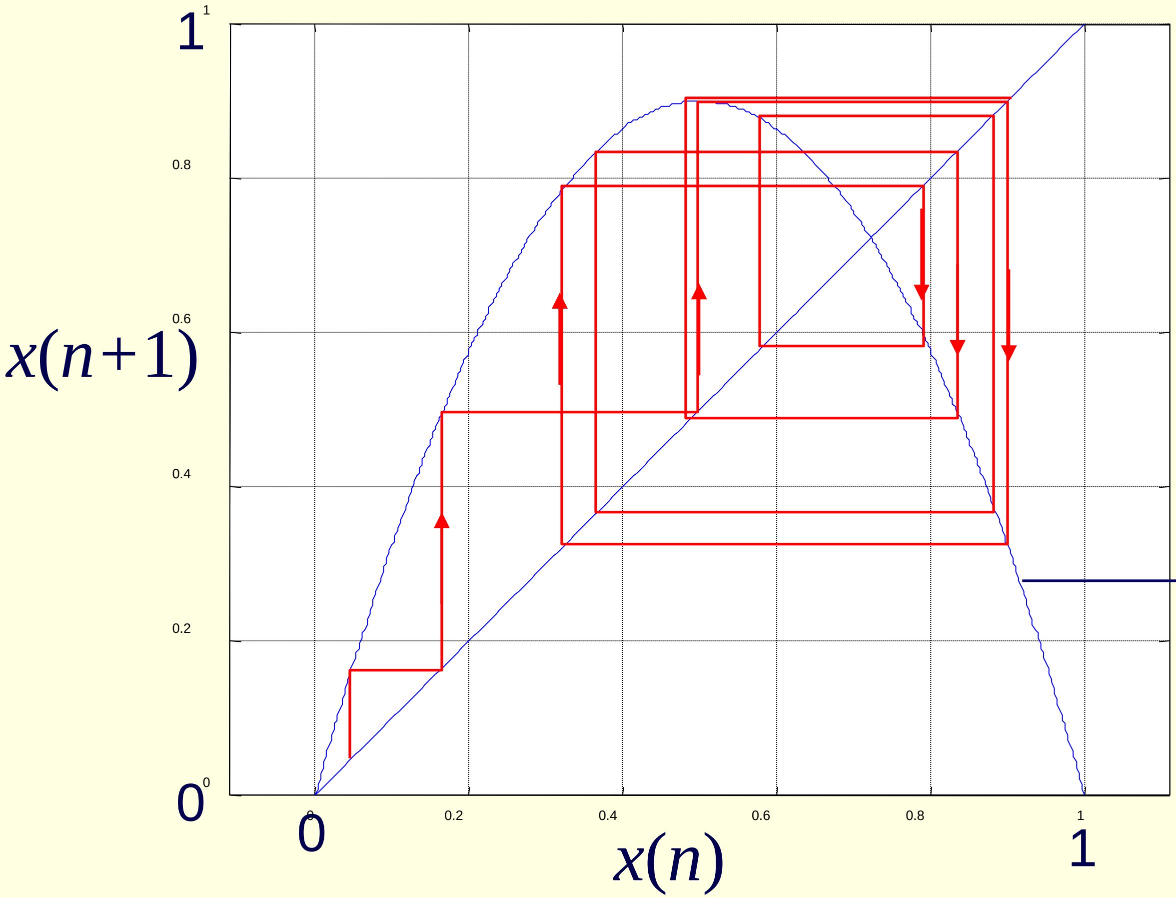
Rozważmy prosty układ dynamiczny:

$$\boxed{x(n+1) = ax(n)(1 - x(n))}$$

Symulujący zmianę liczności x populacji w ograniczonym środowisku. Liczność ta jest znormalizowana do zakresu $\langle 0,1 \rangle$, zatem parametr a przyjmuje wartości z zakresu $\langle 0,4 \rangle$.

Jest to tzw. **równanie logistyczne**.

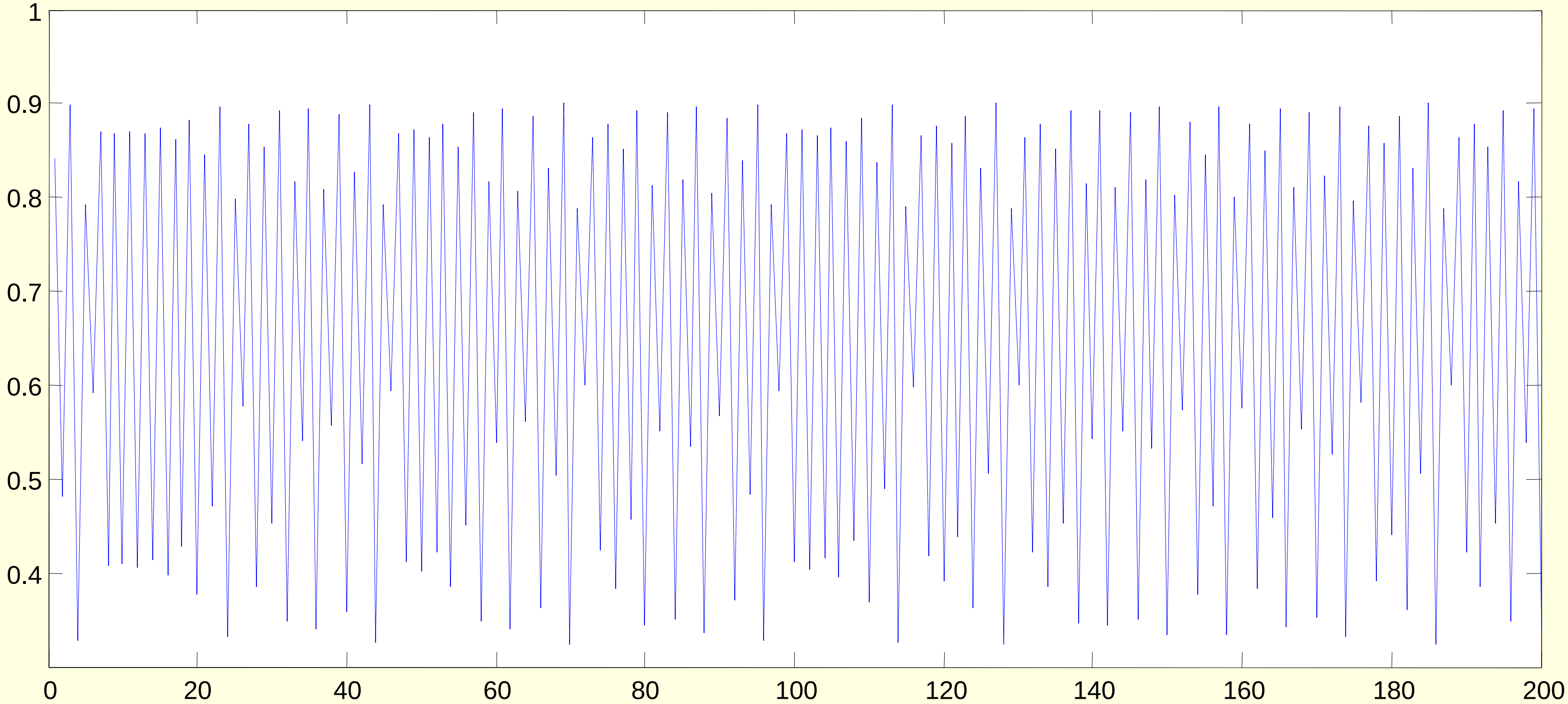
Chaos deterministyczny



tzw. odwzorowanie kwadratowe

$$x(n+1) = 3.6x(n)(1 - x(n))$$

Chaos deterministyczny



Chaos deterministyczny

Właściwości procesów charakteryzujących się deterministycznym ruchem chaotycznym:

- mogą być generowane przez proste układy dynamiczne*,
- dynamika ma „nieprzewidywalny” charakter losowy (tzw. dziwne atraktory – tj. orbity ograniczone, rozbieżne i nieokresowe) → demo Matlab
- dynamika silnie zależna od warunków początkowych („efekt motyla”)

*) R, May: „ … *lepiej by nam się żyło gdyby uwzględniano fakt, że proste układy dynamiczne niekoniecznie prowadzą do prostej ewolucji dynamicznej*.”

Analiza sygnałów losowych

1. Sygnały losowe a sygnały deterministyczne
2. Momenty statystyczne
3. Funkcja autokorelacji
4. Funkcja korelacji wzajemnej
5. Współczynnik korelacji
6. Gęstość widmowa mocy
7. Rozkłady statystyczne
8. Dynamika systemów