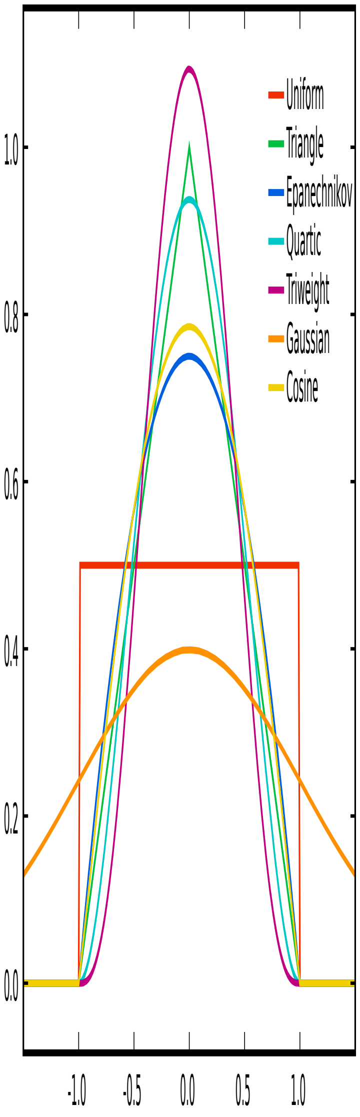
Politechnika Wrocławska



Identyfikacja

Wojciech Myszka
Maj 2015

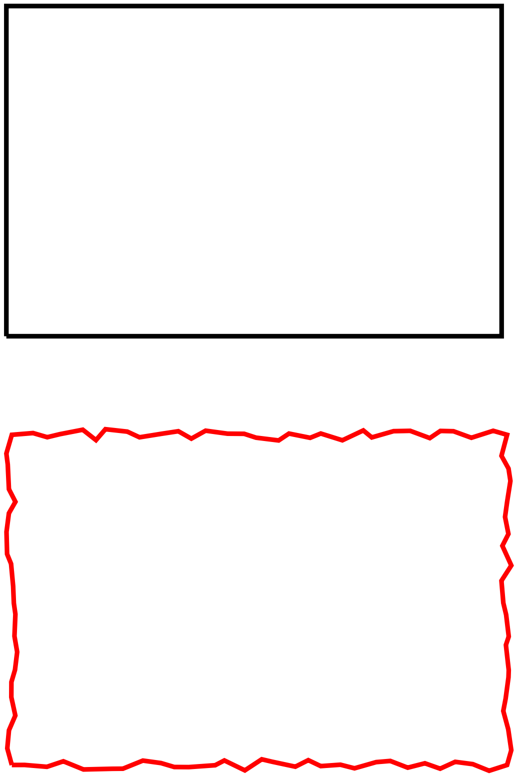
Wstęp

Rozglądamy się

Ten prostokąt („black box”) nazywać będziemy Obiektem (identyfikacji).

Co można z tym zrobić?

Narysować?



Co można z tym zrobić?

Narysować?

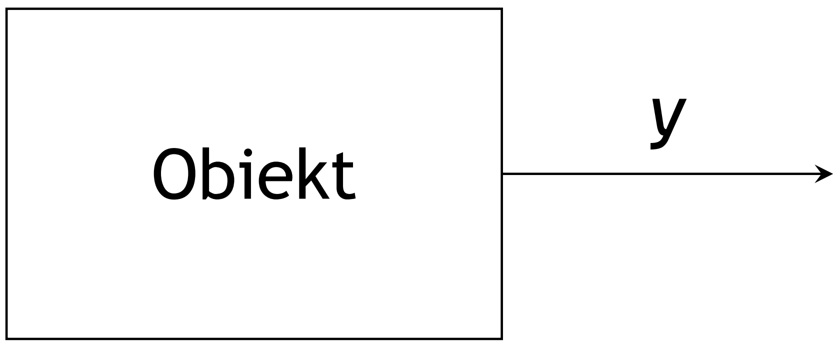
Ten drugi prostokąt to Model na różnych stadiach modelowania.
Naszym dążeniem jest aby Model był jak najbardziej podobny do Obiektu.

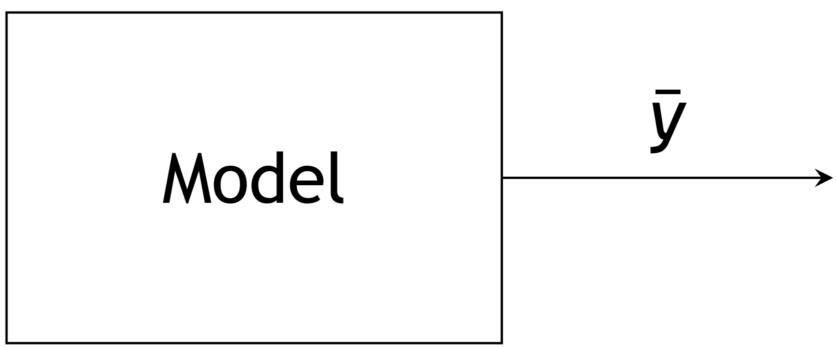
Co jeszcze można?

- Gdy zależy nam na czymś innym niż **wygląd** musimy zmienić zestaw interesujących nas cech charakterystycznych.
- Wyszczególniamy pewne cech, które będziemy obserwować. Bardzo często nazywa się je Wyjściem Obiektu.
- Obiekt nasz wygląda tak:

- y to wyjście Obiektu.

Porównanie Obiektu i Modelu

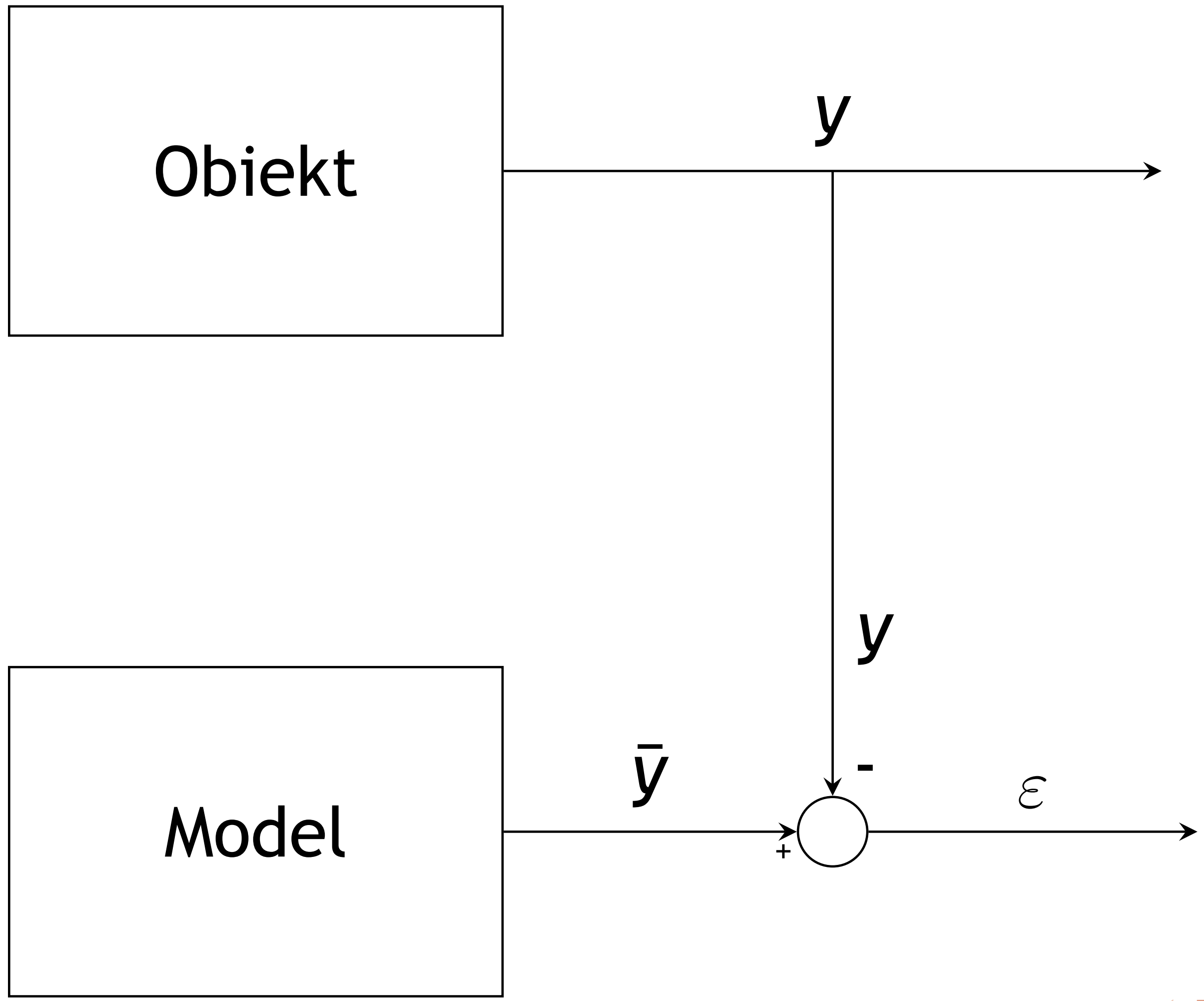
- Aby uzyskać zgodność Obiektu z Modelem, model również musi być wyposażony w wyjście.
- 

- 

- Wyjście modelu to $\bar{y}$.

Porównanie obiektu i modelu

Dokładniej sytuacja przedstawia się tak:



Model

Kilka najprostszych sytuacji:

1. O zmiennej y nie potrafimy nic powiedzieć...
2. Trudno cokolwiek wnioskować o modelu.
3. Możemy założyć, że y jest realizacją zmiennej losowej...
4. Zatem możemy badać pewne statystyczne charakterystyki:
   - średnią,
   - wariancję,
   - momenty,
   - rozkład parwdopodobieństwa
   - ...

i w ten sposób opisać Obiekt

Charakterystyki statystyczne

1. Losowy charakter zmiennej y to bardzo silne założenie. Oprócz zmiennych losowych, możemy również rozpatrywać:
   - zmienne deterministyczne (zawsze zachowują się tak samo),
   - zmienne niedeterministyczne i nielosowe (zmieniają się w sposób nie dający się opisać za pomocą gęstości).
2. Jeszcze gorzej jest gdy charakterystyki zmieniają się w czasie. . .
3. . . . albo obiekt nie spełnia założenia ergodyczności (uśrednianie po realizacjach może być zastąpione uśrednianiem po czasie).

Parametry statystyczne

Załóżmy, że $y_i, i = 1, 2, \ldots, N$, to kolejne realizacje zmiennej losowej Y.
Zazwyczaj korzystamy z następujących estymatorów podstawowych parametrów statystycznych:

- Średnia m_y:

$$m_y = \frac{1}{N} \sum_{i=1}^{N} y_i$$

- Wariancja

$$s_y^2 = \frac{1}{N} \sum_{i=1}^{N} (y_i - m_y)^2 \text{ a czasami } s_y^2 = \frac{1}{N-1} \sum_{i=1}^{N} (y_i - m_y)^2$$

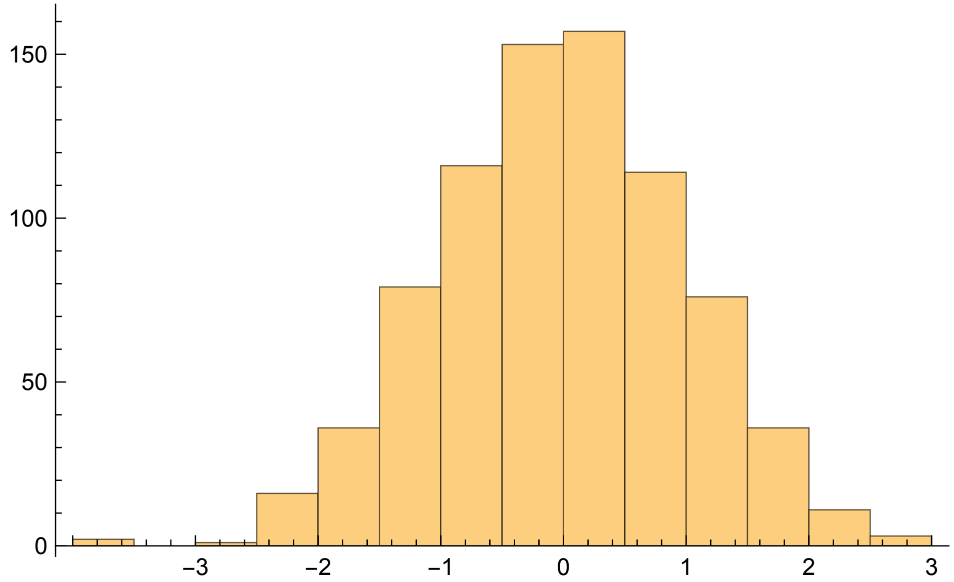
Gęstość

Następnym krokiem może być powiedzenie czegoś o funkcji gęstości.

Gęstość

Procedura aproksymacji funkcji gęstość przez całe lata była taka:

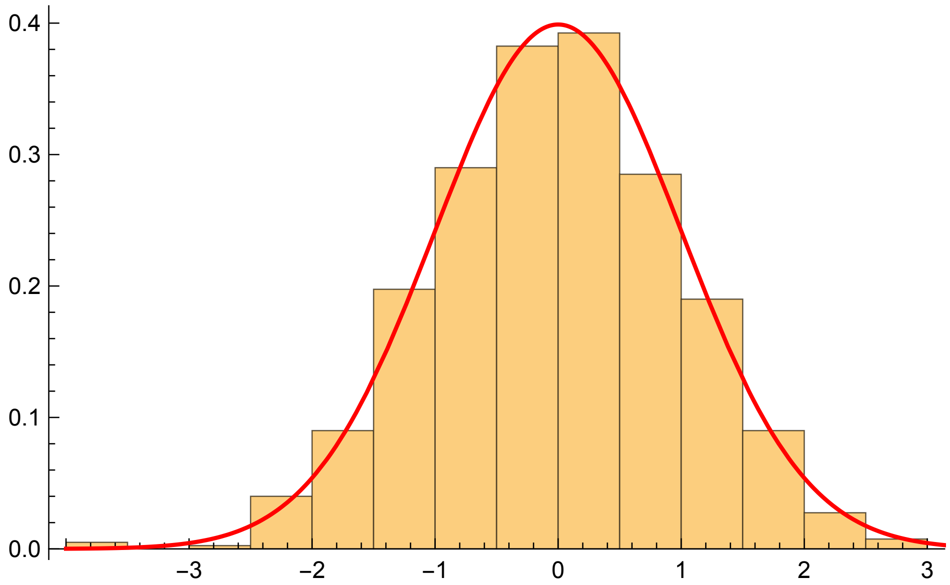
- Na podstawie obserwacji (y_i) przygotowywano histogram (oraz szacowano podstawowe parametry: wartość średnia, dyspersja, . . .)



Gęstość

Procedura aproksymacji funkcji gęstość przez całe lata była taka:

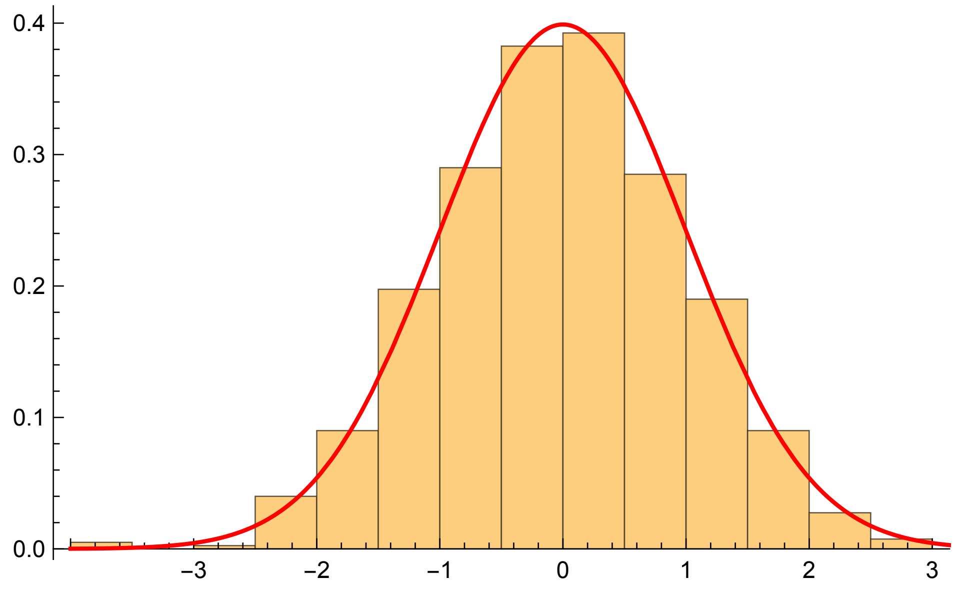
- Następnie dopasowywano funkcję gęstości



Gęstość

Procedura aproksymacji funkcji gęstość przez całe lata była taka:

- Następnie dopasowywano funkcję gęstości



- i przeprowadzano testy statystyczne

Gęstość — estymacja nieparametryczna

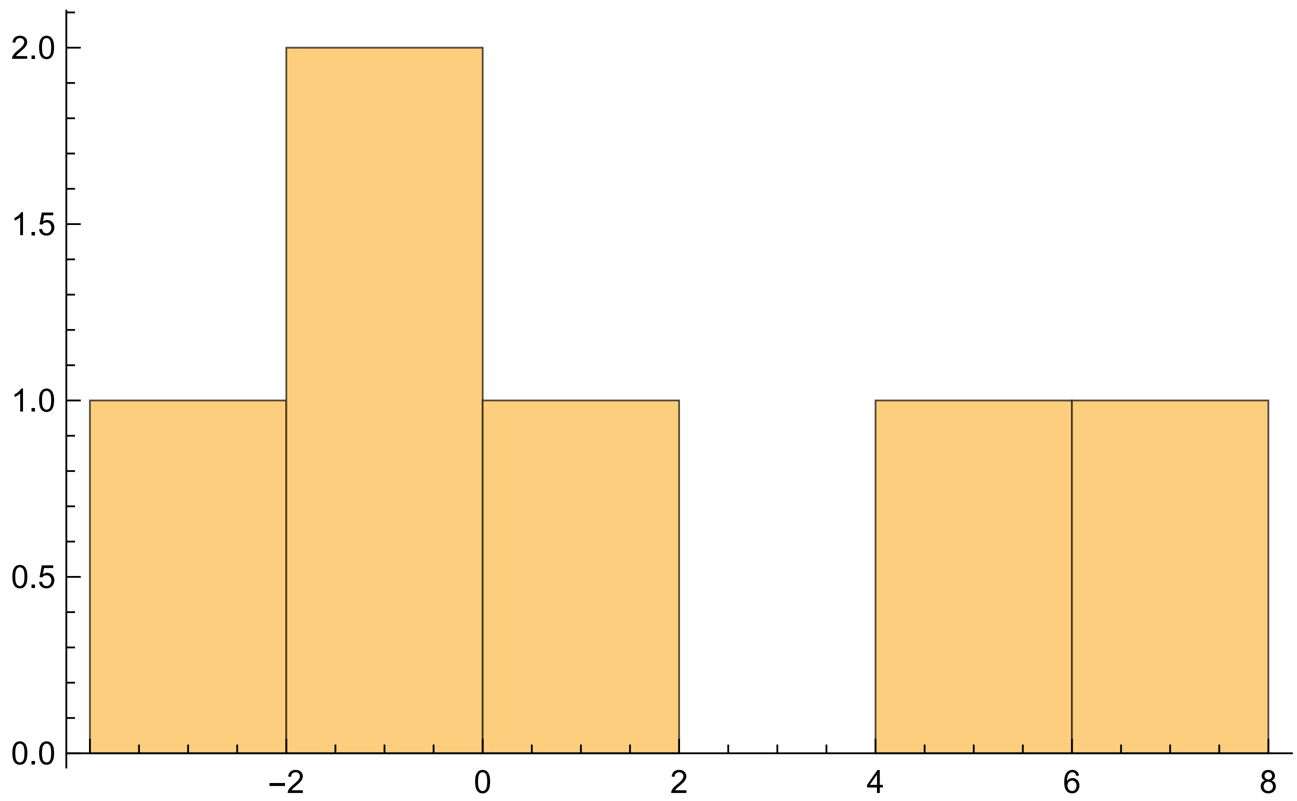
1. Obserwując sposób tworzenia histogramu, ktoś wpadł na prosty pomysł:
2. Gdy $y_1, y_2, \ldots, y_n$ to próbka niezależnych realizacji zmiennej losowej Y o nieznanej funkcji gęstości f,
3. Estymatorem funkcji gęstości będzie:

$$\hat{f}_h(x) = \frac{1}{n}\sum_{i=1}^{n} K_h(x - y_i) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - y_i}{h}\right)$$

4. K to funkcją jądra (o właściwościach funkcji gęstości) a h to parametr wygładzania (ang. *bandwidth parameter*). Będzie o nim nieco później.

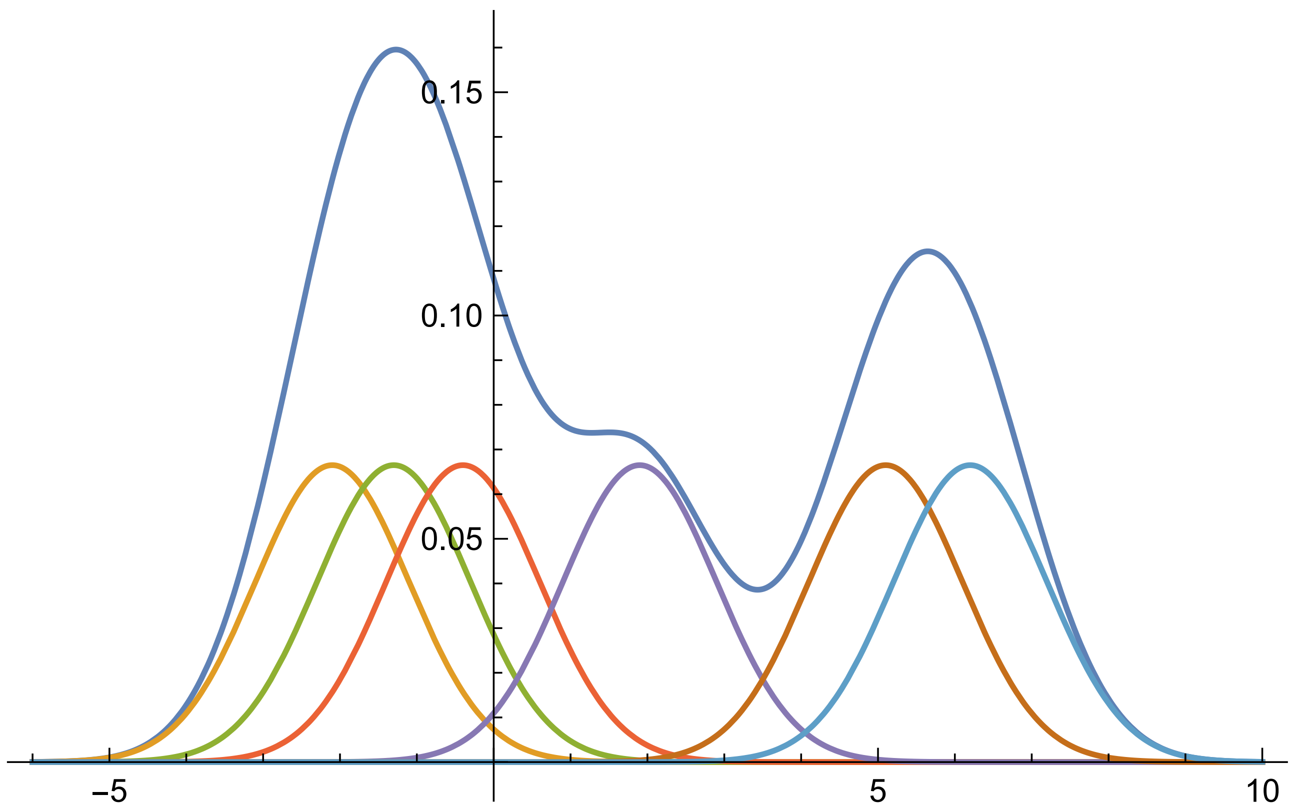
Jak to działa?

Histogram



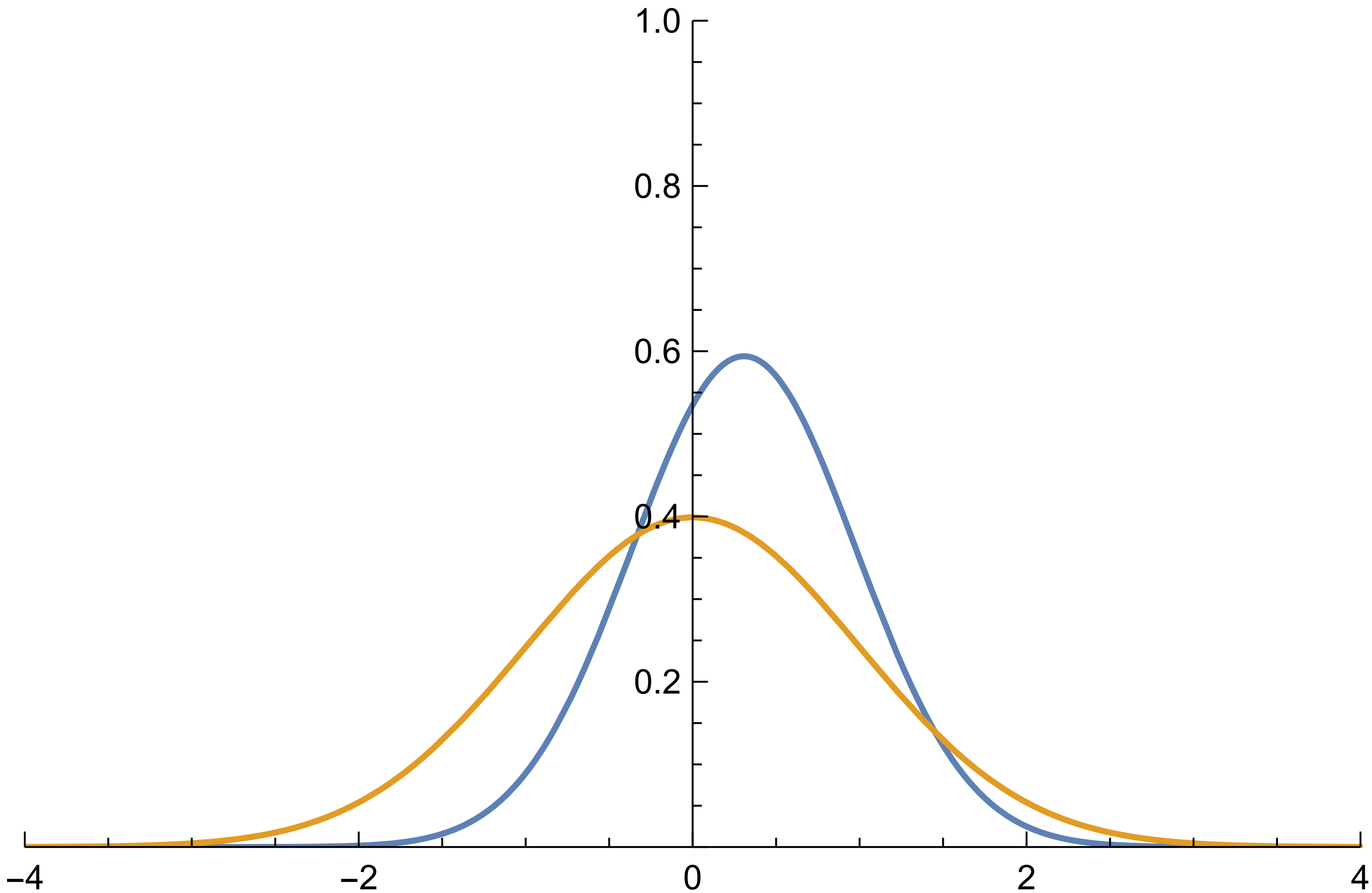
Jak to działa?

Aproksymacja nieparametryczna



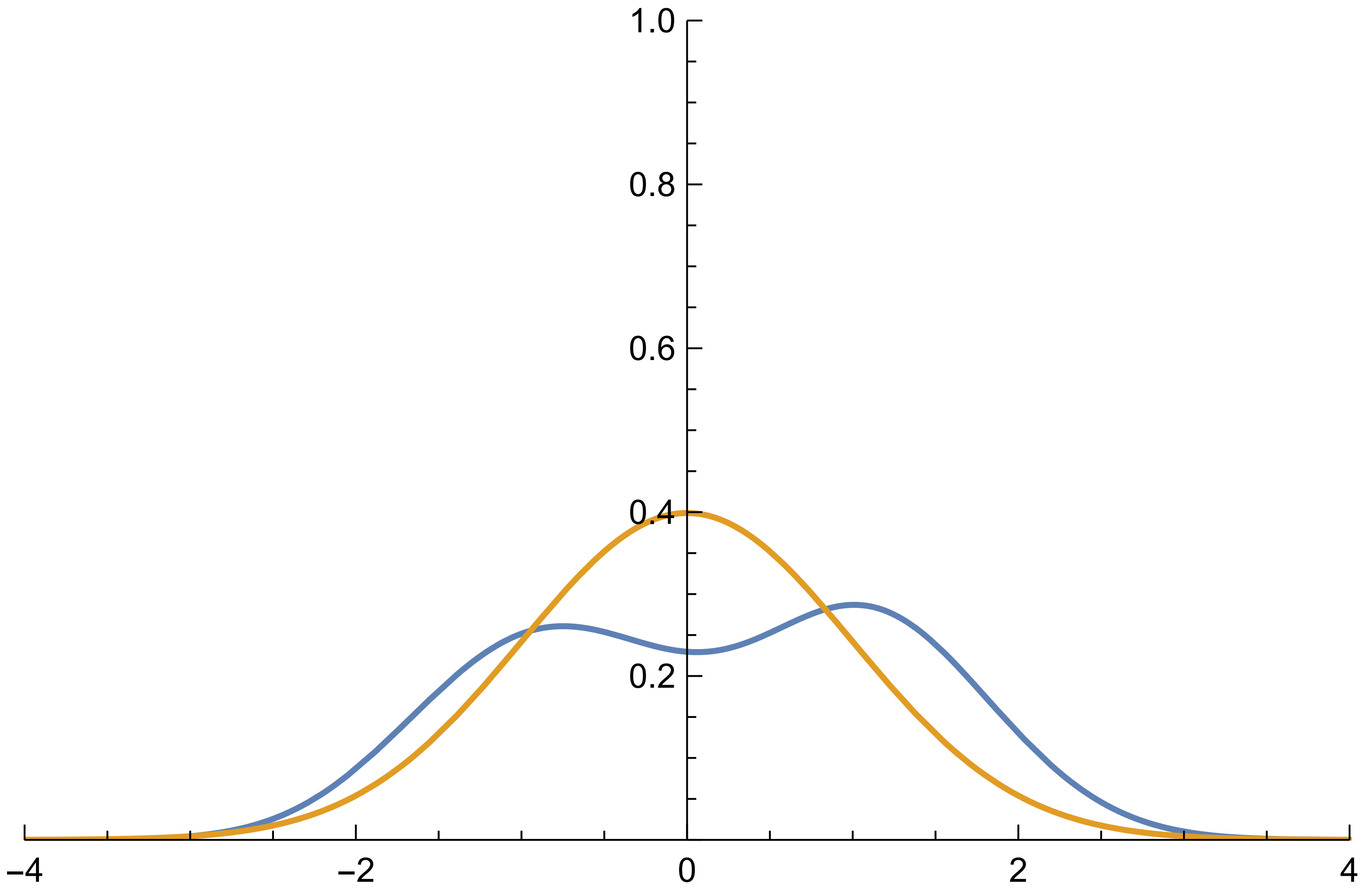
Przykłady użycia różnych funkcji jądra

Gaussa (n=1)



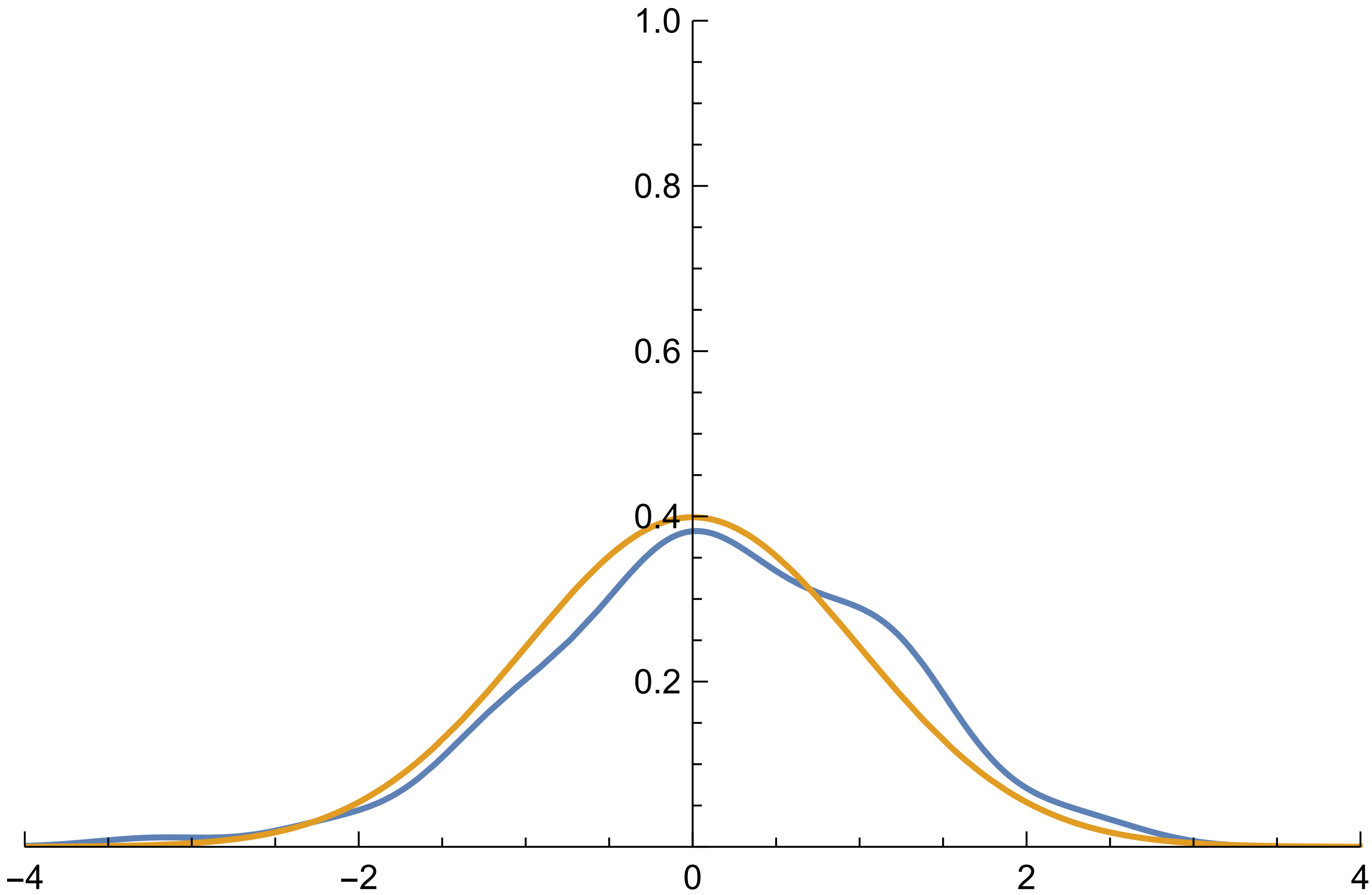
Przykłady użycia różnych funkcji jądra

Gaussa (n=10)

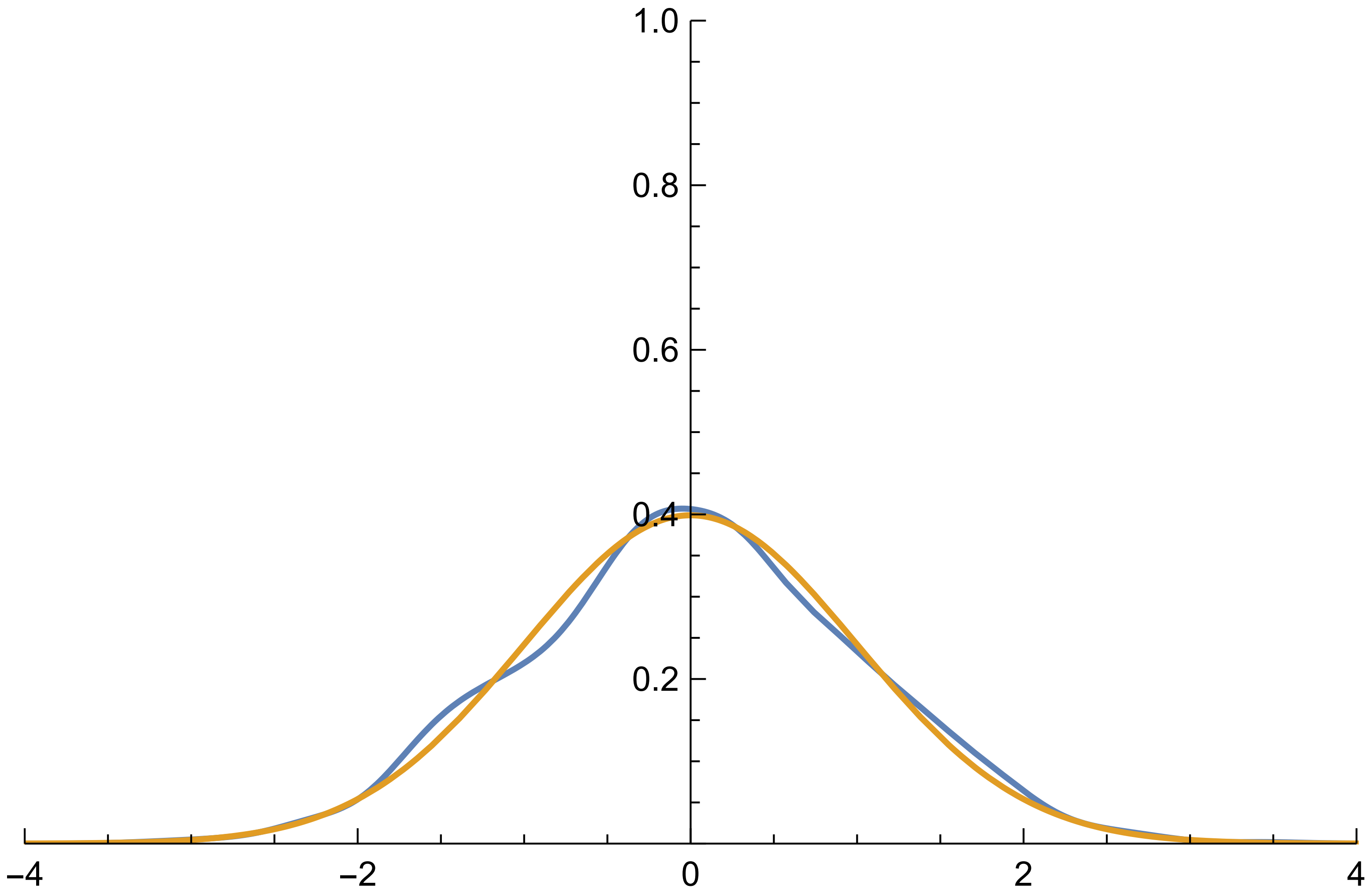


Przykłady użycia różnych funkcji jądra

Gaussa (n=100)

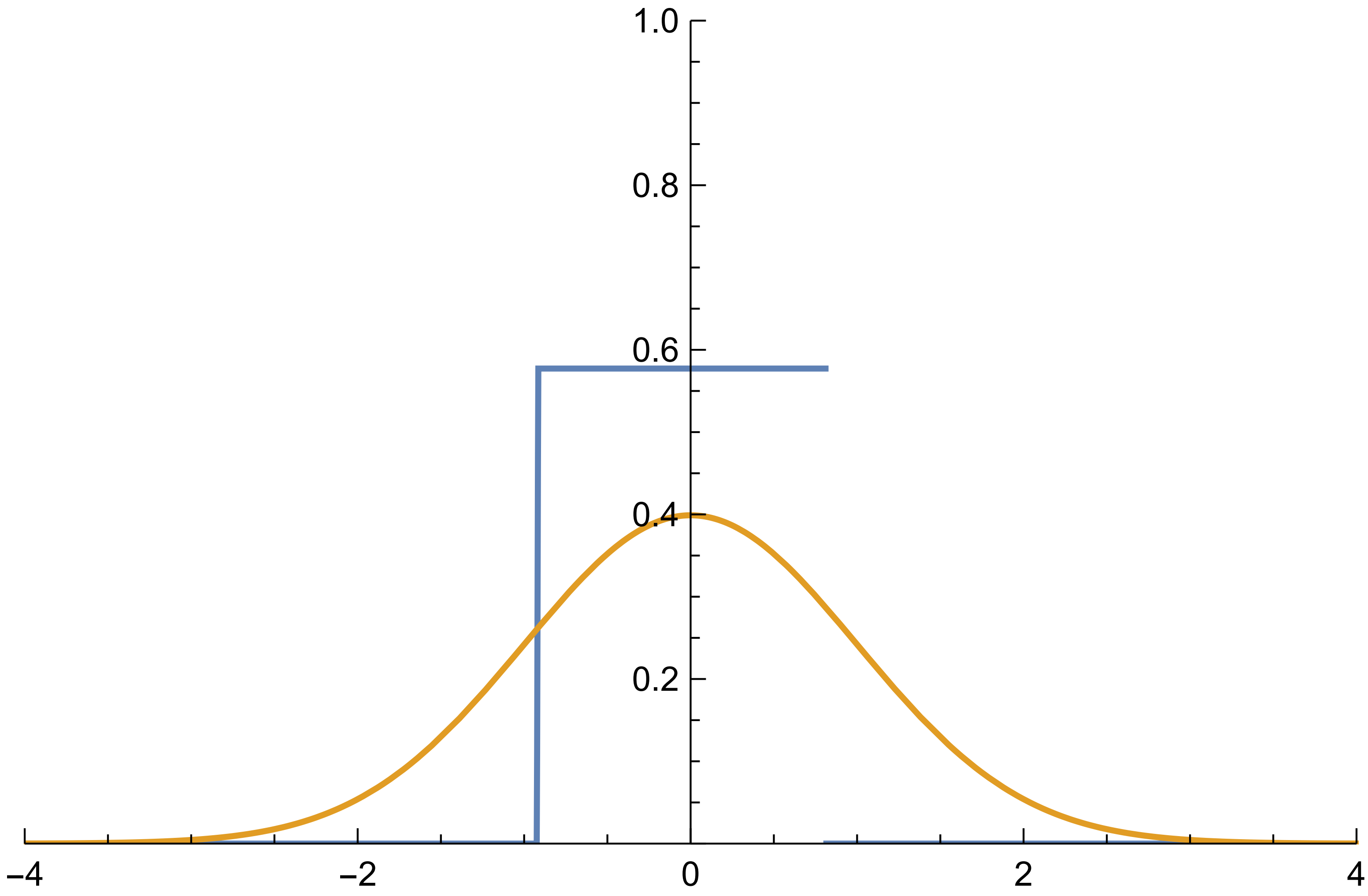


Przykłady użycia różnych funkcji jądra
Gaussa (n=1000)



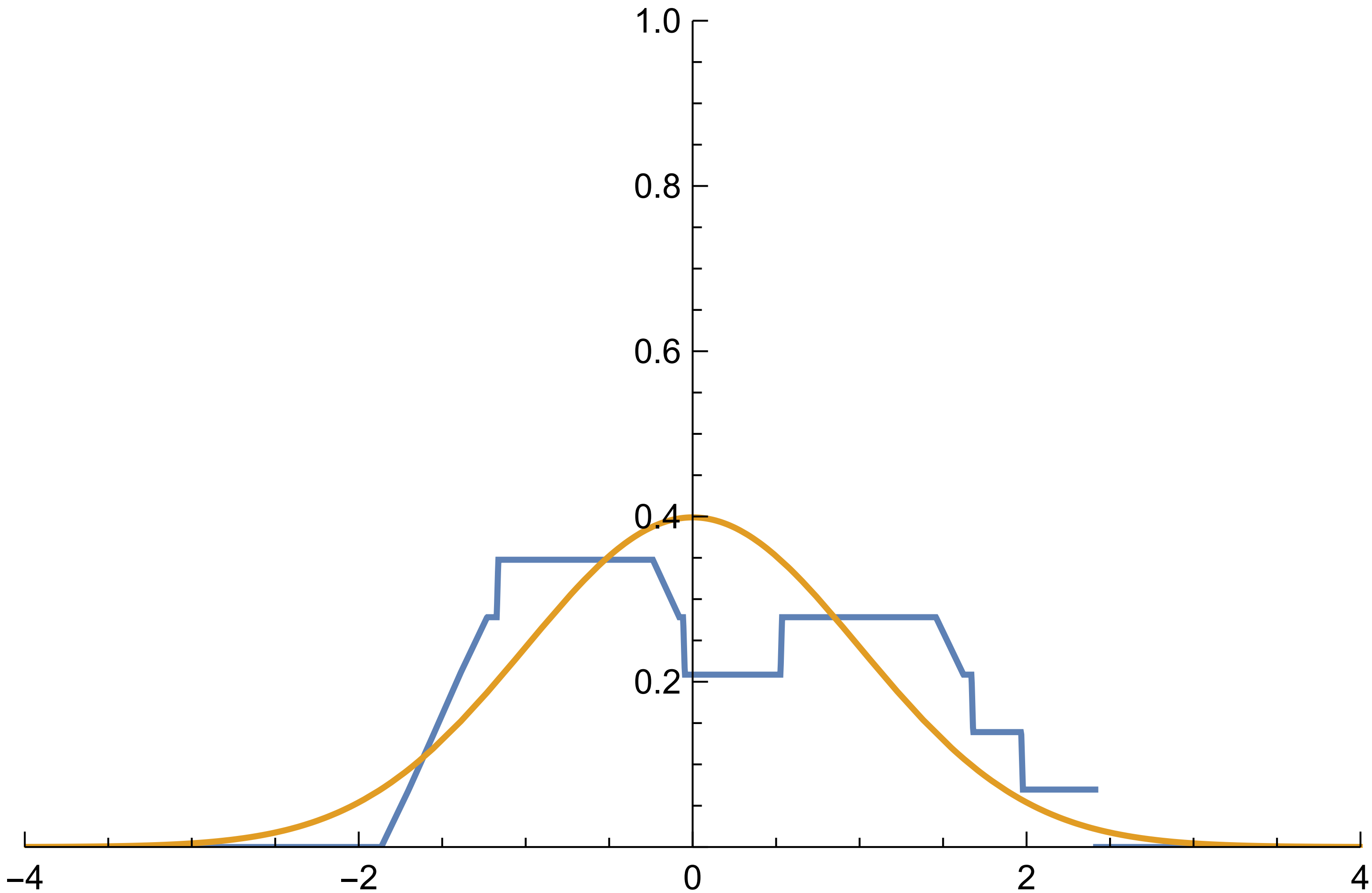
Przykłady użycia różnych funkcji jądra

Prostokątne (n=1)



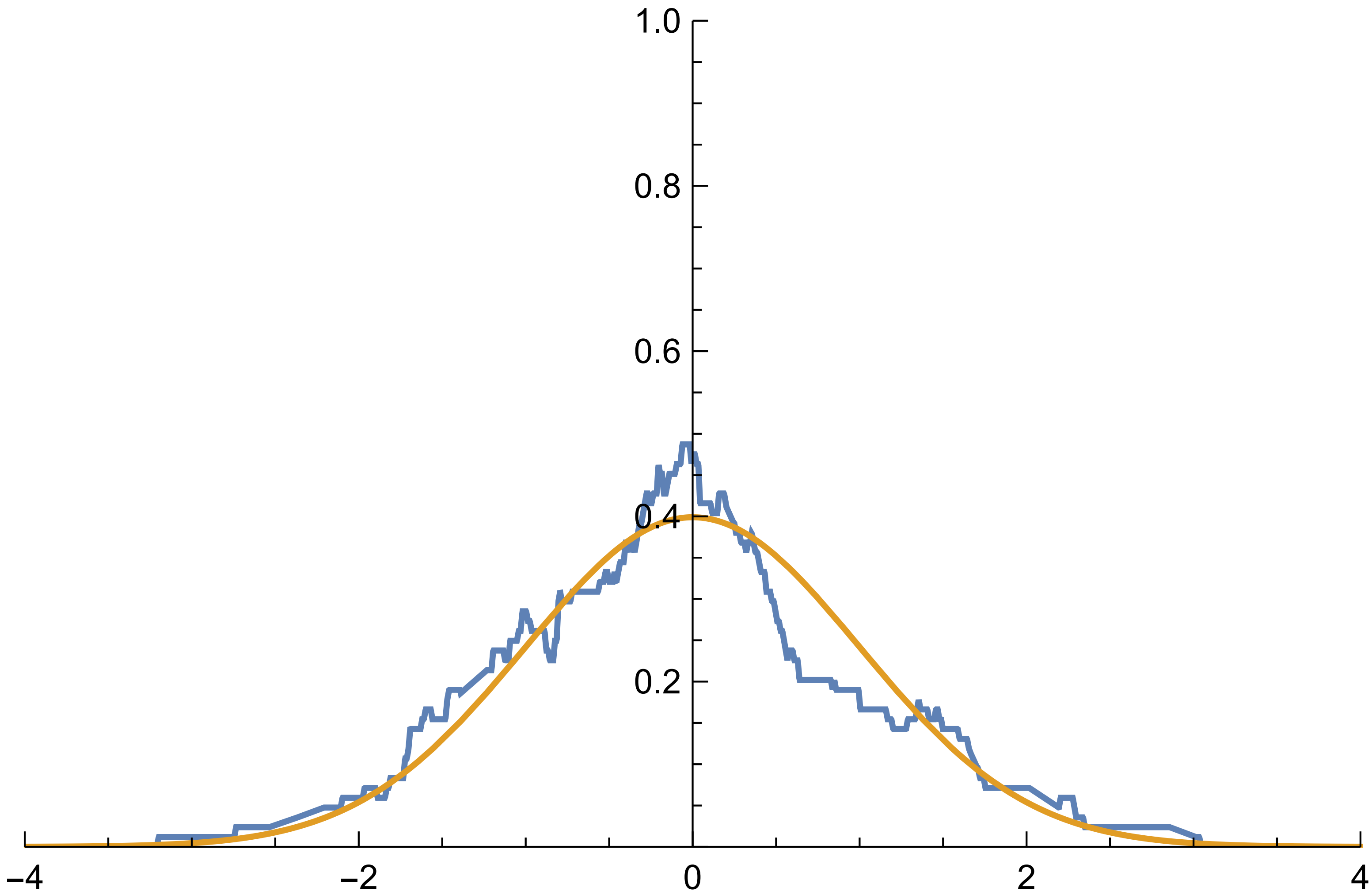
Przykłady użycia różnych funkcji jądra

Prostokątne (n=10)



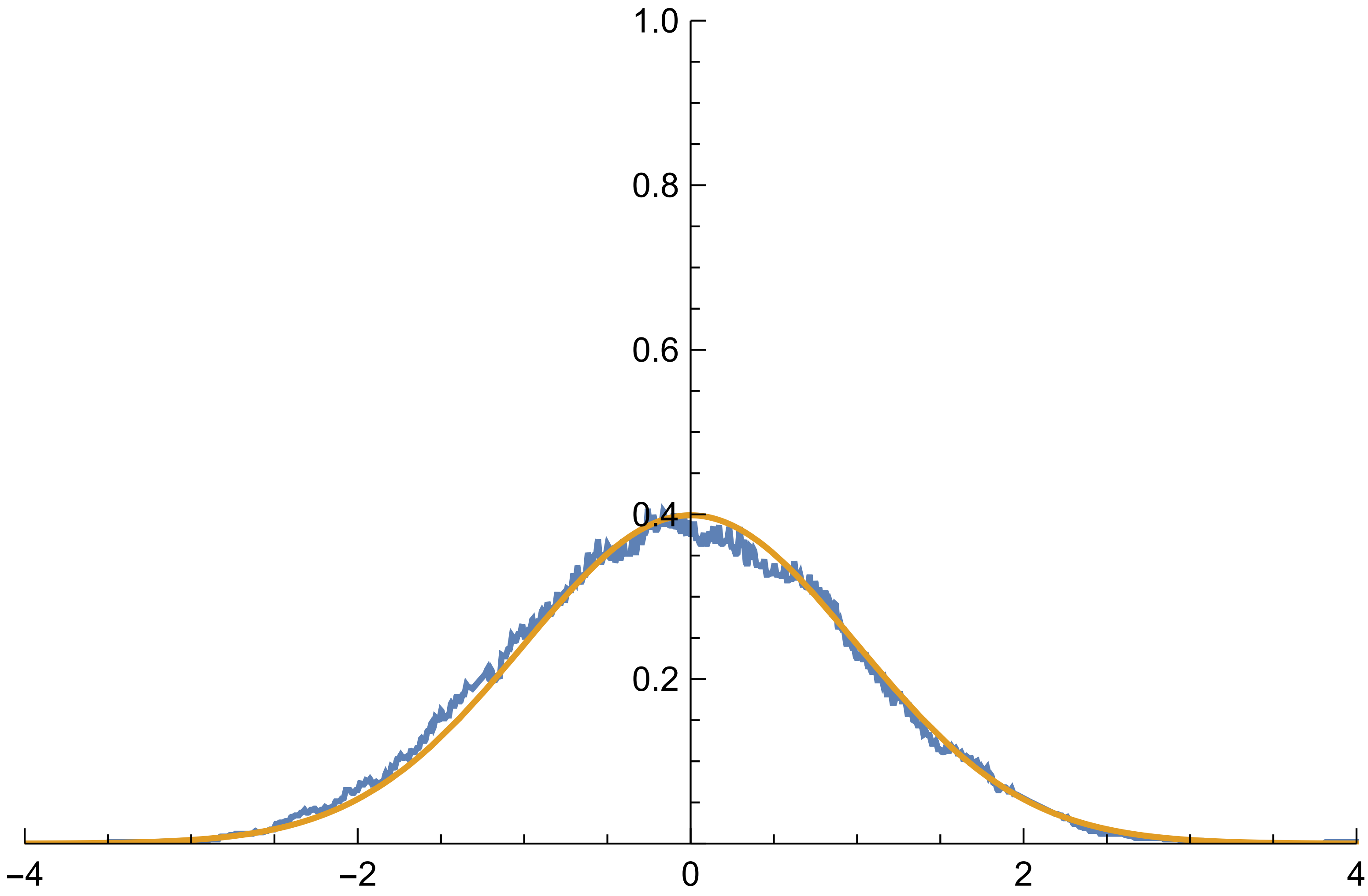
Przykłady użycia różnych funkcji jądra

Prostokątne (n=100)

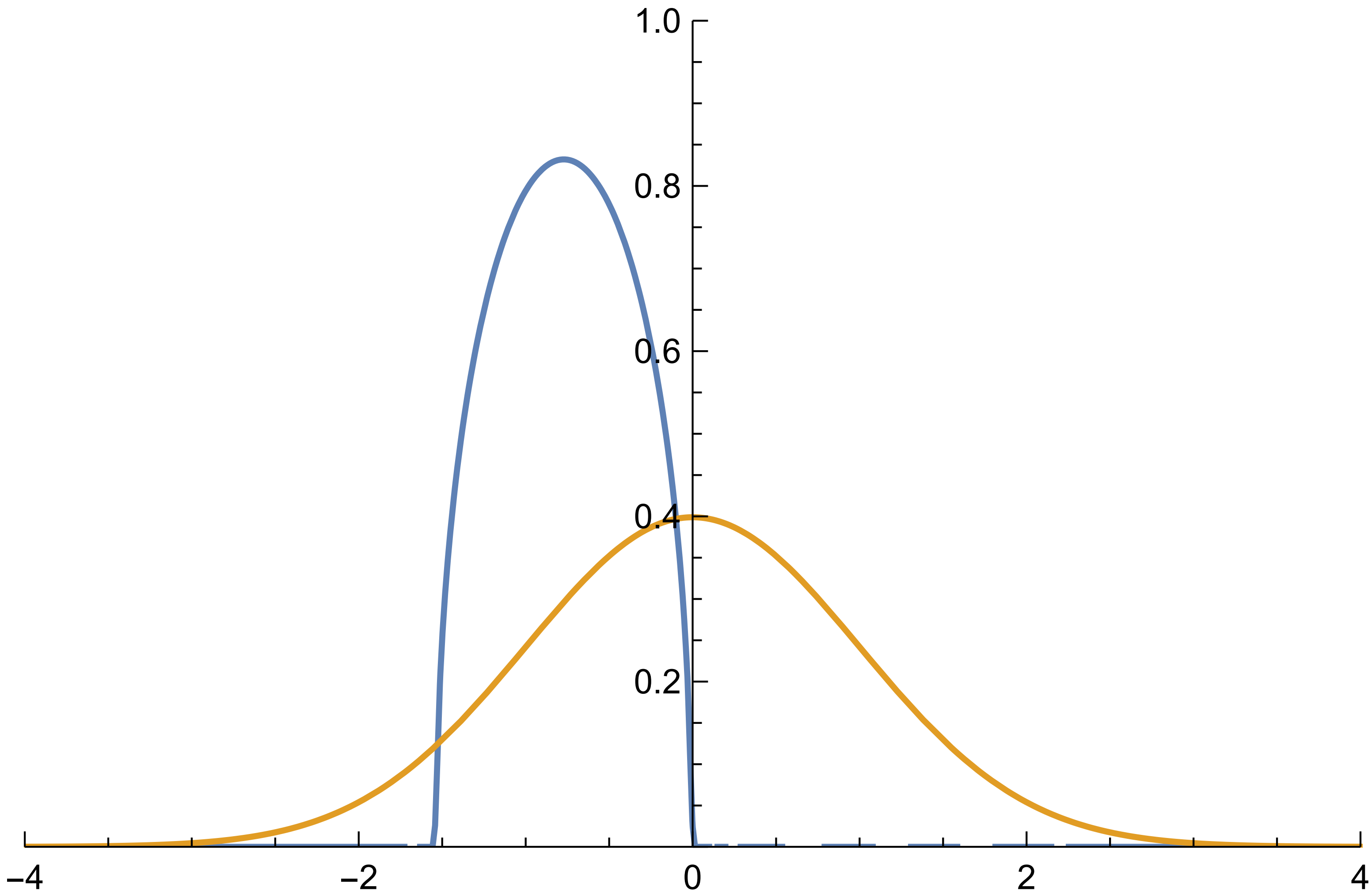


Przykłady użycia różnych funkcji jądra

Prostokątne (n=1000)

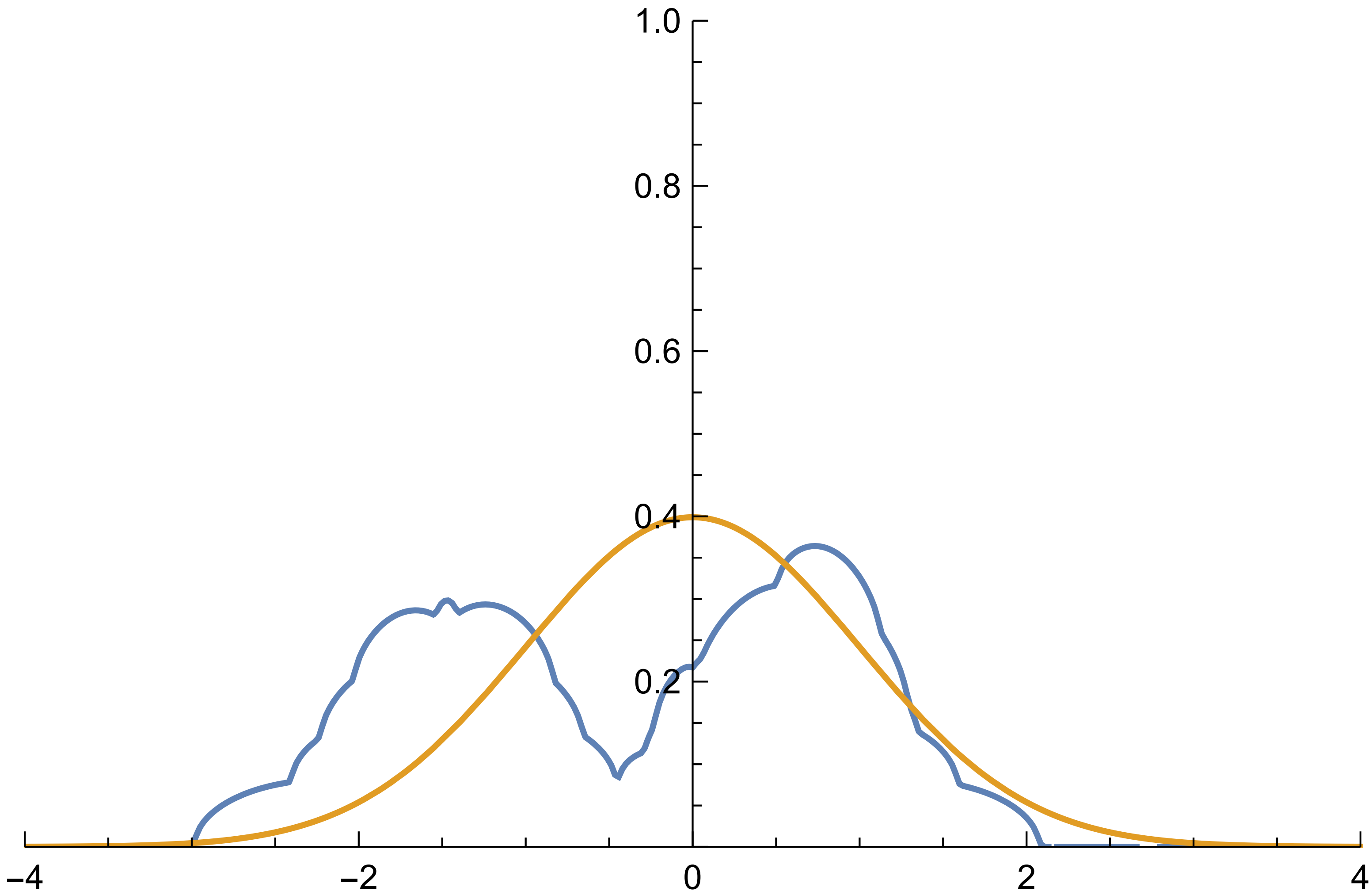


Przykłady użycia różnych funkcji jądra
Półokrągłe (n=1)



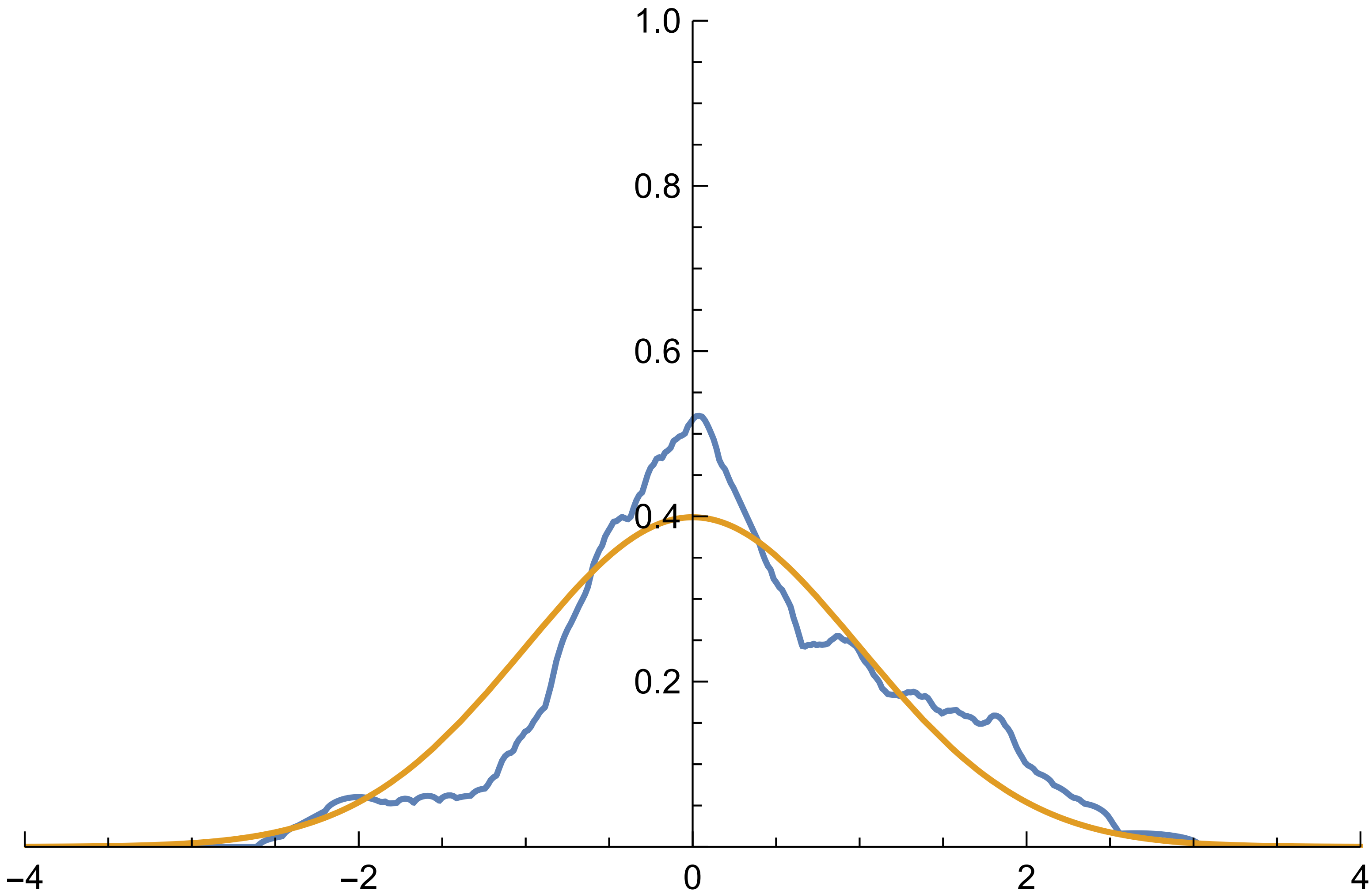
Przykłady użycia różnych funkcji jądra

Półokrągłe (n=10)



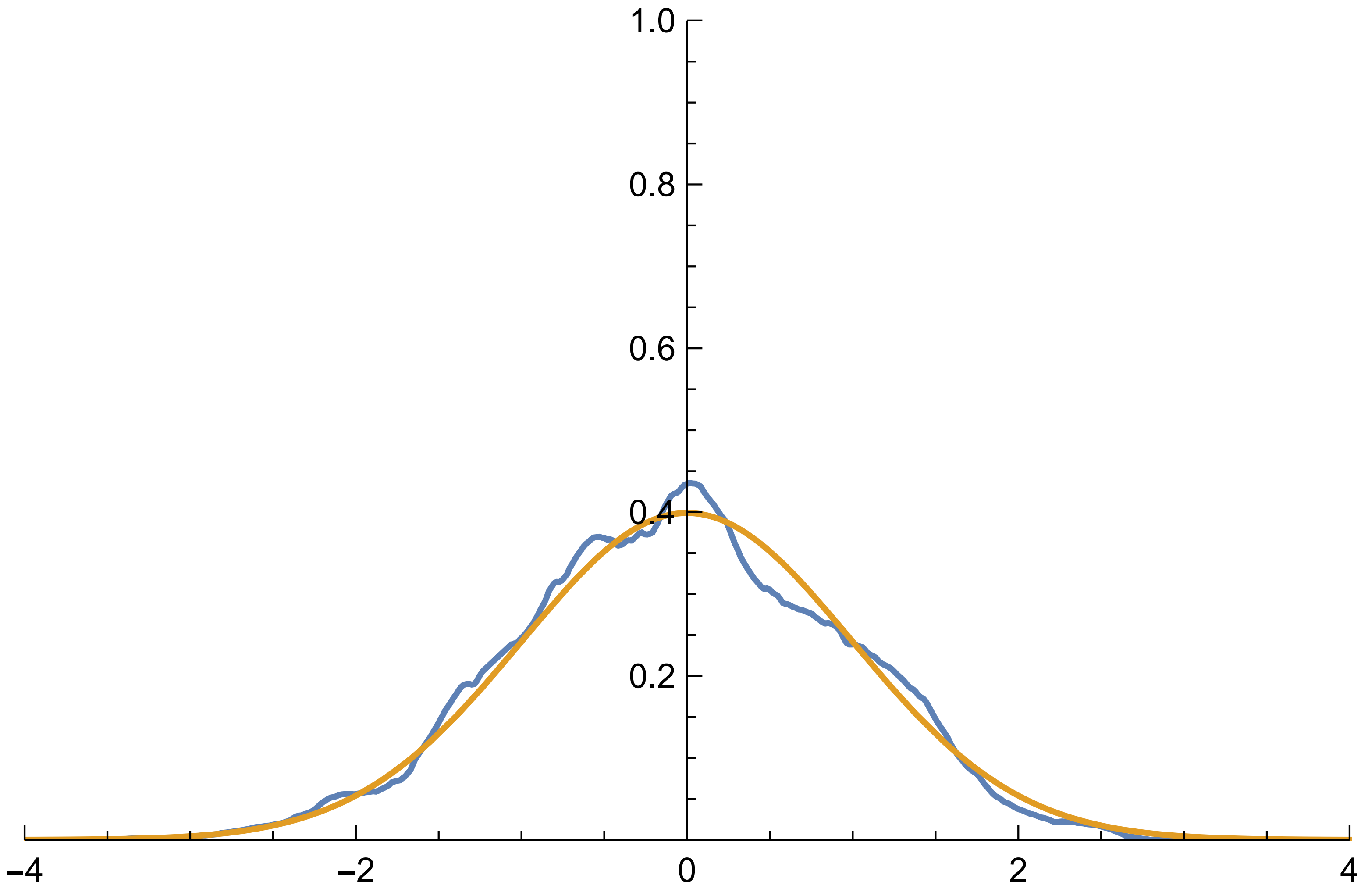
Przykłady użycia różnych funkcji jądra

Półokrągłe (n=100)



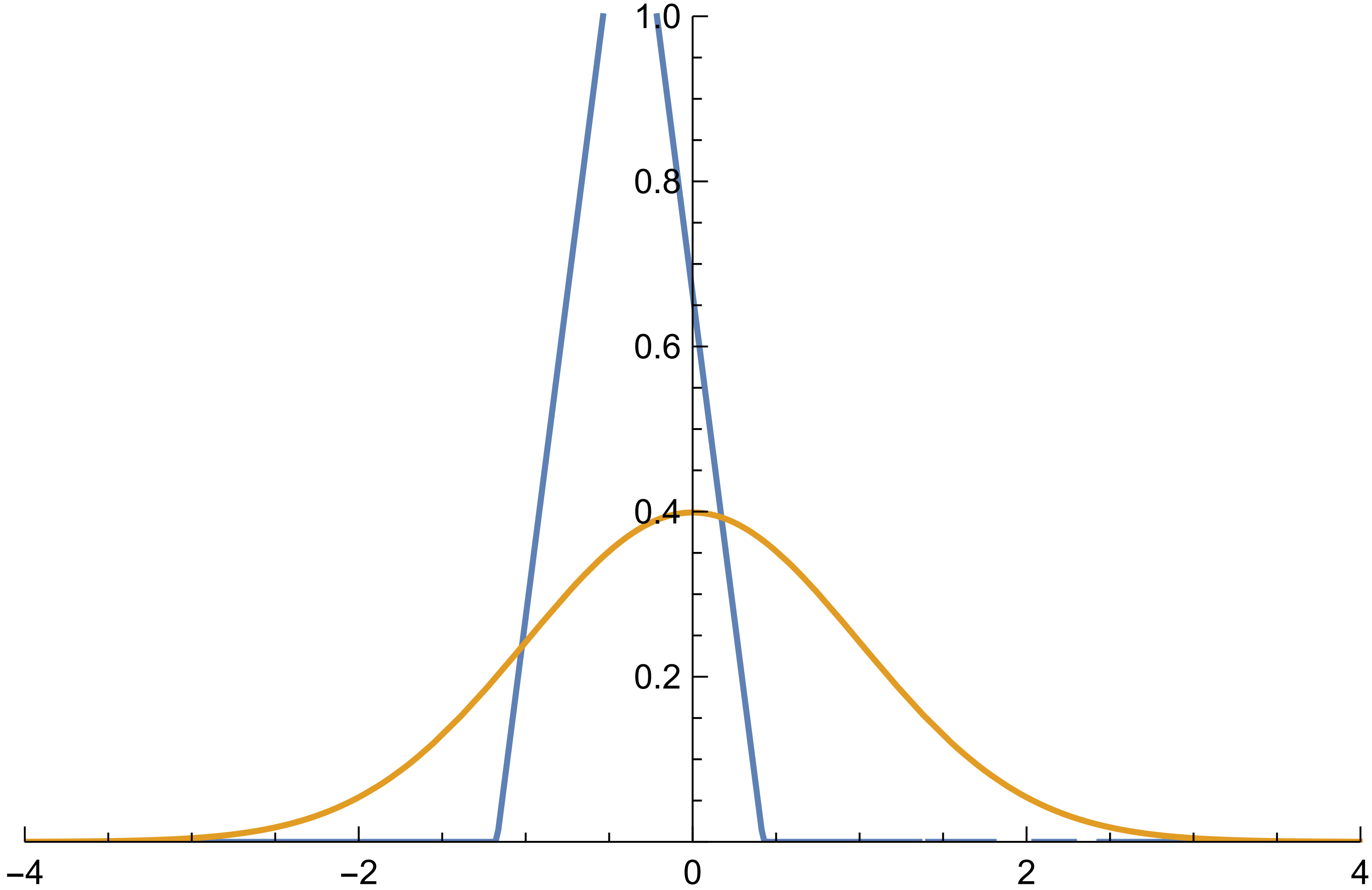
Przykłady użycia różnych funkcji jądra

Półokrągłe (n=1000)



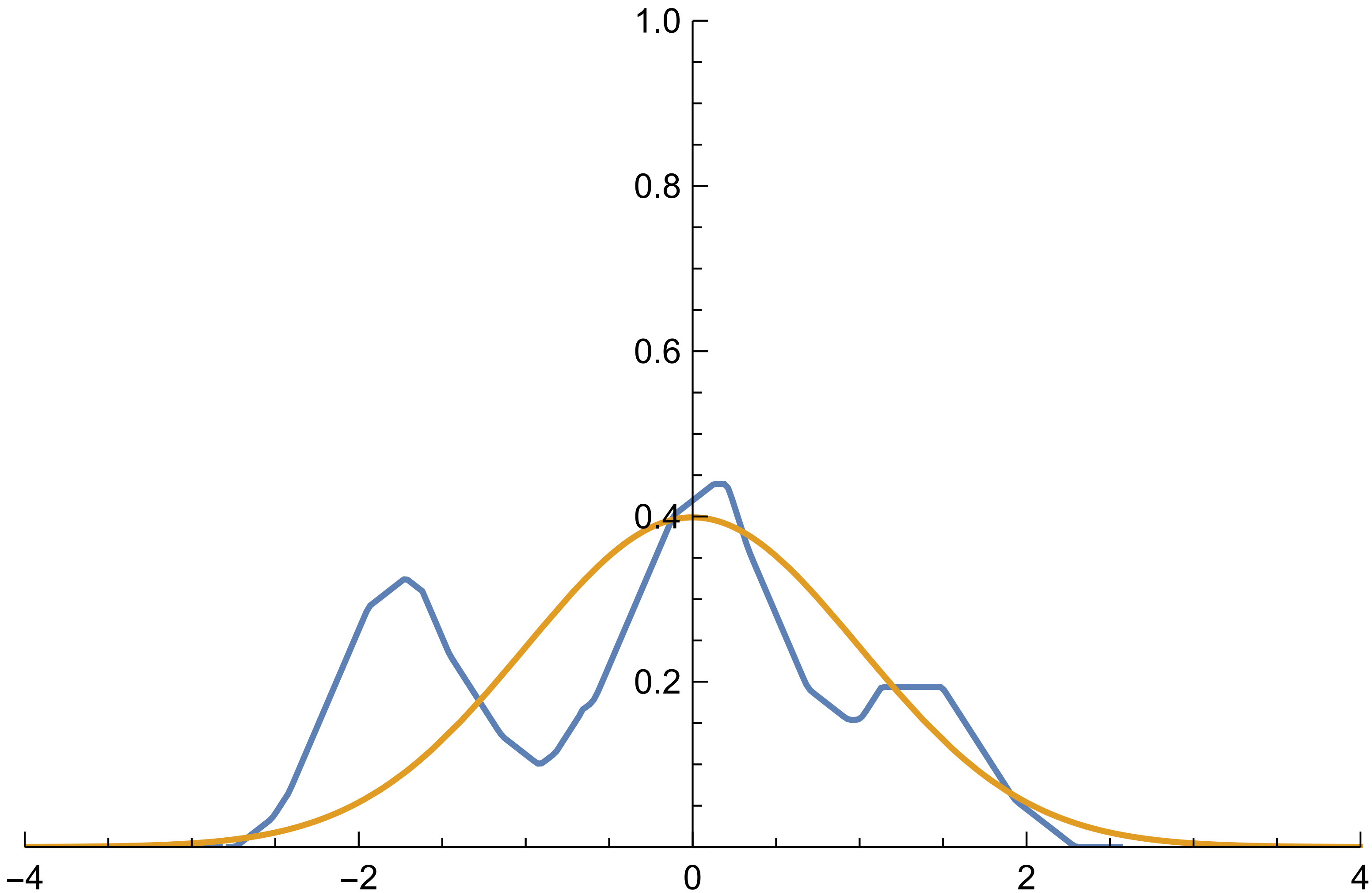
Przykłady użycia różnych funkcji jądra

Trójkątne (n=1)



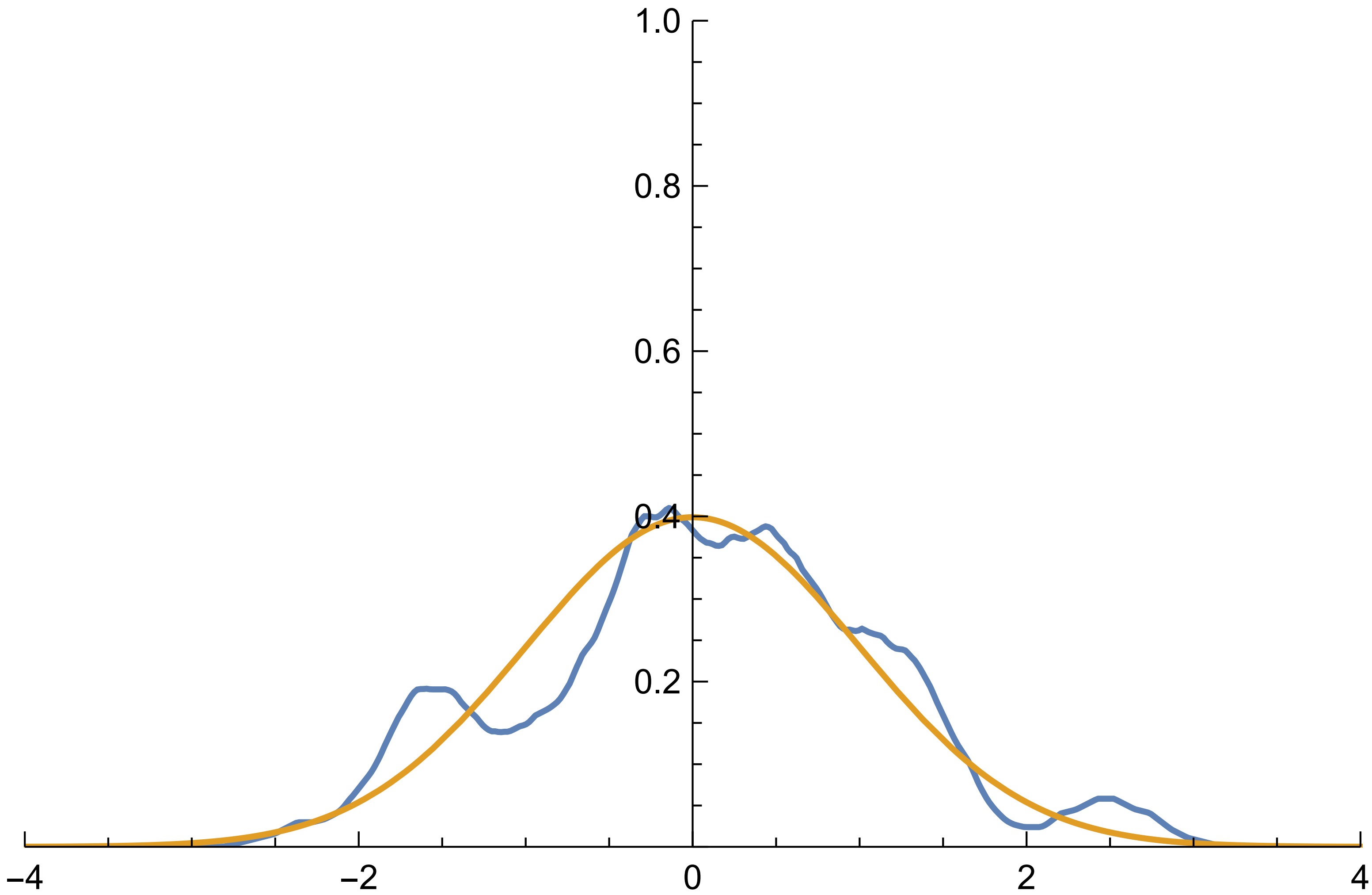
Przykłady użycia różnych funkcji jądra

Trójkątne (n=10)



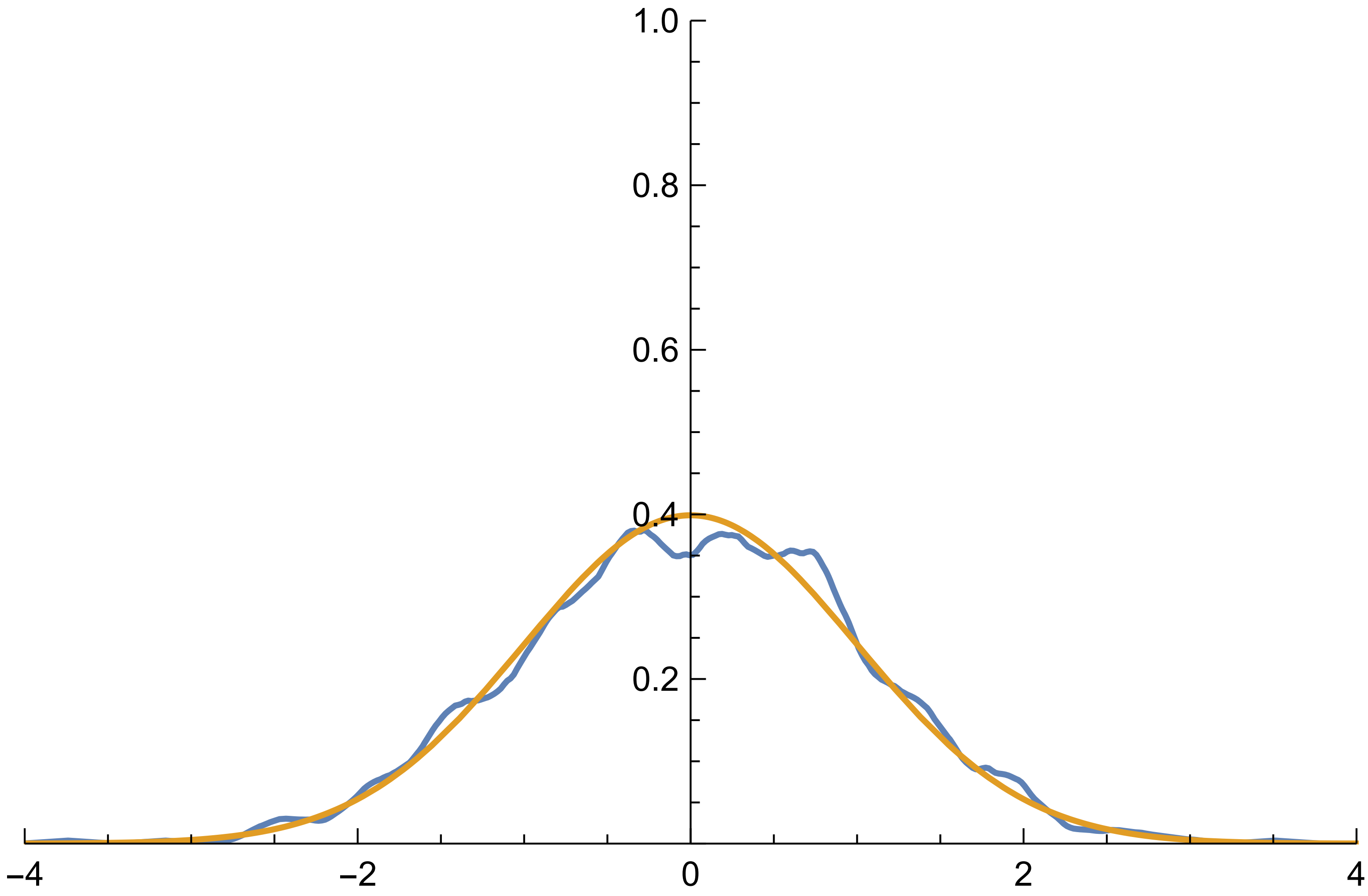
Przykłady użycia różnych funkcji jądra

Trójkątne (n=100)



Przykłady użycia różnych funkcji jądra

Trójkątne (n=1000)

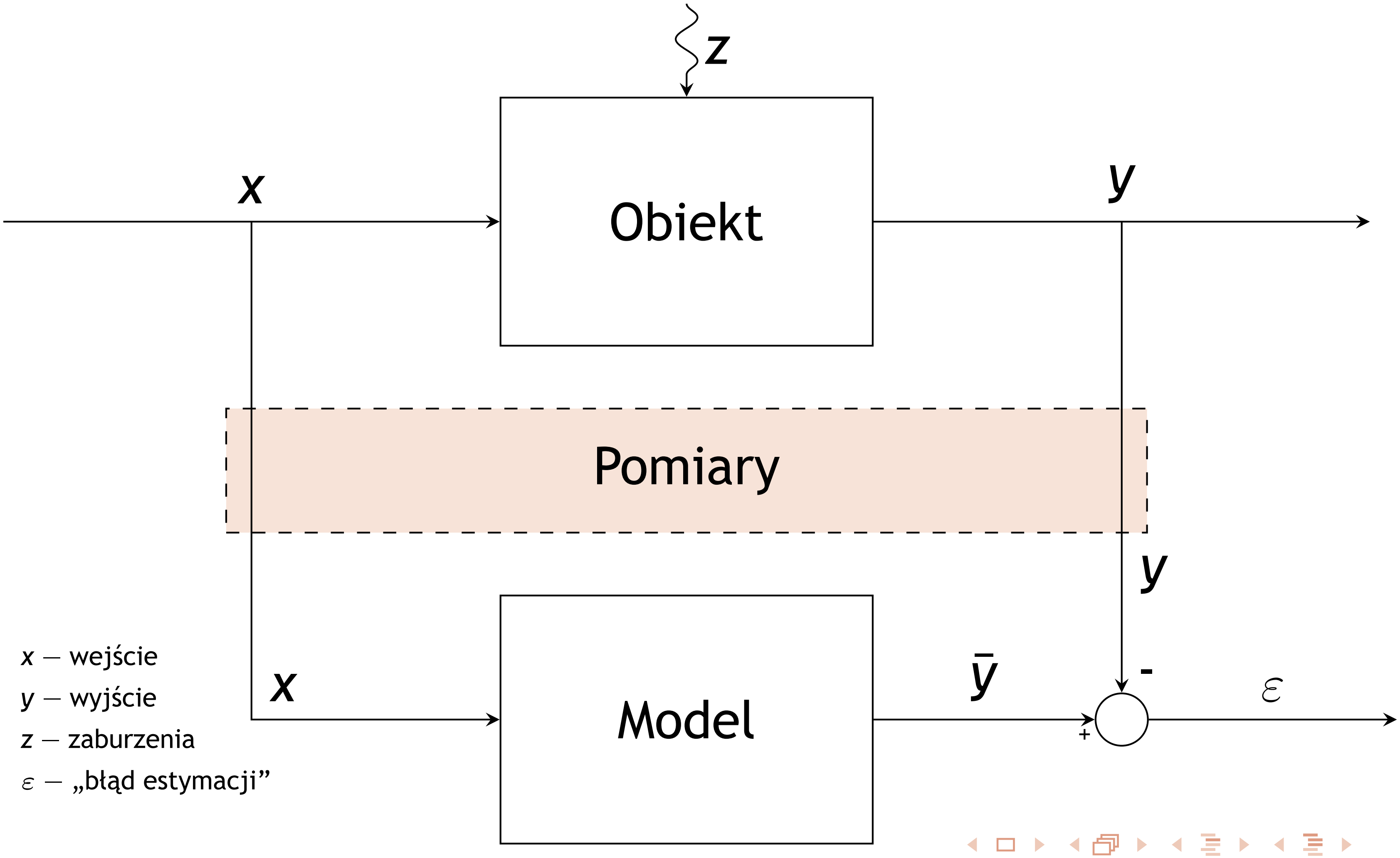


Eksperyment

1. Właściwie trudno tu mówić o prowadzeniu eksperymentu (w obiegowym tego słowa rozumieniu).
2. Rola eksperymentatora sprowadza się do roli **obserwatora**.
3. Ten rodzaj ekspetymentu nazywany również bywa **eksperymentem biernym**.

Prawdziwy eksperyment (czynny): na obiekt możemy wpływać

Nowa sytuacja przedstawia się tak:



x — wejście
y — wyjście
z — zaburzenia
ε — „błąd estymacji”

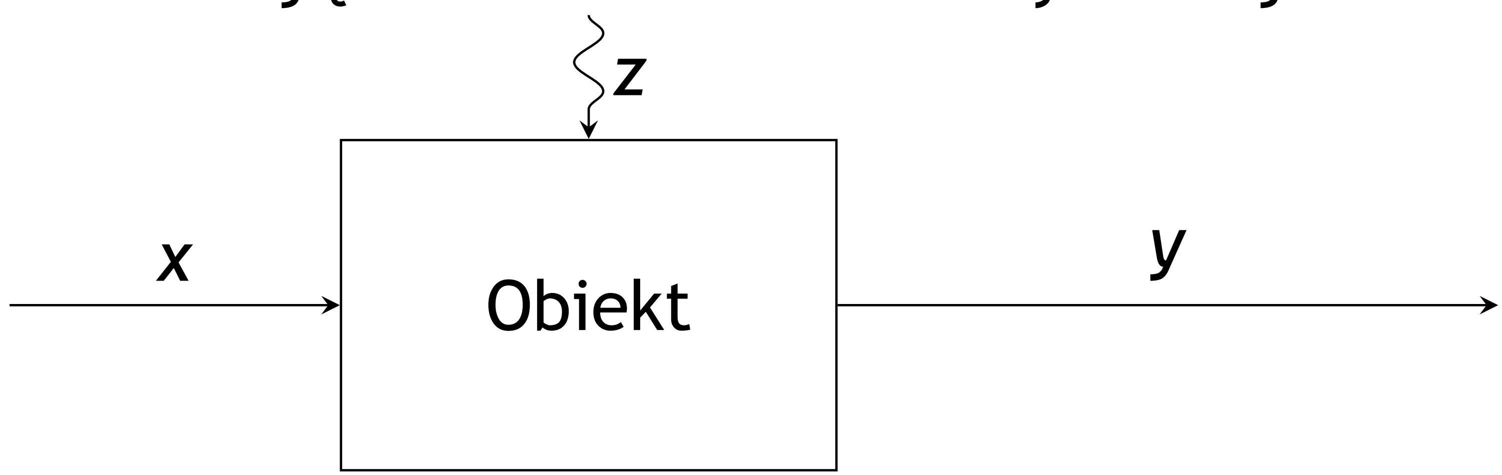
Nazewnictwo

1. SISO – Single Input, Single Output (x i y jednowymiarowe)
2. MISO – Multiple Input, Single Output (X wektor, y jednowymiarowe)
3. MIMO – Multiple Input, Multiple Output (X, Y – wektory)

Zburzenia

Jeżeli chodzi o zaburzenia, to:

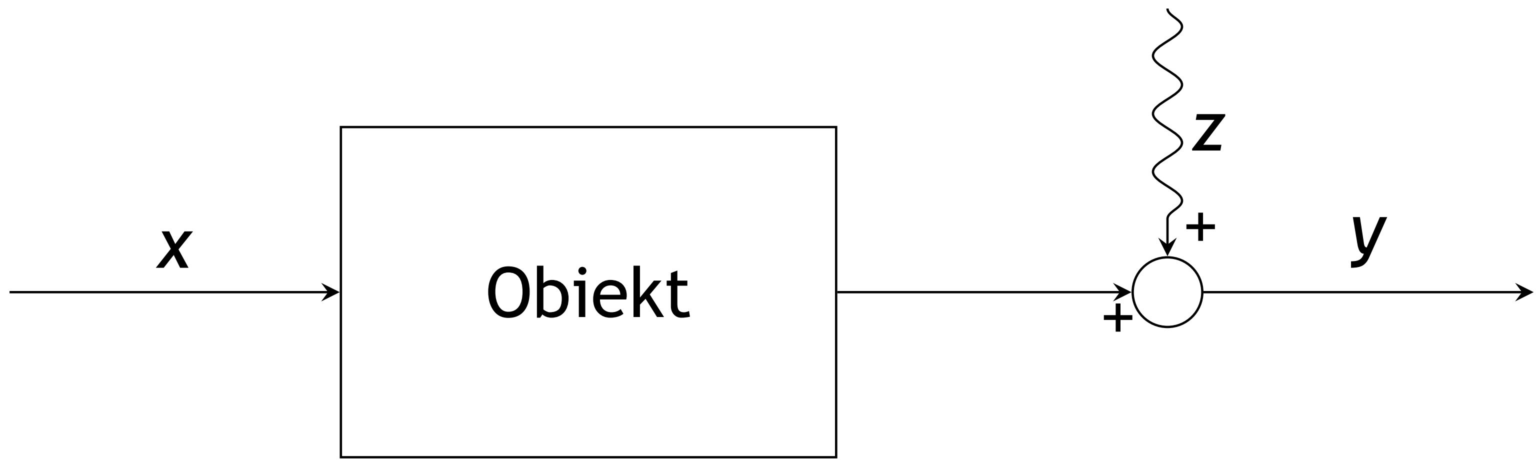
- zakłada się, że są „przyzwoite”:
  - losowe,
  - niezależne,
  - o stałym rozkładzie,
  - najchętniej normalne.
- Oddziałują na obiekt we właściwym miejscu:



Zburzenia

Jeżeli chodzi o zaburzenia, to:

- zakłada się, że są „przyzwoite”:
  - losowe,
  - niezależne,
  - o stałym rozkładzie,
  - najchętniej normalne.
- Oddziałują na obiekt we właściwym miejscu:



Identyfikacja I

- Zakładając, że prawo według którego działa obiekt da się opisać pewną funkcją: $y = f(x)$
- identyfikacja polega na wyznaczeniu takiej funkcji $\bar{y} = \bar{f}(x)$, że różnica $\varepsilon = ||\bar{y} - y||$ jest najmniejsza (używając wybranej metryki $||\cdot||$).
- Dokonuje się tego na podstawie zestawu obserwacji: $\{x_i, y_i\}, i = 1, 2, \ldots, N$
- rozwiązując zadanie optymalizacji:

$$Q = \sum_{i=1}^{N} ||\bar{f}(x_i) - y_i|| \to \text{min!}$$

Identyfikacja II

- Ponieważ tak ogólnie postawione zadanie nie jest łatwe do rozwiązania, zakłada się jakąś parametryczną postać funkcji $\bar{f}(x)$.
  - liniową:
    $$\bar{y} = ax + b$$
    (x i y oraz a i b mogą być wektorami).
  - wielomianową:
    $$\bar{y} = \sum_{j=0}^{n} a_j x^j$$

Identyfikacja III

- czy ogolniej

$$\bar{y} = \sum_{j=0}^{n} a_j f_j(x)$$

- Kwestia doboru modelu (funkcji $f(x)$ jest bardzo ważna (i trochę na poboczu niniejszych rozważań). Najlepiej by było, żeby odzwierciedlała ona zjawiska fizyczne zachodzące w obiekcie. Pojawia się szereg kwestii – fizycznej interpretacji poszczególnych parametrów modelu (w przypadku modelu parametrycznego).

Identyfikacja IV

- Dochodzi do tego jeszcze kwestia taka, że tak y jak i x w realnym świecie mają jakieś wymiary fizyczne. Zatem funkcja f powinna mieć również jakąś interpretacje „fizyczną” (czy, na przykład, ekonomiczną, gdy rozważamy takie procesy).

Model

Z różnych względów, najpopularniejszym modelem jest model liniowy (lub liniowy ze względu na parametry). Do identyfikacji parametrów używa się najbardziej znanych (standardowych) wzorów:

Jeżeli

$$y = \sum_j a_j x_j$$

to możemy to zapisać w postaci wektorowej

$$y = \mathbf{a}'\mathbf{x}$$

Utworzymy macierz

$$X = [\mathbf{x}_1, \mathbf{x}_2 \ldots, \mathbf{x}_N]$$

Model cd

i wektor

$$Y' = [y_1, y_2, \dots, y_N]$$

Przyjmując normę kwadratową możemy kryterium identyfikacji zapisać jako:

$$Q = (a'X - Y')(a'X - Y')' = a'XX'a - a'XY - Y'X'a + Y'Y$$

minimum formy kwadratowej wysnaczyć jest stosunkowo prosto (pochodna po a i miejsce zerowe układu równań liniowych):

$$a = (X'X)^{-1}X'Y$$

Model nieliniowy

Ogólnie, przyjmując ogólniejszy model nieliniowy

$$y = F(x, a)$$

zadanie identyfikacji sprowadza się do minimalizacji

$$Q = \sum_{i=1}^{N} (F(x_i, a) - y_i)^2 \rightarrow \text{min!}$$

Czy da sie problem rozwiązać w przypadku ogólnym? To znaczy poszukać funkcji

$$y = F(x)$$

Identyfikacja nieparametryczna

- Idea zaczerpnięta ze statystyki
- Ze statystycznego punktu widzenia interesuje nas coś takiego:
$$\mathrm{E}(Y|X) = m(X)$$
$\bar{y}$ jako warunkowa wartość oczekiwana, przy ustalonej wartości x.
- $$\mathrm{E}(Y|X=x) = \int yf(y|x)dy = \int y\frac{f(x,y)}{f(x)}dy = \frac{\int yf(x,y)dy}{f(x)}$$
Da się to przedstawić jako pewne operacje na funkcjach gęstości

Identyfikacja nieparametryczna (2)

- Jeżeli przyjąć nieparametryczne estymatory gęstości, oraz do estymacji gęstości dwu zmiennych użyć jądra będącego iloczynem jąder dla jednej zmiennej

$$K(x, y) = K(x)K(y)$$

$$\hat{f}(x, y) = n^{-1}h^{-2}\sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right) K\left(\frac{y - y_i}{h}\right)$$

i

$$\hat{f}(x) = n^{-1}h^{-1}\sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right)$$

Identyfikacja nieparametryczna (3)

- Otrzymujemy estymator postaci:

$$\widehat{m}_h(x) = \frac{\sum_{i=1}^{n} K_h(x - x_i) y_i}{\sum_{i=1}^{n} K_h(x - x_i)}$$

- bo $\int yK\left(\frac{y-y_i}{h}\right) dy = y_i$
- Można to zapisać również inaczej: $m(x) = \sum_i w_{in} y_i$ gdzie $w_{in} = \frac{K\left(\frac{x-x_i}{h}\right)}{\sum_{j=1}^{n} K\left(\frac{x-x_j}{k}\right)}$