Zachodniopomorski Uniwersytet Technologiczny w Szczecinie

Projekt i implementacja aplikacji realizującej operacje arytmetyczne na kwantyfikatorach lingwistycznych

Bartosz Taczała

Szczecin, 2009

ZACHODNIOPOMORSKI UNIWERSYTET TECHNOLOGICZNY

Abstract

Wydział Informatyki

Project and implementation of application evaluating operations on linguistic quantifiers

by Bartosz Taczała

Analitic solution of decisive problems, based only on probabilities of event occurence, is extermally difficult and time-consuming and in most cases incorrect due to arithmetic exceptions made in progress of calculations. In This Master Thesis I will introduce author's application, capable of resolving those decision problems, by capably operating of probability distribution functions for probabilities given in problem. The main focus is on the methodology of solving such problem and as a particular example the problem of "Two business" is solved.

Oświadczenie

Oświadczam, ze przedkładaną pracę magisterską kończącą studia napisałem samodzielnie. Oznacza to, ze przy pisaniu pracy poza niezbędnymi konsultacjami, nie korzystałem z pomocy innych osób, a w szczególności nie zlecałem opracowania rozprawy lub jej części innym osobom, ani nie odpisywałem rozprawy lub jej części od innych osób. Potwierdzam tez zgodność wersji papierowej i elektronicznej złożonej pracy. Mam świadomość,że poświadczenie nieprawdy będzie w tym przypadku skutkowało cofnięciem decyzji o wydaniu dyplomu.

Podpis: ________________________________

Data: ________________________________

Spis treści

Abstract **i**

Declaration of Authorship **ii**

Spis rysunków **v**

Spis kodów źródłowych **vii**

Skróty **viii**

Symbole **ix**

1 Wprowadzenie **1**
1.1 Tematyka pracy . . . 1
1.2 Motywacja . . . 2
1.3 Cel pracy . . . 2

2 Część teoretyczna **3**
2.1 Wyjaśnienie podstawowych zagadnień z zakresu statystyki matematycznej 3
2.2 Funkcja gęstości . . . 7
2.2.1 Funkcja gęstości dla sumy zmiennych losowych . . . 8
2.2.2 Funkcja gęstości dla różnicy zmiennych losowych . . . 9
2.2.3 Funkcja gęstości dla ilorazu zmiennych losowych . . . 10
2.2.4 Funkcja gęstości dla iloczynu zmiennych losowych . . . 10
2.3 Wyjaśnienie podstawowych zagadnień z zakresu matematyki rozmytej . . 11
2.3.1 Zmienne lingwistyczne, zbiory rozmyte, zbiory spełniające warunek podziału do jedności . . . 11
2.3.2 Kwantyfikator lingwistyczny . . . 12
2.3.3 kwantyfikator lingwistyczny jako funkcję rozkładu gęstości prawdopodobieństwa? . . . 14
2.4 Aproksymacja funkcji wielomianem algebraicznym . . . 15

3 Część praktyczna **19**
3.1 Opis założeń programu . . . 19
3.1.1 Środowiska działania programu . . . 20
3.2 Opis technologi użytych przy tworzeniu oprogramowania . . . 20
3.2.1 Język programowanie oraz użyte narzędzia . . . 20

3.3 Struktura programu 21
3.3.1 Opis modułów użytych w programie 22
3.3.2 Opis modułów wewnętrznych 22
3.3.2.1 Szczegółowy opis biblioteki fl 22
3.3.2.2 Opis modułu operacji matematycznych 25
3.3.2.3 Opis modułu metod aproksymacji 28
3.3.2.4 Opis części graficznej 29
3.3.3 Opis modułów zewnętrznych 29
3.4 Przykładowe użycie 30

4 Rozwiązanie przykładowego zadania **34**
4.1 Postawienie zadania 34
4.2 Identyfikacja funkcji gęstości 35
4.2.1 Kwantyfikator "duże" 36
4.2.2 Kwantyfikator "małe" 38
4.2.3 Kwantyfikator "bliskie 10 mln" 39
4.2.4 Kwantyfikator "nie bliskie 10 mln" 41
4.2.5 Kwantyfikator "bliskie 20 mln" 42
4.2.6 Kwantyfikator "nie bliskie 20 mln" 43
4.3 Rozwiązanie problemu 45
4.3.1 Przedstawienie analitycznej formy rozkładów gęstości prawdopodobieństwa dla zadania 46
4.3.2 Funkcje rozkładu gęstości prawdopodobieństwa dla biznesu pierwszego 48
4.3.3 Funkcje rozkładu gęstości prawdopodobieństwa dla biznesu drugiego 50
4.3.4 Wartości oczekiwane i błąd 53
4.3.5 Odpowiedź na problem dwóch biznesów 53

Bibliografia **59**

Spis rysunków

2.1 Zbiór rozmyty . . . 12
2.2 Zbiory rozmyte spełniające warunek podziału jedności . . . 12

3.1 Diagram przedstawiający poglądowych schemat struktury aplikacji . . . 21
3.2 Diagram klas dla modułu fl . . . 22
3.3 Rysunek przedstawiający dialog *Nowa funkcja ciągła* . . . 31
3.4 Wykresy funkcji gęstości dla funkcji *gauss(x-4,0,1)* oraz *gauss(x+2,0.5,1)* 31
3.5 Dialog wyboru funkcji do przeprowadzenia operacji na nich . . . 32
3.6 Dialog wyboru funkcji do aproksymacji . . . 32
3.7 Dialog ukazujący szczegóły dowolnej funkcji . . . 33

4.1 Wykres kwantyfikatora duże . . . 36
4.2 Wykres kwantyfikatora *duże* wraz z odpowiadającą mu funkcją rozkładu gęstości prawdopodobieństwa . . . 38
4.3 Wykres kwantyfikatora małe . . . 39
4.4 Wykres kwantyfikatora *małe* wraz z odpowiadającą mu funkcją rozkładu gęstości prawdopodobieństwa . . . 40
4.5 Wykres funkcji gęstości dla kwantyfikatora *bliskie 10 mln* . . . 41
4.6 Wykres funkcji gęstości dla kwantyfikatora *nie bliskie 10 mln* . . . 42
4.7 Wykres kwantyfikatora bliskie 20 mln . . . 43
4.8 Wykres kwantyfikatora *nie bliskie 20 mln* . . . 44
4.9 Wykres funkcji gęstości dla kwantyfikatorów *duże* i *małe* . . . 45
4.10 Wykres funkcji gęstości dla kwantyfikatorów *bliskie 10 mln* i *nie bliskie 10 mln* . . . 46
4.11 Wykres funkcji gęstości dla kwantyfikatorów *bliskie 20 mln* i *nie bliskie 20 mln* . . . 47
4.12 Wykres przedstawiający niedokładność aproksymacji funkcji gęstości *bliskie 10 mln*$\otimes$*duże* . . . 50
4.13 Wykres funkcji gęstości dla kwantyfikatorów *małe* i *nie bliskie 10 mln* oraz funkcji wynikowej dla operacji $\otimes$ między nimi . . . 51
4.14 Wykres przedstawiający niedokładność aproksymacji funkcji gęstości *małe* $\otimes$ *nie bliskie 10 mln* . . . 52
4.15 Wykres funkcji gęstości gp_{DB1} . . . 53
4.16 Wykres przedstawiający niedokładność aproksymacji funkcji gęstości gp_{DB1} 54
4.17 Wykres funkcji gęstości dla kwantyfikatorów *duże* i *nie bliskie 20 mln* oraz funkcji wynikowej dla operacji $\otimes$ między nimi . . . 54
4.18 Wykres przedstawiający niedokładność aproksymacji funkcji gęstości *nie bliskie 20 mln*$\otimes$*duże* . . . 55

4.19 Wykres funkcji gęstości dla kwantyfikatorów *małe* i *nie bliskie 10 mln* oraz funkcji wynikowej dla operacji $\otimes$ między nimi 55
4.20 Wykres przedstawiający niedokładność aproksymacji funkcji gęstości *małe* $\otimes$ *nie bliskie 10 mln* 56
4.21 Wykres funkcji gęstości gp_{DB2} 56
4.22 Wykres przedstawiający niedokładność aproksymacji funkcji gęstości gp_{DB2} 57
4.23 Wykres przedstawiający niedokładność aproksymacji funkcji gęstości gp_{DB2} 58

Spis kodów źródłowych

3.1 Sygnatura klasy FunctionBase 23
3.2 Sygnatura klasy FunctionContinousBase 24
3.3 Sygnatura klasy Function2D::Function2DBase 24
3.4 Sygnatura klasy Function2D::FunctionMixed 26
3.5 Przykłady użycia typów, funkcji dwuwymiarowych, z przestrzeni fl 27
3.6 Podstawowy interfejs dla operacji matematycznych *IOperation* 28
3.7 Podstawowy interfejs dla metod aproksymacji 28

Skróty

STL **S**tandard **T**emplate **L**ibrary

GCC **G**nu **C**ommon **C**compilers

G++ kompilator języka C++ z pakietu GCC

MSVC **M**icro**S**oft **V**isual **C++**

IDE **I**ntegrated **D**evelopment **E**nvironment **GDB** **G**nu **D**e**b**ugger

Symbole

$\oplus$ - Operacja dodowania kwantyfikatorów lingwistycznych

$\ominus$ - Operacja odejmowania kwantyfikatorów lingwistycznych

$\otimes$ - Operacja mnożenia kwantyfikatorów lingwistycznych

$\oslash$ - Operacja dzielenia kwantyfikatorów lingwistycznych

Rozdział 1

Wprowadzenie

1.1 Tematyka pracy

W pracy autor przedstawi metodykę operowania na kwantyfikatorach lingwistycznych, które traktowane będą jako funkcje rozkładu gęstości prawdopodobieństwa. Operowanie to będzie skutkowało w rozwiązywaniu problemów, które nie są przedstawione w sposób dokładny, za pomocą zidentyfikowanych wartości , ale w sposób bardziej intuicyjny dla człowieka, czyli w języku naturalnym. Stawianie problemów w języku naturalnym i ich rozwiązywanie jest dużo bliższe ludziom, gdyż jest to naturalny sposób w jakim ludzie przekazują między sobą informacje, za pomocą terminów 'duży', 'mały', 'średni', a nie 'od 1 do 2 z dokładnością do 0.1', czy '95 %'. Dużo łatwiej jest konstruować i rozumieć sformułowania 'Jacek jest bardzo wysoki', lub 'Jacek ma około 2 metrów', niż 'wzrost Jacka mieści się w granicach 190-200 centymetrów, z błędem około 5 centymetrów'.
Praca zawiera projekt i implementację programu komputerowego, który wykonuje podstawowe działania na funkcjach rozkładu gęstości prawdopodobieństwa, które będą odzwierciedleniem kwantyfikatorów lingwistycznych m.in:

- dodawanie,
- odejmowanie,
- mnożenie,
- dzielenie.

Umiejętne stosowanie takich działań umożliwi nam rozwiązywanie pewnej klasy zadań , w których dane są przedstawione nie za pomocą liczb, ale właśnie za pomocą kwantyfikatorów językowych, naturalnych dla człowieka. Jedno z takich zadań zostanie rozwiązane z ostatnim rozdziale, przy wykorzystaniu programu będącego częścią pracy magisterskiej.

1.2 Motywacja

Motywacją do podjęcia tego tematu, była chęć sprawdzenia jak komputer radzi sobie w dziedzinach matematyki, gdzie częściej niż jawnymi liczbami, operujemy terminami rozmytymi, językowymi. Wiadomym jest, że komputery dużo sprawniej i szybciej radzą sobie z obliczeniami numerycznymi niż ludzie, natomiast nie są pomocne, gdy próbujemy za ich pomocą rozwiązać problem, gdzie terminologią nie są liczby , ale słowa.

1.3 Cel pracy

Celem pracy jest zaimplementowanie aplikacji, która potrafi wykonywać proste operacje na rozkładach gęstości będącymi odzwierciedleniem kwantyfikatorów lingwistycznych. W ostatnim rozdziale autor przedstawi napisany program jako narzędzie do rozwiązania problemu językowego.

Rozdział 2

Część teoretyczna

W poniższym rozdziale zostaną przedstawione oraz wyjaśnione podstawowe pojęcia dotyczące statystyki matematycznej, probabilistyki oraz matematyki rozmytej. Pojęcia te muszą zostać wprowadzone i wytłumaczone, by móc zrozumieć kluczowe dla tej pracy operacje na funkcjach rozkładu gęstości prawdopodobieństwa utożsamionych z kwantyfikatorami lingwistycznymi.

2.1 Wyjaśnienie podstawowych zagadnień z zakresu statystyki matematycznej

Z bardzo szerokiej dyscypliny jaką jest statystyka matematyczna oraz probabilistyka przedstawione i wyjaśnione zostały tylko te pojęcia, które bezpośrednio łączą się z tematem pracy, lub wymagają objaśnia, gdyż będą pomocne przy objaśnieniu innych pojęć. Zanim autor przedstawi pojęcia najważniejsze w kontekście tej pracy, wyjaśnione zostaną pojęcia fundamentalne dla dyscypliny statystyki matematycznej.

W bardzo wielu dziedzinach matematyki, czy innych istnieją pojęcia pierwotne, niedefiniowalne. Dla statystyki pojęciem takim jest pojęcie *zdarzenia elementarnego*. W praktyce *zdarzeniem elementarnym* określamy jednorazowy wynik eksperymentu statystycznego, taki jak rzut kością do gry, czy monetą.

Zbiorem zdarzeń elementarnych (przestrzenią zdarzeń elementarnych), oznaczanym, zwyczajowo przez Ω, nazywamy zbiór wszystkich niepodzielnych wynikami doświadczenia czy obserwacji, *czyli zdarzeń elementarnych*[7]. Jedno zdarzenie określamy przez ω.

Zmienną losowa określamy jednoznaczne przyporządkowanie każdemu zdarzeniu elementarnemu wartości liczbowej, czyli funkcję $X(\omega)$, która dla każdego $\omega \in \Omega$ przyporządkuję jedną i tylko jedną liczbę $X(\omega) = x$[2]. Zmienne losowe dzielimy na:

- *zmienna losowa typu skokowego* - jeśli może przyjmować skończoną lub nieskończoną, ale przeliczalną liczbę wartości[2],
- *zmienna losowa typu ciągłego* - jeśli możliwe jej wartości tworzą przedział ze zbioru liczb rzeczywistych[2].

Tak więc *zmienna losowa* jest to każde przekształcenie zbioru Ω w zbiór R^1[7]. Zmienne losowe począwszy od teraz dla uproszczenia oznaczamy dużymi końcowymi literami X, Y, Z.

Po wyjaśnieniu pojęć elementarnych autor przejdzie do zdefiniowania prawdopodobieństwa oraz rozkładu prawdopodobieństwa.

> Na rodzinie zbiorów S, określamy funkcję rzeczywistą P, o której zakłada się że spełnia następujące aksjomaty:
> AKSJOMAT I. *Dla każdego* $A \in S$
>
> $$P(A) \geqslant 0,$$
>
> *czyi funkcja P jest nieujemna.*
> AKSJOMAT II. *Prawdopodobieństwo zdarzenia pewnego jest równe jedności, czyli*
>
> $$P(\Omega) = 1.$$
>
> AKSJOMAT III. *Dla każdego ciągu $A_1, A_2, ...$ zdarzeń parami rozłącznych($A_i \bigcap A_j = \varnothing$ dla $i, j = 1, 2, ..; i \neq j$ prawdziwa jest równość*
>
> $$P(\bigcup_{i=1}^{\infty}) = \sum_{i=1}^{\infty} P(A_i)$$
>
> Funkcję P nazywamy *rozkładem prawdopodobieństwa na S. Prawdopodobieństwem* zaś od tej pory będziemy nazywać wartość tej funkcji na jednym zdarzeniu.

By zdefiniować dystrybuantę, rozważony zostanie przypadek, gdzie Ω jest jednowymiarową przestrzenią euklidesową, tzn. $\Omega = R^1$. Gdy chcemy wyznaczyć funkcję P w R^1, należy zdefiniować pomocniczą funkcję F określoną w R^1, mającą następujące własności

1. F jest niemalejąca,
2. $$\lim_{x \to -\infty} F(x) = 0, \lim_{x \to \infty} F(x) = 1 \tag{2.1}$$
3. F jest lewostronnie ciągła.

Ogólna postać naszej pomocniczej funkcji F ma postać:

$$F(x) = \int_{-\infty}^{x} f(u)du, \text{ gdzie } f \geqslant 0 \text{ i } \int_{-\infty}^{\infty} f(u)du = 1. \tag{2.2}$$

Warunek, że funkcja F jest niemalejąca, oznacza, że dla każdego $a < b$, całka z funkcji f po dowolnym przedziale $(a; b)$ jest nieujemna. Funkcję P wyznacza za pomocą funkcji F w sposób:

$$P(\langle a; b)) = F(b) - F(a). \tag{2.3}$$

Przy tak zadanej funkcji F wzór 2.3 określa funkcję P dla każdego $A = \langle a; b)$, czyli określa funkcję na zbiorze wszystkich przedziałów. Funkcję F o takich własnościach nazywamy *dystrybuantą*.
Dla tak zadanych definicji *rozkładu prawdopodobieństwa* oraz *dystrybuanty*, można wyprowadzić definicję *rozkładu prawdopodobieństwa zmiennej losowej* oraz *dystrybuanty zmiennej losowej.*

> Dana jest przestrzeń (Ω, S, P) oraz zmienna losowa X. *Rozkładem prawdopodobieństwa zmiennej losowej* nazywamy funkcję P_x utworzoną w następujący sposób: dla każdego $A \in S$
>
> $$P_x(A) = P(\{\omega : X(\omega) \in A\}, \tag{2.4}$$
>
> oraz spełniającą układ aksjomatów I-III.
> *Dystrybuantę zmiennej losowej* wyznaczoną przez P_x określamy jako:
>
> $$F_x(A) = P_x((-\infty, x)). \tag{2.5}$$

Dystrybuanta oraz dystrybuanta zmiennej losowej nie różnią się pojęciowo, różnica w nazwie zaznacza związek ze zmienną losową.

N potrzeby dalszych rozważań wyjaśnione zostaną pojęcia *niezależności zdarzeń*, a następnie *niezależności zmiennych losowych.* Od tej pory wszystkie brane pod uwagę pary zdarzeń i zmiennych losowe uważamy za niezależne.

Niech $P(B) > 0$. Mówimy, że zdarzenie A nie zależy od zdarzenia B, jeśli

$$P(A|B) = P(A). \tag{2.6}$$

Przekształcając 2.6 oraz wzór na prawdopodobieństwo warunkowe[1], otrzymujemy ważną własność:

$$P(A \bigcap B) = P(A)P(B), \tag{2.7}$$

gdy zdarzenia A i B są niezależne.

Natomiast dla zmiennych losowych definicji ich niezależności brzmi następująco:

Zmienne losowe X_1 i X_2 są *niezależne*, jeśli dla dowolnych zbiorów S_1 S_2, zdarzenia $Z_1 : \{\omega : X_1(\omega) \in S_1\}$ oraz $Z_2 : \{\omega : X_2(\omega) \in S_2\}$ spełniają warunek:

$$P(Z_1 \bigcap Z_2) = P(Z_1)P(Z_2). \tag{2.8}$$

Powyższy warunek jest spełniony, wtedy i tylko wtedy, gdy:

$$F(x, y) = F(x_1)F(x_2). \tag{2.9}$$

W związku z tym warunek 2.9 można również traktować jako definicję zmiennych losowych niezależnych.

Z rozkładem zmiennej losowej są związane pewne, charakteryzujące go w syntetyczny sposób wielkości liczbowe. Charakterystyki te są nazywane parametrami zmiennej losowej. Autor przedstawi jeden taki parametr, *wartość oczekiwaną*, który posłuży przy wnioskowaniu w przedstawionym w ostatnim rozdziale przykładzie.

Wartością oczekiwaną *zmiennej losowej* X (ciągłej) nazywamy wyrażenie:

$$E(X) = \int_{-\infty}^{\infty} f(x) * x \, dx, \tag{2.10}$$

gdzie $f(x)$ jest funkcją gęstości[2] [3][2].

---

[1] $P(A|B) = \frac{P(A \bigcap B)}{P(B)} s$

[2] Wyjaśniona w następnym podrozdziale

[3] Może się zdarzyć, że rozkład zmiennej losowej nie ma wartości oczekiwanej. Dzieje się tak wtedy, gdy całka występująca we wzorze 2.10 nie jest bezwzględnie zbieżna

2.2 Funkcja gęstości

Kolejne pojęcia będziemy omawiać tylko dla zmiennej losowej typu ciągłego, gdyż tylko taka jest użyteczna z naszego punktu widzenia. Dla takich losowych zmiennych definiujemy funkcję gęstości prawdopodobieństwa następująco:

> Jeżeli zmienna losowa X, typu ciągłego, o dystrybuancie F oraz istnieje funkcja $f \geqslant 0$, że dla każdego x zachodzi równość:[7]
>
> $$F(x) = \int_{-\infty}^{x} f(u)du, \tag{2.11}$$
>
> wtedy funkcję f nazywamy *funkcją gęstości* lub *funkcją rozkładu gęstości prawdopodobieństwa.*

Dla każdej funkcji f, będącej gęstością zachodzi zatem zależność:

$$\int_{-\infty}^{\infty} f(x)dx = F(\infty) = 1. \tag{2.12}$$

co oznacza: prawdopodobieństwo zdarzenia, gdy zmienna losowa ciągła jest mniejsza od nieskończoności jest zawsze równe jedności. W punktach, w których funkcja f jest ciągła, zależność 2.11 możemy zróżniczkować stronami względem x i wówczas otrzymujemy

$$F^{'}(x) = f(x) \tag{2.13}$$

Z definicji pochodnej funkcji oraz z własności wyrażonej w 2.13, w punktach ciągłości funkcji zachodzi zależność:

$$f(x) = F^{'}(x) = \lim_{\Delta x \to \infty} \frac{F(x + \Delta x) - F(x)}{\Delta x} = \lim_{\Delta x \to \infty} \frac{P(x \leqslant X < x + \Delta x)}{\Delta x}. \tag{2.14}$$

By odnaleźć wzór na obliczanie prawdopodobieństwa $P(x_1 \leqslant X < x_2)$ na podstawie gęstości f, należy rozwiązać:

$$P(x_1 \leqslant X < x_2) = F(x_2) - F(x_1) = \int_{-\infty}^{x_2} f(u)\,du - \int_{-\infty}^{x_1} f(u)\,du = \int_{x_1}^{x_2} f(u)\,du. \tag{2.15}$$

Powyższą formułę odczytuje się następująco : Prawdopodobieństwo zdarzenia, gdy zmienna losowa jest mniejsza niż x_2 a większa niż x_1 wynosi wartość całki z funkcję gęstości na przedziale $(x1; x2)$.

Poniżej zostały przedstawione rozkłady prawdopodobieństwa oraz funkcje gęstości dla zmiennej losowych, będącej wynikiem operacji na dwóch innych, niezależnych zmiennych losowych. Operacje te to dodawanie, odejmowanie, mnożenie oraz dzielenie.Dla tych operacji na zmiennych losowych, wprowadzone zostaną następujące symbole operujące na ich rozkładach:

- $\oplus$ - operacja między rozkładami zmiennej X_1 oraz X_2, dająca rozkład zmiennej $S = X_1 + X_2$,
- $\ominus$ - operacja między rozkładami zmiennej X_1 oraz X_2, dająca rozkład zmiennej $S = X_1 - X_2$,
- $\otimes$ - operacja między rozkładami zmiennej X_1 oraz X_2, dająca rozkład zmiennej $S = X_1 * X_2$,
- $\oslash$ - operacja między rozkładami zmiennej X_1 oraz X_2, dająca rozkład zmiennej $S = X_1/X_2$.

2.2.1 Funkcja gęstości dla sumy zmiennych losowych

Chcemy znaleźć wzór funkcji gęstości dla zmiennej losowej będącej sumą dwóch innych zmiennych losowych, zakładając przy tym ich niezależność. zmiennąS jest ustalona jako suma zmiennych X_1 i X_2.

$$S = X_1 + X_2 \tag{2.16}$$

Aby odnaleźć wzór funkcji gęstości dla takiego rozkładu prawdopodobieństwa zmiennej losowej oraz jej dystrybuanty rozwiązać należy równanie[5]:

$$F(S) = P(S\ leqs) = P(X_1 + X_2 \leqslant s) = \int\limits_0^{x_2}\int\limits_0^{x_1} f(x_1, x_2)dx_1\,dx_2 \tag{2.17}$$

Rozkład łącznych dwóch niezależnych zmiennych losowych wygląda następująco:

$$f(x_1, x_2) = f_1(x_1)f_2(x_2), \tag{2.18}$$

a zmienną x_2 można przedstawić jako:

$$s = x_1 + x_2 \Rightarrow x_2 = s - x_1, \tag{2.19}$$

co po podstawieniu do 2.17 i zmianie granic całkowania daje:

$$F(S) = \int\limits_0^{x_1}\int\limits_0^{x_2} f_1(x_1)f_2(x_2)dx_2\,dx_1 = \int\limits_0^{x_1}\int\limits_0^{s-x_1} f_1(x_1)f_2(x_2)dx_2\,dx_1. \tag{2.20}$$

By znaleźć funkcję gęstości należy zgodnie z równaniem 2.13 zróżniczkować 2.17:

$$f(s) = \frac{\partial}{\partial s}\int_0^{x_1}\int_0^{s-x_1} f_1(x_1)f_2(x_2)dx_2\,dx_1 = \int_0^{x_1} f_1(x_1)(\frac{\partial}{\partial s}\int_0^{s-x_1} f_2(x_2)dx_2)\,dx_1. \tag{2.21}$$

Wartość różniczki w całce wewnętrznej wynosi, zgodnie z podstawowym twierdzeniem rachunku całkowego [4]

$$\frac{\partial}{\partial s}\int_0^{s-x_1} f_2(x_2)dx_2 = \frac{\partial}{\partial s}F_2(s-x_1) = f_2(s-x_1). \tag{2.23}$$

Podstawiając powyższy wynik do równania 2.21 otrzymujemy:

$$F(s) = \int_0^{x_1} f_1(x_1)f_2(s-x_1)\,dx_1, \tag{2.24}$$

co jest wynikową dystrybuantą zmiennej s. Funkcja podcałkowa to szukana funkcja gęstości[5].

2.2.2 Funkcja gęstości dla różnicy zmiennych losowych

Dystrybuanta oraz funkcja gęstości dla różnicy zmiennych losowych liczona jest analogicznie jak dla sumy zmiennych losowych. Głównym równaniem jest[5]:

$$F(S) = P(S\ leqs) = P(X_1 + X_2 \leqslant s) = \int_0^{x_2}\int_0^{x_1} f(x_1, x_2)dx_1\,dx_2, \tag{2.25}$$

gdzie S jest zdefiniowane następująco:

$$S = X_1 - X_2, \tag{2.26}$$

oraz

$$s = x_1 - x_2 \Rightarrow x_2 = x_1 - s. \tag{2.27}$$

Wykonując analogicznie operacja jak dla sumy końcowym wynikiem jest równanie:

$$F(s) = \int_0^{x_1} f_1(x_1)f_2(x_1-s)\,dx_1. \tag{2.28}$$

[4] Twierdzenie można przedstawić w postaci następującej:

$$\frac{\partial}{\partial dt}\int_{-\infty}^{a} f(t)\,dt = f(a) \tag{2.22}$$

2.2.3 Funkcja gęstości dla ilorazu zmiennych losowych

Dystrybuanta dla ilorazu zmiennych losowych liczona jest analogicznie jak dla sumy i różnicy. Głównym równaniem jest[5]:

$$F(S) = P(S\ leqs) = P(X_1 + X_2 \leqslant s) = \int_0^{x_2}\int_0^{x_1} f(x_1, x_2) dx_1\, dx_2, \tag{2.29}$$

gdzie S jest zdefiniowane następująco:

$$S = X_1 * X_2, \tag{2.30}$$

oraz

$$s = x_1 * x_2 \Rightarrow x_2 = \frac{s}{x_1}. \tag{2.31}$$

Wyjątkiem jest operacja liczenia pochodnej z dystrybuanty $F_2(X_2)$:

$$\frac{\partial}{\partial s}\int_0^{\frac{s}{x_1}} f_2(x_2)dx_2 = \frac{\partial}{\partial s}F_2(\frac{s}{x_1}) = f_2(\frac{s}{x_1})\frac{1}{x_1}, \tag{2.32}$$

gdzie należy uwzględnić pochodną funkcji wewnętrznej. Końcowe równanie wygląda zatem następująco:

$$F(s) = \int_0^{x_1} f_1(x_1)f_2(\frac{s}{x_1})\frac{1}{x_1}\, dx_1. \tag{2.33}$$

2.2.4 Funkcja gęstości dla iloczynu zmiennych losowych

Dla iloczynu zmiennych losowych $s = \frac{x_1}{x_2}$, będącego ostatnią operacją dla której podana zostanie dystrybuanta oraz funkcję gęstości. Dla tej operacji pomięte zostaną operacje szczątkowe i przedstawione zostanie jedynie równanie ostateczne[5]:

$$F(s) = \int_0^{x_1} f_1(x_1)f_2(\frac{x_1}{s})x1\, dx_1. \tag{2.34}$$

2.3 Wyjaśnienie podstawowych zagadnień z zakresu matematyki rozmytej

2.3.1 Zmienne lingwistyczne, zbiory rozmyte, zbiory spełniające warunek podziału do jedności

Zanim wyjaśnione zostaną terminy z zakresu kwantyfikatorów należy wprowadzić podstawowe pojęcia dla matematyki liczb rozmytych.

Zmienną lingwistyczną nazywamy wielkość wejściową, wyjściową bądź zmienną stanu, który zamierzamy ocieniać stosując oceny lingwistyczne (przykładem jest np. prędkość statku)[6].
Wartością lingwistyczną jest słowna ocena wielkości lingwistycznej (przykładem jest są terminy: bardzo duży, średni , brzydki, ładne, młody, stary)[6].
Wartości lingwistyczne występują w modelach wraz ze *zmiennymi lingwistycznymi* których dotyczą (np. wysokie ciśnienie powietrza, średnia prędkość samochodu)[6]. *Przestrzeń lingwistyczna zmiennej* jest zbiorem wszystkich wartości lingwistycznych stosowanych do oceny danej zmiennej lingwistycznej. Przestrzeń tak jest zbiorem skończeniewymiarowym[6]. *Przestrzenią numeryczną zmiennej* nazywamy zbiorem wszystkich wartości numerycznych jakie może ona realnie przyjąć w rozpatrywanym systemie, lub też takich wartości, które są istotna dla rozwiązywanego problemu[6].

Zbiorem rozmytym w pewnej numerycznej przestrzeni rozważań X, nazywamy zbiór par:

$$A = \{(\mu_a^*(x), X))\}, \forall x \in X, \tag{2.35}$$

gdzie: μ_a jest funkcją przynależności zbioru rozmytego A, która każdemu elementowi $x \in X$ przypisuje stopień jego przynależności $\mu_a^*(x)$ do zbioru rozmytego A, przy czym:

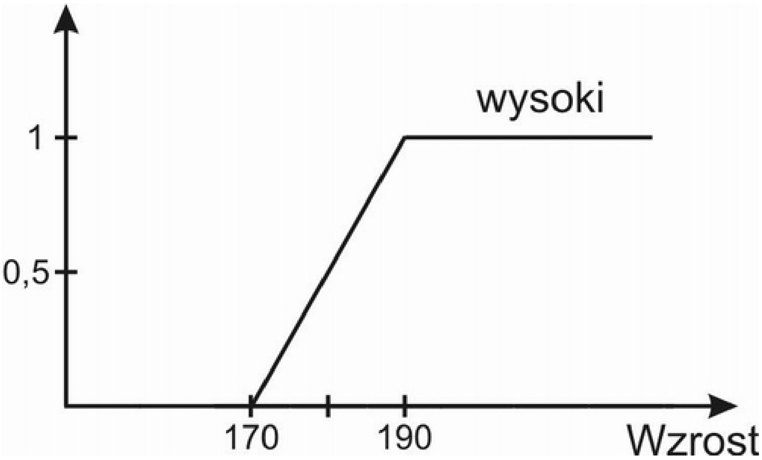
$$\mu_A(x) \in [0, 1]. \tag{2.36}$$

Przykładową interpretację zbioru rozmytego pokazano na rysunku 2.1. Szczególnymi zbiorami rozmytymi na zadanej przestrzeni rozważań, są zbiory rozmyte *spełniające warunek podziału jedności*[5]:

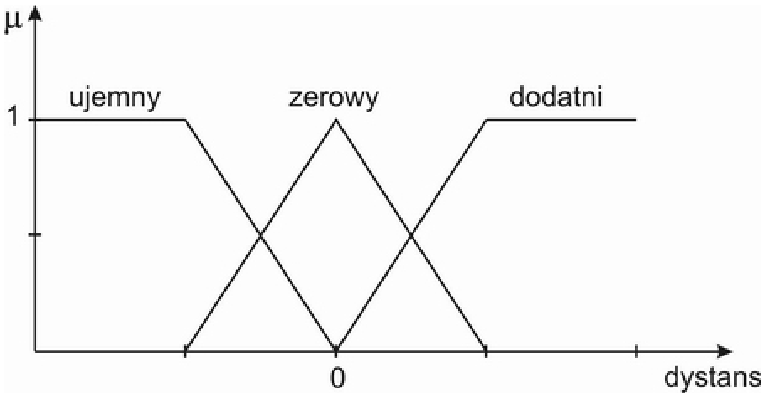
$$\sum_h \mu_{A_h}(x) \equiv 1, \forall x, \tag{2.37}$$

gdzie h oznacza numer zbioru rozmytego[6][4].

[5]przedstawione na rysunku 2.2



RYSUNEK 2.1: Zbiór rozmyty



RYSUNEK 2.2: Zbiory rozmyte spełniające warunek podziału jedności

2.3.2 Kwantyfikator lingwistyczny

Kwantyfikatorem lingwistycznym nazywamy taki kwantyfikator, który jest zdefiniowany przez wyrażenia lingwistyczne (np. 'większość' czy 'około 3')[1].
Wyróżniane są dwa typy kwantyfikowanych stwierdzeń:

- Q X jest A,
- Q B X jest A,

gdzie

- Q - kwantyfikator lingwistyczny,
- X - zbiór, gdzie stopień przynależności dowolnego obiektu do zbioru wynosi 0 lub 1,

- A, B - rozmyte predykaty.

Pierwszy typ oznacza w języku naturalnym: 'Q elementów należących do X spełnia A'. Przykładem jest sformułowanie : *'Większość pracowników jest młoda'*, gdzie:

- Q - oznacza kwantyfikator większość,
- X - oznacza zbiór pracowników,
- A - warunek 'jest młoda'.

Drugi typ kwantyfikatora oznacza: 'Pośród elementów X spełniających B, warunek A jest spełniony przez Q elementów'. Przykładem jest sformułowanie : *'Większość młodych pracowników dobrze zarabia'*, gdzie:

- Q - oznacza kwantyfikator większość,
- B - oznacza warunek 'jest młoda'.
- X - oznacza zbiór pracowników,
- A - oznacza warunek 'dobrze zarabia'.

Reprezentacją kwantyfikatora jest zbiór rozmyty [6]. Funkcja przynależności (dla kwantyfikatora typu 'Q X jest A') w tym zbiorze oznacza stopień spełnienia predykatu A przez j-ty element zbioru X.

Na rysunku 2.2 przedstawiono 3 zbiory rozmyte, które możemy potraktować jako kwantyfikatory lingwistyczne. Praktyczne, intuicyjne, rozumienie takich kwantyfikatorów to odpowiedź na pytanie: co oznacza określenie zmiennej lingwistycznej jako ujemnej ? Tak zdefiniowane kwantyfikatory określają liczbę ujemną i tylko ujemną (nie zerową czy dodatnią) gdy wpada ona do przedziału $[-\infty, -1]$. Gdy natomiast liczba wpada do przedziału $[-1, -0.5]$, można powiedzieć o tej liczbie, że jest ujemna oraz zerowa (mówimy o niej że jest 'zerowa' rzadziej, częściej jest dla nas ujemna). Gdy liczba wpada do przedziału $[-0.5, 0]$ można mówić o niej również , że może być ujemna i może być zerowa (tym razem częściej o niej mówimy że jest zerowa). Analogicznie sytuacja ma się z liczbami określanymi jako zerowe oraz dodatni.

[6] dla kwantyfikatora absolutnego jest nim zbiór rozmyty w przestrzeni rozważań R, dla kwantyfikatora relatywnego, jest nim zbiór rozmyty z przestrzeni zwartej w przedziale $[0, 1]$ [1]

2.3.3 kwantyfikator lingwistyczny jako funkcję rozkładu gęstości prawdopodobieństwa?

Jak rozumieć próbę połączenia dwóch zupełnie różnych terminów z różnych dziedzin matematycznych ? Otóż, przy uważnym zastanowieniu, można dojść do wniosku że oba te terminy nie są od siebie tak dalekie, można powiedzieć nawet że są sobie bliskie, tylko znajdują się w zupełnie różnych działach matematyki.

Autor postara się uzmysłowić tożsamość pojęć na przykładzie. ‘Jaka jest szansa znalezienia na ziemi człowieka o wzroście 192,68 cm ¿. Odpowiedź na to pytanie może być zaskakująca dla większości społeczeństwa, ale szansa ta jest równa zero. W języku naturalnym, wyjaśnia się, że ponieważ wzrost jest wielkością ciągłą to sensowne jest jedynie pytanie o wzrost zawierający się w pewnym przedziale.
W języku statystyki matematycznej znana jest własność, [7]:

$$P(X = x_0) = \lim_{\Delta x \to 0} P(x_0 \leqslant X < x_o + \Delta x) = \tag{2.38}$$

$$P(X = x_0) = \lim_{\Delta x \to 0} P(x_0 \leqslant X < x_o + \Delta x) \tag{2.39}$$

$$= \lim_{\Delta x \to 0} \int_{x_0}^{x_0+\Delta} f(x)dx = \int_{x_0}^{x_0} f(x)dx = 0. \tag{2.40}$$

Nie mówimy zatem o prawdopodobieństwie znalezienia człowieka o wzroście 192,68 centymetrów, ale o prawdopodobieństwie znalezienia człowieka, którego wzrost mieści się w dość bliskim przedziale[8].

Tak sam schemat obowiązuje przy kwantyfikatorach. Stawiając pytanie ‘Jaka jest szansa znalezienia na ziemi człowieka dużego? ‘, szukamy osoby, której wzrost mieści się w pewnym przedziale. Jeżeli badana zmienna mieści się w przedziale, gdzie funkcja przynależności kwantyfikatora ‘duże‘ ma wartość różną od zera, to można stwierdzić, że dana wartość charakteryzuje w jakimś stopniu osobnika jako dużego. Inaczej sformułowany to pytanie brzmi: ‘Jaka jest szansa znalezienia osobnika o wzroście z i sklasyfikowania go jako dużego? ‘ W języku statystyki wyraża się to przez zapis:

$$P(wzrost = duzy|x = z) = ? \tag{2.41}$$

Gdy uda nam się scałkować funkcję gęstości dla takiego prawdopodobieństwa, otrzymamy jego rozkład, który będzie odpowiedzią na postawione przez pytanie.

[7] przedstawiona wraz z dowodem
[8] Oczywiście mowa tutaj jedynie o wielkościach ciągłych oraz o rozkładach zmiennej losowej ciągłej

Warto zaznaczyć, że w tego typu rozważaniach należy tak definiować kwantyfikatory aby odpowiadające im zbiory rozmyte spełniały warunek podziału do jedności. Dla kwantyfikatorów określających wzrost, duży, średni, mały, zachodzić ma następujące równość:

$$\begin{aligned} P(wzrost = duzy|x = z) &+ P(wzrost = sredni|x = z) + \\ &+ P(wzrost = maly|x = z) = 0 \end{aligned}$$

2.4 Aproksymacja funkcji wielomianem algebraicznym

W ostatnim podrozdziale teoretycznym przedstawię metodykę aproksymacji funkcji. Ta operacja będzie potrzeba przy operacjach na funkcjach rozkładów gęstości prawdopodobieństwa. Jak autor wcześniej zaznaczył interesują nas tylko zmienne losowe typu ciągłego, a więc ciągłe rozkłady prawdopodobieństwa oraz ciągłe funkcje gęstości. Jednak, każda operacja na rozkładach gęstości, wykonywana w środowisku komputerowym, w wyniku utworzy funkcję dyskretną. Aproksymacja będzie to metodą umożliwiająca nam powrót do zmiennej ciągłej.
Zadanie: funkcję zadaną w postaci dyskretniej, poprzez zdefiniowanie zbioru punktów w których ta funkcja jest określona, należy przybliżyć funkcją ciągłą na zadanym przedziale. Funkcja ta ma być dana wielomianem m-tego stopnia, gdzie m musi spełniać zależność

$$n > m + 1 \Rightarrow m < n - 1 \tag{2.42}$$

, gdzie:

- n - ilość punktów przez który zadana jest funkcja aproksymowana
- m - stopnień wielomianu

Funkcje bazowe wynikowego wielomianu są postaci :

$$\begin{aligned} U_0(x) &= x^0 \\ U_1(x) &= x^1 \\ \dots\dots\dots\dots & \\ U_m(x) &= x^m. \end{aligned}$$

Funkcja aproksymująca ma postać:

$$F(x) = a_0 + a_1x + a_2x^2 + ... + a_mx^m, \tag{2.43}$$

a odchylenie standardowe wyraża się następująco:

$$S = \sum_{i=1}^{n} (a_0 + a_1 x + a_2 x^2 + ... + a_m x^m - y_i)^2. \tag{2.44}$$

To równanie nazywane jest formą kwadratową i wykazane zostało, że ma tylko jedno minimum. W takim przypadku warunek konieczny istnienia minimum, 2.45 , jest jednocześnie warunkiem koniecznym. Aby znaleźć funkcję aproksymującą należy znaleźć współczynniki $a_0, a_1, a_2, ..., a_m$, takie aby S miało wartość minimalną. W tym celu będziemy przyrównywać pochodne cząstkowe do zera:

$$\begin{aligned}
\frac{\partial S}{\partial a_0} &= 0 \\
\frac{\partial S}{\partial a_1} &= 0 \\
.... & \\
\frac{\partial S}{\partial a_m} &= 0
\end{aligned}$$

Różniczkując 2.44 otrzymujemy układ równań liniowych:

$$\begin{aligned}
2 \sum_{i=n}^{n} (a_0 + a_1 x + a_2 x^2 + ... + a_m x^m - y_i) * 1 &= 0 \\
2 \sum_{i=n}^{n} (a_0 + a_1 x + a_2 x^2 + ... + a_m x^m - y_i) * x_i &= 0 \\
2 \sum_{i=n}^{n} (a_0 + a_1 x + a_2 x^2 + ... + a_m x^m - y_i) * x_i^2 &= 0 \\
.... & \\
2 \sum_{i=n}^{n} (a_0 + a_1 x + a_2 x^2 + ... + a_m x^m - y_i) * x_i^m &= 0
\end{aligned}$$

o m+1 niewiadomych i takiej samej ilości zmiennych. Korzystając z prawa przemienności sumowania otrzymujemy:

$$\begin{aligned}
a_0 \sum_{i=1}^{n} 1 + a_1 \sum_{i=1}^{n} x_i + a_2 \sum_{i=1}^{n} {x_i}^2 + ... + a_m \sum_{i=1}^{n} {x_i}^m &= \sum_{i=1}^{n} y_i \\
a_0 \sum_{i=1}^{n} x_i + a_1 \sum_{i=1}^{n} {x_i}^2 + a_2 \sum_{i=1}^{n} {x_i}^3 + ... + a_m \sum_{i=1}^{n} {x_i}^{m+1} &= \sum_{i=1}^{n} x_i * y_i \\
a_0 \sum_{i=1}^{n} {x_i}^2 + a_1 \sum_{i=1}^{n} {x_i}^3 + a_2 \sum_{i=1}^{n} {x_i}^4 + ... + a_m \sum_{i=1}^{n} {x_i}^{m+2} &= \sum_{i=1}^{n} {x_i}^2 * y_i \\
... & \\
a_0 \sum_{i=1}^{n} {x_i}^m + a_1 \sum_{i=1}^{n} {x_i}^{m+1} + a_2 \sum_{i=1}^{n} {x_i}^{m+2} + ... + a_m \sum_{i=1}^{n} {x_i}^{m+m} &= \sum_{i=1}^{n} {x_i}^m * y_i.
\end{aligned}$$

Wprowadzając oznaczenia:

$$\begin{aligned} S_k &= \sum_{i=1}^{n} {x_i}^k (k = 0, 1, 2, ..., 2m) \\ t_k &= \sum_{i=1}^{n} {x_i}^k y_i (k = 0, 1, 2, ..., 2m) \end{aligned}$$

otrzymujemy równanie macierzowe:

$$\begin{bmatrix} S_0 & S_1 & ... & S_m \\ S_1 & S_2 & ... & S_{m+1} \\ ... & ... & ... & ... \\ S_{m+1} & S_{m+2} & ... & S_{m+m} \end{bmatrix} * \begin{bmatrix} a_0 \\ a_1 \\ ... \\ a_m \end{bmatrix} = \begin{bmatrix} t_0 \\ t_1 \\ ... \\ t_m \end{bmatrix}, \tag{2.45}$$

którego przekształcamy do:

$$\begin{bmatrix} a_0 \\ a_1 \\ ... \\ a_m \end{bmatrix} = \begin{bmatrix} S_0 & S_1 & ... & S_m \\ S_1 & S_2 & ... & S_{m+1} \\ ... & ... & ... & ... \\ S_{m+1} & S_{m+2} & ... & S_{m+m} \end{bmatrix}^{-1} 2 * \begin{bmatrix} t_0 \\ t_1 \\ ... \\ t_m \end{bmatrix}, \tag{2.46}$$

co ogólnie można zapisać jako:

$$A = S^{-1} * T \tag{2.47}$$

Rozwiązanie takiego równanie istnieje wtedy i tylko wtedy, gdy macierz S jest odwracalna. Wartości $a_0, a_1, ..., a_m$ są rozwiązaniem równania 2.46 i stanowią stałe w naszej funkcje aproksymującej 2.43. Aby poznać jakość aproksymacji może posłużyć się następującymi parametrami:

- odchylenie kwadratowe:

$$S = \sum_{i=1}^{n} (a_0 + a_1 x_i + a_2 x_i^+ .. + a_m {x_i}^m - y_i)^2 \tag{2.48}$$

- wariancja:

$$S_{y/x}^2 = \frac{S}{n - (m+1)} \tag{2.49}$$

- odchylenie standardowe:

$$S_{y/x} = \delta = \sqrt{S_{y/x}^2} \tag{2.50}$$

- współczynnik korelacji:

$$r = \sqrt{1 - \frac{S_{y/x}^2}{S_y^2}} \tag{2.51}$$

gdzie:

$${S_y}^2 = \frac{\sum_{i=1}^{n} {y_i}^2 - (\sum_{i=1}^{n} y_i)^2/n}{n-1} \tag{2.52}$$

jest wariancją zmiennej y.

Rozdział 3

Część praktyczna

3.1 Opis założeń programu

Program będący częścią pracy magisterskiej musi spełniać następujące założenia odnośnie oferowanej funkcjonalności:

- wczytywać i operować na funkcjach w postaciach:
  - dyskretnej,
  - ciągłej,
  - mieszanej, czyli funkcja jest postaci ciągłej, ale tylko w ściśle określonych przedziałach, ponadto wzór funkcji w różnych przedziałach może się różnić,
- wyświetlać funkcje na skalowalnym przedziale,
- wykonywać operacje:
  - dodawanie,
  - odejmowanie,
  - mnożenie,
  - dzielenie,

  na funkcjach rozkładów gęstości dla kwantyfikatorów,
- całkować funkcje na zadanym przedziale,
- liczyć wartość oczekiwaną dla funkcji rozkładu gęstości prawdopodobieństwa,
- aproksymować funkcje dyskretne do funkcji ciągłych.

3.1.1 Środowiska działania programu

Jednym z głównych założeń autora przy pisaniu programu, była możliwość uruchomienia go na jak najszerszym zakresie sprzętu komputerowego bez względu na system operacyjny czy architekturę. Założenie zostało zrealizowane w wystarczającym stopniu. Cała aplikacja została napisana w sposób przenośny, tak, że dla każdego środowiska w którym działa kompilator języka C++ oraz dostępne są używane biblioteki program działa w taki sam sposób. Moduł graficzny został napisany z wykorzystaniem bibliotek Qt, dzięki czemu kod zarządzający grafiką jest przenośny między wszystkimi wiodącymi systemami operacyjnymi (Microsoft Windows®, GNU/Linux, Mac OSX®, czy UNIX i inne). Autorowi udało się poprawnie skompilować i uruchomić aplikację na poniższych platformach:

- Microsoft Windows XP®,
- GNU/Linux,
- Mac OSX Leopard,
- Maemo, platforma mobilna, wyposażona w procesor ARM, oraz mobilną edycję systemu GNU/Linux Debian,
- OpenMoko Neo Freerunner, telefon wyposażony w procesor ARM, oraz mobilną wersję GNU/Linux.

Wyżej wymienione systemy stanowią zdecydowaną większość stosowanych systemów operacyjnych na komputerach stacjonarnych oraz przenośnych.

3.2 Opis technologi użytych przy tworzeniu oprogramowania

3.2.1 Język programowanie oraz użyte narzędzia

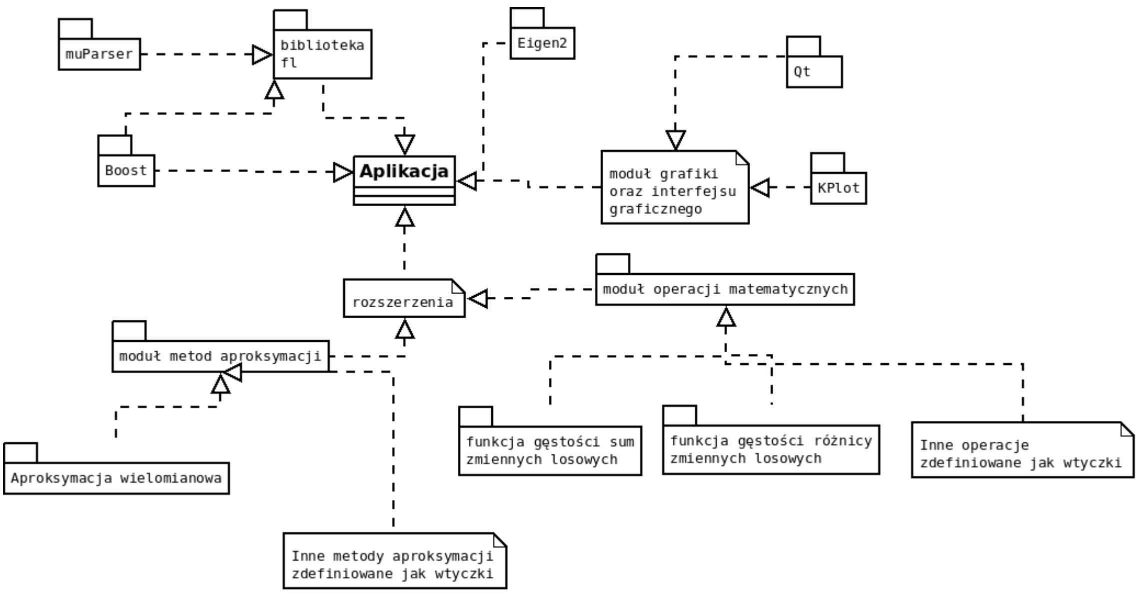
Aplikacja została napisana w jedynie w języku C++, bez wspomagania się innymi dostępnymi językami. Do napisania oprogramowania posłużyły autorowi następujące narzędzie programistyczne:

- MSVC - IDE na platformę Microsoft Windows z wbudowanym kompilatorem,
- KDevelop/4.0 - IDE na platformę GNU/Linux,

- Qt-Creator®, IDE na platformy GNU/Linux, Microsoft Windows oraz Mac OSX,
- CMake - wieloplatformowy system budowy firmy Kitware, generujący Makefile odpowiednie dla danej platformy,
- Git - rozproszony system kontroli wersji,
- Gdb - debugger,
- Vim - Vi Improved, wieloplatformowy edytor tekstowy,
- Umbrello - edytor diagramów UML.

3.3 Struktura programu

Program do pracy magisterskie od początku projektowania, pomyślany został jako modularyzowany. Oznacza to, że istnieje rdzeń aplikacji, a do niego dołączane są na etapie kompilacji oraz linkowania moduły, zarówno zewnętrzne jak i wewnętrzne. Tym samym aplikacja jest rozszerzalna i funkcjonalność może być poszerzana przez dowolnego programistę, bez modyfikowania źródeł aplikacji (poprzez dopisywania wtyczek, bibliotek dynamicznych). Na rysunku 3.1 przedstawiono poglądowy schemat struktury aplikacji, wraz z zaznaczeniem ważniejszych modułów w niej użytych.



Rysunek 3.1: Diagram przedstawiający poglądowych schemat struktury aplikacji

3.3.1 Opis modułów użytych w programie

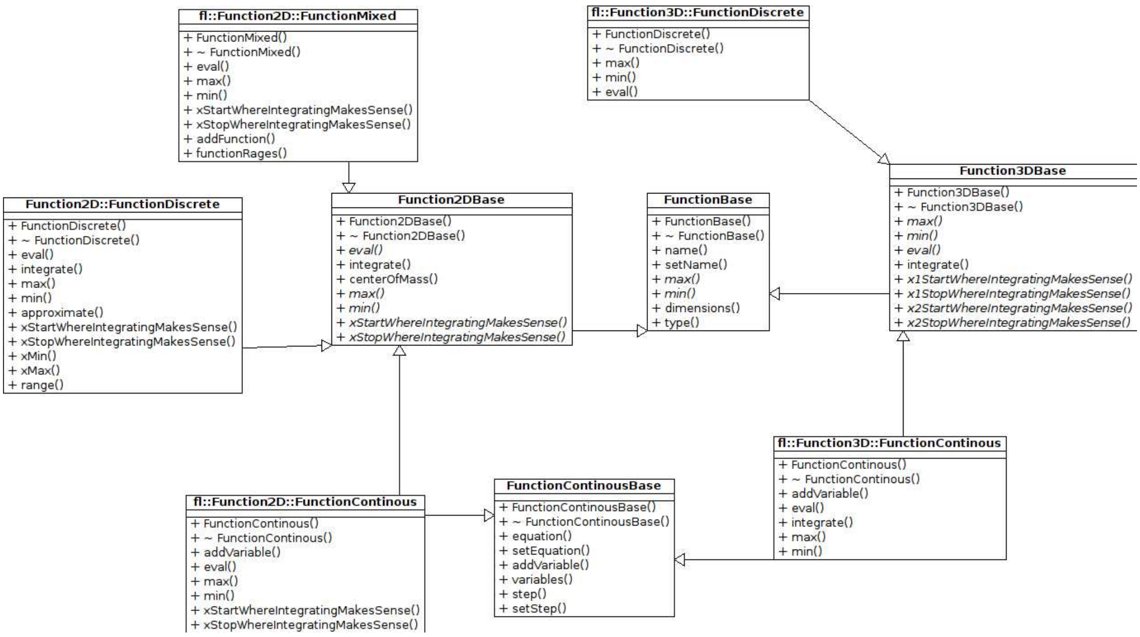
Moduły z których składa się aplikacja, dzielimy na dwie kategorie:

- moduły wewnętrzne - całkowicie napisane przez autora,
- moduły zewnętrzne - moduły napisane przez innych programistów, niektóre zawierają zmiany autora, nie naruszające wymogów licencyjnych.

3.3.2 Opis modułów wewnętrznych

3.3.2.1 Szczegółowy opis biblioteki fl

Moduł **fl** jest biblioteką, która umożliwia programiście operowanie na funkcjach matematycznych. Poprzez operowanie rozumiemy, tworzenie funkcji dyskretnych, ciągłych, mieszanych, liczenie wartości w punkcie, całkowanie numeryczne oraz numeryczne liczenie pochodnych. Na rysunku 3.2 widzimy diagram klas dla tego modułu. Centralnymi



Rysunek 3.2: Diagram klas dla modułu fl

klasami są abstrakcyjne klasy *FunctionBase* o sygnaturze[1], przedstawionej na listingu 3.1 oraz *FunctionContinousBase* o sygnaturze przedstawionej na listingu 3.2[2]. Klasa

[1]Komentarze oraz opis poszczególnych funkcji zostają usunięte z listingów, chyba że komentarz objaśnia mechanizmy nie zawarte w dokumentacji

[2]W poniższym opisie pominięto przestrzeń nazw *fl*, wszystkie niżej przedstawione klasy znajdują się w tej przestrzeni

FunctionBase jest bardzo prosta, posiada pola: nazwa funkcji, jej typ oraz wymiar (czy funkcja jest 2 czy 3 wymiarowa). Oprócz metod - akcesorów do tych pól, oraz konstruktora i wirtualnego konstruktora, widzimy abstrakcyjne metody max() i min(). Każda klasa dziedzicząca po tej klasie, musi implementować te metody (chyba, że ta klasa jest również abstrakcyjna). Klasa *FunctionContinousBase* jest typem podstawowym dla

```
class FunctionBase
{
public:
    enum Type{ // function type
        eDiscrete = 0,
        eContinous,
        eMixed
    } ;
    FunctionBase( const std::string & _functionName ) :
    m_FunctionName(_functionName) {} // default constuctor
    virtual ~FunctionBase() {}
    const std::string & name() const { return m_FunctionName ; }
    void setName ( const std::string & _name ) { m_FunctionName = _name ; }

    virtual double max() const = 0  ;
    virtual double min() const = 0 ;
    virtual int dimensions() { return m_iDimension ; }
    virtual Type type() const { return m_Type ; }
protected:
    std::string m_FunctionName ;
    int m_iDimension ;
    Type m_Type ;
};
```

Listing 3.1: Sygnatura klasy FunctionBase

każdej funkcji, która będzie funkcją ciągłą. Widać, że klasa ta posiada wskaźnik do parsera matematycznego oraz napis będący równaniem funkcji , co nie jest konieczne, gdy mówimy o funkcjach dyskretnych, lub o funkcjach mieszanych [3]
Kolejnym przedstawionym typem jest typ *Function2D::Function2DBase*, którego sygnatura widnieje na listingu 3.3. Typ ten jest podstawą dla każdej funkcji dwu wymiarowej [4]. W klasie *Function2D::Function2DBase* widzimy już operacje wylicznie wartości funkcji w punkcie, całkowania na zadanym przedziale z określonym krokiem czy liczenie środka masy funkcji Analogicznym typem jest klasa *Function3D::Function3DBase*, której sygnatura widnieje na listingu 3.3, będąca podstawą każdej trójwymiarowej funkcji. Od tej pory skupimy się jednak na funkcjach 2 wymiarowych, gdyż każdy typ dla funkcji trójwymiarowej jest analogiczny do typu dla funkcji dwuwymiarowej. [5]

[3] Funkcja mieszana jest traktowane przez bibliotekę jako lista wskaźników do funkcji ciągłych, jednak sama funkcja mieszana nie ma wzoru, ani podpiętego parsera

[4]biblioteka fl operuje tylko na funkcja dwu i trój wymiarowych, dla rozwiązania zadań z dziedziny statystyki, nie ma sensu operawać na funkcja o wyższych stopniach

[5]oczywiście implementacja różni się dla obu typów funkcji, jednak w pracy magisterskiej autor nie skupia się na szczegółach implementacyjnych

```
class FunctionContinousBase
{
public:
    FunctionContinousBase( const std::string & _equation ) :
    m_pParser( mu::Parser::proxyFLParser() ) ,
    m_bMinMaxEval(false)
    {
        ...
    }
    virtual ~FunctionContinousBase() {}
    const std::string & equation() const { return m_functionEquation;}
    void setEquation( const std::string & _equation )
    {
        ...
    }
    virtual void addVariable ( const std::string & varName )
    {
        ...
    }
    std::vector<char> variables() const ;
    double step() const { return m_dStep ; }
    void setStep( double _step ) { m_dStep = _step ; }
protected:
    std::string m_functionEquation ;
    std::auto_ptr<mu::Parser> m_pParser ; // parser matematyczny
    double m_dStep ;
    mutable std::map<std::string,double> m_VariableMap ;
    mutable double m_iMin ;
    mutable double m_iMax ;
    mutable bool m_bMinMaxEval ;
};
```

Listing 3.2: Sygnatura klasy FunctionContinousBase

```
class Function2DBase : public fl::FunctionBase
{
public:
    Function2DBase( const std::string & _functionName ) ;
    virtual ~Function2DBase() {}
    virtual double eval( double point , bool * pCorrect ) const = 0 ;
    virtual double integrate( double start , double stop , double dStep ) const;
    virtual double centerOfMass( double start , double stop , double step ) const ;
    virtual double max() const = 0 ;
    virtual double min() const = 0 ;
    virtual double xStartWhereIntegratingMakesSense() const = 0 ;
    virtual double xStopWhereIntegratingMakesSense() const = 0 ;
};
```

Listing 3.3: Sygnatura klasy Function2D::Function2DBase

Powyżej przedstawione typy są typami abstrakcyjnymi, co oznacza, że nie można stworzyć obiektów tych typów, są one tylko podstawą do typów z nich dziedziczonych i ujednolicają interfejs dla nich.
Nie abstrakcyjne typy danych, które mogą zostać stworzone to:

- Function2D::FunctionDiscrete,
- Function2D::FunctionContinous,
- Function2D::FunctionMixed,

oraz analogicznie:

- Function3D::FunctionDiscrete,
- Function3D::FunctionContinous,
- Function3D::FunctionMixed.

Typ *Function2D::FunctionDiscrete* dziedziczy bezspośrednio z typu *Function2D::Function2DBase*, a pośrednio z *FunctionBase* i jego interfejs nie wprowadza żadnych nowy operacji na tym typie, wszystkie operacje są dziedziczone z typów podstawowych. Klasa *Function2D::FunctionContinous* dziedziczy bezpośrednio z *Function2D::Function2DBase* oraz *FunctionContinousBase* oraz pośrednio z *FunctionBase* i również cały jej interfejs został zawarty w klasach bazowych. Obie te klasy, to jest *Function2D::FunctionDiscrete* oraz *Function2D::FunctionContinous*, dostarczają tylko implementacji swoich interfejsów. Klasa *Function2D::FunctionMixed* wprowadza natomiast nowe operacje oraz typy danych, przedstawione na listingu 3.4 , których trzeba używać, gdy chcemy tworzyć funkcje o zmiennych równaniach, na określonych przedziałach. Klasa ta jest pojemnikiem na obiekty będące typu *Function2D::FunctionContinous*. Dodatkowo w klasie zdefiniowano typ określający rodzaje operatorów porównań[6] oraz typ reprezentujący szczątkową funkcję na zadanym przedziale [7].
Na listingu 3.5, przedstawiono użycie klas *Function2D::FunctionDiscrete*, *Function2D::FunctionContinous* oraz *Function2D::FunctionMixed* w przykładowym zastosowaniu.

3.3.2.2 Opis modułu operacji matematycznych

Moduł operacji matematycznych, jest modułem do którego oddajemy sterowanie, gdy chcemy wykonać jakąś funkcję matematyczną na jednej, dwóch lub więcej funkcjach.

[6] *Function2D::FunctionMixed::Operator*
[7] *Function2D::FunctionMixed::FunctionRange*

```
class FunctionMixed : public fl::Function2D::Function2DBase
{
public:
    enum Operator{
        eGreater=0,
        eGreaterEqual,
        eLess,
        eLessEqual,
        eUnknown
    };
    struct FunctionRange{
        boost::shared_ptr< fl::Function2D::FunctionContinous> m_spFunction ;
        double m_start ;
        double m_stop ;
        Operator m_operatorStart ;
        Operator m_operatorStop ;
    };
public:
    FunctionMixed(const std::string & _functionName );
    virtual ~FunctionMixed(){}
    virtual double eval( double point , bool *pOk ) const ;
    virtual double max() const ;
    virtual double min() const ;
    virtual double xStartWhereIntegratingMakesSense() const ;
    virtual double xStopWhereIntegratingMakesSense() const ;
    void addFunction(  fl::Function2D::FunctionContinous * pFunction,
                       double start,
                       Operator startOperator,
                       double stop,
                       Operator stopOperator);
    const std::vector<FunctionRange> & functionRages() const { return m_Functions ; }
private:
    mutable std::vector<FunctionRange> m_Functions ;
private:
    bool isHere( double point, const FunctionRange & funRange) const ;
};
```

Listing 3.4: Sygnatura klasy Function2D::FunctionMixed

Podstawą tego modułu jest interfejs *IOperation*, przedstawiony na listingu 3.6. Interfejs posiada metody, które pozwalają ustawić funkcje dla danej operacji *IOperation::addFunction*, oraz wykonać operację *IOperation::calculate*. Typ zwracany przez metodę *IOperation::calculate*, jest wskaźnikiem na typ bazowym każdej funkcji z przestrzeni fl. Oba dostać odpowiednią funkcję, należy sprawdzić jej wymiar oraz rodzaj i następnie wykonać rzutowanie do interesującego nas typu. Autor dostarcza cztery operacje matematyczne, to jest:

- operacja $\oplus$ na funkcjach gęstości prawdopodobieństwa,
- operacja $\ominus$ na funkcjach gęstości prawdopodobieństwa,
- operacja $\otimes$ na funkcjach gęstości prawdopodobieństwa,
- operacja $\oslash$ na funkcjach gęstości prawdopodobieństwa.

```
#include <iostream>
#include <boost/assign/std/vector.hpp>
#include <functionException.h>
#include "functiondiscrete.h"
#include "functioncontinous.h"
#include "functionMixed.h"
#include <limits>
using namespace boost::assign; // dla szybkiego przypisania wartości do wektora
int main(int argc, char **argv) {
    try{
        std::vector<double> xs ;
        std::vector<double> ys ;
        xs += 1,2,3,4,5,6 ;
        ys += -40,77,1,1,2,2 ;

        fl::Function2D::FunctionDiscrete f1(xs,ys,"example");
        std::cout << f1.eval(3) << std::endl ;
        std::cout << f1.max() << std::endl ;
        std::cout << f1.min() << std::endl ;

        fl::Function2D::FunctionContinous fC ;
        fC.setEquation("sin(x)");
        fC.addVariable("x");
        std::cout << fC.eval(0);

        fl::Function2D::FunctionContinous fC2 ;
        fC2.setEquation("cos(x)") ;
        fC2.addVariable("x");

        fl::Function2D::FunctionMixed fM ;
        fM.addFunction(fC,
                       -1 * std::numeric_limits<double>::infinity(),
                       fl::Function2D::FunctionMixed::eGreater,
                       0,
                       fl::Function2D::FunctionMixed::eGreaterEqual,
                       );
        fM.addFunction(fC2,
                       0,
                       fl::Function2D::FunctionMixed::eGreater,
                       std::numeric_limits<double>::infinity(),
                       fl::Function2D::FunctionMixed::eGreater,
                       );
        std::cout << fM.eval(-13);
        std::cout << fM.eval(0);
        std::cout << fM.eval(1);
    }
    catch ( fl::FunctionException & e )
    {
        const char *ww = e.what() ;
        std::cout << ww ;
    }
    return 0;
}
```

Listing 3.5: Przykłady użycia typów, funkcji dwuwymiarowych, z przestrzeni fl

```
class IOperation
{
public:
    enum Errors{
        eUndefined=0,
        eSuccess,
        eNotIntegratingToOne
    };
    typedef boost::shared_ptr<const fl::FunctionBase> FunctionBaseShPtr ;
    IOperation() ;
    virtual ~IOperation() {}
    virtual void addFunction( const fl::FunctionBase* pPtr ) ;
    virtual fl::FunctionBase * calculate() = 0 ;
    Errors error () const { return m_error ; }
    virtual std::string operation() const =0;
protected:
    std::vector<FunctionBaseShPtr> m_functions ;
    mutable Errors m_error ;
    std::string m_operation ;
};
```

Listing 3.6: Podstawowy interfejs dla operacji matematycznych *IOperation*

Aby samemu rozszerzyć aplikację o kolejną metodę, należy wydziedziczyć interfejs *IOperation* oraz przedstawić implementację własnej operacji.

3.3.2.3 Opis modułu metod aproksymacji

Moduł dotyczący aproksymacji funkcji jest bardzo podobny do modułu operacji matematycznych, również posiada jeden interfejs, *IApproximation* przedstawiony na listingu 3.7. Autor dostarcza metodę aproksymacji wielomianowej jako rozszerzenie aplikacji. Możliwe jest oczywiście dopisanie innych metod aproksymacji, poprzez wydziedziczenie interfejsu *IApproximation* oraz przedstawienie implementacji.

```
class IApproximation
{
    public:
        IApproximation(const Function2D::FunctionDiscrete::DomainRange & _range) ;
        virtual ~IApproximation() ;
        virtual Function2D::FunctionContinous * approximate() const  = 0 ;
    protected:
        Function2D::FunctionDiscrete::DomainRange m_range ;
};
```

Listing 3.7: Podstawowy interfejs dla metod aproksymacji

3.3.2.4 Opis części graficznej

Aplikacji posiada interfejs graficzny okienkowy , który został stworzony z wykorzystaniem bibliotek Qt. Za wyświetlanie głównego okienka odpowiada klasa *UI::MainWindow*. Klasa ta zarządza nie tylko oknem, ale również panelami wewnątrz niego, paskiem stanu, paskiem menu , oraz gdy trzeba konstruuje okna dialogowe w celu zbierania informacji od użytkownika (wczytanie funkcji z pliku, stworzenie nowej funkcji, aproksymacja funkcji itd). Główny interfejs programu jest typu TDI (ang. Tabbed Document Interface), co oznacza, że użytkownik może tworzyć kilka przesterzni roboczych wewnątrz jednej aplikacji. Przestrzeń robocza, to część programu w której użytkowniki widzi wprowadzone przez siebie funkcje i ma możliwości ich modyfikacji. Interfejs TDI pozwala użytkownikowi uruchomienie tylko jednej instancji aplikacji, przy posiadaniu wielu różnych przestrzeni roboczych, a na każdej różną ilość różnych funkcji.

3.3.3 Opis modułów zewnętrznych

W programie zostały użyte biblioteki zewnętrzne do grafiki, parsowania wyrażeń matematycznych oraz do operowania na macierzach oraz wektorach danych. Autor nie miał intencji pisania lepszych rozwiązań w tych dziedzinach, gdyż obecne dostępne są lepsze szybsze, dokładniejsze oraz przede wszystkim są dokładnie przetestowane przez społeczność programistów. Niektóre zewnętrze biblioteki zostały przez autora zmodyfikowane, aby dostarczyć funkcjonalność której jeszcze nie miał. Wszystkie zewnętrzne źródła zostały wydane na wolnościowych licencjach (GNU GPL lub GNU LGPL), tak że ich licencje w żadnym punkcie nie zostały złamane.
Użyte zostały następujące biblioteki:

- **Qt**[8] w wersji 4.5.3 - wieloplatformowy modularna biblioteka odpowiedzialna za graficzną część programu. To ona wykonuje wszystkie operacje zarządzania okienkami, tworzy zasoby okien, oraz zajmuje się integracją z systemem operacyjnym, tak by aplikacja wyglądała w każdym środowisku w sposób naturalny dla danego środowiska,
- **Boost**[9] w wersji 1.38.0 - wieloplatformowa biblioteka do języka C++, która rozszerza język o dodatkową funkcjonalność. W aplikacji zostały użyte: wskaźniki ze zliczaniem referencji, kontenery na wskaźniki do funkcji, dodatkowe funkcje matematyczne, tablice o statycznym rozmiarze oraz typ danych konwertowalny na każdy dowolny typ,

[8] http://qt.nokia.com/products
[9] http://www.boost.org/

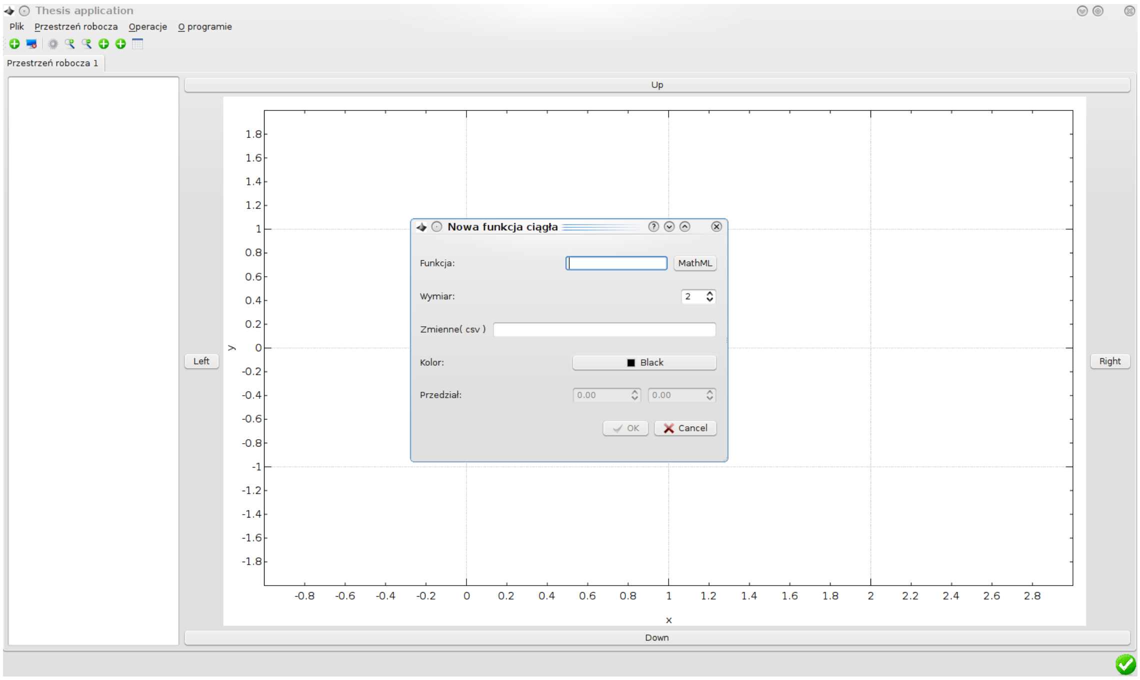
- **muParser**[10] - parser wyrażeń matematycznych wykorzystujący algorytm odwrotnej notacji polskiej. Naniesione zostały na niego drobne zmiany przez autora, aby parser domyślnie używał funkcji matematycznych ze standardu C++98 oraz z biblioteki **Boost::Math**,
- **KPlot** - biblioteka ze środowiska KDE4 przeniesiona do środowiska **Qt**, tak aby była kompilowana w środowiska gdzie nie ma dostępnego KDE4. Naniesione modyfikacje dotyczą również sposobu wyświetlania funkcji,
- **Eigen2**[11] - biblioteka szablonowa używana przy operacjach macierzowych.

3.4 Przykładowe użycie

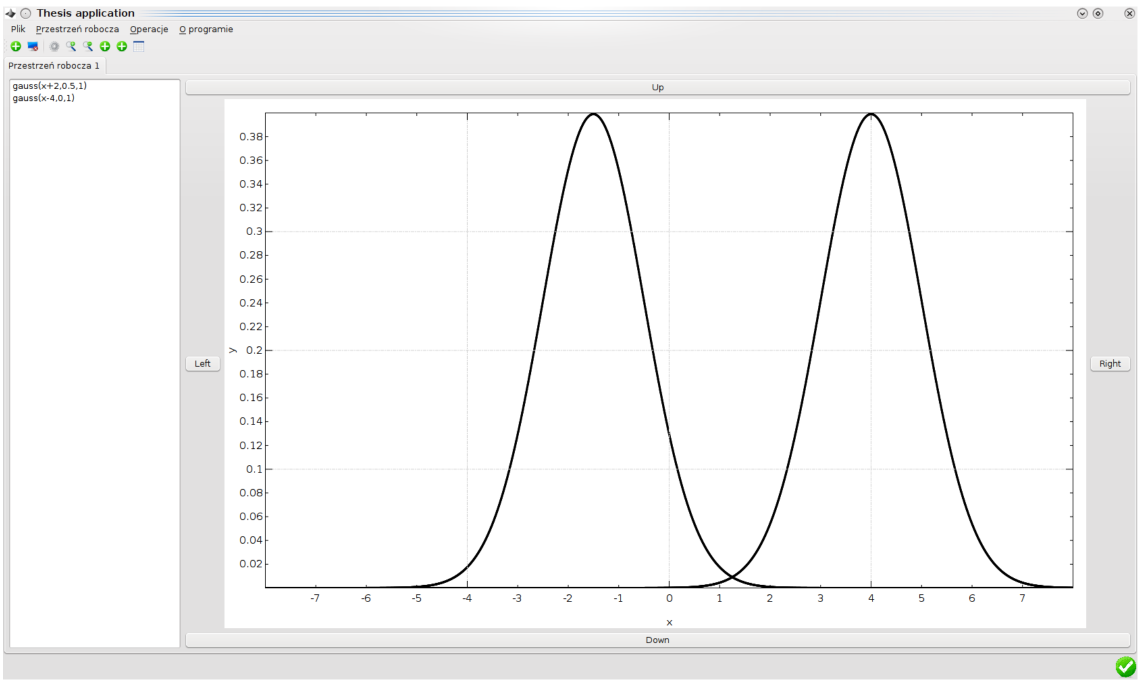
Autor pokaże jak z pomocą programu wykonać prostą operację $\otimes$ między dwoma rozkładami funkcji gęstości prawdopodobieństwa.

Na początek tworzymy nową przestrzeń roboczą wciskając kombinację klawiszy *Ctrl+n*, lub wybierają z menu *Plik* opcje *Nowa przestrzeń robocza*. Następnie, wybieramy z menu Przestrzeń robocza opcje *Nowa funkcja*, a następnie *Nowa funkcja ciągła*. Dialog, który zostanie wyświetlony został przedstawiony na rysunku 3.3. W polu *Funkcja* wpisujemy 'gauss(x-4,0,1)', gdy chcemy operować na rozkładzie normalnym (Gaussowskim), lub inną funkcję będącą funkcją gęstości (całkowalną do jedności). W polu *Zmienne(csv)* wpisujemy zmienne oddzielone przecinkiem, dla funkcji dwuwymiarowej wpisujemy 'x' po czym klikamy 'Ok'. Do przestrzeni roboczej zostanie dodany wykres przedstawiający funkcję gęstości dla rozkładu normalnego. Następnie dodajemy kolejną funkcję, w analogiczny sposób, tym razem określoną wzorem 'gauss(x+2,0.5,1)'. Po przeskalowaniu, w naszej przestrzeni roboczej uzyskujemy wykresy gęstości dla tych dwóch rozkładów, co przedstawiono na rysunku 3.4. Następnie wybieramy pozycję 'Dodaj' z menu 'Operacje' i podmenu 'Operacje na rozkładach gęstości'. Pojawi się dialog przedstawiony na rysunki 3.5. Wybieramy wprowadzone funkcje i w przestrzeni roboczej pojawi się dyskretna funkcja będąca wynikiem tej operacji na zaznaczonych funkcjach. Aby uzyskać końcową funkcję rozkładu gęstości należy wybrać z menu 'Operacje' pozycję aproksymacja (dialog przedstawiony na rysunku 3.6). Z dialogu należy wybrać: funkcję do aproksymacji oraz stopień aproksymacji. Po zaakceptowaniu dialogu w przestrzeni roboczej pojawi się wynikowa funkcja. Klikając w jej nazwę po lewej stronie prawym przyciskiem myszy , oraz wybierając pozycję 'Informacje' z menu kontekstowego można odczytać wykres wynikowej funkcji, co pokazano na rysunku 3.7.

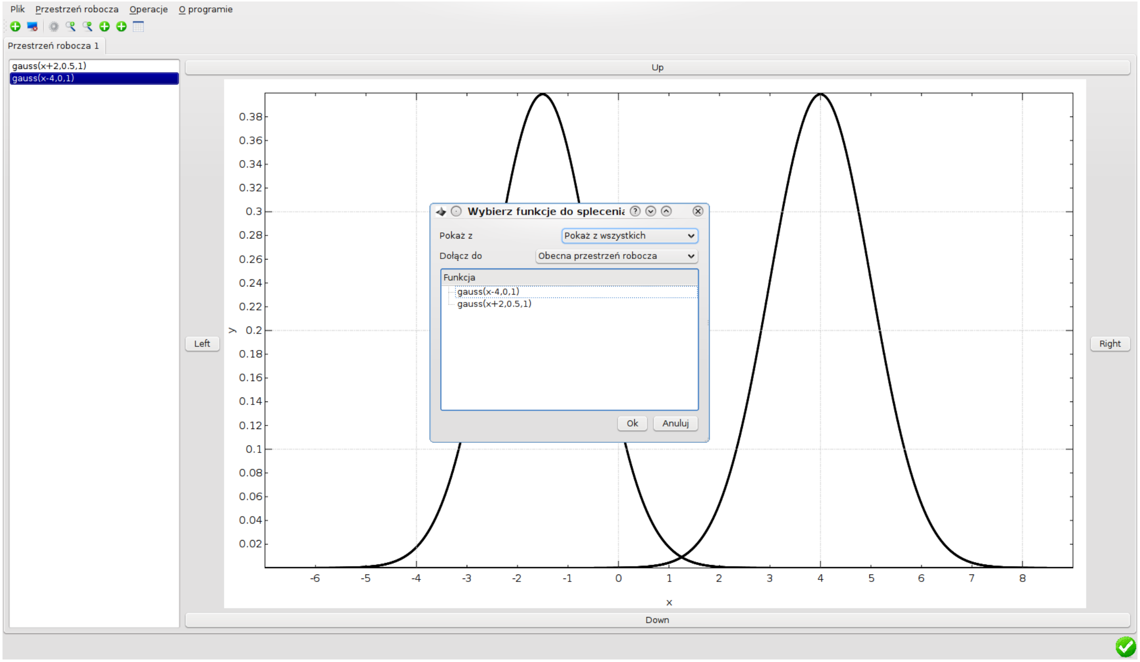
[10] http://muparser.sourceforge.net/
[11] http://eigen.tuxfamily.org/



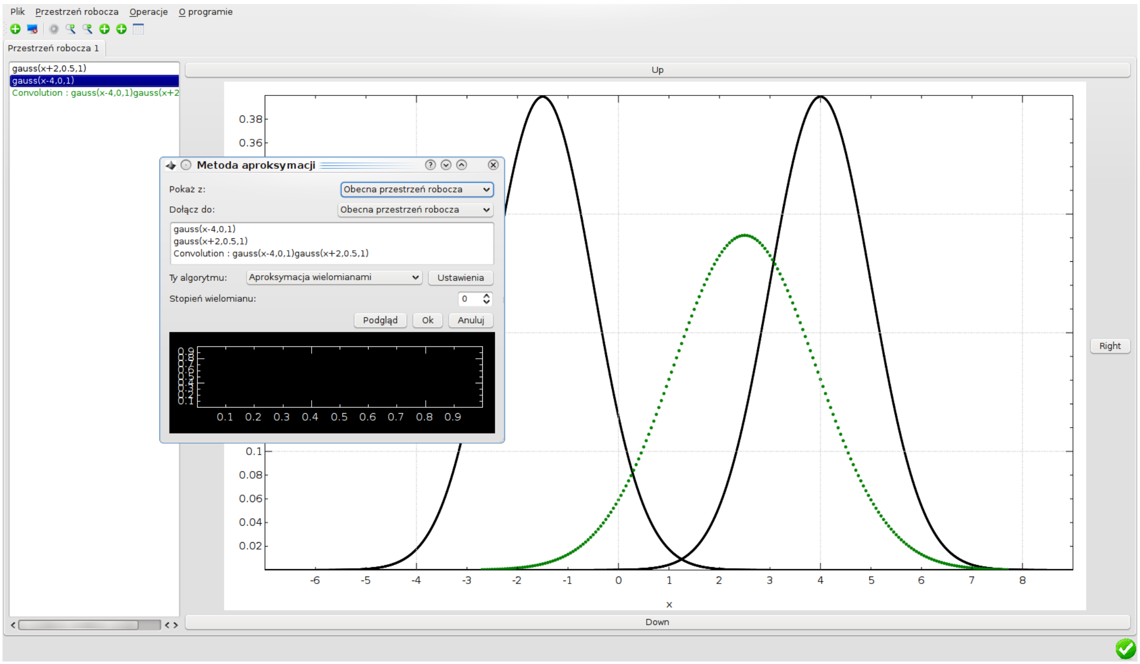
Rysunek 3.3: Rysunek przedstawiający dialog *Nowa funkcja ciągła*



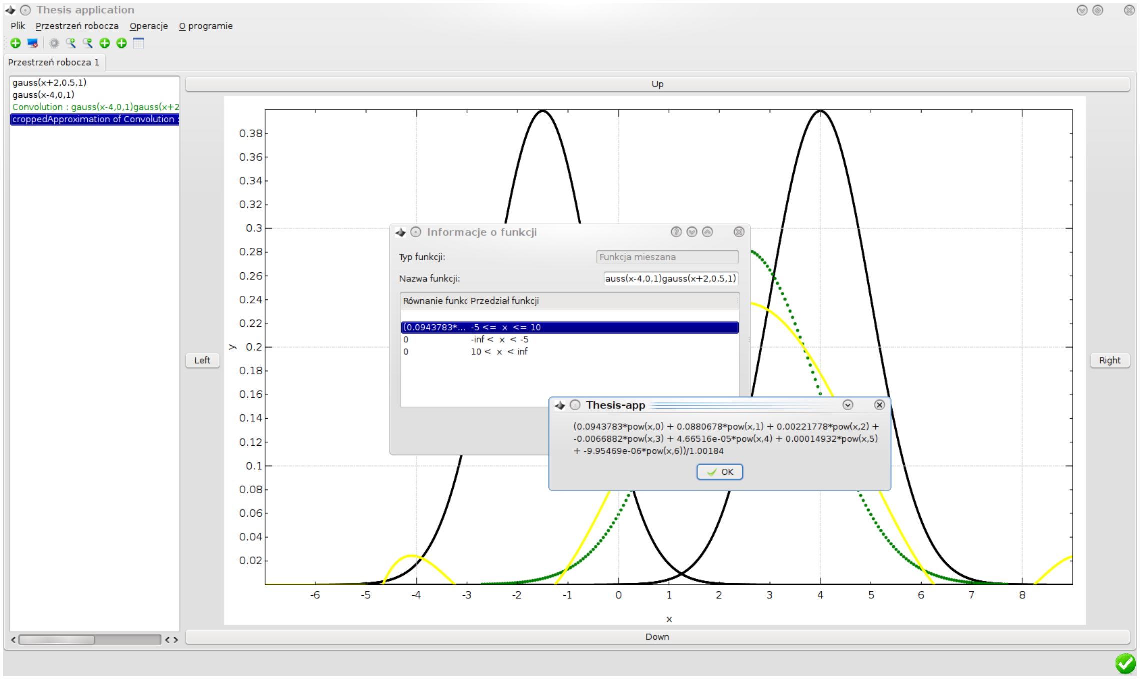
Rysunek 3.4: Wykresy funkcji gęstości dla funkcji *gauss(x-4,0,1)* oraz *gauss(x+2,0.5,1)*



Rysunek 3.5: Dialog wyboru funkcji do przeprowadzenia operacji na nich



Rysunek 3.6: Dialog wyboru funkcji do aproksymacji



Rysunek 3.7: Dialog ukazujący szczegóły dowolnej funkcji

Rozdział 4

Rozwiązanie przykładowego zadania

4.1 Postawienie zadania

Zadaniem, które zostanie rozwiązane za pomocą programu będącego podmiotem pracy magisterskiej, jest problem "dwóch biznesów". Jest to problem decyzyjny, w którym należy podać rozwiązanie opierając się jedynie na informacjach, podanych w języku naturalnym, przedstawionych jako kwantyfikatory lingwistyczne. Problem postawiony jest następująco:

> W biznesie B_1 można zarobić od 0 do 10 milionów złotych. Prawdopodobieństwo dochodu bliskiego 10 milionów złotych jest duże. W biznesie B_2 można zarobić od 0 do 20 milionów złotych. Prawdopodobieństwo dochodu bliskiego 20 milionów złotych jest jednak małe.
> Decyzja: W który z biznesów zainwestować?

Rozwiązaniem tego problemu są wartości oczekiwane dochodów obu biznesów. Aby jednak te wartości obliczyć należy znaleźć rozkłady prawdopodobieństwa dla obu biznesów. W rozdziale 2 przedstawiłem w jaki sposób łączyć kwantyfikatora z rozkładem prawdopodobieństwa, teraz wykorzystamy ten fakt w celu rozpisania funkcji gęstości rozkładów. Zmienna losowa ciągła dla obu biznesów wygląda następująco:

$$S_N = PR_N * DB_n + NIEPR_N * NIEDB_N, \tag{4.1}$$

gdzie:

- S_N - zmienna losowa dla biznesu n-tego,
- PR_N - oznacza zmienną losową prawdopodobieństwa osiągnięcia dochodu DB_N,
- DB_N - oznacza zmienną losową określającą dochód z biznesu n-tego,
- $NIEPR_N$ - oznacza zmienną losową przeciwną do PR_N, czyli przeciwne prawdopodobieństwo osiągnięcia dochodu przeciwnego,
- $NIEDB_N$ - oznacza zmienną losową określającą przeciwny dochód z biznesu n-tego.

Operacje + i * na zmiennych losowych utożsamiamy z operacjami $\oplus$ oraz $\otimes$ na ich rozkładach.

4.2 Identyfikacja funkcji gęstości

Aby podejść do próby rozwiązania tego problemu należy określić kwantyfikatory lingwistyczne

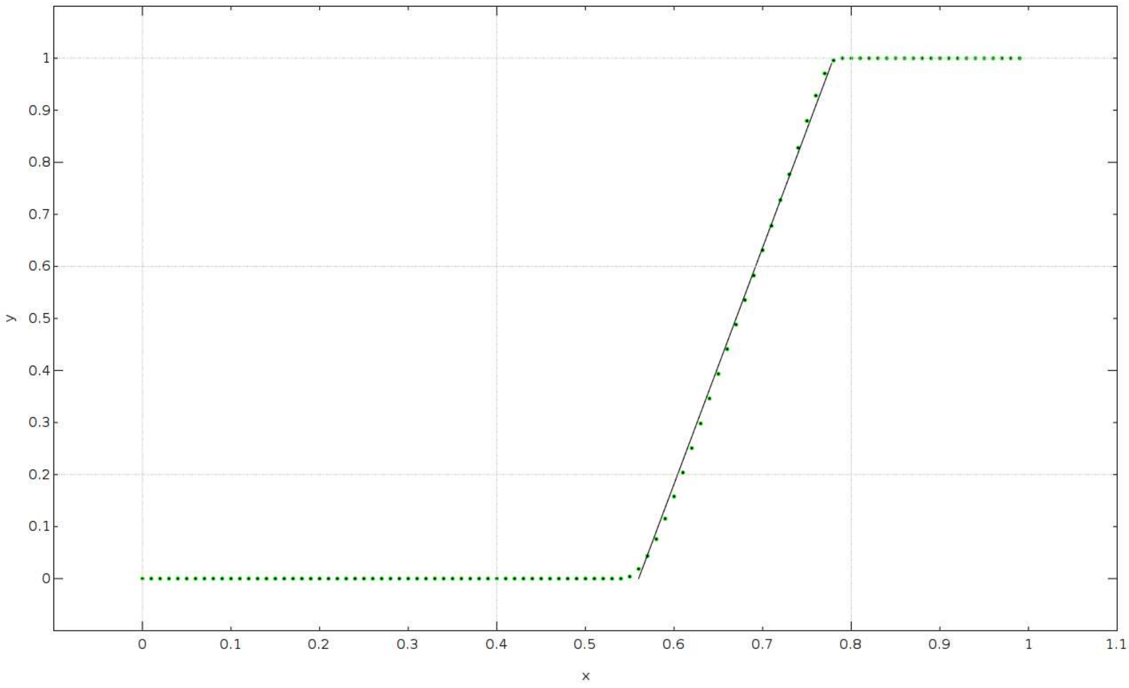
- duże,
- małe (nie duże) ,
- bliskie, dla wartości maksymalnej dochodu z biznesu pierwszego ,
- nie bliskie , dla wartości maksymalnej dochodu z biznesu pierwszego ,
- bliskie, dla wartości maksymalnej dochodu z biznesu drugiego ,
- nie bliskie , dla wartości maksymalnej dochodu z biznesu drugiego.

oraz odpowiadające im funkcje rozkładu gęstości. Kwantyfikatory *duże* i *małe*, zostaną określone na podstawie pracy doktora Marka Landowskiego [3]. Doktor po przeprowadzeniu ankiety wyłuskał dane dotyczące oznaczania danej wartości prawdopodobieństwa jako dużej, małej oraz średniej. W naszym przypadku, interesują nas tylko badania odnośnie pojmowania prawdopodobieństwa jako dużego oraz małego. Kwantyfikatory uzyskane przez doktora Marka Landowskiego posłużą nam do określenia funkcji rozkładu gęstości prawdopodobieństwa. Kwantyfikatory *bliskie 10 mln*, *nie bliskie 10 mln*, *bliskie 20 mln* oraz *nie bliskie 20 mln* są określone poprzez rozciągniecie kwantyfikatorów *małe* i *duże* na przedziały im odpowiadające. [1] Dane otrzymane przez doktora Landowskiego

[1] np. *bliskie 10 mln* na przedział $[0, 10]$

,można przybliżyć funkcjami wysokich rzędów, ale punkty układają się w sposób pozwalający przeprowadzić aproksymację funkcjami pierwszego i zerowego rzędu (funkcjami liniowymi oraz funkcjami stałymi), bez dużej straty dokładności, co pokazano na rysunkach 4.1 4.3.Aby zachować pewne ustalenie, pozostałe kwantyfikatory również będą miały postać funkcji pierwszego i zerowego rzędu. W Aby zidentyfikować poszczególne funkcje rozkładu gęstości prawdopodobieństwa należy przypatrzeć się rozkładowi punktów dla danego kwantyfikatora, oraz wykorzystując właściwość całkowania do jedności funkcji gęstości na określonym przedziale, rozwiązać układ równań, który określi nam funkcję pierwszego rzędu oraz maksymalną wysokość funkcji rozkładu prawdopodobieństwa (maksymalna wysokość kwantyfikatorów wynosi 1, jednak w naszym przypadku, wysokość trzeba przeskalować, aby uzyskać funkcję całkującą się do jedności). Cały algorytm aproksymacji zostanie przedstawiony dla kwantyfikatora *duże*, natomiast dla reszty kwantyfikatorów pominięte zostaną obliczenia i przedstawione zostaną tylko wyniki. W W poniższych obliczeniach oś rzędnych oznacza wartość prawdopodobieństwa, dla kwantyfikatorów *duże* i *małe*, oraz wartość dochodu dla kwantyfikatorów *bliskie 10 mln*, *nie bliskie 10 mln*, *bliskie 20 mln* oraz *nie bliskie 20 mln.*

4.2.1 Kwantyfikator "duże"



Rysunek 4.1: Wykres punktów zebranych dla kwantyfikatora *duże*

Mamy dany zbiór punktów, przedstawiony na wykresie 4.1 i dla tego zbioru, trzeba oszacować funkcję rozkładu gęstości prawdopodobieństwa. Aby to uczynić trzeba odczytać z wykresu punkty gdzie funkcja zaczyna zmieniać swoją postać. W punkcie [0.56,0] funkcja zaczyna rosnąć, w punkcie [0.78,1] funkcja zmienia swoją charakterystykę z liniowej na stałą. Tak więc nasza szukana funkcja będzie w przedziale 0 - 0.56 stała, w przedziale 0.56 - 0.78 przyjmie postać funkcji liniowej, a w przedziale 0.78 - 1 znowu zmienia się w funkcję stałą. Na rysunku zaznaczono również, jak dobrze funkcja liniowa pokrywa się z punktami z przedziału 0.56 - 0.78 i że jest ona wystarczająca do aproksymacji. Przy aproksymacji należy pamiętać, że o ile wartość maksymalna kwantyfikatora wynosi 1 (jakaś wartość przynależy do kwantyfikatora w stopniu maksymalnym), to zaaproksymowana na jej podstawie funkcja gęstości wartość maksymalną będzie miała przeskalowaną tak, aby mogła się całkować do jedności. Szukana funkcja będzie miała zatem postać:

$$f(x) = \begin{cases} ax + b, & \text{dla } 0.56 < x \leqslant 0.78 \\ c, & \text{dla } 0.78 < x < 1. \end{cases} \tag{4.2}$$

Aby znaleźć stałe a, b oraz c rozwiązujemy układ równań:

$$\begin{cases} (1 - 0.78) * c + \frac{1}{2} * (0.78 - 0.56) * c = 1 \\ 0.56 * a + b = 0 \\ 0.78 * a + b = c \end{cases} \Rightarrow \begin{cases} 0 * a + 0 * b + 0.33 * c = 1 \\ 0.56 * a + 1 * b + 0 * c = 0 \\ 0.78 * a + 1 * b - 1 * c = 0. \end{cases} \tag{4.3}$$

Pierwsze równanie z układu 4.3 wynika z faktu, że funkcja rozkładu gęstości prawdopodobieństwa musi się całkować do jedności na całym przedziale. (Poza przedziałem 0.56 - 1 funkcja jest nieokreślona, co oznacza, że pole pod wykresem poza tym przedziałem jest równe 0 i nie musi być brane pod uwagę) Kolejne dwa równania z układu 4.3 wynikają z rozpięcia funkcji liniowej między punktami (0.56 , 0) oraz (0.78 ,c). Trzecie Układ równań, przekształcamy do równania macierzowego:

$$\begin{bmatrix} 0 & 0 & 0.33 \\ 0.56 & 1 & 0 \\ 0.78 & 1 & -1 \end{bmatrix} * \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \tag{4.4}$$

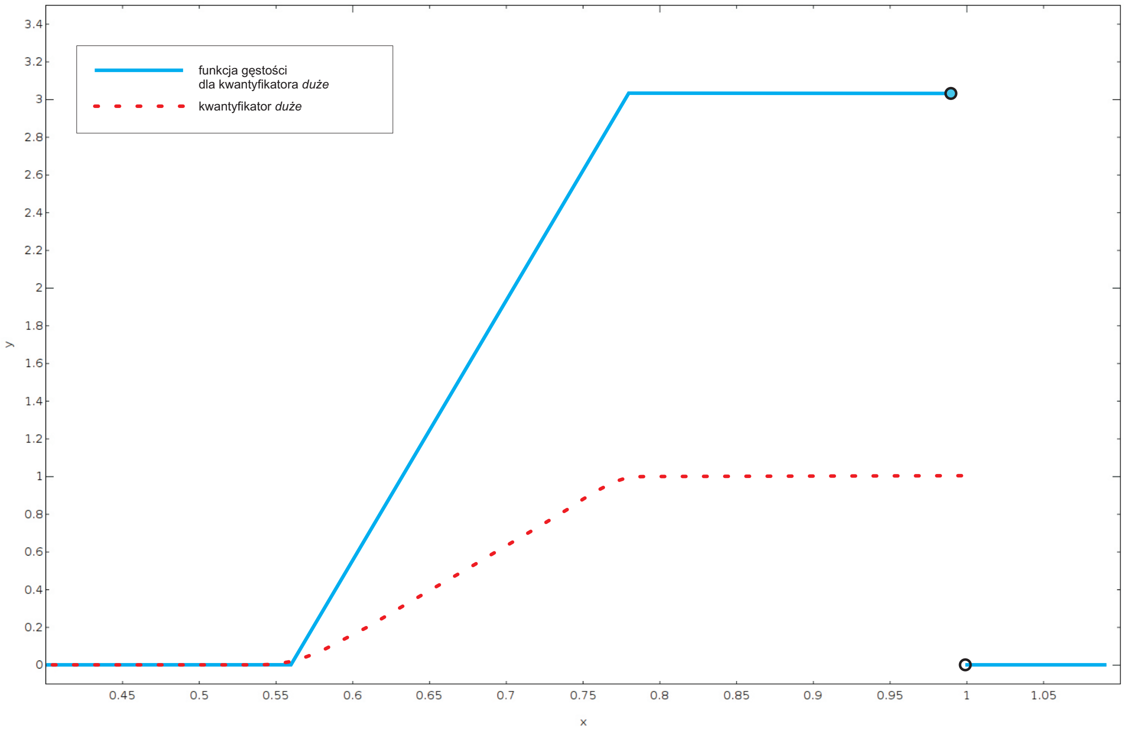
co po przekształceniu daje równanie:

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0.33 \\ 0.56 & 1 & 0 \\ 0.78 & 1 & -1 \end{bmatrix}^{-1} * \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}. \tag{4.5}$$

Rozwiązanie tego równania macierzowego wynosi:

$$a = 13.774, b = -7.713c = 3.030. \tag{4.6}$$

Funkcja rozkładu gęstości prawdopodobieństwa, dla kwantyfikatora *duże*, pokazana na



Rysunek 4.2: Wykres kwantyfikatora *duże* wraz z funkcją rozkładu gęstości prawdopodobieństwa

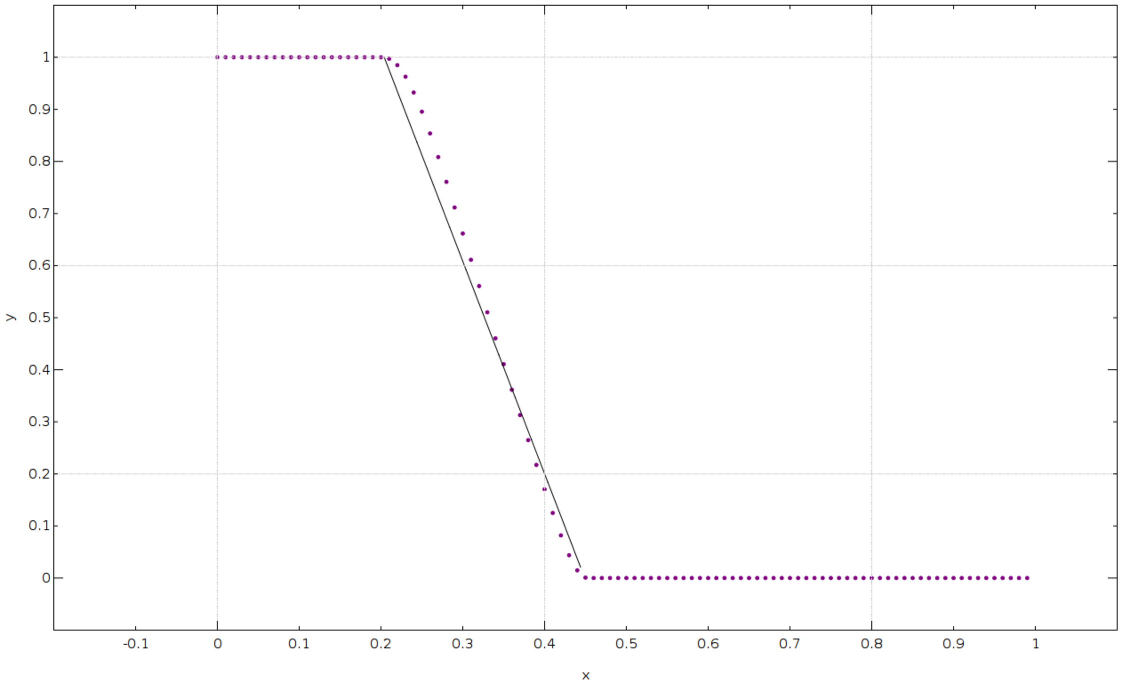
rysunku 4.2 wraz z kwantyfikatorem, ma następujący wzór:

$$f(x) = \begin{cases} 10.331 * x - 5.785, & \text{dla } 0.56 < x \leqslant 0.78 \\ 2.273, & \text{dla } 0.78 < x <= 1. \end{cases} \tag{4.7}$$

4.2.2 Kwantyfikator "małe"

Mamy dany zbiór punktów, przedstawiony na wykresie 4.3 i analogicznie trzeba oszacować funkcję rozkładu gęstości prawdopodobieństwa. Odczytujemy z wykresu punkty gdzie funkcja zaczyna zmieniać swoją postać. W punkcie [0.204,0] funkcja zaczyna maleć, w punkcie [0.454,1] funkcja zmienia swoją charakterystykę z liniowej na stałą. Tak więc nasza szukana funkcja będzie w przedziale 0 - 0.204 stała, w przedziale 0. - 0.454 przyjmie postać funkcji liniowej, a w przedziale 0.454 - 1 znowu zmienia się w funkcję stałą. Aproksymowana funkcja będzie postaci podobnej do funkcji gęstości dla kwantyfikatora *duże* opisanej wzorem 4.2 :

$$f(x) = \begin{cases} ax + b, & \text{dla } 0.204 < x \leqslant 0.454 \\ c, & \text{dla } 0 < x < 0.204. \end{cases} \tag{4.8}$$



RYSUNEK 4.3: Wykres punktów zebranych dla kwantyfikatora *małe*

Aby obliczyć a, b oraz c rozwiązujemy układ analogiczny do układu równań 4.3. Końcowe operacja jest analogiczna do mnożenia macierzy 4.5, i wygląda następująco:

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0.33 \\ 0.455 & 1 & 0 \\ 0.204 & 1 & -1 \end{bmatrix}^{-1} * \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \tag{4.9}$$

a rozwiązaniem tego układu jest trójka:

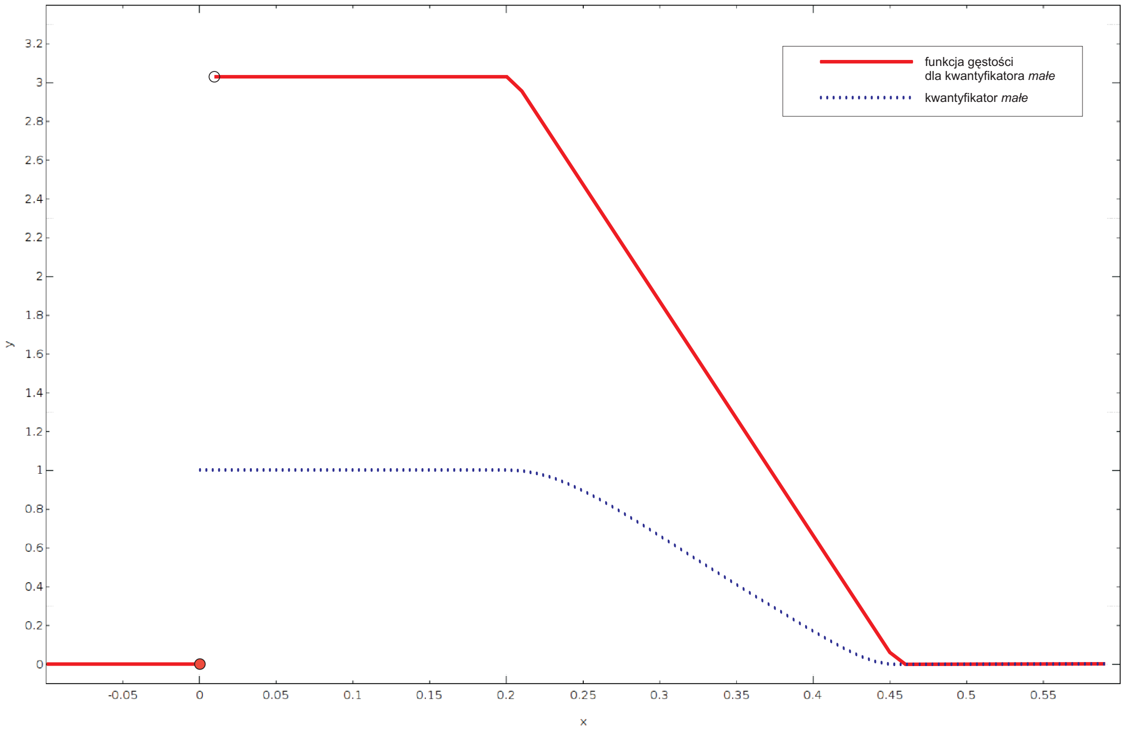
$$a = -8.775, b = 3.993, c = 2.203. \tag{4.10}$$

Funkcja rozkładu gęstości prawdopodobieństwa, dla kwantyfikatora *duże*, pokazana na rysunku 4.4 wraz z kwantyfikatorem, ma następujący wzór:

$$f(x) = \begin{cases} -8.775 * x + 3.993, & \text{dla } 0.204 < x \leqslant 0.454 \\ 2.203, & \text{dla } 0 < x < 0.204. \end{cases} \tag{4.11}$$

4.2.3 Kwantyfikator "bliskie 10 mln"

Kwantyfikator *bliskie 10 mln* jest zadany przez wykres funkcji 4.5. Tym razem, kwantyfikator ten nie jest wynikiem ankiety, tylko jest zadany przez Profesora Piegata(tak



Rysunek 4.4: Wykres kwantyfikatora *małe* wraz z funkcją rozkładu gęstości prawdopodobieństwa

samo jak dla reszty kwantyfikatorów). Aby uzyskać wzór funkcji dla tego i reszty kwantyfikatorów, rozwiązujemy układ równań analogiczny do układu 4.3 o postaci:

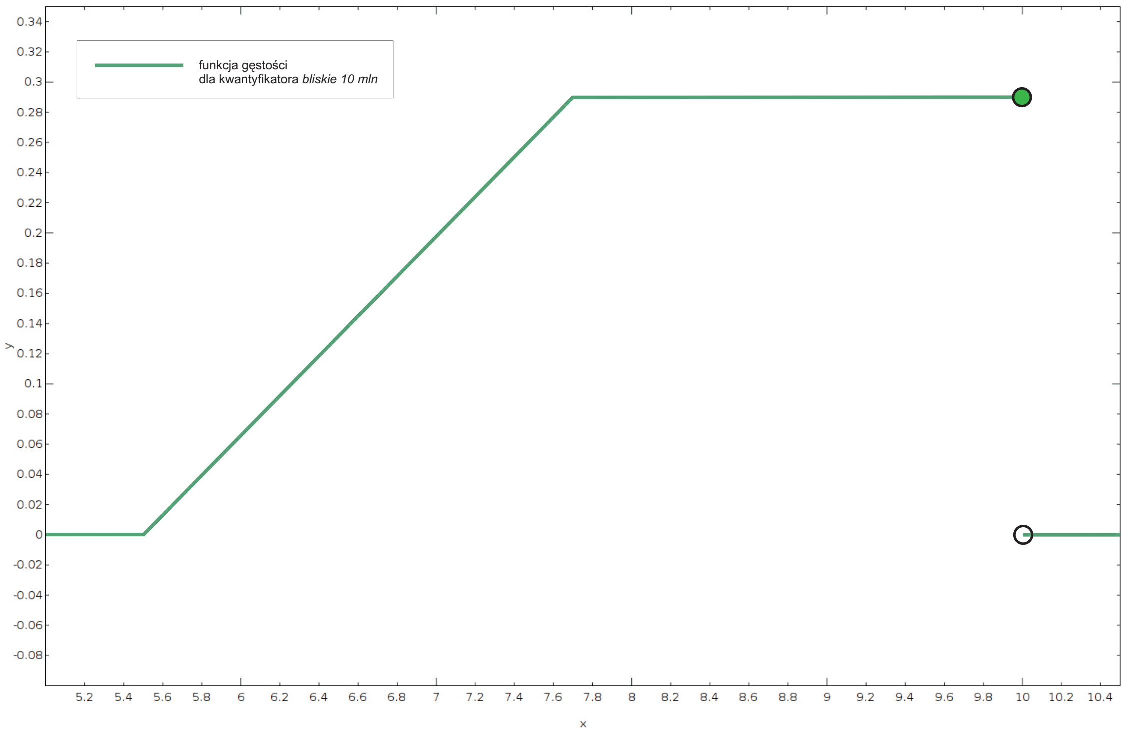
$$\begin{cases} (10-7.7)*c+\frac{1}{2}*(7.7-5.5)*c=1 \\ 5.5*a+b=0 \\ 7.7*a+b=c. \end{cases} \tag{4.12}$$

Końcowa operacja macierzowa wygląda następująco:

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 & 0 & 3.4 \\ 5.5 & 1 & 0 \\ 7.7 & 1 & -1 \end{bmatrix}^{-1} * \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \tag{4.13}$$

jej rozwiązaniem jest:

$$a=0.13, b=-0.73, c=0.29. \tag{4.14}$$



RYSUNEK 4.5: Wykres funkcji gęstości dla kwantyfikatora *bliskie 10 mln*

Funkcja rozkładu gęstości prawdopodobieństwa, która aproksymuje zbiór punktów danych dla kwantyfikatora *bliskie 10 mln* przedstawiona została poniżej:

$$f(x) = \begin{cases} 0.13 * x - 0.73 & , \text{dla } 5.5 < x \leqslant 7.7 \\ 0.29 & , \text{dla } 7.7 < x < 1. \end{cases} \tag{4.15}$$

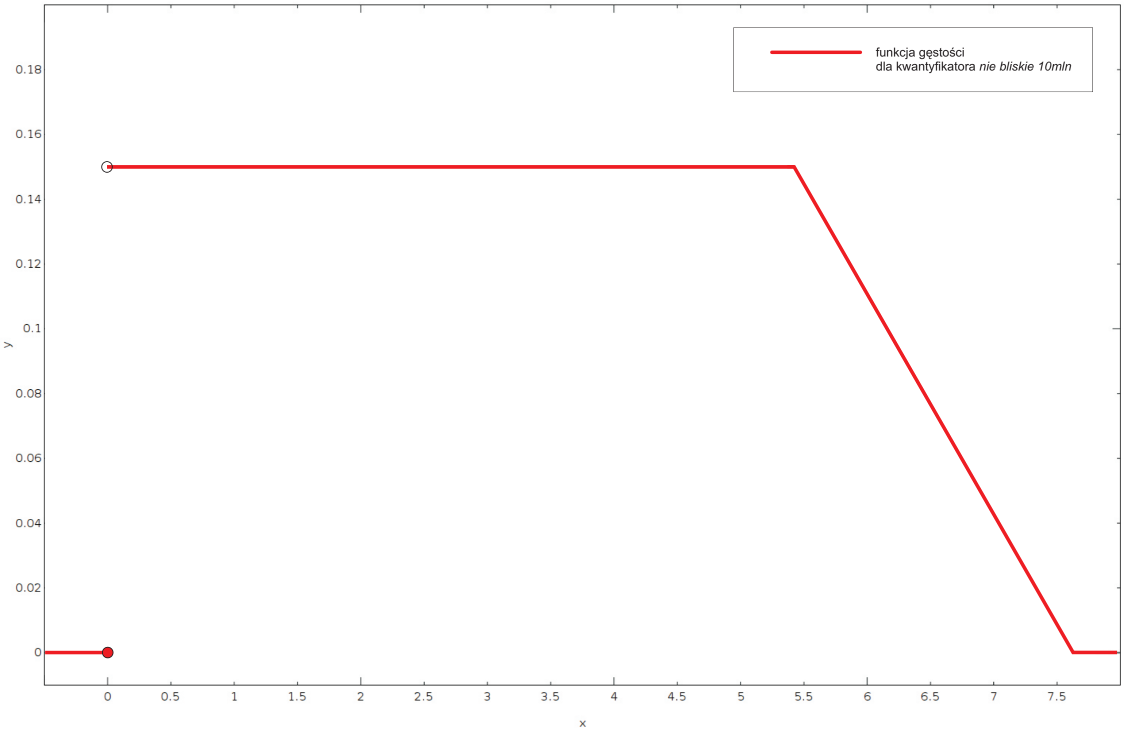
4.2.4 Kwantyfikator "nie bliskie 10 mln"

Kwantyfikator *nie bliskie 10 mln* jest zadany przez wykres funkcji 4.6. Aby uzyskać wzór funkcji dla tego i reszty kwantyfikatorów, rozwiązujemy układ równań analogiczny do układu 4.3 o postaci:

$$\begin{cases} 5.5 * c + \frac{1}{2} * (7.7 - 5.5) * c = 1 \\ 5.5 * a + b = c \\ 7.7 * a + b = 0. \end{cases} \tag{4.16}$$

Finalna operacja macierzowa wygląda następująco:

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 & 0 & 6.6 \\ 5.5 & 1 & -1 \\ 7.7 & 1 & 0 \end{bmatrix}^{-1} * \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}. \tag{4.17}$$



Rysunek 4.6: Wykres funkcji gęstości dla kwantyfikatora *nie bliskie 10 mln*

Rozwiązanie tego układu jest trójka:

$$a = -0.06, b = 0.53, c = 0.15. \tag{4.18}$$

Funkcja rozkładu gęstości prawdopodobieństwa, która aproksymuje zbiór punktów danych dla kwantyfikatora *nie bliskie 10 mln* wygląda zatem następująco:

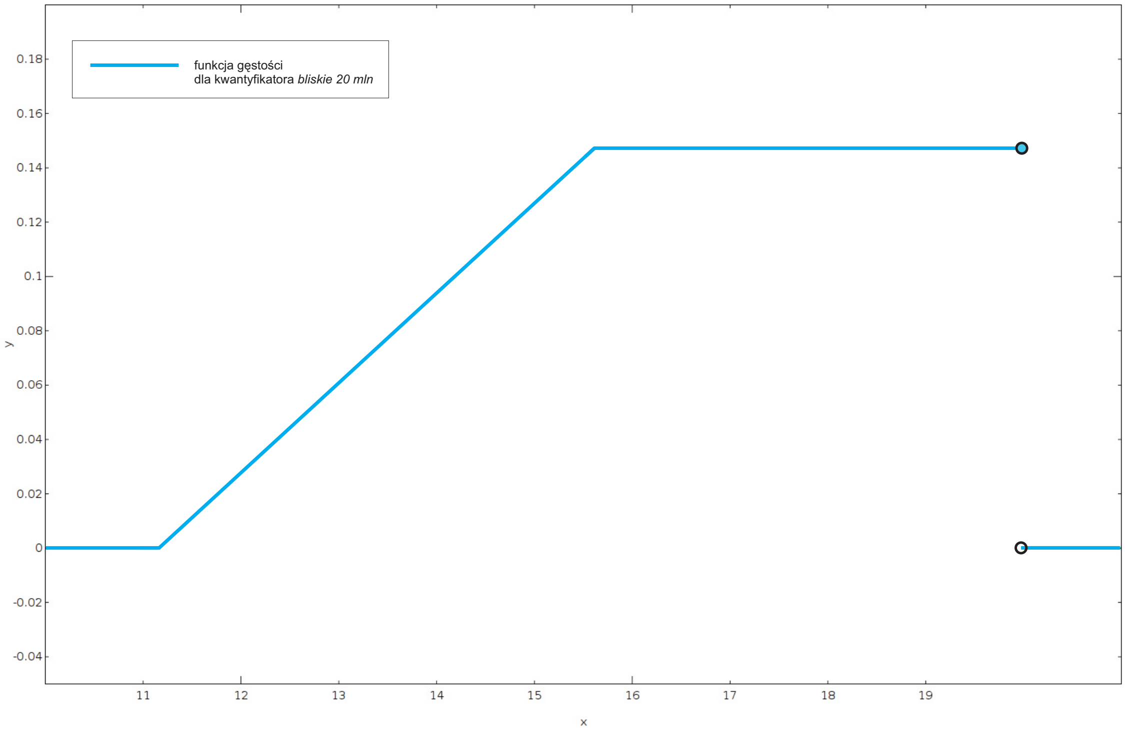
$$f(x) = \begin{cases} -0.06 * x + 0.53, & \text{dla } 5.5 < x \leqslant 7.7 \\ 0.15, & \text{dla } 0 < x < 5.5. \end{cases} \tag{4.19}$$

.

4.2.5 Kwantyfikator "bliskie 20 mln"

Kwantyfikator *bliskie 20 mln* jest zadany przez wykres funkcji 4.7. Aby uzyskać wzór funkcji dla tej funkcji gęstości, rozwiązujemy układ równań analogiczny do układu 4.3 o postaci:

$$\begin{cases} (20 - 15.4) * c + \frac{1}{2} * (15.4 - 11) * c = 1 \\ 11 * a + b = 0 \\ 15.4 * a + b = c \end{cases} \tag{4.20}$$



Rysunek 4.7: Wykres funkcji gęstości dla kwantyfikatora *bliskie 20 mln*

. Finalna operacja macierzowa wygląda następująco:

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 & 0 & 6.8 \\ 11 & 1 & 0 \\ 15.4 & 1 & -1 \end{bmatrix}^{-1} * \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \tag{4.21}$$

rozwiązaniem jest trójka:

$$a = 0.033, b = -0.368, c = 0.147. \tag{4.22}$$

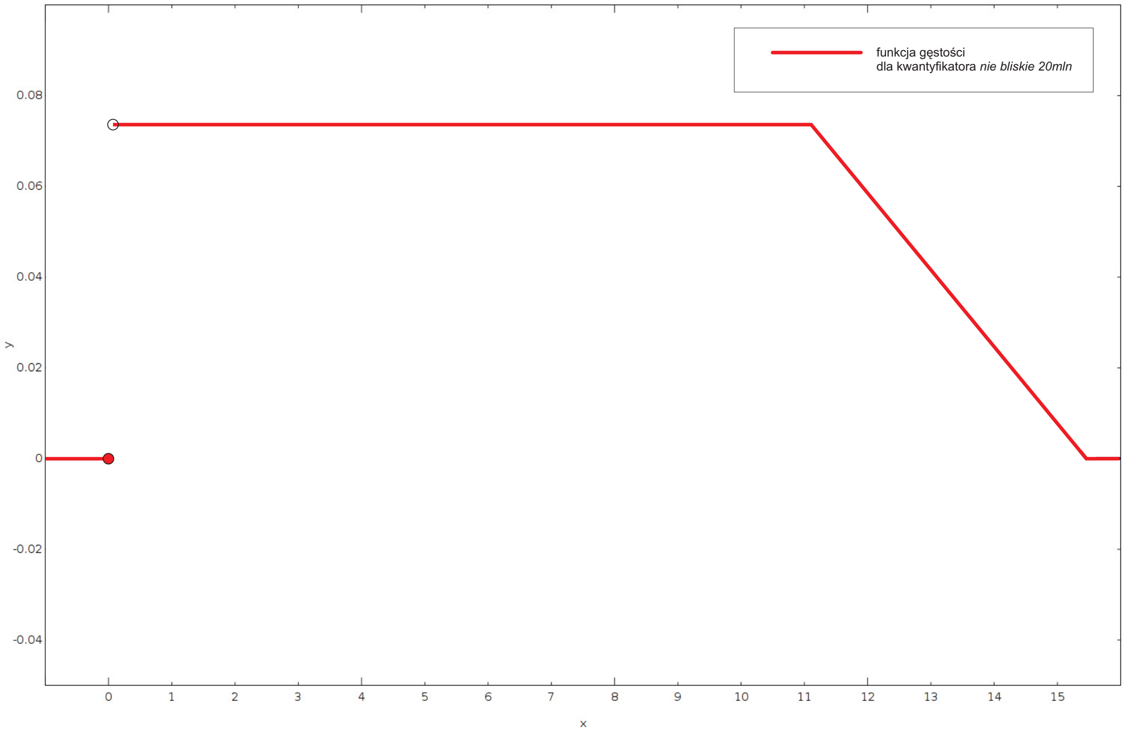
Funkcja rozkładu gęstości prawdopodobieństwa, która aproksymuje zbiór punktów danych dla kwantyfikatora *bliskie 20 mln* wygląda zatem następująco:

$$f(x) = \begin{cases} 0.02 * x - 0.22, & \text{dla } 11 < x \leqslant 15.4 \\ 0.09, & \text{dla } 15.4 < x < 20. \end{cases} \tag{4.23}$$

.

4.2.6 Kwantyfikator "nie bliskie 20 mln"

Kwantyfikator *nie bliskie 20 mln* jest zadany przez wykres funkcji 4.8. Aby uzyskać wzór



Rysunek 4.8: Wykres funkcji gęstości dla kwantyfikatora *nie bliskie 20 mln*

funkcji dla tego kwantyfikatora, rozwiązujemy układ równań analogiczny do układu 4.3 o postaci:

$$\begin{cases} 11 * c + \frac{1}{2} * (15.4 - 11) * c = 1 \\ 11 * a + b = c \\ 15.4 * a + b = 0. \end{cases} \tag{4.24}$$

Finalna operacja macierzowa:

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 & 0 & 13.2 \\ 11 & 1 & -1 \\ 15.4 & 1 & 0 \end{bmatrix}^{-1} * \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \tag{4.25}$$

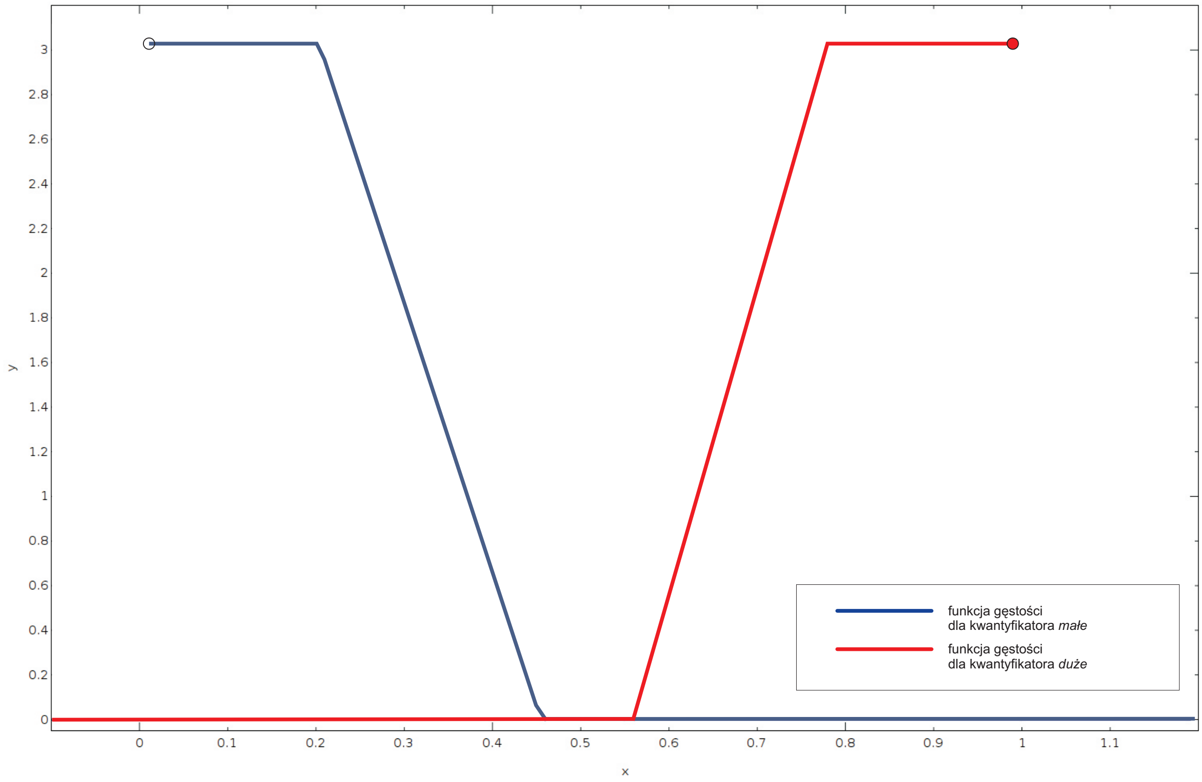
oraz jej rozwiązanie

$$a = -0.017b = 0.263, c = 0.075. \tag{4.26}$$

Funkcja rozkładu gęstości prawdopodobieństwa, która aproksymuje zbiór punktów danych dla kwantyfikatora *nie bliskie 20 mln* wygląda zatem następująco:

$$f(x) = \begin{cases} -0.017 * x + 0.263, & \text{dla } 11 < x \leqslant 15.4 \\ 0.075, & \text{dla } 0 < x < 11. \end{cases} \tag{4.27}$$

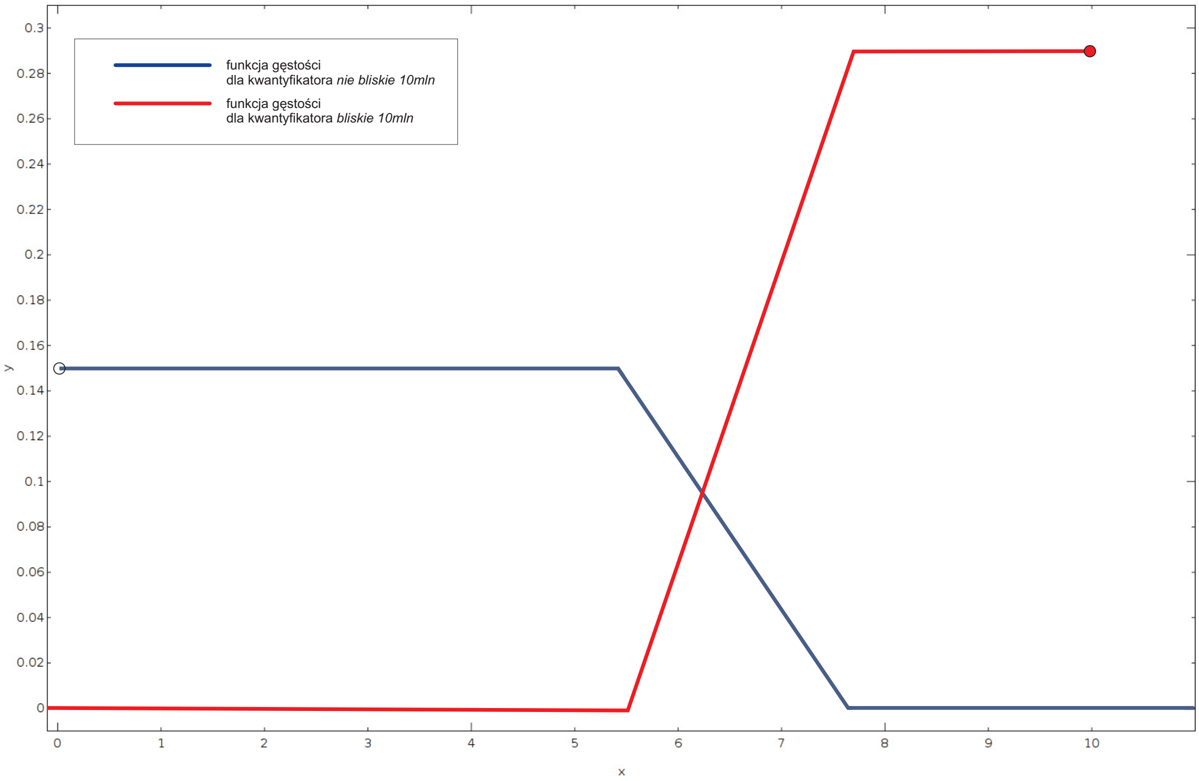
Poniżej, pokazano parami, funkcje rozkładu gęstości prawdopodobieństwa dla przeciwstawnych kwantyfikatorów. Funkcję rozkładu gęstości prawdopodobieństwa dla kwantyfikatorów *duże małe* pokazano na wykresie 4.9, dla kwantyfikatorów *bliskie 10 mln* i *nie bliskie 10 mln*, pokazano na wykresie 4.10 a dla kwantyfikatorów *bliskie 20 mln* i *nie bliskie 20 mln* na wykresie 4.11.



Rysunek 4.9: Wykres funkcji gęstości dla kwantyfikatorów *duże* i *małe*

4.3 Rozwiązanie problemu

Po identyfikacji funkcji rozkładu gęstości prawdopodobieństwa, można przystąpić do rozwiązania zadanego problemu. Jako odpowiedź na postawiony w zadaniu problem decyzyjny postaramy się podać rozkłady gęstości prawdopodobieństwa dla biznesu pierwszego oraz biznesu drugiego, oraz policzyć wartość oczekiwaną z wynikowych rozkładów. Wartość oczekiwana oraz odchylenie standardowe z tych rozkładów będą stanowiły dla nas podstawę do ostatecznego podjęcia decyzji, który biznes, ze statystycznego punktu widzenia, jest bardziej korzystny.



RYSUNEK 4.10: Wykres funkcji gęstości dla kwantyfikatorów *bliskie 10 mln* i *nie bliskie 10 mln*

4.3.1 Przedstawienie analitycznej formy rozkładów gęstości prawdopodobieństwa dla zadania

Ogólna funkcja rozkładu gęstości przyjmuje postać:

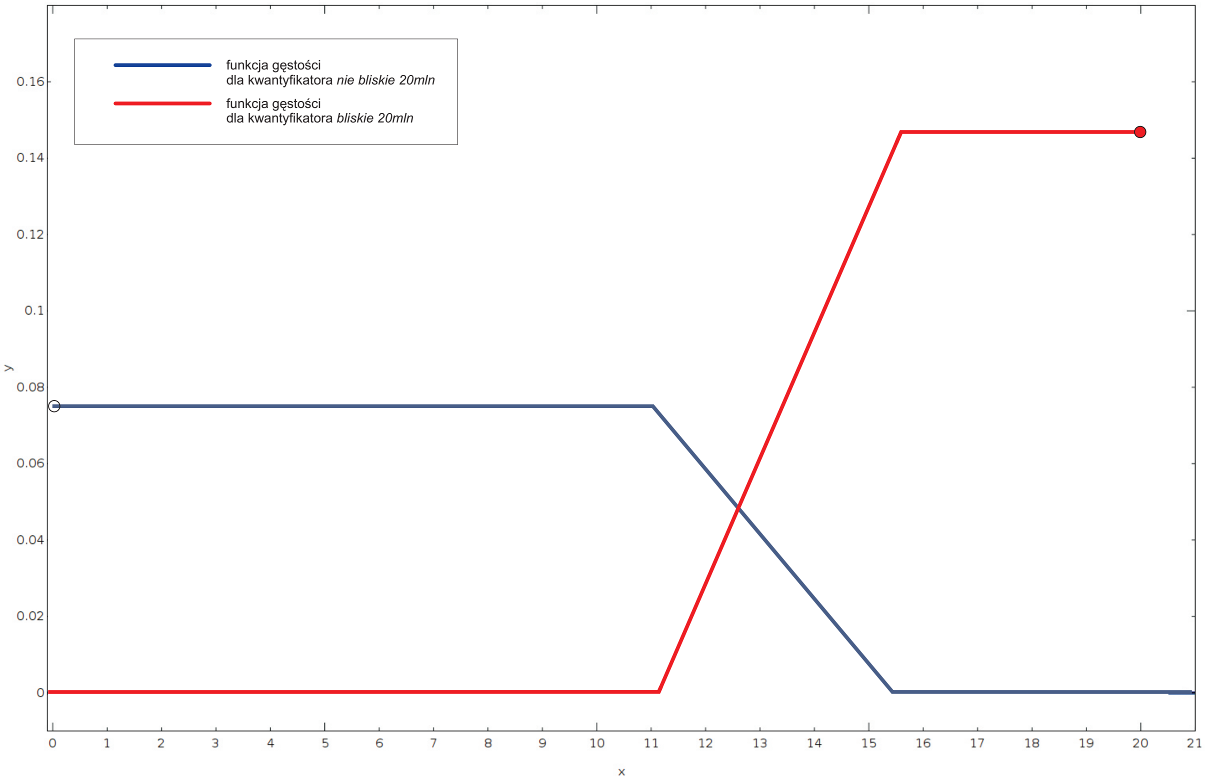
$$gp_{DBn} = f(DBn, p), \tag{4.28}$$

gdzie

- gp_{DBn} - oznacza gęstość rozkładu prawdopodobieństwa dochodu z biznesu n-tego
- DBn - oznacza dochód z biznesu n-tego
- p - oznacza prawdopodobieństwo

Funkcję rozkładu gęstości prawdopodobieństwa dla biznesu pierwszego przedstawia się następująco:

$$gp_{DB1} = f(DB1, p) = fb10mln(p) \otimes D(p) \oplus fnb10mln(p) \otimes M(p), \tag{4.29}$$



RYSUNEK 4.11: Wykres funkcji gęstości dla kwantyfikatorów *bliskie 20 mln* i *nie bliskie 20 mln*

gdzie

- $fb10mln(p)$ - oznacza funkcję rozkładu gęstości dla kwantyfikatora *bliskie 10 mln*
- $D(p)$ - oznacza funkcję rozkładu gęstości dla kwantyfikatora *duże*
- $fnb10mln(p)$ - oznacza funkcję rozkładu gęstości dla kwantyfikatora *nie bliskie 10 mln*
- $M(p)$ - oznacza funkcję rozkładu gęstości dla kwantyfikatora *małe*

Analogicznie, by określamy funkcję rozkładu gęstości prawdopodobieństwa dla biznesu drugiego:

$$gp_{DB2} = f(DB1, p) = fb20mln(p) \otimes M(p) \oplus fnb20mln(p) \otimes D(p), \tag{4.30}$$

gdzie

- $fb20mln(p)$ - oznacza funkcję rozkładu gęstości dla kwantyfikatora *bliskie 20 mln*
- $D(p)$ - oznacza funkcję rozkładu gęstości dla kwantyfikatora *duże*

- $fnb20mln(p)$ - oznacza funkcję rozkładu gęstości dla kwantyfikatora *nie bliskie 20 mln*
- $M(p)$ - oznacza funkcję rozkładu gęstości dla kwantyfikatora *małe*

Gdy uda nam się ustalić obie funkcje, należy policzyć wartość oczekiwaną takiego rozkładu oraz odchylenie standardowe, co będzie stanowiło dla nas statystyczną odpowiedź.

4.3.2 Funkcje rozkładu gęstości prawdopodobieństwa dla biznesu pierwszego

Funkcja gęstości rozkładu prawdopodobieństwa jest dana wzorem 4.29. Zanim zostanie przedstawiona końcowa funkcje gęstości, pokazane zostaną wyniki posiłkowe poszczególnych operacji. Pierwszą operacją jest operacja:

$$biznes1f1(x) = fb10mln(p) \otimes D(p), \tag{4.31}$$

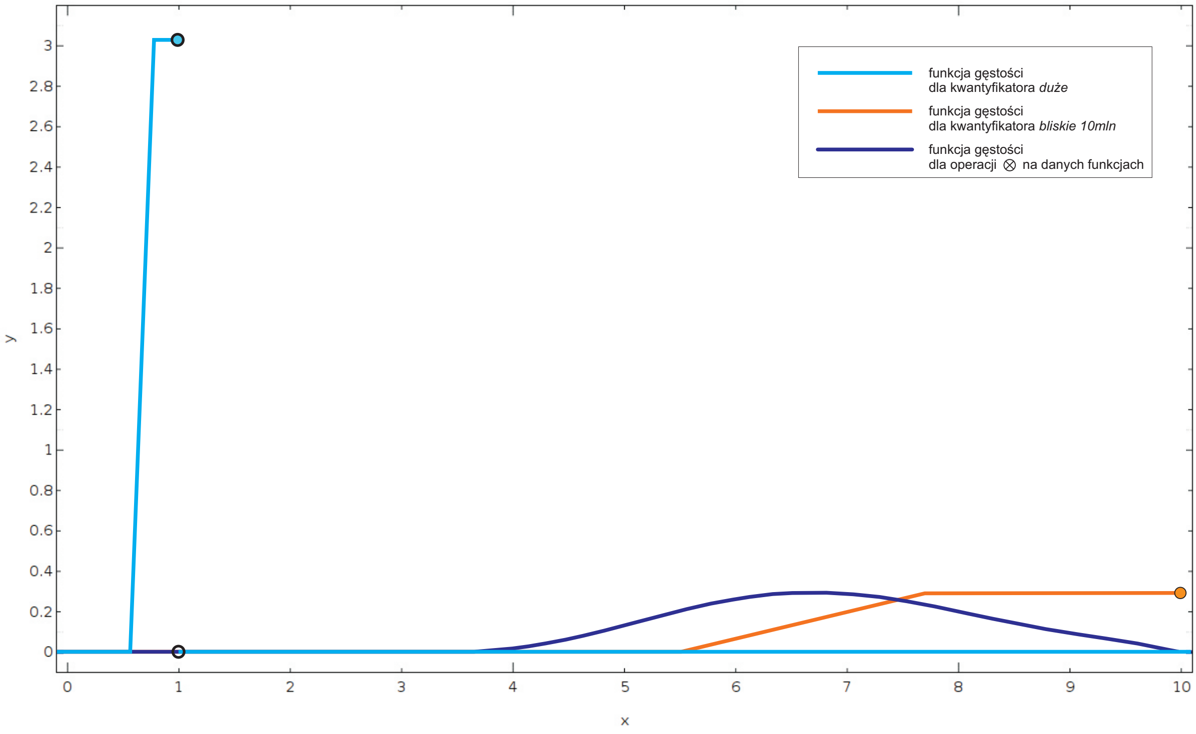
która zachodzi między funkcjami rozkładów gęstości prawdopodobieństwa dla kwantyfikatorów *bliskie 10 mln* i *duże*.Operacja ta $\otimes$ została opisana w rozdziale 2. Wynik tego działania został przedstawiony na wykresie 4.3.2. Funkcja, która powstaje w wyniku tej operacji, wyliczona w programie jest funkcją dyskretną określoną w danych punktach. Ponieważ, nas interesują jedynie funkcje ciągłe, funkcja ta musi zostać zaaproksymowana na zadanym przedziale (pamiętamy o tym że funkcja będąca wynikiem aproksymacji, również musi być funkcją gęstość, co oznacza jej całkowalność do 1) . Wykres 4.12 przedstawia funkcję dyskretną oraz jej aproksymację do wielomianu 6-tego stopnia. Stopnie wielomianu, w tej operacji przybliżenia, oraz w pozostałych, zostały ustalone empirycznie poprzez minimalizację odchylenia odchyłki funkcji od jej przybliżenia. Równanie funkcji zostało przedstawione na wzorze[2] 4.32:

$$biznes1f1(x) = \begin{cases} \frac{-7.2+8.8-4.2x^2+1.0x^3-0.1x^6}{0.98}, & \text{dla } 3.08 < x \leqslant 9.98 \\ 0, & \text{dla } -\infty < x < 3.08 \\ 0, & \text{dla } 9.98 < x < \infty \end{cases} . \tag{4.32}$$

Jak widać na wykresie 4.12, aproksymacja ta jest bardzo dokładna na przedziale i jej błąd jest niewielki.
Następnym działaniem jest:

$$biznes1f2(x) = fnb10mln(p) \otimes M(p). \tag{4.33}$$

[2]program podaje dużo dokładniejsze wartości stałych w aproksymowanych wielomianach, ale autor zdecydował się podawać ich zaokrągloną postać, w celach estetycznych



Wynik tej operacji został przedstawiony na wykresie 4.13. Tak jak w przypadku operacji 4.31 tak w tym przypadku powstała funkcja jest funkcją dyskretną wymagającą aproksymacji na przedziale. Równanie funkcji zostało przedstawione na wzorze 4.34, a niedokładności aproksymacji na wykresie 4.14.

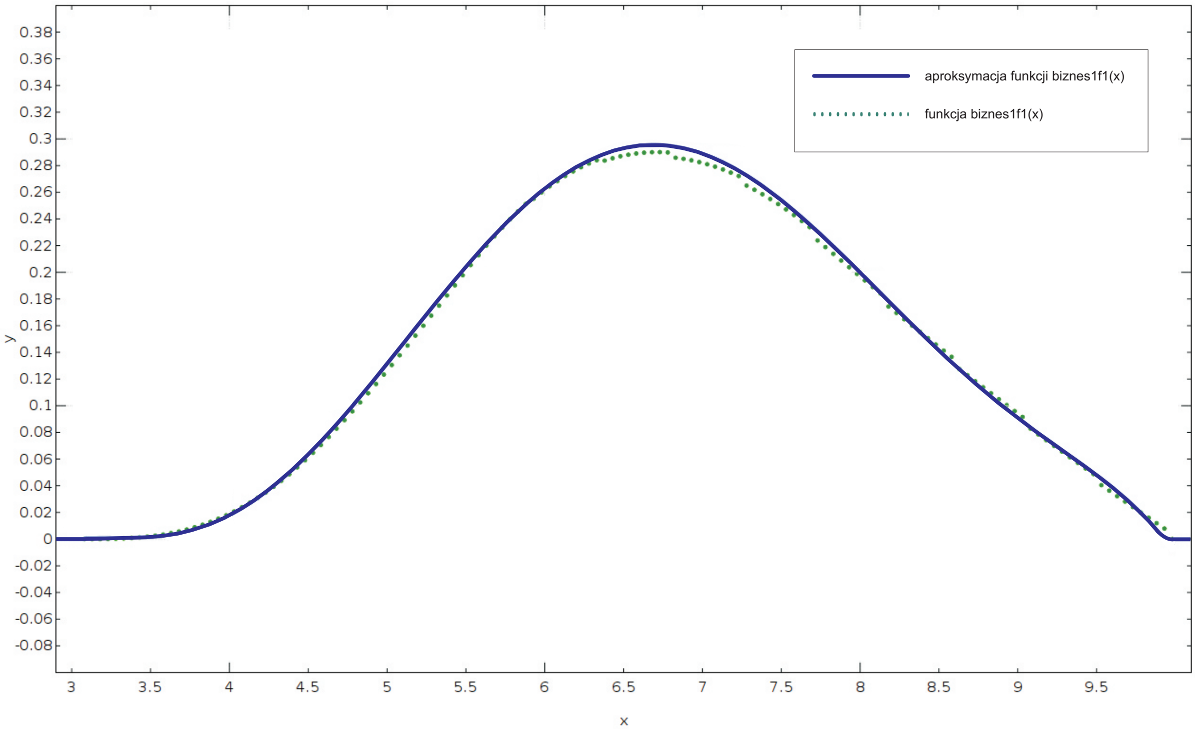
$$biznes1f2(x) = \begin{cases} \frac{2.5-13.4x+40.2x^2-64.1x^3+57.4x^4-30.0x^5+9x^6-1.4x^7+0.09x^8}{0.97}, & \text{dla } 0 < x \leqslant 3.45 \\ 0, & \text{dla } -\infty < x < 0 \\ 0, & \text{dla } 3.45 < x < \infty \end{cases} \tag{4.34}$$

Ostatnią operacją jest splecenie funkcji pośrednich, uzyskanych w operacjach 4.31 oraz 4.33 :

$$gp_{DB1} = biznes1f1(x) \oplus biznes1f2(x) = fb10mln(p) \otimes D(p) \oplus fnb10mln(p) \otimes M(p), \tag{4.35}$$

o wzorach 4.32 i4.34. Wynik tego działania pokazany jest na wykresie 4.15, a niedokładność aproksymacji na wykresie 4.16. Wynikowa funkcja gp_{DB1}, ma wzór:

$$gp_{DB1} = \begin{cases} \frac{2.5-1.5x+0.2x^2+0.01x^3-0.007x^4+0.0006x^5-0.001x^6}{0.9}, & \text{dla } 3.18 < x \leqslant 13.38 \\ 0, & \text{dla } -\infty < x < 3.18 \\ 0, & \text{dla } 13.38 < x < \infty \end{cases} . \tag{4.36}$$



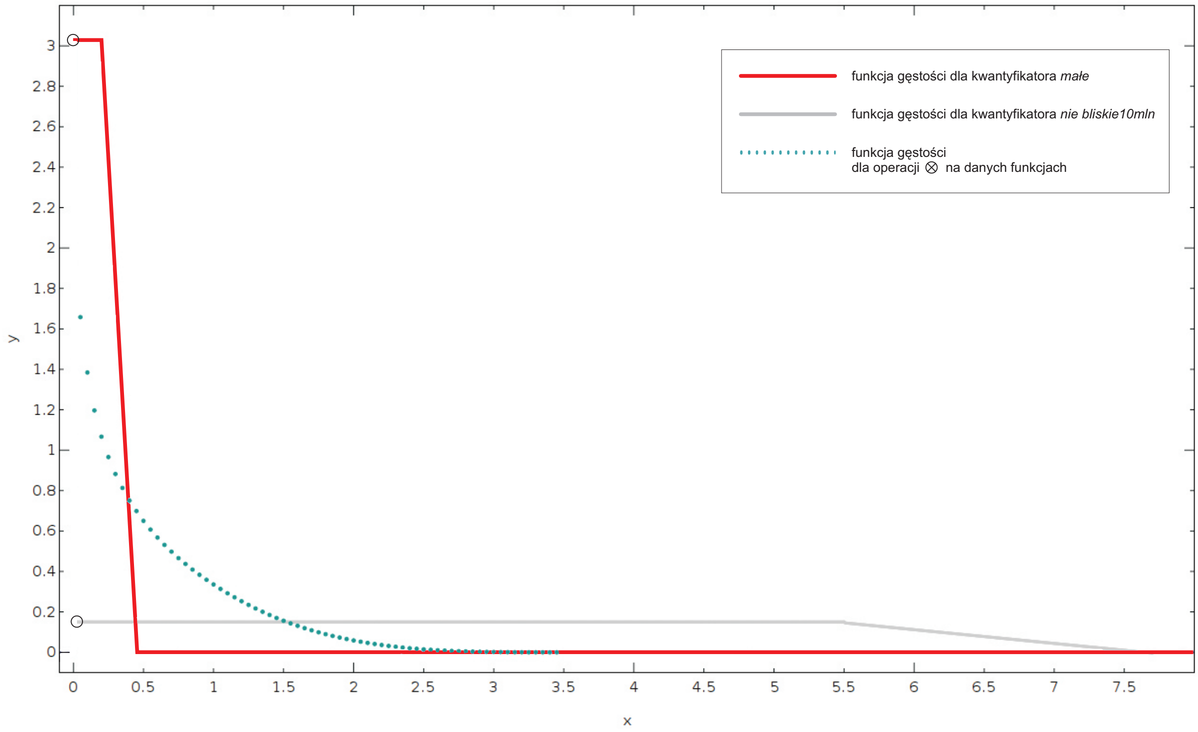
RYSUNEK 4.12: Wykres przedstawiający niedokładność aproksymacji funkcji gęstości *bliskie 10 mln*$\otimes$ *duże*

4.3.3 Funkcje rozkładu gęstości prawdopodobieństwa dla biznesu drugiego

Funkcja gęstości rozkładu prawdopodobieństwa dla biznesu drugiego jest dana wzorem 4.30. Tak jak w przypadku obliczeń funkcji gęstości prawdopodobieństwa dla biznesu pierwszego, tak w tym przypadku obliczenia zostaną podzielone na trzy operacje. Pierwszym działaniem jest:

$$biznes2f1(x) = fnb20mln(p) \otimes D(p), \tag{4.37}$$

która zachodzi między funkcjami rozkładów gęstości prawdopodobieństwa dla kwantyfikatorów *nie bliskie 20 mln* i *duże*.Wynik tego działania został przedstawiony na wykresie 4.17, a niedokładność odwzorowania została uwidoczniona na wykresie 4.18. Równanie



RYSUNEK 4.13: Wykres funkcji gęstości dla kwantyfikatorów *małe* i *nie bliskie 10 mln* oraz funkcji wynikowej dla operacji ⊗ między nimi

funkcji zostało przedstawione na wzorze 4.38:

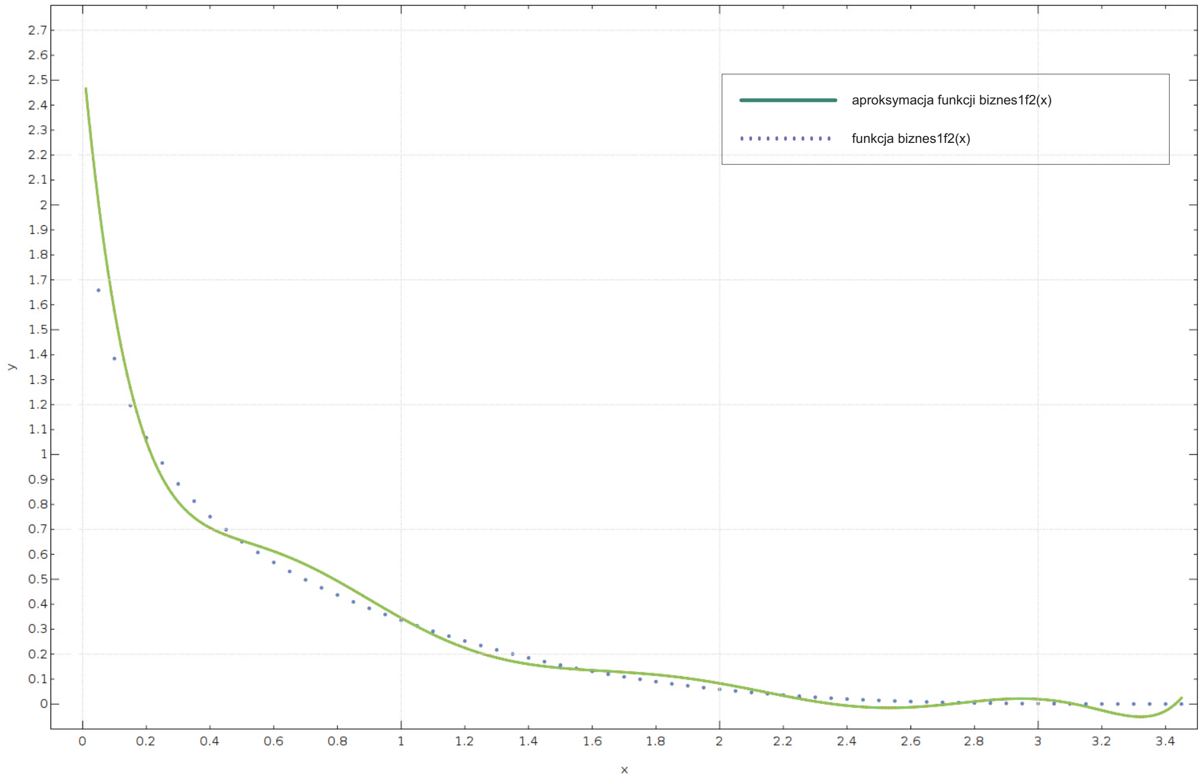
$$biznes2f1(x) = \begin{cases} \frac{0.08+0.01x-0.01x^2+0.003x^3-0.0003x^4}{0.98} & \text{, dla } 0.5 < x \leqslant 15.35 \\ 0, & \text{dla } -\infty < x < 0.5 \\ 0, & \text{dla } 15.35 < x < \infty \end{cases} . \tag{4.38}$$

Następnym działaniem jest:

$$biznes2f2(x) = fb20mln(p) \otimes M(p). \tag{4.39}$$

Wynik tej operacji został przedstawiony na wykresie 4.19. Równanie zaaproksymowanej funkcji zostało przedstawione na wzorze 4.40, a niedokładności aproksymacji na wykresie 4.20.

$$biznes1f2(x) = \begin{cases} \frac{0.1-0.02x+0.03x^2-0.01x^3+0.001x^4}{0.95} & \text{dla } 0.1 < x \leqslant 9.05 \\ 0, & \text{dla } -\infty < x < 0.1 \\ 0, & \text{dla } 9.05 < x < \infty \end{cases} \tag{4.40}$$



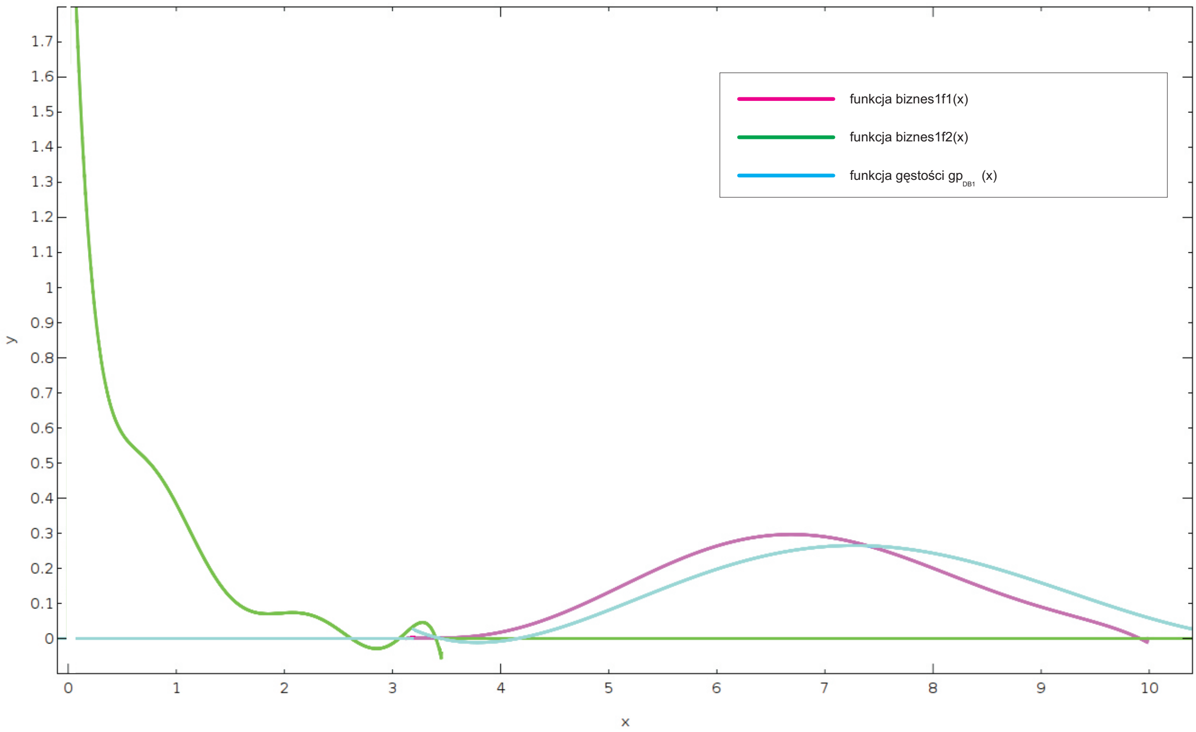
RYSUNEK 4.14: Wykres przedstawiający niedokładność aproksymacji funkcji gęstości *małe* $\otimes$ *nie bliskie 10 mln*

Ostatnią operacją jest splecenie funkcji pośrednich, uzyskanych w operacjach 4.37 oraz 4.39 :

$$gp_{DB2} = biznes2f1(x) \oplus biznes2f2(x) = fnb20mln(p) \otimes D(p) \oplus fb20mln(p) \otimes M(p), \tag{4.41}$$

o wzorach 4.38 i4.40. Wynik tego działania pokazany jest na wykresie 4.21, a niedokładność aproksymacji na wykresie 4.22. Wynikowa funkcja gp_{DB2}, ma wzór:

$$gp_{DB2} = \begin{cases} \frac{-0.001+0.01x+0.001x^2+-0.0003x^3+1.37708e-05*x^4}{1.0002}, & \text{dla } 0.25 < x \leqslant 24.35 \\ 0, & \text{dla } -\infty < x < 3.18 \\ 0, & \text{dla } 13.38 < x < \infty \end{cases} . \tag{4.42}$$



Rysunek 4.15: Wykres funkcji gęstości gp_{DB1}

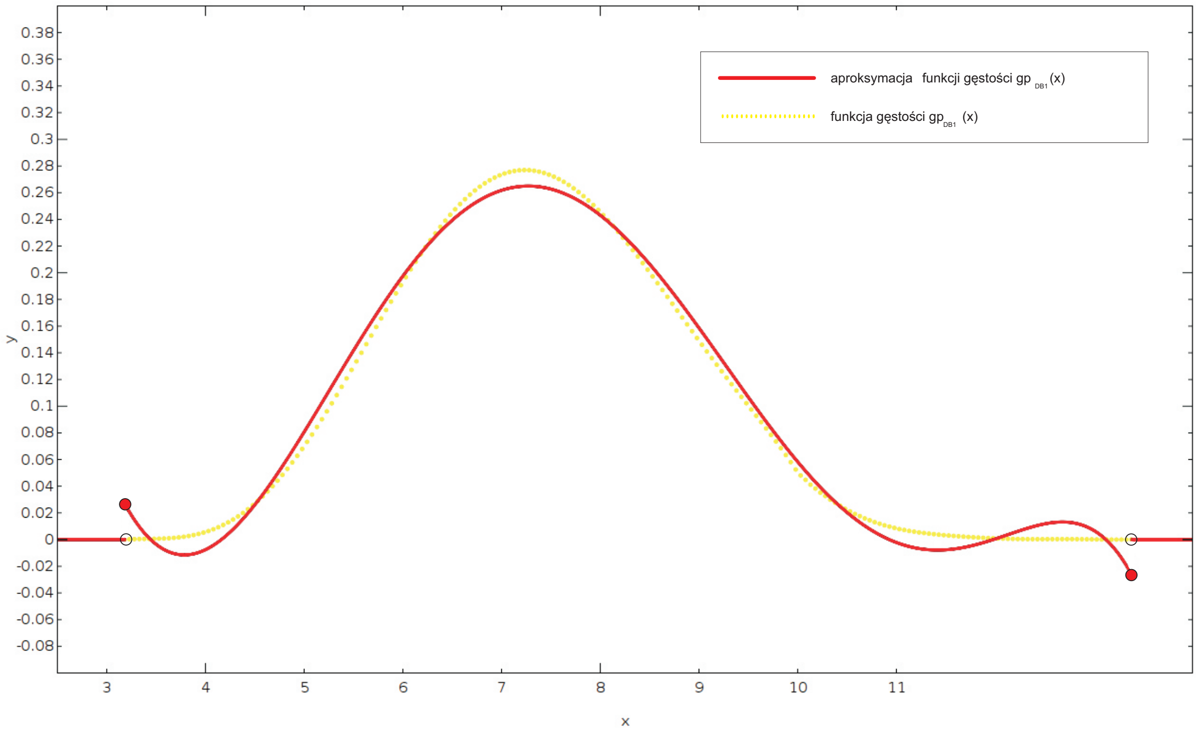
4.3.4 Wartości oczekiwane i błąd

Z funkcji rozkładu gęstości prawdopodobieństwa biznesu pierwszego, danej wzorem 4.36, policzona wartość oczekiwana wynosi ≈ 7.4, natomiast z funkcji rozkładu gęstości prawdopodobieństwa biznesu drugiego, danej wzorem 4.36, wyliczona wartość oczekiwana wynosi ≈ 8.46. Na rysunku 4.23 przedstawiono oba funkcje rozkładów gęstości dla obu biznesów.

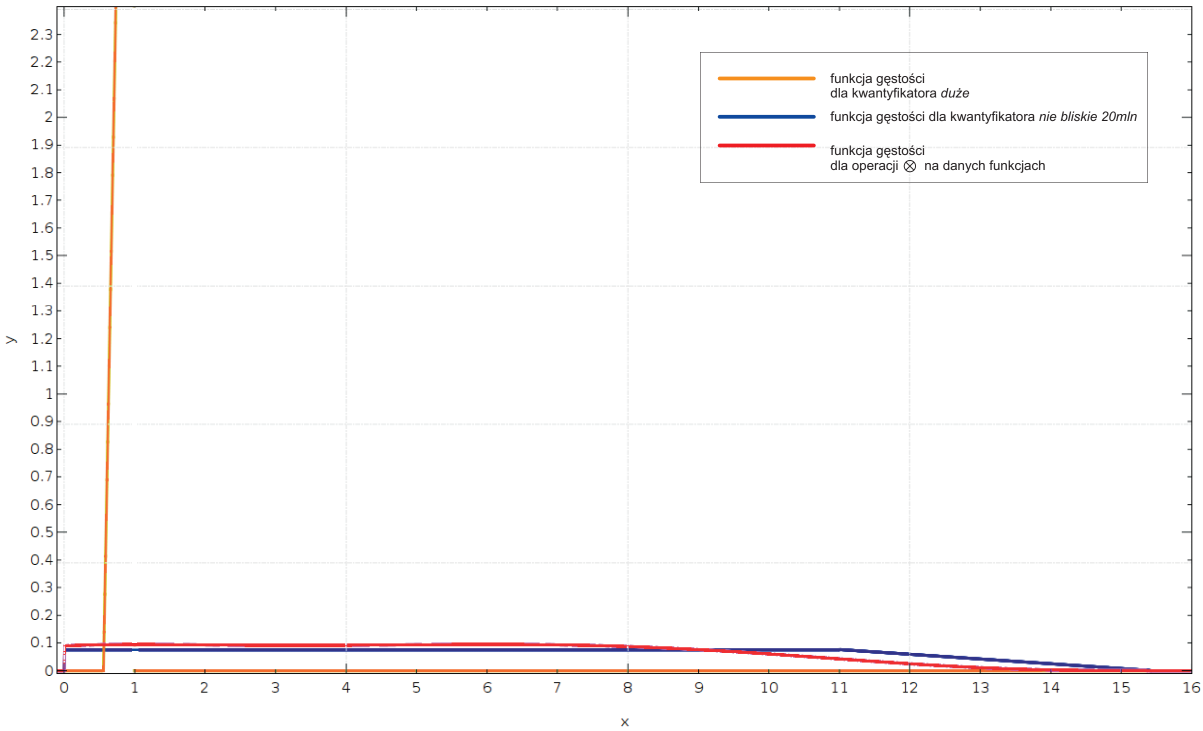
Przedstawione wyżej wartości są wartościami uśrednionymi. Zależnie od wyboru stopnia wielomianów, którymi aproksymowaliśmy wynikowe funkcje, wartość oczekiwana obu końcowych rozkładów ulegała zmianie. Wartość oczekiwana biznesu pierwszego zawsze należała do przedziału $[7.34, 7.4]$, natomiast dla biznesu drugiego $[8.38, 8.55]$. Błąd wyniku dla pierwszego biznesu wynosił zatem ≈ 1 % i był taki sam jak dla biznesu drugiego. Błąd na takim poziomie nie wpływa na końcowy wynik.

4.3.5 Odpowiedź na problem dwóch biznesów

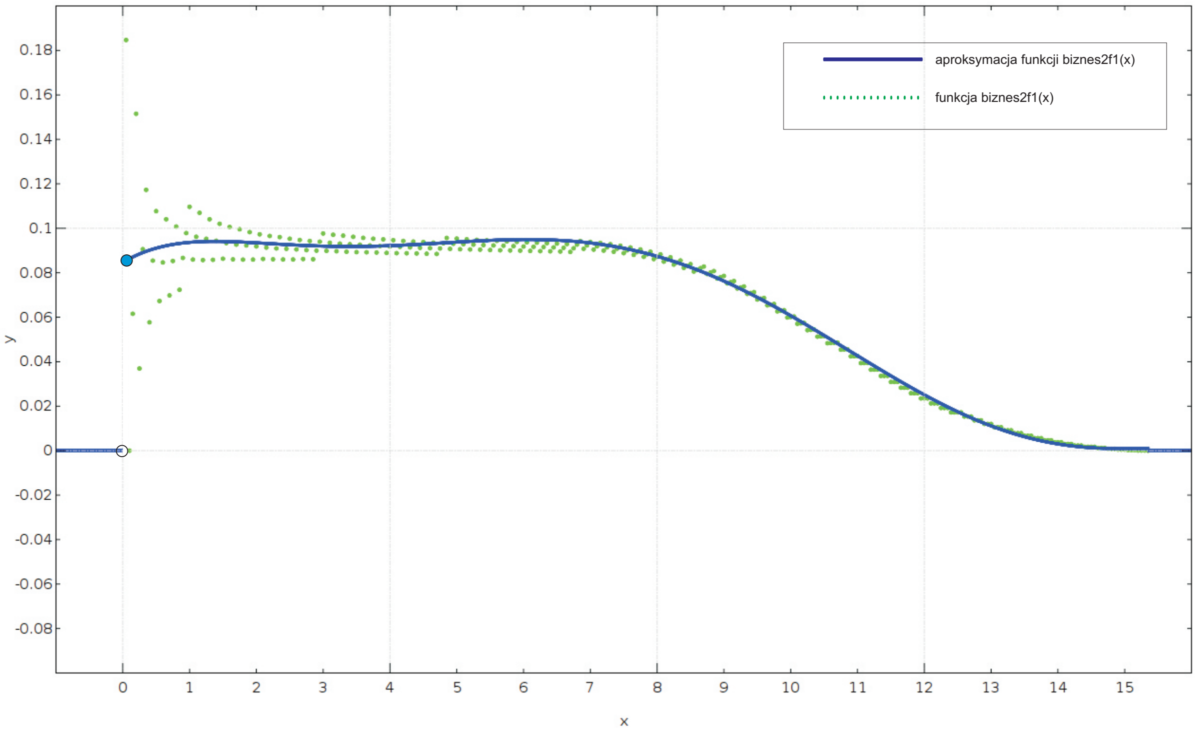
Ze statystycznego punktu widzenia, zdecydowanie lepiej jest wybrać biznes drugi, ponieważ wartość oczekiwana dochodu z niego jest wyższa ≈ 1 mln złotych.



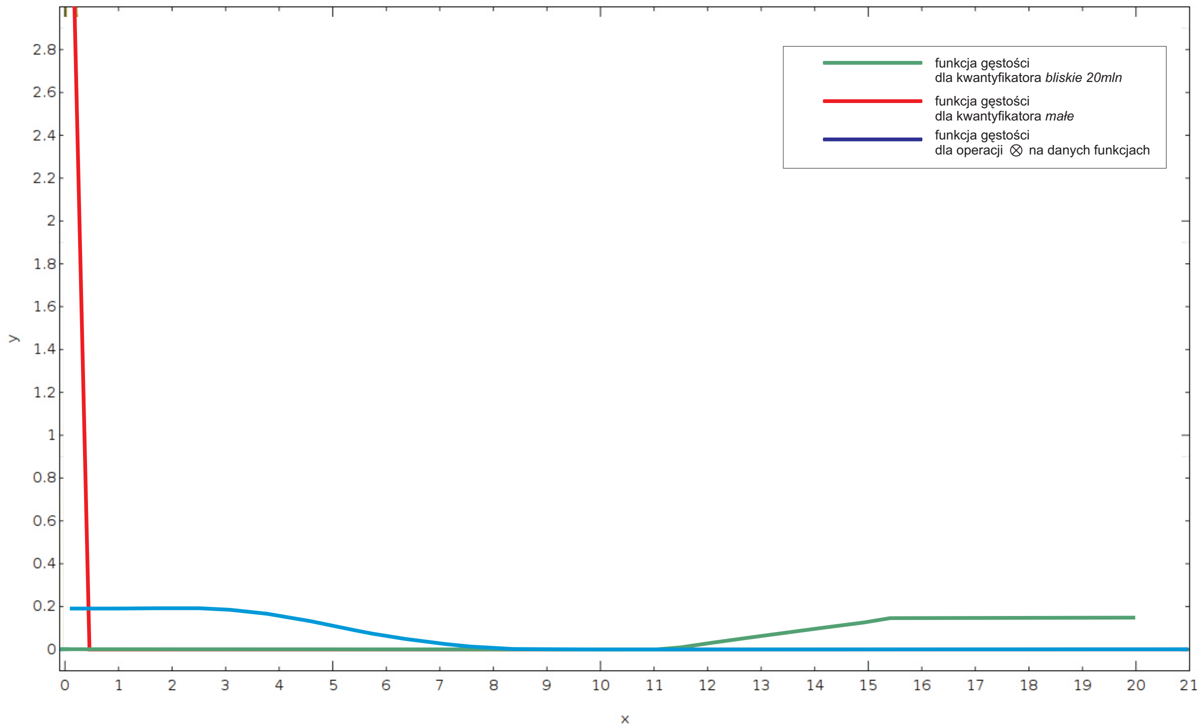
Rysunek 4.16: Wykres przedstawiający niedokładność aproksymacji funkcji gęstości gp_{DB1}



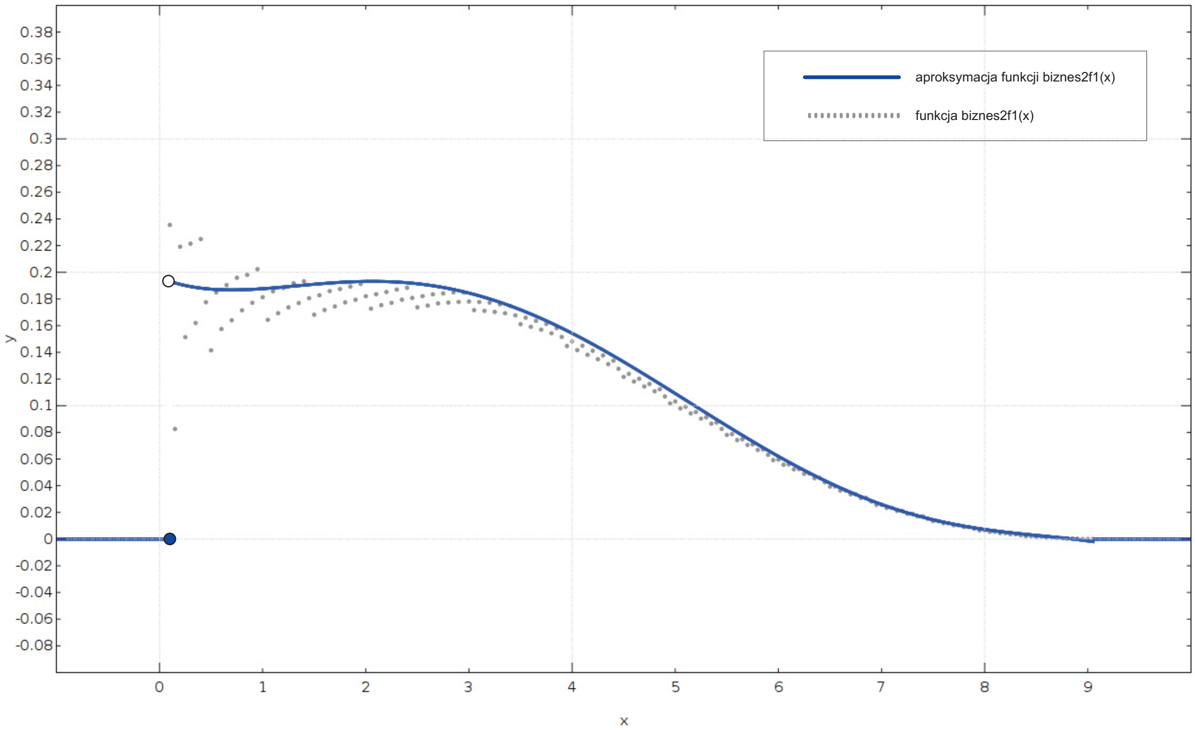
Rysunek 4.17: Wykres funkcji gęstości dla kwantyfikatorów *duże* i *nie bliskie 20 mln* oraz funkcji wynikowej dla operacji $\otimes$ między nimi



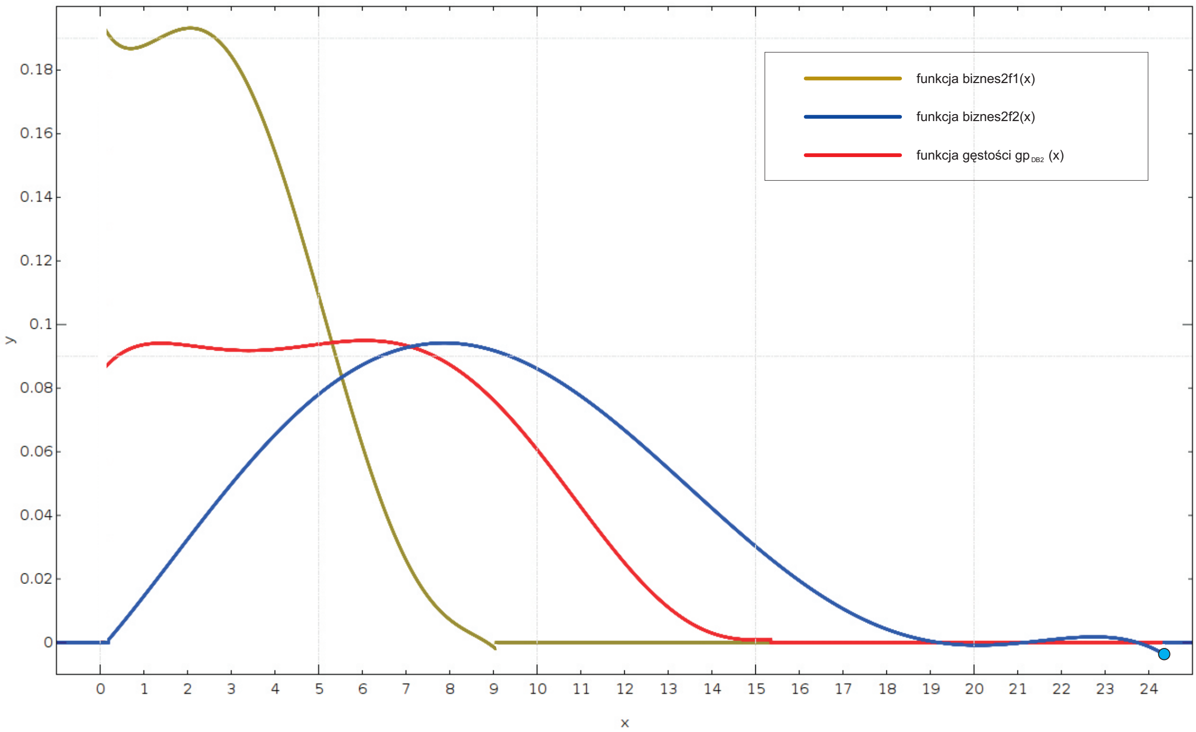
Rysunek 4.18: Wykres przedstawiający niedokładność aproksymacji funkcji gęstości *nie bliskie 10 mln*⊗ *duże*



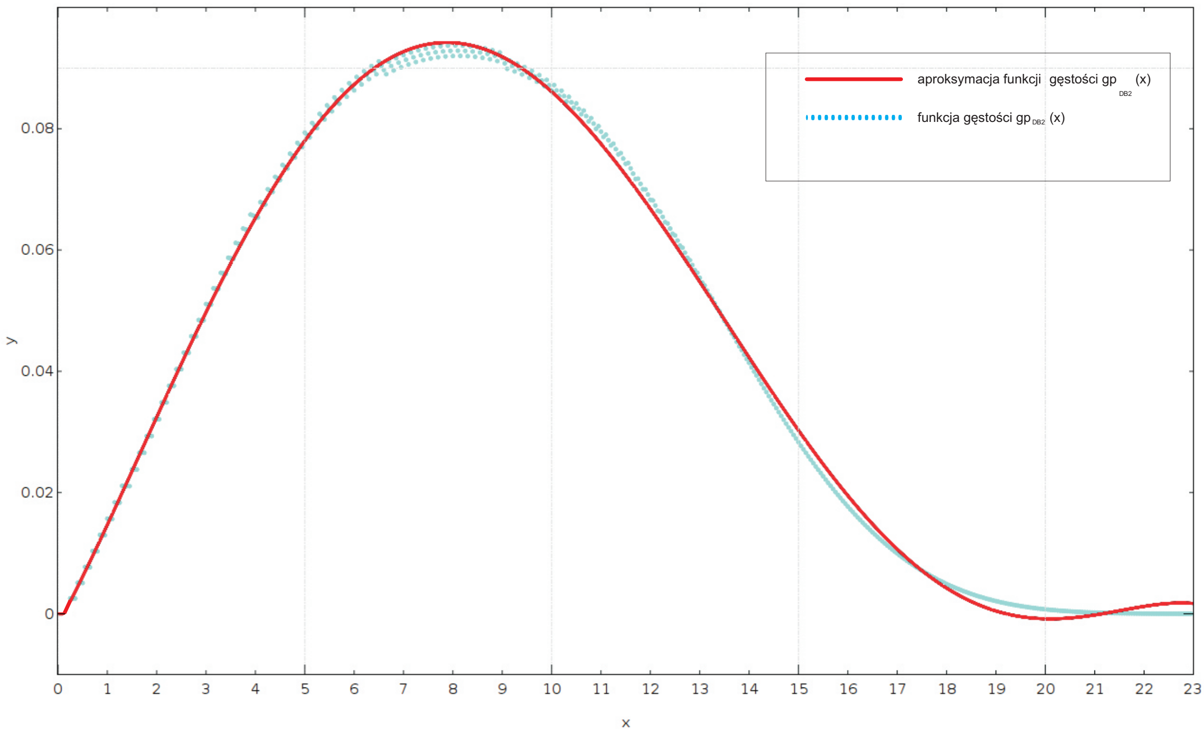
Rysunek 4.19: Wykres funkcji gęstości dla kwantyfikatorów *małe* i *nie bliskie 10 mln* oraz funkcji wynikowej dla operacji ⊗ między nimi



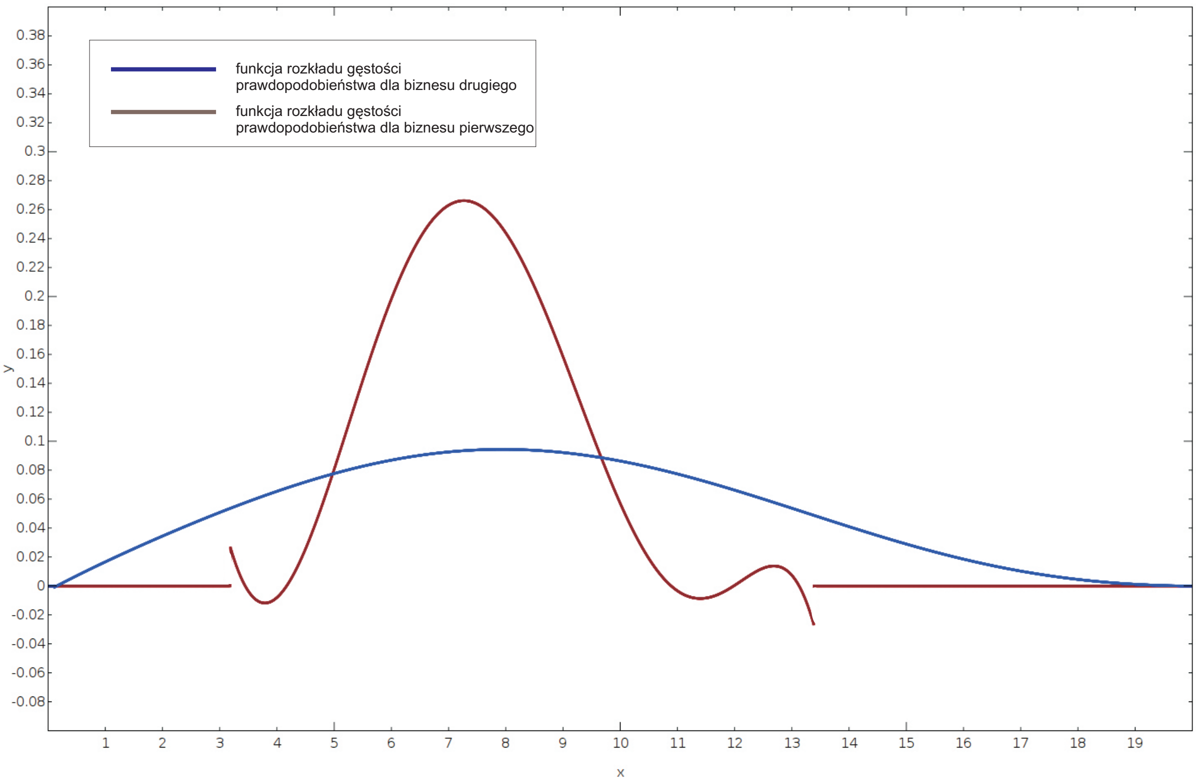
Rysunek 4.20: Wykres przedstawiający niedokładność aproksymacji funkcji gęstości *małe* $\otimes$ *nie bliskie 10 mln*



Rysunek 4.21: Wykres funkcji gęstości gp_{DB2}



RYSUNEK 4.22: Wykres przedstawiający niedokładność aproksymacji funkcji gęstości gp_{DB2}



RYSUNEK 4.23: Wykres przedstawiający niedokładność aproksymacji funkcji gęstości gp_{DB2}

Bibliografia

[1] Oscar Castillo, Patricia Melin, Oscar Montiel Ross, Roberto Sepulveda Cruz, Witold Pedrycz. *Theoretical Advances and Applications of Fuzzy Logic and Soft Computing.* Springer, 2007.

[2] Janina Jóźwiak, Jarosław Podgórski. *Statystyka od podstaw.* Państwowe Wydawnictwo EkonomicznePress, 1995.

[3] Marek Landowski. *Metoda identyfikacji probabilistycznych modeli ludzkich percepcji wyrażonych w formie słowników kwantyfikatorów lingwistycznych.* Praca doktorska, Zachodniopomorski Uniwersytet Technologiczny, 2009.

[4] Cornelius T. Leondes. *Fuzzy Theory Systems, Volume 2.* Academic Press, 1999.

[5] Prof. Dmitry Panchenko. Functions of random values, 2005.

[6] Andrzej Piegat. *Modelowanie i sterowanie rozmyte.* Akademicka Oficyna Wydawnicza EXIT, 1999.

[7] Agnieszka Plucińska. *Probabilistyka.* Wydawnictwo Naukowo-Techniczne, 2000.