MWS LAB.1

Piotr Mikołajczyk

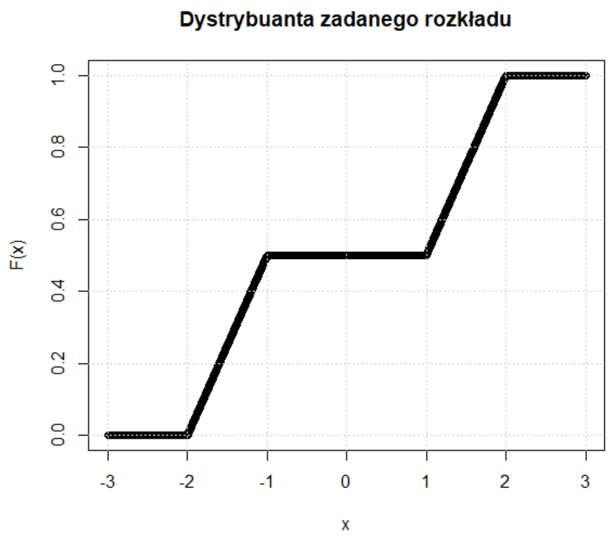
11.04.2021

Zad1

a) Wyznacz i narysuj dystrybuantę tego rozkładu

Na poniższych rysunkach przedstawiono zadaną funkcje gęstości rozkładu prawdopodobieństwa. Ponieważ zadana funkcja składa się z dwóch prostokątów o jednostajnej amplitudzie, o podstawie 1, przyjęto wartość 0.5, aby w konsekwencji otrzymać prawdopodobieństwo 1, dla całkowania całego zadanego rozkładu.

a)



b)

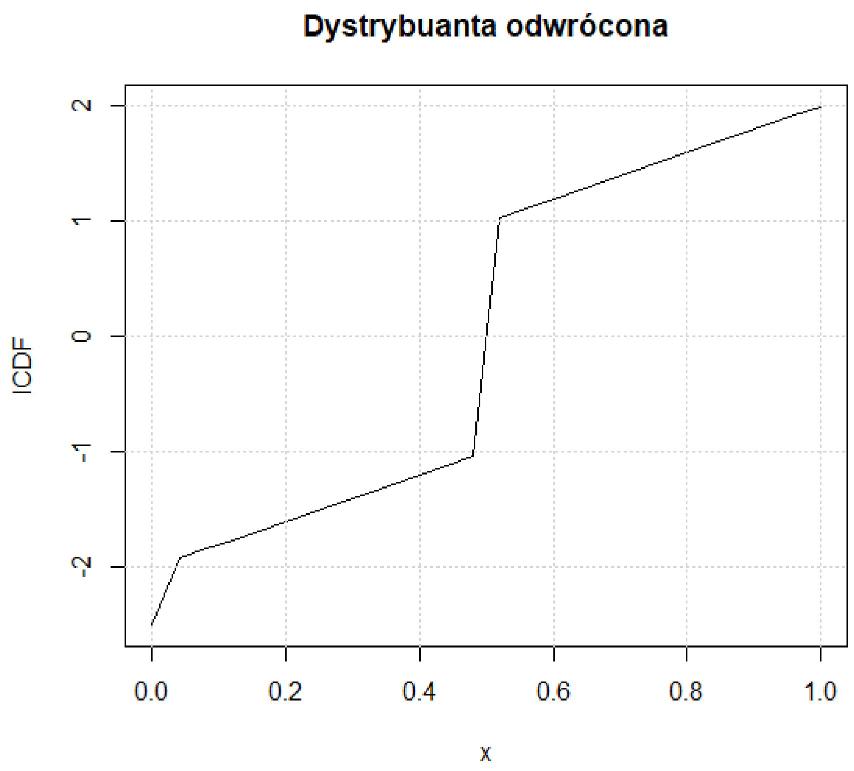
Rys.1 – a) Funkcja rozkładu gęstości prawdopodobieństwa , b) Dystrybuanta rozkładu z rysunku a)

Wnioski:

Kształt dystrybuanty jest typowy dla rozkładu jednostajnego ciągłego. Ze względu na to że rozkład podzielony jest na dwa równe prostokątne części rozkładu składający się łącznie na cały rozkład. W obszarze ich pola powierzchni, dystrybuanta rośnie, natomiast w obszarze gęstości prawdopodobieństwa wynoszącego 0, dystrybuanta nie zmienia swojej wartości i jest stała.

b) Wygeneruj 1000-elementową próbę losową z tego rozkładu

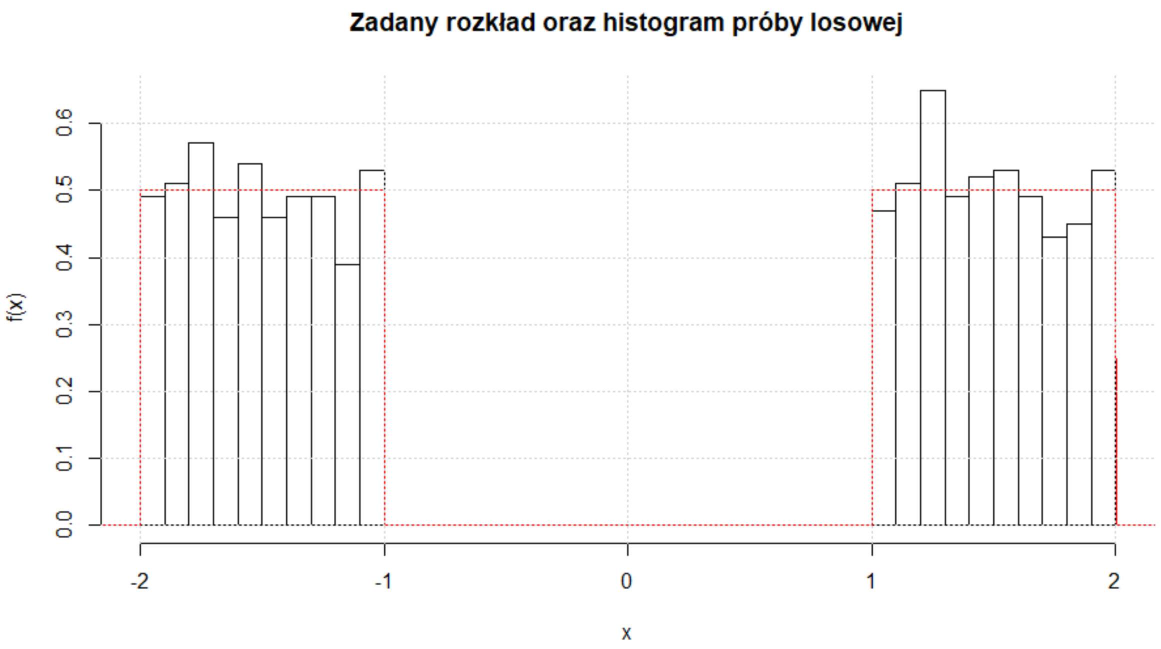
Do wygenerowania 1000 próbek w zakresie funkcji rozkładu prawdopodobieństwa, wykorzystano funkcje runif. Następnie w celu uzyskania próby losowej z tego rozkładu, wygenerowaną próbę randomową z zakresu 0 – 1, wykorzystano jako argument dla odwróconej dystrybuanty, którą uzyskano za pomocą funkcji approxfun , zamieniając oś y z x dystrybuanty pierwotnej zadanego rozkładu, uzyskując odpowiednio próbę losową.



Rys.2 – Dystrybuanta odwrócona

c) Narysuj histogram tej próby i porównaj go z wykresem funkcji gęstości prawdopodobieństwa.

Rysunek 3 przedstawia znormalizowane porównanie funkcji gęstości prawdopodobieństwa i wylosowane próbki z tego rozkładu w postaci histogramu.

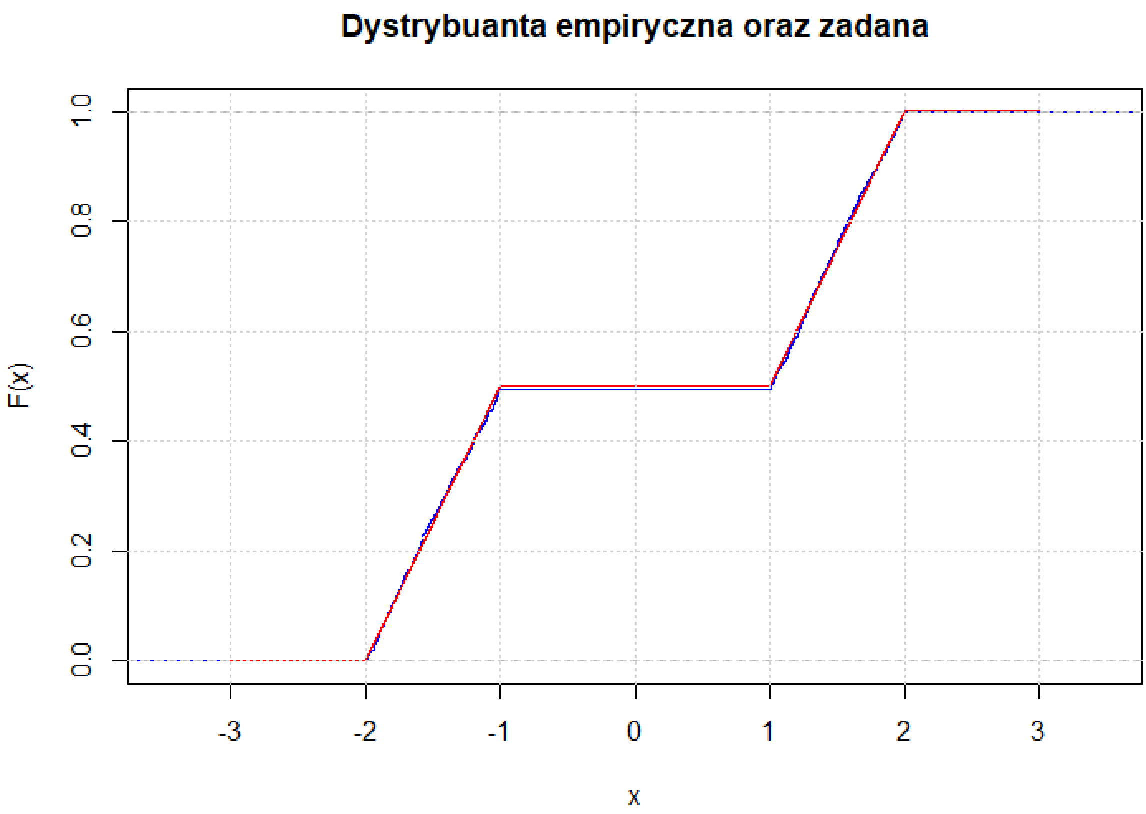


Rys.3 – Zadany rozkład prawdopodobieństwa oraz histogram próby losowej z zadanego rozkładu

Wnioski:

Na rysunku 3 przedstawiono czerwoną linią zadany w przykładzie pierwszym rozkład prawdopodobieństwa, natomiast czarną linią histogram próby losowej tego rozkładu. Widać że wylosowana próba zachowuje się prawidłowo, tzn. próbki istnieją tylko w przedziale w którym jest prawdopodobne ich wystąpienie. Mediana tych próbek wynosi średnio tyle, ile amplituda gęstości rozkładu oryginalnego. Przez zastosowanie funkcji approx do wygenerowania odwrotnej dystrybuanty, możliwe jest pojawienie się próbek po za granicami rozkładu ze względu na przybliżenie.

d) Dla wygenerowanej w poprzednim punkcie próby losowej narysuj wykres dystrybuanty empirycznej i porównaj go z dystrybuantą z punktu a.



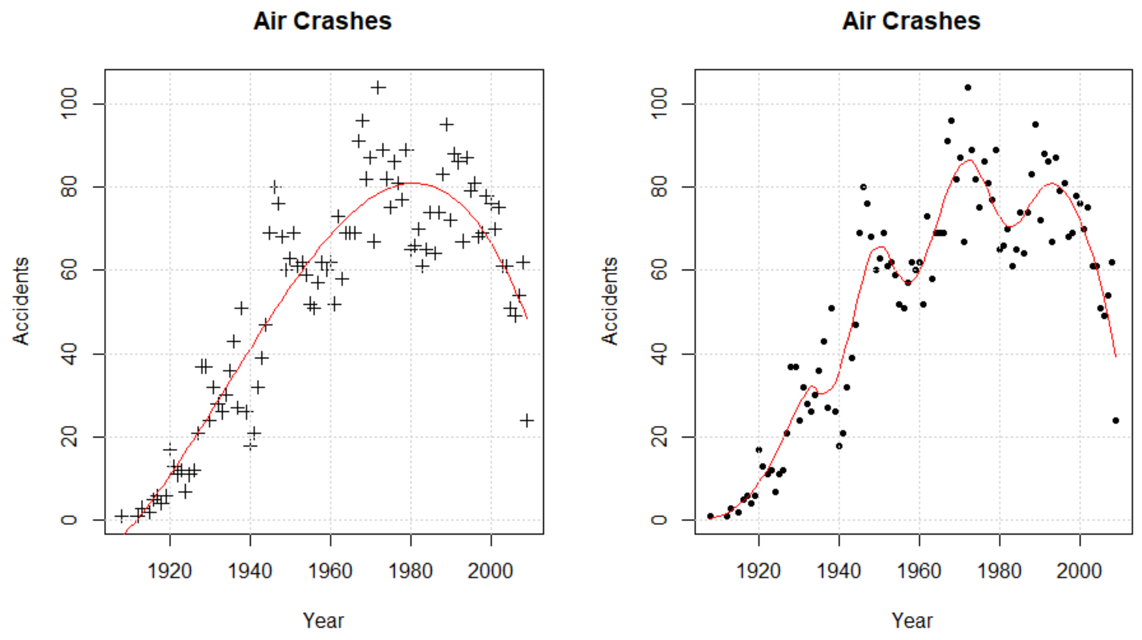
Rys.4 – Porównanie dystrybuanty empirycznej (niebieski kolor) oraz zadanego rozkładu (czerwony kolor).

Wnioski:

Dystrybuanta z próby losowej z rozkładu pokrywa się z dystrybuantą samego rozkładu gęstości prawdopodobieństwa. Dla coraz większej n-elementowej próby losowej rozkładu, kształt zbliżałby się coraz bardziej do kształtu dystrybuanty samego rozkładu.

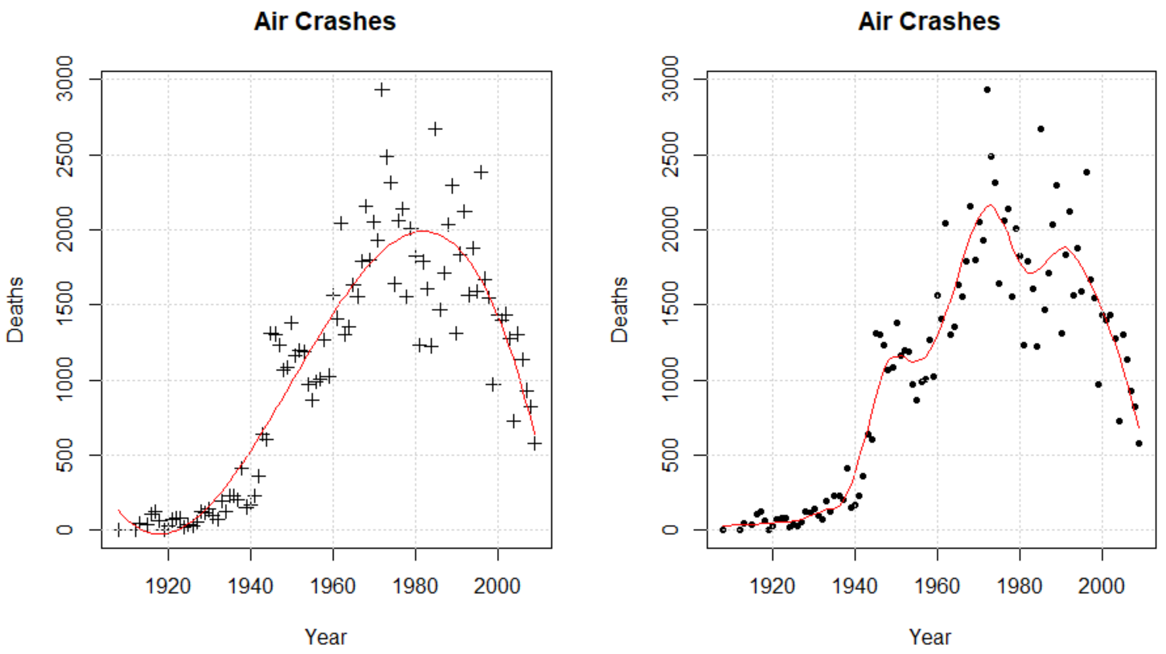
Zad2

a) b) Narysuj jak zmieniała się liczba wypadków i ich ofiar śmiertelnych w kolejnych latach. Do otrzymanych wykresów należy dorysować gładką krzywą ilustrującą te same zależności, jednak nie rozpraszające uwagi krótkotrwałymi wahaniami.



a) b)

Rys. 5 – a) Liczba wypadków lotniczych w latach uśredniona metodą regresji liniowej, b)) Liczba wypadków lotniczych w latach uśredniona metodą regresji lokalnej).



a) b)

Rys. 6 – a) Liczba śmierci przez wypadki lotnicze w latach uśredniona metodą regresji liniowej, b)) Liczba śmierci przez wypadki lotnicze w latach uśredniona metodą regresji lokalnej).

Wnioski:

Na rysunku 5a) oraz 6a) wykorzystano metodę aproksymacji wielomianowej (funkcja lm) – w tym wypadku 3 rzędu, do wyrysowania krzywej ukazującej przybliżoną średnią z wypadków lotniczych na przestrzeni lat, natomiast na rysunku 5b) oraz 6b), wykorzystano inną funkcje również do aproksymacji wielomianowej z tym że lokalnej (loess oraz predict) – pozwoliło to narysować bardziej szczegółową krzywą niż w przypadku pierwszej funkcji. Pierwsza metoda zachowuje się bardziej globalnie, dzięki czemu uwidacznia bardziej trend zachowania się funkcji, pozwalając sobie jednak na lokalnie większe odchylenia od wartości prawdziwych. Druga funkcja, ze względu na jej lokalny charakter, jest bardziej odporna na zmiany, odczytywanie z niej wartości jest bliższe prawdy natomiast mniej uwypuklony jest w niej trend. Całościowo funkcja loess, wydaje się być właściwsza do zadanych danych katastrof lotniczych.