
Raport 2
Symulacje Komputerowe

Natalia Lach 262303, Alicja Myśliwiec 262275

Matematyka Stosowana
Wydział Matematyki Politechniki Wrocławskiej

Spis treści

1. Wstęp	2
2. Zadanie 1	2
2.1. Cel zadania	2
2.2. Teoria	2
2.2.1. Proces ryzyka	2
2.2.2. Jednorodny proces Poissona	3
2.2.3. Opis algorytmu symulacji jednorodnego procesu Poissona	3
2.2.4. Algorytm generowania procesu ryzyka	4
2.2.5. Moment ruiny	4
2.2.6. Algorytm wyznaczania prawdopodobieństwa ruiny w skończonym czasie	5
2.2.7. Wzór Pollaczka-Chinczyna	5
2.2.8. Algorytm wyznaczania prawdopodobieństwa ruiny w nieskończonym czasie	5
2.3. Część I	6
2.4. Część II	12
3. Zadanie 2	13
3.1. Cel zadania	13
3.2. Teoria	13
3.2.1. Proces Wienera/Ruch Browna	13
3.2.2. Ruch Browna startujący z x	13
3.2.3. Prawo iterowanego logarytmu	13
3.2.4. Średni czas wyjścia z przedziału $[a, b]$	14
3.2.5. Algorytm generowania procesu Wienera	14
3.3. Wizualizacja trajektorii	14
3.4. Oszacowanie średniego czasu wyjścia z przedziału $[a, b]$	15
3.5. Oszacowanie prawdopodobieństwa wyjścia przez granicę b	16
4. Podsumowanie	17

1. Wstęp

W niniejszym raporcie przedstawione zostaną pojęcia procesu Ryzyka oraz ruchu Browna (inaczej też nazywanego procesem Wienera). Przedstawione zostaną odpowiednie algorytmy oraz pojęcia teoretyczne, dzięki którym możliwe będzie rozwiązanie postawionych przed nami zadań dotyczących wyżej wymienionych procesów. Wszelkie symulacje zostały wykonane w języku programowania Python.

2. Zadanie 1

2.1. Cel zadania

W części pierwszej zadania, celem było dopasowanie modelu klasycznego procesu Ryzyka do otrzymanych danych. Należało zatem odnaleźć wszystkie parametry procesu, w tym zastanowić się z jakiego rozkładu pochodzą zmienne losowe odpowiadające wielkości strat. Ponadto trzeba było odnaleźć wartość intensywności jednorodnego procesu Poissona λ oraz parametru θ , który odpowiada za wysokość premii. W ramach zadania opisano dokładnie schemat szukania wyżej wymienionych parametrów oraz algorytm symulowania procesu Ryzyka, a także jednorodnego procesu Poissona.

W części drugiej zadania należało oszacować prawdopodobieństwo ruiny otrzymanego procesu Ryzyka dla dwóch skończonych horyzontów czasowych, a także w nieskończonym czasie. Zadanie wykonane zostało na podstawie wzoru Pollaczka–Chinczyna oraz symulacji aproksymujących ów prawdopodobieństwa.

2.2. Teoria

2.2.1. Proces ryzyka

Proces ryzyka to proces stochastyczny opisujący kapitał firmy ubezpieczeniowej. Jest on dany wzorem

$$R(t) = u_0 + c(t) - \sum_i^{N(t)} X_i, \quad (1)$$

gdzie

- u_0 - kapitał początkowy,
- $c(t)$ - funkcja premii (czyli przychody ze sprzedaży polis),
- $N(t)$ - proces liczący straty,
- X_i zmienne losowe iid. reprezentujące wysokość wypłacanych odszkodowań (strat dla firmy).

Ważne założenia: $\mathbb{E}X_i = \mu$, $X_i > 0$, $X_i \perp\!\!\!\perp N(t)$.

Można także mówić o tak zwanym klasycznym procesie Ryzyka. Jest to taki proces, dla którego funkcja premii jest funkcją liniową, to jest

$$R(t) = u_0 + ct - \sum_i^{N(t)} X_i, \quad (2)$$

z oznaczeniami jak wyżej, z czego w tym wypadku $c = \text{const}$ jest dodatnią stałą określającą wielkość liniowego przyrostu dochodów. Ów stałą będziemy odtąd nazywać po prostu premią.

Co więcej, rozważaną stałą można wyrazić za pomocą zależności

$$c = (1 + \theta)\mu\lambda, \quad (3)$$

gdzie θ - tak zwany narzut, μ - wartość oczekiwana zmiennych X_i odpowiadających i-tej szkodzi oraz λ - intensywność jednorodnego procesu Poissona. Gdy $X_i \sim \text{Exp}(\beta)$ oraz $\mathbb{E}X_i = \frac{1}{\beta}$, wzór ten można przekształcić do następującej postaci.

$$c = (1 + \theta)\frac{\lambda}{\beta} \quad (4)$$

2.2.2. Jednorodny proces Poissona

Aby poprawnie przeprowadzić symulację procesu ryzyka, należy wiedzieć czym jest jednorodny proces Poissona z intensywnością λ oraz jak go generować. Jedną z definicji jest następująca.

Jednorodnym procesem Poissona z intensywnością $\lambda > 0$ nazywamy proces liczący $\{N(t), t \geq 0\}$, który spełnia następujące warunki.

1. $N(0) = 0$
2. $N(t)$ ma niezależne przyrosty
3. $N(t)$ ma stacjonarne przyrosty
4. $N(t) \sim \mathcal{Pois}(\lambda t)$

Jedną z ważniejszych własności ów procesu dotyczy czasów oczekiwania na kolejny skok, a mianowicie dla jednorodnego procesu Poissona z intensywnością λ zmienne losowe T_i , oznaczające czasy oczekiwania na kolejny skok, są iid. oraz

$$T_i \sim \mathcal{Exp}(\lambda). \quad (5)$$

Drugą, równie ważną własnością procesu, która ułatwia przeprowadzanie jego symulacji, jest własność dotycząca samych momentów skoków.

Warunkowy rozkład wektora momentów skoków na odcinku $[0, T]$ (gdzie T - horyzont czasowy) pod warunkiem $N(t) = I$ jest równy rozkładowi wektora statystyk pozycyjnych $(U_{1:I}, U_{2:I}, \dots, U_{I:I})$, gdzie U_1, \dots, U_I to zmienne losowe iid. o rozkładzie jednostajnym $U_i \sim \mathcal{U}(0, 1)$.

Każda z własności przyda się w późniejszych częściach projektu.

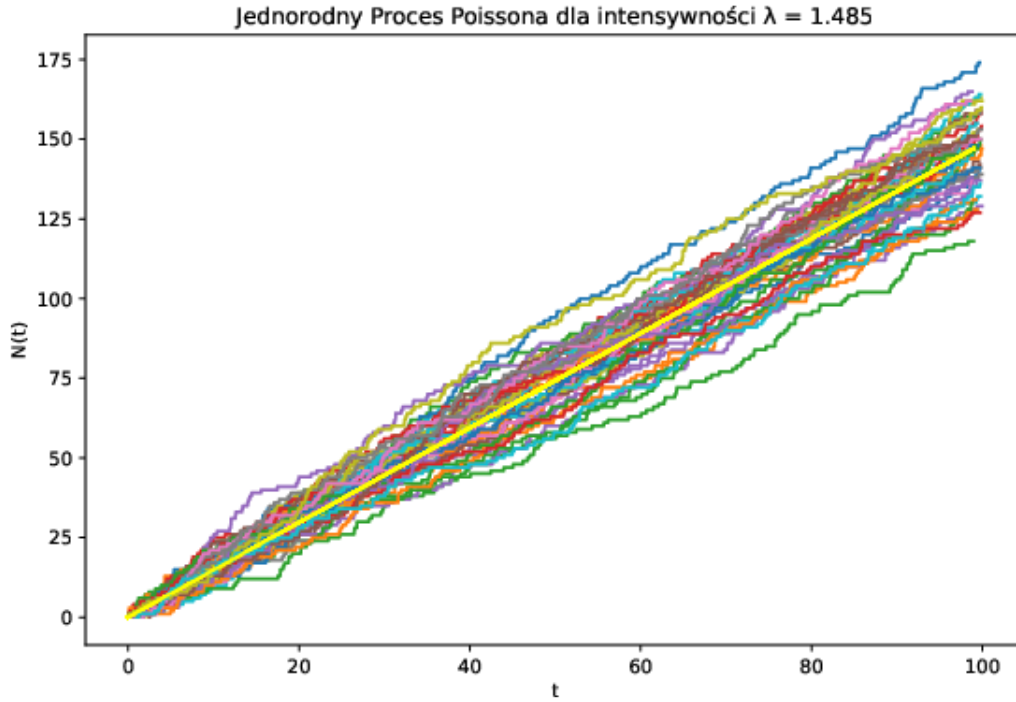
2.2.3. Opis algorytmu symulacji jednorodnego procesu Poissona

Wykorzystując własności jednorodnego procesu Poissona możemy skonstruować algorytm jego symulowania. Załóżmy, że mamy daną intensywność $\lambda > 0$ oraz horyzont czasowy T . Schemat postępowania jest następujący.

1. Wygeneruj ilość skoków $I \sim \mathcal{Pois}(\lambda T)$
2. Jeśli $I = 0$ zakończ procedurę. W przeciwnym razie wygeneruj U_1, \dots, U_I , gdzie U_i muszą być iid. oraz $U_i \sim \mathcal{U}(0, T)$.
3. Posortuj U_1, \dots, U_I , otrzymując statystykę pozycyjną $U_{1:I}, U_{2:I}, \dots, U_{I:I}$
4. Wstaw $S_i = U_{i:I}$, $i = 1, 2, \dots, I$. Zwróć wektor momentów skoków (S_1, \dots, S_I) .
5. Na podstawie wektora momentów skoków stwórz wektor o tej samej długości, przedstawiający ilość skoków, która wystąpiła do danego momentu skoku, to znaczy $[1, 2, \dots, I]$.
6. Używając funkcji schodkowej *step*, narysuj jednorodny proces Poissona.

Jest to oczywiście jedna z wielu metod generowania rozważanego procesu. Została ona wybrana ze względu na jej prostotę wykonania. Aby przetestować powyższy algorytm, wysymulowano 50 trajektorii procesu o intensywności $\lambda = 1.485$ oraz przedstawiono je na wykresie. Ponadto naniesiono średnią procesu. Jest wartość wynosi $\mathbb{E}[N(t)] = \lambda t = 1.485t$. Powyższy wzór wynika z faktu, że $N(t) \sim \mathcal{Pois}(\lambda t)$. Ponadto w późniejszych częściach projektu, to jest w sekcji 2.3, zostanie także uzasadnione, czemu wzięto taką a nie inną wartość λ .

Wykres wysymulowanych trajektorii przedstawiono poniżej.



Rys. 1: Trajektorie jednorodnego procesu Poissona z $\lambda = 1.485$ oraz naniesioną średnią procesu.

Wykres sugeruje, że podany wcześniej algorytm rzeczywiście generuje jednorodny proces Poissona o danej intensywności, gdyż każda z trajektorii oscyluje wokół średniej procesu.

2.2.4. Algorytm generowania procesu ryzyka

Algorytm generowania procesu Ryzyka opiera się w głównej mierze na generowaniu jednorodnego procesu Poissona. Korzystając z przykładowego algorytmu przedstawionego w sekcji 2.2.3, można w łatwy sposób uzyskać wartości $R(t)$ w kolejnych momentach czasu.

1. Generuj jednorodny proces Poissona z intensywnością λ na odcinku $[0, T]$.
2. Generuj $X_i, i = 1, \dots, N(t)$ z odpowiedniego rozkładu strat.
3. Korzystając ze wzoru 1., wyznacz kolejne wartości procesu ryzyka.

2.2.5. Moment ruiny

W części drugiej zadania wykorzystywane zostaną poniższe wzory. Moment ruiny, zdefiniowany jako pierwszy moment, w którym wartość $R(t)$ osiągnęła wartość poniżej zera, czyli

$$\tau(u) = \inf\{t \geq 0 : R(t) < 0\}. \quad (6)$$

Na tej podstawie można zdefiniować prawdopodobieństwo ruiny w skończonym czasie, czyli

$$\Psi(u, T) = P(\tau(u) < T), \quad (7)$$

gdzie T to horyzont czasowy.

Dodatkowo rozpatrywać można także prawdopodobieństwo ruiny w nieskończonym czasie, czyli

$$\Psi(u) = P(\tau(u) < \infty). \quad (8)$$

2.2.6. Algorytm wyznaczania prawdopodobieństwa ruiny w skończonym czasie

Szacowanie prawdopodobieństwa ruiny dla skończonego horyzontu czasowego jest bardzo proste i opiera się na symulacji Monte Carlo. Schemat postępowania jest następujący.

1. Generuj N trajektorii klasycznego procesu ryzyka na $[0, T]$, oznaczanych jako $R^1(t), \dots, R^N(t)$, gdzie N - liczba symulacji Monte Carlo.
2. Wyznacz ilość procesów, podczas których wartość $R(t)$ spadła poniżej 0, to znaczy $n = \{i \in \{1, \dots, N\} : \min_{t \in [0, T]} R^i(t) < 0\}$
3. Wstaw $\Psi(u, T) = \frac{n}{N}$

2.2.7. Wzór Pollaczka-Chinczyna

Podczas aproksymacji prawdopodobieństwa ruiny w nieskończonym czasie, można skorzystać ze wzoru Pollaczka-Chinczyna podanego poniżej.

$$\Psi(u) = \frac{\theta}{\theta + 1} \sum_{n=0}^{\infty} \left(\frac{1}{1 + \theta} \right)^n B_n(u), \quad (9)$$

gdzie $B_n(u) = P(Y_1 + \dots + Y_n > u)$ oraz Y_i to zmienne losowe iid. o gęstości $f(x) = \frac{1 - F_X(x)}{\mathbb{E}X}$, z czego X_i to zmienne losowe przedstawiające wartość i -tej szkody w procesie ryzyka oraz $F_X(x)$ - dystrybuanta zmiennych losowych X_i . Gdy X_i są z rozkładu wykładniczego o parametrze β , wzór ten można przekształcić do jawnej postaci.

$$\Psi(u) = \frac{1}{1 + \theta} \exp \left(-\beta \mu \frac{\theta}{1 + \theta} \right) \quad (10)$$

Z powyższego wzoru będziemy korzystać w sekcji 2.4.

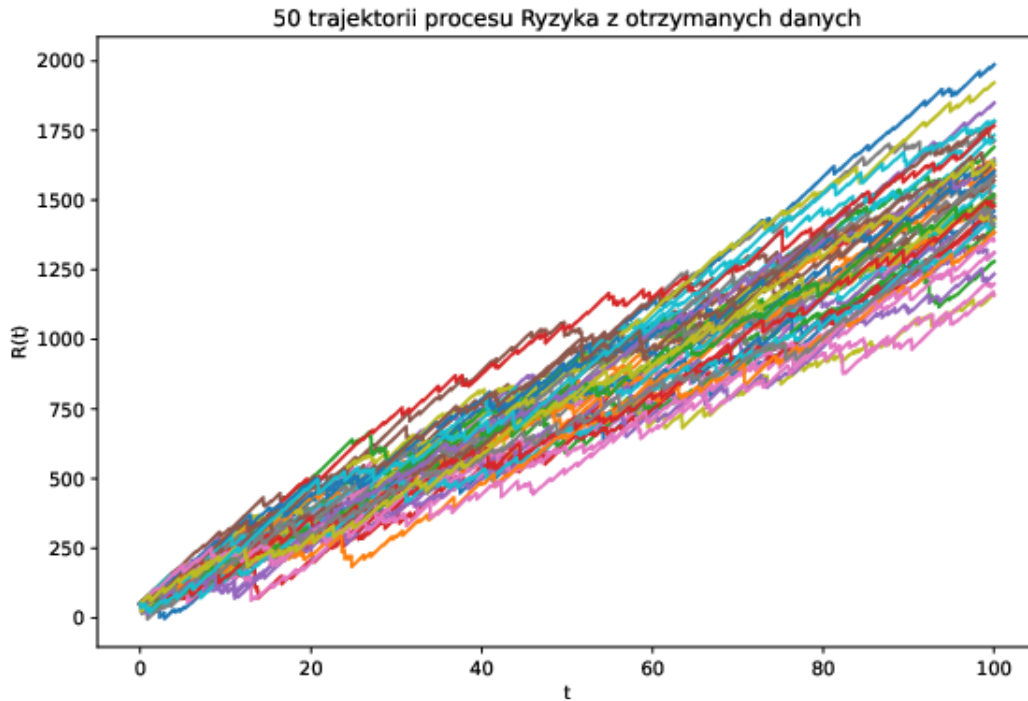
2.2.8. Algorytm wyznaczania prawdopodobieństwa ruiny w nieskończonym czasie

Szacowanie prawdopodobieństwa ruiny dla nieskończonego horyzontu czasowego nie jest już tak proste i intuicyjne, jak ów proceder dla pewnego skończonego horyzontu czasowego. Algorytm symulowania rozważanego zagadnienia opiera się w dużej mierze na wzorze Pollaczka-Chinczyna oraz symulacjach Monte Carlo. Schemat postępowania jest następujący.

1. Generuj $K \sim \mathcal{Geom}(\frac{\theta}{1+\theta})$.
2. Generuj Y_1, \dots, Y_K - zmienne losowe iid. o gęstości ze wzoru (9), czyli $f(x) = \frac{1 - F_X(x)}{\mathbb{E}X}$
3. Jeśli $Y_1 + \dots + Y_K > u$, wstaw $Z(i) = 1$, w przeciwnym razie $Z(i) = 0$.
4. Powtórz kroki (1)-(3) N razy (N powtórzeń Monte Carlo).
5. Wstaw $\Psi(u) = \frac{Z(1) + \dots + Z(N)}{N}$

2.3. Część I

Do wykonania zadania otrzymano zestaw danych zawierających 50 trajektorii procesu Ryzyka, wysymulowanych na odcinku $[0,100]$ z krokiem czasowym $h = 10^2$, z pewnymi parametrami β, λ, θ . Wszystkie trajektorie przedstawiono na poniższym wykresie.



Rys. 2: Trajektorie procesu Ryzyka na podstawie otrzymanych danych

Należy teraz odnaleźć wartości powyższych parametrów. Wpierw odszukajmy stałą c widoczną we wzorze 1, zwaną premią.

Każda z trajektorii ma w sobie pewną ilość odcinków, na których następuje liniowy wzrost procesu. Nie występują wtedy żadne straty. Jeśli więc rozważymy początkowy bieg trajektorii aż do wystąpienia pierwszej straty, możemy otrzymać wyrażenie na prostą funkcję liniową.

$$R(t) = u_0 + ct$$

Na tej podstawie łatwo jest odnaleźć szukaną wartość premii, gdyż wystarczy skorzystać ze wzoru

$$c = \frac{R(t+h) - R(t)}{h},$$

gdzie h to czas pomiędzy dwoma kolejnymi krokami czasowymi, a t jest mniejsze od momentu wystąpienia pierwszego skoku.

Każdy z procesów w czasie $t = 0$ osiąga wartość 50, a w następnych momentach kolejno $R(0.01) = 50.3$, $R(0.02) = 50.6$ i tak dalej aż do wystąpienia pierwszego skoku. Na tej podstawie można zatem wyliczyć szukaną wartość premii c i wynosi ona

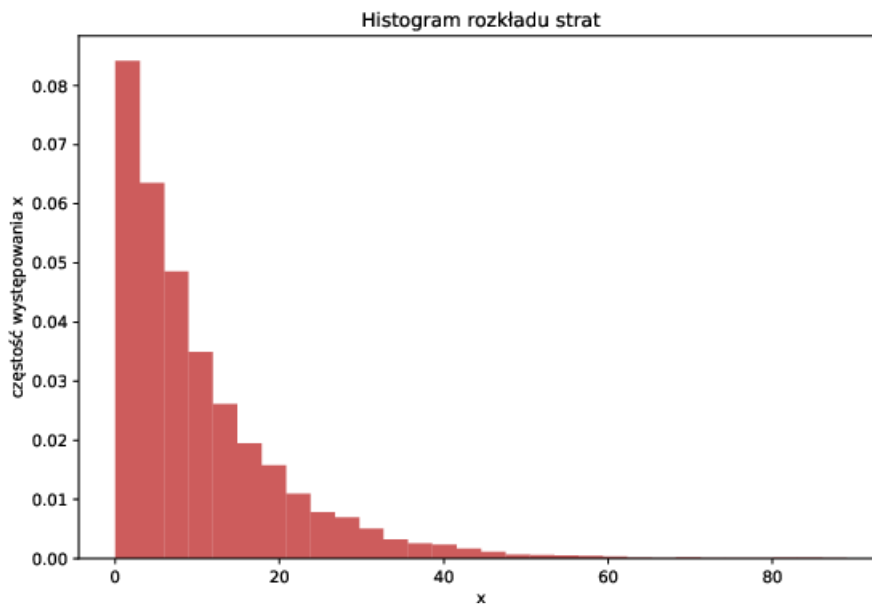
$$c = 30.$$

Następnym krokiem poszukiwania wartości parametrów rozważanego procesu Ryzyka niech będzie odnalezienie rozkładu strat, to znaczy rozkładu zmiennej losowej X_i , oznaczającej wielkość i -tej szkody. Aby to zrobić, należy najpierw uzyskać dane dotyczące wielkości strat w otrzymanych trajektoriach procesu Ryzyka.

Schemat postępowania dla każdej z trajektorii przedstawiony jest poniżej.

1. Stwórz pustą listę, w której przechowywane będą wartości strat.
2. W pętli po wszystkich wartościach procesu sprawdź, czy $R(t+h) = R(t) + ch$, gdzie h - krok czasowy.
3. Jeśli równość jest prawdziwa - przejdź do kolejnego kroku w pętli.
Gdy równość jest nieprawdziwa - wartość skoku w tym momencie jest równa $R(t) - R(t+h) + ch$.
Dodaj ją do stworzonej wcześniej listy. Przejdź do kolejnego kroku pętli.
4. Po przejściu przez całą pętlę, zwróć listę wartości strat.

Postępując zgodnie ze schematem otrzymano zbiór wartości strat ze wszystkich 50 trajektorii procesu Ryzyka dla $h = 10^2$ oraz z pewnymi parametrami β, λ, θ . Wszystkie trajektorie przedstawiono na poniższym wykresie.



Rys. 3: Histogram rozkładu strat pozyskanych z otrzymanych danych

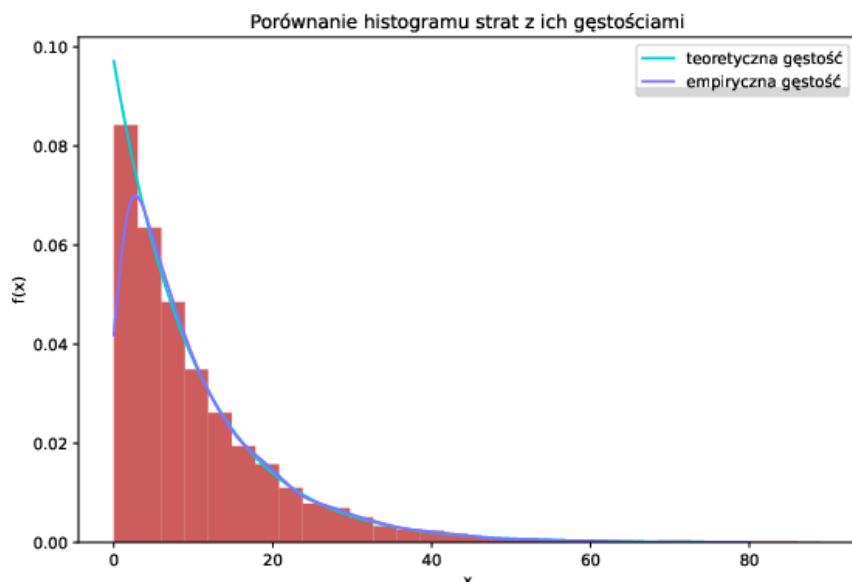
Powyższy wykres ma postać przypominającą pewien inny bardzo znany rozkład - rozkład wykładniczy z parametrem β . Aby oszacować ów parametr, można skorzystać z estymatora wyznaczonego metodą momentów danego zależnością

$$\hat{\beta} = \frac{1}{\frac{1}{n} \sum_i^n X_i}. \quad (11)$$

Po zastosowaniu powyższego wzoru otrzymano

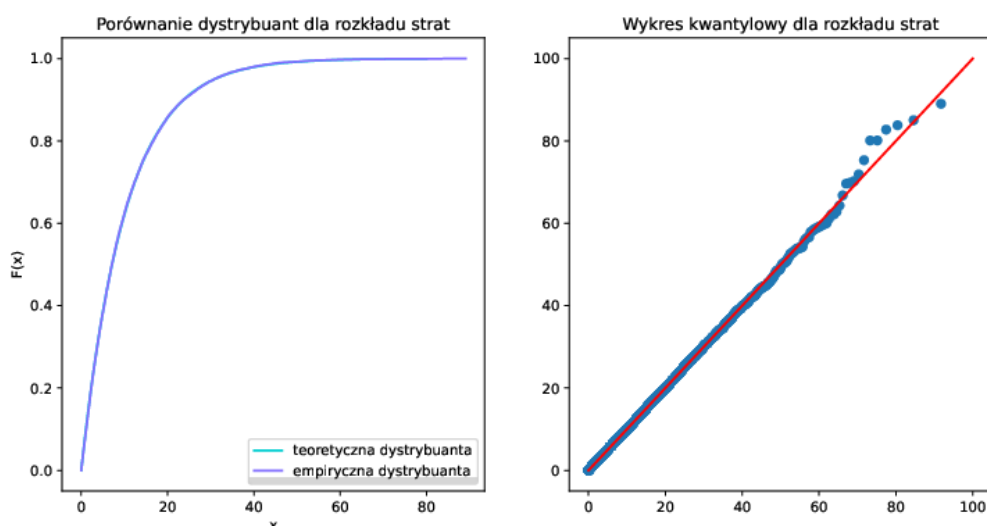
$$\beta \approx 0.097.$$

Aby teraz sprawdzić postawioną hipotezę, zestawiono ze sobą histogram częstości oraz gęstości teoretyczną i empiryczną rozważanego rozkładu.



Rys. 4: Histogram rozkładu strat pozyskanych z otrzymanych danych wraz z gęstościami empiryczną i teoretyczną

Wykres sugeruje, że straty są rzeczywiście zmiennymi losowymi z rozkładu wykładniczego o parametrze $\beta = 0.097$. Teoretyczna gęstość prawie idealnie przybliża wartości histogramu. Gęstość empiryczna w okolicach $x \approx 5$ zaczyna się trochę oddalać od swojego teoretycznego odpowiednika, lecz wynika to z niedokładności metody numerycznej stosowanej do jej wyznaczenia. Poza tym odcinkiem jest ona zbieżna z resztą porównywanych z nią funkcji. W ramach upewnienia się co do poprawności postawionej hipotezy wygenerowano także porównanie dystrybuant empirycznej z teoretyczną oraz wykres kwantylowy.



Rys. 5: Porównanie dystrybuant oraz wykres kwantylowy dla rozkładu strat

I tym razem wykresy sugerują że dane pochodzą z rozważanego rozkładu wykładniczego. Dystrybuanty empiryczna i teoretyczna są niemal identyczne. Wykres kwantylowy także potwierdza nasze założenia. Poza kilkoma wartościami odstającymi dla dalszych wartości, kwantyle teoretyczne i empiryczne są ze sobą niemal równe.

Ponadto wykonano test Kołmogorowa-Smirnowa, aby ostatecznie potwierdzić, czy dane rzeczywiście pochodzą z rozkładu wykładniczego. Skorzystano z wbudowanych funkcji jednego z modułów języka Python i przeprowadzono symulację dla p -wartości rozważanej hipotezy. Otrzymano wynik

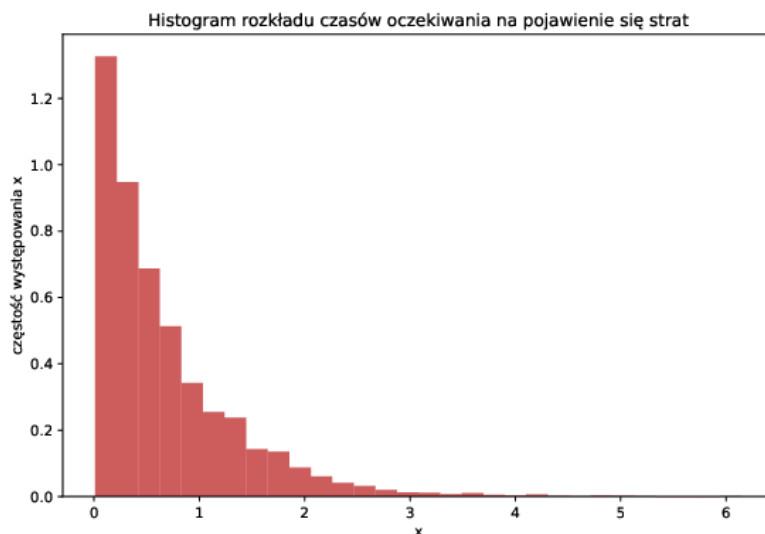
$$p\text{-wartość} \approx 0.83.$$

Jej wartość jest większa od 0.1, co pozwala nam zaakceptować hipotezę zerową, to znaczy potwierdza postawioną przez nas hipotezę o pochodzeniu danych. Wysunąć można zatem wniosek, że rozkład strat rozważanych procesów Ryzyka jest wykładniczy z parametrem $\beta = 0.097$ oraz z jego własności $\mathbb{E}X_i = \frac{1}{\beta} = 10.3$.

Skoro uzyskaliśmy postać rozkładu strat procesu Ryzyka, to kolejnym krokiem może być znalezienie parametru λ odpowiadającego intensywności jednorodnego procesu Poissona. Aby to zrobić, skorzystamy z własności ów procesu, która mówi o tym, że czasy oczekiwania na kolejny skok są z rozkładu wykładniczego o parametrze λ . Wpierw należy wyznaczyć momenty pojawiania się strat w każdej trajektorii procesu Ryzyka, a później czasy oczekiwania pomiędzy każdym z nich. Schemat postępowania jest bardzo podobny do szukania wartości strat.

1. Stwórz pustą listę, w której przechowywane będą wartości czasów oczekiwania na kolejny skok.
2. W pętli po wszystkich wartościach procesu sprawdź, czy $R(t+h) = R(t) + ch$, gdzie h - krok czasowy.
3. Jeśli równość jest prawdziwa - przejdź do kolejnego kroku w pętli.
Gdy równość jest nieprawdziwa - dodaj do listy moment pojawienia się i -tej straty t_i , czyli $t_i = t + h$. Przejdź do kolejnego kroku pętli.
4. Po przejściu przez całą pętlę, otrzymano listę z momentami pojawienia się strat t_i . Stwórz kolejną pustą listę, służącą do przechowywania danych dotyczących czasów oczekiwania na dany skok T_i .
5. W kolejnej pętli (po właśnie co wyznaczonych momentach strat) wyznacz czas oczekiwania $T_i = t_{i+1} - t_i$. Dodaj wartość do utworzonej listy.
6. Zwróć listę z czasami oczekiwania na pojawienie się kolejnych strat.

Ponownie, korzystając ze schematu, uzyskano dane dotyczące czasów oczekiwania na kolejne straty we wszystkich 50 trajektoriach procesu Ryzyka i przedstawiono je w postaci histogramu częstości.

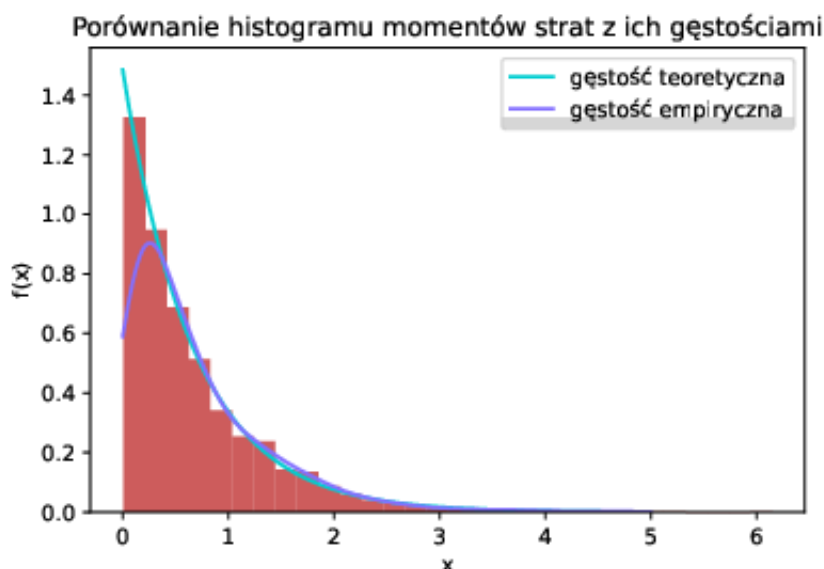


Rys. 6: Histogram częstości dla rozkładu czasów oczekiwania

Tym razem wiadomym jest fakt, że dane pochodzą z rozkładu wykładniczego z pewnym parametrem λ . Korzystając ponownie z estymatora danego wzorem 11, uzyskano wartość parametru i wynosi ona

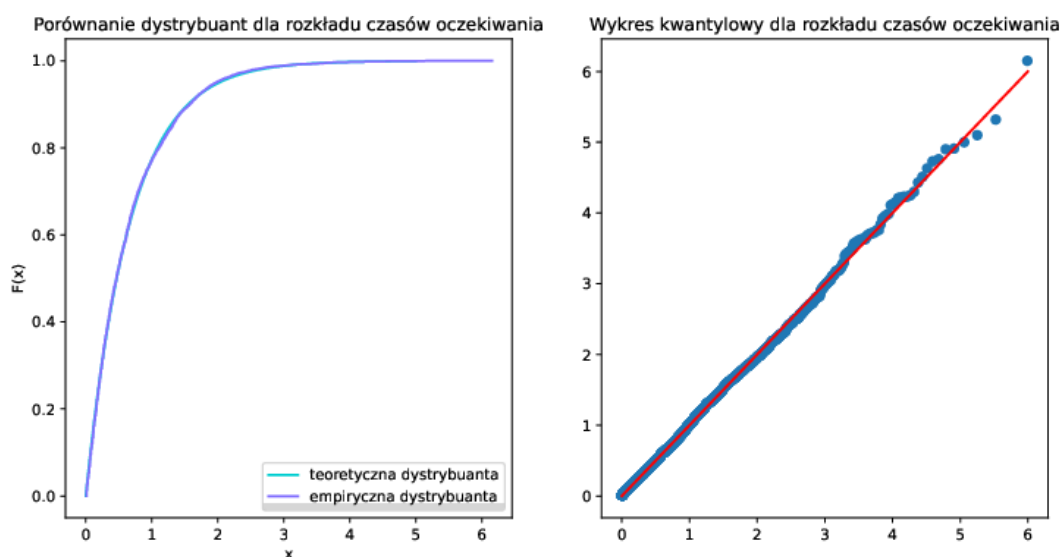
$$\lambda \approx 1.485.$$

Wykorzystując tę wiedzę, wygenerowano wykresy przedstawiające ponownie histogram częstości, ale tym razem z naniesionymi gęstościami teoretyczną i empiryczną rozpatrywanego rozkładu.



Rys. 7: Porównanie gęstości oraz histogramu dla rozkładu czasów oczekiwania

I tym razem powyższy wykres sugeruje poprawność naszych rozważań. Gęstość teoretyczna jest bardzo zbliżona do swojego empirycznego odpowiednika, jakim jest histogram. Jedyna rozbieżność występuje na pewnym początkowym odcinku dla gęstości empirycznej, lecz ponownie wynika ona z błędów numerycznych. Poza tym jest ona niemal identyczna w porównaniu do reszty. Wygenerowano także zestawienie dystrybuant oraz wykres kwantylowy.



Rys. 8: Porównanie dystrybuant oraz wykres kwantylowy dla rozkładu czasów oczekiwania

Ponownie, wykresy sugerują poprawność postawionej przez nas hipotezy. Obie dystrybanty są niemal identyczne, a wykres kwantylowy układu się liniowo.

Z powyższych rozważań i otrzymanych wyników można zatem wysunąć wniosek, że czasy oczekiwania na kolejne pojawienie się strat w omawianym procesie Ryzyka są z rozkładu wykładniczego o parametrze $\lambda = 1.485$. Ze wspomnianej wcześniej własności uzyskujemy tym samym wartość intensywności jednorodnego procesu Poissona, czyli

$$\lambda = 1.485.$$

Ostatnim krokiem do wykonania zadania, jest znalezienie wartości parametru θ odpowiadającego za narzut w procesie Ryzyka. Korzystając z otrzymanych wcześniej wyników oraz ze wzoru 11, możemy łatwo wyznaczyć tę wartość, gdyż jest ona dana zależnością

$$\theta = \frac{c\beta}{\lambda} - 1.$$

Podstawiając znane nam już wartości otrzymujemy

$$\theta \approx 0.96.$$

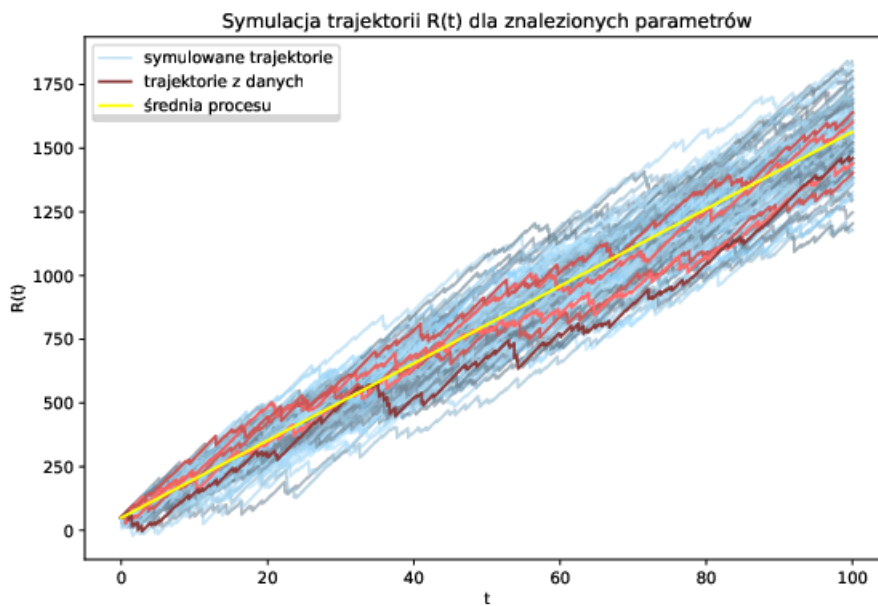
Uzyskaliśmy tym samym wszystkie interesujące nas wartości parametrów rozważanego procesu Ryzyka. Wynoszą one kolejno

$$c = 30, \quad \beta = 0.097, \quad \lambda = 1.485, \quad \theta = 0.96.$$

Teraz należało wysymulować 100 trajektorii procesu Ryzyka ze znalezionymi przez nas parametrami i porównać je z kilkoma trajektoriami spośród tych otrzymanych, a także średnią procesu. W tym celu wyznaczmy zatem wartość oczekiwaną procesu $R(t)$.

$$\mathbb{E}[R(t)] = \mathbb{E}\left[u_0 + ct - \sum_i^{N(t)} X_i\right] = u_0 + ct - \mu\lambda t = u_0 + t(c - \mu\lambda) \quad (12)$$

Uzyskaliśmy w ten sposób wzór funkcji liniowej przybliżającej średnią procesu. Podstawiając znane nam wartości parametrów do wzoru 12 i algorytmu generowania procesu Ryzyka, przeprowadzono symulację i otrzymano następujące wyniki.



Rys. 9: Symulacja 100 trajektorii procesu Ryzyka, wraz ze średnią procesu oraz kilkoma trajektoriami z podanych danych.

Wykres na rysunku 9 sugeruje, że udało się poprawnie wyznaczyć wszystkie parametry otrzymanego procesu Ryzyka. Wszystkie trajektorie oscylują wokół średniej procesu i są one do siebie dość zbliżone.

2.4. Część II

Zanim przejdziemy do szacowania prawdopodobieństwa ruiny symulowanych procesów ryzyka, znajdziemy ów prawdopodobieństwo dla otrzymanego zestawu danych. Schemat postępowania jest analogiczny do tego podanego w sekcji 2.2.6 z pominięciem pierwszego kroku. Otrzymano

$$\Psi(u, T) = 0.04.$$

W naszym przypadku dane składały się z $N = 50$ trajektorii oraz kapitał początkowy wynosił $u = 50$, a $T = 100$.

Przeprowadźmy teraz własne symulacje trajektorii procesu ryzyka z wyznaczonymi w poprzedniej sekcji parametrami. Oszacowano prawdopodobieństwo ruiny na skończonym horyzoncie czasowym $T_1 = 100$ oraz $T_2 = 200$ za pomocą algorytmu z sekcji 6. Za liczbę powtórzeń Monte Carlo wzięto kilka wartości $N = 50$, $N = 100$, $N = 1000$, $N = 10000$ i wyniki przedstawiono w poniższej tabeli.

	$T_1 = 100$	$T_2 = 200$
$N = 50$	0.04	0.04
$N = 100$	0.04	0.05
$N = 1000$	0.055	0.046
$N = 10000$	0.047	0.048

Tab. 1: Tabela szacowanych wartości prawdopodobieństwa ruiny dla różnej ilości powtórzeń Monte Carlo na skończonych przedziałach czasowych

Wyniki w tabeli dla $T = 100$ oraz $N = 50$ powtórzeń Monte Carlo pokrywają się z wynikiem dla otrzymanych wcześniej danych, to jest

$$\Psi(u, T) = 0.04.$$

Świadczy to o poprawności wykonanych obliczeń i przeprowadzonych symulacji. Ponadto możemy zauważyć, że wartości Ψ wahają się pomiędzy 0.4 a 0.5. Dla kolejnych wartości N , Ψ osiąga dość zbliżone wartości dla obu horyzontów czasowych. Może to świadczyć o tym, że następuje na tyle duży stały wzrost, że jeśli firma nie zbankrutuje na początku swojego istnienia, to jest mała szansa na to, żeby w przyszłości wygenerowało się na tyle strat, by jej fundusze spadły poniżej zera po $T_1 = 100$.

Oszacowano także wartość prawdopodobieństwa dla nieskończonego horyzontu czasowego. W tym przypadku skorzystano ze wzoru Pollaczka-Chinczyna, a dokładniej ze wzoru 10, gdyż rozkład strat w otrzymanych danych okazał się być wykładniczy. Wartość szacowanego prawdopodobieństwa wynosi

$$\Psi(u) \approx 0.0474.$$

Jest to wartość bardzo zbliżona do tych otrzymanych w tabeli 2. Może to świadczyć o poprawności wysuniętych przed chwilą wniosków. Wsymulujmy jednak wartości $\Psi(u)$ korzystając z algorytmu podanego w sekcji 2.2.8. Ponownie wzięto kilka wartości $N = 50$, $N = 100$, $N = 1000$, $N = 10000$, a kapitał początkowy wynosi $u = 50$. Wyniki przedstawiono w poniższej tabeli.

	$\Psi(u = 50)$
$N = 50$	0.08
$N = 100$	0.050
$N = 1000$	0.058
$N = 10000$	0.0478

Tab. 2: Tabela szacowanych wartości prawdopodobieństwa ruiny dla różnej ilości powtórzeń Monte Carlo na nieskończonym przedziale czasowym

Wartości przedstawione oscylują wokół swojej teoretycznej wartości $\Psi(u) \approx 0.0474$. Im większa ilość powtórzeń Monte Carlo tym jest są one sobie bliższe, a także nie występują aż tak duże wahania pomiędzy osiąganymi wartościami dla danego N . Ogółem można wysunąć wnioski, że teoria zgodziła się z praktyką, a co za tym idzie, wszelkie wyprowadzenia i symulacje zostały przeprowadzone poprawnie.

3. Zadanie 2

3.1. Cel zadania

Celem zadania jest oszacowanie średniego czasu w jakim trajektoria ruchu Browna startującego z x opuści odpowiednio dobrany przedział $[a, b]$. Ponadto należy oszacować prawdopodobieństwo zjawiska takiego, że ów przedział został opuszczony poprzez górną granicę b . Oba zagadnienia zostaną zwizualizowane oraz do wysumulowanych punktów pomiarowych zostaną wyznaczone krzywe najlepszego dopasowania.

3.2. Teoria

3.2.1. Proces Wienera/Ruch Browna

Procesem Wienera (Ruchem Browna) nazywamy taki proces stochastyczny $B(t), t > 0$, który spełnia następujące warunki.

1. $B(0) = 0$
2. $B(t)$ ma niezależne przyrosty
3. $B(t)$ ma stacjonarne przyrosty
4. $B(t) \sim \mathcal{N}(0, t)$
5. $B(t)$ ma ciągłe trajektorie

3.2.2. Ruch Browna startujący z x

Proces oznaczony jako B_t^x nazywać będziemy ruchem Browna startującym z $x \in \mathbb{R}$. Jest to równoznaczne z

$$B_t^x = B_t^0 + x, \quad (13)$$

gdzie B_t^0 jest klasycznym ruchem Browna.

3.2.3. Prawo iterowanego logarytmu

Jest to jedna z asymptotycznych własności omawianego procesu. Wiadomo, że

$$\limsup_{t \rightarrow \infty} \frac{B(t)}{\sqrt{2t \log \log t}} = 1 \quad \wedge \quad \liminf_{t \rightarrow \infty} \frac{B(t)}{\sqrt{2t \log \log t}} = -1. \quad (14)$$

3.2.4. Średni czas wyjścia z przedziału $[a, b]$

Zajwisko to opisuje nam funkcja $x \mapsto \mathbb{E}\tau^x$, dla $x \in (a, b)$, gdzie

$$\tau^x := \inf \{t \geq 0 : B_t^x \notin [a, b]\}. \quad (15)$$

Wzór ten jest analogiczny do zdefiniowanego momentu pierwszego uderzenia w punkt a , tj.

$$\tau^x := \inf \{t \geq 0 : B_t^0 = a\}. \quad (16)$$

Można również szacować prawdopodobieństwo nastąpienia opisywanego wyjścia przez jedną z granic. Zapis takiego prawdopodobieństwa dla granicy b zapiszemy następująco.

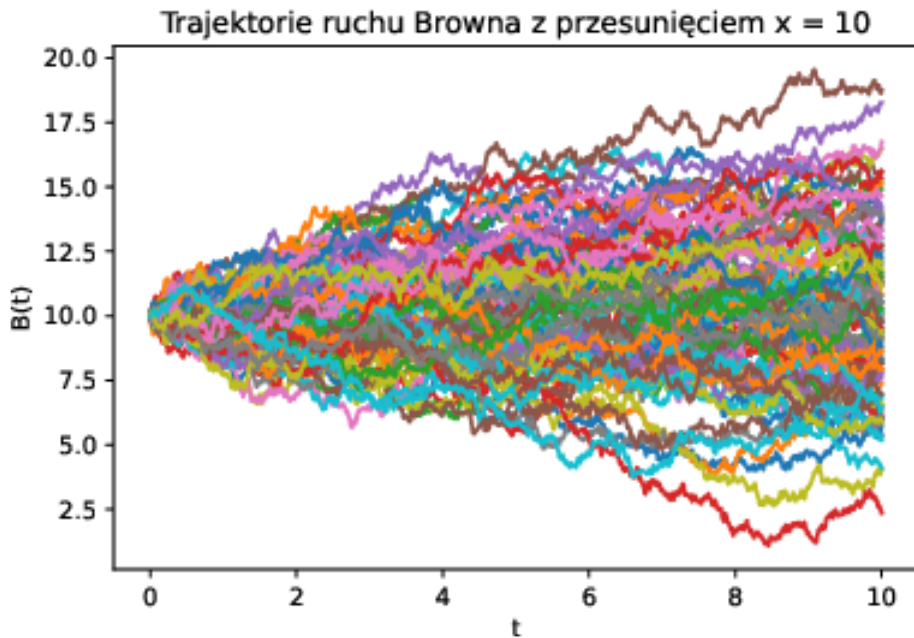
$$\mathbb{P}(B_{\tau^x}^x = b) \quad (17)$$

3.2.5. Algorytm generowania procesu Wienera

Symulując ruch Browna, chcemy wygenerować wektor $[B(t_0), B(t_1), \dots, B(t_n)]$. Ustalony przedział czasowy T podzielimy na n odcinków długości $h = \frac{T}{n}$. Oznaczamy $t_i = ih$, $i = 0, 1, \dots, n$.

1. $B(t_0) = 0$
2. Generujemy $\xi_i \sim \mathcal{N}(0,1)$
3. Rekurencyjnie wyznaczamy $B(t_{i+1}) = B(t_i) + h^{\frac{1}{2}} \cdot \xi_i$

3.3. Wizualizacja trajektorii



Rys. 10: Symulacja $n = 1000$ trajektorii procesu Wienera z przesunięciem $x = 10$.

Analizując sytuację na powyższym rysunku, dochodzimy do wniosku, iż sam proces nie jest wrażliwy na przesunięcie układu. Z uwagi na zachowanie poszczególnych trajektorii, punkty a i b zostaną wybrane symetrycznie względem punktu startu procesu, w tym przypadku będzie to symetrycznie względem punktu $x = 10$.

3.4. Oszacowanie średniego czasu wyjścia z przedziału $[a, b]$

Na podstawie wzoru (15), można oszacować wartość, której trajektorie nie powinny raczej przekraczać. Dla klasycznego ruchu Browna wynosi ona

$$\pm \sqrt{2t \log \log t}, \quad (18)$$

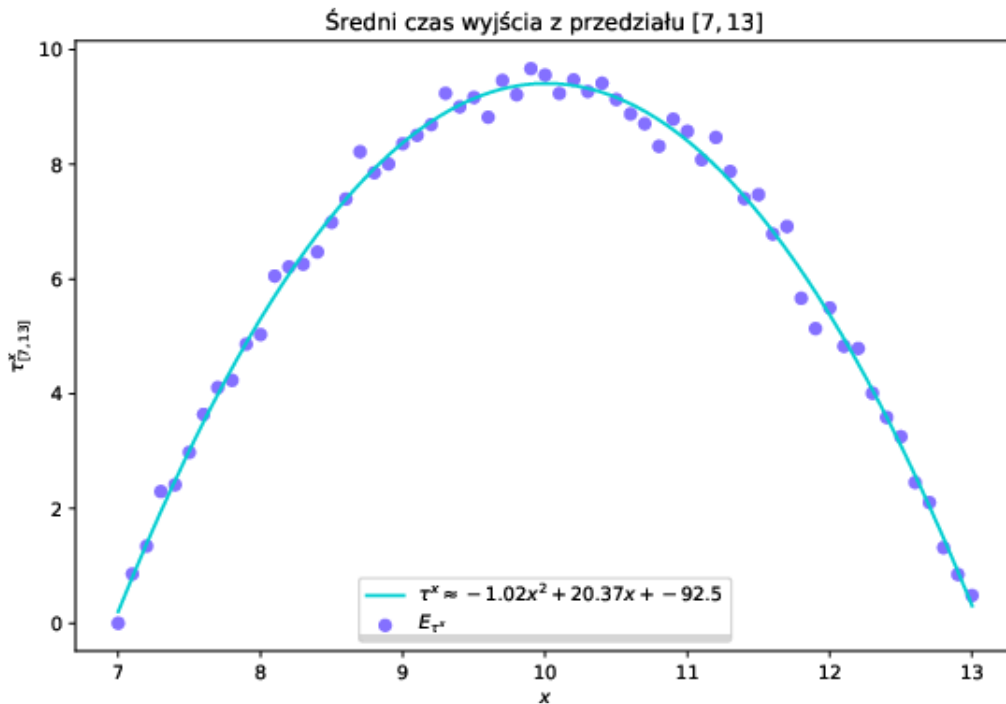
a dla przesuniętego procesu zachowa się ona podobnie osiągając

$$\pm \sqrt{2t \log \log t} + x. \quad (19)$$

Niech $t = 10$, $x = 10$. Wartości otrzymane na podstawie wzoru (19) wynoszą około 5.92 oraz 14.084, dlatego też wybór przedziału $[a, b]$ postawmy na $a = 7$, $b = 13$. Symulacja została przeprowadzona dla $N = 1000$ kroków Monte Carlo. Schemat postępowania jest następujący.

1. Generuj trajektorie przesuniętego o x ruchu Browna.
2. Przy każdym kolejnym kroku rekurencyjnym sprawdzaj, czy wartość B_t^x przekroczyła którąś z wartości granicznych w badanym przedziale. Jeśli tak, zapisz otrzymaną wartość do listy.
3. Powtórz symulacje N razy.
4. Wyznacz średnią z otrzymanych wartości

Dopasowywanie krzywych, zarówno dla średnich czasów opuszczenia przedziału, jak i prawdopodobieństwa, że stanie się to przez granicę górną, zostanie wyznaczone przy pomocy biblioteki **scipy** w języku Python oraz modułu **optimize.curve_fit**. Opiera się ona na metodzie najmniejszych kwadratów ¹.



Rys. 11: Średni czas opuszczenia przedziału $[7, 13]$ przez trajektorie dla przesuniętego procesu Wienera o $x = 10$ wraz z dopasowaną krzywą.

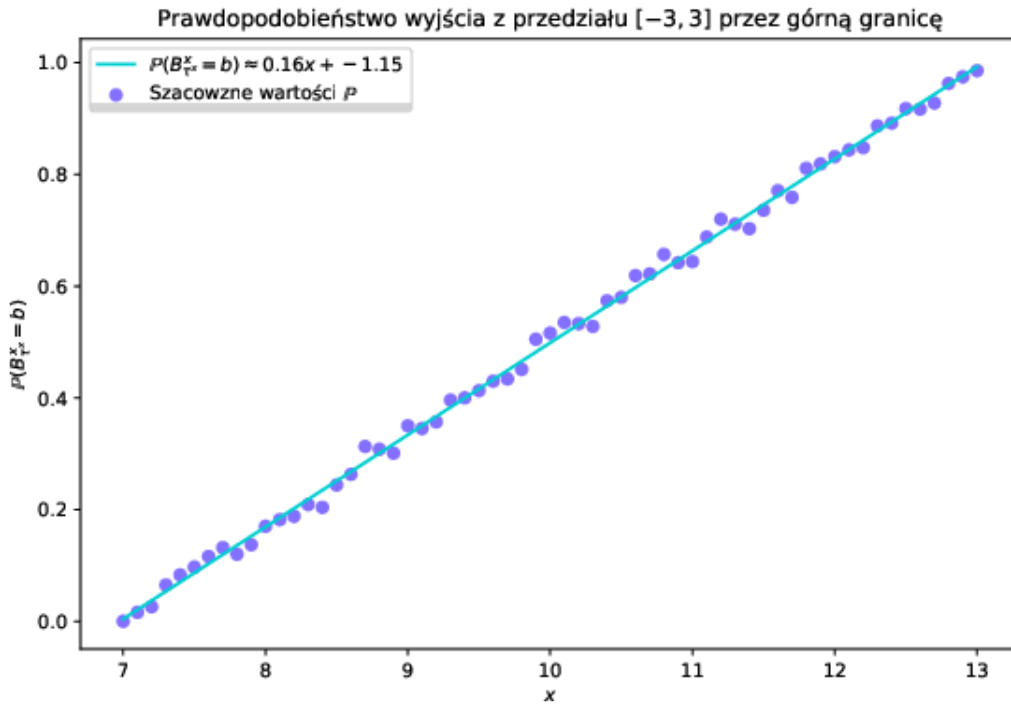
¹ https://linuxhint.com/scipy-optimize-curve_fit/

Otrzymane wyniki nie są zaskakujące biorąc pod uwagę zachowanie trajektorii zauważalne na rysunku 10. Wraz ze wzrostem odległości od punktu startowego dla omawianego procesu, średni czas opuszczenia przedziału maleje.

Charakterystyka wykresu od razu przypominała funkcję kwadratową, dlatego też w celu najlepszego dopasowania posłużono się funkcją $y = ax^2 + bx + c$. Dopasowanie to wydaje się dość dobre, a otrzymane wcześniej wspomnianym module wartości współczynników wyniosły odpowiednio $a = -1.02, b = 20.37$ i $c = -92.5$. Biorąc pod uwagę widoczne miejsca zerowe, bez dokładnych obliczeń można spróbować dopasować krzywą $y = -(x - 7)(x - 13)$, co po wyznaczeniu postaci ogólnej, jest bardzo zbliżone do wygenerowanej krzywej najlepszego dopasowania.

3.5. Oszacowanie prawdopodobieństwa wyjścia przez granicę b

W celu oszacowania takiego prawdopodobieństwa, we wcześniejszym algorytmie symulacyjnym należy sprecyzować, że to właśnie górna granica została przekroczona.



Rys. 12: Prawdopodobieństwo opuszczenia przedziału $[7, 13]$ przez górną granicę przez trajektorię dla przesuniętego procesu Wienera o $x = 10$ wraz z dopasowaną krzywą.

Otrzymane wartości \mathbb{P} dla $x \in (7, 3)$, ustawiły się niemalże liniowo. Z tego też powodu poszukiwaną krzywą dopasowania była funkcja $y = ax + b$. Współczynniki wyniosły $a = 0.16$ oraz $b = -1.15$. Ponownie, zachowanie prawdopodobieństwa w zależności od x było przewidywalne pod względem monotoniczności. Wraz ze zmniejszaniem odległości do górnej granicy prawdopodobieństwo przekroczenia jej rosło. Warto nadmienić również, że z uwagi na wybór symetrycznego przedziału względem punktu startowego, dla tego ów punktu $\mathbb{P}(B_{\tau^x}^x = b) \approx \mathbb{P}(B_{\tau^x}^x = a)$ i wynoszą około $\frac{1}{2}$.

4. Podsumowanie

W powyższym raporcie, przeprowadzona została analiza dwóch różnych procesów.

W zadaniu pierwszym analizowany był proces Ryzyka. Wspomniane zostały wszystkie definicje dotyczące tego zagadnienia, a także algorytmy umożliwiające jego symulowanie. Zadanie w głównej mierze polegało na dopasowaniu modelu klasycznego procesu Ryzyka do otrzymanych danych. Wy-tłumaczono schemat postępowania odnajdywania kolejnych parametrów omawianego zjawiska. Po odnalezieniu wszystkich wartości wysymulowano przykładowe trajektorie procesu i były one zgodne z tymi otrzymanymi. Podobnie było w przypadku wyznaczania prawdopodobieństw ruiny w danym horyzoncie czasowym. Wartości szacunkowe ów zagadnienia otrzymane za pomocą przeprowadzonych symulacji były bardzo zbliżone do wartości prawdopodobieństwa ruiny na podstawie danych z zadania, w tym także dla przypadku nieskończonego horyzontu czasowego. Może to świadczyć o tym, że największe prawdopodobieństwo osiągnięcia przez proces wartości mniejszej od 0 występuje na początku.

W zadaniu drugim przyjrzelśmy się procesowi Wienera (inaczej ruch Browna). Korzystając z przedstawionych zagadnień teoretycznych, byliśmy w stanie symulacyjnie wyznaczyć zadane nam problemy, takie jak średni czas opuszczenia przedziału przez trajektorię, czy oszacowanie prawdopodobieństwa, że nastąpi to przez jedną z granic ów przedziału. Badając przedziały o granicach równie oddalonych od punktu x_0 , doszliśmy do wniosku, iż wartość oczekiwana malała wraz ze zmniejszaniem odległości od granicy, a największą wartość przyjmuje dla x -ów w okolicy punktu, z którego proces startuje. Szacując prawdopodobieństwo przekroczenia górnej granicy takiego przedziału, wywnioskowaliśmy, iż liniowo rośnie ono wraz ze wzrostem x .

Ogólnie rzecz biorąc, wszelkie wykonane obliczenia oraz symulacje wydają się poprawne. W każdym wypadku otrzymano wyniki zgodne z przewidywaniami.