

Symulacje komputerowe

Raport 2

Temat: Proces Rzyzka i ruch Browna

Imię i Nazwisko prowadzącego kurs: **Dr Michał Balcerak**

Imię i Nazwisko, nr indeksu	Kacper Budnik, 262286 Szymon Malec, 262276
Wydział:	Wydział matematyki, W13
Dzień i godzina zajęć:	Wtorek, 11 ¹⁵
Kod grupy ćwiczeniowej:	T00-70d
Data oddania raportu:	26.06.2022
Ocena końcowa	

Adnotacje i uwagi:

1. Wstęp

Drugi raport z symulacji komputerowych został podzielony na dwa zadania. Celem pierwszego z nich było dopasowanie modelu Ryzyka do otrzymanych danych oraz oszacowanie prawdopodobieństwa ruiny w skończonym i nieskończonym czasie. W tym celu:

- Znaleźliśmy rozkład zmiennej losowej X_i oznaczającej wielkość i -tej szkody.
- Wyestymowaliśmy pozostałe parametry procesu:
 1. λ – intensywność jednorodnego procesu Poissona,
 2. θ – parametr odpowiedzialny za wysokość premii.
- Oszacowaliśmy prawdopodobieństwo ruiny dla czasu:
 1. skońzonego przy pomocy metody Monte Carlo,
 2. nieskońzonego przy pomocy algorytmu Pollaczka-Chinczyna.

Drugie zadanie polegało natomiast na oszacowaniu średniego czasu wyjścia ruchu Browna z zadanego przedziału w zależności od punktu startowego. Dodatkowo mieliśmy oszacować prawdopodobieństwo wyjścia przez dany koniec odcinka. Na koniec, przy pomocy pakietów matematycznych, dopasowaliśmy funkcję do wygenerowanych metodą Monte Carlo danych.

2. Zadanie 1

2.1. Jednorodny proces Poissona

Jednorodnym procesem Poissona z intensywnością $\lambda > 0$ nazywamy proces liczący $\{N_t\}_{t \geq 0}$, który spełnia:

- $N_0 = 0$,
- N_t ma niezależne przyrosty,
- N_t ma stacjonarne przyrosty,
- $N_t \sim \text{Poiss}(\lambda t)$.

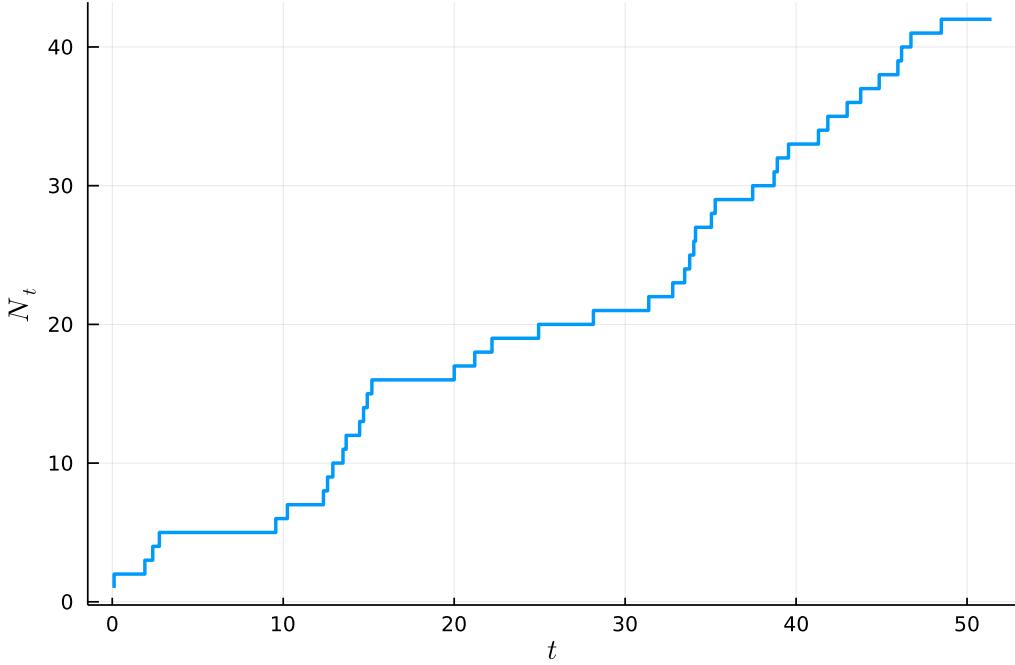
Proces ten służy do zliczania wystąpień pewnych zdarzeń losowych, których częstość występowania (intensywność) jest cały czas taka sama.

Do symulacji procesu Poissona przydaje się jedna z jego własności, którą jest to, że odstępy czasowe pomiędzy skokami mają rozkład $\mathcal{Exp}(\lambda)$. Oznacza to, że jeśli chcemy wygenerować jedną trajektorię na przedziale $[0, T]$, to wystarczy generować kolejne $X_i \sim \mathcal{Exp}(\lambda)$, aż ich suma przekroczy T .

Algorytm

1. Definiuj $I = 0$, $t = 0$.
2. Generuj $X \sim \mathcal{Exp}(\lambda)$.
3. $t = t + X$. Jeśli $t > T$, przerwij i zwróć $\{S_i\}$. W przeciwnym razie $I = I + 1$, $S_I = t$.
4. Wróć do kroku 2.

Otrzymane z powyższego algorytmu $\{S_i\}$ to czasy oczekiwania na i-ty skok od startu trajektorii.



Rysunek 1: Przykładowa trajektoria procesu Poisson z intensywnością $\lambda = 1$.

2.2. Klasyczny proces Ryzyka

Klasycznym procesem Ryzyka nazywamy proces stochastyczny opisujący kapitał firmy ubezpieczeniowej postaci

$$R_t = u + ct - \sum_{i=1}^{N_t} X_i,$$

gdzie

- $u > 0$ – kapitał początkowy,
- ct – stałe przychody ($c > 0$),
- N_t – jednorodny proces Poissona z intensywnością $\lambda > 0$,
- X_i – ciąg niezależnych zmiennych losowych o tym samym rozkładzie opisujący straty, $\mathbb{E}X_i = \mu$, $X_i > 0$.

Proces ten odgrywa ważną rolę w branży ubezpieczeniowej, ponieważ symulacja tego procesu pozwala oszacować prawdopodobieństwo bankructwa firmy oraz jej średni zysk. Dzięki temu jesteśmy w stanie określić jaka powinna być cena ubezpieczenia, by takie prawdopodobieństwo nie było zbyt wysokie, a zysk zadowalający.

Parametrem odpowiadającym za przychody jest stała c . Do odpowiedniego doboru tego parametru wykorzystuje się wzór

$$c = (1 + \theta)\mu\lambda \tag{1}$$

gdzie $\theta > 0$ jest parametrem odpowiadającym za wysokość premii.

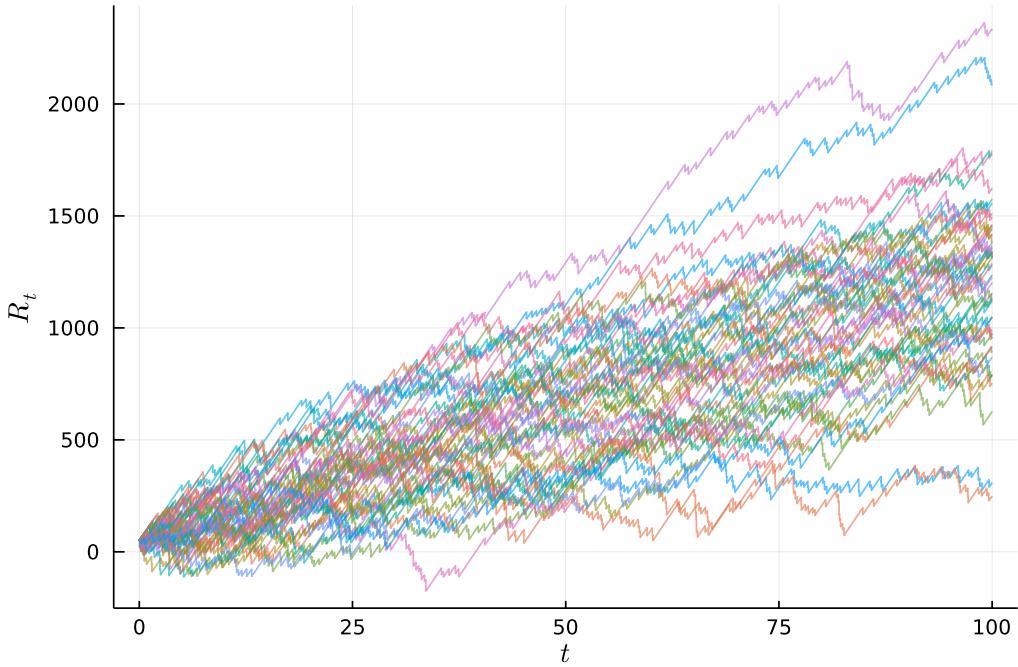
Algorytm

1. Generuj N_t na przedziale $[0, T]$.
2. Generuj X_1, \dots, X_{N_t} .
3. Zwróć $R_t = u + ct - \sum_{i=1}^{N_t} X_i$.

2.3. Dopasowanie modelu do danych

Zastosujemy teraz proces Ryzyka w praktyce. Postaramy się jak najlepiej dopasować model klasycznego procesu Ryzyka do danych, które zostały nam udostępnione przez prowadzącego zajęcia laboratoryjne. Dane te opisują 50 trajektorii pewnego klasycznego procesu Ryzyka. Naszym celem jest wyznaczenie parametrów tego procesu na podstawie danych, by móc następnie przeprowadzać jego symulacje.

Badany proces został wysymulowany na odcinku $[0, 100]$ z krokiem czasowym $h = 0,01$. Kapitał początkowy wynosi $u = 50$.



Rysunek 2: Trajektorie 50 realizacji klasycznego procesu Ryzyka zawartych w danych.

2.3.1. Parametr c

Możemy wyznaczyć go w bardzo prosty sposób. Wiedząc, że jest to współczynnik nachylenia funkcji liniowej ct , wystarczy, że policzymy o ile funkcja ta wzrasta w jednym kroku czasowym (obliczymy różnicę dwóch sąsiadujących danych) i wartość tę podzielimy przez długość kroku. Ważne, by upewnić się, czy w danym kroku nie nastąpił spadek. Dla pewności możemy policzyć c dla

kilku kroków i sprawdzić, czy otrzymane wartości są identyczne. Wartość, którą otrzymaliśmy z naszych danych to $c = 56,25$.

2.3.2. Parametr λ

W przypadku tego parametru nie jesteśmy w stanie wyznaczyć jego dokładnej wartości, zatem skorzystamy z estymatora. Wykorzystamy fakt, że jeśli proces Poissona ma intensywność λ , to odstępy czasowe między kolejnymi skokami mają rozkład $\mathcal{Exp}(\lambda)$. Wydzielamy więc wszystkie czasy oczekiwania ze wszystkich trajektorii do jednego zbioru danych. Robimy to sumując wszystkie kroki czasowe następujące po każdym spadku aż do następnego spadku, przy czym dodajemy także ostatni krok spadkowy. Należy zaznaczyć w tym miejscu, że pozwalamy sobie tutaj na pewne uproszczenie, ponieważ przyjmujemy, że spadek miał miejsce dokładnie pod koniec kroku spadkowego, podczas gdy w rzeczywistości miał on miejsce w jego trakcie.

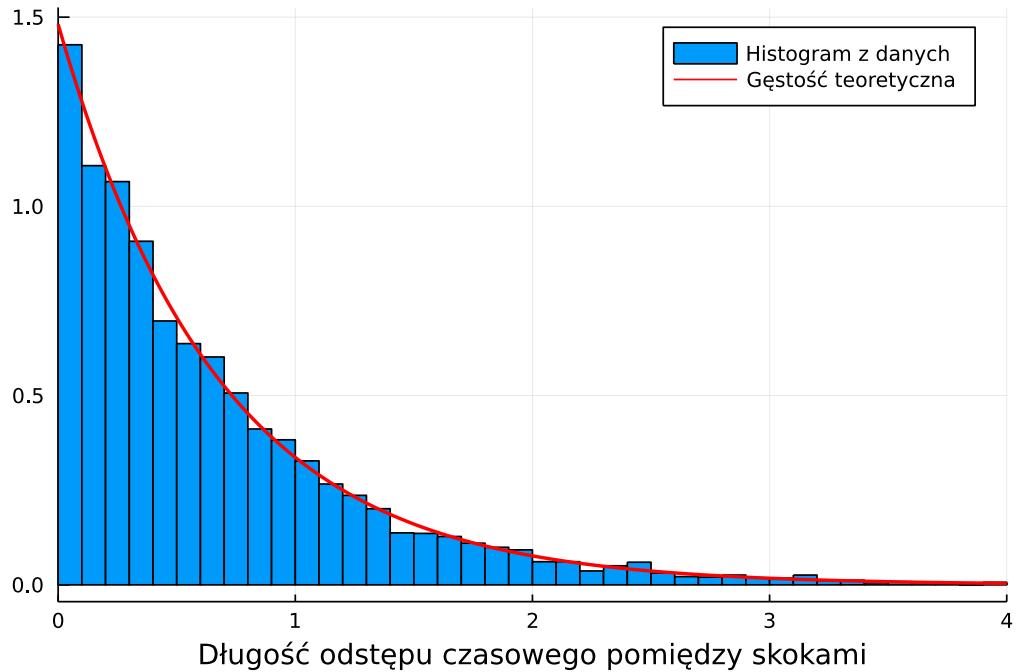
By wyestymować parametr λ , skorzystamy z faktu, że dla $\tau \sim \mathcal{Exp}(\lambda)$ wartość oczekiwana jest równa

$$\mathbb{E}\tau = \frac{1}{\lambda}.$$

Najlepszym estymatorem wartości oczekiwanej jest średnia z próby. Oznaczmy przez τ_1, \dots, τ_n odstępy czasowe pozyskane z naszych danych. Estymatorem λ będzie więc

$$\hat{\lambda} = \frac{1}{\frac{1}{n} \sum_{i=1}^n \tau_i} = \frac{n}{\sum_{i=1}^n \tau_i}.$$

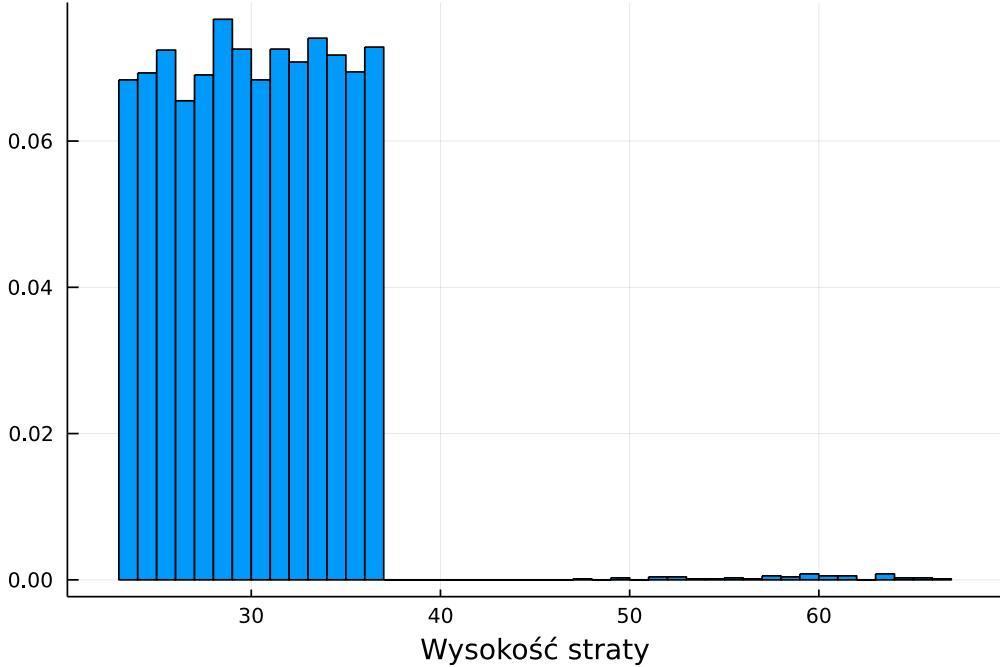
Po obliczeniu otrzymaliśmy $\hat{\lambda} \approx 1,48$. Poniżej możemy zobaczyć, że krzywa gęstości teoretycznej rozkładu $\mathcal{Exp}(\hat{\lambda})$ wyraźnie pokrywa się z histogramem z naszych danych.



Rysunek 3: Porównanie unormowanego histogramu z odstępów czasowych pomiędzy skokami z gęstością teoretyczną rozkładu $\exp(-\lambda)$.

2.3.3. Straty X_i

Niezbędną informacją potrzebną do dopasowania modelu do badanego procesu Ryzyka jest rozkład strat X_i . Aby pozyskać dane opisujące straty, obliczamy różnicę sąsiadujących danych w każdym kroku spadkowym. Dodatkowo, ponieważ nasze dane opisują trajektorie z dokładnością do kroku czasowego, to należy do otrzymanych wartości dodać jeszcze przyrost z przychodów w tym kroku. W naszym przypadku przyrost ten będzie wynosił $c \cdot h = 0,56625$.



Rysunek 4: Unormowany histogram ze wszystkich strat pozyskanych z 50 trajektorii.

Na powyższym histogramie widzimy, że zdecydowana większość danych znajduje się w okolicach 30. Kształt histogramu sugeruje, że mają one rozkład jednostajny. Możemy jednak zauważać, że pojawia się niewielka ilość odstających większych wartości. Stąd możemy przypuszczać, że jest to kompozycja rozkładów. Zdefiniujmy zatem jej gęstość jako

$$f_X(x) = p_1 f_1(x) + p_2 f_2(x), \quad (2)$$

gdzie p_1 i p_2 są pewnymi prawdopodobieństwami ($p_1 + p_2 = 1$), a f_1 i f_2 są gęstościami skomponowanych rozkładów.

Zajmijmy się najpierw znalezieniem rozkładu dominujących danych (o gęstości f_1). Jak zostało wspomniane, histogram sugeruje, że mają one rozkład jednostajny. Odseparujemy te wartości definiując zbiór

$$K_1 = \{X_i : X_i < 40\}.$$

Założymy, że dane te pochodzą z rozkładu $\mathcal{U}(a, b)$. Estymatorami parametrów a i b wyznaczonymi metodą Największej Wiarygodności są

$$\hat{a} = \min(K_1) \approx 23, \quad \hat{b} = \max(K_1) \approx 37.$$

Skorzystamy z testu Kołmogorowa-Smirnowa w celu weryfikacji hipotezy

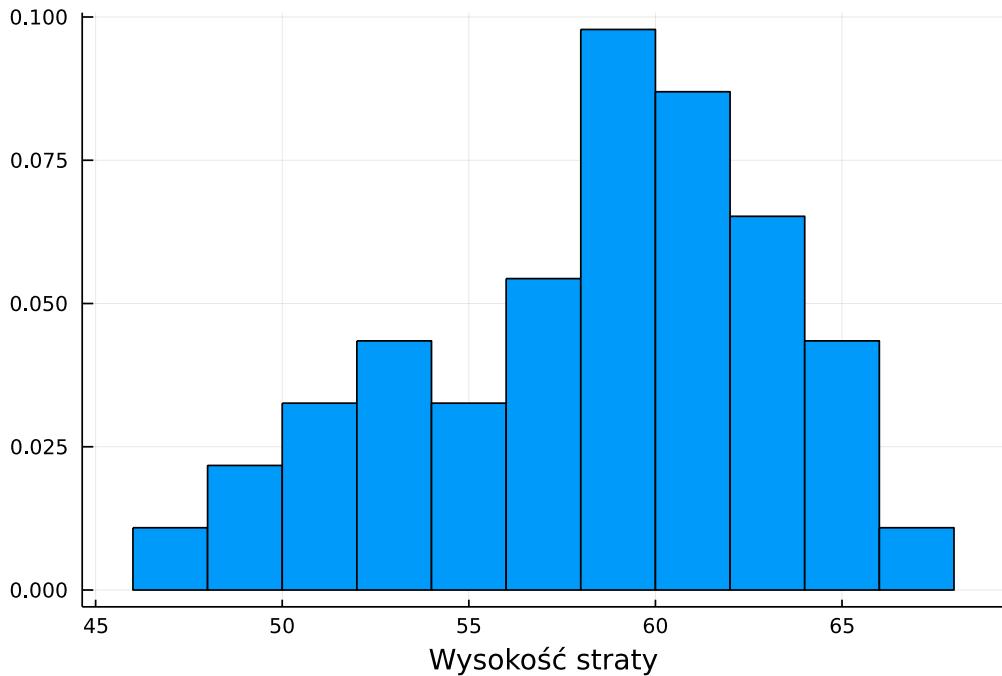
$$H_0 : X \sim \mathcal{U}(23, 37) \quad \forall_{X \in K_1}.$$

Otrzymaliśmy p-wartość równą 0,2449, zatem możemy śmiało przyjąć tę hipotezę.

Pozostało nam wyznaczyć rozkład wartości odstających. Ponieważ jest ich zaledwie 46 na 7358 wszystkich danych, to nie jesteśmy w stanie jednoznacznie stwierdzić do jakiego rozkładu one należą. Jednakże, ze względu na tak niewielką ilość, nie jest to na tyle istotne, ponieważ wartości te mają marginalny wpływ na całkowity rozkład. Ważne jest, by w naszym modelu z taką samą częstością pojawiały się takie właśnie ekstremalne spadki, a ich wartość była zbliżona do tych, które zawierają dane.

Zdefiniujmy zbiór wartości odstających jako

$$K_2 = \{X_i : X_i > 40\}.$$



Rysunek 5: Unormowany histogram ze strat większych od 40.

Na powyższym histogramie zauważmy, że dane zachowują się w miarę podobny sposób jak rozkład normalny. Przyjmijmy więc, że

$$X \sim \mathcal{N}(m, \sigma^2) \quad \forall_{X \in K_2}.$$

Estymatorami parametrów m i σ^2 są oczywiście

$$\widehat{m} = \frac{1}{\#K_2} \sum_{X \in K_2} X \approx 58,5, \quad \widehat{\sigma}^2 = \frac{1}{\#K_2 - 1} \sum_{X \in K_2} (X - \widehat{m})^2 \approx 4,88.$$

Dla tego rozkładu również przeprowadziliśmy test Kołmogorowa-Smirnowa i otrzymaliśmy p-wartość równą 0,7586, więc możemy zaakceptować nasze założenie.

Policzmy jeszcze współczynniki p_1 i p_2 we wzorze (2). Możemy je wyestymować w prosty sposób:

$$\widehat{p}_1 = \frac{\#K_1}{\#K_1 + \#K_2} \approx 0,9937, \quad \widehat{p}_2 = \frac{\#K_2}{\#K_1 + \#K_2} \approx 0,0063.$$

Na koniec znajdziemy wartość parametru μ . Wystarczy, że obliczymy średnią ze wszystkich strat, czyli

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n X_i \approx 30,23.$$

2.3.4. Parametr θ

Znając już wszystkie parametry z prosty sposób możemy wyliczyć wartość parametru θ odpowiadającego za wysokość premii. Przekształcając wzór (1) dostaniemy

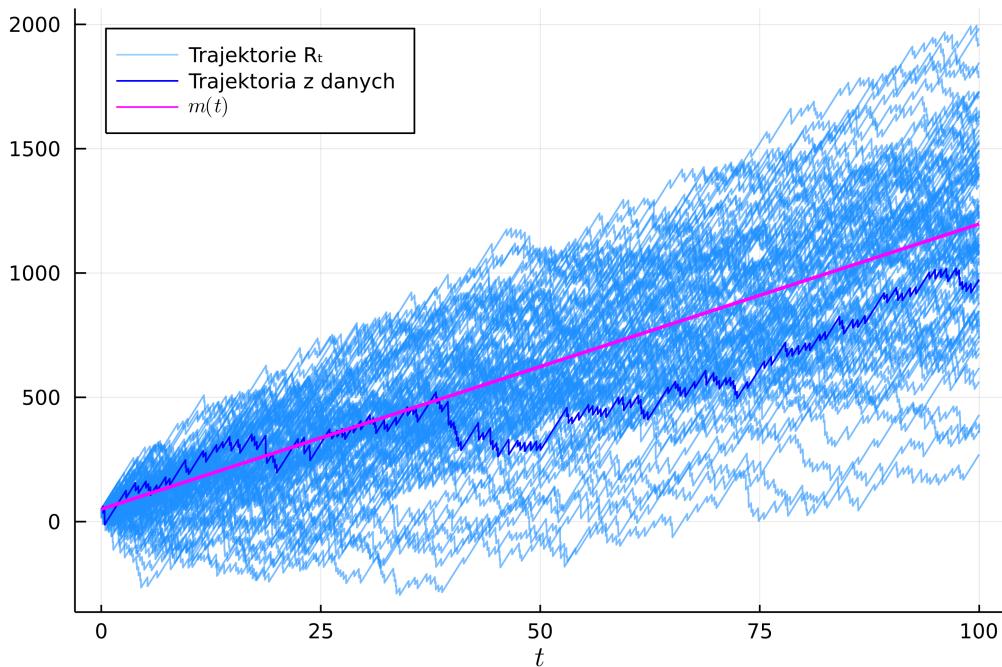
$$\theta = \frac{c}{\mu\lambda} - 1.$$

Po podstawieniu wartości otrzymujemy $\theta \approx 0,256$.

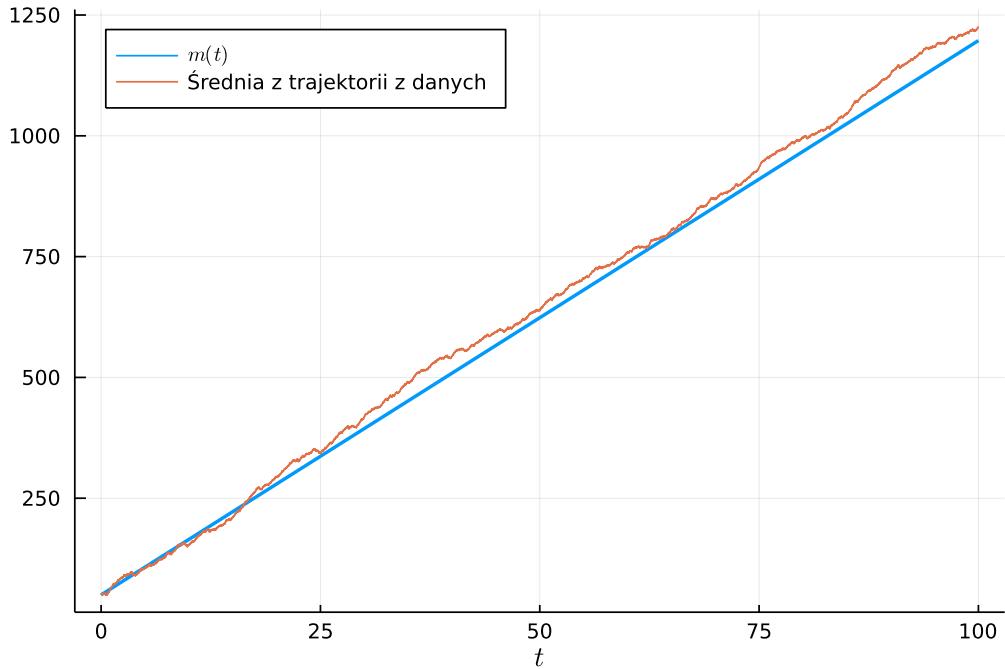
2.3.5. Gotowy model

Po wyznaczeniu wszystkich potrzebnych parametrów i znalezieniu rozkładu strat możemy sprawdzić działanie modelu. Zanim to zrobimy, wyznaczmy średnią procesu. Jest funkcja liniowa postaci

$$\begin{aligned} m(t) &= \mathbb{E}R_t = \mathbb{E} \left[u + ct - \sum_{i=1}^{N_t} X_i \right] = u + ct + \mathbb{E} \left[\mathbb{E} \left[\sum_{i=1}^{N_t} X_i | N_t \right] \right] \\ &= u + ct - \mu\lambda t = u + (1 + \theta)\mu\lambda t - \mu\lambda t = u + \mu\lambda\theta t. \end{aligned}$$



Rysunek 6: Wykres 100 trajektorii klasycznego modelu Ryzyka z wyznaczonymi parametrami wraz z jedną trajektorią z danych oraz średnią procesu wyznaczoną analitycznie.



Rysunek 7: Porównanie funkcji średniej $m(t)$ wyznaczonego procesu oraz średniej z trajektorii z danych.

Jak widzimy na Rys. 6, wysymulowane trajektorie oscylują wokół średniej procesu. Pojedyncza trajektoria pochodząca z danych wydaje się zachowywać podobnie do tych wysymulowanych. Z kolei na Rys. 7 można zobaczyć, że funkcja średniej $m(t)$ dla naszego procesu pokrywa się ze średnią z trajektorii, które zawierają dane.

2.4. Prawdopodobieństwo ruiny

Gdy już udało nam się wydobyć z danych wszystkie potrzebne informacje aby dopasować do nich nasz model, możemy rozpocząć analizę wyznaczonego procesu. Naszym celem jest obliczenie prawdopodobieństwa ruiny (bankructwa) firmy, dla której może być to najcenniejsza informacja, ponieważ pokazuje jakie są szanse, że firma odniesie sukces.

2.4.1. Prawdopodobieństwo ruiny w skończonym czasie

Definiujemy moment ruiny jako

$$\tau(u) = \inf\{t \geq 0 : R_t < 0\},$$

gdzie u oznacza kapitał początkowy. Prawdopodobieństwo ruiny w skończonym czasie T definiujemy poprzez

$$\Psi(u, T) = \mathbb{P}(\tau(u) < T).$$

Prawdopodobieństwo to da się wyznaczyć tylko w przypadku, gdy X_i mają rozkład wykładniczy, co nas nie dotyczy. Skorzystamy więc z symulacji Monte Carlo.

Algorytm

1. Generuj N trajektorii $R_t^{(1)}, \dots, R_t^{(N)}$ procesu ryzyka na $[0, T]$.
2. Wyznacz $n = \#\left\{i = 1, \dots, N : \min_{t \in [0, T]} R_t^{(i)} < 0\right\}$.
3. Zwróć $\Psi(u, T) = \frac{n}{N}$.

Wygenerowaliśmy więc $N = 10\ 000$ trajektorii na przedziałach $[0, 100]$ i $[0, 200]$, po czym sprawdziliśmy ile z nich spadło w dowolnym momencie poniżej zera. Otrzymaliśmy

$$\Psi(50, 100) = 0,4091, \quad \Psi(50, 200) = 0,4194.$$

2.4.2. Prawdopodobieństwo ruiny w nieskończonym czasie

Definiujemy prawdopodobieństwo ruiny w nieskończonym czasie jako

$$\Psi(u) = \mathbb{P}(\tau(u) < \infty).$$

Analityczne wyznaczenie $\Psi(u)$ także nie jest możliwe w naszym przypadku. Nie możemy również skorzystać z metody Monte Carlo, ponieważ czas jest tym razem nieskończony. Posłużymy się więc wzorem Pollaczka-Chinczyna, który wygląda następująco:

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

gdzie $B_n(u) = \mathbb{P}(Y_1 + \dots + Y_n > u)$ i Y_i to zmienne losowe iid o gęstości $f(x) = \frac{1-F_X(x)}{\mu}$, $x > 0$, a $F_X(x)$ jest dystrybuantą rozkładu strat X_i . Wykorzystamy pewien fakt wynikający z tego wzoru, a konkretnie to, że

$$\Psi(u) = \mathbb{P}(Y_1 + \dots + Y_K > u),$$

gdzie $K \sim \text{Geom}\left(\frac{\theta}{1+\theta}\right)$.

Algorytm

1. Definiuj $Z = 0$.
2. Generuj $K \sim \text{Geom}\left(\frac{\theta}{1+\theta}\right)$.
3. Generuj Y_1, \dots, Y_K z rozkładu o gęstości $f(x) = \frac{1-F_X(x)}{\mu}$, $x > 0$.
4. Jeśli $Y_1 + \dots + Y_K > u$, to $Z = Z + 1$.
5. Kroki 2, 3, 4 powtórz N razy.
6. Zwróć $\Psi(u) = \frac{Z}{N}$.

Zauważmy, że do tej symulacji potrzebna jest nam gęstość $f(x) = \frac{1-F_X(x)}{\mu}$. Ponieważ F_X , czyli dystrybuanta rozkładu zmiennych X_i jest kompozycją dwóch innych dystrybuant (podobnie jak dla gęstości we wzorze (2)), w tym dystrybuanty rozkładu normalnego, wyznaczenie gęstości f będzie problematyczne. Wiemy jednak, że współczynnik p_2 jest bardzo mały, a więc możemy pozwolić sobie na małe uproszczenie i w symulacji uwzględnić tylko dystrybuantę rozkładu jednostajnego, czyli

$$F_1(x) = \begin{cases} 0, & x < 23 \\ \frac{x-23}{14}, & 23 \leq x < 37 \\ 1, & 37 \leq x \end{cases}$$

Wtedy musimy także przyjąć wartość teoretyczną μ dla tego rozkładu tzn. $\mu = 30$. Po takim uproszczeniu otrzymamy

$$f(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{30}, & 0 \leq x < 23 \\ \frac{37-x}{420}, & 23 \leq x < 37 \\ 0, & 37 \leq x \end{cases}.$$

Zmienne o takiej gęstości już o wiele łatwiej wygenerować (np. metodą akceptacji i odrzucenia). Ostatecznie po wykonaniu algorytmu dla $N = 1\ 000\ 000$ powtórzeń, otrzymaliśmy

$$\Psi(50) \approx 0,4206.$$

Wynik ten może być odrobinę zaniżony przez to, że nie uwzględniliśmy odstających wartości. Mimo to jest on większy od poprzednich prawdopodobieństw dla skończonego czasu, co jest naturalne, ponieważ im więcej czasu minie, tym więcej trajektorii może jeszcze spaść do zera.

3. Zadanie 2

3.1. Proces Wienera

Procesem Wienera (Ruchem Browna) nazywamy proces $\{W_t\}_{t \geq 0}$, który spełnia następujące własności:

- $W_0 = 0$,
- W_t ma niezależne przyrosty,
- W_t ma stacjonarne przyrosty,
- $W_t \sim \mathcal{N}(0, t)$,
- W_t ma ciągłe trajektorie.

Algorytm

Naszym celem jest wygenerowanie wektora $[W_{t_0}, W_{t_1}, \dots, W_{t_n}]$, gdzie

$$t_i = ih, \quad i \in [n] = \{0, 1, \dots, n\}, \quad h = T/n.$$

W tym celu stosujemy poniższy algorytm

1. Generuj realizację zmiennych ξ_i iid. $\mathcal{N}(0,1)$, $i \in [n-1]$,
2. $W_{t_0} = W(0) = 0$,
3. $W_{t_{i+1}} = W_{t_i} + \sqrt{h}\xi_i$, $i \in [n-1]$.

3.2. Czas wyjścia

Czasem wyjścia tego procesu z ustalonego przedziału $[a, b]$ nazywamy zmieniąną

$$\tau = \inf \{t \geq 0 : W_t \notin [a, b]\}.$$

Niech $\{B_t^x\}_{t \geq 0}$ będzie ruchem Browna startującym z $x \in \mathbb{R}$. Zdefiniujmy teraz zmienną

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

Łatwo można pokazać, że $B_t^{x+y} \in [a,b] \iff B_t^x \in [a-y, b-y]$. Dodatkowo, ponieważ ruch Browna jest $\frac{1}{2}$ -samopodobny ($W_t \stackrel{d}{=} \sqrt{\Delta} W_{t/\Delta}$), zachodzi równość

$$B_t^x = W_t + x \stackrel{d}{=} \Delta W_{t/\Delta^2} + x = \Delta \left(W_{t/\Delta} - \frac{x}{\Delta} \right) = \Delta B_{t/\Delta^2}^{x/\Delta}.$$

Korzystając z tych dwóch własności można udowodnić, że

$$\begin{aligned} \tau^x &\stackrel{d}{=} \inf \left\{ \Delta^2 \frac{t}{\Delta^2} \geq 0 : \Delta B_{t/\Delta^2}^{x/\Delta} \notin [a, b] \right\} = \Delta^2 \inf \left\{ t^* \geq 0 : B_{t^*}^{x/\Delta} \notin \left[\frac{a}{\Delta}, \frac{b}{\Delta} \right] \right\} \\ &= \Delta^2 \inf \left\{ t^* \geq 0 : B_{t^*}^{(x-s)/\Delta} \notin \left[\frac{a-s}{\Delta}, \frac{b-s}{\Delta} \right] \right\}. \end{aligned}$$

Wybierając teraz $\Delta = (b-a)$ oraz $s = a$ otrzymujemy

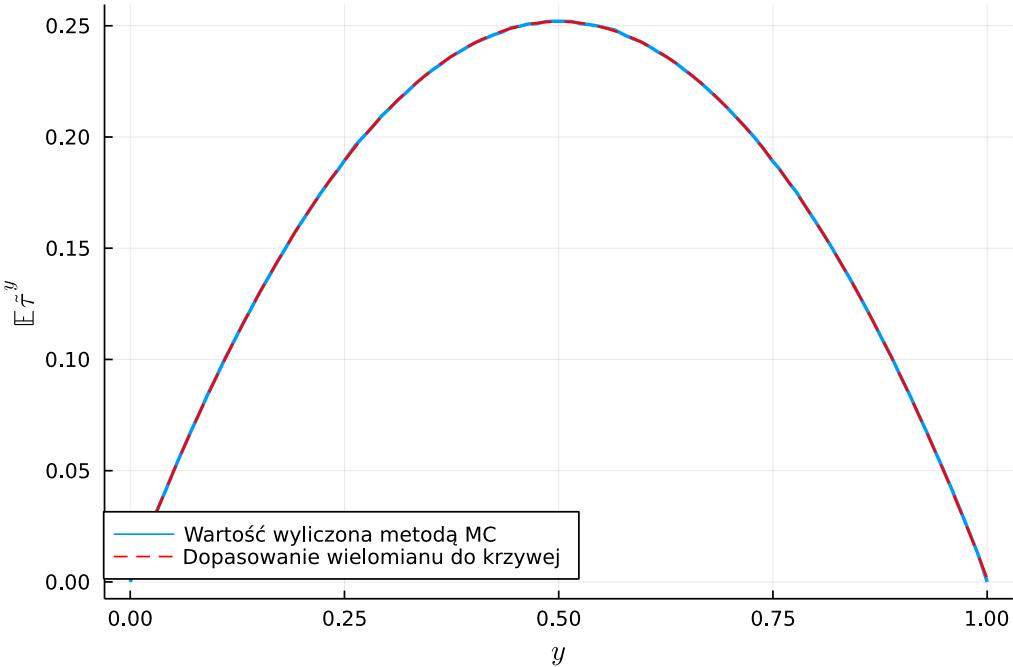
$$\tau^x \stackrel{d}{=} (b-a)^2 \inf \{t \geq 0 : B_t^y \in [0,1]\} = (b-a)^2 \tilde{\tau}^y, \quad (3)$$

gdzie $\tilde{\tau}^y$ jest czasem wyjścia z przedziału $[0, 1]$, jeśli zaczniemy z punktu $y = \frac{x-a}{b-a}$, czyli z punktu dzielącego ten odcinek w stosunku takim samym jak x dzieli odcinek $[a, b]$.

Dzięki tym przekształceniom widzimy, że $\mathbb{E}\tau^x$ zwiększa się z kwadratem długości przedziału jaki rozważamy oraz zależy ona od odległości x od jednego końca przedziału względem drugiego. Z tego powodu później będziemy rozważać jedynie czas wyjścia z przedziału $[0, 1]$.

3.3. Średni czas wyjścia procesu Wienera

Interesuje nasz średni czas wyjścia procesu Wienera w zależności od wybranego punktu startowego x . Tą wielkość oszacowaliśmy poprzez obliczenie średniej w wygenerowanej przez nas próbie. Metoda jest poprawna, ponieważ średnia w próbie jest dobrym estymatorem wartości oczekiwanej oraz można udowodnić, że $\mathbb{E}\tau^x < \infty$ ^[1]. W wyniku symulacji otrzymaliśmy poniższy wykres.



Rysunek 8: Oszacowanie wartości oczekiwanej czasu wyjścia w zależności od punktu początkowego.

Kształt wykresu mógł sugerować, że funkcja $f(y) = \mathbb{E}\tilde{\tau}^y$ jest wielomianem. Z pomocą pakietów matematycznych udało nam się dopasować funkcję

$$f(y) = -1,0012y^2 + 1,0012y + 0,0016, \quad (4)$$

która jak widzimy na wykresie, dobrze przybliża otrzymaną krzywą.

Wartość oczekiwana czasu wyjścia dla standardowego procesu Wienera jest znana^[1]

$$\mathbb{E}\tau = ab,$$

gdzie τ jest czasem wyjścia z przedziału $[-a, b]$. Stąd, uwzględniając przesunięcie początkowe naszego procesu, otrzymujemy

$$\mathbb{E}\tilde{\tau}^y = y(1 - y).$$

Stosując niewielkie przybliżenie (do 2-giego miejsca po przecinku) do funkcji (4), otrzymujemy wartość teoretyczną. Dodatkowo, korzystając teraz z (3) oraz $y = (x - a)/(b - a)$ możemy uogólnić wzór na dowolny przedział $[a, b]$, otrzymując

$$\mathbb{E}\tau^x = (b - a)^2 \frac{x - a}{b - a} \left(1 - \frac{x - a}{b - a}\right) = -(x - b)(x - a).$$

3.4. Prawdopodobieństwo wyjścia przez b

Chcieliśmy jeszcze sprawdzić, jak często wyjście nastąpi przez większy z końców przedziału, czyli

$$\mathbb{P}(B_{\tau_x}^x = b).$$

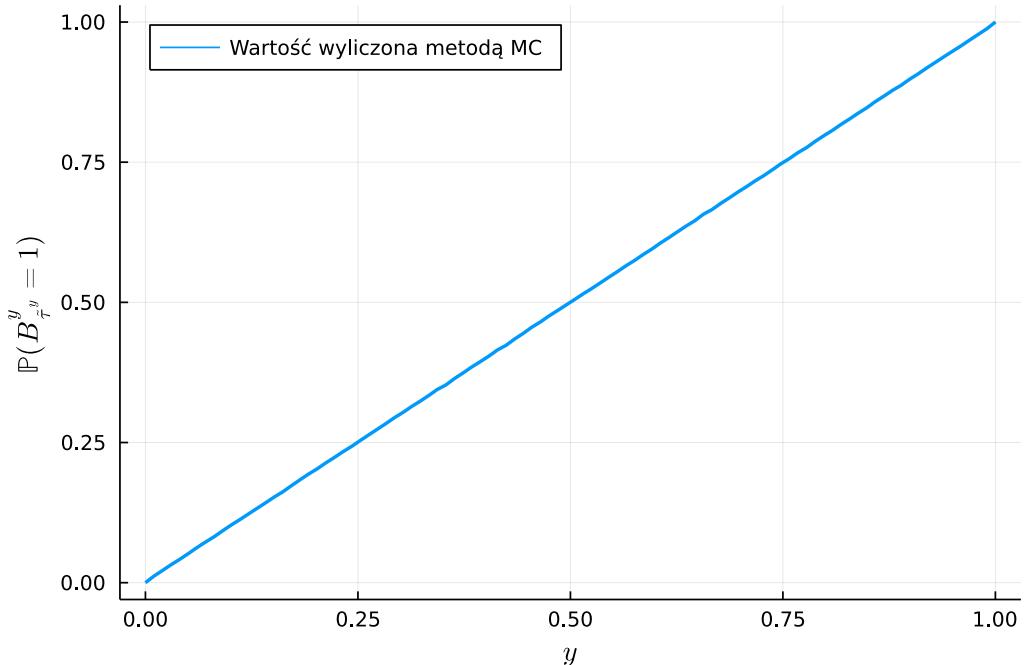
Korzystając z własności udowodnionych wcześniej, przekształciliśmy powyższy wzór do postaci

$$\mathbb{P}(B_{\tilde{\tau}^y}^y = 1),$$

gdzie $\tilde{\tau}^y$ jest czasem wyjścia procesu z przedziału $[0, 1]$. By oszacować to prawdopodobieństwo, ponownie skorzystaliśmy z metody Monte Carlo. Niestety generując trajektorię procesu generujemy punkty dyskretne, sprawdzając czy w danym momencie $B_t^x \geq b$. Prawdopodobieństwo, że trafimy dokładnie w punkt b jest zatem znikome. Ponieważ $B_{\tau^x}^x$ może przyjmować jedynie dwie wartości, a oraz b ($b > a$), możemy przekształcić to prawdopodobieństwo, otrzymując

$$\mathbb{P}(B_{\tau^x}^x = b) = \mathbb{P}(B_{\tilde{\tau}^y}^y \geq 1). \quad (5)$$

Wykorzystując powyższy wzór, byliśmy w stanie wygenerować poniższy wykres.



Rysunek 9: Prawdopodobieństwo wyjścia przez punkt 1.

Na wykresie widoczna jest zależność liniowa $\mathbb{P}(B_{\tilde{\tau}^y}^y \geq 1) = y$ dla $y \in [0, 1]$. Korzystając z równości (5) oraz $y = (x - a)/(b - a)$, szukane prawdopodobieństwo możemy przybliżyć funkcją

$$\mathbb{P}(B_{\tau^x}^x = b) = \frac{x - a}{b - a}. \quad (6)$$

Podobnie jak przy wyznaczaniu wartości oczekiwanej, znany jest nam wynik analityczny^[1] dla procesu Wienera startującego w 0

$$\mathbb{P}(W_\tau = b) = \frac{a}{b + a},$$

gdzie τ to czas wyjścia tego procesu z przedziału $[-a, b]$. Uwzględniając przesunięcie punktu początkowego szukane przez nas prawdopodobieństwo wynosi dokładnie

$$\mathbb{P}(B_{\tau^x}^x = b) = \mathbb{P}(W_\tau = b - x) = \frac{x - a}{(b - x) + (x - a)} = \frac{x - a}{b - a}.$$

Powyższy wzór pokrywa ze wzorem (6) uzyskanym przy pomocy symulacji.

4. Podsumowanie

W zadaniu 1, korzystając z różnych metod statystycznych, udało nam się dopasować model klasycznego procesu Ryzyka do danych, co potwierdziło między innymi porównanie funkcji średniej procesu ze średnią z trajektorii z danych. Następnie wykorzystaliśmy wyznaczony model do oszacowania prawdopodobieństwa ruiny. Dokonaliśmy tego korzystając z metody Monte Carlo, a w przypadku nieskończonego czasu, ogromnie przydatny okazał się wzór Pollaczka-Chinczyna. Zgodnie z intuicją okazało się, że wraz ze zwiększeniem czasu, prawdopodobieństwo ruiny rośnie, jednak nie może przekroczyć tego dla nieskończonego czasu.

Celem drugiego zadania było oszacowanie dwóch funkcji. Zanim do tego przystąpiliśmy, pokazaliśmy, że rozpatrywany przedział mogliśmy wybrać swobodnie. W prosty sposób można uogólnić otrzymane wyniki na dowolny inny przedział. Obie funkcje oszacowaliśmy metodą Monte Carlo. W ten sposób otrzymaliśmy szukane funkcje

$$\mathbb{E}\tau^x = -(x - b)(x - a) \quad \text{oraz} \quad \mathbb{P}(B_{\tau^x}^x = b) = \frac{x - a}{b - a}.$$

Otrzymane wyniki bardzo dobrze odzwierciedlają wartości teoretyczne.

Dodatkowo podczas symulacji zauważaliśmy, jak ważne jest generowanie procesu Wienera z odpowiednio małym krokiem czasowym. Wraz ze wzrostem wielkości kroku, rósł średni czas wyjścia procesu. By uzyskać dokładne dane, generowaliśmy proces z krokiem $h = 10^{-4}$.

Źródła

[1] [https://galton.uchicago.edu/~lalley/Courses/313/
BrownianMotionCurrent.pdf](https://galton.uchicago.edu/~lalley/Courses/313/BrownianMotionCurrent.pdf)