

Wybrane aspekty podstaw rachunku prawdopodobieństwa i statystyki

Metoda Monte Carlo jest metodą stochastyczną, której podstawą działania jest losowość opisana przez rachunek prawdopodobieństwa i statystykę. Omówimy teraz podstawowe zagadnienia z obu tych działów matematyki, będą się one przewijać w trakcie całego wykładu.

Co nam będzie potrzebne

- definicja prawdopodobieństwa i jego własności
- miara prawdopodobieństwa (funkcja gęstości prawdopodobieństwa, dystrybuanta)
- metody wyznaczania momentów i ich interpretacja (wartość oczekiwana, wariancja, odchylenie standardowe)
- fundament metody MC: prawo wielkich liczb oraz centralne twierdzenie graniczne
- metody określania poziomu wiarygodności wyniku eksperymentu numerycznego (przedziały ufności, określanie minimalnej liczności populacji)

Literatura

- M.H. Kalos, P.A. Whitlock „Monte Carlo Methods”
- W.L. Dunn, J.K. Shultis „Exploring Monte Carlo Methods”
- A. Haghigat „Monte Carlo Methods for particle transport”

Zmienna losowa – to wynik zdarzenia losowego, który może mieć wymiar

- liczbowy (wynik rzutu kostką, liczba zliczeń w detektorze, wartość zmiennej o rozkładzie ciągłym)

$$X \in \{1, 2, 3, 4, 5, 6\} \quad X \in [a, b]$$

- logiczny/binarny (losowanie bitów 0-1, prawda/fałsz, rzut monetą)

$$X \in \{0, 1\} \quad X \in \{\text{prawda, fałsz}\} \quad X \in \{\text{orzeł, reszka}\}$$

- nieliczbowy (losowanie karty z tali, kolorów), ale z możliwością przypisania wynikowi unikatowej wartości liczbowej (1,2,..)

$$X \in \{\blacksquare, \color{red}\square, \color{green}\square, \color{blue}\square, \dots\} \implies \{1, 2, 3, \dots\}$$

$$X \in \{2\clubsuit, 3\clubsuit, 4\clubsuit, \dots\} \implies \{1, 2, 3, \dots, \}$$

Zazwyczaj interpretacja pojedynczego wyniku zdarzenia losowego nie ma znaczenia. Istotny (statystycznie) jest opis wyniku serii zdarzeń losowych, do tego wykorzystamy narzędzia jakie daje nam statystyka - wnioskowanie na podstawie informacji zawartych w generowanej przy użyciu Monte Carlo próbce danych.

Zjawisko losowe – proces stochastyczny generujący ciąg **zmiennych losowych** $\{X_1, X_2, X_3, \dots\}$, których wartości nie możemy przewidzieć, ale możemy przypisać im określone prawdopodobieństwo wystąpienia $\{p_1, p_2, p_3, \dots\}$.

$$\underbrace{X_1}_{p_1}, \underbrace{X_2}_{p_2}, \underbrace{X_3}_{p_3}, \dots, \underbrace{X_n}_{p_n} \quad P\{X_i\} = p_i$$

Prawdopodobieństwo p_i jest granicą, do której dąży asymptotycznie częstość wystąpień zmiennej X_i

$$\lim_{N \rightarrow \infty} \frac{n_i}{N} = p_i$$

Uwaga: należy pamiętać o różnicy pomiędzy tymi wielkościami

- prawdopodobieństwo jest określone w sposób ścisły (dokładny)
- częstość wystąpień to zmienna losowa i podlega pewnym fluktuacjom statystycznym
- równość tylko w granicy, której nigdy na komputerze nie osiągniemy

- prawdopodobieństwo zdarzenia losowego ma określone własności

$$0 \leq p_i \leq 1$$

$p_i = 0 \implies$ zdarzenie nigdy nie zajdzie

$p_i = 1 \implies$ zdarzenie pewne

- prawdopodobieństwo otrzymania jakiegokolwiek wyniku losowego (normalizacja prawdopodobieństwa)

$$\sum_{i=1}^N p_i = 1$$

- prawdopodobieństwo zdarzenia losowego a prawdopodobieństwo jego braku

$$P\{X_i\} = p_i \quad P\{\bar{X}_i\} = 1 - p_i$$

- prawdopodobieństwo wystąpienia dwóch zdarzeń losowych

$$P\{X_i \wedge X_j\} \leq p_i + p_j$$

- dwa zdarzenia losowe wzajemnie się wykluczają gdy

$$P\{X_i \wedge X_j\} = 0 \quad P\{X_i \vee X_j\} = p_i + p_j$$

- **prawdopodobieństwo łączne wystąpienia dwóch zdarzeń** (pr. łączne – „**joint probability**”)

$P\{X_i, X_j\} = p_{ij}$ - wystąpienie zdarzenia i-tego zależy o wystąpienia j-tego i vice versa

$p_{ij} = p_i p_j$ - zdarzenia niezależne, każde opisywane swoim prawdopodobieństwem niezależnie od drugiego

przykład:

zdarzenia niezależne w rzucie kostką:

wylosowanie „2” oraz „5”

zdarzenie zależne w rzucie kostką:

suma dwóch wylosowanych liczb jest nieparzysta
- jeśli pierwsza wylosowana jest parzysta
to druga musi być nieparzysta i *vice versa*

- **prawdopodobieństwo zredukowane** („**marginal probability**”) - $p(i)$

$p(i)$ – **prawdopodobieństwo zredukowane** określa wystąpienie zdarzenia i-tego niezależnie od wystąpienia innych zdarzeń mających na nie wpływ (zredukowaliśmy informację o innych zdarzeniach – **marginalizacja pozostałych zdarzeń**)

$$p_{ij} = p_{ij} \frac{\sum_k p_{ik}}{\sum_k p_{ik}} = \left(\sum_k p_{ik} \right) \left(\frac{p_{ij}}{\sum_k p_{ik}} \right) = p(i) \left(\frac{p_{ij}}{\sum_k p_{ik}} \right) \implies p(i) = \sum_k p_{ik}$$

Oczywiście nadal obowiązuje warunek normalizacji prawdopodobieństwa (wystąpienie jakiegokolwiek zdarzenia jest pewne)

$$\sum_i p(i) = \sum_i \sum_k p_{ik} = 1$$

- **prawdopodobieństwo warunkowe (zależne)**
określa prawdopodobieństwo wystąpienia zdarzenia X_j
pod warunkiem że wcześniej wystąpiło zdarzenie X_i

$$p(j|i) = \frac{p_{ij}}{\sum_k p_{ik}} \quad \begin{array}{l} \text{- licznik określa prawdopodobieństwo łączne } X_j \text{ i } X_i \\ \text{- mianownik określa całkowite prawdopodobieństwo} \\ \text{wystąpienia } X_i \end{array}$$

lub dla większej liczby zdarzeń: $p(j|i_1, i_2, i_3, \dots, i_n)$

uwaga 1: $p_{ij} = p_{ji}$ - to informacja globalna o możliwości wystąpienia dwóch zdarzeń
(niezależnie od kolejności – więc uwzględnia wszystkie możliwe wystąpienia obu zdarzeń)

uwaga 2: w ogólności musimy założyć, że

$$p(j|i) \neq p(i|j)$$

co wynika z definicji (mianowniki mogą być różne)

$$p(j|i) = \frac{p_{ij}}{\sum_k p_{ik}} = \frac{p_{ij}}{p(i)} \quad p(i|j) = \frac{p_{ij}}{\sum_k p_{jk}} = \frac{p_{ij}}{p(j)}$$

uwaga 3: jeśli wystąpiło zdarzenie X_i to wystąpienie dowolnego innego zdarzenia zależnego X_j jest pewne
(rozszerzony warunek normalizacji)

$$\sum_j p(j|i) = \sum_j \frac{p_{ij}}{\sum_k p_{ik}} = \frac{\sum_j p_{ij}}{\sum_k p_{ik}} = 1$$

przykład:

rzut kostką – wyrzuciliśmy „2”, jakie jest prawdopodobieństwo, że po wyrzuceniu kolejnej liczby X_j ich suma będzie nie większa niż 8?

Dla ciągu **dyskretnych zmiennych losowych** i przypisanych im prawdopodobieństw

$$\{X_1, X_2, \dots, X_n\} \rightarrow \{P\{X_1\}, P\{X_2\}, \dots, P\{X_n\}\}$$

wielkości istotne statystycznie to

- **wartość oczekiwana (przeciętna) zmiennej losowej**

$$E(X) \equiv \sum_i P\{X = x_i\}x_i = \sum_i p_i x_i = \mu$$

inne stosowane oznaczenie to

$$\langle x \rangle = \mu$$

jeśli zmienna losowa jest argumentem funkcji

$$g(x_i) = g_i$$

to wartość funkcji staje się także zmienną losową (dokonujemy transformacji zmiennej)
i możemy wyznaczyć jej wartość oczekiwaną

$$E(g(X)) = \sum_i p_i g(x_i)$$

przykład: rzut kostką

$$X_i \in \{1, 2, 3, 4, 5, 6\}$$

$$g(x) = x^2$$

$$E(g(x)) = ?$$

Własności wartości oczekiwanej

- funkcja stała $g(x) = \text{const} = C \quad \rightarrow \quad \langle g(x) \rangle = \sum_i p_i C = C \sum_i p_i = C$

- rozdzielność względem sumy

$$a_1, a_2 = \text{const}, \quad \rightarrow \quad \langle a_1 g_1(X) + a_2 g_2(X) \rangle = a_1 \langle g_1(X) \rangle + a_2 \langle g_2(X) \rangle$$

- dla funkcji liniowej

$$g(x) = a \cdot x, \quad \rightarrow \quad \langle g(X) \rangle = g(\langle X \rangle)$$

- n-ty moment zmiennej losowej**

$$\langle X^n \rangle \equiv \sum_i p_i x_i^n, \quad n = 1, 2, 3, \dots$$

- n-ty moment centralny zmiennej losowej**

$$\langle (X - \mu)^n \rangle \equiv \sum_i p_i (x_i - \mu)^n, \quad n = 1, 2, 3, \dots$$

drugi moment centralny to **wariancja** i ma szczególne znaczenie dla analizy wyników

$$\begin{aligned} \text{var}\{X\} &= \langle (X - \mu)^2 \rangle = \langle (X - \langle X \rangle)^2 \rangle = \langle X^2 - 2X \underbrace{\langle X \rangle}_{const} + \underbrace{\langle X \rangle^2}_{const} \rangle = \langle X^2 \rangle - 2\langle X \rangle \langle X \rangle + \langle X \rangle^2 \\ &= \langle X^2 \rangle - \langle X \rangle^2 \end{aligned}$$

częściej oznaczamy ją jako

$$\sigma^2 = \text{var}\{X\}$$

miarą rozrzutu zmiennej losowej X wokół jej wartości oczekiwanej jest **odchylenie standardowe** σ

$$\sigma = \sqrt{\text{var}\{X\}}$$

W metodzie MC zawsze staramy minimalizować odchylenie standardowe,
 najprościej: zwiększając liczbę losowań/powtórzeń,
 lub bardziej roztropnie: modyfikując metodę
 (szczegóły w dalszej części wykładu)

Identycznie możemy określić wariancję zmiennej zdefiniowanej jako funkcja,
 której argumentem jest zmienna losowa

$$\text{var}\{g(X)\} = \langle (g(X) - \langle g(X) \rangle)^2 \rangle = \langle g^2(X) \rangle - \langle g(X) \rangle^2$$

Elementy pojedynczego ciągu zmiennych losowych $\{X_i\}$ możemy wykorzystać do wyznaczenia wartości dwóch funkcji

$$g_1(x) \neq g_2(x)$$

Uwaga:

- wartości obu funkcji będą losowe ale mogą być od siebie zależne – bo korzystamy z tego samego ciągu argumentów
- wartości $\{g_1(X)\}$ oraz $\{g_2(Y)\}$ byłyby niezależne gdybyśmy je utworzyli z dwóch rozdzielnych ciągów $\{X_i\}$ oraz $\{Y_i\}$

Z definicji wartości oczekiwanej wynika

$$\langle a_1 g_1(X) + a_2 g_2(Y) \rangle = a_1 \langle g_1(X) \rangle + a_2 \langle g_2(Y) \rangle$$

niezależnie od tego czy oba ciągi $\{X\}$ są identyczne czy różne.

Policzmy wariancję sumy (**wariancja jest operacją nieliniową**)

$$\text{var}\{a_1 g_1(X) + a_2 g_2(Y)\} = a_1^2 \text{var}\{g_1(X)\} + a_2^2 \text{var}\{g_2(Y)\} + 2a_1 a_2 \langle g_1(X) g_2(Y) \rangle - 2a_1 a_2 \langle g_1(X) \rangle \langle g_2(Y) \rangle$$

widzimy że dla $\{X\} \neq \{Y\} \quad \rightarrow \quad \langle g_1(X) g_2(Y) \rangle = \langle g_1(X) \rangle \langle g_2(Y) \rangle$

$$\text{var}\{a_1 g_1(X) + a_2 g_2(Y)\} = a_1^2 \text{var}\{g_1(X)\} + a_2^2 \text{var}\{g_2(Y)\}$$

ale dla $\{X\}=\{Y\}$ dwa ostatnie wyrazy nie kasują się - **dwa ostatnie wyrazy są miarą niezależności/zależności obu zmiennych losowych**

$$\langle XY \rangle = \sum_i \sum_j p_{ij} x_i y_j$$

dla **zmiennych niezależnych** dostajemy warunek

$$p_{ij} = p_{x,i} \cdot p_{y,j} \quad \rightarrow \quad \langle XY \rangle = \sum_i p_{x,i} x_i \sum_j p_{y,j} y_j = \langle X \rangle \langle Y \rangle$$

- miarą (nieunormowaną) stopnia zależności zmiennych jest **kowariancja**

$$\text{cov}\{X, Y\} = \langle XY \rangle - \langle X \rangle \langle Y \rangle$$

$$\text{cov}\{X, Y\} = 0 \quad \implies \quad \{X\}, \{Y\} - \text{zmiennne niezależne}$$

$$\text{cov}\{X, X\} = \text{var}\{X\}$$

uwaga: zerowa kowariancja nie gwarantuje niezależności obu zmiennych

przykład:

$$X \in \{-1, 0, 1\}, \quad Y = X^2$$

$$\langle X \rangle = 0, \quad \langle Y \rangle - \text{nieistotne}$$

$$\langle XY \rangle = \langle X^3 \rangle = 0 \quad - \text{bo wkłady od wartości dodatnich i ujemnych kasują się}$$

$$\text{cov}\{X, Y\} = \langle XY \rangle - \langle X \rangle \langle Y \rangle = 0$$

- a przecież Y jest funkcją X,
- wynik może prowadzić do błędnej interpretacji (**niezależność X i Y**)
- źródłem problemu jest w tym przypadku symetria ciągu i wzoru $\text{cov}()$

- unormowaną miarą stopnia zależności/korelacji zmiennych jest **współczynnik korelacji**

$$r(X, Y) = \frac{\text{cov}\{X, Y\}}{\sqrt{\text{var}\{X\}\text{var}\{Y\}}}, \quad r(X, Y) \in [-1, 1]$$

uwaga: ponieważ kowariancja może dawać ujemny wkład do wariancji sumy dwóch zmiennych losowych, można ten fakt wykorzystać w metodzie Monte Carlo do redukcji błędu (niepewności) wyniku – metoda zmiennych antytetycznych

Rozkład Bernoulliego

Rozkład Bernoulliego opisuje proces, którego wynik może przyjmować jedynie dwie wartości (np.: rzut monetą, przejście cząstki przez tarczę)

$$X \in \{0, 1\} \qquad (X \in \{a, b\})$$

obu zmiennym możemy przypisać prawdopodobieństwo wylosowania

$$P\{0\} = 1 - p = q, \quad P\{1\} = p, \quad p + q = 1$$

wówczas **rozkład prawdopodobieństwa** opisuje

$$p_r(n) = p^n(1 - p)^{1-n}, \quad n = 0, 1$$

Wartość oczekiwana zmiennej (pierwszy moment)	$\langle X \rangle = (1 - p) \cdot 0 + p \cdot 1 = p$
drugi moment	$\langle X^2 \rangle = p$
wariancja	$var\{X\} = \langle X^2 \rangle - \langle X \rangle^2 = p - p^2 = p(1 - p)$

Różnica pomiędzy **prawdopodobieństwem wylosowania zmiennej** a **rozkładem prawdopodobieństwa**

- prawdopodobieństwo wylosowania oznacza jak często pojawi się w wyniku dana zmienna losowa, gdy liczba losowań będzie dążyła do ∞ , dotyczy tylko tej jednej zmiennej (bardzo ograniczona informacja) i stanowi tylko wycinek informacji o rozkładzie
- rozkład prawdopodobieństwa określa prawdopodobieństwo wylosowania dowolnej zmiennej, zawiera pełną informację o rozkładzie
- z rozkładu prawdopodobieństwa możemy wyznaczyć prawdopodobieństwo wylosowania dowolnej zmiennej, informacje o pozostałych zmiennych musimy **zmarginalizować/zredukować**

Rozkład dwumianowy – bazuje na rozkładzie Bernoulliego

wartości zmiennej: $X \in \{0, 1\}$ 0 - porażka (reszka), 1 - sukces (orzeł)

rozkład prawdopodobieństwa: $P\{X_0\} = q = 1 - p$, $P\{X_1\} = p$, $P\{X_0\} + P\{X_1\} = 1$

$\langle X \rangle = p$ $\langle X^2 \rangle = p$ $var\{X\} = \langle X^2 \rangle - \langle X \rangle^2 = p - p^2 = p(1 - p)$

Definiujemy nową zmienną losową $Y = \sum_{i=1}^N X_i$, $Y \in \{0, 1, 2, \dots, N\}$ - czyli w ciągu N zmiennych mamy n „1” oraz (N-n) „0”

Dyskretny rozkład prawdopodobieństwa nowej zmiennej - **rozkład dwumianowy**

$$P\{Y = n\} = \binom{N}{n} p^n (1 - p)^{N-n}$$

- n jedynek więc reszta ciągu to (N-n) zer
- **dłaczego pojawił się tu symbol Newtona?**

zobaczymy możliwe realizacje dla krótkiego ciągu (N,n)=(4,2)

$$\{0, 0, 1, 1\}, \quad \{0, 1, 0, 1\}, \quad \{1, 0, 0, 1\}, \quad \{1, 0, 1, 0\}, \quad \{1, 1, 0, 0\}, \quad \{0, 1, 1, 0\} \quad \binom{4}{2} = 6$$

każdy z tych ciągów jest inny – rozróżnialny i musimy ten fakt uwzględnić

wartość oczekiwana zmiennej Y $\langle Y \rangle = \sum_{n=0}^N y_n P\{Y_n\} = \sum_{n=0}^N n \binom{N}{n} p^n (1 - p)^{N-n}$

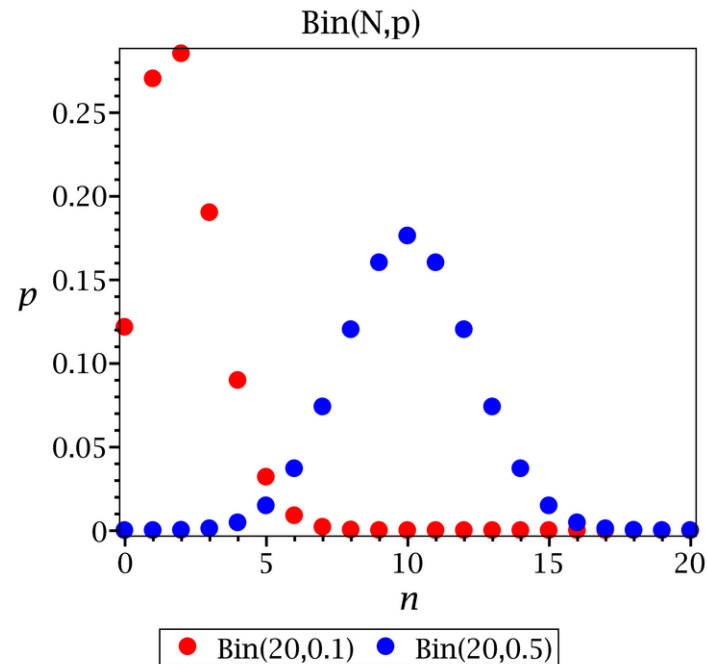
każda zmienna w sumie jest niezależna o identycznej wartości oczekiwanej

$$= \sum_{i=0}^N \langle X_i \rangle = \sum_{i=0}^N p = Np$$

podobnie liczymy **wariancję** pamiętając, że zmienne X_i tworzące ciąg $\{X_1, X_2, \dots\}$ są **niezależne** (z definicji to suma wariancji)

$$\text{var}\{Y\} = \langle (Y - Np)^2 \rangle = \sum_{i=1}^N \text{var}\{X_i\} = \sum_{i=1}^N \underbrace{p(1-p)}_{\text{Bernoulli}} = Np(1-p)$$

Przykładowy rozkład prawdopodobieństwa dla $\text{Bin}(N,p)$



Ponieważ wartość oczekiwana to $\langle Y \rangle = Np$ więc maksimum rozkładu lokalizuje się w jej bezpośrednim otoczeniu.

Rozkład geometryczny

- przyjmujemy jak w rozkładzie dwumianowym $X \in \{0, 1\}$, $P\{X_0\} = q$, $P\{X_1\} = p$, $q + p = 1$
- generujemy ciągi w których (n-1) pierwszych zmiennych to „0” a n-ta zmienna to „1” - wtedy kończymy generowanie ciągu

$$Y_n = \{0, 0, 0, \dots, 1\}$$

- prawdopodobieństwo wygenerowania ciągu o długości n

$$P\{Y = n\} = q^{n-1}p, \quad n = 1, 2, 3, \dots$$

- wartość oczekiwana

$$\langle X \rangle = \sum_{i=1}^{\infty} n q^{n-1} p = \frac{p}{(1-q)^2} = \frac{1}{p}$$

(korzystamy z tożsamości)

$$\sum_{n=1}^{\infty} n x^{n-1} = \frac{d}{dx} \left(\sum_{n=0}^{\infty} x^n \right) = \frac{d}{dx} \left(\frac{1}{1-x} \right) = \frac{1}{(1-x)^2}$$

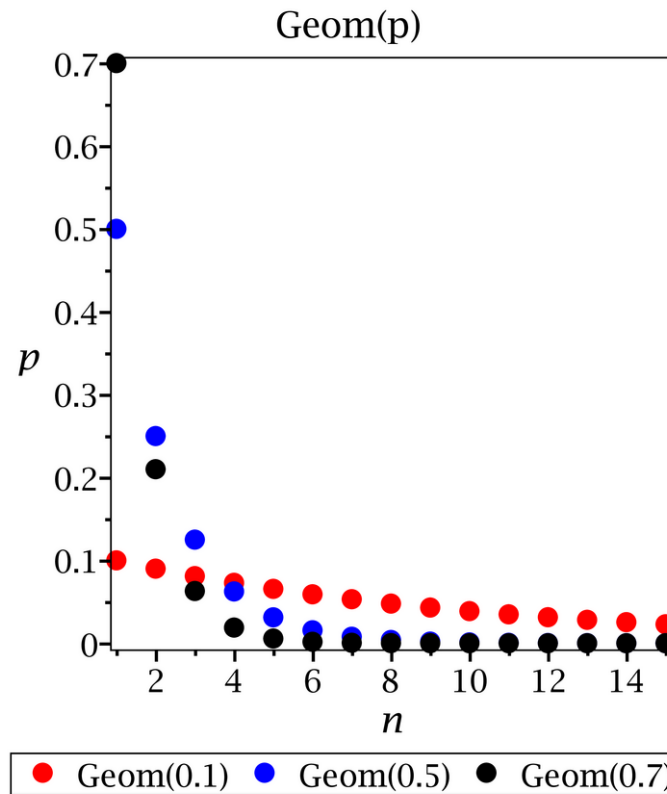
- drugi moment zmiennej X

$$\langle X^2 \rangle = \sum_{n=1}^{\infty} n^2 q^{n-1} p = \sum_{n=1}^{\infty} (n(n-1)q^{n-2} \cdot qp + nq^{n-1}p) = \left(\frac{2}{p^2} - \frac{1}{p} \right)$$

- tu szereg różniczkujemy dwukrotnie

- wariancja zmiennej X

$$\text{var}\{X\} = \langle X^2 \rangle - \langle X \rangle^2 = \frac{1}{p^2} - \frac{1}{p}$$

Przykładowy rozkład geometryczny $G(p)$ 

- jeśli prawdopodobieństwo sukcesu jest małe to musimy generować długie ciągi aby osiągnąć sukces
- im wyższe p tym krótsze (statystycznie) będą generowane ciągi

Rozkład Poissona

Rozkład ten opisuje zdarzenia o bardzo małym prawdopodobieństwie sukcesu ($p \ll 1$).

przykład: liczba rejestrowanych rozpadów promieniotwórczych w równych przedziałach czasu dla materiału mało aktywnego (pojedyncze impulsy)

Punktem wyjścia jest rozkład dwumianowy z warunkami

$$N \gg 1, \quad n \ll N, \quad p \ll 1, \quad \left(p = \frac{m}{N}\right)$$

$$\begin{aligned} p_r(n) &= \binom{N}{n} p^n (1-p)^{N-n} = \frac{\overbrace{(N-1)(N-2)\dots(N-n+1)}^{\approx N^n}}{n!} \left(\frac{m}{N}\right)^n \left(1 - \frac{m}{N}\right)^{N-n} \\ &\approx \frac{N^n}{n!} \left(\frac{m}{N}\right)^n \left(1 - \frac{m}{N}\right)^{N-n} \\ &= \frac{m^n}{n!} \left(1 - \frac{m}{N}\right)^N \left(1 - \frac{m}{N}\right)^{-n} \end{aligned}$$

rozwińnięcie w szereg Taylora funkcji eksponencjalnej

$$e^{-m} = e^{-\frac{m}{N}N} = \left(e^{-\frac{m}{N}}\right)^N = \left(\sum_{k=0}^N \frac{1}{k!} \frac{(-m)^k}{N^k}\right)^N = \left(1 - \frac{m}{N} + \underbrace{\frac{1}{2} \left(\frac{m}{N}\right)^2 + \dots}_{\approx 0 \iff N \gg m}\right)^N \approx \left(1 - \frac{m}{N}\right)^N$$

czyli możemy użyć przybliżenia

$$\lim_{N \rightarrow \infty} \left(1 - \frac{m}{N}\right)^N = e^{-m}$$

Drugi wyraz zapisujemy w jawnej postaci

$$\left(1 - \frac{m}{N}\right)^{-n} = \frac{1}{\left(1 - \frac{m}{N}\right)^n} = \frac{1}{1 - n\frac{m}{N} + \frac{n(n-1)}{2!}\frac{m^2}{N^2} - \dots + \frac{(-m)^n}{N^n}} \approx 1$$

ze względu na założenia:

$$n, m \sim 1$$

$$\frac{n}{N} \ll 1$$

uwzględniając oba wyniki dostajemy wyrażenie definiujące rozkład Poissona

$$p_r(n) = \frac{m^n}{n!} e^{-m}$$

- m to parametr rozkładu

wartość oczekiwana rozkładu

$$\langle n \rangle = m$$

$$\langle n \rangle = \sum_{n=0}^N p_r(n) n = \sum_{n=1}^N \frac{m^n}{n!} e^{-m} n = m e^{-m} \sum_{k=0}^{N-1} \frac{m^k}{k!} \approx m e^{-m} e^m = m$$

drugi moment

$$\langle n^2 \rangle = m^2 + m$$

$$\begin{aligned} \langle n^2 \rangle &= \sum_{n=0}^N p_r(n) n^2 = \sum_{n=1}^N \frac{m^n}{n!} e^{-m} n^2 = m e^{-m} \sum_{n=1}^N \frac{m^{n-1}}{(n-1)!} n \\ &= m e^{-m} \sum_{k=0}^N \frac{m^k}{k!} (k+1) = m e^{-m} (m+1) e^m = m^2 + m \end{aligned}$$

wariancja rozkładu Poissona

$$\text{var}\{n\} = \langle n^2 \rangle - \langle n \rangle^2 = m$$

Dystrybuanta $F(x)$ i funkcja gęstości prawdopodobieństwa $f(x)$ rozkładu

Dla rozkładu dyskretnego definiowaliśmy

$$X \in \{X_1, X_2, \dots, X_n\} \quad P\{X_i\} = p_i$$

możemy zatem określić prawdopodobieństwo wylosowania zmiennej z podzbioru

$$P\{X \in \{X_1, X_2, \dots, X_k\}\} = P\{X \leq X_k\} = p_1 + p_2 + \dots + p_k = \sum_{i=1}^k p_i, \quad k \leq n$$

funkcję określającą tak skumulowane prawdopodobieństwo nazywamy **dystrybuantą rozkładu**

$$\underbrace{F(x_k) \equiv P\{X \leq X_k\}}_{\text{dyskretny}} \quad \vee \quad \underbrace{F(x_k) \equiv P(X \leq x_k)}_{\text{ciągły}}$$

własności dystrybuanty

$$F(x) \in [0, 1] \quad - \text{ unormowana do 1$$

$$F(x + \delta x) \geq F(x), \quad \delta x > 0 \quad - \text{ niemalejąca (i nieujemna)}$$

$$x \in (-\infty, \infty) \implies \begin{cases} \lim_{x \rightarrow -\infty} F(x) = 0 \\ \lim_{x \rightarrow +\infty} F(x) = 1 \end{cases} \quad x \in [a, b] \implies \begin{cases} F(a) = 0 \\ F(b) = 1 \end{cases}$$

z definicji dystrybuanty wynika też

$$x_2 > x_1 \implies P(x_1 \leq X \leq x_2) = P(X \leq x_2) - P(X \leq x_1) = F(x_2) - F(x_1)$$

Dla zmiennej losowej o rozkładzie ciągłym

$$X \in [a, b] \quad \vee \quad X \in (-\infty, \infty)$$

ostatnia własność pozwala obliczyć zmianę prawdopodobieństwa w niewielkim wycinku przestrzeni $(x, x+dx)$

$$P(X \leq x + \Delta x) - P(X \leq x) = F(x + \Delta x) - F(x) = \Delta F(x) \quad / : \Delta x$$

$$\lim_{\Delta x \rightarrow 0} \frac{\Delta F}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{F(x + \Delta x) - F(x)}{\Delta x} = \frac{dF}{dx} = f(x) \geq 0$$

Pochodna dystrybuanty określa lokalną gęstość prawdopodobieństwa, natomiast zależność funkcyjną $f(x)$ nazywamy **funkcją gęstości prawdopodobieństwa**

Obie wielkości są powiązane więc możemy wyznaczyć $F(x)$ z $f(x)$

$$\frac{dF}{dx} = f(x) \quad \implies \quad dF = f(x)dx \quad \implies \quad \int dF = F(x_k) = \int_{-\infty}^{x_k} f(x)dx$$

$$F(x_k) = \int_{-\infty}^{x_k} f(x)dx$$

Funkcja gęstości prawdopodobieństwa jednoznacznie definiuje rozkład prawdopodobieństwa oraz sposób jego wyznaczania (dystrybuanta) dlatego stanowi miarę prawdopodobieństwa.

W rozkładzie dyskretnym całka zastępowana jest sumą, a ciągła funkcja $f(x)$ dyskretnymi wartościami p_i lub możemy się posłużyć funkcją **delta Diraca**

$$f(x) = \sum_i \delta(x - x_i) p_i$$

wówczas możemy wykonać całkowanie

$$F(x_k) = \int_{-\infty}^{x_k} f(x) dx = \sum_i p_i \int_{-\infty}^{x_k} \delta(x - x_i) dx = \sum_{i \leq k} p_i$$

Własności fgp

- jest nieujemna

$$f(x) = \frac{dF}{dx} = \frac{dP}{dx} \implies dP(x) = f(x) dx \implies f(x) \geq 0, \quad x \in (-\infty, \infty)$$

- jest unormowana

$$\lim_{x_k \rightarrow \infty} (F(x_k) - F(-x_k)) = \int_{-\infty}^{\infty} f(x) dx = 1$$

- wartości $f(x)$ mogą być nieograniczone (przykład: delta Diraca)

$$f(x) = \delta(x) \implies \int_{-\infty}^{\infty} \delta(x) dx = 1$$

Możemy też definiować wielowymiarowe zmienne losowe (wektory losowe)

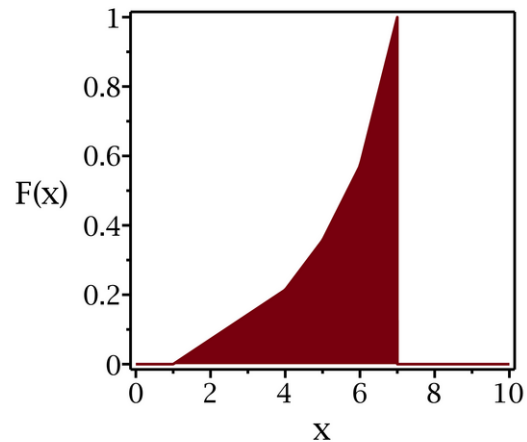
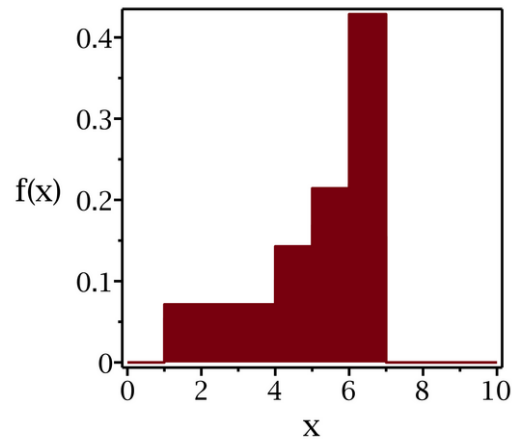
$$\vec{X} = [X_1, X_2, \dots, X_n] \quad X_i \in [a_i, b_i]$$

i dla takich zmiennych wielowymiarowych określać fgp z warunkiem normalizacji

$$f(\vec{x}) = f(x_1, x_2, \dots, x_n) \geq 0 \quad \implies \quad \int_{a_1}^{b_1} dx_1 \dots \int_{a_1}^{b_1} dx_n f(x_1, x_2, \dots, x_n) = 1$$

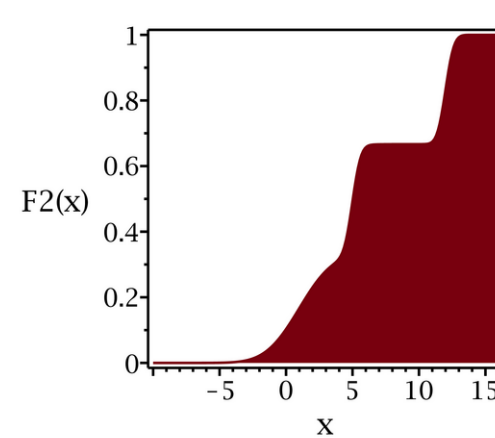
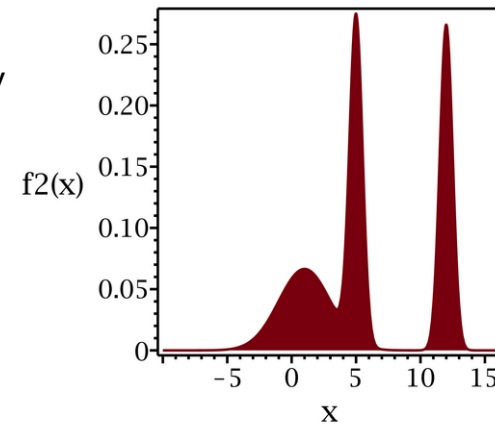
rozkład
skokowy

$x \in [1, 7]$



rozkład
nieograniczony

$x \in (-\infty, \infty)$



- wartość oczekiwana (pierwszy moment) zmiennej losowej o rozkładzie ciągłym

$$E(X) = \langle X \rangle \equiv \int_{-\infty}^{\infty} x dF(x) = \int_{-\infty}^{\infty} x f(x) dx$$

- drugi moment

$$E(X^2) = \langle X^2 \rangle = \int_{-\infty}^{\infty} x^2 f(x) dx$$

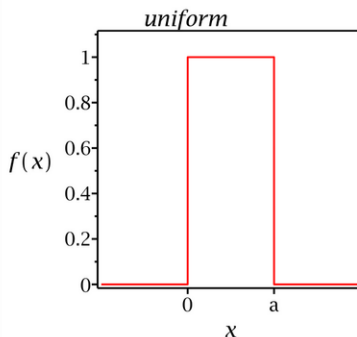
- wariancja

$$\text{var}\{X\} = E((X - E(X))^2) = E(X^2) - E(X)^2 = \langle X^2 \rangle - \langle X \rangle^2$$

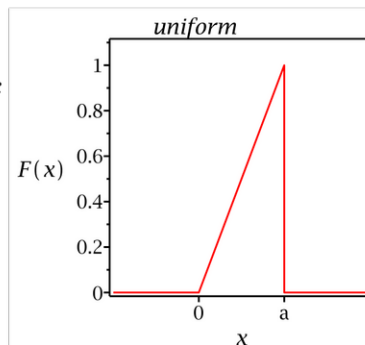
Przykładowe rozkłady prawdopodobieństwa

$$f(x) = \frac{1}{a}$$

$$x \in [0, a]$$



$$F(x) = x$$

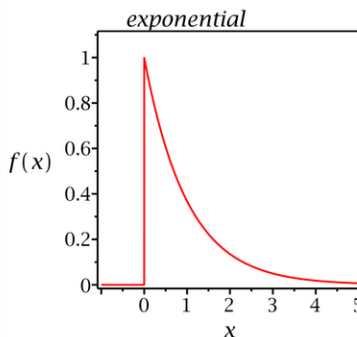


$$\langle X \rangle = \frac{a}{2}$$

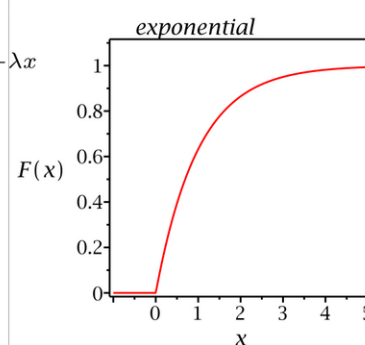
$$\text{var}\{X\} = \frac{1}{12}a$$

$$f(x) = \lambda e^{-\lambda x}$$

$$x \in [0, \infty)$$



$$F(x) = 1 - e^{-\lambda x}$$

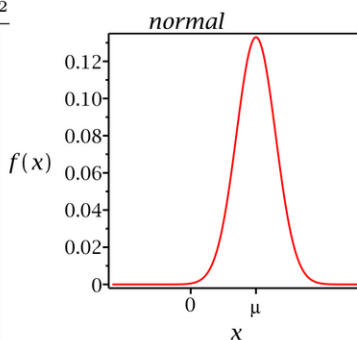


$$\langle X \rangle = \frac{1}{\lambda}$$

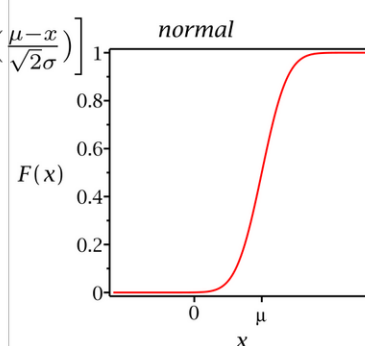
$$\text{var}\{X\} = \frac{1}{\lambda^2}$$

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$x \in (-\infty, \infty)$$



$$F(x) = \frac{1}{2} \left[1 - \text{erf}\left(\frac{\mu-x}{\sqrt{2}\sigma}\right) \right]$$

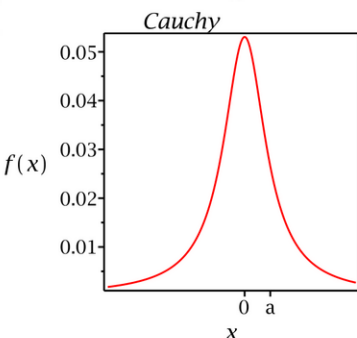


$$\langle X \rangle = \mu$$

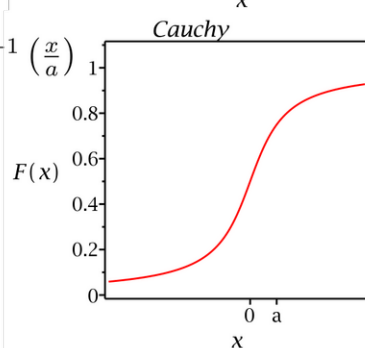
$$\text{var}\{X\} = \sigma^2$$

$$f(x) = \frac{a}{a^2+x^2}$$

$$x \in (-\infty, \infty)$$



$$F(x) = \frac{1}{2} + \frac{1}{\pi} \tan^{-1}\left(\frac{x}{a}\right)$$



$$\langle X \rangle = ?$$

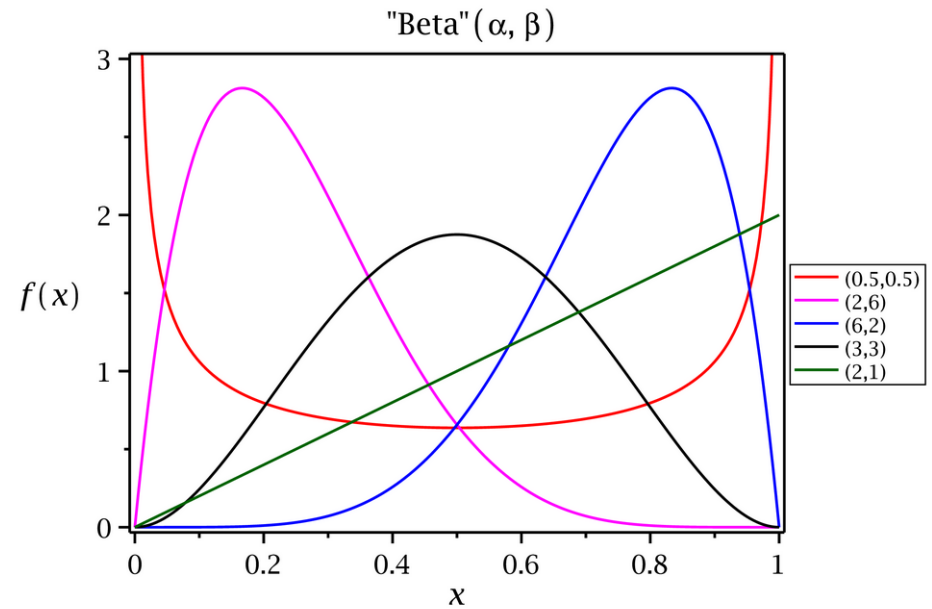
$$\text{var}\{X\} = \infty$$

Rozkład beta - dwuparametryczny

$$f(x, \alpha, \beta) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)}$$

$$B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha + \beta)} \quad \text{- beta Eulera}$$

$$\Gamma(x) \quad \text{- gamma Eulera}$$



Rozkład dzięki możliwości manipulacji dwoma parametrami jest bardzo elastyczny - możemy dopasować jego kształt do potrzeb symulacji Monte Carlo.

Zakres jest ograniczony do $[0,1]$ tylko pozornie.

Rozkład możemy przesunąć i rozciągnąć/ścisnąć wykonując transformację liniową

$$y = ax + b \quad \implies \quad x = \frac{y - b}{a} \quad y \in [b, b + a]$$

b – przesunięcie

a – skalowanie długości przedziału

Wartość funkcji jako zmienna losowa

Wiemy, że funkcja której argumentem jest zmienna losowa też staje się zmienną losową. Jak w takim przypadku wyznaczyć wartość oczekiwaną?

$$y = y(x) \quad \Longrightarrow \quad Y = y(X)$$

zastosujemy ścisłą definicję wartości oczekiwanej zmiennej Y

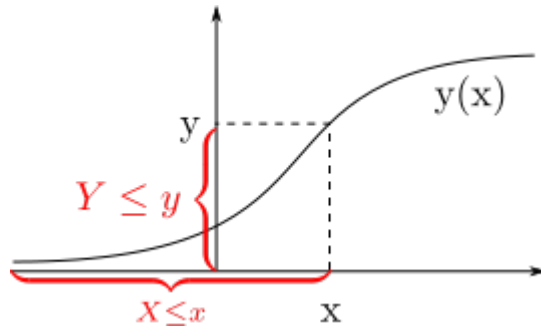
$$E(Y) = \int_{-\infty}^{\infty} y f_y(y) dy$$

pojawił się problem – nie znamy (zazwyczaj) funkcji gęstości prawdopodobieństwa $f_y(y)$

$$f_y(y) = ?$$

to częsty przypadek w symulacjach MC i musimy zastanowić się nad rozwiązaniem tego problemu

1-przypadek: funkcja niemalejąca



$$X \leq x \iff y(X) \leq y(x)$$

wobec tego oczekujemy spełnienia relacji

$$P(y(X) = Y \leq y) = P(X \leq x)$$

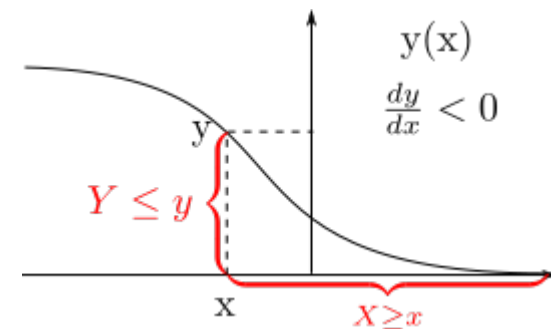
$$F_y(y) = F_x(x)$$

$$dF_y(y) = dF_x(x)$$

$$f_y(y) \frac{dy}{dx} = f_x(x)$$

$$f_y(y) = f_x(x) \frac{dx}{dy}, \quad \frac{dy}{dx} > 0$$

2-przypadek: funkcja nierosnąca



$$X \geq x \iff y(X) \leq y(x)$$

$$P(y(X) = Y \leq y) = P(X \geq x) = 1 - P(X < x)$$

$$F_y(y) = 1 - F_x(x)$$

$$dF_y(y) = -dF_x(x)$$

$$f_y(y) \frac{dy}{dx} = -f_x(x)$$

$$f_y(y) = -f_x(x) \frac{dx}{dy}, \quad \frac{dy}{dx} < 0$$

Oba przypadki możemy wyrazić za pomocą wyrażenia

$$f_y(y) = f_x(x) \left| \frac{dy}{dx} \right|^{-1} > 0$$

Warunek

$$f_y(y) = f_x(x) \left| \frac{dy}{dx} \right|^{-1} > 0$$

alternatywnie możemy też zapisać

$$|f_y(y)dy| = |f_x(x)dx| \quad \implies \quad f_y(y)|dy| = f_x(x)|dx|$$

Całkując obie fgp zakładamy, że $dx = |dx|$, $dy = |dy|$

$$E(Y) = \int_{-\infty}^{\infty} y f_y(y) dy = \int_{-\infty}^{\infty} y(x) f_x(x) \left| \frac{dy}{dx} \right|^{-1} dy = \int_{-\infty}^{\infty} y(x) f_x(x) dx$$

- pierwszy moment zmiennej Y

$$E(Y) = \langle Y \rangle = \int_{-\infty}^{\infty} y(x) f_x(x) dx$$

- drugi moment zmiennej Y

$$E(Y^2) = \langle Y^2 \rangle = \int_{-\infty}^{\infty} y^2(x) f_x(x) dx$$

- wariancja zmiennej Y

$$E((Y - E(Y))^2) = E(Y^2) - [E(Y)]^2 = \langle Y^2 \rangle - \langle Y \rangle^2$$

Przykład transformacji fgp – rozkład Cauchy-ego

$$f_x(x) = \frac{4}{\pi} \frac{1}{1+x^2}, \quad x \in (0, 1)$$

definiujemy funkcję $y(x)$ i nową zmienną losową

$$y(x) = \frac{1}{x}, \quad y \in (1, \infty)$$

$$\frac{dy}{dx} = -\frac{1}{x^2} = -y^2$$

$$f_y(y) = \frac{4}{\pi} \frac{1}{1+(y^{-1})^2} \frac{1}{y^2} = \frac{4}{\pi} \frac{1}{1+y^2}, \quad y \in (1, \infty)$$

uzyskaliśmy podobną zależność funkcyjną, ale to przypadek

Przykład transformacji fgp – rozkład normalny $N(0,1)$

$$f_x(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right), \quad x \in -\infty, \infty$$

wykonujemy transformację liniową

$$y(x) = \sigma x + \mu \quad \Longrightarrow \quad x = \frac{y - \mu}{\sigma}$$

$$f_y(y) = f_x\left(\frac{y - \mu}{\sigma}\right) \frac{1}{\sigma} = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(y - \mu)^2}{2\sigma^2}\right), \quad y \in (-\infty, \infty)$$

otrzymaliśmy rozkład normalnyo innych paramaterach $N(\mu, \sigma)$

Przykład transformacji fgp – rozkład jednorodny U(0,1) – przypadek 1

$$f_x(x) = 1, \quad x \in [0, 1]$$

$$y(x) = x^n \quad \Longrightarrow \quad x = y^{1/n}$$

$$f_y(y) = \left| \frac{1}{n} \right| y^{\frac{1}{n}-1}, \quad \begin{cases} y \in (0, 1) & \iff n > 0 \\ y \in (1, \infty) & \iff n < 0 \end{cases}$$

rozkład f_y może mieć osobliwość, np. dla $n > 1$ ale nadal jest całkowny

$$n = 2 \implies f_y(y) = \frac{1}{2\sqrt{y}} \quad E(Y) = \int_0^1 y f_y(y) dy = \int_0^1 \frac{1}{2} \sqrt{y} dy = \frac{1}{3}$$

Przykład transformacji fgp – rozkład jednorodny U(0,1) – przypadek 2

$$f_x(x) = 1, \quad x \in [0, 1]$$

$$y(x) = -\ln(x) \quad \Longrightarrow \quad x = e^{-y}, \quad y \in (0, \infty)$$

$$f_y(y) = f_x(x) \left| \frac{dy}{dx} \right|^{-1} = x = e^{-y}$$

To będzie jeden ze sposobów generowania zmiennej o rozkładzie eksponencjalnym → wykład: generatory liczb pseudolosowych.

Rozkład zmiennej losowej w dwóch wymiarach

Rozkład dwuwymiarowy definiujemy za pomocą dystrybuanty rozkładu

$$F(x, y) = P(X \leq x, Y \leq y)$$

o funkcji gęstości prawdopodobieństwa

$$f(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y}$$

- n-ty moment zmiennej losowej X (Y analogicznie) oraz jej wariancja

$$E(X^n) = \langle X^n \rangle = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^n f(x, y) dx dy \quad \text{var}\{X\} = \langle X^2 \rangle - \langle X \rangle^2$$

- kowariancja $\text{cov}(X, Y)$ i współczynnik korelacji liczymy z definicji

$$\text{cov}(X, Y) = \langle XY \rangle - \langle X \rangle \langle Y \rangle \quad r_{\text{corr}} = \frac{\text{cov}(X, Y)}{\sqrt{\text{var}\{X\} \text{var}\{Y\}}}$$

- dowolna funkcja $g(x, y)$ jest nową zmienną losową o n-tym momencie

$$g = g(x, y) \quad \implies \quad E([g(X, Y)]^n) = \langle [g(X, Y)]^n \rangle = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [g(x, y)]^n f(x, y) dx dy$$

$$\text{var}\{G\} = \langle G^2 \rangle - \langle G \rangle^2, \quad G = g(X, Y)$$

- rozkład możemy zredukować tj. usunąć informację o jednej zmiennej np. Y

$$F(x) = P(X \leq x) = \int_{-\infty}^x \int_{-\infty}^{\infty} f(x', y) dx' dy = \int_{-\infty}^x f_x(x') dx'$$

w ten sposób otrzymaliśmy zredukowaną fgp – tylko da zmiennej X

$$f_x(x) = \int_{-\infty}^{\infty} f(x, y) dy$$

wykorzystajmy ją do innego sformułowania fgp

$$f(x, y) = \frac{f(x, y)}{\int_{-\infty}^{\infty} f(x, y) dy} \int_{-\infty}^{\infty} f(x, y) dy = \frac{f(x, y)}{f_x(x)} f_x(x) = f(y|x) f_x(x)$$

fgp rozkładu jest iloczynem dwóch rozkładów jednowymiarowych

$$f(x, y) = f(y|x) f_x(x)$$

rozkładu warunkowego

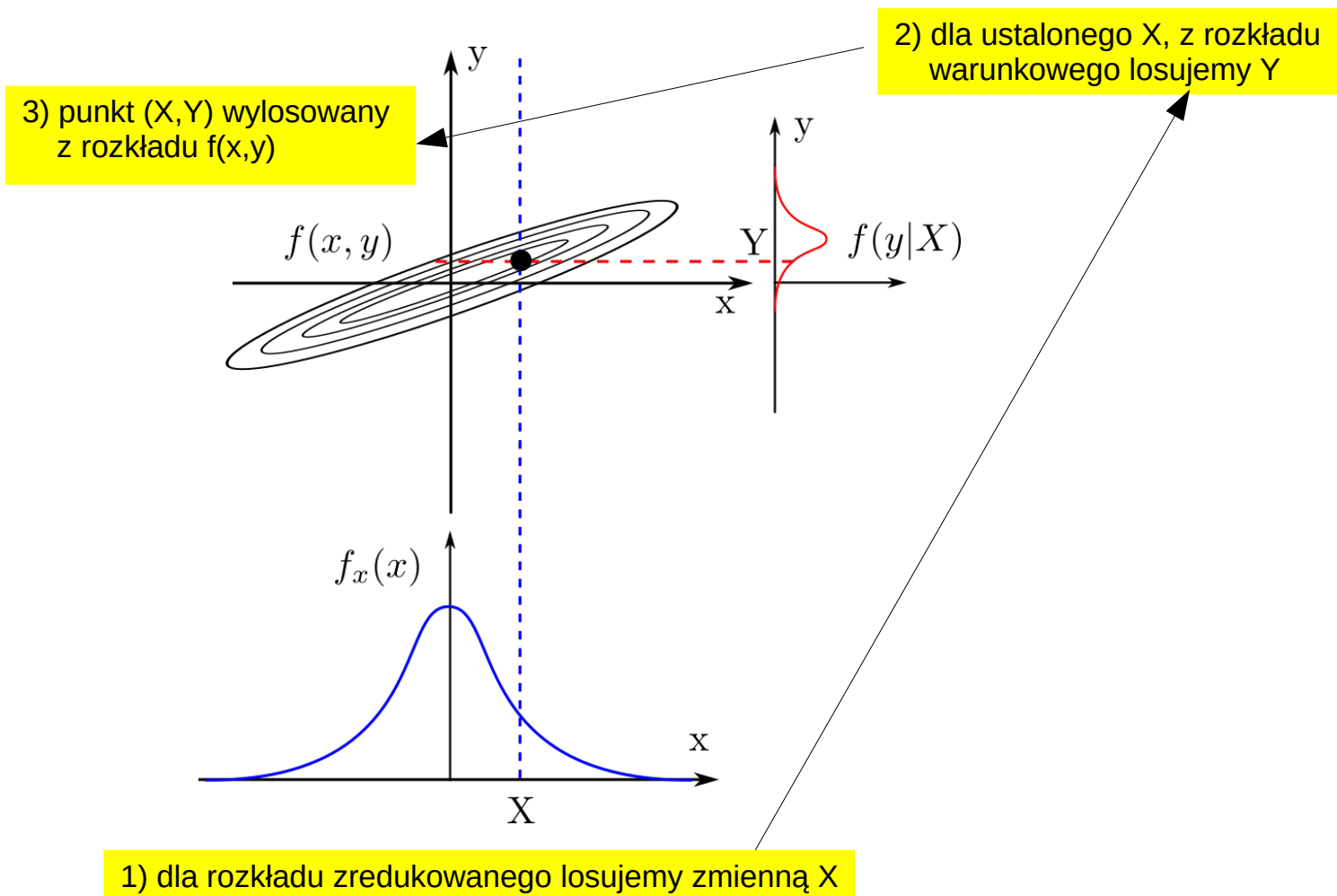
$$f(y|x) = \frac{f(x, y)}{f_x(x)}$$

rozkładu zredukowanego

$$f_x(x) = \int_{-\infty}^{\infty} f(x, y) dy$$

Taka separacja daje korzyści wówczas, gdy konieczne staje się losowanie zmiennych X i Y z fgp $f(x, y)$, wówczas losujemy najpierw X z rozkładu f_x a następnie Y z rozkładu warunkowego.

Przykład generowania zmiennej losowej z rozkładu $f(x,y)$ przy pomocy rozkładu zredukowanego f_x i rozkładu warunkowego $f(y|x)$



Teraz przejdziemy do omówienia dwóch twierdzeń będących fundamentem metody **Monte Carlo**

- prawa wielkich liczb
- centralnego twierdzenia granicznego (CTG)

Słaba forma prawa wielkich liczb

Rozważamy ciąg N zmiennych losowych

$$\{X_1, X_2, \dots, X_N\}$$

nieskorelowanych

$$\sigma^2(X_i + X_j) = \sigma_{X_i}^2 + \sigma_{X_j}^2 \quad \implies \quad \langle X_i X_j \rangle = \langle X_i \rangle \langle X_j \rangle \quad \implies \quad \text{cov}(X_i, X_j) = 0$$

Oznaczmy wartość oczekiwaną zmiennej losowej jako

$$\mu = \langle X \rangle = E(X)$$

jej estymatorem dla próbki N liczb jest średnia arytmetyczna

$$\bar{X} = \frac{1}{N} \sum_{i=1}^N X_i$$

Zgodnie ze słabym sformułowaniem prawa wielkich liczb relacja pomiędzy nimi jest następująca

$$\bigwedge_{\epsilon > 0} \lim_{N \rightarrow \infty} P\{|\bar{X} - \mu| \geq \epsilon\} = 0$$

- różnica między wartością oczekiwaną a średnią arytmetyczną znika w granicy nieskończonego N
- dla skończonych, ale dużych N różnica ta może być jednak duża, co wynika z faktu że

$$0 < P \ll 1$$

małe P oznacza tylko tyle, że takich przypadków będzie „niewiele” – ile? tego nie wiemy bo zakładamy, że N jest duże

Dowód twierdzenia

- zakładamy, że zmienna losowa X ma rozkład opisany $f(x)$ oraz skończone wartości μ i σ^2
- definiujemy nową zmienną losową (nieujemną) i liczbę dodatnią

$$g = g(x) \geq 0, \quad w \in R, \quad w > 0$$

- wartość oczekiwana $\langle g(x) \rangle$

$$\begin{aligned} \langle g(x) \rangle &= \int_{-\infty}^{\infty} g(x) f(x) dx = \int_{\{x:g(x) \geq w\}} g(x) f(x) dx + \int_{\{x:g(x) < w\}} g(x) f(x) dx \\ &\geq \int_{\{x:g(x) \geq w\}} g(x) f(x) dx \geq \int_{\{x:g(x) \geq w\}} w f(x) dx = w P\{g(x) \geq w\} \end{aligned}$$

$$\langle g(x) \rangle \geq w P\{g(x) \geq w\} \longrightarrow P\{g(x) \geq w\} \leq \frac{\langle g(x) \rangle}{w}$$

$g(x)$ i w przyjmujemy w postaci

$$g(x) = (x - \mu)^2$$

$$w = \lambda^2 \sigma^2$$

$$P\{(x - \mu)^2 \geq \lambda^2 \sigma^2\} = P\{|x - \mu| \geq \lambda \sigma\} \leq \frac{\langle (x - \mu)^2 \rangle}{\lambda^2 \sigma^2} = \frac{1}{\lambda^2}$$

- we wzorze występuje zmienna losowa x - musimy ją zastąpić średnią arytmetyczną (twierdzenie)
- średnia arytmetyczna jest uśrednioną sumą zmiennych losowych więc sama jest zmienną losową - jaką ma wartość oczekiwaną i wariancję?

- wartość oczekiwana średniej arytmetycznej (zakładamy, że zmienne mają identyczny rozkład: $\mu_i = \mu$)

$$\langle \bar{x} \rangle = \left\langle \frac{1}{N} \sum_i x_i \right\rangle = \frac{1}{N} \sum_i \langle x_i \rangle = \frac{1}{N} \sum_i \mu_i = \frac{N\mu}{N} = \mu \quad - \text{ jest identyczna jak zmiennej } X$$

- wariancja średniej arytmetycznej

$$\begin{aligned} \sigma_{\bar{x}}^2 &= \langle \bar{x}^2 \rangle - \langle \bar{x} \rangle^2 = \left\langle \frac{1}{N} \sum_i x_i \cdot \frac{1}{N} \sum_j x_j \right\rangle - \left\langle \frac{1}{N} \sum_i x_i \right\rangle \left\langle \frac{1}{N} \sum_j x_j \right\rangle \\ &= \frac{1}{N^2} \left[\sum_i \langle x_i^2 \rangle + \sum_i \sum_{j \neq i} \langle x_i x_j \rangle - \sum_i \langle x_i \rangle^2 - \sum_i \sum_{j \neq i} \langle x_i \rangle \langle x_j \rangle \right] \\ &= \frac{1}{N^2} \left[\sum_i \underbrace{(\langle x_i^2 \rangle - \langle x_i \rangle^2)}_{\sigma_i^2} + \sum_i \sum_{j \neq i} \underbrace{(\langle x_i x_j \rangle - \langle x_i \rangle \langle x_j \rangle)}_{\text{cov}(x_i, x_j)=0} \right] \\ &= \frac{1}{N^2} \sum_i \sigma_i^2 \end{aligned}$$

$$\sigma_{\bar{x}}^2 = \frac{1}{N^2} \sum_i \sigma_i^2 = \frac{N\sigma^2}{N^2} = \frac{\sigma^2}{N}, \quad (\sigma_i = \sigma)$$

Otrzymaliśmy wynik, który należy zapamiętać

- ze względu na interpretację wyniku MC
- stanowią część dowodu dla CTG

$$\langle \bar{x} \rangle = \mu, \quad \sigma_{\bar{x}}^2 = \frac{\sigma^2}{N}, \quad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{N}}$$

We wzorze określającym prawdopodobieństwo

$$P\{|x - \mu| \geq \lambda\sigma\} \leq \frac{\langle (x - \mu)^2 \rangle}{\lambda^2 \sigma^2} = \frac{1}{\lambda^2}$$

podstawiamy kolejno

$$(1) x \rightarrow \bar{x}, \quad (2) \sigma^2 \rightarrow \sigma_{\bar{x}}^2, \quad (3) \lambda\sigma_{\bar{x}} = \epsilon \quad (4) \frac{1}{\lambda^2} = \frac{\sigma_{\bar{x}}^2}{\epsilon^2} = \frac{\sigma^2}{N\epsilon^2}$$

$$P\{|\bar{x} - \mu| \geq \epsilon\} \leq \frac{\sigma^2}{N\epsilon^2}$$

- w granicy $N \rightarrow \infty$ prawa strona dąży do zera
- tempo zbieżności jest odwrotnie proporcjonalne do N co jest charakterystyczne dla metody Monte Carlo
- istotna jest też wartość wariancji rozkładu zmiennej X

warunek ten możemy też zapisać w postaci alternatywnej

$$P\{|\bar{x} - \mu| < \epsilon\} = 1 - \frac{\sigma^2}{N\epsilon^2}$$

- σ^2 to wariancja zmiennej losowej X

- w tej postaci prawo wielkich liczb wyraża zbieżność wartości średniej do wartości oczekiwanej dla dowolnego $\epsilon > 0$
- wynik istotny dla MC bo zawiera jawną relację pomiędzy wariancją a liczbą losowań N
- wielkość ϵ wyraża nasze oczekiwania co do spodziewanej ($P < 1$ nie daje pewności) maksymalnej różnicy wartości średniej i oczekiwanej
- w praktyce ϵ i σ są ze sobą powiązane, bo w MC to wariancja jest istotna (możemy ją na bieżąco kontrolować)

$$\epsilon = \gamma\sigma \quad \implies \quad P\{|\bar{x} - \mu| < \gamma\sigma\} = 1 - \frac{1}{N\gamma^2}$$

- twierdzenie możemy wykorzystać do szacowania wartości N dla zadanego przedziału ufności

Przykład. Wyznaczyć minimalne N spełniające warunki

$$(1) \quad P(|\bar{x} - \mu| < \epsilon) = p = 0.99$$

$$(2) \quad \epsilon_1 = 0.1\sigma, \quad \epsilon_2 = 0.001\sigma$$

$$P(|\bar{x} - \mu| \geq \epsilon) \leq \frac{\sigma^2}{N\epsilon^2} \quad \implies \quad P(|\bar{x} - \mu| < \epsilon) = p = 1 - \frac{\sigma^2}{N\epsilon^2}$$

$$N_1 = \frac{\sigma^2}{(1-p)\epsilon_1^2} = \frac{\sigma^2}{0.01 \cdot 0.01\sigma^2} = 10^4$$

$$N_2 = \frac{\sigma^2}{(1-p)\epsilon_2^2} = \frac{\sigma^2}{0.01 \cdot 10^{-6}\sigma^2} = 10^8$$

- żądając dużej dokładności wyniku musimy przeprowadzić dużą liczbę losowań/powtórzeń w MC a to jest czasochłonne
- czy istnieje rozwiązanie tego problemu?
w pewnych przypadkach tak – należy zmodyfikować algorytm MC by zmniejszyć wariancję 100 razy, wówczas pierwszy warunek będzie odpowiadał drugiemu, ale dla znacznie mniejszego N (wykład z całkowania MC – metody redukcji wariancji)

Centralne twierdzenie graniczne

Dla zbioru N **niezależnych** zmiennych losowych o rozkładach opisywanych fgp $f_i(x)$, wartościach własnych μ_i oraz wariancjach σ_i , ich wartość średnia jest zmienną losową

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i$$

o poniższych własnościach

(1) wartość oczekiwana zmiennej $\langle \bar{x} \rangle = \frac{1}{N} \sum_{i=1}^N \mu_i$

(2) wariancja zmiennej $\sigma_{\bar{x}}^2 = \frac{1}{N^2} \sum_{i=1}^N \sigma_i^2$

(1) i (2) udowodniliśmy przy omawianiu prawa wielkich liczb

(3) dla $N \rightarrow \infty$ rozkład **zmiennej losowej - wartości średniej** dąży do rozkładu normalnego o wartości oczekiwanej i wariancji tej zmiennej

$$\lim_{N \rightarrow \infty} f_{\bar{x}}(\bar{x}) = \frac{1}{\sigma_{\bar{x}} \sqrt{2\pi}} \exp \left[-\frac{(\bar{x} - \langle \bar{x} \rangle)^2}{2\sigma_{\bar{x}}^2} \right]$$

- przeprowadzając eksperyment MC wielokrotnie dostaniemy wartości średnie, które są zmiennymi losowymi
- wartości średnie mają rozkład normalny
- wiemy jak minimalizować wariancję wartości średniej – wystarczy zwiększać liczbę losowań
- wynik MC ma zatem jednoznaczną interpretację – określony rozkład wyniku i sposób minimalizacji wariancji