Lista 6

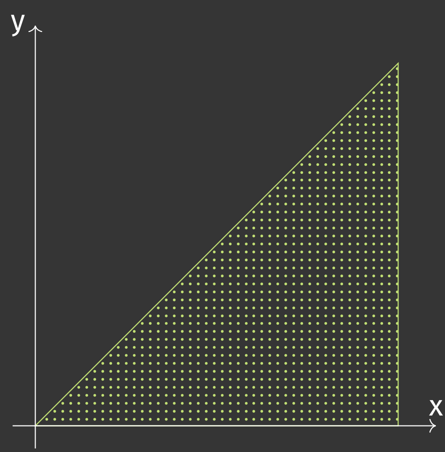
Rachunek Prawdopodobieństwa 1R

Weronika Jakimowicz

Zadanie 1. Niech (X, Y) będzie 2-wymiarową zmienną losową o rozkładzie zadanym gęstością $f(x, y) = C(x + y)$ dla $0 \le y \le x \le 1$ i $f(x, y) = 0$ poza tym zbiorem. Znajdź wartość C. Znajdź rozkłady brzegowe. Czy zmienne losowe X i Y są niezależne?

Mamy dane

$$f(x,y) = \begin{cases} C(x+y) & 0 \le y \le x \le 1 \\ 0 & \text{wpp} \end{cases}$$



i w pierwszej kolejności pytamy o wartość zmiennej C. Wiemy, że

$$\int_{\mathbb{R}^2} f(x,y) = 1,$$

a ponieważ my żyjemy w świecie trójkąta pod $y = x$, to mamy:

$$1 = \int_0^1 \int_0^x C(x+y)dydx = C\int_0^1 (x^2 + \frac{x^2}{2})dx = C(\frac{1}{3} + \frac{1}{6})$$

czyli z moich bardzo precyzyjnych kalkulacji wynika, że $C = 2$.

Teraz pora na rozkłady brzegowe.

$$\mathbb{P}[X \in A] = \mathbb{P}\left[(X,Y) \in A \times \mathbb{R}\right] = \int_{A\times\mathbb{R}} f(x,y)dydx = \int_A \int_{\mathbb{R}} 2(x+y)dydx =$$
$$= \int_A \int_0^x 2(x+y)dydx = \int_A 3x^2 dx$$

$$\mathbb{P}[Y \in B] = \mathbb{P}\left[(X,Y) \in \mathbb{R} \times B\right] = \int_{\mathbb{R}\times B} f(x,y)dydx = \int_{\mathbb{R}} \int_B f(x,y)dydx = \int_B \int_{\mathbb{R}} f(x,y)dxdy =$$
$$= \int_B \int_y^1 2(x+y)dxdy = \int_B [1 + 2y - 3y^2]dy$$

Na pytanie, czy są to zmienne niezależne odpowiadamy patrząc na gęstości tych dwóch zmiennych losowych. Żeby były niezależne, musiałoby zachodzić

$$f(x,y) = f_X(x)f_Y(y)$$

Tutaj mamy

$$f_X(x)f_Y(y) = 3x^2(1 + 2y - 3y^2) \neq 2(x+y)$$

więc są bardzo zależne.

Zadanie 2. Zmienna losowa (X, Y) ma rozkład z gęstością $g(x, y) = C \cdot xy \cdot \mathbb{1}_{[0,1]^2}$.

(a) Wyznaczyć C.

(b) Obliczyć $\mathbb{P}[X + Y \le 1]$

(c) Wyznaczyć rozkład zmiennej losowej $\frac{X}{Y}$.

(d) Czy zmienne X i Y są niezależne?

(e) Czy $\frac{X}{Y}$ i Y są niezależne?

(a) To będzie całeczką $\star$

$$1 = \int\int_{\mathbb{R}^2} g(x, y)\, dxdy = C \int_0^1 \int_0^1 xydxdy = C \int_0^1 y\frac{1}{2}\, dy = \frac{C}{4} \implies C = 4$$

(b) Ustalmy sobie najpierw s i niech X = s. Wtedy $X + Y \leq 1 \iff Y \leq 1 - s$. Takie prawdopodobieństwo liczymy w następujący sposób:

$$\mathbb{P}[X = s, Y \leq 1 - s] = \int_0^{1-s} g(s, t)\, dt.$$

Super, ale s może być dowolnym punktem z przedziału [0, 1], więc my chcemy zliczyć wartości dla każdego takiego s. W tym pomaga nam kolejna całka:

$$\mathbb{P}[X + Y \leq 1] = \int_0^1 \mathbb{P}[X = s, Y \leq 1 - s] = \int_0^1 \int_0^{1-s} 4 \cdot st\, dtds$$

(c) Policzymy dystrybuantę. Do tego potrzebuję się zastanowić, kiedy $\frac{X}{Y} \leq t$? Ustalmy sobie Y = s, wtedy $\frac{X}{Y} = \frac{X}{s} \leq t \implies X \leq ts$

$$\mathbb{P}\left[\frac{X}{Y} \leq t\right] = \int_0^1 \mathbb{P}[X \leq ts, Y = s]\, ds = \int_0^1 \int_0^{ts} \mathbb{P}[X = p, Y = s]\, dpds =$$
$$= \int_0^1 \int_0^{ts} 4ps\, dpds = 2\int_0^1 t^2 s^2\, ds = \frac{2}{3}t^2$$

(d) Do sprawdzania, czy zmienne X i Y są niezależne potrzebujemy znać rozkłady brzegowe, czyli rzuty na X i na Y. Będę liczyć dystrybuantę tych rzutów:

$$\mathbb{P}[X \leq t] = \mathbb{P}\left[X \leq t, Y \in [0, 1]\right] =$$
$$= \int_0^t \int_0^1 g(p, s)\, dsdp =$$
$$= 2\int_0^t p\, dp = t^2$$

$$\mathbb{P}[Y \leq t] = \int_0^t \int_0^1 g(p, s)\, dsdp = t^2$$

Teraz sprawdzamy, czy $\mathbb{P}[X \leq x, Y \leq y] = \mathbb{P}[X \leq x]\,\mathbb{P}[Y \leq y]$

$$\mathbb{P}[X \leq x, Y \leq y] = \int_0^x \int_0^y g(s, t)\, dtds = \int_0^x 4s \int_0^y t\, dtds =$$
$$= \int_0^x 2sy^2\, ds = x^2y^2$$

Czyli zmienne są niezależne.

(e) Zmienne są niezależne, jeśli $\mathbb{P}[A]\,\mathbb{P}[B] = \mathbb{P}[A \cap B]$. U nas niech

$$A = \{\omega \;:\; \frac{X(\omega)}{Y(\omega)} \leq t\} \quad \left[\mathbb{P}[A] = \mathbb{P}\left[\frac{X}{Y} \leq t\right]\right]$$

$$B = \{\omega \;:\; Y(\omega) \leq s\} \quad [\mathbb{P}[B] = \mathbb{P}[Y \leq s]]$$

$$A \cap B = \{\omega \ : \ \frac{X(\omega)}{Y(\omega)} \leq t \text{ i } Y(\omega) \leq s\} = \{\omega \ : \ X(\omega) \leq tY(\omega) \text{ i } Y(\omega) \leq s\}$$

Pierwsze dwa prawdopodobieństwa już mamy, zostaje nam obliczyć prawdopodobieństwo przekroju.

$$\begin{aligned}\mathbb{P}[A \cap B] = \mathbb{P}[X \leq tY, Y \leq s] &= \int_0^s \mathbb{P}[X \leq ty, Y = y] \ dy = \\ &= \int_0^s \int_0^{ty} \mathbb{P}[X = x, Y = y] \ dxdy = \int_0^s \int_0^{ts} 4xy \ dxdy = \\ &= 2\int_0^s yt^2y^2 \ dy = \frac{1}{2}t^2s^4\end{aligned}$$

No i wyszło, że są zależne [co jest dość rozsądnym wynikiem].

Zadanie 3. Niech X i Y będą niezależnymi zmiennymi losowymi, których rozkład jest zadany gęstością $2x \cdot \mathbb{1}_{[0,1]}$. Znaleźć prawdopodobieństwo, że

(a) $X + Y < \frac{1}{2}$

(b) $XY < \frac{1}{2}$

(c) $|X - Y| < \frac{1}{2}$

(d) $X^2 + y^2 \leq \frac{1}{2}$

Zmienne są niezależne, więc gęstość $f_{(X,Y)}(x, y) = f_X(x) \cdot f_Y(y)$.

(a) Robimy tym samym trikiem co wcześniej, tzn. $X + Y < \frac{1}{2}$, ustalamy x takie, że $X = t$ i mamy $Y < \frac{1}{2} - t$. Zauważamy jeszcze, że $t \in [0, \frac{1}{2}]$, Czyli dostajemy

$$\begin{aligned}\mathbb{P}\left[X + Y < \frac{1}{2}\right] &= \int_0^{\frac{1}{2}} \mathbb{P}\left[X = x, Y < \frac{1}{2} - x\right] \ dx = \int_0^{\frac{1}{2}} \int_0^{\frac{1}{2}-x} \mathbb{P}[X = x, Y = y] \ dydx = \\ &= \int_0^{\frac{1}{2}} \int_0^{\frac{1}{2}-x} 4xy \ dydx\end{aligned}$$

(b)

$$\mathbb{P}\left[XY < \frac{1}{2}\right] = \int_0^1 \mathbb{P}\left[X = x, Y < \frac{1}{2x}\right] \ dx = \int_0^1 \int_0^{\frac{1}{2x}} 4xy \ dydx$$

(c)

$$\mathbb{P}\left[|X - Y| < \frac{1}{2}\right] = \mathbb{P}\left[-\frac{1}{2} < X - Y < \frac{1}{2}\right] = \int_0^1 \mathbb{P}\left[X = x, x - \frac{1}{2} < Y < x + \frac{1}{2}\right] \ dx$$

a to już można sobie scałkować jak ma się ochotę.

(d) Tutaj zauważmy, że $X \leq \frac{1}{\sqrt{2}} = a$, bo inaczej wyjdziemy poza zakres.

$$\begin{aligned}\mathbb{P}\left[X^2 + Y^2 \leq \frac{1}{2}\right] &= \int_0^a \mathbb{P}\left[X = x, Y^2 \leq \frac{1}{2} - x^2\right] \ dx = \int_0^a \mathbb{P}\left[X = x, Y \leq \sqrt{\frac{1}{2} - x^2}\right] \ dx = \\ &= \int_0^a \int_0^{\sqrt{\frac{1}{2}-x^2}} 4xy \ dydx\end{aligned}$$

Zadanie 4. Niech $X_1, \ldots, X_n$ będą niezależnymi zmiennymi losowymi o rozkładzie wykładniczym z parametrem 1. Znajdź rozkład $Y = \min_{1 \leq i \leq n} X_i$. Czy X_n i Y są niezależne?

Mają rozkład wykładniczy, więc funkcja gęstości X_i to

$$f_i(x) = e^{-x}\mathbb{1}_{[0,\infty)}$$

Aby zmienne były niezależne, musimy mieć

$$\mathbb{P}[X_n = x]\,\mathbb{P}[Y = y] = \mathbb{P}[X_n = x, Y = y]$$

Policzmy najpierw funkcję gęstości Y. Niech k takie, że $X_k = Y$, wtedy:

$$\mathbb{P}[Y = y] = \mathbb{P}\left[X_k = y\right]\prod_{i\neq k}\mathbb{P}\left[X_i \geq y\right] = e^{-y}\prod_{i\neq k}\int_y^\infty e^{-x}\,dx = e^{-y}\prod_{i\neq k} e^{y} = e^{(n-2)y}$$

k jest tutaj nieważne tak naprawdę.

Jeśli x = y, wtedy

$$\mathbb{P}[X_n = x, Y = x] = \mathbb{P}[X_n = x]\prod_{i<n}\mathbb{P}\left[X_i \geq x\right] = e^{-x}e^{(n-1)x} = e^{(n-2)x}$$

ale

$$\mathbb{P}[X_n = x]\,\mathbb{P}[Y = x] = e^{-x}e^{(n-2)x} = e^{(n-3)x}$$

Zadanie 5. Zmienne losowe X i Y są niezależne. Pokaż, że jeżeli X nie ma atomów, to $P(X = Y) = 0$.

Ponieważ X nie ma atomów, to zbiór $\{x \;:\; \mathbb{P}[X = x] > 0\} = \emptyset$.

Szukamy tak naprawdę $\mathbb{P}[X = t, Y = t]$ po wszystkich t. Ponieważ zdarzenia są niezależne, to:

$$\mathbb{P}[X = Y] = \int_{-\infty}^{\infty}\mathbb{P}[X = t, Y = t]\,dt = \int \mathbb{P}[X = t]\,\mathbb{P}[Y = t]\,dt = \int 0dt = 0,$$

gdyż X jest bezatomowa.

Zadanie 6. Zmienne X i Y są niezależne. X ma rzokład jednostajny na przedziale $[0, 1]$, a Y ma rozkład zadany przez $\mathbb{P}[Y = -1] = \frac{1}{3}$, $\mathbb{P}[Y = 2] = \frac{2}{3}$.

(a) Oblicz $\mathbb{P}[3X < Y]$

(b) Wyzacz rozkład zmiennej XY

(a) W dyskretnym rozkładzie (jaki ma Y) dodajemy wartości, więc zrobimy to najpierw, a potem policzymy po X.

$$\mathbb{P}[3X < Y] = \frac{1}{3}\mathbb{P}[3X < -1] + \frac{2}{3}\mathbb{P}[3X < 2] = \int_0^{\frac{2}{3}}\frac{2}{3}dx = \frac{4}{9}$$

(b)

$$\mathbb{P}[XY < t] = \frac{1}{3}\mathbb{P}\left[X\cdot(-1) < t\right] + \frac{2}{3}\mathbb{P}[X\cdot 2 < t] = \frac{t}{3}$$

i obcinam nadmiar ponad 1.

Zadanie 7. Pokaż, że zmienne losowe $X_1, ..., X_n$ o gęstościach $f_1, ..., f_n$ są niezależne $\iff$ zmienna $X = (X_1, ..., X_n)$ ma gęstość

$$f(x_1, ..., x_n) = f_1(x_1)...f_n(x_n)$$

$$\text{Niezależne} \iff f(x_1, ..., x_n) = f_1(x_1)...f_n(x_n)$$

$\Longrightarrow$

Nirch $T_i \subseteq \mathbb{R}$, wtedy z niezależności zmiennych mamy:

$$\mathbb{P}[X_1 \in T_1, ..., X_n \in T_n] = \int_{T_1} ... \int_{T_n} f(x_1, ..., x_n)dx_1...dx_n =$$
$$= \int_{T_1} f_1(x_1)dx_1... \int_{T_n} f_n(x_n)dx_n = \mathbb{P}[X_1 \in T_1]...\mathbb{P}[X_n \in T_n]$$

rozpisując krok po kroku:

$$\mathbb{P}[X_1 \in T_1]...\mathbb{P}[X_n \in T_n] = \int_{T_1} f_1(x_1)dx_1... \int_{T_n} f_n(x_n)dx_n = \int_{T_1} f_1(x_1) \int_{T_2} f_2(x_2)dx_2dx_1... \int_{T_n} f_n(x_n)dx =$$
$$= \int_{T_1}\int_{T_2} f_1(x_1)f_2(x_2)dx_2dx_1... \int_{T_n} f_n(x_n)dx_n = ... = \int_{T_1} ... \int_{T_n} f_1(x_1)...f_n(x_n)dx_1...dx_n =$$
$$= \int_{T_1} ... \int_{T_n} f(x_1, ...x_n)dx_1...dx_n = \mathbb{P}\left[X_1 \int T_1, ..., X_n \in T_n\right]$$

Ponieważ dzieje się tak dla dowolnych T_i, to funkcje pod całką muszą się równać (prawie wszędzie?). Czyli

$$f(x_1, ..., x_n) = f_1(x_1)...f_n(x_n).$$

$\Longleftarrow$

Wychodzimy z tego, że

$$f(x_1, ..., x_n) = f_1(x_1)...f_n(x_n).$$

Wybierając dowolne $T_i \subseteq \mathbb{R}$ i całkując obie strony dostajemy:

$$\int_{T_1} ... \int_{T_n} f(x_1, ..., x_n)dx_1...dx_n = \int_{T_1} ... \int_{T_n} f_1(x_1)...f_n(x_n)dx_1...dx_n =$$
$$= \int_{T_1} ... \int_{T_{n-1}} f_1(x_1)...f_{n-1}(x_{n-1})dx_1...dx_n \int_{T_n} f_n(x_n)dx_n = ...$$
$$= \int_{T_1} f_1(x_1)dx_1... \int_{T_n} f_n(x_n)dx_n$$

Prawa strona równania to iloczyn $\mathbb{P}[X_1 \in T_1]...\mathbb{P}[X_n \in T_n]$, a lewa to $\mathbb{P}[X_1 \in T_1, ..., X_n \in T_n]$ i znowu dzieje się tak bez względu na wybór T_i, czyli mamy równość i zmienne są niezależne.

Zadanie 8. Z odcinka $[0, 1]$ losujemy niezależnie w sposób jednostajny liczby $X_1, X_2, ...$ Uzasadnij, że z prawdopodobieństwem 1 ciąg $\{X_n\}$ jest gęsty w odcinku $[0, 1]$.

Weźmy sobie dowolną kulę na odcinku $[0, 1]$ o promieniu r i środku x: $B_r(x)$. Prawdopodobieństwo, że ani jedna ze zmiennych w nią trafi wynosi $1 - 2r$ (tutaj $r \le \frac{1}{2}$). Losujemy niezależnie, więc zmienne są niezależne. Jeśli rozważymy pierwsze n zmiennych, to prawdopodobieństwo, że ani jedna z nich wpadnie w $B_r(x)$ wynosi:

$$\mathbb{P}\left[X_1 \in B_r(x)^c, X_2 \in B_r(x)^c, ...X_n \in B_r(x)^c\right] = \mathbb{P}\left[X_1 \in B_r(x)^c\right]...\mathbb{P}\left[X_n \in B_r(x)^c\right] = (1 - 2r)^n$$

Chcemy użyć lematu Borela-Cantelliego, więc sprawdzamy sumę:

$$\sum_{n\in\mathbb{N}} (1 - 2r)^n = \frac{1}{1 - (1 - 2r)} = \frac{1}{2r} < \infty$$

czyli z prawdopodobieństwem 1 skończenie wiele zmiennych nie trafi do $B_r(x)$, czyli nieskończenie wiele z nich do $B_r(x)$ trafi. Tak się dzieje dla każdej kuli, więc z prawdopodobieństwem 1 przetniemy dowolną kulę - tworzy się gęsty podzbiór $[0, 1]$.

Zadanie 9. Zmienne losowe X i Y są niezależne i mają rozkłady wykładnicze z parametrami λ i μ odpowiednio. Znajdź rozkład zmiennej losowej X + Y.

Co wiemy? Że gęstość X to $f_X(t) = \lambda e^{-\lambda t}$ gdy $t \geq 0$, a gęstość Y to $f_Y(t) = \mu e^{-\mu t}$ gdy $t \geq 0$. Dalej, wiem że

$$f(t_x, t_y) = f_X(t_x) f_Y(t_y)$$

a poszukuję $\mathbb{P}[X + Y = t]$

Skrypt mówi, że sploty is the way (ale miałam nawet ten sam pomysł!). Nie mogę puścić całki aż do ∞, bo wtedy mi się zeruje e^{t-s} dla $s > t$. Czyli:

$$\mathbb{P}[X + Y = t] = \int_0^t \mathbb{P}[X = a, Y = t - a]\, da,$$

a w mowie skryptowej:

$$\mathbb{P}[X + Y = t] = f_x \star f_y(t) = \int_0^t f_x(t - s) f_y(s) ds = \lambda\mu \int_0^t e^{-\lambda(t-s)} e^{-\mu s} ds =$$
$$= \lambda\mu e^{-\lambda t} \int_0^t e^{\lambda s - \mu s} ds = \lambda\mu e^{-\lambda t} \int_0^t e^{s(\lambda-\mu)} ds = \frac{\lambda\mu}{\lambda - \mu}[e^{-\lambda t} - e^{-\mu t}]$$

Zadanie 10. Zmienne losowe X i Y są niezależne i mają rozkład wykładniczy z parametrem 1. Udowodnić, że zmienne $\frac{X}{Y}$ oraz $X + Y$ są niezależne.

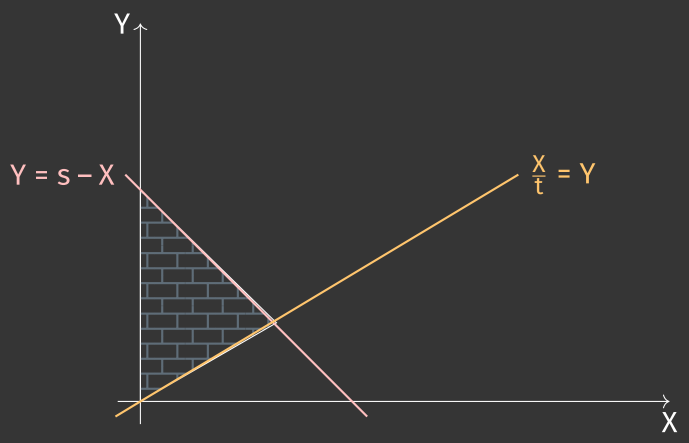
To lecimy od znalezienia rozkładu $\frac{X}{Y}$, potem $X + Y$ i na końcu rozkład $(\frac{X}{Y}, X + Y)$.

$$\mathbb{P}\left[\frac{X}{Y} \leq t\right] = \mathbb{P}[X \leq tY] = \int_0^\infty \mathbb{P}[X \leq ty, Y = y]\ dy = \int_0^\infty \int_0^{ty} \mathbb{P}[X = x, Y = y]\ dxdy =$$
$$= \int_0^\infty \int_0^{ty} e^{-x} e^{-y}\ dxdy = \int_0^\infty e^{-y}[1 - e^{-ty}]\ dy = 1 - \frac{1}{1+t} = \frac{t}{1+t}$$

$$\mathbb{P}[X + Y \leq t] = \int_0^t \mathbb{P}[X + Y = s]\ ds = \int_0^t \int_0^s \mathbb{P}[X = x, Y = s - x]\ dxds = \int_0^t \int_0^s e^{-x} e^{x-s}\ dxds =$$
$$= \int_0^t \int_0^s e^{-s}\ dxds = \int_0^t s e^{-s} ds = 1 - e^{-t}(t + 1)$$

$$\mathbb{P}\left[\frac{X}{Y} \leq t\right] \mathbb{P}[X + Y \leq s] = \frac{t}{1+t}[1 - e^{-s}(1 + s)]$$

$$\frac{X}{t} \leq Y$$
$$Y \leq s - X$$



$$\mathbb{P}\left[\frac{X}{Y} \le t, X+Y \le s\right] = \int_0^{\frac{st}{s+1}} e^{-x}\mathbb{P}\left[\frac{x}{t} \le Y, Y \le s-x\right]\, dx = \int_0^{\frac{st}{s+1}} e^{-x}\int_{\frac{x}{t}}^{s-x} e^{-y}\, dydx = \frac{t}{t+1}[1-e^{-s}(1+s)]$$

Zadanie 11. Zmienne losowe $X_1, ...X_n$ są niezależne i mają rozkłady Poissona z parametrami λ_i. Pokaż, że $X_1 + ... + X_n$ ma rozkład Poissona z parametrem $\lambda_1 + ... + \lambda_n$.

To jest rozkład dyskretny. Gęstość X_i to

$$f_i(k) = \lambda_i^k \frac{e^{-\lambda}}{k!}$$

Dowód przez indukcję:

Dla $n = 1$ nie ma co robić, jeśli $n = 2$, to

$$\mathbb{P}[X_1 + X_2 = t] = \sum_{x_1=0}^{t} \mathbb{P}[X_1 = x_1, X_2 = t - x_1] = \sum_{x_1=0}^{t} \lambda_1^{x_1}\frac{e^{-\lambda_1}}{x_1!}\lambda_2^{t-x_1}\frac{e^{-\lambda_2}}{(t-x_1)!} =$$

$$= \frac{e^{-(\lambda_1+\lambda_2)}}{t!}\sum_{x_1=0}^{t}\binom{t}{x_1}\lambda_1^{x_1}\lambda_2^{t-x_1} = (\lambda_1+\lambda_2)^t\frac{e^{-(\lambda_1+\lambda_2)}}{t!}$$

czyli to co chcemy.

Niech $\lambda = \sum \lambda_i$.

$$\mathbb{P}[X_1 + ... + X_n = t] = \sum_{x_1+...+x_n=t} \lambda_1^{x_1}...\lambda_n^{x_n}\frac{e^{-\lambda}}{x_1!...x_n!} = \frac{e^{-\lambda}}{t!}\sum_{x_1+...+x_n}\binom{t}{x_1,...,x_n}\lambda_1^{x_1}...\lambda_n^{x_n} = \frac{e^{-\lambda}}{t!}\lambda^t$$

Zadanie 12. Załóżmy, że X_1 i X_2 są niezależnymi zmiennymi losowymi o rozkładach odpowiednio $N(m_1\sigma_1)$ i $N(m_2, \sigma_2)$. Oblicz rozkład zmiennej losowej $X_1 + X_2$.

Funkcja gęstości w rozkładzie normalnym:

$$f_i(x) = \frac{1}{\sigma_i\sqrt{2\pi}}e^{-\frac{(x-m_i)^2}{2\sigma_i^2}}$$

$$\mathbb{P}[X_1 + X_2 = t] = \int_{\mathbb{R}} \mathbb{P}[X_1 = x, X_2 = t-x]\, dx = \int_{\mathbb{R}} \mathbb{P}[X_1 = x]\mathbb{P}[X_2 = t-x]\, dx =$$

$$= \int_{\mathbb{R}} \frac{1}{\sigma_1\sqrt{2\pi}}e^{-\frac{(x-m_1)^2}{2\sigma_1^2}}\frac{1}{\sigma_2\sqrt{2\pi}}e^{-\frac{(t-x_1-m_1)^2}{2\sigma_2^2}}\, dx =$$

$$= \frac{1}{2\pi\sigma_1\sigma_2}\int_{\mathbb{R}} \exp\left[-\frac{\sigma_2^2(x-m_1)^2 + \sigma_1^2(t-x-m_2)^2}{2\sigma_1^2\sigma_2^2}\right]\, dx$$

O JEZU NIE?

Zadanie 13. Niech $\{X_n\}_{n\in\mathbb{N}}$ będzie ciągiem niezależnych zmiennych losowych o rozkładzie $N(0,1)$. Uzasadnij, że z prawdopodobieństwem 1 istnieje nieskończenie wiele indeksów n takich, że

$$|X_{2n} - X_{2n+1}| \le \frac{1}{n}.$$

Szacujemy prawdopodobieństwo z zadania:

$$\mathbb{P}\left[|X_{2n} - X_{2n+1}| \leq \frac{1}{n}\right] \geq \mathbb{P}\left[X_1 = 0, -\frac{1}{n} \leq X_{2n} \leq \frac{1}{n}\right] = \int_{-\frac{1}{n}}^{\frac{1}{n}} \frac{1}{2\pi} e^{-\frac{y^2}{2}} \, dydx$$

Wykres funkcji $e^{-\frac{x^2}{2}}$ jest dla ujemnych x. W dodatku, funkcja ta jest parzysta.

Czyli możemy oszacować:

$$\int_{-\frac{1}{n}}^{\frac{1}{n}} \frac{1}{2\pi} e^{-\frac{y^2}{2}} \, dy \geq \frac{1}{2\pi} \cdot \frac{2}{n} \cdot e^{-\frac{1}{2n^2}} = \frac{1}{n\pi} e^{-\frac{1}{2n^2}}$$

Zmienne zawierające informacje o tym, kiedy $X_{2n} = 0 \ \wedge \ |X_{2n+1}| \leq \frac{1}{n}$ są niezależne (to widać z rozpisania wyżej), więc chcę użyć lematu Borela-Cantellego.

$$\sum \mathbb{P}\left[|X_{2n} - X_{2n+1}| \leq \frac{1}{n}\right] \geq \sum \mathbb{P}\left[X_{2n} = 0, |X_{2n+1}| \leq \frac{1}{n}\right] \geq \sum \frac{1}{n\pi} e^{-\frac{1}{2n^2}} \geq$$
$$\geq \sum \frac{1}{n\pi} \geq \sum \frac{1}{n} = \infty$$

No to super! Z prawdopodobieństwem 1 nieskończenie wiele razy zdarzy się, że $X_{2n} = 0$ i $|X_{2n+1}| \leq \frac{1}{n}$, czyli tym bardziej $|X_{2n} - X_{2n+1}| \leq \frac{1}{n}$ zdarzy się nieskończenie wiele razy z prawdopodobieństwem 1.

Zadanie 14. Niech $\{X_i\}_{i=1,\dots,5}$ będzie ciągiem niezależnych zmiennych losowych.

(a) Czy zmienne losowe $X_1 + X_2$ oraz $X_3 + X_4X_5$ są niezależne?

(b) Czy zmienne losowe X_1, X_1X_2 są niezależne?

(a) TAK, SĄ NIEZALEŻNE

$$\begin{aligned}
\mathbb{P}[X_1 + X_2 = t, X_3 + X_4X_5 = s] &= \int_{\mathbb{R}} \mathbb{P}[X_1 = a, X_2 = t - a, X_3 + X_4X_5] \, da = \\
&= \int_{\mathbb{R}} \int_{\mathbb{R}} \mathbb{P}[X_1 = a, X_2 = t - a, X_3 = b, X_4X_5 = s - b] \, dadb = \\
&= \int \int \int \mathbb{P}\left[X_1 = a, X_2 = t - a, X_3 = b, X_4 = c, X_5 = \frac{s-b}{c}\right] \, dadbdc = \\
&= \int \int \int \mathbb{P}[X_1 = a] \mathbb{P}[X_2 = t - a] \mathbb{P}[X_3 = b] \mathbb{P}[X_4 = c] \mathbb{P}\left[X_5 = \frac{s-b}{c}\right] \, dadbdc = \\
&= \int \int \mathbb{P}\left[X_3 = b, X_4 = c, X_5 = \frac{s-b}{c}\right] \, dbdc \int \mathbb{P}[X_1 = a, X_2 = t - a] \, da = \\
&= \int \mathbb{P}[X_3 = b, X_4X_5 = s - b] \, db \mathbb{P}[X_1 + X_2 = t] = \\
&= \mathbb{P}[X_3 + X_4X_5 = s] \mathbb{P}[X_1 + X_2 = t]
\end{aligned}$$

(b) Odp.: WELES JEST CZARNY

Jeśli popatrzymy na zmienne losowe takie, że $\mathbb{P}\left[X_i = 1\right] = \frac{1}{2}, \mathbb{P}\left[X_i = -1\right] = \frac{1}{2}$, to zmienne X_1 i X_1X_2 są faktycznie niezależne:

$$\mathbb{P}[X_1 = 1, X_1X_2 = 1] = \mathbb{P}[X_1 = 1, X_2 = 1] = \frac{1}{4} = \frac{1}{2}\left(\frac{1}{4} + \frac{1}{4}\right) = \mathbb{P}[X_1 = 1]\,\mathbb{P}[X_1X_2 = 1]$$

$$\mathbb{P}[X_1 = -1, X_1X_2 = 1] = \mathbb{P}[X_1 = -1, X_2 = -1] = \frac{1}{4} = \frac{1}{2}\left(\frac{1}{4} + \frac{1}{4}\right) = \mathbb{P}[X_1 = -1]\,\mathbb{P}[X_1X_2 = 1]$$

i tak dalej.

Jeśli natomiast popatrzymy na zmienne o rozkładzie wykładniczym z parametrem 1, to mamy

$$\mathbb{P}[X_1X_2 = y] = \int_0^\infty \mathbb{P}\left[X_1 = t, X_2 = \frac{y}{t}\right] dt = \int_0^\infty e^{-t}e^{-b\frac{y}{t}}\,dt =$$
$$= \int_0^\infty e^{-t-\frac{y}{t}}\,dt$$

$$\mathbb{P}[X_1 = x, X_1X_2 = y] = \mathbb{P}\left[X_1 = x, X_2 = \frac{y}{x}\right] = \mathbb{P}[X_1 = x]\,\mathbb{P}\left[X_2 = \frac{y}{x}\right] =$$
$$= e^{-x}e^{-\frac{y}{x}}$$

Czyli, żeby zmienne były niezależne, musielibyśmy mieć

$$\mathbb{P}[X_1X_2 = y] = e^{-\frac{y}{x}},$$

a my mamy, że

$$\mathbb{P}[X_1X_2 = y] = \int_0^\infty e^{-\frac{x^2+y}{x}}\,dx \neq e^{-\frac{y}{x}}$$

Czyli tutaj nie są niezależne.