Lista 6

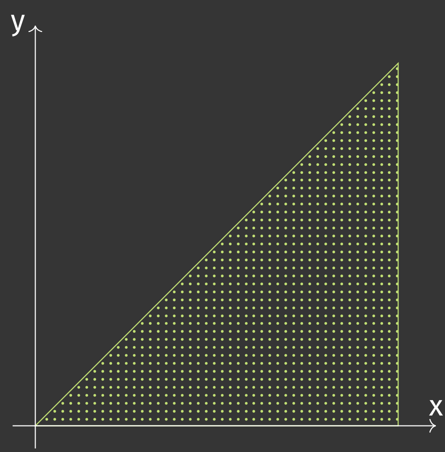
Rachunek Prawdopodobieństwa 1R

Weronika Jakimowicz

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

Mamy dane

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



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 $\mathbb{P}[X + Y \leq 1]$

(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 $\frac{X}{Y} = \frac{X}{s} \leq t \implies X \leq ts$

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

(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:

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

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

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

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

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 4. Niech $X_1, ..., X_n$ 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:

$$f_i(t) = \lambda_i e^{-\lambda_i t}$$

jeśli $t \geq 0$, wpp. mamy 0.

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:

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

gdyż X jest bezatomowa.

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

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

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

$\implies$

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

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

rozpisując krok po kroku:

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

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:

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

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 $r \leq \frac{1}{2}$). 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:

$$\begin{aligned}\mathbb{P}[X + Y = t] = f_X \star f_Y(t) &= \int_0^t f_X(t - s)f_Y(s)ds = \lambda\mu \int_0^t e^{-\lambda(t-s)}e^{-\mu s}ds = \\ &= \lambda\mu e^{-\lambda t} \int_0^t e^{\lambda s - \mu s}ds = \lambda\mu e^{-\lambda t} \int_0^t e^{s(\lambda - \mu)}ds = \frac{\lambda\mu}{\lambda - \mu}[e^{-\lambda t} - e^{-\mu t}]\end{aligned}$$