Rachunek prawdopodobieństwa i statystyka

Suma zmiennych losowych

1. Prosty dowód równości $\mathrm{E}(aX+b) = a\mathrm{E}(X)+b$.

   Zachodzą równości:
   $$\int_{\mathbb{R}} f(x)\,dx = 1, \quad \mathrm{E}(X) = \int_{\mathbb{R}} x\, f(x)\,dx.$$
   Jest zatem $\mathrm{E}(aX+b) = \int_{\mathbb{R}} (ax+b)\, f(x)\,dx = a\int_{\mathbb{R}} x\, f(x)\,dx + b\int_{\mathbb{R}} f(x)\,dx = a\mathrm{E}(X)+b.$
   Można powiedzieć ogólniej: jeżeli $Y = g(X)$, to $\mathrm{E}(Y) = \mathrm{E}(g(X)) = \int_{\mathbb{R}} g(x)\, f(x)\,dx.$

2. Definicja dystrybuanty zmiennej losowej.

   Definicja 1. *Niech X będzie zmienną losową. Dystrybuantą $F_X(t)$ nazywamy funkcję określoną jako $F_X(t) = P(X \leqslant t)$.*

   Jeżeli wiadomo o jakiej zmiennej losowej mówimy, to używamy oznaczenia $F(t)$. W wypadku dyskretnym $F_X(t) = \sum\limits_{x_i \leqslant t} p_i$, w wypadku zmiennej typu ciągłego $F_X(t) = \int_{-\infty}^{t} f(x)\,dx$.

3. Nawiązując do wypadku ciągłego i dystrybuanty $F(t)$ powołajmy się na tzw. "pierwsze główne twierdzenie rachunku całkowego".

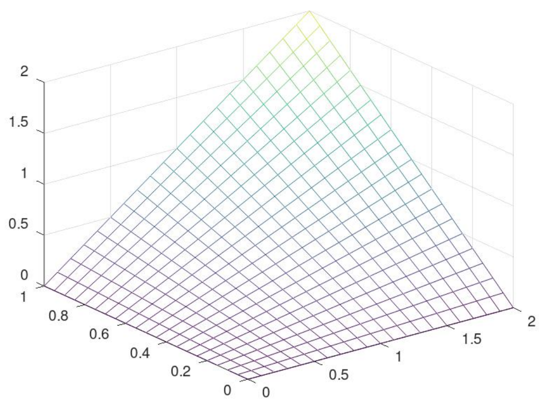
   Twierdzenie 1.
   Funkcja $f(x)$ jest całkowalna na zbiorze $\mathbb{R}$. Niech $F(t) = \int_{-\infty}^{t} f(x)\,dx$. Jest wówczas:

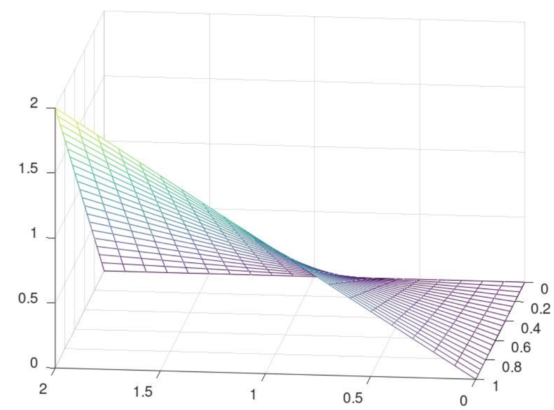
   (a) funkcja $F(t)$ jest ciągła w $\mathbb{R}$,
   (b) funkcja $F(t)$ jest różniczkowalna w każdym punkcie t ciągłości funkcji $f(t)$, i w tychże punktach jest $F'(t) = f(t)$.

   Z czysto praktycznego (obliczeniowego) punktu widzenia interesuje nas podpunkt (b).

4. Dodatkowy przykład na zamianę zmiennych.
   Weźmy pod uwagę funkcję $f(x,y) = xy$ na obszarze $[0,2]\times[0,1]$. Sprawdzamy na początek, że
   $$\int_0^2\int_0^1 xy\,dy\,dx = \int_0^2 \frac{xy^2}{2}\bigg|_{y=0}^{1} dx = \int_0^2 \frac{x}{2}dx = \frac{x^2}{4}\bigg|_0^2 = 1.$$
   Wprowadźmy nowe zmienne $S = X+Y,\ T = X-Y$. Przekształcenie odwrotne to $X = \dfrac{S+T}{2}$ oraz $Y = \dfrac{S-T}{2}$. Obliczenie $J \equiv J_{xy}$ daje
   $$J_{xy} = \begin{vmatrix} \frac{\partial x}{\partial s} & \frac{\partial x}{\partial t} \\ \frac{\partial y}{\partial s} & \frac{\partial y}{\partial t} \end{vmatrix} = \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{vmatrix} = -\frac{1}{2}. \tag{1}$$
   Zauważmy też, że
   $$J_{st} = \begin{vmatrix} \frac{\partial s}{\partial x} & \frac{\partial s}{\partial y} \\ \frac{\partial t}{\partial x} & \frac{\partial t}{\partial y} \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix} = -2 = \frac{1}{J_{xy}}.$$

Wykres gęstości:



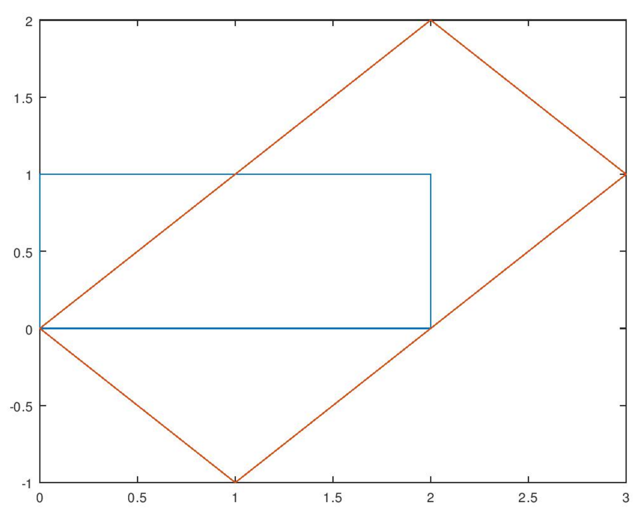


Mamy więc dwa elementy dla funkcji $f(x(s,t), y(s,t))$:

(a) $g(s,t) = f(x(s,t), y(s,t)) = \dfrac{s^2 - t^2}{4}$,

(b) $|J_{xy}| = \dfrac{1}{2}$

Do określenia pozostaje jeszcze znalezienie obszaru całkowania. Wprawdzie $s \in [0,3],\ t \in [-1,2]$ ale $\neg\left((s,t) \in [0,3] \times [-1,2]\right)$. Na rysunku poniżej zaznaczono obszary całkowania dla funkcji $f(x,y),\ g(s,t)$.
Rozpoczynamy od stwierdzenia, że $s \in [0,3]$. Dla ustalonego s należy znaleźć przedział zmienności dla t. Wówczas przejdziemy do obliczenia całki $\int_0^3 \int_{l(s)}^{h(s)} g(s,t)\, |J|\, dt\, ds$. Symbole $l(s), h(s)$ oznaczają granice całkowania po zmiennej t przy ustalonej wartości s.



$$\begin{cases} 0 < x < 2 \\ 0 < y < 1 \end{cases}, \quad \begin{cases} 0 < \frac{s+t}{2} < 2 \\ 0 < \frac{s-t}{2} < 1 \end{cases}, \quad \begin{cases} -s & < t < & 4-s \\ s-2 & < t < & s \end{cases}.$$

W podsumowaniu: dla ustalonego $s \in [0,3]$ jest $t \in [\max\{-s, s-2\}, \min\{4-s, s\}]$. Pod całką występuje wyrażenie $\dfrac{s^2 - t^2}{8}\, dt\, ds$, natomiast całki oznaczone są następujące

$$\int_0^1 \int_{-s}^{s} \ldots + \int_1^2 \int_{s-2}^{s} \ldots + \int_2^3 \int_{s-2}^{4-s} \ldots = 1.$$

5. Suma zmiennych

Gęstością dwuwymiarowej zmiennej losowej (X, Y) jest funkcja $f(x, y) = xy$ na obszarze $[0, 2] \times [0, 1]$. Znajdziemy rozkład (gęstość) zmiennej $S = X + Y$.

(a) Od zmiennej (X, Y) przechodzimy do zmiennej (S, T), gdzie $S = X + Y, \;\; T = Y$,

$$\begin{cases} S = X + Y \\ T = Y \end{cases}, \quad \begin{cases} X = S - T \\ 0 < Y = T \end{cases}, \quad J = \begin{vmatrix} 1 & -1 \\ 0 & 1 \end{vmatrix} = 1.$$

Stąd $g(s, t) = st - t^2$.

(b) Wyznaczamy gęstość brzegową $f_S(s)$ zmiennej (S, T). $f_S(s) = \int_{\mathbb{R}} g(s, t)\, dt = \frac{st^2}{2} - \frac{t^3}{3} + C.$

Pytanie: Jak dla ustalonego $s \in [0, 3]$ zmienia się t?

$$\begin{cases} 0 < X < 2 \\ 0 < Y < 1 \end{cases}, \quad \begin{cases} 0 < S - T < 2 \\ 0 < T < 1 \end{cases}, \quad \begin{cases} S - 2 < T < S \\ 0 < T < 1 \end{cases},$$

Wynika stąd, że $t \in (\max\{s - 2, 0\}, \min\{s, 1\})$, co daje następujące wyrażenia dla gęstości:

$$\begin{cases} s \in [0, 1] \quad t \in [0, s] & f_S(s) = \int_0^s g(s, t)\, dt = \frac{s^3}{6}, \\ s \in [1, 2] \quad t \in [0, 1] & f_S(s) = \int_0^s g(s, t)\, dt = \frac{s}{2} - \frac{1}{3}, \\ s \in [2, 3] \quad t \in [s - 2, 1] & f_S(s) = \int_0^s g(s, t)\, dt = \frac{s}{2} - \frac{1}{3} - \frac{s(s-2)^2}{2} + \frac{(s-2)^3}{3}. \end{cases},$$

$f_S(s)$ jest gęstością, ponieważ $\int_0^1 g_S(s)\, ds + \int_1^2 g_S(s)\, ds + \int_2^3 g_S(s)\, ds = \frac{1}{24} + \frac{5}{12} + \frac{13}{24} = 1.$

Witold Karczewski