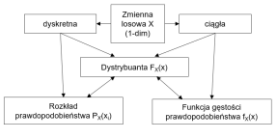
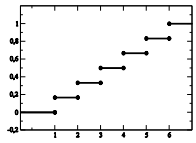
Poprzedni wykład

- Niezależność zdarzeń
- Zmienna losowa $\forall s \in S \xrightarrow{X(s)} x \in S_X$



- Dystrybuanta $F_X(x) = P(X \le x)$



1

Dystrybuanta - własności

- $0 \le F_X(x) \le 1$ (gdyż $F_X(x)$ jest sumą prawdopodobieństw)
- $\lim_{x \to \infty} F_X(x) = 1$ (gdyż zdarzenie x≤+∞ jest pewne)
- $\lim_{x \to -\infty} F_X(x) = 0$ (gdyż zdarzenie poniżej min{x} jest niemożliwe)
- $F_X(x)$ jest niemalejąca tzn. $x_1 < x_2 \Rightarrow F_x(x_1) \le F_x(x_2)$ gdyż zdarzenie $\{X \le x_1\} \subset \{X \le x_2\}$
- $F_X(x)$ jest ciągła prawostronnie $\lim_{\delta \to 0^+} (F_X(x+\delta) - F_X(x)) = 0$
- $F_X(x)$ jest ciągła dla ciągłych zmiennych losowych
- $F_X(x)$ jest bezwymiarowa

2

Związek dystrybuanty z prawdopodobieństwem

- Tw. $\boxed{P(a < X \le b) = F_X(b) - F_X(a)}$
- Dow:

Zdarzenie $\{X \le b\} \equiv \{X \le a\} \cup \{a < X \le b\}$ (rozłączne) $/P(\,)$

$$P(\{X \le b\}) = P(\{X \le a\}) + P(\{a < X \le b\})$$
$$F_X(b) = F_X(a) + P(\{a < X \le b\})$$
$$P(\{a < X \le b\}) = F_X(b) - F_X(a)$$

3

Związek dystrybuanty z prawdopodobieństwem - wnioski

- Wniosek 1

  Weźmy $a = x - \delta, \quad b = x$

  $P(x - \delta < X \le x) = F_X(x) - F_X(x - \delta) \quad / \lim_{\delta \to 0}$

  $P(X = x) = F_X(x) - \lim_{\delta \to 0} F_X(x - \delta)$

  ale dla ciągłej zmiennej losowej $\lim_{\delta \to 0} F_X(x - \delta) = F_X(x)$

  zatem $\forall x: \quad P(X = x) = F_X(x) - F_X(x) = 0$
- Wniosek 2 (z Wniosku 1)

  $P(a \le X \le b) = P(a < X \le b) = P(a \le X < b) = P(a < X < b)$

  (związek z prawdopodobieństwem geometrycznym)

4

Dystrybuanta warunkowa

Syt:

X-zmienna losowa,

A – zdarzenie takie, że P(A)>0

- **Dystrybuantą warunkową** zmiennej losowej X pod warunkiem, że zaszło zdarzenie A nazywamy funkcję

$$F_X(x \mid A) = \frac{P(\{X \le x\} \cap A)}{P(A)}$$

5

Funkcja rozkładu prawdopodobieństwa

- Syt:

  X – **dyskretna** zmienna losowa

  S_X – odpowiedni zbiór wartości zmiennej X
- **Funkcją rozkładu prawdopodobieństwa** $P_X(x_k)$ nazywamy funkcję przyporządkowującą wartościom zmiennej losowej prawdopodobieństwo jej wystąpienia

$$\forall k: \; P_X(x_k) = P(X = x_k)$$

6

Funkcja rozkładu prawdopodobieństwa - własności

- $P_X(x_k) \geq 0$ (własność prawdopodobieństwa)
- $\sum_k P_X(x_k) = 1$ gdyż $P(S_X) = P(S) = 1$
- dla $x_1 < x_2 < \ldots$ oraz $k > 1$ mamy

  $P_X(x_k) = F_X(x_k) - F_X(x_{k-1})$
- $P_X(x_k)$ jest bezwymiarowe

7

Funkcja gęstości prawdopodobieństwa

- Syt:

  X – **ciągła** zmienna losowa o różniczkowalnej dystrybuancie

  S_X – odpowiedni zbiór wartości zmiennej X
- **Funkcją gęstości prawdopodobieństwa** $f_X(x)$ nazywamy funkcję $\boxed{f_X(x) = \frac{d}{dx} F_X(x)}$

$$P(x < X \leq x + \varepsilon) = F_X(x+\varepsilon) - F_X(x) =$$
$$= \frac{F_X(x+\varepsilon) - F_X(x)}{\varepsilon}\,\varepsilon \xrightarrow{\lim_{\varepsilon \to 0}} \frac{dF_X(x)}{dx}\,\varepsilon$$

8

Funkcja gęstości prawdopodobieństwa - własności

- $f_X(x) \geq 0$
- $P(x < X \leq x + dx) = P(x \leq X < x + dx) = \ldots = f_X(x)dx$
- $\int_{-\infty}^{x} f_X(t)dt = \int_{-\infty}^{x} \frac{d}{dt} F_X(t)dt = F_X(t)\Big|_{-\infty}^{x} = F_X(x) - 0 = F_X(x)$
- $\int_{-\infty}^{+\infty} f_X(t)dt = F_X(t)\Big|_{-\infty}^{+\infty} = 1 - 0 = 1$
- $P(a < X < b) = F_X(b) - F_X(a) = \int_a^b f_X(t)dt$
- $[\, f_X(t)\,] = \frac{1}{[x]}$ (wymiar)

9

Funkcja gęstości prawdopodobieństwa - własności

- Uwaga 1:

  Jeżeli zmienna losowa X przybiera wartości tylko z pewnego obszaru to poza tym obszarem definiujemy $f_X(x)=0$
- Uwaga 2:

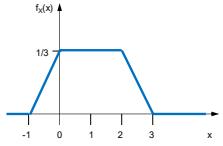
  Zarówno funkcja rozkładu prawdopodobieństwa jak i funkcja gęstości prawdopodobieństwa muszą być znormalizowane, gdy tak nie jest należy przeprowadzić normalizację $f_X(x) \to N \cdot f_X(x)$
- **Warunkową funkcją gęstości prawdopodobieństwa** nazywamy funkcję

$$f_X(x \mid A) = \frac{d}{dx} F_X(x \mid A)$$

10

Przykład: FGP (trapez) i odwracanie dystrybuanty

Syt: Zmienna losowa X opisywana jest funkcją gęstości prawdopodobieństwa w kształcie trapezu



$$f_X(x) = \begin{cases} 0 & dla\ x \in (-\infty, -1] \\ \frac{1}{3}x + \frac{1}{3} & dla\ x \in (-1, 0] \\ \frac{1}{3} & dla\ x \in (0, 2] \\ \frac{-1}{3}x + 1 & dla\ x \in (2, 3] \\ 0 & dla\ x \in (3, +\infty] \end{cases}$$

- Wysokość trapezu obliczamy z warunku normalizacji

  $Pole\ trapezu = \frac{(a+b)\cdot h}{2} = \frac{(4+2)\cdot h}{2} = 3h \;\rightarrow\; 1 = 3h \;\rightarrow\; h = \frac{1}{3}$
- Funkcję gęstości prawdopodobieństwa obliczamy wiedząc przez jakie punkty przechodzą poszczególne odcinki

11

Przykład: FGP (trapez) i odwracanie dystrybuanty

- Obliczamy dystrybuantę $F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt$

$$f_X(x) = \begin{cases} 0 & dla\ x \in (-\infty, -1] \\ \frac{1}{3}x + \frac{1}{3} & dla\ x \in (-1, 0] \\ \frac{1}{3} & dla\ x \in (0, 2] \\ \frac{-1}{3}x + 1 & dla\ x \in (2, 3] \\ 0 & dla\ x \in (3, +\infty] \end{cases}$$

$dla\ x \in (-\infty, -1]$: $F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt = \int_{-\infty}^{x} 0\,dt = 0$

$dla\ x \in (-1, 0]$: $F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt = \int_{-\infty}^{-1} 0\,dt + \int_{-1}^{x} \frac{1}{3}t + \frac{1}{3}\,dt = 0 + \left(\frac{1}{6}t^2 + \frac{1}{3}t\right)\Big|_{-1}^{x} =$

$= \frac{1}{6}x^2 + \frac{1}{3}x - \frac{1}{6}(-1)^2 - \frac{1}{3}(-1) = \frac{1}{6}x^2 + \frac{1}{3}x + \frac{1}{6}$

$dla\ x \in (0, 2]$: $F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt = \int_{-\infty}^{0} f_X(t)\,dt + \int_{0}^{x} \frac{1}{3}\,dt = F_X(0) + \int_{0}^{x} \frac{1}{3}\,dt =$

$= \left(\frac{1}{6}\cdot 0^2 + \frac{1}{3}\cdot 0 + \frac{1}{6}\right) + \left(\frac{1}{3}t\right)\Big|_{0}^{x} = \frac{1}{6} + \frac{1}{3}x - \frac{1}{3}\cdot 0 = \frac{1}{3}x + \frac{1}{6}$

$dla\ x \in (2, 3]$: $F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt = \int_{-\infty}^{2} f_X(t)\,dt + \int_{2}^{x} \frac{-1}{3}t + 1\,dt = F_X(2) + \int_{2}^{x} \frac{-1}{3}t + 1\,dt =$

$= \left(\frac{1}{3}\cdot 2 + \frac{1}{6}\right) + \left(\frac{-1}{6}t^2 + t\right)\Big|_{2}^{x} = \frac{5}{6} - \frac{1}{6}x^2 + x - \left(\frac{-1}{6}\cdot 2^2\right) - 2 = -\frac{1}{6}x^2 + x - \frac{1}{2}$

$dla\ x \in (3, +\infty]$: $F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt = \int_{-\infty}^{3} f_X(t)\,dt + \int_{3}^{x} 0\,dt = F_X(3) + 0 = -\frac{1}{6}\cdot 3^2 + 3 - \frac{1}{2} = 1$

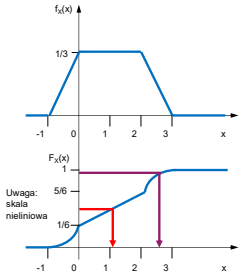
12

Przykład: FGP (trapez) i odwracanie dystrybuanty

- Generacja liczb pseudolosowych: odwracanie dystrybuanty

$$f_x(x)=\begin{cases}0 & dla\ x\in(-\infty,-1]\\ \frac{1}{3}x+\frac{1}{3} & dla\ x\in(-1,0]\\ \frac{1}{3} & dla\ x\in(0,2]\\ \frac{-1}{3}x+1 & dla\ x\in(2,3]\\ 0 & dla\ x\in(3,+\infty]\end{cases}$$

$$F_x(x)=\begin{cases}0 & dla\ x\in(-\infty,-1]\\ \frac{1}{6}x^2+\frac{1}{3}x+\frac{1}{6} & dla\ x\in(-1,0]\\ \frac{1}{3}x+\frac{1}{6} & dla\ x\in(0,2]\\ \frac{-1}{6}x^2+x-\frac{1}{2} & dla\ x\in(2,3]\\ 1 & dla\ x\in(3,+\infty]\end{cases}$$



13

Przykład: FGP (trapez) i odwracanie dystrybuanty

- Łącznie $F_x^{-1}(x)$

$$F_x^{-1}(x)=\begin{cases}\sqrt{6x}-1 & dla\ x\in(0,\frac{1}{6}]\\ 3x-\frac{1}{2} & dla\ x\in(\frac{1}{6},\frac{5}{6}]\\ 3-\sqrt{6-6x} & dla\ x\in(\frac{5}{6},1]\end{cases}$$

Generator liczb pseudolosowych o badanej $f_X(x)$:

- Losujemy jednorodnie liczbę *x* z przedziału (0,1)
- Obliczamy $y := F_x^{-1}(x)$ zgodnie z powyższym wzorem
- Zwracamy *y*
- Liczby y będą miały zadaną funkcję gęstości prawdopodobieństwa w przedziale (-1,3)

Kiedy metoda odwracania dystrybuanty nie zadziała (od razu) ?

14

Odwracanie dystrybuanty jako metoda generowania liczb pseudolosowych

- Syt: dana funkcja gęstości prawdopodobieństwa $f_X(x)$, chcemy uzyskać liczby pseudolosowe przez nią opisywane.
- Obliczamy dystrybuantę $F_X(x)$.
- Znajdujemy funkcję odwrotną do dystrybuanty $F^{-1}{}_X(x)$.
- Losujemy liczbę *y* z przedziału <0,1> używając „standardowego" (zaimplementowanego) generatora liczb pseudolosowych.
- Obliczmy $z=F^{-1}{}_X(y)$.
- Tak otrzymane liczby opisane są funkcją gęstości prawdopodobieństwa $f_X(x)$

15

Globalny opis dystrybuanty i funkcji gęstości prawdopodobieństwa

16

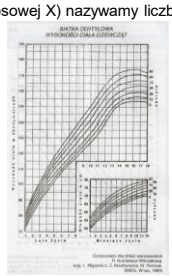
Globalny opis dystrybuanty i funkcji gęstości prawdopodobieństwa

- **Kwantylem rzędu p** (dla zmiennej losowej X) nazywamy liczbę x_p: $F_X(x_p)=p \quad (0\le p\le 1)$

W szczególności:

- **Mediana** to kwantyl rzędu ½
- **Kwartyle** to kwantyle
  rzędu ¼ (pierwszy kwartyl),
  rzędu ½ (drugi kwartyl),
  rzędu ¾ (trzeci kwartyl)
- **Percentyle** to kwantyle
  rzędu 0.01, 0.02, …, 0.99

- **Moda** – wartość najbardziej prawdopodobna czyli x: max{ $f_X(x)$ }



17