Materiały do wykładu PMPiSt 12.11.2020
tryb ON-LINE

Temat. Rozkłady ciągłe – przykłady. Parametry rozkładów.

Zaczniemy od przykładu.

P1. Niech $X \in U([0,1])$. Znaleźć rozkład $Y = \ln(X+1)$.
Z założenia mamy MPK (Ω, Σ, P), na którym określona jest zmienna losowa X, a więc i Y.

Dalej

$$F_Y(t) = P(\{\omega \in \Omega : Y(\omega) < t\}), \quad t \in \mathbb{R}.$$

Zauważmy, że $\forall_{\omega \in \Omega}\ X(\omega) \in [0,1]$, a więc

$X(\omega) + 1 \geq 1$, co z własności funkcji logarytm oznacza, że $\forall_{\omega \in \Omega}\ \cancel{X}(\omega) \geq 0$.

Dalej dla $t \leq 0$ $F_Y(t) = 0$ (dlaczego?).
Możemy więc założyć, że $t > 0$. Wtedy

$$F_Y(t) = P(\{\omega \in \Omega : \ln(X(\omega)+1) < t\}) =$$
$$= P(\{\omega \in \Omega : X(\omega)+1 < e^t\}) = P(\{\omega \in \Omega : X(\omega) < e^t - 1\})$$

Oznacza to, iż

$$F_Y(t) = \begin{cases} 0, & t \leq 0 \\ F_X(e^t - 1), & t > 0 \end{cases}.$$

Wniosek: Y ma rozkład ciągły (dlaczego?)

Dla wyznaczenia rozkładu znajdźmy f. gęstości Y,

$$f_Y(t) = F'_Y(t) = \begin{cases} 0, & t \leq 0 \\ \frac{d}{dt} F_X(e^t - 1), & t > 0 \end{cases}$$

Skorzystajmy z reguły różniczkowania f. złożonej:

Dla $t > 0$

$$f_Y(t) = F'_X(e^t - 1) \cdot e^t = f_X(e^t - 1) e^t.$$

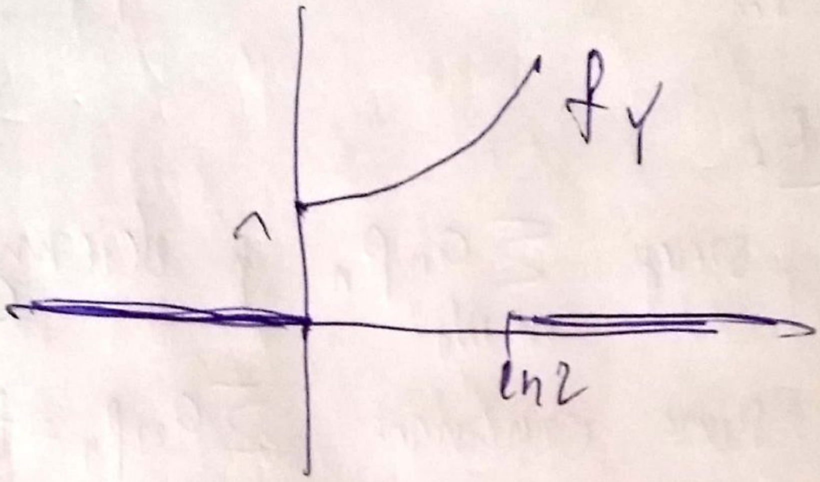
Ale z założenia $f_X(u) = \begin{cases} 1 & u \in [0,1] \\ 0, & u \notin [0,1] \end{cases}$

$0 \leq u = e^t - 1 \leq 1 \iff 1 \leq e^t \leq 2 \iff$

$0 \leq t \leq \ln 2$

Oznaczmy, ...

$$f_Y(t) = \begin{cases} e^t & t \in [0, \ln 2] \\ 0, & t \notin [0, \ln 2] \end{cases}$$



P2. Niech $X \in \mathcal{U}([0,1])$. Dla $\lambda > 0$ definiujmy

$$Y = -\frac{1}{\lambda} \ln(1-X).$$

Wyznaczyć rozkład Y

Z założenia X (a więc i Y) określone są na pewnym MPK (Ω, Σ, P).

Ponieważ $P(\{\omega \in \Omega : X(\omega) \in [0,1)\}) = 1$ (dlaczego?)

zmienna losowa Y dana wzorem

$$Y(\omega) = -\frac{1}{\lambda} \ln(1 - X(\omega))$$ określona jest poprawnie.

Ponadto, $\forall \omega \in \Omega$ $Y(\omega) > 0$, co oznacza, że

$$F_Y(t) = 0 \text{ dla } t \le 0.$$

Niech $t > 0$. Wtedy

$$F_Y(t) = P(\{\omega \in \Omega : Y(\omega) < t\}) = P(\{\omega \in \Omega : -\tfrac{1}{\lambda}\ln(1 - X(\omega)) < t\})$$

$$= P(\{\omega \in \Omega : \ln(1 - X(\omega)) > -\lambda t\}) =$$

$$= P(\{\omega \in \Omega : 1 - X(\omega) > e^{-\lambda t}\}) =$$

$$= P(\{\omega \in \Omega : X(\omega) < 1 - e^{-\lambda t}\}) = F_X(1 - e^{-\lambda t}).$$

Mamy zatem:

$$F_Y(t) = \begin{cases} 0, & t \le 0 \\ F_X(1 - e^{-\lambda t}), & t > 0, \end{cases}$$

co oznacza, że Y ma rozkład ciągły.

Ponadto $f_Y(t) = 0$, $t \le 0$, oraz dla $t > 0$

$$f_Y(t) = \frac{d}{dt} F_X(1 - e^{-\lambda t}) = f_X(1 - e^{-\lambda t}) \cdot (-e^{-\lambda t}(-\lambda))$$

$$= f_X(1-e^{-\lambda t})\lambda e^{-\lambda t}.$$

Mamy więc

$$f_Y(t) = \begin{cases} 0, & t \leq 0 \\ \lambda e^{-\lambda t}, & 0 \leq 1-e^{-\lambda t} \leq 1 \end{cases}$$

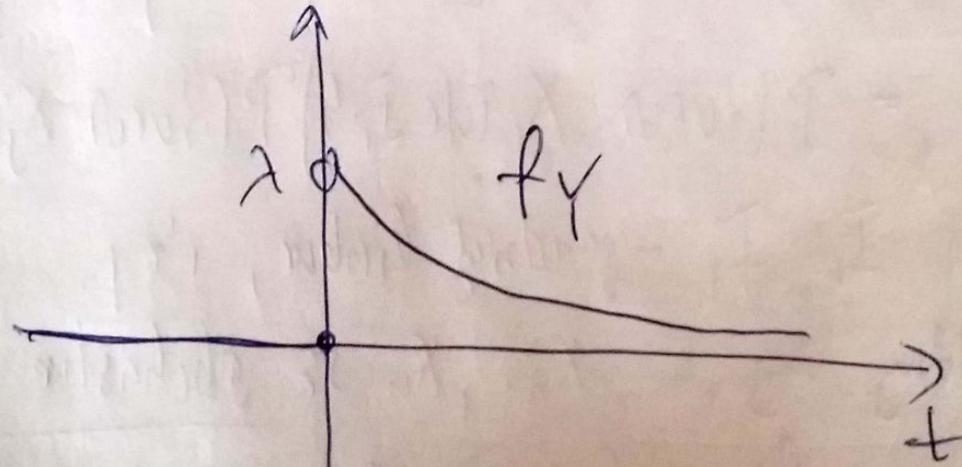
Ale

$$0 \leq 1-e^{-\lambda t} \leq 1 \iff$$

$$0 \leq e^{-\lambda t} \leq 1 \iff t > 0,$$

zatem

$$f_Y(t) = \begin{cases} 0, & t \leq 0 \\ \lambda e^{-\lambda t}, & t > 0 \end{cases}$$



P3. Znaleźć dystrybuantę zmiennej losowej z P2.

Z postaci f_Y, $F_Y(t)=0$ dla $t\le 0$ (dlaczego?).

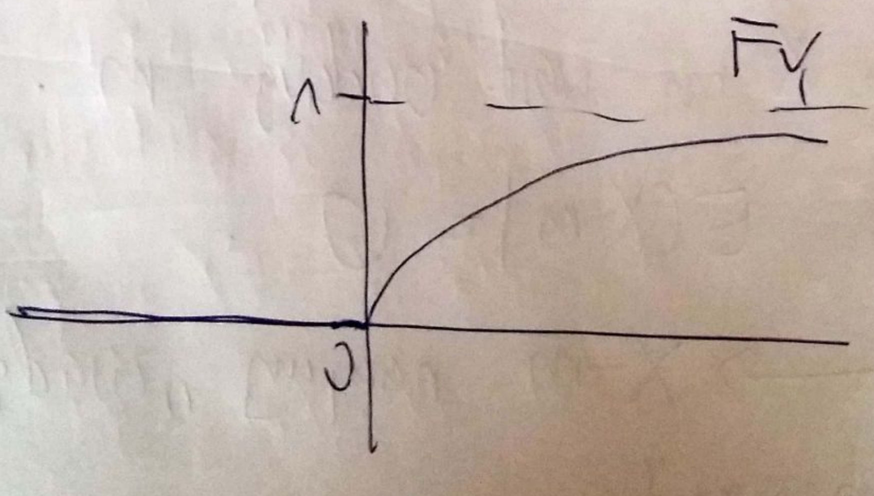
Niech $t>0$. Wtedy

$$F_Y(t)=\int_{-\infty}^{t} f_Y(u)\,du=\int_{-\infty}^{0} 0\,du+\int_{0}^{t} f_Y(u)\,du$$

$$=\int_0^t \lambda e^{-\lambda u}\,du=\lambda\left(-\frac{1}{\lambda}e^{-\lambda u}\right)\Big]_0^t=1-e^{-\lambda t}.$$

Mamy zatem

$$F_Y(t)=\begin{cases}0, & t\le 0\\ 1-e^{-\lambda t}, & t>0\end{cases}$$



Def

Mówimy, że $X \in W(\lambda)$, (rozkład wykładniczy z parametrem $\lambda > 0$), jeśli

$$f_X(t) = \begin{cases} 0, & t \leq 0 \\ \lambda e^{-\lambda t}, & t > 0 \end{cases}$$

Parametry rozkładu prawdopodobieństwa.

Przypomnijmy, mówiąc o zmiennej losowej X mamy na myśli dwie sytuacje: albo F_X jest schodkowa (X ma rozkład dyskretny), albo F_X jest różniczkowalna (X ma rozkład ciągły).

Uwaga.

Nie są to jedyne rozkłady. Ze względów technicznych na naszym kursie ograniczymy się tylko do tych dwóch.

Niech X będzie zmienną losową. Wtedy można próbować wykonać na X operację całkowania po Ω względem P, czyli

$$X \longrightarrow \int_\Omega X\,dP$$

Jeśli wynikiem jest liczba, to jej wartość oznaczamy EX, czyli

$$X \longrightarrow \int_\Omega X\,dP = EX \in \mathbb{R}$$

Uwaga. Symbol „E" pierwotnie oznaczał „nadzieję matematyczną" (z franc. „ESPOIRE").
Współcześnie „E" to expectation

Liczbę EX - wartość oczekiwana X, nazywamy też „średnią" rozkładu - oznaczamy przez m (mean).
Nazywamy też pierwszym momentem rozkładu i oznaczamy m_1. Mamy więc

$$\int_\Omega X\,dP = EX = m = m_1$$

$X \to EX$ jako operacja całkowania ma własności operacji liniowej, czyli

FAKT 1 (liniowość)

Niech X_1, X_2 mają swoje wartości oczekiwane. Wtedy dla dowolnych $a_1, a_2 \in \mathbb{R}$, zm. losowa $a_1 X_1 + a_2 X_2$ też ma wartość oczekiwaną i zachodzi równość

$$E(a_1 X_1 + a_2 X_2) = a_1 EX_1 + a_2 EX_2 .$$

Uwaga.
Jeśli X jest osobliwa, czyli $\underset{\omega \in \Omega}{\forall} X(\omega) = c \in \mathbb{R}$

to $EX = c$.

Problem. Jak dla danego rozkładu X obliczyć EX?

I Przypadek rozkładu dyskretnego.

Załóżmy: $X(\Omega) = \{a_n \in \mathbb{R},\ n \in \mathbb{N}_0 \subset \mathbb{N}\}$

$$p_n = P(\{\omega \in \Omega : X(\omega) = a_n\}),\ n \in \mathbb{N}_0.$$

FAKT.

Jeśli szereg $\sum_{n \in \mathbb{N}_0} a_n p_n$ jest zbieżny, to jego suma oznaczana $\sum_{n \in \mathbb{N}_0} a_n p_n = EX$.

Uwagi

1°. Gdy X ma rozkład skończony, to EX istnieje zawsze

$$(*)\quad EX = \sum_{n=1}^{k} a_n p_n, \text{ gdzie}$$

$$X(\Omega) = \{a_1, a_2, \ldots, a_k\}$$

Wtedy liczbę $(*)$ nazywamy też „średnią ważoną" (wagi pełnią prawdop. p_n).

W szczególności, gdy X ma rozkład <u>jednostajny</u>, czyli $p_1 = p_2 = \ldots = p_k = \frac{1}{k}$, to

$$(\ast\ast) \quad EX = \frac{1}{k}\sum_{n=1}^{k} a_n$$ – taka średnia

nazywa się <u>arytmetyczna</u>.

<u>Uwaga</u>. Między innymi z tego powodu EX nazywamy <u>średnią</u>.

<u>Przykłady</u>.

1°. X ma d_X: $\begin{array}{c|c} 1 & 0 \\ \hline p & q \end{array}$ $\quad q = 1-p$

$$EX = 1\cdot p + 0\cdot q = p$$

2°. $X \in B(n,p)$. Wtedy

$$EX = \sum_{k=0}^{n} k\binom{n}{k}p^k(1-p)^{n-k} = \sum_{k=1}^{n}\frac{n!}{(k-1)!(n-k)!}p^k(1-p)^{n-k}$$

Podstawiając $k-1=j$ otrzymamy

$$EX = \sum_{j=0}^{n-1}\frac{n!}{j!(n-1-j)!}p^{j+1}(1-p)^{n-1-j} =$$

$$= np \sum_{j=0}^{n-1} \frac{(n-1)!}{j!(n-1-j)!} p^j (1-p)^{n-1-j} = np(p+1-p)^{n-1}.$$

Zatem $\underline{EX = np}$.

Uwaga.

Można inaczej. Potrzebny jest do tego następujący fakt

FAKT 3 (o faktoryzacji $X \in B(n,p)$)

$$X = X_1 + X_2 + \ldots + X_n,$$

gdzie każda X_j ma rozkład $\begin{array}{c|c} 1 & 0 \\ \hline p & 1-p \end{array}$,

oraz $P(\{\omega\in\Omega : X_i(\omega)\in I_i \wedge X_j(\omega)\in I_j\})$

$= P(\{\omega\in\Omega : X_i(\omega)\in I_i\})\,P(\{\omega\in\Omega : X_j(\omega)\in I_j\})$,

gdzie I_i, I_j – przedziały liczbowe, $i \neq j$.

Mówimy wtedy, że $X_1, \ldots, X_n$ są stochastycznie niezależne.

Wtedy, z liniowości operacji $X \to EX$ i przykładu 2^0,

$$EX = EX_1 + EX_2 + \ldots + EX_n = \underbrace{p + p + \ldots + p}_{n} = np.$$

3. $X \sim P(\lambda)$, $\lambda > 0$.

Teraz mamy wzory $\sum_{k \geq 0} k e^{-\lambda} \frac{\lambda^k}{k!} =$

$$= e^{-\lambda} \sum_{k \geq 1} k \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k \geq 1} \frac{\lambda^k}{(k-1)!} \quad (***)$$

Podstawiamy ~~k-1~~ $k-1=j$

$$(***) = e^{-\lambda} \sum_{j \geq 0} \frac{\lambda \cdot \lambda^j}{j!} = \lambda e^{-\lambda} \underbrace{\sum_{j \geq 0} \frac{\lambda^j}{j!}}_{= e^{\lambda}} = \lambda e^{-\lambda} e^{\lambda} = \lambda$$

Zatem X ma wartość oczekiwaną

$$EX = \lambda.$$

FAKT 4 (Zasada centrowania).

Niech X ma wart. oczekiwaną m.

Wtedy $E(X-m) = 0$.

$X \longrightarrow X-m$ nazywamy „zasadą centrowania”