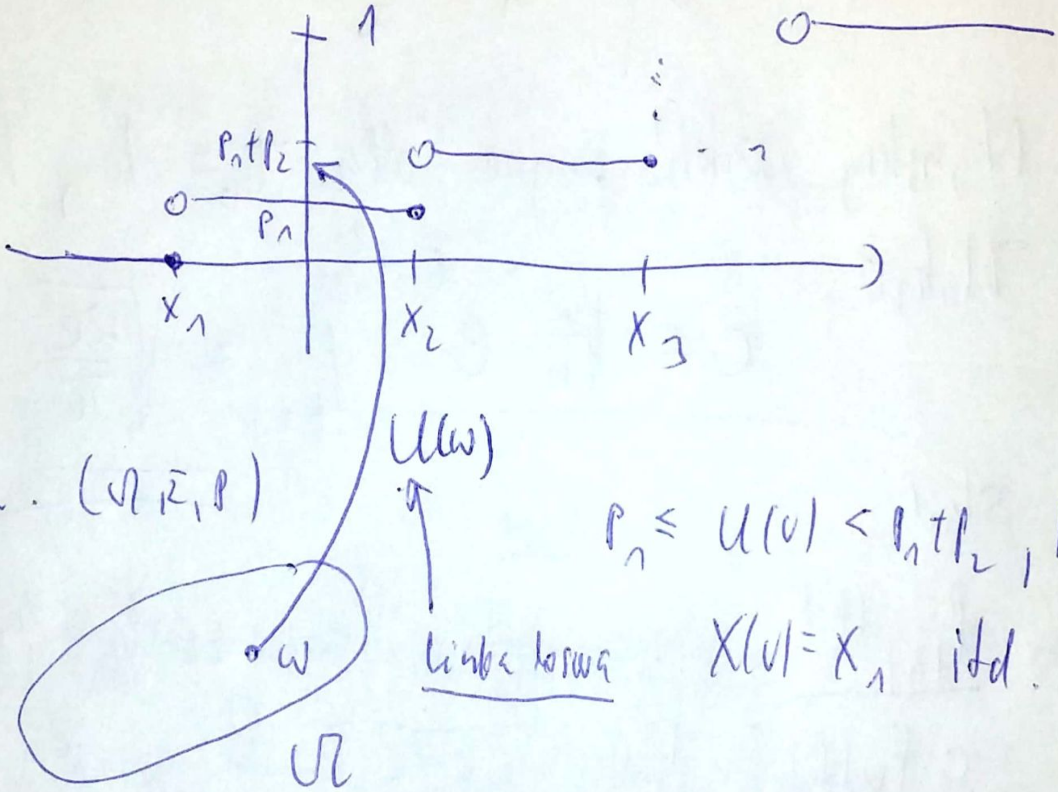
PSK - zajęcia zaplanowane na ... -

Temat. Generowanie rozkładów prawdop. c.d.

(II) Przypadek rozkładu ciągłego

Przypomnijmy ideę generowania rozkładu X dla p. dyskretnego.

Jeśli $P(\{\omega\in\Omega : X(\omega)=x_i\})=p_i$, czyli



Z.L. $(\Omega, \mathcal{F}, P)$

$p_1 \leq U(\omega) < p_1+p_2$, to

$X(\omega)=x_1$ itd.

Z tego powodu poznaliśmy ten rozwiązały MOO - metodą odwr. odwrotnego.

Zaadaptujmy tę metodę do przypadku rozkładu ciągłego.

Założenie X ma rozkład ciągły, a więc

$X(\Omega)=I$ - przedział (lewi może nieograniczony)

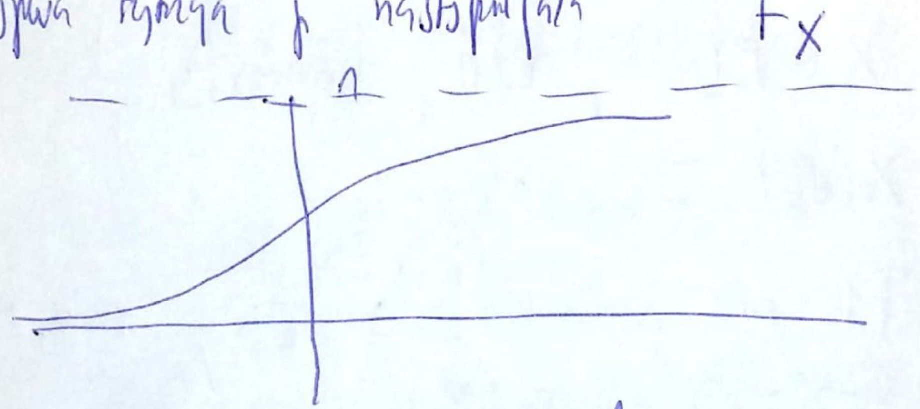
Wtedy dla

$$\mathbb{R} \ni t \longrightarrow F_X(t) = P(\{\omega \in \Omega : X(\omega) < t\}) \in [0,1],$$

dystrybuanta F_X jest różniczkowalna, gdzie

$$\mathbb{R} \ni t \to \frac{d}{dt} F_X(t) = f_X(t)$$ – funkcja gęstości rozkładu

Typowa sytuacja jest następująca



rys 1.

Wtedy F_X jako rosnąca jest odwracalna. Dlatego dla F_X

$F_X : \mathbb{R} \longrightarrow [0,1]$ istnieje funkcja odwrotna

$F_X^{-1} : [0,1] \longrightarrow \mathbb{R}$, czyli

$$\forall t \in \mathbb{R} \quad F_X^{-1}(F_X(t)) = t$$

Niech $U \in \mathcal{J}([0,1])$ – zm. losowa o rozkładzie jednostajnym, a więc

$$\forall_{0 \le a < b \le 1} \quad P(\{\omega \in \Omega : a \le U(\omega) < b\}) = b - a$$

Bierzemy złożenie

$$\begin{tikzcd} \Omega \ni \omega \arrow[r] \arrow[rr, bend right, "Y"'] & U(\omega) \ (\in [0,1]) \arrow[r] & F_X^{-1}(U(\omega)) \in \mathbb{R} \end{tikzcd}$$

Wyznaczmy dystrybuantę F_Y zmiennej losowej $Y = F_X^{-1} \circ U$

$$F_Y(t) = P(\{\omega \in \Omega : Y(\omega) < t\}) =$$

$$= P(\{\omega \in \Omega : F_X^{-1}(U(\omega)) < t\}) =$$

$$= P(\{\omega \in \Omega : U(\omega) < F_X(t)\}) =$$

$$= P(\{\omega \in \Omega : 0 \leq U(\omega) < F_X(t)\}) =$$

$$= F_X(t) - 0 = F_X(t) .$$

Zatem $d(Y) = d(X)$.

<u>WNIOSEK</u>. Możemy generować wartości $X(\omega) \in \mathbb{R}$ za pomocą dystrybuanty F_X i generatora liczb (pseudo)losowych!

P1. Wygenerujmy $X(\omega)$, jeśli

$$F_X(t) = \begin{cases} t^n, & t \in [0,1] \\ 1 & t \geq 1 \\ 0 & t \leq 0 \end{cases}$$

Oznacza to, iż $X(\Omega) = [0,1]$ - skoncentrowany p na $[0,1]$. Dlatego ograniczmy F_X do $[0,1]$.

Jeśli $F_X(t) = u$, dla $t \in [0,1]$, czyli

$$t^n = u, \text{ to } t = \sqrt[n]{u}.$$

Dlatego $F_X^{-1} \circ U(\omega) = \sqrt[n]{U(\omega)}$.

Generując teraz liczby (pseudo)losowe $U(\omega)$, mamy wartości (losowe) $X(\omega)$ w postaci $\sqrt[n]{U(\omega)}$.

P.2. Niech $X \in W(\lambda)$, $\lambda > 0$ - rozkład wykładniczy,

czyli

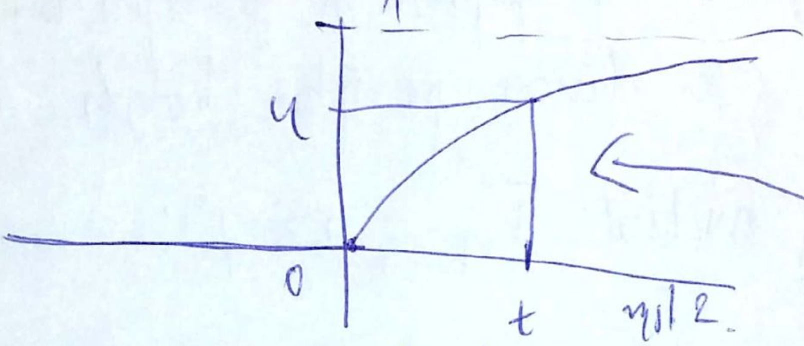
$$f_X(t) = \begin{cases} \lambda e^{-\lambda t}, & t \geq 0 \\ 0, & t \leq 0 \end{cases}$$

Zatem $X(\Omega) = [0, +\infty)$

Jak wiemy, wtedy

$$\mathbb{R} \ni t \longrightarrow F_X(t) = P(\{\omega\in\Omega : X(\omega) < t\})$$

$$= \begin{cases} 1-e^{-\lambda t}, & t \geq 0 \\ 0, & t \leq 0. \end{cases}$$



rys. 2.

Dla uproszczenia, niech dalej $\underline{\lambda = 1}$. Z powyższego ograniczmy F_X do $[0,+\infty)$. Wtedy

$$[0,+\infty) \ni t \longrightarrow F_X(t) = 1-e^{-t} = u \in [0,1)$$

$$\equiv 1-u = e^{-t} \equiv t = -\ln(1-u),\ u\in[0,1).$$

Ale, jeśli $U(\omega)$ jest liczbą (pseudo)losową, to $1-U(\omega)$ również. Stąd wynikiem generowania $X(\omega)$ jest $\underline{-\ln U}$.

I ogólnie, dla $\lambda > 0$, $\underline{\underline{-\frac{1}{\lambda}\ln U}}$

W teorii rozkładu ciągłych częściej przedstawiany je za pomocą funkcji gęstości, aniżeli dystrybuanty.

Pokażmy teraz metodę generowania $X(\omega)$ bezpośrednio z f_X. Nazywa się ona metodą eliminacji* (ME) i należy do J. von Neumanna.

Przedstawienie ME.

Niech (jak wcześniej) X, której wartości chcemy generować ma f. gęstości f_X.

Przypuśćmy, że umiemy generować (np. metodą MOO) wartości zm. losowej Y o f. gęstości f_Y, gdzie

$$\frac{f_X(t)}{f_Y(t)} \le c \quad , t \in \mathbb{R} \quad \text{dla pewnej } c \in \mathbb{R}_+$$

Wtedy stosując algorytm ME przedstawiony poniżej, wygenerujemy X:

Krok 1: generujemy Y o gęstości f_Y

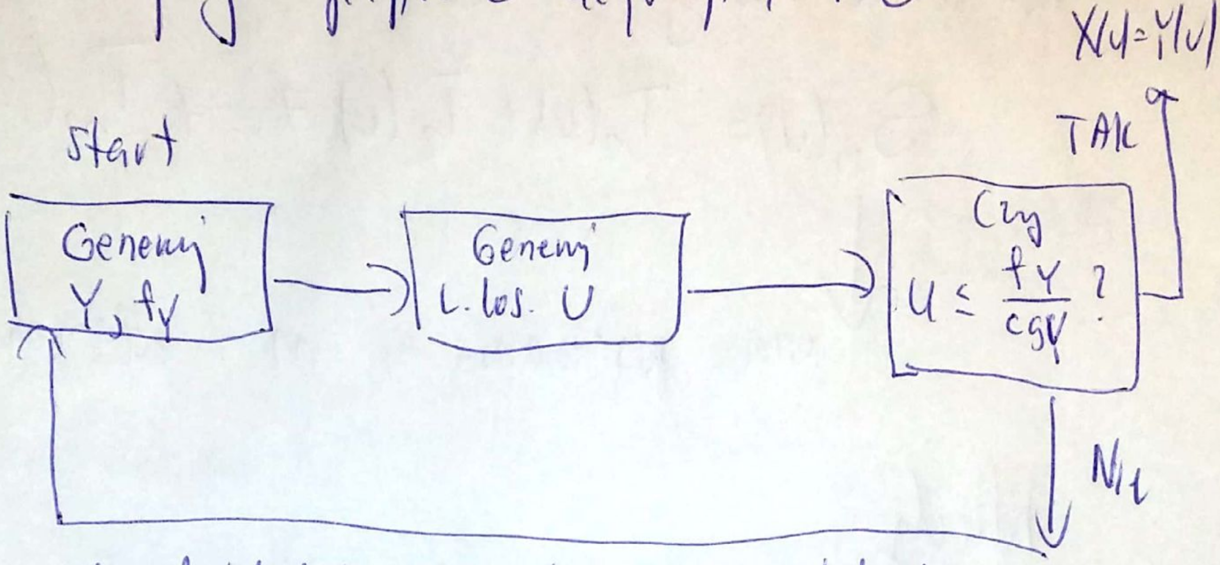
Krok 2: generujemy l. (pseudo)losową $U(\omega)$

Krok 3: Jeśli $U(\omega) \le \frac{f_X(t)}{c f_Y(t)}$, $X(\omega) = Y(\omega)$.

W przeciwnym razie

* Rejection Method (z ang.)

Zilustrujmy graficznie algorytm ME



Wyjaśnijmy to dokładniej na kolejnym przykładzie.

<u>P.3</u>. Użyjemy ME do X z f. gęstości

$$f_X(t) = \begin{cases} 20t(1-t)^3, & t \in (0,1) \\ 0, & t \notin (0,1) \end{cases}$$

Z założenia $X(\Omega) = (0,1)$ - jest skoncentrowana na $(0,1)$.

<u>Bierzemy</u> $Y \in \mathcal{U}([0,1])$, $f_Y(t) = \begin{cases} 1, & t \in [0,1] \\ 0, & t \notin [0,1] \end{cases}$

Wyznaczamy najmniejsze $c \in \mathbb{R}_+$, takie, że

$$\frac{f_X(t)}{f_Y(t)} \le c, \quad t \in [0,1]$$

W tym celu bierzemy funkcję

$$h(t) = \frac{f_X(t)}{g_X(t)} = 20t(1-t)^3, \quad t \in [0,1]$$

i metodą rachunku różniczkowego wyznaczmy jej wartość największą.

Mamy kolejno:

$$h'(t) = 20(1-t)^3 + 60t(1-t)^2 \cdot (-1)$$

$$= 20(1-t)^2[1-t-3t] =$$

$$= 20(1-t^2)(1-4t).$$

Wyznaczmy zbiór punktów osobliwych: są to końce przedziału $[0,1]$ i rozwiązanie równania:

$$t \in (0,1); \quad h'(t) = 0 \iff t = 1/4.$$

Mamy $\{0, 1/4, 1\}$.

Bierzemy $h(0) = 0$, $h(1) = 0$,

$$h\left(\tfrac{1}{4}\right) = 5 \cdot \left(\tfrac{3}{4}\right)^3 = \frac{135}{64} = c.$$

Stąd $\forall\, t \in [0,1]$

$$\frac{f_X(t)}{c\,g_X(t)} = \frac{20t(1-t)^3}{\frac{135}{64}} = \underline{\frac{256}{27}t(1-t)^3}$$

Generujemy $X(U)$:

~~Krok~~. Ponieważ Y ma rozkład $\mathcal{U}([0,1])$, wynikiem generowania Y jest liczba pseudolosowa.

Dlatego generowanie X ma postać:

<u>Krok 1</u> generuj dwie liczby U_1, U_2

<u>Krok 2</u>. jeśli $U_2 \leq \frac{256}{27} U_1 (1-U_2)^3$,

to STOP i $X(U) = U_1$.

W przeciwnym razie wracamy <u>do Kroku 1</u>.

<u>ZADANIE 1</u>

Napisać procedurę realizującą <u>P.3</u>.

<u>ZADANIE 2</u>

Zaimplementować wynik z <u>P.1</u>.

Pk. Przypadek $X \in N(0,1)$

Jak wiemy, w tym przypadku X ma gęstość:

$$f_X(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}, \quad t \in \mathbb{R}$$

Bierzemy $|X|$ - wartość bezwzględną. Z symetrii wykresu

$$f_{|X|}(t) = \begin{cases} \frac{2}{\sqrt{2\pi}} e^{-t^2/2}, & t \in [0, +\infty) \\ 0, & t < 0, \end{cases}$$

bowiem $|X|(\Omega) = [0, +\infty)$.

Za rozkład bazowy Y bierzemy: $Y \in W(\lambda)$, $\lambda = 1$,

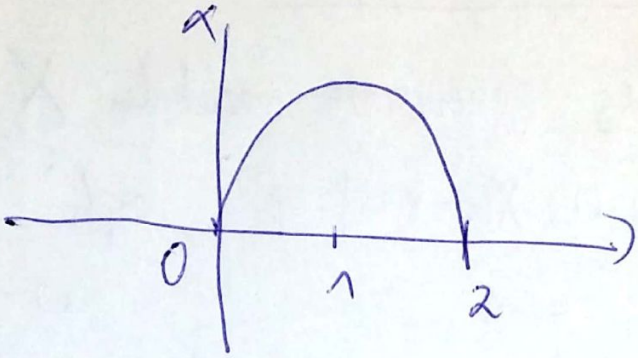
czyli $f_Y(t) = \begin{cases} e^{-t}, & t \geq 0 \\ 0, & t < 0. \end{cases}$

Mamy: $$\frac{f_{|X|}(t)}{f_Y(t)} = \frac{\frac{2}{\sqrt{2\pi}} e^{-t^2/2}}{e^{-t}} = \sqrt{\frac{2}{\pi}}\, e^{t - t^2/2}, \quad t \geq 0$$

Aby znaleźć stałą c - największą wartość tego ilorazu weźmy trójmian kwadratowy

- 11 -

$$h(t) = t - \frac{t^2}{2} = t(1 - t/2)$$



Największą wartość przyjmuje dla $t_0 = 1$, $h(1) = 1/2$.

Dlatego

$$c = \sqrt{\frac{2}{\pi}}\, e^{t - t^2/2}\Big|_{t=1} = \underline{\sqrt{\frac{2e}{\pi}}}.$$

Stąd

$$\frac{f_{|X|}(t)}{c\, f_Y(t)} = \sqrt{\frac{\pi}{2e}}\, \frac{\frac{2}{\sqrt{2\pi}} e^{-t^2/2}}{e^{-t}} = e^{t - t^2/2 - 1/2} = e^{-\frac{(t-1)^2}{2}},$$

dlatego algorytm AE ma postać:

<u>Krok 1</u>: generuj $Y(\omega) \sim W(\lambda)$, $\lambda = 1$ (jak w P2)

<u>Krok 2</u>: generuj liczbę losową $U(\omega)$

<u>Krok 3</u>: Jeśli $U(\omega) \le e^{-\frac{(Y(\omega)-1)^2}{2}}$, przyjmij $|X|(\omega) = Y(\omega)$.

W przeciwnym razie idź do <u>Krok 1</u>.

ZADANIE 3.

Zaimplementuj poniższy algorytm.

Proces Poissona i sposób jego generowania.

Zacznijmy od przypomnienia rozkładu Poissona i jego interpretacji.

1°. $X \in \mathcal{P}(\lambda)$, $\lambda > 0$, oznacza, że

$X(\Omega) = \mathbb{Z}_+ \cup \{0\}$

$P(\{\omega \in \Omega : X(\omega) = k\}) = e^{-\lambda} \frac{\lambda^k}{k!}$, $k = 0, 1, 2 \ldots$

$EX = \lambda = Var(X)$

2. Niech dla $\omega \in \Omega$, $X(\omega)$ # zaobserwowanych zdarzeń.

Załóżmy, że zdarzenia te możemy podzielić na $r \geq 2$ typów, gdzie dla i-tego typu prawdop. pojawienia się takiego zdarzenia wynosi p_i oraz $p_1 + p_2 + \ldots + p_r = 1$.

Niech $X_i(\omega)$ - # zaobserwowanych zdarzeń typu i.

Wtedy $X(\omega) = X_1(\omega) + X_2(\omega) + \ldots + X_r(\omega)$,

gdzie $X_i \in \mathcal{P}(\lambda_i)$, $\lambda_i = \lambda p_i$

Jeśli dodatkowo zdarzenia pojawiają się niezależnie, to X_{Λ}, X_{γ} są niezależne.

Przypuśćmy teraz, że zdarzenia pojawiają się losowo i obserwujemy ich liczbę wraz z upływem czasu $t \geq 0$.

Dokładniej, niech

$$X_t(\omega) = \#\ \text{zdarzeń, które pojawiły się w przedziale czasowym } [0,t).$$

Jeśli:

a) $X_0(\omega) = 0$, $\omega \in \Omega$

b) X_t oraz $X_{t+s} - X_t$ $\forall s > 0$ są niezależne, co oznacza, że liczby zdarzeń zdarzeń zachodzących na rozłącznych przedziałach czasowych są niezależne

c) rozkład liczby zdarzeń na danym przedziale zależy tylko od jego długości, a nie lokalizacji, czyli $X_{t+s} - X_t$ ma rozkład niezależny od t, a tylko od s

d) $\dfrac{P(\{\omega : X_h(\omega) = 1\})}{h} \xrightarrow[h \to 0]{} \lambda$

co oznacza, że jeśli przedział jest dostatecznie krótki, to prawdop. zaobserwowania pojedynczego zdarzenia $\approx \lambda h$

e) $$\frac{P(\{\omega : X_h(\omega) \geq 2\})}{h} \xrightarrow[h \to 0]{} 0 ,$$

natomiast prawdop. zaobserwowania dwóch i więcej ≈ 0

Wtedy mówimy, że $(X_t)_{t \geq 0}$ jest jednorodnym procesem Poissona o współczynniku λ
intensywności

Jeśli $(X_t)_{t \geq 0}$ jak wyżej, to:

Fakt 1 $\quad X_t \in P(\lambda t) \, , \; t > 0$

Fakt 2 $\quad X_{t+s} - X_t \in P(s\lambda) , \; s > 0 .$

Zdefiniujmy teraz proces pomocniczy $(T_n)_{n \geq 1}$, gdzie: $T_1(\omega)$ jest czasem pojawienia się pierwszego zdarzenia, które jest rejestrowane procesem Poissona $(X_t)_{t \geq 0}$.

W ogólności, T_n oznacza czas jaki upłynął pomiędzy pojawieniem się zdarzenia $n-1$ a n.

Jeśli mamy np. $T_1(\omega) = 5$, $T_2(\omega) = 15$, to:

pierwsze zdarzenie pojawiło się po upływie 5 jednostek czasu, a drugie po upływie kolejnych 15.

Wyznaczmy rozkład T_n, $n \geq 1$:

Zauważmy, że $\omega \in \{\omega \in \Omega : T_1(\omega) > t\} \equiv$

$X_t(\omega) = 0$. Ale $X_t \in \mathcal{P}(\lambda t)$, więc

$$P(\{\omega \in \Omega : T_1(\omega) > t\}) = P(\{\omega \in \Omega : X_t(\omega) = 0\})$$

$$= e^{-\lambda t},$$ co dowodzi, że

$$T_1 \in W(\lambda) \text{ oraz } ET_1 = \frac{1}{\lambda}.$$

Wyznaczmy teraz rozkład T_2:

$$P(\{T_2(\omega)>t\} \mid \{\omega\in\Omega: T_1(\omega)=s\})$$

$$= P(\{\text{zaszło } 0 \text{ zdarzeń w przedziale } (s,s+t]\} \mid \{\omega\in\Omega: T_1(\omega)=s\})$$

$$\parallel$$

$$X_{s+t}(\omega)-X_s(\omega)$$

$$= P(\{\omega\in\Omega: X_{s+t}(\omega)-X_s(\omega)=0 \mid X_s(\omega)=1\})$$

$$= \frac{\cancel{P(\{\omega\in\Omega: X_s(\omega)=1\})} \cdot P(\{\omega\in\Omega: X_{s+t}(\omega)-X_s(\omega)=0\})}{\cancel{P(\{\omega\in\Omega: X_s(\omega)=1\})}}$$

$$= P(\{\omega\in\Omega: X_{s+t}(\omega)-X_s(\omega)=0\})$$

Ale $X_{s+t}-X_s$ ma rozkład X_t, dlatego

$$P(\{\omega\in\Omega: T_2(\omega)>t\}) = P(\{\omega\in\Omega: X_t(\omega)=0\})$$
$$= e^{-\lambda t},$$ co

dowodzi, że $T_2\in W(\lambda)$ i $ET_2=\frac{1}{\lambda}$.

I ogólnie: Ciąg zmiennych $(T_n)_{n\geq 1}$ jest niezależny, $T_n\in W(\lambda)$, $n\geq 1$.

Na końcu mamy wziąć

$$S_n(\omega) = T_1(\omega) + T_2(\omega) + \dots + T_n(\omega)$$

$\downarrow$ czas pojawienia się n - zdarzeń.

Wtedy

$$P(\{\omega\in\Omega : S_n(\omega) \leq t\}) = P(\{\omega\in\Omega : X_t(\omega) \geq n\})$$

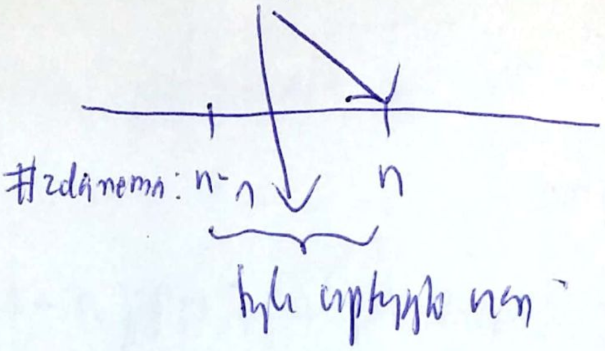
$$= \sum_{j\geq n} e^{-\lambda t}\frac{(\lambda t)^j}{j!}$$

<u>Generowanie procesu (jednorodnego) Poissona.</u>

Niech dany będzie proces (jednorodny) Poissona $(X_t)_{t\geq 0}$. Jak wiemy z przyrostów z nim stowarzyszonych jest ciąg zmiennych losowych $(T_n)_{n\geq 1}$, które są niezależne o rozkładzie λ - współczynnik intensywności $(X_t)_{t\geq 0}$, gdzie $T_n \in W(\lambda)$ oraz $T_n(\omega)$ określa czas jaki upłynął od zdarzenia n-1 nastąpiło n'te <u>zdarzenia n-1 nastąpiło n'te</u>.

Aby generować $(X_t)_{t \geq 0}$ wystarczy wygenerować T_n, co przedstawiamy poniżej:

$$\Omega \ni \omega \longrightarrow T_n(\omega)$$



$$X_{T_1(\omega)+\dots+T_n(\omega)}(\omega) = n$$

W tym celu generujemy n - liczb (pseudo)losowych $U_1(\omega), U_2(\omega), \dots, U_n(\omega)$. Ponieważ $U_j \in \mathcal{U}([0,1])$, $-\frac{1}{\lambda}\ln U_j \in W(\lambda)$.

Bierzemy $T_j(\omega) = -\frac{1}{\lambda}\ln U_j(\omega)$, $j = 1, 2, \dots, n$.

$\downarrow$

tyle czasu upłynie pomiędzy pojawieniem się zdarzenia $j-1$ oraz j.

Mamy więc algorytm dla $(X_t)_{t \ge 0}$, dla $t \le T$.

Krok 1: $t=0$, $\#(t)=0$ → liczba pojawiających się zdarzeń

Krok 2: generuj u (pseudo)losowy $U(0,1)$

Krok 3. $t = t - \frac{1}{\lambda} \ln U(u)$. Jeśli $t > T$, STOP

Krok 4. $\#(t) = \#(t) + 1$, $S(\#(t)) = t$

Krok 5 Idź do kroku 2.

Zadanie 4.

Zaimplementować powyższy algorytm ($\lambda = 2.0$)