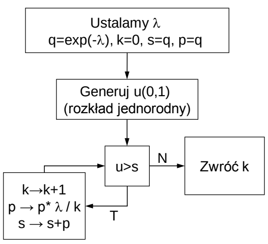
Rozkład Poissona – realizacja numeryczna

$$P_X(k)=\frac{\lambda^k}{k!}e^{-\lambda}$$



$$
\begin{aligned}
&dla\ x<0 && F_X(x)=0\\
&dla\ 0\le x<1 && F_X(x)=P(k=0)=e^{-\lambda}\\
&dla\ 1\le x<2 && F_X(x)=P(k=0)+P(k=1)=e^{-\lambda}+\lambda e^{-\lambda}\\
&dla\ 2\le x<3 && F_X(x)=e^{-\lambda}+\lambda e^{-\lambda}+\frac{1}{2}\lambda^2e^{-\lambda}\\
&dla\ 3\le x<4 && F_X(x)=e^{-\lambda}+\lambda e^{-\lambda}+\frac{1}{2}\lambda^2e^{-\lambda}+\frac{1}{6}\lambda^3e^{-\lambda}
\end{aligned}
$$

$$
\begin{aligned}
&dla\ k=0 && q=e^{-\lambda} \quad p=e^{-\lambda} \quad s=e^{-\lambda}\\
&dla\ k=1 && p=\frac{p\lambda}{k}=\lambda e^{-\lambda} \quad s=e^{-\lambda}+\lambda e^{-\lambda}\equiv F_X(1)\\
&dla\ k=2 && p=\frac{1}{2}\lambda^2e^{-\lambda} \quad s=e^{-\lambda}+\lambda e^{-\lambda}+\frac{1}{2}\lambda^2e^{-\lambda}\equiv F_X(2)\\
&dla\ k=3 && p=\frac{1}{2}\lambda^2e^{-\lambda}\frac{\lambda}{3}=\frac{1}{6}\lambda^3e^{-\lambda} \quad s=F_X(2)+\frac{1}{6}\lambda^3e^{-\lambda}\equiv F_X(3)
\end{aligned}
$$

Istnieją związki pomiędzy rozkładami;

- Przykład: **rozkład Poissona jako granica rozkładu dwumianowego** dla dużych n, małych p, i n·p=const= λ . Stąd wynika przydatność rozkładu Poissona do opisu „zjawisk rzadkich".

Generatory rozkładu jednorodnego

- **Rozkład jednorodny** na przedziale (a,b)
  E(X)=(b+a)/2 var(X)=(b-a)²/12
- Realizacja numeryczna:
- Najbezpieczniej używać sprawdzonych procedur (np. biblioteki CERN)
- 1) Metoda kongruencyjna – generator liniowy

$$x_{n+1}=(\underbrace{a_0x_n+a_1x_{n-1}+\ldots+a_kx_{n-k}}_{k+1\ poprzednich}+a_{k+1})(\bmod M) \qquad a_i,x_i\in[0,M)$$

K=0 – generatory multiplikatywne $x_{n+1}=(a_0x_n)(\bmod M)$

K=1 – generatory Fibonacciego $x_{n+1}=(a_0x_n+a_1x_{n-1})(\bmod M)$

Generatory mieszane $x_{n+1}=(a_0x_n+a_1)(\bmod M)$

(W języku C funkcja srand daje a_1= liczba sekund od 1.1.1970)

Np. a_0=69069 a_1=1 M=2^{32} (Marsaglia 1972)

Generatory rozkładu jednorodnego

- Przykład - generator mieszany:
  $x_n=(5x_{n-1}+1)(\bmod 16)$

Okres = 16

$x_1=x_{17}$=10

- Dane wejściowe: x_0=5:
- Kolejne otrzymywane liczby:
  10, 3, 0, 1, 6, 15, 12, 13, 2, 11, 8, 9, 14, 7, 4, 5 10, 3, 0, 1, 6, 15, 12, 13, 2, 11, 8, 9, 14, 7, 4, 5.
- Po podzieleniu przez 16:
- 0.6250, 0.1875, 0.0000, 0.0625, 0.3750, 0.9375, 0.7500, 0.8125, 0.1250, 0.6875, 0.5000, 0.5625, 0.8750, 0.4375, 0.2500, 0.3125, 0.6250, 0.1875, 0.0000, 0.0625, 0.3750, 0.9375, 0.7500, 0.8125, 0.1250, 0.6875, 0.5000, 0.5625, 0.8750, 0.4375, 0.2500, 0.3125.

Generatory rozkładu jednorodnego

- Uwagi do implementacji: musimy zachowywać odpowiednią precyzję i unikać przepełnienia

$$x_n=(7^5x_{n-1}+1)(\bmod 2^{31}-1)$$

- 2^{31}-1 = 2147483647 - liczba pierwsza
- 7^5 = 16807 jest jednym z 534600000 pierwotnych pierwiastków czyli w praktyce daje maksymalny okres.
- Iloczyn a*x_{n-1} może osiągnąć 16807× 2147483647≈ 1.03× 2^{45} → potrzeba 46-bitowych integer
- Pomocny może być wzór: (algorytm Schrage'a)

$$\boxed{ax \bmod m = g(x)+mh(x)}$$

$$g(x)\equiv a(x \bmod q)-r(x \operatorname{div} q)$$

$$h(x)\equiv (x \operatorname{div} q)-(ax \operatorname{div} m)$$

$$q\equiv m \operatorname{div} a$$

$$r=m \bmod a$$

- Zadania testowe: np. poprawna implementacja daje dla x_0=1: x_{10000}= 1043618065

Generatory rozkładu jednorodnego

- 2) Uogólnione generatory Fibonacciego
  (przypominają ciąg Fibonacciego $f_0=f_1=1$ $f_n=f_{n-2}+f_{n-1}$ dla n>1)

$$x_n=(a_0x_{n-2}+a_1x_{n-1})(\bmod M)\rightarrow$$

$$x_n=(a_0x_{n-r}+a_1x_{n-s})(\bmod M)\rightarrow$$

$$x_n=(a_0x_{n-r}\oplus a_1x_{n-s})(\bmod M)$$

gdzie $\oplus\equiv +,-,\cdot,xor,\ldots$ jest działaniem modulo M (nie może wyprowadzać poza zakres [0,M)

Inny zapis: $F(r,s,\oplus)$

Np.

F(17,5,+) dla M=2^{32} ma okres $(2^{17}-1)2^{31}$

F(17,5,*) dla M=2^{32} ma okres $(2^{17}-1)2^{29}$

Generatory rozkładu jednorodnego

- 3) Generatory oparte na mnożeniu z przeniesieniem
  Inicjujemy podając wartości $x_1,x_2,\ldots,x_k\in(0,M)$ i c=0 c-parametr przeniesienia.

$$x_n=(c+a_1x_{n-1}+\ldots+a_kx_{n-k})(\bmod M) \qquad a_i,x_i\in[0,M)$$

gdzie w kolejnych iteracjach

$$c\rightarrow Int[(c+a_1x_{n-1}+a_2x_{n-2}+\ldots+a_kx_{n-k})/M]$$

Np.

$$x_n=(c+1941x_{n-1}+1860x_{n-2}+1812x_{n-3}+1776x_{n-4}+1492x_{n-5}+$$
$$+1215x_{n-6}+1066x_{n-7}+12013x_{n-8}) \bmod 2^{16}$$

- 4) Generatory nieliniowe np. $x_n=(a_1(x_{n-1})^{-1}+c)(\bmod M)$

x_n nie zależy od poprzednich x_k → dobry do programowania równoległego

$$x_n=(a(n+b)+c)^{-1}(\bmod M)$$

$$x_n=(x_{n-1})^2(\bmod M)$$

Generatory rozkładu jednorodnego

- 5) Generatory oparte na rejestrach przesuwnych

$$x_n = (a_1 x_{n-1} + \ldots + a_k x_{n-k}) (\mathrm{mod}\, 2) \qquad a_i, x_i \in \{0,1\}$$

a	b	$a\ xor\ b$
0	0	0
1	0	1
0	1	1
1	1	0

Ponieważ (a+b) mod 2 = a *xor* b to

$$x_n = x_{n-j1}\ xor\ x_{n-j2}\ xor \ldots xor\ x_{n-jl} \qquad \begin{pmatrix} a_{jm} = 1\ dla\ m \le l \\ a_{jm} = 0\ dla\ m > l \end{pmatrix}$$

Tworzymy ciąg bitów x_n i z niego liczby pseudolosowe

$$u_i = 0.x_{i\cdot s+1} x_{i\cdot s+2} \ldots x_{i\cdot s+L} \quad dla\ s \le L\ oraz\ i = 0,1,\ldots$$

gdzie s i L to wybrane parametry.

Np. L=8,s=3

$$u_0 \overset{i\cdot s=0}{=} 0.x_1 x_2 \ldots x_8$$

$$u_1 \overset{i\cdot s=3}{=} 0.x_{3+1} x_{3+2} \ldots x_{3+8} = 0.x_4 x_5 \ldots x_{11}$$

$$u_2 \overset{i\cdot s=6}{=} 0.x_{6+1} x_{6+2} \ldots x_{6+8} = 0.x_7 x_8 \ldots x_{14}$$

Konstrukcja ta zwana jest również schematem Tausworha.

Generatory rozkładu jednorodnego

- Przykład: generator MZT (Marsaglia, Zaman, Tsang 1990), okres 2^{144}.
  Składa się z dwóch generatorów:
  1) F(97,33,⊕) o okresie 2^{120} generuje liczby $V_n \in [0,1)$

$$V_n = V_{n-97} \oplus V_{n-33} \qquad x \oplus y \equiv \begin{cases} x - y & dla \quad x \ge y \\ x - y + 1 & dla \quad x < y \end{cases}$$

Inicjalizacja polega na nadaniu wartości liczbom $V_1,\ldots,V_{97}$. Robimy to generując ich bity b_n generatorami liczb całkowitych

$$y_n = (y_{n-3} \cdot y_{n-2} \cdot y_{n-1}) \,\mathrm{mod}(179)$$

$$z_n = (52 z_{n-1} + 1) \,\mathrm{mod}(169)$$

$$b_n = \begin{cases} 0 & dla \quad (y_n \cdot z_n) \,\mathrm{mod}(64) < 32 \\ 1 & dla \quad (y_n \cdot z_n) \,\mathrm{mod}(64) \ge 32 \end{cases}$$

Czyli użytkownik podaje $y_1, y_2, y_3 \in \{1,2,\ldots,178\}$, $z_1 \in \{1,2,\ldots,168\}$

Generatory rozkładu jednorodnego

2) Generator liniowy o okresie 2^{24}-3 generuje liczby $c_n \in (0,1)$

$$C_n = C_{n-1} \oplus a, \quad n \ge 2$$

gdzie a=7654321/16777216, C_1=362436/16777216,

$$x \oplus y \equiv \begin{cases} x - y & dla \quad x \ge y \\ x - y + 16777213/16777216 & dla \quad x < y \end{cases}$$

Ostatecznie generator MZT zwraca liczbę

$$U_n = V_n \oplus C_n$$

gdzie ponownie

$$x \oplus y \equiv \begin{cases} x - y & dla \quad x \ge y \\ x - y + 1 & dla \quad x < y \end{cases}$$

Powyższe przykłady i więcej informacji: R.Wieczorkowski, R.Zieliński „Komputerowe generatory liczb losowych",WN-T 1997

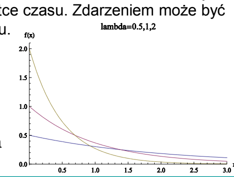
Rozkład wykładniczy (ekspotencjalny)

- Jest to rozwinięcie rozkładu geometrycznego na ciągłe zmienne losowe

$$f_X(x) = \lambda\, e^{-\lambda x} \qquad x \in [0, +\infty)$$

λ – parametr rozkładu, λ >0

Gdy x interpretujemy jako czas to rozkład wykładniczy opisuje prawdopodobieństwo zajścia zdarzenia, które zachodzi ze stałym prawdopodobieństwem w jednostce czasu. Zdarzeniem może być przejście układu do nowego stanu.



Zastosowania:
czas dostępu do serwera
promieniowanie kosmiczne
czas wezwania karetki pogotowia
czas obsługi klienta w banku

Rozkład wykładniczy -własności

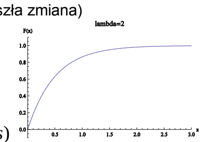
- Normalizacja

$$\int_0^{+\infty} \lambda\, e^{-\lambda x} dx = \lambda \int_0^{+\infty} e^{-\lambda x} dx = \lambda \frac{1}{(-\lambda)} e^{-\lambda x}\Big|_0^{+\infty} = (-1)\left(e^{-\lambda\infty} - e^{-\lambda 0}\right) = (-1)(0-1) = 1$$

- Dystrybuanta (prawdopodobieństwo, że zaszła zmiana)

$$F_X(x) = \int_0^x \lambda\, e^{-\lambda t} dt = (-1)\left(e^{-\lambda x} - e^{-\lambda 0}\right) = 1 - e^{-\lambda x}$$

$$\rightarrow \quad P_X(X > x) = 1 - F_X(x) = e^{-\lambda x}$$



- Posiada własność braku pamięci

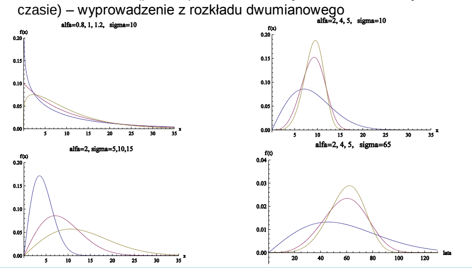
$$\forall s,t > 0: \quad P_X(X > t + s \mid X > t) = P_X(X > s)$$

(dowód analog. do rozkładu geometrycznego, zastosowanie: metoda datowania węglem ^{14}C)

- Wartość oczekiwana E(X)=1/ λ var(X)=1/(λ²)

Rozkład Weibulla

- Rozszerzenia rozkładu wykładniczego to
- **Rozkład Erlanga** (ile trzeba czekać na n-te zdarzenie)
- **Rozkład Weibulla** (prawdopodobieństwo zajścia zdarzenia nie jest stałe w czasie) – wyprowadzenie z rozkładu dwumianowego



Rozkład Weibulla

- Funkcja gęstości prawdopodobieństwa
  $f_X(x) = \alpha\sigma^{-\alpha} x^{\alpha-1} e^{-(x/\sigma)^\alpha} \qquad x \in [0,+\infty), \quad \alpha > 0, \sigma > 0$
- Parametr α decyduje o kształcie, σ o położeniu i wysokości (skali).
  dla α=1 - rozkład wykładniczy,
  dla α=2,3 - rozkład zbliżony do normalnego
- Dystrybuanta
  $F_X(x) = 1 - e^{-(x/\sigma)^\alpha}$
- Wartość oczekiwana i wariancja
  $E(X) = \sigma\Gamma(1+\alpha^{-1}) \qquad \mathrm{var}(X) = \sigma^2\left\{\Gamma(1+2\alpha^{-1}) - (\Gamma(1-\alpha^{-1}))^2\right\}$
- Funkcja Gamma Eulera

$$\Gamma(p) = \int_0^{+\infty} x^{p-1} e^{-x} dx \qquad \Gamma(1) = 1 \qquad \Gamma(0.5) = \sqrt{\pi}$$

$$\Gamma(p) = (p-1)\Gamma(p-1) \qquad \Gamma(n) = (n-1)! \quad n \in N$$

Rozkład Weibulla

- $\int_0^{\sigma} f_X(x)\,dx \approx 0.6321$

  $\longrightarrow$ 63,21% populacji umiera do czasu σ (niezależnie od α).
- Zastosowania: czas życia (ubezpieczenia), zużywanie się elementów w technice, czas dostawy produktu (logistyka), rozkład siły wiatru.
- Generowanie:
  1) Jeżeli zmienna losowa X ma rozkład wykładniczy o parametrze σ to zmienna losowa $Y=X^{(1/\alpha)}$ ma rozkład Weibulla $f_Y(y)$ o parametrach α i σ .
  2) Jeżeli zmienna losowa X ma rozkład jednorodny na przedziale (0,1) to zmienna losowa $Y=\sigma(-\ln(X))^{(1/\alpha)}$ ma rozkład Weibulla $f_Y(y)$ o parametrach α i σ .