Poprzednim razem

Rozkłady dyskretne

- Dwumianowy

$$P_X(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

- Geometryczny

$$P_X(k) = (1-p)^{k-1} p$$

- Poissona

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

1

Przykład: **rozkład Poissona jako granica rozkładu dwumianowego** dla dużych n, małych p, i n·p=const= λ

Rozkład dwumianowy: $B_k^n \equiv P_n(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$ $P_X(k) = \frac{\lambda^k}{k!} e^{-\lambda}$

dla k=0

$$B_0^n = \binom{n}{0} p^0 (1-p)^{n-0} = (1-p)^n = \left(1-\tfrac{\lambda}{n}\right)^n \quad \rightarrow \quad \lim_{n\to\infty} B_0^n = \lim_{n\to\infty} \left(1-\tfrac{\lambda}{n}\right)^n = e^{-\lambda} = P_X(X=0)$$

dla dowolnych *k* i *k+1*

$$\frac{B_{k+1}^n}{B_k^n} = \frac{\binom{n}{k+1} p^{k+1} (1-p)^{n-(k+1)}}{\binom{n}{k} p^k (1-p)^{n-k}} = \frac{n!k!(n-k)!}{(k+1)!(n-k-1)!n!} p(1-p)^{-1} = \frac{n-k}{k+1} \cdot \frac{p}{1-p} = \frac{np-pk}{(k+1)\cdot(1-p)} =$$

$$= \frac{\lambda - \frac{\lambda}{n} k}{(k+1)\cdot(1-\frac{\lambda}{n})} = \frac{\lambda(1-\frac{k}{n})}{(k+1)\cdot(1-\frac{\lambda}{n})} \quad \rightarrow \quad \lim_{n\to\infty} \frac{B_{k+1}^n}{B_k^n} = \lim_{n\to\infty} \frac{\lambda(1-\frac{k}{n})}{(k+1)\cdot(1-\frac{\lambda}{n})} = \frac{\lambda}{k+1}$$

stąd $\lim_{n\to\infty} \frac{B_1^n}{B_0^n} = \frac{\lim_{n\to\infty} B_1^n}{\lim_{n\to\infty} B_0^n} = \frac{\lambda}{1} = \lambda \quad \rightarrow \quad \lim_{n\to\infty} B_1^n = \lambda \lim_{n\to\infty} B_0^n = \lambda e^{-\lambda} = P_X(X=1)$

$$\lim_{n\to\infty} \frac{B_2^n}{B_1^n} = \frac{\lim_{n\to\infty} B_2^n}{\lim_{n\to\infty} B_1^n} = \frac{\lambda}{1+1} = \frac{\lambda}{2} \quad \rightarrow \quad \lim_{n\to\infty} B_2^n = \frac{\lambda}{2} \lim_{n\to\infty} B_1^n = \frac{\lambda^2 e^{-\lambda}}{2} = P_X(X=2)$$

$$\rightarrow \quad \lim_{n\to\infty} B_{k+1}^n = \frac{\lambda}{k+1} \lim_{n\to\infty} B_k^n = \frac{\lambda \cdot \lambda^k e^{-\lambda}}{(k+1)\cdot k!} = \frac{\lambda^{k+1} e^{-\lambda}}{(k+1)!} = P_X(X=k+1)$$

opis „zjawisk rzadkich”

2

Zmienna losowa ciągła

- Rozkład jednorodny

$$f_X(x) = \begin{cases} \frac{1}{b-a} & dla\ x \in [a,b] \\ 0 & dla\ x \notin [a,b] \end{cases}$$

3

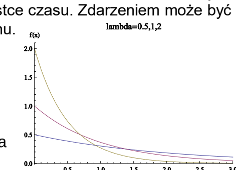
Rozkład wykładniczy (eksponencjalny)

- 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

4

Rozkład wykładniczy -własności

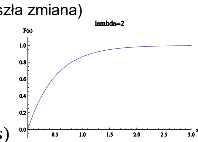
- Normalizacja

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

- 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/(λ^2)

5

Rozkład wykładniczy – związek z próbą Bernoulliego

Prawdopodobieństwo braku sukcesu w kolejnych n próbach (= same porażki)

$$Q_n = \underbrace{(1-p)\cdot(1-p)\cdot\ldots\cdot(1-p)}_{n\ razy} = (1-p)^n$$

Zał: proces zachodzi ze stałą częstością w czasie tzn. E(X)~t (ozn: E(X)=λt)
Wiemy, że w n-elementowej próbie Bernoulliego E(X)=np.

$$np = \lambda t \quad \rightarrow \quad p = \tfrac{\lambda t}{n}$$

$$Q_n = (1-p)^n = (1-\tfrac{\lambda t}{n})^n \quad \rightarrow \quad \lim_{n\to\infty} Q_n = \lim_{n\to\infty} (1-\tfrac{\lambda t}{n})^n = e^{-\lambda t}$$

Prawdopodobieństwo wystąpienia co najmniej jednego sukcesu do chwili t

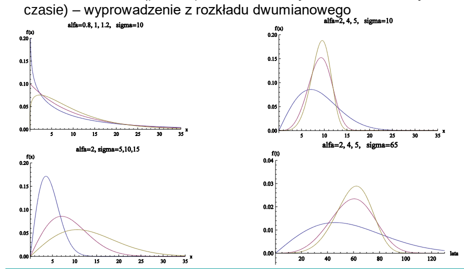
$$F(t) = 1 - Q_n = 1 - e^{-\lambda t}$$

$$\rightarrow \quad f_T(t) = \frac{dF(t)}{dt} = -(-\lambda)e^{-\lambda t} = \lambda e^{-\lambda t}$$

6

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



7

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

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

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

8

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 .

9

Rozkład normalny N(μ,σ^2)

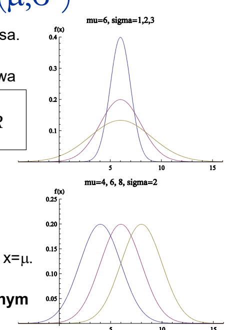
- Zwany jest również rozkładem Gaussa.
- Jest to najważniejszy z rozkładów.
- Funkcja gęstości prawdopodobieństwa

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{\frac{-(x-\mu)^2}{2\sigma^2}} \qquad x\in R$$

- Rozkład ten jest unormowany:

$$\int_{-\infty}^{+\infty}\frac{1}{\sigma\sqrt{2\pi}}e^{\frac{-(x-\mu)^2}{2\sigma^2}}dx = 1$$

- Rozkład jest symetryczny względem x=μ.
- Dla μ=0 i σ=1 nazywany jest **standardowym rozkładem normalnym** (standardowym rozkładem Gaussa).



10

Dystrybuanta rozkładu normalnego N(0,1)

- Dystrybuanta rozkładu normalnego
  Nie potrafimy policzyć analitycznie. Można tego uniknąć licząc ją numerycznie, korzystając z tablic dla N(0,1) lub korzystając z przybliżonych wzorów dla N(0,1)

$$F_X^{(1)}(x) = \frac{1}{2}\pm\frac{1}{2}\sqrt{1-e^{\frac{-2x^2}{\pi}}}$$

$$F_X^{(2)}(x) = \frac{1}{2}\pm\frac{1}{2}\sqrt{1-e^{\frac{-2x^2}{\pi}}-\frac{2(\pi-3)}{3\pi^2}x^4e^{\frac{-x^2}{2}}}$$

  gdzie górny znak odpowiada x>0 a dolny x≤0.

- Średnia arytmetyczna $F_X^{(3)}(x) \equiv (F_X^{(1)}(x)+F_X^{(2)}(x))/2$
  przybliża prawdziwą dystrybuantę z dokładnością 0.1%.
- Wzory te mogą posłużyć za punkt wyjścia do metody odwracania dystrybuanty (numerycznego) i napisaniu generatora liczb pseudolosowych o rozkładzie N(0,1)
- Funkcja błędu Gaussa bywa zaimplementowana w niektórych kompilatorach (Java)

$$erf(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}dt \qquad \text{wtedy} \qquad F_X(x) = 0.5\left(1+erf(x/\sqrt{2})\right)$$

11

E(X) i var(X) rozkładu normalnego N(μ,σ^2)

- Wartość oczekiwana rozkładu normalnego $N(\mu,\sigma^2):\ f_X(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{\frac{-(x-\mu)^2}{2\sigma^2}}$

$$E(X) = \int_{-\infty}^{+\infty} f_X(x)x\,dx = \int_{-\infty}^{+\infty}\frac{1}{\sigma\sqrt{2\pi}}e^{\frac{-(x-\mu)^2}{2\sigma^2}}x\,dx = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{\frac{-(x-\mu)^2}{2\sigma^2}}x\,dx = \left|\begin{matrix}\frac{x-\mu}{\sigma}\equiv u\\ x=\sigma u+\mu\\ dx=\sigma\,du\end{matrix}\right| =$$

$$= \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{\frac{-u^2}{2}}(\sigma u+\mu)\sigma\,du = \frac{\sigma}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{\frac{-u^2}{2}}u\,du + \frac{\mu}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{\frac{-u^2}{2}}\,du = \frac{\sigma}{\sqrt{2\pi}}\cdot 0+\mu\cdot 1 = \mu$$

- Wariancja rozkładu normalnego

$$\mathrm{var}(X) = E\left((X-E(X))^2\right) = \int_{-\infty}^{+\infty} f_X(x)(x-\mu)^2\,dx = \int_{-\infty}^{+\infty}\frac{1}{\sigma\sqrt{2\pi}}e^{\frac{-(x-\mu)^2}{2\sigma^2}}\frac{(x-\mu)^2}{\sigma^2}\sigma^2\,dx = \left|\begin{matrix}\frac{x-\mu}{\sigma}\equiv u\\ x=\sigma u+\mu\\ dx=\sigma\,du\end{matrix}\right| =$$

$$= \frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{\frac{-u^2}{2}}u^2\,du = \left|\int_{-\infty}^{+\infty}e^{-a^2t^2}t^2\,dt = \frac{\sqrt{\pi}}{2a^3}\right| = \frac{\sigma^2}{\sqrt{2\pi}}\cdot\frac{\sqrt{\pi}}{2\left(\frac{1}{\sqrt{2}}\right)^3} = \sigma^2$$

- Odchylenie standardowe rozkładu normalnego $\sigma(X) = \sqrt{\mathrm{var}(X)} = \sqrt{\sigma^2} = \sigma$

12

Rozkład normalny -własności

- Rolę rozkładu normalnego podkreśla Centralne Twierdzenie Graniczne.
- Zastosowanie: tam gdzie szereg czynników ma wpływ na wielkość X i wpływają na nią w mniej więcej jednakowym stopniu. Jest szeroko stosowany w naukach społecznych, ekonomicznych, biologicznych (dla X lub ln(X)), np. wzrost, ciśnienie krwi, inteligencja, wielkość rocznych opadów deszczu jak również przy szacowaniu niepewności przypadkowych, testach statystycznych. Często jest używany w sytuacjach gdy nie znamy prawdziwego rozkładu. Nie mniej rzadko występuje naprawdę.
- Generowanie, przykładowe metody:

  1) Suma zmiennych o rozkładzie jednorodnym

  2) Transformacja Box-Mullera

  3) Metoda biegunowa (Marsaglia), v_1 i v_2 mają rozkład N(0,1)

  $$x, y \in (-1,1) \rightarrow z \equiv x^2 + y^2$$

  $$z < 1 \;\Rightarrow\; v_1 \equiv x\sqrt{-2z^{-1}ln(z)} \quad i \quad v_2 \equiv y\sqrt{-2z^{-1}ln(z)}$$

Rozkład normalny -własności

- Wartość oczekiwana i wariancja: $\boxed{E(X) = \mu \qquad var(X) = \sigma^2}$
- **Standaryzacja dowolnego rozkładu** pozwala przechodzić pomiędzy rozkładami normalnymi:

  Jeżeli zmienna losowa X ma rozkład N(μ,σ^2) to zmienna losowa Y=(X-μ)/σ ma rozkład N(0,1) (dowód na ćwiczeniach).