

Poprzednim razem

Rozkłady dyskretne

Dwumianowy

$$P_X(k) = \binom{n}{k} p^k (1-p)^{n-k} \quad k=\{0,1,2,\dots,n\}$$

Geometryczny

$$P_X(k) = (1-p)^{k-1} p \quad k=\{1,2,3,\dots\}$$

Poissona

$$P_X(k) = \frac{\lambda^k}{k!} e^{-\lambda} \quad k=\{0,1,2,\dots\}$$

1

Zmienna losowa ciągła

Rozkład jednorodny

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

2

Rozkład wykładniczy (eksponencjalny)

- Jest to rozwinięcie rozkładu geometrycznego na ciągle zmienne losowe

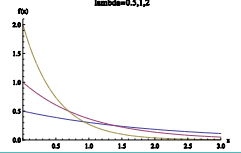
$$f_X(x) = \lambda e^{-\lambda x} \quad x \in [0, +\infty)$$

λ – parametr rozkładu, $\lambda > 0$

Gdy x interpretujemy jako czas to rozkład wykładniczy opisuje prawdopodobieństwo tego, że na zajście zdarzenia będziemy czekać czas x . Zakładamy, że zdarzenia 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



3

Rozkład wykładniczy -własności

Normalizacja

$$\int_0^{+\infty} \lambda e^{-\lambda x} dx = \lambda \int_0^{+\infty} e^{-\lambda x} dx = \lambda \left[-\frac{1}{\lambda} e^{-\lambda x} \right]_0^{+\infty} = (-1)(e^{-\lambda \cdot \infty} - e^{-\lambda \cdot 0}) = (-1)(0 - 1) = 1$$

- Dystrybucja (prawdopodobieństwo, że zaszła zmiana)

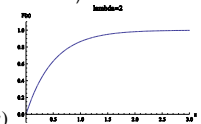
$$F_X(x) = \int_0^x \lambda e^{-\lambda t} dt = (-1)(e^{-\lambda x} - e^{-\lambda \cdot 0}) = 1 - e^{-\lambda x}$$

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

- Posiada własność braku pamięci

$\forall s, t > 0: P_X(X > t+s | X > t) = P_X(X > s)$
(dowód analog. do rozkładu geometrycznego, zastosowanie: metoda datowania węglem ^{14}C)

- Wartość oczekiwana $E(X) = 1/\lambda$ $\text{var}(X) = 1/(\lambda^2)$



4

Rozkład wykładniczy – związek z próbą Bernoulliego

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

$$Q_n = (1-p) \cdot (1-p) \cdot \dots \cdot (1-p) = (1-p)^n$$

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

$$np = \lambda t \rightarrow p = \frac{\lambda t}{n}$$

$$Q_n = (1-p)^n = \left(1 - \frac{\lambda t}{n}\right)^n \rightarrow \lim_{n \rightarrow \infty} Q_n = \lim_{n \rightarrow \infty} \left(1 - \frac{\lambda t}{n}\right)^n = e^{-\lambda t}$$

Prawdopodobieństwo wystąpienia co najmniej jednego sukcesu do chwili t

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

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

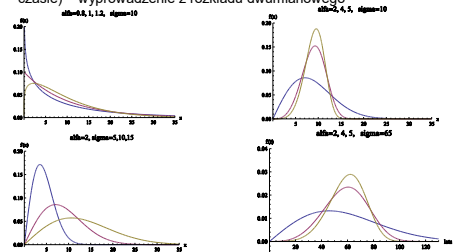
5

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



6

Rozkład Weibulla

- Funkcja gęstości prawdopodobieństwa

$$f_X(x) = \alpha \sigma^{-\alpha} x^{\alpha-1} e^{-(x/\sigma)^\alpha} \quad x \in [0, \infty), \quad \alpha > 0, \sigma > 0$$
- Parametr α decyduje o kształcie, σ o położeniu i wysokości (skali).
 dla $\alpha=1$ - rozkład wykładniczy,
 dla $\alpha=2,3$ - rozkład zbliżony do normalnego
- Dystrybucja

$$F_X(x) = 1 - e^{-(x/\sigma)^\alpha}$$
- Wartość oczekiwana i wariancja

$$E(X) = \sigma \Gamma(1 + \alpha^{-1}) \quad \text{var}(X) = \sigma^2 \{ \Gamma(1 + 2\alpha^{-1}) - (\Gamma(1 + \alpha^{-1}))^2 \}$$
- Funkcja Gamma Eulera

$$\Gamma(p) = \int_0^\infty x^{p-1} e^{-x} dx \quad \Gamma(1) = 1$$

$$\Gamma(p) = (p-1)\Gamma(p-1) \quad \Gamma(0.5) = \sqrt{\pi}$$

$$\Gamma(n) = (n-1)! \quad n \in \mathbb{N}$$

7

Rozkład Weibulla

- $\int_0^\sigma f_X(x) dx \approx 0.6321$
 $\rightarrow$ 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:
 - 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 .
 - 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 .

8

Rozkład normalny $N(\mu, \sigma^2)$

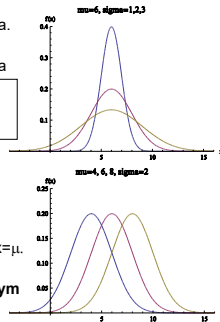
- Zwany jest również rozkładem Gaussa.
- Jest to najważniejszy z rozkładów.
- Funkcja gęstości prawdopodobieństwa

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

- Rozkład ten jest unimodalny:

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

- Rozkład jest symetryczny względem $x=\mu$.
- Dla $\mu=0$ i $\sigma=1$ nazywany jest **standardowym rozkładem normalnym** (standardowym rozkładem Gaussa).



9

Dystrybucja rozkładu normalnego $N(0,1)$

- Dystrybucja 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 $N(0,1)$

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

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

gdzie górny znak odpowiada $x>0$ a dolny $x\leq 0$.

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

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

10

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:
 - Suma zmiennych o rozkładzie jednorodnym
 - Transformacja Box-Mullera
 - Metoda biegunowa (Marsaglia), v_1 i v_2 mają rozkład $N(0,1)$

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

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

11

Rozkład normalny -własności

- Wartość oczekiwana i wariancja: $E(X) = \mu \quad \text{var}(X) = \sigma^2$

- Standaryzacja dowolnego rozkładu** pozwala przechodzić pomiędzy rozkładami normalnymi:
 Jeżeli zmienna losowa X ma rozkład $N(\mu, \sigma^2)$ to zmienna losowa $Y = (X-\mu)/\sigma$ ma rozkład $N(0,1)$ (dowód na ćwiczeniach).

- „Reguła trzech sigma” :

k	$P(\mu - k\sigma < X < \mu + k\sigma)$	k	Surprisal #bits $S = \log_2(1/P)$ dla $P=0.5$ (rzut monetą)
1	0.6826895	0.50	0.6745
2	0.9544997	0.80	1.2816
3	0.9973002	0.90	1.6449
4	0.9999367	0.95	1.96
5	0.9999994	0.99	2.5758
		1	1.66
		3	8.53
		5	20.67

12

Rozkład normalny -własności

- Szerokość poławkowa FWHM (full width at half maximum)

$$f_X(x_{1/2}) = 0.5 \cdot f_X^{\max} = 0.5 \cdot f_X(\mu) = \frac{1}{2\sigma\sqrt{2\pi}}$$

$$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(\mu-x_{1/2})^2}{2\sigma^2}} = \frac{1}{2\sigma\sqrt{2\pi}}$$

$$x_{1/2}^{(\pm)} = \mu \pm \sigma\sqrt{2\ln(2)}$$

$$FWHM \equiv \Gamma = x_{1/2}^{(+)} - x_{1/2}^{(-)} =$$

$$= 2\sigma\sqrt{2\ln(2)} \approx 2.355\sigma$$

$$\rightarrow \sigma = \frac{\Gamma}{2.355} = 0.425\Gamma$$

