Poprzednim razem

- Transformacje $P_X(x)$ i $f_X(x)$
- Próba Bernoullego
- Rozkład dwumianowy (dwumienny, Bernoullego)

  X - liczba sukcesów w n próbach Bernoullego;
  $S_X=\{0,1,\ldots,n\}$
  p – prawdopodobieństwo sukcesu w pojedynczej próbie

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

  Wartość oczekiwana w rozkładzie dwumianowym *E(X)=np*.
  Wariancja w rozkładzie dwumianowym *var(X)=np(1-p)*

Rozkład geometryczny

- Zmienną losową jest ilość prób potrzebna do uzyskania sukcesu w nieskończenie długiej próbie Bernoulliego.
  Zbiór wartości $S_X=\{1,2,3,\ldots\}$
  Rozkład prawdopodobieństwa zmiennej X, przyjmującej wartości $k\in S_X$ jest postaci

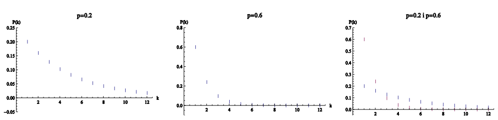
  $$P_X(k)\equiv P(X=k)=P(\{\underbrace{\bar{A}\bar{A}\bar{A}\ldots\bar{A}}_{k-1\ razy}A\})=(1-p)^{k-1}p\equiv q^{k-1}p$$

  gdzie A oznacza uzyskanie sukcesu w pojedynczej próbie.
- Normalizacja:

  $$\sum_{k=1}^{\infty}P_X(k)=\sum_{k=1}^{\infty}q^{k-1}p=p\sum_{k=1}^{\infty}q^{k-1}=p\underbrace{(1+q+q^2+\ldots)}_{|q|<1}=\frac{p}{1-q}=\frac{p}{p}=1$$

- Wartość oczekiwana w rozkładzie geometrycznym $E(X)=p^{-1}$
- Wariancja w rozkładzie geometrycznym $var(X)=(1-p)p^{-2}$

Rozkład geometryczny



- Dystrybuanta (dla uproszczenia zapisu tylko w argumentach całkowitych)

  $$F_X(n)=\sum_{k=1}^{n}P_X(k)=p\underbrace{(1+q+q^2+\ldots+q^{n-1})}_{|q|<1}=p\frac{1-q^n}{1-q}=\frac{p(1-q^n)}{p}=1-q^n$$

  czyli

  $$P(X>n)=1-F_X(n)=1-(1-q^n)=q^n$$

Rozkład geometryczny

- Zmienna losowa X o rozkładzie geometrycznym posiada własność „braku pamięci"

  $$\forall k,j\in\{1,2,\ldots\}:\quad P_X(X>k+j\,|\,X>j)=P(X>k)$$

  czyli wcześniejsze wyniki nie wpływają na następne.
  Dowód:

  $$P(X>k+j\,|\,X>j)=\frac{P(\{X>k+j\}\cap\{X>j\})}{P(X>j)}=\frac{P(X>k+j)}{P(X>j)}=$$

  $$=\frac{1-F_X(k+j)}{1-F_X(j)}=\frac{1-(1-q^{k+j})}{1-(1-q^j)}=\frac{q^{k+j}}{q^j}=q^k=1-(1-q^k)=1-F_X(k)=P(X>k)$$

  Jest to jedyny dyskretny rozkład prawdopodobieństwa o tej własności.

- Istnieje również możliwość zdefiniowania rozkładu geometrycznego jako rozkładu zmiennej – liczby rzutów przed osiągnięciem pierwszego sukcesu. Wtedy $S_X=\{0,1,2,\ldots\}$ i $P_X(k)=(1-p)^k p$.

Rozkład Poissona

- Opisuje prawdopodobieństwo zajścia k sukcesów w określonym przedziale czasu, jeżeli znamy średnią częstotliwość wystąpienia sukcesu, a czas oczekiwania na sukces podlega rozkładowi wykładniczemu.

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

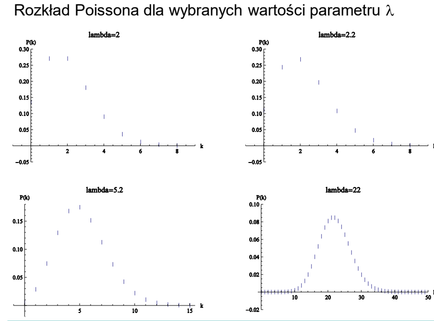
  gdzie λ to parametr rozkładu, a $k=\{0,1,2,\ldots\}$

  Wartość oczekiwana w rozkładzie Poissona $E(X)=\lambda$
  Wariancja w rozkładzie Poissona $var(X)=\lambda$

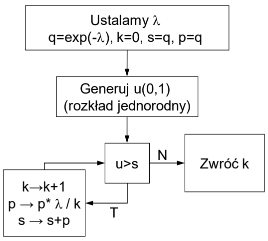
  Zastosowania: liczba ludzi pochodzących do kasy w supermarkecie w określonym przedziale czasu (np.2 minut), liczba wejść na stronę internetową w określonym przedziale czasu, liczba rozpadów jader atomowych, liczba mutacji DNA, liczba wypadków lotniczych w ciągu roku, liczba sprzedanych sztuk drogiego towaru, (duża próbka i rzadkie zjawisko)

Rozkład Poissona

Rozkład Poissona dla wybranych wartości parametru λ



Rozkład Poissona – realizacja numeryczna



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

$dla\ x < 0 \quad F_X(x) = 0$

$dla\ 0 \le x < 1 \quad F_X(x) = P(k=0) = e^{-\lambda}$

$dla\ 1 \le x < 2 \quad F_X(x) = P(k=0) + P(k=1) = e^{-\lambda} + \lambda e^{-\lambda}$

$dla\ 2 \le x < 3 \quad F_X(x) = e^{-\lambda} + \lambda e^{-\lambda} + \frac{1}{2}\lambda^2 e^{-\lambda}$

$dla\ 3 \le x < 4 \quad F_X(x) = e^{-\lambda} + \lambda e^{-\lambda} + \frac{1}{2}\lambda^2 e^{-\lambda} + \frac{1}{6}\lambda^3 e^{-\lambda}$

$dla\ k = 0 \quad q = e^{-\lambda} \quad p = e^{-\lambda} \quad s = e^{-\lambda}$

$dla\ k = 1 \quad p = \frac{p\lambda}{k} = \lambda e^{-\lambda} \quad s = e^{-\lambda} + \lambda e^{-\lambda} \equiv F_X(1)$

$dla\ k = 2 \quad p = \frac{1}{2}\lambda^2 e^{-\lambda} \quad s = e^{-\lambda} + \lambda e^{-\lambda} + \frac{1}{2}\lambda^2 e^{-\lambda} \equiv F_X(2)$

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

Istnieją związki pomiędzy rozkładami;

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

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

dla k=0

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

dla dowolnych *k* i *k+1*

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

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

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

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

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

opis „zjawisk rzadkich”

Zmienna losowa ciągła

- Rozkład jednorodny

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

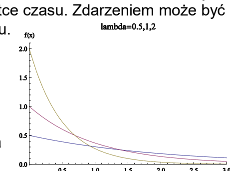
Rozkład wykładniczy (eksponencjalny)

- 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

Rozkład wykładniczy -własności

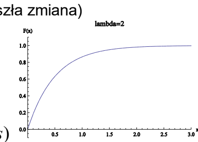
- Normalizacja

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

- Dystrybuanta (prawdopodobieństwo, że zaszła zmiana)

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

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



- 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 wykładniczy – związek z próbą Bernoulliego

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

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

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.

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

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

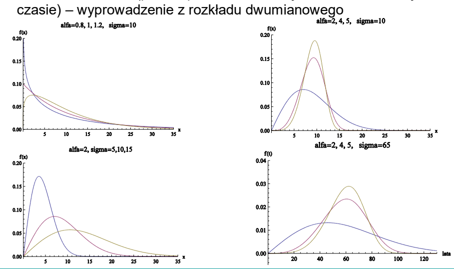
Prawdopodobieństwo wystąpienia co najmniej jednego sukcesu do chwili t

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

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

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
-

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

$\longrightarrow$ 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.