

Dyskretne rozkłady prawdopodobieństwa

- Rozkład zero-jedynkowy (rozkład Bernoulliego)** o parametrze p
Zmienna losowa to ilość sukcesów w jednokrotnym powtórzeniu eksperymentu w którym możliwe są tylko dwa wyniki
 $S_x = \{0, 1\}$
Rozkład prawdopodobieństwa $P(x) = \begin{cases} 1-p & \text{dla } x=0 \\ p & \text{dla } x=1 \end{cases}$
- Rozkład dwupunktowy** o parametrze p
Zmienna losowa przyjmuje dwie dowolne wartości, jedną z prawdopodobieństwem p , drugą z prawdopodobieństwem $1-p$.
- Rozkład wielopunktowy** o parametrach $p_i, i=1, 2, \dots, n$
przy czym $p_1 + p_2 + \dots + p_n = 1$.
Zmienna losowa przyjmuje skończoną (n) ilość dyskretnych wartości.
- Rozkład wielopunktowy** o parametrach $p_i, i=1, 2, \dots$
przy czym $p_1 + p_2 + \dots = 1$.
Zmienna losowa przyjmuje nieskończoną ilość dyskretnych wartości

1

Próba Bernoulliego i rozkład dwumienny

- Próba Bernoulliego** to sekwencja powtórzeń n razy **tego samego** eksperymentu losowego, w wyniku którego możemy uzyskać jeden spośród dwóch wyników, zwanych jako „sukces” i „porażka”.
Muszą być spełnione dwa warunki:
1. Powtórzenia są niezależne
2. Prawdopodobieństwo sukcesu jest takie samo we wszystkich powtórzeniach.

Przykład: n -krotny rzut monetą.

- Rozkład dwumienny (dwumianowy)** (zwany w Polsce **rozkładem Bernoulliego**)
Zmienna losowa X oznacza liczbę osiągniętych sukcesów w n -elementowej próbie Bernoulliego.
 $S_x = \{0, 1, 2, \dots, n\}$

Rozkład prawdopodobieństwa zmiennej X , przyjmującej wartości $k \in S_x$ jest postaci

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

2

Rozkład dwumienny

Uzasadnienie:

Niech zdarzenie P oznacza osiągnięcie sukcesu w pojedynczym eksperymencie.

Niech zdarzenie P oznacza osiągnięcie porażki w pojedynczym eksperymencie.

Wynik próby (zdarzenie A) opisujemy jako n -elementowy ciąg wyników pojedynczych eksperymentów

np. $A = \{S, P, S, S, \dots, P, P, S\}$

Prawdopodobieństwo uzyskania pojedynczego ciągu w którym wystąpiło k sukcesów to $P(A) = p^k (1-p)^{n-k}$ (niezależność prób)

Ilość ciągów sprzyjających wystąpieniu k sukcesów to $\binom{n}{k}$ (kombinacja bez powtórzeń)

Zatem (rozłączność ciągów)

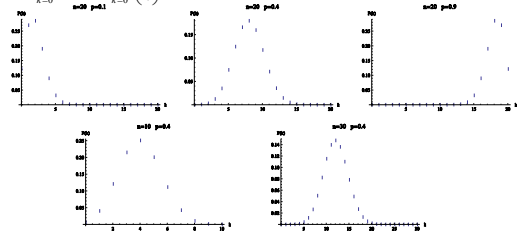
$$P_x(k) = \binom{n}{k} P(A) = \binom{n}{k} p^k (1-p)^{n-k}$$

3

Rozkład dwumienny

- Rozkład ten jest poprawnie unormowany:

$$\sum_{k=0}^n P_x(k) = \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} = (p + (1-p))^n = 1^n = 1$$



4

Rozkład dwumienny

- Wartość oczekiwana w rozkładzie dwumianowym $E(X) = np$.
- Wariancja w rozkładzie dwumianowym $var(X) = np(1-p)$
- Dystrybucja $F_x(x) = \sum_{k=0}^x P_x(k)$ (dystrybucja jest określona na zbiorze ciągłym).
- Moda – zależy od n i p , mogą to być dwie wartości k (jest tak dla $(n+1)p$ całkowitego, wtedy moda to $k=(n+1)p$ i $k=(n+1)p-1$).
- Rozkład wielomianowy** – wiele możliwych wyników pojedynczego eksperymentu, zachodzących z dowolnymi prawdopodobieństwami. Zachowujemy stałość prawdopodobieństw w kolejnych próbach i niezależność prób. Opisujemy $P(k_1, k_2, \dots)$.

5

Rozkład geometryczny

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

$$P_x(k) \equiv P(X=k) = P(\overline{A} \overline{A} \dots \overline{A} A) = (1-p)^{k-1} p \equiv q^{k-1} p$$

gdzie A oznacza uzyskanie ^{$k-1$ razy} sukcesu w pojedynczej próbie.

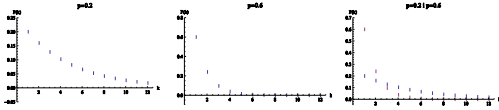
- Normalizacja:

$$\sum_{k=1}^{\infty} P_x(k) = \sum_{k=1}^{\infty} q^{k-1} p = p \sum_{k=1}^{\infty} q^{k-1} = p(1 + q + q^2 + \dots) = \frac{p}{1-q} = \frac{p}{p} = 1$$

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

6

Rozkład geometryczny



- Dystrybucja (dla uproszczenia zapisu tylko w argumentach całkowitych)

$$F_X(n) = \sum_{k=1}^n P_X(k) = p(1+q+q^2+\dots+q^{n-1}) = p \frac{1-q^n}{1-q} = \frac{p(1-q^n)}{p} = 1-q^n$$

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”

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

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 $S_X = \{0, 1, 2, \dots\}$ i $P_X(k) = (1-p)^k p$.

7

8

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 $k = \{0, 1, 2, \dots\}$

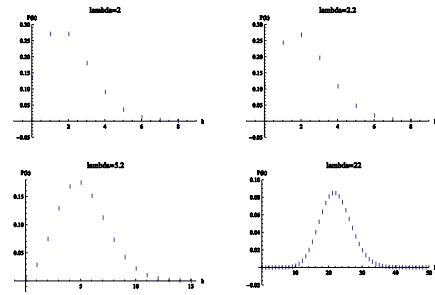
Wartość oczekiwana w rozkładzie Poissona $E(X) = \lambda$

Wariancja w rozkładzie Poissona $\text{var}(X) = \lambda$

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 jąder atomowych, liczba mutacji DNA, liczba wypadków lotniczych w ciągu roku, liczba sprzedanych sztuk drogiego towaru, (duża próbka i rzadkie zjawisko)

Rozkład Poissona

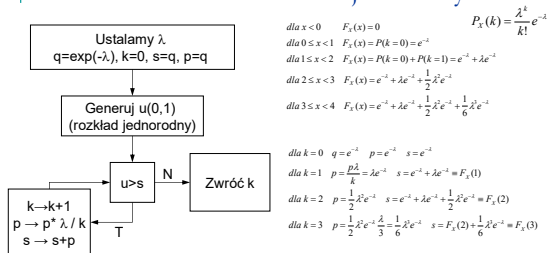
Rozkład Poissona dla wybranych wartości parametru λ .



9

10

Rozkład Poissona – realizacja numeryczna



Przykład: rozkład Poissona jako granica rozkładu dwumianowego dla dużych n, małych p, i $n \cdot p = \text{const} = \lambda$

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 $k=0$ $B_0^n = \binom{n}{0} p^0 (1-p)^n = (1-p)^n = (1 - \frac{\lambda}{n})^n \rightarrow \lim_{n \rightarrow \infty} B_0^n = \lim_{n \rightarrow \infty} (1 - \frac{\lambda}{n})^n = e^{-\lambda} = P_X(X=0)$

dla dowolnych k i $k+1$

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

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

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

$$\lim_{n \rightarrow \infty} \frac{B_k^n}{B_{k-1}^n} = \frac{\lambda}{k} \rightarrow \lim_{n \rightarrow \infty} B_k^n = \frac{\lambda}{k} \lim_{n \rightarrow \infty} B_{k-1}^n = \frac{\lambda^2 e^{-\lambda}}{k} = P_X(X=2)$$

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

opis „zjawisk rzadkich”

11

12