Dyskretne zmienne losowe

dr Mariusz Grządziel

Katedra Matematyki, Uniwersytet Przyrodniczy we Wrocławiu

Wykład 5; r. akad. 2022/2023

Definicja 1

Zmienną losową nazywamy dyskretną (skokową), jeśli zbiór jej wartości $x_1, x_2, \ldots,$ można ustawić w ciąg.

Zmienna losowa X, która przyjmuje wszystkie wartości z zadanego przedziału (a, b), nie jest zmienną losową dyskretną, ponieważ elementów tego przedziału nie da się ustawić w ciąg ("ponumerować").
- G. Cantor 1873— twierdzenie— wszystkich liczb rzeczywistych nie da się ustawić w ciąg.

Rozkład dyskretnej zmiennej losowej

Zbiór wartości dyskretnej zmiennej losowej X— ciąg $x_1, x_2, \ldots,$ (skończony lub nieskończony).
Rozkład zmiennej losowej dyskretnej X jest określony przez nieujemne liczby $p_1, p_2, \ldots$ spełniające warunki:

$$\sum p_i = 1, \tag{1}$$

$$p_i = P(X = x_i). \tag{2}$$

Dyskretne zmienne losowe— przykłady

Przykłady:

- Rozkład U, sumy oczek w dwukrotnym rzucie kostką (patrz poprzedni wykład);
- Rozkład Z, gdzie Z oznacza liczbę rzutów monetą, po której po raz pierwszy wypada orzeł (zdarzeniu polegającemu na tym, że orzeł wypadnie już w pierwszym rzucie, odpowiada wartość zmiennej Z równa 0).

z niezależności zdarzeń:

$$P(Z = k) = \left(\frac{1}{2}\right)^{k+1}, \quad k = 0, 1, 2, \ldots .$$

Zmienna losowa Z przykład dyskretnej zmiennej losowej, dla której zbiór wartości: $\{0, 1, 2, \ldots\}$ nie jest skończony.

Rzuty osobiste— przykład

Niech X- liczba trafień w wykonywanym przez koszykarza A rzucie osobistym.Niech: T odpowiada trafieniu do kosza, C odpowiada chybieniu.
Przestrzeń zdarzeń elementarnych: $\mathcal{S} = \{C, T\}$.
Niech X- liczba trafionych rzutów. Zmienna X jest funkcją określoną na $\mathcal{S}$;

$$X(C) = 0, \quad X(T) = 1.$$

Zakładamy, że prawdopodobieństwo trafienia wynosi 0,9.
Rozkład zmiennej losowej X można przedstawić przy pomocy tabelki:

k	0	1
$P(X = k)$	0,1	0,9

Liczba trafień Y w dwóch rzutach

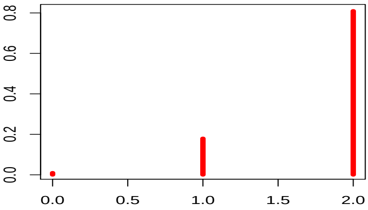
Niech Y- liczba trafień w dwóch wykonywanych przez koszykarza A rzutach osobistych.
Przyjmujemy, że prawdopodobieństwo trafienia w jednym rzucie osobistym wynosi 0,9 i zdarzenie trafienia/chybienia w drugim rzucie jest niezależne od analogicznego zdarzenia w pierwszym rzucie.
Można pokazać, że:

$$P(Y = 0) = \left(\frac{1}{10}\right)^2 = 0{,}01,$$
$$P(Y = 1) = 2 \times \frac{1}{10} \times \frac{9}{10} = 0{,}18,$$
$$P(Y = 2) = \left(\frac{9}{10}\right)^2 = 0{,}81.$$

Liczba trafień Y w dwóch rzutach— c.d.

Rozkład można przedstawić w postaci tabelki lub wykresu słupkowego:

k	0	1	2
$P(X = k)$	0,01	0,18	0,81



Rozkład dwumianowy

Symbol Newtona $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ jest równy liczbie podzbiorów k-elementowych zbioru n-elementowego ($0 \leqslant k \leqslant n$).

Definicja 2

Mówimy, że zmienna losowa X ma rozkład dwumianowy z parametrami $n \in \mathbb{N}$ i $0 < p < 1$, co w skrótowo zapisujemy $X \sim Bin(n, p)$ (lub $X \sim Bin(n; p)$), jeśli

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, 2, \ldots, n.$$

„Dziesięciokrotny rzut monetą”— przykład

Niech V oznacza liczbę orłów otrzymanych w dziesięciokrotnym rzucie monetą (zakładamy, że moneta jest „rzetelna”, tj. prawdopodobieństwo otrzymania orła jest równe $\frac{1}{2}$ oraz że wyniki kolejnych rzutów są od siebie niezależne). Chcemy obliczyć prawdopodobieństwo;

$$P(V \geqslant 9).$$

Rozwiązanie $V \sim Bin(10; 0{,}5)$,

$$\begin{aligned} P(V \geqslant 9) =& P(V = 9) + P(V = 10) = \\ =& \binom{10}{9} (0{,}5)^9 (0{,}5)^1 + \binom{10}{10} (0{,}5)^{10} (0{,}5)^0 = \\ =& \frac{10}{1024} + \frac{1}{1024} = \frac{11}{1024}. \end{aligned}$$

Pojęcie dystrybuanty rozkładu

W obliczeniach podobnych do tych z poprzedniego przykładu użyteczne może się okazać pojęcie dystrybuanty zmiennej losowej.

Definicja 3

Niech X będzie dowolną zmienną losową. Dystrybuantą zmiennej losowej X nazywamy funkcję F określoną jako:

$$F(x) = P(X \leqslant x).$$

Uwaga. W powyższej definicji nie zakładamy, że zmienna X jest dyskretna.

„Dziesięciokrotny rzut monetą"—c.d.

V-liczba wyrzuconych orłów w dziesięciokrotnym rzucie monetą;

$$P(V \geqslant 9) = P(V = 9) + P(V = 10) = F_V(10) - F_V(8)$$

gdzie F_V jest dystrybuantą zmiennej losowej V.
Obliczenia wykonane w R-rze:

```
> pbinom(10,10,0.5)- pbinom(8,10,0.5)
[1] 0.01074219
```

pbinom- pierwsza litera odpowiada „dystrybuancie", binom odpowiada rodzajowi rozkładu (ang. *binomial*- dwumianowy). Korzystając z polecenia *pbinom* można obliczać wartości dystrybuanty rozkładu $Bin(n, p)$ dla dużych wartości n.

Wartość oczekiwana zmiennej losowej dyskretnej

Definicja 4

Dla zmiennej losowej dyskretnej X wartość oczekiwana, jeśli istnieje, jest liczbą określoną wzorem

$$EX = \sum_i x_i p_i,$$

w którym sumowanie obejmuje wszystkie wartości zmiennej X.

Uwaga Wartość oczekiwana zmiennej losowej nazywana jest literaturze także wartością średnią zmiennej losowej. W definicji tej zakładamy, że $\sum_i x_i p_i$ jest liczbą skończoną (w przypadku, gdy liczba składników jest nieskończona, zakładamy zbieżność sumy do granicy skończonej).
Wartość oczekiwana może być interpretowana jako "środek ciężkości" układu punktów materialnych $x_1, x_2, \ldots$ o wagach $p_1, p_2, \ldots$.

Wariancja zmiennej losowej

Definicja 5

Wariancję zmiennej losowej X określamy wzorem

$$VarX = E(X-\mu)^2,$$

gdzie $\mu = EX$.

Wariancja jest równa wartości oczekiwanej kwadratu odchylenia wartości zmiennej losowej od swojej wartości przeciętnej.

Uwaga: W definicji tej nie zakładamy, że zmienna losowa X jest dyskretna. Zakładamy natomiast istnienie wartości oczekiwanej $E(X-\mu)^2$.

Odchylenie standardowe

Odchylenie standardowe zmiennej losowej X, oznaczane przez DX, definiujemy jako pierwiastek kwadratowy wariancji X. Odchylenie standardowe zmiennej losowej X czesto jest też oznaczane grecką literą σ.
Można pokazać, że dla $a \geqslant 0$:

$$D(aX + b) = aDX. \tag{3}$$

Niezależność zmiennych losowych

Definicja 6

Mówimy, że zmienne losowe X i Y są niezależne, jeżeli

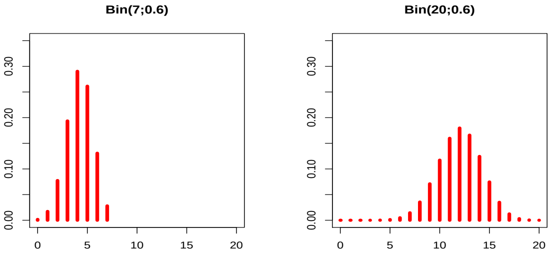
$$P(X \in [a, b] \wedge Y \in [c, d]) = P(X \in [a, b]) \times P(Y \in [c, d])$$

dla dowolnych przedziałów $[a, b]$ *i* $[c, d]$.

Intuicyjnie: niezależne zmienne losowe odpowiadają realizacje liczbowe niezależnych zmiennych losowych.

Wartość oczekiwane i wariancje rozkładu dwumianowego

Można pokazać, że wartość oczekiwana zmiennej losowej X, $X \sim Bin(n, p)$ wynosi np, a wariancja tej zmiennej jest równa $Var(Y) = np(1 - p)$.



Rysunek: Wykresy słupkowe przedstawiające rozkłady $Bin(7; 0{,}6)$ oraz $Bin(20; 0{,}6)$

Wariancja $X_1 \sim Bin(7; 0{,}6)$: $7 \times 0{,}6 \times 0{,}4 = 1{,}68$
Wariancja $X_2 \sim Bin(20; 0{,}6)$: $20 \times 0{,}6 \times 0{,}4 = 4{,}8$
X_2 jest bardziej „rozproszona"!

Lektura uzupełniająca

T. Bednarski, Elementy matematyki w naukach ekonomicznych. Oficyna ekonomiczna. Kraków 2004, str. 228–234.

Koronacki, J., Mielniczuk, J. Statystyka dla studentów kierunków technicznych i przyrodniczych. WNT. Warszawa 2001, podrozdział 2.2.1, str. 94–110.