Seria 5. Zmienne losowe dyskretne

1. Definicja zmiennej losowej

Definicja 1 *Niech* $(\Omega, \mathcal{F}, \mathbf{P})$ *będzie przestrzenią probabilistyczną. Dowolne odwzorowanie mierzalne* $X : \Omega \to \mathbb{R}^n$ *przyjmujące jedynie przeliczalnie wiele wiele wartości nazywamy zmienną losową dyskretną. Wystarczy zakładać, że dla każdego punktu* $t \in \mathbb{R}$, $X^{-1}(\{t\}) \in \mathcal{F}$. *Zauważmy ponadto, że istnieje zbiór przeliczalny* N *taki, że* $\mathbf{P}(X \in N) = 1$.

Definicja 2 *Rozkładem prawdopodobieństwa na* $\mathbb{R}^n$ *nazywamy dowolną miarę probabilistyczną* μ *(to znaczy* $\mu(\mathbb{R}^n) = 1$*) na* $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$.

Definicja 3 *Rozkładem prawdopodobieństwa zmiennej losowej dyskretnej* X *o wartościach w zbiorze przeliczalnym* N *nazywamy miarę probabilistyczną* μ_X *zadaną wzorem*

$$\mu_X(\{t\}) = \mathbf{P}(X = t), \ \ \forall\, t \in N.$$

2. Dystrybuanta

Definicja 4 *Dystrybuantą rozkładu* μ *na* $\mathbb{R}$ *nazywamy funkcję* $F : \mathbb{R} \to \mathbb{R}$ *zadaną wzorem* $F(t) := \mu((-\infty, t])$. *Jeśli rozkład wyznaczony jest przez zmienną losową* X *dystrybuantę nazywamy* F_X.

Zad 1 *Pokaż, że dystrybuanta* F_X *jest ciągła w punkcie* t *wtedy i tylko wtedy, gdy* $\mathbf{P}_X(\{t\}) = 0$.

Zad 2 *Wykonujemy doświadczenia Bernouliego, aż do chwili otrzymania pierwszego sukcesu. Niech* X *oznacza liczbę wykonanych doświadczeń. Wyznacz rozkład* X, *wyznacz dystrybuantę* F_X.

Zad 3 *Rzucamy dwiema kostkami. Niech* X *będzie większą z liczb uzyskanych na poszczególnych kostkach. Obliczyć* $\mathbf{P}(X \in [2, 4])$.

Zad 4 *Udowodnij, że* F_X *ma własności*

- F_X *jest niemalejące, to znaczy jeśli* $s \leqslant t$, *to* $F_X(s) \leqslant F_X(t)$;
- $\lim_{t\to\infty} F_X(t) = 1$, $\lim_{t\to-\infty} F_X(t) = 0$;
- *Pokaż, że* F_X *jest prawostronnie ciągła. To znaczy* $\lim_{s\to t_+} F_X(s) = F(t)$.

Zad 5 *Pokaż, że istnieją dwie różne zmienne losowe* X, X', *które mają ten sam rozkład. To znaczy* $\mu_X = \mu_{X'}$

Zad 6 *Udowodnij, że dwie zmienne o identycznych dystrybuantach mają identyczne rozkłady.*

Zad 7 *Wykaż, że rozkład geometryczny ma własność braku pamięci to znaczy*

$$\mathbf{P}(X > t + s | X > t) = \mathbf{P}(X > s), \ \ \forall\, s, t \in \mathbb{N}.$$

3. Charakterystyki zmiennych

Definicja 5 *Wartością oczekiwaną zmiennej losowej dyskretnej* X *nazywamy*

$$\mathbb{E}X = \sum_{x \in N} x\mu_X(\{x\}) = \sum_{x\in N} x\mathbf{P}(X = x),$$

gdzie N *jest przeliczalnym zbiorem wartości zmiennej* X.

Definicja 6 *Wariancją zmiennej losowej dyskretnej* X *nazywamy*

$$\mathbf{V}arX = \mathbf{D}^2X = \sum_{x\in N} (x - \mathbb{E}X)^2 \mu_X(\{x\}) = \sum_{x\in N} (x - \mathbb{E}X)^2 \mathbf{P}(X = x).$$

Definicja 7 *Zmienne dyskretne* $X_1, ..., X_n$ *określone na* $(\Omega, \mathcal{F}, \mathbf{P})$ *nazywamy niezależnymi jeśli*

$$\mathbf{P}(X = x_1, ..., X = x_n) = \mathbf{P}(X = x_1)...\mathbf{P}(X = x_n), \ \ \forall\, x_1 \in N_1, x_2 \in N_2, ..., x_n \in N_n.$$

Twierdzenie 1 *Dla dowolnych zmiennych dyskretnych* $X_1, ..., X_n$, $\mathbb{E}(X_1 + ... + X_n) = \mathbb{E}X_1 + ...\mathbb{E}X_n$. *Jeśli zmienne dyskretne* $X_1, ..., X_n$ *są niezależne, to* $\mathbf{D}^2(X_1 + ... + X_n) = \mathbf{D}^2X_1 + ... + \mathbf{D}^2X_n$.

Zad 8 *Przykład, że Twierdzenie 1 przydaje się. Niech* X *oznacza sumę oczek uzyskaną w stukrotnym rzucie kostką. Pokaz, że* $\mathbb{E}X = 350$.

Zad 9 *Niech* X *ma rozkład jednopunktowy* δ_a *zadany wzorem* $\mathbf{P}(X = a) = 1$. *Pokaż, że* $\mathbb{E}X = a$, $\mathbf{D}^2X = 0$.

Zad 10 *Niech* X *ma rozkład Bernouliego* $\mathcal{B}(n, p)$ *zadany wzorem* $\mathbf{P}(X = k)\binom{n}{k}p^kq^{n-k}$, $p + q = 1$. *Pokaz, że* $\mathbb{E}X = np$, $\mathbf{D}^2X = npq$.

Zad 11 *Wykonujemy doświadczenia Bernouliego, aż do chwili otrzymania pierwszego sukcesu. Niech* X *oznacza liczbę wykonanych doświadczeń. Pokaż, że* $\mathbb{E}X = 1/p$, $\mathbf{D}^2X = q/p^2$.

Zad 12 *Przypuśćmy, że zmienna losowa ma rozkład Poissona* $\mathcal{P}(\lambda)$ *zadany wzorem* $\mathbf{P}(X = k) = \frac{\lambda^k}{k!}e^{-\lambda}$, $k = 1, 2, .., \lambda > 0$. *Pokaż, że* $\mathbb{E}X = \lambda$, $\mathbf{D}^2X = \lambda$.

4. Wzór Poissona

Twierdzenie 2 *Jeśli* $n \to \infty$, $p_n \to 0$, $np_n \to \lambda > 0$, *to*

$$lim_{n\to\infty}\binom{n}{k}p_n^k(1-p)n)^{n-k} = \frac{\lambda^k}{k!}e^{-\lambda}.$$

Na ile dobre jest oszacowanie Poissona

Twierdzenie 3 *Niech* $X_1, X_2, ..., X_n$ *będą niezależnymi zmiennymi losowymi o tym samym rozkładzie:* $\mathbf{P}(X_i = 1) = p$, $\mathbf{P}(X_i = 0) = q$ $p + q = 1$, $i \in 1, 2, ..., n$. *Oznaczmy* $\lambda = np$, $\pi_k = \frac{\lambda^k}{k!}e^{-\lambda}$ *dla* $k = 0, 1, 2, ...$, $S_n = X_1 + ... + X_n$. *Dla każdego zbioru borelowskiego* $B \subset \mathbb{R}$ *mamy*

$$|\mathbf{P}(S_n \in B) - \sum_{k\in B\cap\mathbb{N}} \pi_k| \leqslant \frac{\lambda^2}{n}.$$

Wynika stąd, że przybliżenie Poissona jest dobre jeśli $np_n^2 \to 0$.

Zad 13 *P-stwo* p *trafienia szóstki w Lotto jest równe* $1/\binom{49}{6} \simeq 7\cdot10^{-8}$. *Ilu szóstek można się spodziewać w każdym tygodniu, jeśli grający wypełniają kupony całkowicie losowy i niezależnie od siebie, a kuponów jest* $n = 10^7$.

Zad 14 *Tekst broszury zawiera* $n = 100000$ *znaków. W trakcie pisania każdy znak może zostać błędnie wprowadzony z p-stwem* 0.0001. *Z kolei redaktor znajduje każdy z błędów z p-stwem* 0.9 *po czym tekst wraca do autora, który znajduje każdy z pozostałych błędów z p-stwem* 0.5. *Jaka jest szansa, że po obu korektach broszura będzie zawierała nie więcej niż* 3 *błędy?*

Zad 15 *Ile średnio rodzynków powinno zawierać ciastko, żeby z p-stwem* 0.99 *dane ciastko zawierło przynajmniej jeden rodzynek?*

Zad 16 *W Warszawie na Ursynowie ginie średnio* 7 *samochodów tygodniowo. Jaka jest szansa, że jutro będzie dzień bez kradzieży, przy założeniu stałej intensywności złodziei.*