Rachunek Prawdopodobieństwa

Wydział Elektroniki

WYKŁAD III

1. Zmienne losowe

Zauważmy, że często z wynikiem doświadczenia losowego związana jest pewna liczba rzeczywista albo ciąg takich liczb. Pokazują to niektóre przykłady, które dotychczas analizowaliśmy. Wykonując wielokrotne rzuty monetą symetryczną, pytaliśmy na przykład ile razy w ustalonej liczbie rzutów pojawi się orzeł albo w którym rzucie po raz pierwszy wypadnie orzeł. Te *losowe* wartości liczbowe są funkcjami określonymi na przestrzeni Ω, które będziemy nazywać *zmiennymi losowymi.*

1.1. **Definicja i przykłady.** Niech $(\Omega, \mathcal{F}, \mathbb{P})$ będzie przestrzenią probabilistyczną.

Definicja. Funkcję $X : \Omega \to \mathbb{R}$, która dla każdego przedziału $(a, b]$ spełnia warunek

$$\{\omega \in \Omega : X(\omega) \in (a, b]\} \in \mathcal{F},$$

nazywamy **zmienną losową** o wartościach w $\mathbb{R}$.

Funkcja $\mathbb{P}$ nie pojawiła się bezpośrednio w definicji zmiennej losowej. Najczęściej jednak będziemy pytać właśnie o prawdopodobieństwo tego, że wartości zmiennej losowej należą do jakiegoś podzbioru $\mathbb{R}$. Przypomnijmy, że prawdopodobieństwo $\mathbb{P}$ jest określone tylko dla zbiorów z sigma-ciała $\mathcal{F}$. Stąd właśnie warunek występujący w definicji zmiennej losowej.

Będziemy używać skróconego zapisu

$$\mathbb{P}(X \in A) = \mathbb{P}(\{\omega \in \Omega : X(\omega) \in A\}).$$

Przykład. Przykłady zmiennych losowych.

- Waga urodzeniowa dziecka: wartościami są liczby z przedziału $(0, \infty)$;
- Płeć dziecka: wartościami mogą być liczby 0 - dziewczynka, 1 - chłopiec;
- Liczba oczek przy jednokrotnym rzucie kostką: wartościami są liczby ze zbioru $\{1, 2, 3, 4, 5, 6\}$;
- Czas oczekiwania na autobus na przystanku;
- Liczba osób urodzonych z wadą genetyczną w danej populacji.

1.2. **Rozkład i dystrybuanta zmiennej losowej.** Zmienna losowa umożliwia naturalny *transport* prawdopodobieństwa $\mathbb{P}$ na prostą rzeczywistą $\mathbb{R}$.

Przypomnijmy, że **sigma-ciałem zbiorów borelowskich** $\mathcal{B}(\mathbb{R})$ nazywamy najmniejszą rodzinę podzbiorów zbioru liczb rzeczywistych $\mathbb{R}$, posiadającą własności z definicji sigma-ciała $\mathcal{F}$, która zawiera wszystkie przedziały.

Definicja. Rozkładem prawdopodobieństwa (lub krótko: **rozkładem**) zmiennej losowej X o wartościach rzeczywistych nazywamy funkcję zbioru μ_X określoną na $\mathcal{B}(\mathbb{R})$ wzorem

$$\mu_X(A) \stackrel{def}{=} \mathbb{P}(X \in A), \quad A \in \mathcal{B}(\mathbb{R}).$$

Uwaga. Funkcja zbioru μ_X posiada takie same własności jak $\mathbb{P}$, tzn. jest unormowana $\mu_X(\mathbb{R}) = 1$ oraz przeliczalnie addytywna: dla dowolnych zbiorów $A_1, A_2, A_3, \ldots$ z rodziny $\mathcal{B}(\mathbb{R})$, parami rozłącznych, zachodzi równość

$$\mu_X\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu_X(A_i).$$

Oznacza to, że μ_X także jest prawdopodobieństwem, zaś trójka $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu_X)$ jest nową przestrzenią probabilistyczną.

Ten *transport prawdopodobieństwa* przynosi wymierną korzyść. Zamiast analizować wyjściową, często bardzo abstrakcyjną, przestrzeń probabilistyczną $(\Omega, \mathcal{F}, \mathbb{P})$, możemy pracować w konkretnej przestrzeni $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu_X)$. Rozkład zmiennej X zawiera pełną informację o zbiorze jej wartości i o tym, jak rozłożone jest prawdopodobieństwo na tym zbiorze. Wszystko to jest także zakodowane w dystrybuancie tego rozkładu, która jest obiektem znacznie bardziej przyjaznym.

Definicja. Dystrybuantą rozkładu μ_X (zmiennej losowej X) nazywamy funkcję $F_X : \mathbb{R} \to \mathbb{R}$ daną wzorem

$$F_X(t) \stackrel{def}{=} P(X \le t) = \mu_X((-\infty, t]), \quad t \in \mathbb{R}.$$

Podstawowe własności dystrybuanty są opisane w nastepującym twierdzeniu.

Twierdzenie 1. *Dystrybuanta F_X rozkładu zmiennej X ma następujące własności:*

(1) *F_X jest funkcją niemalejącą,*
(2) *F_X jest prawostronnie ciągła,*
(3) *$\lim_{t\to-\infty} F_X(t) = 0$ oraz $\lim_{t\to\infty} F_X(t) = 1$.*

Ponadto, każda funkcja $F : \mathbb{R} \to \mathbb{R}$ spełniająca trzy powyższe warunki jest dystrybuantą pewnego rozkładu. Jeżeli dwie dystrybuanty są sobie równe, to odpowiadają temu samemu rozkładowi.

Inne ważne własności dystrybuanty:

- $\mathbb{P}(X \le t) = F_X(t)$ oraz $\mathbb{P}(X > t) = 1 - F_X(t)$;
- $\mathbb{P}(X < t) = \lim_{s\to t-} F_X(s)$ oraz $\mathbb{P}(X \ge t) = 1 - \lim_{s\to t-} F_X(s)$;
- $\mathbb{P}(s < X \le t) = F_X(t) - F_X(s)$ itp.

Przykład. Gracz rzuca kostką do gry. Gdy wyrzuci 5 oczek, wygrywa 10 złotych, gdy wyrzuci liczbę oczek podzielną przez 3, wygrywa 5 złotych. W pozostałych przypadkach traci 1 złoty. Niech X oznacza wygraną gracza. Wyznacz dystrybuantę X (ten przykład omówimy dokładnie podczas naszego videospotkania).

2. Rozkłady dyskretne

2.1. **Definicja i podstawowe własności.** Bardzo ważnym typem zmiennych losowych są zmienne *dyskretne*, których rozkłady skupione są na zbiorach przeliczalnych. Przypomnijmy, że zbiór $S \subset \mathbb{R}$ nazywamy przeliczalnym, jeżeli jego elementy można ustawić w ciąg (być może skończony).

Definicja. Rozkład μ_X nazywamy **dyskretnym**, jeżeli istnieje zbiór przeliczalny $S \subset \mathbb{R}$ taki, że $\mu_X(S) = 1$.

Uwaga. Zmienną X nazwiemy dyskretną, jeżeli μ_X jest dyskretny.

Jeżeli X ma rozkład dyskretny, to istnieje ciąg liczbowy $(x_n)_{n\in T} \subset \mathbb{R}$ (skończony lub nie) taki, że

$$\mathbb{P}(X = x_n) =: p_n > 0, \qquad \sum_{n\in T} p_n = 1.$$

Jeśli zmienna X przyjmuje nieskończenie wiele wartości, to $T = \mathbb{N}_+$. Jeżeli zbiór wartości X jest skończony, to $T = \{1, 2, \dots, N\}$, gdzie N jest liczbą tych wartości.

Zauważmy, że rozkład X jest jednoznacznie wyznaczony przez ciąg wartości $(x_n)_{n\in T}$ oraz ciąg prawdopodobieństw $(p_n)_{n\in T}$, z jakimi te wartości są przyjmowane.

Uwaga. Wyznaczenie $\mu_X(A) = \mathbb{P}(X \in A)$ polega na identyfikacji wszystkich tych spośród liczb x_n, które należą do zbioru A i obliczeniu sumy odpowiadających im prawdopodobieńtw p_n, tzn.

$$\mu_X(A) = \mathbb{P}(X \in A) = \sum_{\{n\in T:\, x_n\in A\}} p_n.$$

Uwaga. Dystrybuanta rozkładu dyskretnego jest funkcją schodkową, która w punktach $x_1, x_2, x_3, \dots$ ma skoki wysokości odpowiednio $p_1, p_2, p_3, \dots$.

2.2. **Najważniejsze rozkłady dyskretne na $\mathbb{R}$.** Omówimy teraz cztery podstawowe przykłady rozkładów dyskretnych.

(1) Pojedyncza próba Bernoulliego i rozkład dwupunktowy. Wyobraźmy sobie doświadczenie losowe, w którym możliwe są dwa różne wyniki. Nazwijmy je umownie *sukcesem* i *porażką*. Przypuśćmy, że sukces pojawia się w naszym doświadczeniu z prawdopodobieństwem $p \in (0, 1)$. Wówczas prawdopodobieństwo porażki to $q = 1 - p \in (0, 2)$. Takie doświadczenie losowe nazywane jest **próbą Bernoulliego**.

Przykładami takich doświadczeń są:

- rzut monetą: sukces = wypadnięcie orła, porażka = wypadnięcie reszki, $p = 1/2$;
- rzut kostką: sukces = wypadnięcie szóstki, porażka = niewypadnięcie szóstki, $p = 1/6$;
- gra w totolotka: sukces = trafienie szóstki, porażka = nietrafienie szóstki, $p = 1/13983816$.

Z wynikiem tego doświadczenia możemy powiązać bezpośrednio wartości liczbowe: np. sukces to 1, a porażka to 0. Prowadzi to wprost do określenia zmiennej losowej X o **rozkładzie dwupunktowym**:

$$\mu_X(\{1\}) = \mathbb{P}(X = 1) = p, \quad \mu_X(\{0\}) = \mathbb{P}(X = 0) = q = 1 - p.$$

(2) Schemat Bernoulliego i rozkład dwumianowy. Przypuśćmy, że powtarzamy doświadczenie losowe z punktu (1) n-krotnie, gdzie $n \in \mathbb{N}_+$, w sposób niezależny, przy czym w każdej pojedynczej próbie prawdopodobieństwo sukcesu wynosi niezmiennie $p \in (0, 1)$ (a porażki $q = 1 - p$). Takie doświadczenie losowe nazywane jest **schematem Bernoulliego**.

Rozważając schemat Bernoulliego możemy określać różne zdarzenia losowe i pytać o ich prawdopodobieństwa. Podstawowe pytanie brzmi tak: *jakie jest prawdopodobieństwo, że k spośród takich n prób zakończy się sukcesem?* Oczywiście nasze k może być tu jedną z liczb $0, 1, \ldots, n$.

Wynikami naszego doświadczenia są n-elementowe ciągi składające się z *sukcesów* i *porażek*. Nasze pytanie dotyczy zaś tylko takich ciągów, w których *sukces* występuje k-krotnie, a *porażka* $n - k$ krotnie. Ponieważ wszystkie próby wykonujemy w sposób niezależny, prawdopodobieństwo uzyskania każdego takiego ustalonego ciągu wynosi $p^k(1 - p)^{n-k}$, a wszystkich takich ciągów mamy $\binom{n}{k}$ (wybieramy k spośród n miejsc, na których występuje sukces). Wobec tego szukane prawdopodobieństwo wynosi:

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

To rozważanie prowadzi nas do określenia nowej zmiennej losowej X, która opisuje liczbę sukcesów w n próbach Bernoulliego. Dokładniej, mamy

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

Taki rozkład nazywany jest **rozkładem dwumianowym** z parametrami $p \in (0, 1)$ i $n \in \mathbb{N}_+$, i jest oznaczany przez $B(n, p)$ (ang. binomial distribution). Zauważmy, że rozkład dwupunktowy jest szczególnym przypadkiem rozkładu dwumianowego, gdzie $n = 1$ (ozn. $B(1, p)$).

Przykład. Co jest bardziej prawdopodobne: wygrać z równorzędnym przeciwnikiem 3 partie z 4, czy 5 z 8?

Przykład. Wśród kłębków pewnej partii bawełny znajduje się 30% kolorowych. Oblicz prawdopodobieństwo, że wśród 10 losowo wybranych kłębków znajdują się co najwyżej 3 kolorowe.

(3) Nieskończony ciąg prób Bernoulliego i rozkład geometryczny. Pomyślmy, że wykonujemy kolejne próby Bernoulliego (z prawdopodobieństwem sukcesu równym $p \in (0, 1)$).

Możemy zapytać: *jakie jest prawdopodobieństwo, że sukces pojawi się po raz perwszy w próbie o numerze k?* Jest jasne, że k może być tutaj dowolną liczbą naturalną dodatnią.

Rozumując podobnie jak poprzednio, możemy stwierdzić, że nasze pytanie dotyczy ciągu długości k, który na $k - 1$ pierwszych pozycjach ma same *porażki* i ma dokładnie jeden *sukces* na pozycji k. Prawdopodobieństwo jego wystąpienia jest więc równe

$$(1 - p)^{k-1} p.$$

Pozwala nam to na określenie zmiennej losowej X, która opisuje numer próby w ciągu prób Bernoulliego, w której po raz pierwszy pojawi się sukces (albo, innymi słowy, liczbę prób w schemacie Bernoulliego, które trzeba wykonać, aby pojawił się pierwszy sukces). Dokładniej,

$$\mu_X(\{k\}) = \mathbb{P}(X = k) = (1 - p)^{k-1} p, \quad k = 1, 2, 3, \ldots.$$

Taki rozkład nazywany jest **rozkładem geometrycznym** z parametrem $p \in (0, 1)$. Zauważmy, że jest on skupiony na zbiorze liczb naturalnych dodatnich, a więc na zbiorze przeliczalnym nieskończonym (inaczej niż w dwóch poprzednich przykładach).

Przykład. Przykład pojawił się już w Rozdziale 2.3 (c) Wykładu I oraz w zad. 8 na liście 1.

(4) Rozkład Poissona. Drugim ważnym przykładem zmiennej dyskretnej X, która przyjmuje wartości ze zbioru przeliczalnego nieskończonego jest zmienna losowa o **rozkładzie Poissona** z parametrem $\lambda > 0$, tzn.

$$\mu_X(\{k\}) = \mathbb{P}(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}, \quad k = 0, 1, 2, \ldots.$$

Rozkład ten oznaczamy przez $Poiss(\lambda)$. Warto w tym miejscu zwrócić uwagę na bezpośredni związek tego rozkładu z reprezentacją szeregową funkcji wykładniczej

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}.$$

Następujące twierdzenie pokazuje, że przy pewnych założeniach rozkład Poissona jest rozkładem granicznym dla ciągu rozkładów dwumianowych. Wykorzystuje się go powszechnie do modelowania tzw. zjawisk „rzadkich" – wypadków drogowych, awarii, pożarów, wygranych w Lotto itp.

Twierdzenie 2 (Poisson)**.** *Jeżeli $n \to \infty$, $p_n \to 0$ i jednocześnie $np_n \to \lambda > 0$, to dla każdego ustalonego $k = 0, 1, 2, \ldots$ mamy*

$$\binom{n}{k} p_n^k (1 - p_n)^{n-k} \longrightarrow e^{-\lambda} \frac{\lambda^k}{k!}.$$

Jeśli X_n oznacza liczbę sukcesów w n próbach Bernoulliego z prawodpodobienstwem sukcesu $p_n \in (0, 1)$, a Y jest zmienną losową o rozkładzie Poissona z parametrem $\lambda > 0$, to powyższe twierdzenie orzeka, że

$$\lim_{n\to\infty} \mathbb{P}(X_n = k) = \mathbb{P}(Y = k), \quad k = 0, 1, 2, \ldots,$$

o ile $p_n \to 0$ oraz $np_n \to \lambda > 0$, gdy $n \to \infty$. Zdarzenia $\{X_n = k\}$ oraz $\{Y = k\}$ mogą zostać tutaj zastąpione przez $\{X_n \in A\}$ i $\{Y \in A\}$, dla dowolnego zbioru borelowskiego $A \subset \mathbb{R}$. Powyższa zbieżność ma znaczenie praktyczne: oznacza, że rozkład Poissona może być przybliżany rozkładami dwumianowymi.

Twierdzenie Poissona ma bardzo elementarny dowód, który wykorzystuje definicję liczby e poznaną na kursie z analizy. Można je jednak także uzyskać jako wniosek z jeszcze ogólniejszego twierdzenia, które podamy poniżej.

Twierdzenie 3. *Jeśli X_n oznacza liczbę sukcesów w n próbach Bernoulliego z prawodpodobienstwem sukcesu $p \in (0, 1)$, a Y_n jest zmienną losową o rozkładzie Poissona z parametrem $\lambda_n = np$, to dla każdego zbioru borelowskiego $A \subset \mathbb{R}$ zachodzi nierówność*

$$|\mathbb{P}(X_n \in A) - \mathbb{P}(Y_n \in A)| \leq np^2.$$

Powyższe twierdzenie ma duże znaczenie praktyczne: nie tylko implikuje twierdzenie Poissona (a więc pozwala przybliżać rozkład Poissona rozkładami dwumianowym), ale, co ważniejsze, pozwala także przybliżać rozkład dwumianowy $B(n, p)$ rozkładem Poissona z parametrem $\lambda_n = np$. Stwierdza, że błąd takiego przybliżenia jest nie większy niż

$$np^2 = \frac{\lambda_n^2}{n} = p\lambda_n.$$

Takie przybliżenie może być zatem efektywnie wykorzystane w przypadku, gdy liczba np^2 jest mała (jak w przykładzie poniżej).

Przykład. Niech $X \sim B(n, p)$, gdzie $n = 100$ oraz $p = 0,01$. Oblicz $\mathbb{P}(X > 2)$ i porównaj z przybliżeniem Poissona.

Mamy

$$\begin{aligned} \mathbb{P}(X > 2) &= 1 - (\mathbb{P}(X = 0) + \mathbb{P}(X = 1) + \mathbb{P}(X = 2)) \\ &= 1 - ((0,99)^{100} + 100 \cdot (0,01) \cdot (0,99)^{99} + \frac{100 \cdot 99}{2}(0,01)^2(0,99)^{98}) \approx 0,0794. \end{aligned}$$

Z drugiej strony, dla $Y \sim Poiss(\lambda)$ z parametrem $\lambda = np = 1$ uzyskujemy

$$\begin{aligned} \mathbb{P}(Y > 2) &= 1 - (\mathbb{P}(Y = 0) + \mathbb{P}(Y = 1) + \mathbb{P}(Y = 2)) \\ &= 1 - e^{-1}\left(\frac{1^0}{0!} + \frac{1^1}{1!} + \frac{1^2}{2!}\right) = 1 - \frac{5}{2e} \approx 0,0803 \end{aligned}$$

Błąd przybliżenia jest rzeczywiście mniejszy niż $np^2 = 100(0,01)^2 = 0,01$.

3. Rozkłady absolutnie ciagłe

3.1. Definicja i własności rozkładów absolutnie ciągłych. Niech X będzie zmienną losową o wartościach rzeczywistych określoną na przestrzeni probabilistycznej $(\Omega, \mathcal{F}, \mathbb{P})$. Przypomnijmy, że rozkład zmiennej X oznaczamy przez μ_X.

Definicja. Jeżeli istnieje funkcja nieujemna i całkowalna f na $\mathbb{R}$ taka, że

$$\mu_X(A) = \int_A f(x)dx, \quad A \in \mathcal{B}(\mathbb{R}),$$

to mówimy, że μ_X jest **rozkładem absolutnie ciągłym**. Funkcję f nazywamy wtedy **gęstością rozkładu** μ_X.

Własności gęstości i rozkładów absolutnie ciągłych:

- $\mu_X(\mathbb{R}) = \int_{-\infty}^{\infty} f(x)dx = 1$;
- $\mathbb{P}(X = a) = 0$, dla dowolnego $a \in \mathbb{R}$;
- $\mathbb{P}(X < a) = \mathbb{P}(X \leq a) = \int_{-\infty}^{a} f(x)dx$, dla dowolnego $a \in \mathbb{R}$;
- $\mathbb{P}(X > a) = \mathbb{P}(X \geq a) = \int_{a}^{\infty} f(x)dx$, dla dowolnego $a \in \mathbb{R}$;
- $\mathbb{P}(a < X < b) = \ldots = \mathbb{P}(a \leq X \leq b) = \int_a^b f(x)dx$, dla dowolnych $a < b$;
- $F_X(t) = \int_{-\infty}^{t} f(x)dx$, $t \in \mathbb{R}$;
- Jeżeli F_X jest ciągła wszędzie i różniczkowalna poza skończoną liczbą punktów, to rozkład μ_X jest absolutnie ciągły z gęstością $f(x) = F_X'(x)$ (tam, gdzie pochodna istnieje).

3.2. Najważniejsze przykłady rozkładów absolutnie ciągłych.

(1) Rozkład jednostajny na odcinku $[a, b] \subset \mathbb{R}$, $a < b$ (ozn. $U[a, b]$) ma gęstość

$$f(x) = \frac{1}{b-a}\mathbf{1}_{[a,b]}(x).$$

(2) Rozkład normalny (Gaussa) z parametrami (m, σ^2), $m \in \mathbb{R}$, $\sigma^2 > 0$ (ozn. $N(m, \sigma^2)$) ma gęstość

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right).$$

Standardowy rozkład normalny $N(0, 1)$ ma zatem gęstość

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right).$$

Dystrybuanta standardowego rozkładu normalnego jest funkcją specjalną i jest oznaczana przez

$$\Phi(t) = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx, \quad t \in \mathbb{R}.$$

Jej wartości można odczytać z tablic rozkładu normalnego. Zauważmy, że $\Phi(0) = 1/2$.

(3) Rozkład wykładniczy z parametrem $\lambda > 0$ (ozn. $Exp(\lambda)$) ma gęstość

$$f(x) = \lambda e^{-\lambda x}\mathbf{1}_{(0,\infty)}(x).$$

(4) Rozkład gamma z parametrami $\alpha > 0$ i $\beta > 0$ ma gęstość

$$f(x) = \frac{\alpha^\beta}{\Gamma(\beta)} x^{\beta-1} e^{-\alpha x}\mathbf{1}_{(0,\infty)}(x).$$

Funkcja Gamma Eulera:

$$\Gamma(\beta) = \int_0^\infty x^{\beta-1}e^{-x}dx, \quad \beta > 0.$$

Zachodzi równość $\Gamma(n) = (n-1)!$.

Przykład. Czas oczekiwania na tramwaj na przystanku ma rozkład wykładniczy z parametrem $\lambda = 15$. Jakie jest prawdopodobieństwo, że będziemy czekać na przyjazd tramwaju dłużej niż 10 min? Jakie jest prawdopodobieństwo, że tramwaj przyjedzie między 10 a 20 minutą czekania? Omówimy to dokładnie 15 kwietnia.