Wstęp do rachunku prawdopodobieństwa
7. Parametry rozkładów dyskretnych

Twierdzenie. *Dla dowolnych zmiennych losowych $X_1, X_2, \ldots, X_n$ zachodzi:*

$$\mathbb{E}(X_1 + X_2 + \ldots + X_n) = \mathbb{E}X_1 + \mathbb{E}X_2 + \ldots + \mathbb{E}X_n.$$

Twierdzenie. *Niech $X_1, X_2, \ldots, X_n$ będzie ciągiem zmiennych losowych. Jeśli dla każdego $i = 1, 2, \ldots, n$ istnieje $VarX_i$, wówczas*

$$Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} VarX_i + 2 \sum_{1 \leqslant i < j \leqslant n} Cov(X_i, X_j).$$

W szczególności, jeśli zmienne losowe $X_1, X_2, \ldots, X_n$ są **parami nieskorelowane***, tzn. dla każdych $1 \leqslant i < j \leqslant n$ mamy $Cov(X_i, X_j) = \rho(X_i, X_j) = 0$, to*

$$Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} VarX_i.$$

Twierdzenie (Własności wariancji)**.** *Jeśli wariancja $VarX$ zmiennej losowej X istnieje, to dla dowolnej stałej $a \in \mathbb{R}$ zachodzi*

$$Var(aX) = a^2\, VarX$$

oraz

$$Var(X + a) = VarX.$$

Przykład. *Załóżmy, że zmienna losowa X przyjmuje wartości $0, 1, \ldots, n$ i może zostać przedstawiona w postaci*

$$X = X_1 + X_2 + \ldots + X_n,$$

gdzie każda ze zmiennych losowych X_i, $i = 1, 2, \ldots, n$, ma rozkład dwupunktowy o zbiorze atomów $\{0, 1\}$ z $\mathbb{P}(X_i = 1) = p_i$. Takie zmienne losowe X_i nazywamy **zmiennymi losowymi indykatorowymi***. Ponadto załóżmy, że zmienne te są* **parami niezależne***, czyli dla każdego $1 \leqslant i < j \leqslant n$ zmienne losowe X_i i X_j są niezaeżne, co z kolei implikuje $Cov(X_i, X_j) = 0$. Wówczas możemy wyznaczyć wartość oczekiwaną oraz wariancję zmiennej losowej X w prosty sposób odwołując się do rozkładów zmiennych indykatorowych X_i. Mianowicie:*

$$\mathbb{E}X = \mathbb{E}X_1 + \mathbb{E}X_2 + \ldots + \mathbb{E}X_n = p_1 + p_2 + \ldots + p_n$$

oraz

$$VarX = VarX_1 + VarX_2 + \ldots + VarX_n = (p_1 - p_1^2) + (p_2 - p_2^2) + \ldots + (p_n - p_n^2).$$

Dodatek A. Zadania na ćwiczenia

Zadanie A.1. Wiedząc, że $\mathbb{E}X = 1$ i $\mathrm{Var}X = 5$, znajdź $\mathbb{E}(2 + X)^2$ oraz $\mathrm{Var}(4 + 3X)$.

Zadanie A.2. Roztargniona sekretarka włożyła losowo 10 zaadresowanych listów do 10 zaadresowanych kopert. Oblicz wartość oczekiwaną i wariancję liczby listów, które trafiły do swoich adresatów.

Zadanie A.3. Rzucamy 100 razy trzema kostkami. Wyznacz wartość oczekiwaną i wariancję sumy wyrzuconych oczek.

Zadanie A.4. Niech X będzie liczbą jedynek, a Y liczbą dwójek otrzymanych w wyniku n rzutów wyważoną kostką. Oblicz $\rho(X, Y)$.

Dodatek B. Zadania domowe

Zadanie B.1. Załóżmy, że zmienna losowa X ma rozkład dwumianowy $Bin(100, 1/5)$. Ile wynosi wartość oczekiwana oraz wariancja zmiennej losowej $Y = X/10 + 10$?

Zadanie B.2. Wyznacz $\mathrm{Cov}(X, Y)$, jeśli $\mathrm{Var}X = 3$, $\mathrm{Var}Y = 2$, a $\mathrm{Var}(X + 2Y) = 15$. Czy można rozstrzygnąć, czy zmienne losowe X i Y są niezależne?

Zadanie B.3. Łucznik strzela do tarczy n razy. Za każdym razem trafia niezależnie za i punktów, $1 \leqslant i \leqslant 10$, z prawdopodobieństwem $1/10$. Wyznacz wartość oczekiwaną i wariancję liczby punktów, które uzyska.

Zadanie B.4. W urnie jest 6 losów o wartościach: $1, 1, 1, 1, 2, 2$. Losujemy z urny 2 losy jednocześnie. Niech X będzie największą z wylosowanych wartości, a Y sumą wartości wylosowanych losów. Oblicz $\mathrm{Cov}(X, Y)$.

Zadanie B.5. Mamy do dyspozycji po 100 kul w kolorach: czerwony, zielony i niebieski. Wrzucamy losowo po 3 z tych kul do 100 urn tak, że wykorzystujemy wszystkie kule. Wyznacz wartość oczekiwaną liczby urn z kulami w trzech różnych kolorach.

Zadanie B.6. Rozkład hipergeometryczny dotyczy eksperymentu, w którym losujemy r–elementową próbkę z m–elementowej populacji, w której znajduje się n wyróżnionych elementów (np. w urnie znajduje się m kul, z czego dokładnie n jest białych, losujemy z urny r kul i interesuje nas ile z wylosowanych kul jest białych). Wówczas prawdopodobieństwo, że w r–elementowej próbce znajdzie się dokładnie k wyróżnionych elementów, gdzie $\max(0, n + r - m) \leqslant k \leqslant \min(n, r)$, wynosi

$$p_k = \frac{\binom{n}{k}\binom{m-n}{r-k}}{\binom{m}{r}}.$$

Niech X będzie zmienną losową o rozkładzie hipergeometrycznym z parametrami m, n, r. Wyznacz wartość oczekwianą zmiennej losowej X.

Zadanie B.7. Rzucono dwa razy kostką. Niech X będzie sumą, a Y różnicą liczb oczek otrzymanych za pierwszym i drugim razem. Bez wyznaczania rozkładów brzegowych zmiennych losowych X i Y, oblicz $\mathrm{Cov}(X, Y)$ oraz $\mathrm{Var}(X + Y)$.

Odpowiedzi

B.1 $\mathbb{E}Y = 12, \mathrm{Var}Y = 0{,}16$

B.2 $\mathrm{Cov}(X, Y) = 1$, nie są niezależne

B.3 Wartość oczekiwana: $5{,}5n$; Wariancja $8{,}25n$

B.4 $\mathrm{Cov}(X, Y) = \frac{4}{15}$

B.5 $\frac{100^4}{\binom{300}{3}}$

B.6 $\mathbb{E}X = \frac{rn}{m}$

B.7 $\mathrm{Cov}(X, Y) = 0, \mathrm{Var}(X + Y) = 11\frac{2}{3}$