# Kule i urny

**Zadanie 1.** Wyznacz wartość oczekiwaną i wariancję rozkładu Poissona na podstawie funkcji tworzącej wyliczonej na wykładzie.

**Zadanie 2** (Ostatni krok dowodu twierdzenia 5.13 z wykładu)**.** Niech $X_1, \ldots, X_n$ będą zmiennymi niezależnymi z rozkładem Poissona o parametrze $\lambda = \ln n + c$ dla pewnej ustalonej stałej $c$. Niech $X = X_1 + \cdots + X_n$. Zatem, $X$ ma rozkład Poissona z parametrem $m = n\lambda$. Niech $\mathcal{E}$ będzie zdarzeniem, że dla każdego $i \in [n]$ mamy $X_i > 0$. Wykaż, że

$$\lim_{n\to\infty} \left| \mathrm{P}\left(\mathcal{E} \mid |X - m| \leqslant \sqrt{2m \ln m}\right) - \mathrm{P}\left(\mathcal{E} \mid X = m\right) \right| = 0.$$

**Zadanie 3.** Niech zmienna $X_i^{(m)}$ oznacza liczbę kul w $i$-tym kubełku, przy założeniu, że rzucamy $m$ kul do $n$ kubełków w sposób jednostajny i niezależny. Niech zmienne $Y_1^{(m)}, \ldots, Y_n^{(m)}$ będą niezależnymi zmiennymi z rozkładu Poissona o średniej $\frac{m}{n}$. Podczas wykładu widzieliśmy, że dla każdej nieujemnej funkcji $f$ zachodzi

$$\mathrm{E}\left(f\left(Y_1^{(m)}, \ldots, Y_n^{(m)}\right)\right) \geqslant \mathrm{E}\left(f\left(X_1^{(m)}, \ldots, X_n^{(m)}\right)\right) \mathrm{P}\left(\sum_{i=1}^{n} Y_i^{(m)} = m\right)$$

(i) Pokaż, że jeśli $\mathrm{E}\left(f\left(X_1^{(m)}, \ldots, X_n^{(m)}\right)\right)$ jest rosnąca względem $m$, to przy założeniu nieujemności $f$ zachodzi:

$$\mathrm{E}\left(f\left(Y_1^{(m)}, \ldots, Y_n^{(m)}\right)\right) \geqslant \mathrm{E}\left(f\left(X_1^{(m)}, \ldots, X_n^{(m)}\right)\right) \mathrm{P}\left(\sum_{i=1}^{n} Y_i^{(m)} \geqslant m\right)$$

(ii) Wykaż, że jeśli $\mathrm{E}\left(f\left(X_1^{(m)}, \ldots, X_n^{(m)}\right)\right)$ jest malejąca względem $m$ to

$$\mathrm{E}\left(f\left(X_1^{(m)}, \ldots, X_n^{(m)}\right)\right) \leqslant 2 \cdot \mathrm{E}\left(f\left(Y_1^{(m)}, \ldots, Y_n^{(m)}\right)\right).$$

**Zadanie 4.** Niech $X$ będzie zmienną losową z rozkładu Poissona o wartości oczekiwanej $\mathrm{E}(X) = \mu$. Zmienna $X$ reprezentuje liczbę błędów w pewnej książce. Każdy błąd jest niezależnie od pozostałych błędem gramatycznym z prawdopodobieństwem $p$ albo leksykalnym z prawdopodobieństwem $1-p$. Niech $Y$ określa liczbę błędów gramatycznych, zaś $Z$ liczbę błędów leksykalnych ($X = Y + Z$). Pokaż, że $Y$ jest zmienną losową z rozkładu Poissona o wartości oczekiwanej $p\mu$, zaś $Z$ jest zmienną losową z rozkładu Poissona o wartości oczekiwanej $(1-p)\,\mu$. Dodatkowo pokaż, że zmienne $Y$ oraz $Z$ są niezależne.

**Zadanie 5.** Niech $X$, $Y$ i $Z$ będą niezależnymi zmiennymi o rozkładzie Poissona. Niech $E(X) = 1$, $E(Y) = 2$ i $E(Z) = 3$. Oblicz $P(X + Y + Z \leqslant 2)$.

**Zadanie 6.** Niech $X$ jest zmienną Poissona z wartością oczekiwaną $\mu \in \{1, 2, \ldots\}$.

(i) Pokaż, że $\mathrm{P}\left(X=\mu+h\right) \geqslant \mathrm{P}\left(X=\mu-h-1\right)$ dla $0 \leqslant h \leqslant \mu-1$.
(ii) Na bazie poprzedniego podpunktu, wyargumentuj że $\mathrm{P}\left(X \geqslant \mu\right) \geqslant \frac{1}{2}$.

**Zadanie 7.** Rzucamy $m$ kul w sposób jednostajny i niezależny do $n$ urn, ponumerowanych $0, \ldots, n-1$. Mówimy, że w indeksie $i$ jest $k$-luka, jeśli wszystkie urny o indeksach $i, i+1, \ldots, i+k-1$ są puste.

(i) Wyznacz wartość oczekiwaną liczby $k$-luk.
(ii) Wyznacz nierówność Chernoffa dla liczby $k$-luk.

**Zadanie 8.** Otrzymujemy w prezencie $X$ kul, gdzie $X$ ma rozkład Poissona z parametrem 2022. Dla każdej kuli rzucamy sześciościenną kostką i jeśli wyrzuciliśmy 1 to malujemy kulę na biało, jeśli wyrzuciliśmy 2 to malujemy kulę na czarno, jeśli zaś wyrzuciliśmy coś innego to wyrzucamy kulę do śmieci.

1. Z jakim prawdopodobieństwem będziemy mieć więcej czarnych kul niż białych?
2. Jaki jest rozkład liczby czarnych kul przy założeniu, że nie uzyskaliśmy żadnej białej kuli?
3. Jaka jest oczekiwana liczba białych kul przy założeniu, że czarnych kul jest dwa razy więcej od białych?

Uwaga: w zadaniach typu "przeczytaj" należy być gotowym do sprawnego (ale zwięzłego) zreferowania danej sekcji przy tablicy.

**Zadanie 9.** Przeczytaj sekcję 5.5.1 *Chain Hashing* i 5.5.2 *Hashing: Bit Strings* (przykład zastosowania modelu kul i urn).

**Zadanie 10.** Rozwiąż zadanie 5.22 z książki. Uwaga: w podpunkcie (b) powinno być $\lfloor \log n \rfloor$ zamiast $\log n$.

**Zadanie 11.** Przeczytaj sekcję 5.6.1 *Random Graph Models* (modele losowych grafów, podobieństwo do modelu kul i urn). Wypowiedz i udowodnij Lemat 5.14 dla malejących własności grafów.

**Zadanie 12.** Dla jakiego $p \in [0,1]$ oczekiwana liczba kopii pografu indukowanego $K_{3,3}$ w grafie $G$ wylosowanym z modelu $G_{n,p}$ (definicja w książce) jest największa? Możesz założyć, że $n$ jest wystarczająco duże, wynik może zależeć od $n$.

## Zadanie bonusowe do spisania (2pkt)

Do zrozumienia treści poniższych zadań wymagane jest zrobienie Zadania 12.

**Problem 1.** Rozwiąż zadanie 5.18 z książki.

Dwa kolejne problemy to zastosowania powyższego.

**Problem 2.** Udowodnij twierdzenie 5.15 z książki.

**Problem 3.** Co można powiedzieć o oczekiwanej liczbie kopii pografu indukowanego $K_{3,3}$ w grafie $G$ wylosowanym z modelu $G_{n,N}$?

**Problem 4.** Przeczytaj i zrozum materiał z sekcji 5.6.2 *Application: Hamiltonian Cycles in Random Graphs*. W ramach weryfikacji należy napisać (indywidualne) oświadczenie, że się przeczytało i zrozumiało materiał.

**Problem 5.** Zauważ, że Algorytm 5.1 działa w czasie wielomianowym. Ponadto, gdy Algorytm 5.1 kończy się sukcesem to zwraca on poprawny cykl Hamiltona. Wiemy, że problem znajdowania cyklu Hamiltona jest NP-zupełny, więc jest wysoce spodziewane, że istnieją instancje, gdzie Algorytm 5.1 się myli, to znaczy, zwraca porażkę mimo, że graf zawiera cykl Hamiltona. Skonstruuj dowolnie dużą taką instancje.

**Problem 6.** Rozwiąż zadanie 5.27 z książki.

Każdy z trzech problemów jest warty $\frac{1}{3}$ punkta (czyli łącznie można zdobyć 2pkt). Przypominamy, że rozwiązania spisane **na komputerze** należy wysyłać na adres podany na stronie internetowej kursu w terminie do niedzieli 19.11.2023 do godziny 23:59.