Funkcje zmiennych losowych

11.1. Niech X i Y będą niezależne, przy czym

$$P\{X = 0\} = P\{X = 1\} = 0.5 \quad \text{oraz} \quad P\{Y < x\} = x \ (0 < x < 1).$$

Znaleźć rozkłady zmiennych losowych
a. $Z_1 = Y + X$;
b. $Z_2 = Y + \frac{1}{2}X$;
c. $Z_3 = YX$.

11.2. Znaleźć rozkład sumy niezależnych zmiennych losowych X i Y takich, że X ma rozkład jednostajny w przedziale $(-h, h)$, a Y ma dystrybuantę $F(y)$.

11.3. Wielkość $Q_X(l) = \sup_x P\{x \leq X \leq x + l\}$ nazywamy funkcją koncentracji zmiennej losowej X. Udowodnić, że dla dowolnego Y niezależnego od X funkcja koncentracji sumy $X + Y$ spełnia warunek $Q_{X+Y}(l) \leq Q_X(l)$.

11.4. Udowodnić, że jeżeli zmienna losowa X ma gęstość $p_X(x)$, to dla dowolnej zmiennej losowej Y niezależnej od X suma $X + Y$ ma także funkcję gęstości $p_{X+Y}(x)$, przy czym $p_{X+Y}(x) \leq \sup_x p_X(x)$.

11.5. Niech zmienna losowa X ma rozkład o gęstości $p_X(x)$. Znaleźć funkcję gęstości zmiennych losowych
a. $Y = aX + b$, gdzie a i b są rzeczywiste;
b. $Y = X^{-1}$;
c. $Y = \cos X$;
d. $Y = f(X)$, gdzie $f(x)$ jest funkcją ciągłą i monotoniczną.

11.6. Udowodnić, że dla dowolnej zmiennej losowej X o ciągłej dystrybuancie $F(x)$ dla $0 \leq x \leq 1$ mamy $P\{F(X) < x\} = x$.

11.7. Zmienna losowa skokowa X ma *rozkład Poissona*

$$P\{X = m\} = \frac{\lambda^m e^{-\lambda}}{m!}, \qquad (m = 0, 1, 2, \ldots).$$

Niech M oznacza średnią arytmetyczną z N niezależnych realizacji zmiennej X.
a. Znaleźć średnią i wariancję zmiennej M.
b. Znaleźć rozkład prawdopodobieństwa zmiennej M.
c. Narysować wykres dla wyników punktu b tego zadania przy $\lambda = 1$ i $N = 3$, i $N = 10$.

11.8. Zmienna losowa X ma *rozkład Cauchy'ego* dany przez gęstość

$$p(x) = \frac{c}{1 + x^2}.$$

Niech M będzie średnią arytmetyczną N niezależnych zmiennej X. Znaleźć
a. wartość c;
b. gęstość rozkładu M;
c. prawdopodobieństwo, że każda z dwóch niezależnych realizacji X będzie mniejsza od 1.

11.9. Niech $p_X(x)$, $p_Y(y)$, $p_{X+Y}(z)$ będą gęstościami zmiennych losowych zmiennych losowych X, Y, $X+Y$. Udowodnić, że jeżeli X i Y są niezależne, to

$$p_{X+Y}(z) = \int_{-\infty}^{\infty} p_X(z - y) p_Y(y) dy = \int_{-\infty}^{\infty} p_X(x) p_Y(z - x) dx.$$

11.10. Gęstość niezależnych zmiennych losowych X i Y równa się
a. $p_X(x) = p_Y(x) = \begin{cases} 0, & x \leq 0, \\ ae^{-ax}, & x > 0; \end{cases}$
b. $p_X(x) = p_Y(x) = \begin{cases} 0, & x \leq 0, x \geq a, \\ \frac{1}{a}, & 0 < x < a; \end{cases}$
c. $p_X(x) = p_Y(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right)$.
Wyznaczyć gęstość rozkładu zmiennej losowej $Z = X/Y$.

11.11. Znaleźć rozkład iloczynu niezależnych czynników X i Y mając dane ich dystrybuanty $F_1(x)$ i $F_2(y)$.

11.12. Zmienne losowe $X_1, X_2, \dots, X_n, \dots$ są niezależne i mają jednakowy rozkład jednostajny w przedziale $[0,1]$. Niech Y będzie zmienną losową równą liczbie całkowitej k, dla której po raz pierwszy suma $S_k = X_1 + X_2 + \cdots + X_k$ przekracza wartość 1. Udowodnić, że $EY = e$.

11.13. Dany jest ciąg niezależnych zmiennych losowych $\{X_i\}$ przyjmujących wartości 0 i 1 z prawdopodobieństwami 0.5. Znaleźć rozkład zmiennej losowej

$$X = \sum_{i=1}^{\infty} \frac{X_i}{2^i}.$$

11.14. Dla przeprowadzenia obliczeń tzw. metodą Monte Carlo potrzeba często ciągu niezależnych zmiennych losowych o rozkładzie normalnym. W maszynach cyfrowych za pomocą metod teorioliczbowych lub za pomocą urządzeń fizycznych można generować ciągi niezależnych zmiennych losowych $X_1, X_2, \dots$ o rozkładzie jednostajnym w przedziale $[0,1]$. Można pokazać, że istnieje taka funkcja $\psi(x)$, że zmienne losowe $Y_i = \psi(X_i)$ mają rozkłady normalne. Niewygodnie jest jednakże konstruować ciąg zmiennych o rozkładzie normalnym wychodząc z ciągu $\{X_i\}$ i funkcji ψ, gdyż do zapamiętania funkcji ψ potrzeba dużej ilości miejsc w pamięci maszyny. Zazwyczaj postępuje się następująco. Rozbija się ciąg $\{X_i\}$ na pary. Z każdej pary z par (X_i, X_{i+1}) za pomocą przekształcenia

$$\varphi = 2\pi X_i, \;\; z = -\ln X_{i+1}, \;\; r = \sqrt{2\pi}, \;\; Y_i = r\cos\varphi, \;\; Y_{i+1} = r\sin\varphi$$

otrzymuje się ciąg wzajemnie niezależnych zmiennych o rozkładzie normalnym. W związku z tym pojawiają się następujące zagadnienia:

a. znaleźć funkcję ψ;
b. pokazać, że z ma rozkład wykładniczy;
c. pokazać, że Y_i i Y_{i+1} są niezależne i mają rozkład normalny o parametrach 0 i 1.

11.15. Na odcinek $[0,1]$ rzucono n punktów. Zakładając, że punkty zostały rzucone losowo (tzn. położenie każdego z nich ma rokład jednostajny na przedziale $[0,1]$ i są one wzajemnie niezależne) znaleźć

a. gęstość rozkładu $Z_1 = \max(X_1, X_2, \dots, X_n)$;
b. gęstość rozkładu k–tego punktu od lewej strony;
c. łączny rozkład położenia k–tego i m–tego punktu licząc od lewej strony $(k < m)$;
d. gęstość rozkładu $Z_2 = \max_i X_i - \min_i X_i$.

11.16. Niech X_i $(i = 1, 2, \dots, n)$ będą niezależnymi zmiennymi losowymi o jednakowym rozkładzie. Wiadomo, że mają one rozkład jednostajny w pewnym przedziale, który jednakże jest nieznany. Można zbudować różne estymatory dla położenia środka tego przedziału; np.

$$a_1 = \frac{1}{n}\sum_{i=1}^{n} X_i, \quad a_2 = \frac{1}{2}\left(\max_{1\le i\le n} X_i + \min_{1\le i\le n} X_i\right).$$

Udowodnić, że estymatory te są *nieobciążone*, tzn., że $Ea_1 = Ea_2 = a$, gdzie a jest środkiem przedziału. Udowodnić, że $D^2 a_1 > D^2 a_2$, tzn., że estymator a_2 jest efektywniejszy od estymatora a_1.

11.17. Niech zmienne losowe X_i $(i = 1, 2, \dots, n)$ będą niezależne i mają jednakowy rozkład o dystrybuancie

$$F(x) = \begin{cases} 0, & \text{dla } x \le 0, \\ 1 - e^{-x}, & \text{dla } x > 0. \end{cases}$$

Znaleźć rozkład zmiennej losowej $Z_n = \max_{1\le i\le n} X_i$ i udowodnić, że można wybrać takie stałe a_n, że rozkłady zmiennych $Z_n - a_n$ zmierzają do rozkładu granicznego.

11.18. Niech X_i $(i = 1, 2, \dots, n)$ będą niezależnymi zmiennymi losowymi o jednakowym rozkładzie o dystrybuancie $F(x)$. Oznaczmy przez $\nu(x)$ ilość tych X_i, które są mniejsze od x. Udowodnić, że rozkład

$$D_n = \sup_x \left| \frac{\nu(x)}{n} - F(x) \right|$$

nie zależy od $F(x)$. Znaleźć rozkład $\nu(x)$.
Uwaga. Fakt, że rozkład D_n nie zależy od rozkładu $F(x)$ ma ważne zastosowanie w statystyce. Opierają się na nim kryteria Kołmogorowa dla odchylenia empirycznej funkcji rozkładu od dystrybuanty teoretycznej.