

WSTĘP DO RACHUNKU PRAWDOPODOBIEŃSTWA
12. NIERÓWNOŚCI PROBABILISTYCZNE

DODATEK A. ZADANIA NA ĆWICZENIA

Zadanie A.1. Czas pracy X pewnego urzędnika (liczony w dniach) ma rozkład dany dystrybuantą $F(x) = 1 - e^{-x/400}$ dla $x \geq 0$ i $F(x) = 0$ dla $x \leq 0$.

(a) Z jakim rozkładem prawdopodobieństwa mamy do czynienia?

Na podstawie dystrybuanty widzimy, że zmienna losowa X ma rozkład wykładniczy z parametrem $\lambda = \frac{1}{400}$.

(b) Znajdź prawdopodobieństwo, że urządzenie będzie działało przez rok (nieprzestępny).

Urządzenie będzie działało przez rok (nieprzestępny), jeżeli będzie działało przez co najmniej 365 dni. Zatem musimy obliczyć $\mathbb{P}(X \geq 365)$, korzystając np. z dystrybuanty zmiennej losowej X :

$$\mathbb{P}(X \geq 365) = 1 - \mathbb{P}(X < 365) = 1 - \mathbb{P}(X \leq 365) = 1 - F(365) = 1 - (1 - e^{-365/400}) = e^{-73/80} \approx 0,4$$

(wykorzystaliśmy tutaj fakt, że podana dystrybuanta jest funkcją ciągłą).

(c) Wyznacz gęstość f_X zmiennej losowej X .

Gęstość f zmiennej losowej X wyznaczymy korzystając ze wzoru $f(x) = F'(x)$. Mamy zatem:

$$f(x) = \begin{cases} \frac{1}{400} e^{-x/400} & \text{dla } x \geq 0, \\ 0 & \text{w pozostałych przypadkach.} \end{cases}$$

(d) Ile wynosi wartość oczekiwana oraz wariancja zmiennej losowej X ?

Ponieważ zmienna losowa X ma rozkład wykładniczy z parametrem $\lambda = \frac{1}{400}$, możemy od razu wyznaczyć jej wartość oczekiwaną oraz wariancję:

$$\mathbb{E}X = \frac{1}{\lambda} = 400, \quad \text{Var}X = \frac{1}{\lambda^2} = 400^2 = 160000.$$

(e) Oblicz dokładną wartość $\mathbb{P}(X \geq 1000)$, a następnie oszacuj to prawdopodobieństwo, korzystając z nierówności Markowa, a potem z nierówności Czebyszewa.

Korzystając z dystrybuanty zmiennej losowej X otrzymujemy:

$$\mathbb{P}(X \geq 1000) = 1 - \mathbb{P}(X < 1000) = 1 - F(1000) = 1 - (1 - e^{-1000/400}) = e^{-5/2} \approx 0,082.$$

Następnie porównamy ten wynik z oszacowaniami, które dostajemy z nierówności Markowa oraz Czebyszewa-Bienaymé. Zwróćmy najpierw uwagę, że zmienna losowa X przyjmuje wartości ujemne z zerowym prawdopodobieństwem (bo $f(x) = 0$ dla $x < 0$). Oznacza to, że $\mathbb{P}(X \geq 1000) = \mathbb{P}(|X| \geq 1000)$ oraz $\mathbb{E}X = \mathbb{E}|X|$. Zatem z nierówności Markowa dostajemy:

$$\mathbb{P}(X \geq 1000) = \mathbb{P}(|X| \geq 1000) \leq \frac{\mathbb{E}|X|}{1000} = \frac{400}{1000} = \frac{2}{5} = 0,4.$$

Oszacujemy teraz to prawdopodobieństwo korzystając z nierówności Czebyszewa-Bienaymé. Zanim to zrobimy zauważmy, że zdarzenie $X - \mathbb{E}X \geq 600$ jest podzbiorem zdarzenia $|X - \mathbb{E}X| \geq 600$. Rzeczywiście

$$|X - \mathbb{E}X| \geq 600 \iff X - \mathbb{E}X \geq 600 \text{ lub } X - \mathbb{E}X \leq -600.$$

Mamy zatem:

$$\mathbb{P}(X \geq 1000) = \mathbb{P}(X - \mathbb{E}X \geq 1000 - 400) \leq \mathbb{P}(|X - \mathbb{E}X| \geq 600) \leq \frac{\text{Var}X}{600^2} = \frac{160000}{360000} = \frac{4}{9} \approx 0,44.$$

Jak widać w tym przypadku oba oszacowania są dalekie od właściwego wyniku i porównywalnie dobre, choć oszacowanie z nierówności Markowa jest nieznacznie lepsze.

Zadanie A.2. Na pewnym poznańskim skrzyżowaniu zdarza się średnio 0,9 wypadków rocznie. Oblicz przybliżoną wartość prawdopodobieństwa, że w przyszłym roku na tym skrzyżowaniu zajdą co najmniej trzy wypadki, a następnie oszacuj to prawdopodobieństwo używając nierówności Markowa i nierówności Czebyszewa.

Niech X będzie zmienną losową oznaczającą liczbę wypadków w przyszłym roku. Musimy najpierw ustalić, jaki rozkład ma ta zmienna losowa. Wiemy, że wypadki na skrzyżowaniu zdarzają się tylko od czasu do czasu, a ich średnia to 0,9 wypadków na rok, więc możemy przyjąć, że jest to zdarzenie „rzadkie”. Co za tym idzie możemy założyć, że X ma rozkład Poissona z parametrem 0,9 (bo przyjmujemy, że średnia liczba wypadków jest wartością oczekiwaną $\mathbb{E}X = \lambda$ zmiennej losowej X). W dalszej części zadania skorzystamy z własności rozkładu Poissona.

Policzymy najpierw przybliżoną wartość prawdopodobieństwa, że na tym skrzyżowaniu zajdą co najmniej trzy wypadki. Zauważmy, że dla $X \sim \text{Po}(0,9)$ mamy:

$$\mathbb{P}(X = k) = \frac{0,9^k}{k!} e^{-0,9}.$$

Zatem

$$\begin{aligned} \mathbb{P}(X \geq 3) &= 1 - \mathbb{P}(X < 3) = 1 - (\mathbb{P}(X = 0) + \mathbb{P}(X = 1) + \mathbb{P}(X = 2)) \\ &= 1 - \left(\frac{(0,9)^0}{0!} e^{-0,9} + \frac{(0,9)^1}{1!} e^{-0,9} + \frac{(0,9)^2}{2!} e^{-0,9} \right) = 1 - 2,305e^{-0,9} \approx 0,063. \end{aligned}$$

Teraz oszacujemy to prawdopodobieństwo z nierówności Markowa. Podobnie jak w poprzednim zadaniu możemy założyć, że zmienna losowa X nie przyjmuje ujemnych wartości, zatem $\mathbb{P}(X \geq 3) = \mathbb{P}(|X| \geq 3)$ oraz $\mathbb{E}|X| = \mathbb{E}X = 0,9$. W związku z tym mamy następujące oszacowanie:

$$\mathbb{P}(X \geq 3) \leq \frac{0,9}{3} = 0,3.$$

Zanim skorzystamy z nierówności Czebyszewa-Bienaymé zauważmy, że wariancja zmiennej losowej X dana jest wzorem $\text{Var}X = \lambda = 0,9$. A zatem:

$$\mathbb{P}(X \geq 3) = \mathbb{P}(X - \mathbb{E}X \geq 3 - 0,9) \leq \mathbb{P}(|X - \mathbb{E}X| \geq 2,1) \leq \frac{\text{Var}X}{2,1^2} = \frac{0,9}{4,41} = \frac{10}{49} \approx 0,204.$$

Widzimy, że po raz kolejny oszacowania z nierówności nie dały nam aż tak dobrych wyników, natomiast w tym przypadku lepsze okazało się oszacowanie z nierówności Czebyszewa-Bienaymé.

Zadanie A.3. *Pięć stacji pomiarowych na pięciu różnych kontynentach rejestruje wysokoenergetyczne cząstki kosmiczne (Przyjmujemy, że pomiary stacji są niezależne). Liczba cząstek zarejestrowanych rocznie przez każdą ze stacji ma rozkład Poissona ze średnią 3,6. Znajdź prawdopodobieństwo, że dokładnie trzy z pięciu stacji zarejestruje więcej niż 3 cząstki w następnym roku.*

Zastanówmy się najpierw jakie są szanse na to, że jedna stacja zarejestruje więcej niż 3 cząstki w następnym roku. Niech Y będzie zmienną losową zliczającą cząstki zarejestrowane przez pojedynczą stację w następnym roku. Wiemy, że Y ma rozkład Poissona z parametrem $\lambda = 3,6$ (bo średnia liczba zarejestrowanych cząstek to 3,6). Zatem dla $k = 0, 1, 2, \dots$ mamy:

$$\mathbb{P}(Y = k) = \frac{3,6^k}{k!} e^{-3,6}.$$

Korzystając z tych informacji obliczamy interesujące nas prawdopodobieństwo:

$$\begin{aligned} \mathbb{P}(Y > 3) &= 1 - \mathbb{P}(Y \leq 3) = 1 - (\mathbb{P}(Y = 0) + \mathbb{P}(Y = 1) + \mathbb{P}(Y = 2) + \mathbb{P}(Y = 3)) \\ &= 1 - e^{-3,6} \left(\frac{3,6^0}{0!} + \frac{3,6}{1!} + \frac{3,6^2}{2!} + \frac{3,6^3}{3!} \right) = 1 - 18,856e^{-3,6} \approx 0,485. \end{aligned}$$

W zadaniu pytamy o szanse na to, że dokładnie trzy z pięciu stacji zarejestrują więcej niż 3 cząstki. O tym doświadczeniu losowym możemy myśleć w następujący sposób: jest to pięciokrotne powtórzenie eksperymentu losowego (po jednym powtórzeniu na każdą stację), w którym interesuje nas czy zajdzie sukces (czyli stacja zarejestruje więcej niż 3 cząstki w następnym roku) czy też porażka (czyli stacja zarejestruje co najwyżej 3 cząstki w przyszłym roku). Ponadto możemy założyć, że stacje rejestrują cząstki niezależnie. Zatem mamy do czynienia ze schematem Bernoulliego, gdzie mamy $n = 5$ prób, a prawdopodobieństwo sukcesu w jednej próbie to $p = \mathbb{P}(Y > 3)$. Oznacza to, że szukane prawdopodobieństwo jest równe:

$$\tau_3 = \binom{5}{3} p^3 (1-p)^2 \approx \binom{5}{3} (0,485)^3 (1 - 0,485)^2 \approx 0,302.$$

Zadanie A.4. W pewnym urzędzie czas obsługi petenta (liczony w minutach) jest zmienną losową X o parametrach $\mathbb{E}X = 45$ i $\text{Var}X = 15$. Zakładamy, że czasy obsługi różnych petentów są niezależne.

- (a) Załóżmy, że w urzędzie pojawiło się 50 petentów. Oszacuj z dołu szansę, że średni czas ich obsługi (czyli średnia arytmetyczna z czasów ich obsługi) będzie wynosił od 35 do 55 minut (nierówności ostre)?

Niech Y będzie zmienną losową oznaczającą łączny czas obsługi 50 petentów. Chcemy oszacować prawdopodobieństwo, że średni czas obsługi będzie wynosił od 35 do 55 minut, czyli $\mathbb{P}\left(35 < \frac{Y}{50} < 55\right)$. Znamy tylko parametry związane z czasem obsługi pojedynczego petenta, więc przedstawimy zmienną losową Y jako sumę pomocniczych zmiennych losowych X_1, \dots, X_{50} , gdzie dla $i = 1, \dots, 50$ zmienna losowa X_i oznacza czas obsługi i -tego petenta.

Musimy teraz obliczyć parametry zmiennej losowej Y . Zgodnie z treścią zadania każda ze zmiennych losowych X_i ma wartość oczekiwaną równą 45 i wariancję równą 15. Zatem z liniowości wartości oczekiwanej otrzymujemy:

$$\mathbb{E}Y = \mathbb{E}X_1 + \dots + \mathbb{E}X_{50} = 50 \cdot 45 = 2250.$$

Ponadto zakładamy, że czasy obsługi petentów są niezależne, skąd otrzymujemy:

$$\text{Var}Y = \text{Var}(X_1 + \dots + X_{50}) = \text{Var}X_1 + \dots + \text{Var}X_{50} = 50 \cdot 15 = 750.$$

Teraz możemy już oszacować interesujące nas prawdopodobieństwo korzystając z nierówności Czebyszewa-Bienaymé:

$$\begin{aligned} \mathbb{P}\left(35 < \frac{Y}{50} < 55\right) &= \mathbb{P}(1750 < Y < 2750) = \mathbb{P}(1750 - 2250 < Y - \mathbb{E}Y < 2750 - 2250) \\ &= \mathbb{P}(-500 < Y - \mathbb{E}Y < 500) = \mathbb{P}(|Y - \mathbb{E}Y| < 500) = 1 - \mathbb{P}(|Y - \mathbb{E}Y| \geq 500) \\ &\geq 1 - \frac{\text{Var}Y}{500^2} = 1 - \frac{750}{250000} = 1 - \frac{3}{1000} = 0,997. \end{aligned}$$

- (b) Oszacuj, ilu petentów w urzędzie wystarczy, by z prawdopodobieństwem co najmniej 0,9 średni czas obsługi należał do przedziału $(44, 46)$.

Podobnie jak w poprzednim podpunkcie niech zmienna losowa Y będzie łącznym czasem obsługi petentów, a zmienne losowe X_1, \dots, X_n będą czasami obsługi poszczególnych petentów. Wtedy $Y = X_1 + \dots + X_n$, a dla każdego $i = 1, \dots, n$ mamy $\mathbb{E}X_i = 45$ oraz $\text{Var}X_i = 15$. W związku z tym parametry zmiennej losowej Y wynoszą:

$$\mathbb{E}Y = \mathbb{E}X_1 + \dots + \mathbb{E}X_n = n \cdot 45 = 45n$$

oraz

$$\text{Var}Y = \text{Var}(X_1 + \dots + X_n) = \text{Var}X_1 + \dots + \text{Var}X_n = 15n.$$

Naszym celem jest znalezienie takiej wartości liczby n , dla której prawdopodobieństwo $\mathbb{P}\left(\frac{Y}{n} \in (44, 46)\right)$ będzie równe co najmniej 0,9. W tym celu najpierw oszacujemy to prawdopodobieństwo korzystając z nierówności Czebyszewa-Bienaymé:

$$\begin{aligned} \mathbb{P}\left(\frac{Y}{n} \in (44, 46)\right) &= \mathbb{P}\left(44 < \frac{Y}{n} < 46\right) = \mathbb{P}(44n < Y < 46n) = \mathbb{P}(44n - 45n < Y - \mathbb{E}Y < 46n - 45n) \\ &= \mathbb{P}(-n < Y - \mathbb{E}Y < n) = \mathbb{P}(|Y - \mathbb{E}Y| < n) = 1 - \mathbb{P}(|Y - \mathbb{E}Y| \geq n) \geq 1 - \frac{\text{Var}Y}{n^2} \\ &= 1 - \frac{15n}{n^2} = 1 - \frac{15}{n}. \end{aligned}$$

Zatem wystarczy dobrać n spełniające poniższą nierówność

$$1 - \frac{15}{n} \geq 0,9.$$

Otrzymujemy $n \geq 150$. Zauważmy, że zgodnie z obliczeniami powyżej oznacza to, że jeśli w urzędzie pojawi się 150 (lub więcej) petentów, ich średni czas obsługi będzie mieścił się w przedziale $(44, 46)$ z prawdopodobieństwem co najmniej 0,9. Co więcej, to prawdopodobieństwo będzie rosło wraz z n .

Zadanie A.5. X_1, X_2, \dots, X_{20} są zmiennymi losowymi (niekoniecznie niezależnymi) o rozkładzie Poissona $Po(1)$.

(a) Użyj nierówności Markowa do oszacowania z góry $\mathbb{P}(X \geq 30)$, dla $X = \sum_{i=1}^{20} X_i$.

Musimy najpierw obliczyć $\mathbb{E}X$. Każda ze zmiennych losowych X_1, \dots, X_{20} ma rozkład Poissona z parametrem $\lambda = 1$, więc ich wartości oczekiwane również wynoszą 1. Zatem z liniowości wartości oczekiwanej otrzymujemy:

$$\mathbb{E}X = \mathbb{E}X_1 + \dots + \mathbb{E}X_{20} = 20 \cdot 1 = 20.$$

Zauważmy jeszcze, że zmienna losowa X jest sumą zmiennych losowych o rozkładach Poissona, a więc przyjmujących wartości ujemne z zerowym prawdopodobieństwem. Zatem zmienna losowa X także przyjmuje wartości ujemne z zerowym prawdopodobieństwem. Możemy teraz oszacować szukane prawdopodobieństwo korzystając z nierówności Markowa:

$$\mathbb{P}(X \geq 30) = \mathbb{P}(|X| \geq 30) \leq \frac{\mathbb{E}|X|}{30} = \frac{20}{30} = \frac{2}{3}.$$

(b) Załóżmy teraz, że X_1, X_2, \dots, X_{20} są niezależne. Czy można uzyskać lepsze oszacowanie prawdopodobieństwa $\mathbb{P}(X \geq 30)$ korzystając z nierówności Czebyszewa–Bienaimé?

Wariancja każdej ze zmiennych losowych X_1, \dots, X_{20} wynosi $\lambda = 1$. Zatem jeśli założymy, że zmienne losowe X_1, \dots, X_{20} są niezależne, to:

$$\text{Var}X = \text{Var}X_1 + \dots + \text{Var}X_{20} = 20 \cdot 1 = 20.$$

Możemy teraz oszacować szukane prawdopodobieństwo z nierówności Czebyszewa–Bienaimé:

$$\mathbb{P}(X \geq 30) = \mathbb{P}(X - \mathbb{E}X \geq 30 - 20) \leq \mathbb{P}(|X - \mathbb{E}X| \geq 10) \leq \frac{\text{Var}X}{10^2} = \frac{20}{100} = \frac{1}{5}.$$

Po raz kolejny oszacowanie z nierówności Czebyszewa–Bienaimé okazuje się być lepsze niż oszacowanie z nierówności Markowa.