RP dla bioinformatyki, seria I

1. Z liczb $\{1, 2, 3, \ldots, 98, 99\}$ losujemy jedną. Jakie jest prawdopodobieństwo, że będzie ona podzielna przez 2? Przez 3? Przez 2 i przez 3? Przez 2 lub przez 3? Przez 2, ale nie przez 3?

2. Które z poniższych zdarzeń jest najbardziej prawdopodobne?

- W najbliższym losowaniu Lotto padnie szóstka 1, 2, 3, 4, 5, 6.
- W najbliższym losowaniu Lotto padnie ta sama szóstka co w poprzednim losowaniu.
- W najbliższych dwóch losowaniach Lotto padnie taka sama szóstka.

3. Rzucamy trzy razy sześcienną kostką. Jakie jest prawdopodobieństwo, że

a) pewien wynik wystąpi dokładnie dwa razy?
b) wyrzucimy dokładnie dwie szóstki?

4. Pięćdziesięciu harcerzy ustawiło się w losowej kolejności w szeregu. Jakie jest prawdopodobieństwo

a) tego, że Antek stoi obok Bartka?
b) tego, żecały trójosobowy zastęp Żubrów stoi razem?
c) tegoo, że Antek stoi przed Bartkiem, a Bartek przed Czarkiem? („Przed”, ale niekoniecznie „bezpośrednio przed”).

5. Losujemy 13 kart ze standardowej talii 52 kart. Jakie jest prawdopodobieństwo

(1) tego, że każda karta jest innej wartości?
(2) tego, że nie mamy żadnej figury?
(3) układu kolorów $4 - 4 - 3 - 2$?
(4) układu kolorów $4 - 4 - 4 - 1$?
(5) tego, że mamy dokładnie dwa piki?
(6) tego, że mamy co najmniej dwa piki?
(7) tego, że nie mamy żadnego pika?
(8) tego, że nie mamy (co najmniej) jednego z kolorów?

6. Które z poniższych zdarzeń jest najbardziej prawdopodobne?

- Uzyskanie co najmniej jednej szóstki w sześciu rzutach kostką.
- Uzyskanie co najmniej dwóch szóstek w 12 rzutach kostką.
- Uzyskanie co najmniej trzech szóstek w 18 rzutach kostką.

7. Dziesięć osób usiadło losowo przy okrągłym stole. (Jakie są możliwe modele?)

a) Jakie jest prawdopodobieństwo tego, że Alicja siedzi obok Barbary?
b) Jakie jest prawdopodobieństwo tego, że Barbara siedzi między Alicją a Celiną?

8. Losujemy 5 kart ze standardowej talii 52 kart. Jakie jest prawdopodobieństwo, że otrzymaliśmy

a) pokera?
b) karetę?
c) fulla?
d) kolor?
e) strita?
f) trójkę?

g) dwie pary?
h) parę?
i) wysoką kartę?

9. a) Ile różnych słów (niekoniecznie sensownych) można ułożyć przestawiając litery słowa SZCZEBRZESZYN?

b) Na 8 karteczkach napisano litery A, A, A, R, S, W, W, Z. Układamy je w losowej kolejności. Jakie jest prawdopodobieństwo tego, że powstanie napis WARSZAWA?

10. a) Jakie jest prawdopodobieństwo, że ktoś z grupy 25 studentów ma urodziny tego samego dnia, co prowadzący ćwiczenia?

b) Jakie jest prawdopodobieństwo tego, że w grupie 25 studentów (co najmniej) dwie osoby mają urodziny tego samego dnia?

11. a) Na ile sposobów n zbojców może podzielić między sobą m złotych monet? Wszystkie monety mają ten sam nominał, więc zakładamy, że są nierozróżnialne.

b) Na ile sposobów n zbójców może podzielić między sobą k kamieni szlachetnych? Kamienie są różnych kolorów i rozmiarów, więc są jak najbardziej rozróżnialne.

12. Opiekun roku ocenia szanse studenta na zdanie egzaminów z AM i GALu. Na podstawie danych z lat poprzednich wiadomo, że szansa zdania egzaminu z GALu to 0.8, szansa zdania co najmniej jednego egzaminu to 0.9, a szansa zdania obydwu – 0.5. Jaka jest szansa zdania egzaminu z AM?

13. Losujemy 6 z 49 liczb. Jakie jest prawdopodobieństwo, że wśród wylosowanych liczb nie będzie dwóch kolenych?

14. a) W n urnach rozmieszczono losowo n kul. Jakie jest prawdopodobieństwo tego, że żadna urna nie jest pusta?

b) W n urnach rozmieszczono losowo n kul. Jakie jest prawdopodobieństwo tego, że dokładnie jedna urna będzie pusta?

c) W n urnach rozmieszczono losowo $n+1$ kul. Jakie jest prawdopodobieństwo tego, że żadna urna nie będzie pusta?

15. Rzucamy symetryczną monetą do chwili wypadnięcia orła.

a) Opisać przestrzeń probabilistyczną dla tego doświadczenia
b) Udowodnić, że p.n. będziemy rzucać skończenie wiele razy.
c) Jaka jest szansa, że liczba rzutów będzie większa niż 5?
d) Jaka jest szansa, że liczba rzutów będzie parzysta?

16. Z kwadratu $[0,1]^2$ wybrano losowo punkt o współrzędnych (X,Y). Jakie jest prawdopodobieństwo, że

a) $\max\{X,Y\} \leqslant 1/2$;
b) $\min\{X,Y\} \leqslant 1/2$;
c) $X^2 + Y^2 \leqslant 1/4$;
d) $|X - Y| \leqslant 1/2$;
e) X jest liczbą wymierną.

17. Patyk łamiemy na trzy kawałki. (Jak?)

a) Jakie jest prawdopodobieństwo, że z tych kawałków można zbudować trójkąt?
b) Jakie jest prawdopodobieństwo, że środkowa część jest najkrótsza?

RP dla bioinformatyki, seria II

1. Gramy w brydża (jest przed licytacją).

a) Wśród pierwszych siedmiu kart nie mamy asa. Jakie jest prawdopodobieństwo tego, że nie mamy żadnego asa?

b) Podczas rozdawania kart podejrzeliśmy, że naszą pierwszą kartą jest as pik. Jakie jest prawdopobieństwo tego, że będziemy mieli co najmniej dwa asy? Co jeśli nie dostrzegliśmy koloru naszego asa?

c) Mamy dwa asy. Jakie jest prawdopodobieństwo tego, że nasz partner nie ma żadnego asa?

d) Nie mamy ma żadnego asa. Jakie jest prawdopodobieństwo tego, że nasz partner ma dwa asy?

2. Z miejsca zbrodni udało się pobrać próbkę DNA należącą do sprawcy. Analiza laboratoryjna w połączeniu z danymi statystycznymi sugeruje, że DNA jednej osoby na 7000 pasuje do pobranej próbki. Jakiś czas później zaaresztowano podejrzanego. Po wykonaniu odpowienich badań okazało się, że jego DNA pasuje do próbki pobranej z miejsca zbrodni. Przedyskutować prawdziwość następujących stwierdzeń.

a) Prawdopodobieństwo, że podejrzany jest *niewinny* to zaledwie 1/7000.

b) Przecież w mieście, w którym popełniono zbrodnię żyje 700 tysięcy osób, więc aż 100 z nich ma pasujące DNA, więc prawdopodobieństwo, że podejrzany jest *winny* to co najwyżej 1/100.

3. Gramy w brydża. Na stole i ręku mamy łącznie 7 pików. Co jest bardziej prawdopodobne: podział pików u przeciwników $4-2$ czy $3-3$?

4. W pewnym mieście 85% taksówek jest koloru zielonego, a 15% – niebieskiego. Świadek nocnego wypadku zakończonego ucieczką taksówki z miejsca zdarzenia, twierdzi, że taksówka była niebieska. Z przeprowadzonych eksperymentów wynika, że w trudnych warunkach oświetleniowych świadek poprawnie rozpoznaje kolor w 80% przypadków (a w pozostałych 20% myli się). Jaka jest szansa, że w wypadku uczestniczyła niebieska taksówka?

5. W urnie są trzy kule czarne i pewna liczba kul białych. Bolek dorzucił do urny jedną kulę losowego koloru (białą lub czarną). Lolek losuje jedną kulę i jest ona czarna. Jakie jest prawdopodobieństwo tego, że Bolek dorzucił do urny czarną kulę?

6. W populacji jest 15% dyslektyków. Jeśli w teście diagnostycznym uczeń popełni 6 lub więcej błędów, to zostaje uznany za dyslektyka. Każdy dyslektyk na pewno popełni co najmniej 6 błędów w takim teście, ale również nie-dyslektyk może popełnić więcej niż 5 błędów z prawdopodobieństwem 1/10. Jaś popełnił w teście 6 błędów.

a) Jakie jest prawdopodobieństwo tego, że jest dyslektykiem?

b) Jakie jest prawdopodobieństwo tego, że w kolejnym teście popełni co najmniej 6 błędów?

7. a) Uczestnik teleturnieju stoi przed trzema zasłoniętymi bramkami. Za jedną z nich jest nagroda główna (samochód), za pozostałymi dwoma – koza. Uczestnik wybiera jedną z bramek. Spośród dwóch pozostałych bramek prowadzący odsłania jedną i pokazuję, że jest za nią koza. Uczestnik teleturnieju może w tym momencie zmienić swój wybór bramki. Czy powininen to zrobić?

b) A co jeśli bramek jest 100, samóchód jest jeden, kóz jest 99, a prowadzący odsłania 98 bramek ukazując 98 kóz?

8. W urnie jest b białych i c czarnych kul. Powtarzamy następującą operację: losujemy jedną kulę, sprawdzamy jej kolor, a następnie wrzucamy ją z powrotem do urny i dorzucamy jeszcze a kul tego samego koloru.

a) Jakie jest prawdopodobieństwo wylosowania za drugim razem kuli białej?

b) Jakie jest pradopodobieństwo wylosowania za n-tym razem kuli białej?

c) Jakie jest prawdopodobieństwo wylosowania dokładnie k białych kul w n losowaniach?

9. Wiadomo, że 10% populacji choruje na chorobę wieńcową, którą można wykryć przy pomocy testu wysiłkowego. Czuość tego testu wynosi 65%, a swoistość 85% (tzn. test daje wynik dodatni u 65% osób chorych i u 15% osób zdrowych). Wyznaczyć prawdopodobieństwo tego, że:

a) u losowo wybranej osoby próba wysiłkowa prowadzi do poprawnej diagnozy.

b) osoba z ujemnym wynikiem testu jest zdrowa.

c) osoba z dodatnim wynikiem testu jest chory.

d) osoba wybrana losowo spośród personelu szpitala, która trzy razy powtórzyła test i trzykrotnie uzyskała wynik dodatni, jest chora.

10. W pewnej miejscowości są dwie szkoły podstawowe – nr 1 i nr 2. Na egzaminie na koniec roku najgorzej wypadło zadanie o procentach. Okazało się, że procent dziewczynek w szkole nr 1, które potrafiły je rozwiązać, jest większy niż procent dziewczynek w szkole nr 2, które potrafiły to zrobić. Podobnie było w przypadku chłopców. Czy znaczy to, że „statystyczne dziecko” ze szkoły nr 1 lepiej wypadło w umiejętności radzenia sobie z procentami niż „statystycznego dziecko” ze szkoły nr 2?

11. Z odcinka $[0, 1]$ wybieramy losowo dwie liczby. Obliczyć prawdopodobieństwo tego, że niewiększa z nich przekracza $1/3$, jeśli wiadomo, że suma liczb jest co najmniej jeden.

12. Pewna rodzina ma dwójkę dzieci. Jaka jest szansa, że ma dwóch synów? Jaka jest szansa, że ma dwóch synów, jeśli wiemy, że

a) starsze dziecko to syn?

b) losowo wybrane dziecko okazało się być synem?

c) co najmniej jedno z dzieci to syn? (Skąd to wiemy?)

d) co najmniej jedno z dzieci to syn, który urodził sie w środę?

e) co najmniej jedno z dzieci to syn, który ma na drugię imię Piotr? A jeśli ma on drugie imię Eustachy? (Przedyskutować możliwe problemy).

RP dla bioinformatyki, seria III

1. Rzucamy dwiema sześciennymi kostkami. Czy zdarzenia „na pierwszej kostce wypadła nieparzysta liczba oczek”, „na drugiej kostce wypadła nieparzysta liczba oczek”, „suma oczek z obu kostek była nieparzysta” są parami niezależne? Czy są niezależne (łącznie)?

2. Rzucamy trzema sześciennymi kostkami o następujących ściankach:

$$X: \quad 2,2,4,4,9,9; \qquad Y: \quad 1,1,6,6,8,8; \qquad Z: \quad 3,3,5,5,7,7.$$

Rozważmy następujące zdarzenia:

$$A = \{\text{na kostce } X \text{ wypadła większa liczba niż na kostce } Y\},$$
$$B = \{\text{na kostce } Y \text{ wypadła większa liczba niż na kostce } Z\},$$
$$C = \{\text{na kostce } Z \text{ wypadła większa liczba niż na kostce } X\}.$$

Jakie są ich prawdopodobieństwa? Czy zdarzenia te są niezależne?

3. Rzucono n kostkami do gry. Niech dla $1 \leqslant i \leqslant n$, A_i oznacza zdarzenie „na i-tej kostce wypadła szóstka, i niech A_{n+1} to zdarzenie „suma oczek na wszystkich kostkach jest podzielna przez 6. Wykazać, że dowolnych n zdarzeń spośród $A_1, \ldots, A_{n+1}$ jest niezależnych, ale zdarzenia $A_1, \ldots, A_{n+1}$ nie są niezależne.

4. Rzucamy monetą (niekoniecznie symetryczną). Wykazać, że p.n. wyrzucimy nieskończenie wiele razy wzorzec ORRROO.

5. Jaka jest najbardziej prawdopodobna liczba sukcesów w schemacie Bernoulliego? Ile wynosi odpowiednie prawdopodobieństwo, gdy $n = 100$, $p = 1/2$?

6. Rzucamy n razy monetą, na której orzeł wypada z prawdopodobieństwem $p \in (0,1)$. Podać zwięzły wzór na prawdopodobieństwo uzyskania parzyście wielu orłów. Jaką wartość otrzymamy dla $n = 20$ i $p = 1/10$?

7. a) Jakie jest prawdopodobieństwo, że w ciągu prób Bernoulliego a sukcesów pojawi się przed b porażkami?

b) Dwóch równych graczy jest zmuszonych przerwać grę do 4 wygranych przy stanie 2:1. Jak należy podzielić pulę?

RP dla bioinformatyki, seria IV

0. Znaleźć wartości oczekiwane i wariancje dyskretnych rozkładów, które pojawiły się na wykładzie, a dla których te parametry rozkładu nie zostały policzone.

1. a) Dla jakich wartości parametrów $a, b \in \mathbb{R}$ funkcja

$$F(t) = F_{a,b}(t) = \begin{cases} 0 & \text{dla } y < 0, \\ 1/4 & \text{dla } 0 \leqslant t < 1, \\ a & \text{dla } 1 \leqslant t < 2, \\ 1/2 & \text{dla } 2 \leqslant t < 3, \\ b & \text{dla } t = 3, \\ 1 & \text{dla } t > 3, \end{cases}$$

jest dystrybuantą pewnego rozkładu prawdopodobieństwa?

b) Jaka jest średnia tego rozkładu?

c) Dla jakich $a, b \in \mathbb{R}$ wariancja tego rozkładu jest najmniejsza?

2. W klasie 1A jest 28 uczniów, w klasie 1B – 30 uczniów, a w 1C – 32 uczniów. Wybieramy losowo ucznia z klasy pierwszej i oznaczamy przez X liczność klasy do której chodzi. Jaka jest wartość oczekiwana X? A jaka wariancja?

3. W urnie znajduje się jedna kula biała i jedna czarna. Wykonujemy następujący ciąg losowań: losujemy jedną kulę z urny, oglądamy ją, wrzucamy ją z powrotem i dorzucamy jedną kulę czarną. Czynność powtarzamy do momentu wylosowania kuli białej. Niech X oznacza liczbę losowań. Wyznaczyć rozkład zmiennej losowej X. Jaka jest jest wartość oczekiwana?

4. Tasujemy starannie nową talię, w której karty ułożone są w pewnej konkretnej kolejności. Ile kart znajdzie się średnio na tej samej pozycji co wyjściowa? Jaka jest wariancja?

5. Wykazać, że rozkład geometryczny (w której bądź wersji) ma własność braku pamięci: jeśli X ma rozkład geometryczny, to

$$\mathbb{P}(X > m + n | X > m) = \mathbb{P}(X > n).$$

6. Rzucamy trzy razy prawidłową monetą. Niech X oznacza różnicę między liczbą wyrzuconych reszek i wyrzuconych orłów. Wyznaczyć $\mathrm{Var}(X)$.

7. Nieprzyjaciel ma n czołgów. Ich liczba jest nam nieznana, ale wiemy, że są oznaczone numerami $1, 2, \ldots, n$. Zwiadowcy zaobserwowali grupę k czołgów; największym z numerów na czołgach było m. Analitycy mówią, że należy zinterpretować tę informację następująco: przy obserwacji losowo wybranych k czołgów wartość oczekiwana największego numeru powinna być równa (w przybliżeniu) m. Ile czołgów ma nieprzyjaciel?

8. Wybieramy losowo podzbiór zbioru $\{1, 2, \ldots, n\}$ (wybór każdego podzbioru jest tak samo prawdopodobny). Niech X oznacza moc wylosowanego podzbioru. Podać „zwięzły” wzór na $\mathbb{E}X$.

9. Niech $\lambda > 0$ i niech $p_n := \frac{\lambda}{n}$. Obliczyć granicę

$$\lim_{n\to\infty} \binom{n}{k} p_n^k (1 - p_n)^{n-k}.$$

10. Danych jest sto liczb rzeczywistych (niekoniecznie różnych). Jacek mówi nam, że po usunięciu pewnej z tych liczb mediana pozostałych 99 to 74. Placek inny mówi nam, że po usunięciu innej z tych liczb mediana pozostałych 99 to 68. Jaka jest mediana wszystkich 100 liczb?

11. Zmienna losowa X przyjmuje wartości w zbiorze liczb naturalnych (z zerem). Wykazać, że

$$\mathbb{E}X = \sum_{k=1}^{\infty} \mathbb{P}(X \geqslant k).$$

12. Zmienne losowe $X_1, \ldots, X_n$ są niezależne przy czym $X_i \sim \mathrm{Poiss}(\lambda_i)$. Jaki rozkład ma $X_1 + \cdots + X_n$?

RP dla bioinformatyki, seria V

1. a) Zmienna losowa X jest całkowalna z kwadratem ($\mathbb{E}X^2 < \infty$). Definiujemy

$$f(t) = \mathbb{E}(X-t)^2, \quad t \in \mathbb{R}.$$

Ile jest równe $\min_{\mathbb{R}} f$ i dla jakiego argumentu jest przyjmowane?

b*) Zmienna losowa X jest całkowalna ($\mathbb{E}|X| < \infty$). Definiujemy

$$h(t) = \mathbb{E}|X-t|, \quad t \in \mathbb{R}.$$

Dla jakich argumentów jest przyjmowane $\min_{\mathbb{R}} h$?

2. Co jest większe: prawdopodobieństwo, że kwadrat losowo wybranej liczby z przedziału $[0, 10]$ będzie większy niż 25, czy prawdopodobieństwo, że pierwiastek kwadratowy losowo wybranej liczby z przedziału $[0, 100]$ będzie większy niż 5?

3. Dla jakich wartości parametrów $a, b \in \mathbb{R}$ następujące funkcje

$$f(x) = ax^2\mathbf{1}_{[0,1]}(x),$$
$$g(x) = (ax^2+b)\mathbf{1}_{[0,1]}(x)$$

są gęstościami pewnych rozkładów prawdopodobieństwa?

4. Dla jakich wartości parametrów $a_1, a_2, b, c, d \in \mathbb{R}$ funkcja

$$F(t) = \begin{cases} 0 & \text{dla } t < 0, \\ a_1 & \text{dla } t = 0, \\ a_2 + \frac{1}{2}t^2 & \text{dla } t \in (0,1), \\ b & \text{dla } t \in [1,2), \\ ct + d & \text{dla } t \in [2,3), \\ 1 & \text{dla } t \geqslant 3, \end{cases}$$

jest dystrybuantą pewnego ciągłego rozkładu prawdopodobieństwa? Jakim wzorem dana jest gęstość?

5. Zmienna losowa G ma rozkład $\mathcal{N}(0,1)$, a liczby a, b są rzeczywiste. Jaką gęstość ma zmienna $aG + b$? Jak jej dystrybuanta wyraża się w terminach dystrybuanty G?

6. Zmienne losowe $X_1, \ldots, X_n$ są niezależne i mają rozkłady wykładcznicze z parametrami $\lambda_1, \ldots, \lambda_n$ odpowiednio. Jaki rozkład ma zmienna $Y = \min\{X_1, \ldots, X_n\}$?

7. Zmienna losowa X ma rozkład $\mathcal{N}(0,1)$. Wyznaczyć dystrybuantę i gęstość zmiennych losowych $Y = e^X$ (rozkład log-normalny) oraz $Z = X^2$.

8. Zmienna losowa X ma rozkład o gęstości $g(x) = \frac{3}{7}x^2\mathbf{1}_{[1,2]}(x)$. Dla jakiej funkcji $h\colon \mathbb{R} \to \mathbb{R}$ zmienna losowa $h(X)$ ma rozkład wykładniczy z parametrem 1?

9. Załóżmy, że X jest zmienną losową skupioną na $(0, \infty)$, która ma ciągłą dystrybuantę i własność braku pamięci (co to znaczy?). Wykazać, że X ma rozkład wykładniczy.

10. Zmienna losowa X ma rozkład o gęstości $g(x) = \frac{1}{2}\sin(x)\mathbf{1}_{[0,\pi]}(x)$. Jakim wzorem dana jest jej dystrybuanta? Ile są równe $\mathbb{E}X$, $\mathbb{E}\cos(X)$, $\mathbb{E}\sin(X)$?

11. Dla jakiego $c > 0$ funkcja $g(x) = \frac{1}{c(1+x^2)}, x \in \mathbb{R}$, jest gęstością? Wykazać, że zmienna losowa X o takiej gęstości nie ma wartości oczekiwanej. Jaki rozkład ma zmienna $1/X$?

12. Mówimy, że zmienna losowa X ma rozkład Γ z parametrem $a > 0$ (ozn. $X \sim \Gamma(a)$), jeśli X ma gęstość

$$\frac{1}{\Gamma(a)} x^{a-1} e^{-x} \mathbf{1}_{\{x \geqslant 0\}},$$

gdzie $\Gamma(a) = \int_0^\infty x^{a-1} e^{-x} dx$ to stała normująca.

a) Wykazać, że jeśli zmienne $X_1, \ldots, X_n$ są niezależne oraz $X_i \sim \Gamma(a_i)$, to $X_1 + \cdots + X_n \sim \Gamma(a_1 + \cdots + a_n)$.

b) Wykazać, że jeśli zmienne $X_1, \ldots, X_n$ są niezależne oraz każda ma rozkład wykładniczy z parametrem λ, to $X_1 + \cdots + X_n \sim \frac{1}{\lambda} Y$, gdzie $Y \sim \Gamma(n)$. W zależność od konwencji rozkład ten oznacza się $\Gamma(n, \lambda)$ lub $\Gamma(n, 1/\lambda)$.

c) Zmienne losowe $G_1, \ldots, G_n$ są niezależne i każda ma rozkład $\mathcal{N}(0, 1)$. Wykazać, że $X_1^2 + \cdots + X_n^2 \sim 2Z$, gdzie $Z \sim \Gamma(n/2)$. Rozkład ten nazywamy rozkładem chi-kwadrat o n stopniach swobody (ozn. χ_n^2).

13. Wektor losowy G o wartościach w $\mathbb{R}^n$ ma rozkład standardowy normalny, $G \sim \mathcal{N}(0, \mathrm{Id})$; A to pewna macierz $m \times n$, $b \in \mathbb{R}^m$. Jaką gęstość ma wektor losowy $AG + b$? Jaka jest jego macierz kowariancji?

14. Zmienne losowe X i Y są niezależne i mają rozkład $\mathcal{N}(0, 1)$. Jaki rozkład (łączny) ma wektor losowy $(X + Y, X - Y)$? Co można powiedzieć o jego współrzędnych?

15. a) Wektor losowy (X, Y) ma (łączny) rozkład gaussowski (dwuwymiarowy, ale niekoniecznie $\mathcal{N}(0, I)$). Wykazać, że jego współrzędne są niezależne wtedy i tylko wtedy gdy są nieskorelowane.

b) Niech X i r będą niezależne, $X \sim \mathcal{N}(0, 1)$, $\mathbb{P}(r = 1) = \frac{1}{2} = \mathbb{P}(r = -1)$; $Y = rX$. Sprawdzić, że $Y \sim \mathcal{N}(0, 1)$, X i Y są nieskorelowane, ale X i Y nie są niezależne, a także łączny rozkład wektora (X, Y) nie jest gaussowski.

16. Niech $U = (U_1, \ldots, U_n)$ będzie wektorem statystyk pozycyjnych dla rozkładu jednostajnego na $[0, 1]$: zmienne losowe $X_1, \ldots, X_n$ są niezależne i każda z nich ma rozkład jednostajny na $[0, 1]$,

$$\begin{aligned} U_1 &= \min\{X_1, \ldots, X_n\}, \\ U_k &= k\text{-ta najmniejsza wartość wśród } X_1, \ldots, X_n, \\ U_n &= \max\{X_1, \ldots, X_n\}. \end{aligned}$$

Wykazać, że U ma rozkład jednostajny na $\Delta_n = \{x \in \mathbb{R}^n : 0 \leqslant x_1 \leqslant \cdots \leqslant x_n \leqslant 1\}$. Wyznaczyć $\mathbb{E}U_k$.

17. Niech X, Y_n to niezależne zmienne losowe, takie że $X \sim \mathcal{N}(0, 1)$ i $Y_n \sim \chi_n^2$. Zdefiniujmy

$$T_n = \sqrt{n} \frac{X}{\sqrt{Y_n}}$$

– rozkład tej zmiennej to rozkład t-Studenta[1]. Jaka jest gęstość tego rozkładu? Dla jakich n istnieje jego wartość oczekiwana, a dla jakich wariancja?

[1] William Sealy Gosset (13.06.1876 – 16.09.1937), angielski statystyk, publikował pod pseudonimem Student.

RP dla bioinformatyki, seria VI

1. Wykazać, że rozkłady brzegowe rozkładu ciągłego są ciągłe. Wykazać, że odwrotna implikacja nie musi zachodzić.

2. Wektor losowy (X, Y) ma rozkład zadany przez

$$\mathbb{P}((X,Y) = (k,l)) = \frac{4kl}{n^2(n+1)^2}, \quad k, l = 1, \ldots, n.$$

a) Obliczyć $\mathbb{P}(X + Y = n + 1)$.
b) Wyznaczyć rozkład (brzegowy) zmiennej losowej X.
c) Czy zmienne losowe X, Y są niezależne?
d) Znaleźć macierz kowariancji wektora (X, Y). Obliczyć $\mathrm{Cov}(X+2Y+3, 4X+5Y+6)$.

3. Wektor losowy (X, Y) ma gęstość

$$g(x,y) = \begin{cases} Kxy & \text{dla } x > 0, y > 0, x^2 + y^2 \leqslant 1, \\ 0 & \text{w przeciwnym przypadku.} \end{cases}$$

a) Wyznaczyć stałą K i obliczyć $\mathbb{P}(X \leqslant Y)$, $\mathbb{P}(X + Y \leqslant 1)$.
b) Czy X i Y są niezależne?
c) Znaleźć gęstość X oraz gęstość Y.
d) Obliczyć $\mathrm{Cov}(X, Y)$.
e) Znaleźć rozkład zmiennej losowej $X^2 + Y^2$ i wyznaczyć jej gęstość (o ile istnieje).

4. Dysponujemy wynikami dwóch niezależnych pomiarów pewnej wielkości fizycznej. Błędy poszczególnych pomiarów to zmienne losowe X_1, X_2 o średniej 0 i wariancjach σ_1^2, σ_2^2 odpowiednio. Jako ostateczny wynik chcemy przyjąć średnią ważoną z wyników poszczególnych pomiarów. Jakie wagi należy przyjąć by błąd średniokwadratowy (czyli wariancja ostatecznego wyniku) był jak najmniejszy?

5. Zmienna losowa X jest nieujemna i całkowalna, a M jest (pewną) jej medianą, tzn.

$$\mathbb{P}(X \geqslant M) \geqslant 1/2, \qquad \mathbb{P}(X \leqslant M) \geqslant 1/2.$$

Wykazać, że $M \leqslant 2\mathbb{E}X$. Sformułować odpowiednią wersję tego zadania dla innych kwantyli.

6. a) Obliczyć $\mathbb{E}e^{\lambda G}$ dla $G \sim \mathcal{N}(0,1)$, $\lambda \in \mathbb{R}$.
b) Udowodnić, że $\mathbb{P}(G \geqslant t) \leqslant e^{-t^2/2}$ dla $t \geqslant 0$.
c) Niech $X \sim \mathcal{N}(a, \sigma)$. Sprawdzić, że $\mathbb{P}(a - 3\sigma \leqslant X \leqslant a + 3\sigma) \geqslant 0,97$.
d) Uzasadnić (najlepiej bez rachunków), że dla $t \geqslant 0$

$$\mathbb{P}(G \geqslant t) = \frac{1}{\sqrt{2\pi}} \int_t^\infty e^{-x^2/2} dx \leqslant \frac{1}{\sqrt{2\pi} t} e^{-t^2/2}$$

i wywnioskować, że wręcz $\mathbb{P}(a - 3\sigma \leqslant X \leqslant a + 3\sigma) \geqslant 0,997$.

RP dla bioinformatyki, seria VII

1. Pewien tekst liczy 100000 znaków. Podczas wprowadzania go do komputera prawdopodobieństwo iż dany znak zostanie wpisany błędnie wynosi $0,0001$.

a) Jakie jest w przybliżeniu prawdopodobieństwo, że w tekście są co najmniej 2 błędy? Podać oszacowanie błędu.

b) Tekst został sczytany przez redaktora, który każdy błąd wychwutuje z prawdopodobieństwem 90%. Jakie jest w przybliżeniu prawdopodobieństwo, że w tekście są co najmniej 2 błędy? Podać oszacowanie błędu.

2. Z odcinka $[0,2]$ wybieramy losowo 100 punktów. Co o prawdopodobieństwie tego, że co najmniej jeden z nich będzie należał do odcinka $[0,\frac{1}{4}]$ orzeka twierdzenie Poissona?

3. W pewnej fabryce wyrabiane śrubki podlegają statystycznej kontroli jakości. Prawdopodobieństwo, że wyprodukowana śrubka jest wadliwa wynosi $0,015$.

a) Pudełko zawiera 100 śrubek. Jakie jest prawdopodobieństwo, że wśród nich nie ma ani jednej wadliwej? Porównać dokładny wynik z przybliżeniem z twierdzenia Poissona.

b) Ile śrubek powinno zawierać pudełko, żeby prawdopodobieństwo znalezienia w nim co najmniej 100 dobrych śrubek było nie mniejsze niż $0,8$? Znaleźć odpowiedź stosując odpowiednie przybliżenia, a następnie wykonać sprawdzenie.

4. Pewien tekst został sprawdzony niezależnie przez dwóch korektorów. Pierwszy z nich znalazł 185 błędów, a drugi – 151. Błędów wskazanych przez obydwu było 98. Jak spisali się korektorzy?

5. Przeprowadzamy ankietę z wysoce krępującym pytaniem. W celu zapewnienia sobie niezakłamanych odpowiedzi, a ankietowanym – poczucia prywatności, prosimy ich o potajmne rzucenie kostką przed udzieleniem odpowiedzi:

- jeśli wypadnie 1 lub 2, mają odpowiedzieć „tak”,
- jeśli wypadnie 3 lub 4, mają odpowiedzieć „nie”,
- jeśli wypadnie 5 lub 6, mają odpowiedzieć zgodnie z prawdą na krępujące pytanie.

Czy z tak przeprowadzonej ankiety można wyciągnąć jakiekolwiek wnioski?

6. Zmienne losowe $X_1, X_2, \ldots$ są niezależne i mają ten sam rozkład: $\mathbb{P}(X_1 = 3) = \mathbb{P}(X_1 = -3) = \frac{1}{2}$. Obliczyć

$$\lim_{n\to\infty} \mathbb{P}(X_1 + \cdots + X_n \geqslant n),$$
$$\lim_{n\to\infty} \mathbb{P}(|X_1 + \cdots + X_n| \leqslant n^2).$$

7. a) Powtarzamy niezależnie pewne doświadczenie losowe i otrzymujemy próbkę, tj. ciąg $X_1, X_2, \ldots$ niezależnych zmiennych losowych o nieznanej nam dystrybuancie F. Niech

$$F_n(t) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\{X_i \leqslant t\}}, \qquad t \in \mathbb{R}.$$

Wykazać, że dla każdego $t \in \mathbb{R}$ mamy $\mathbb{P}(F_n(t) \to F(t)) = 1$.

b*) Wykazać, że mamy wręcz $\mathbb{P}(\sup_{t\in\mathbb{R}} |F_n(t) - F(t)| \to 0) = 1$.

8. Funkcja $f\colon [0,1] \to \mathbb{R}$ spełnia $\int_0^1 |f(x)|dx < \infty$. Zmienne losowe $X_1, X_2, \ldots$ są niezależne, mają ten sam rozkład o gęstości g i przyjmują wartości tylko w przedziale $[0,1]$. Zbadać zbieżność wyrażenia

$$\frac{1}{n}\sum_{i=1}^{n}\frac{f(X_i)}{g(X_i)}.$$

9. Obliczyć granice

$$\lim_{n\to\infty}\int_0^1\ldots\int_0^1 \frac{x_1^2+\cdots+x_n^2}{n}\,dx_1\ldots dx_n,$$
$$\lim_{n\to\infty}\int_0^1\ldots\int_0^1 \frac{x_1^3+\cdots+x_n^3}{x_1+\cdots+x_n}\,dx_1\ldots dx_n.$$

10. Masę urodzeniową noworodka można z powodzeniem przybliżać rozkładem normalnym.[1] Dla dziewczynek średnia waga to $\mu = 3,232$ kg, zaś wariancja to $\sigma^2 = (0,479 \text{ kg})^2$. Ala przy urodzeniu ważyła $3,800$ kg.

a) Jaki przeciętnie odsetek dziewcząt waży przy urodzeniu tyle lub więcej?

b) Ile musi ważyć nowonarodzona dziewczynka, żeby jej masa znalazła się w 95 centylu?

c) Jakie jest w przybliżeniu „prawdopodobieństwo" tego, że masa nowonarodzonej dziewczynki będzie ujemna?

d) Przyjmijmy, że zarówno masa nowonarodzonej dziewczynki, jak i masa nowonarodzonego chłopca mają *dokładnie* rozkład normalny (z różnymi parametrami). Czy możemy twierdzić, że masa nowonarodzonego dziecka też ma *dokładnie* rozkład normalny?

11. a) Jakie jest prawdopodobieństwo, że rzucając 10.000 razy symetryczną monetą uzyskamy więcej niż 5000 orłów?

b) Prawdopodobieństwo urodzenia się chłopca to $0,517$. Jakie jest prawdopodobieństwo, że wśród 10.000 noworodków liczba chłopców nie przewyższy liczby dziewcząt?

12. W pewnym stanie w wyborach prezydenckich głosuje 6.000.000 osób. Załóżmy, że wyborcy głosują na każdego z dwóch kandydatów losowo i niezależnie z prawdopodobieństwem 1/2. Jaka jest szansa, że różnica między kandydatami będzie nie większa niż 1000 głosów?

13. Stwierdzono, że przeciętnie 30% spośród ogólnej liczby studentów przyjętych na studia kończy je w terminie. Ile trzeba przyjąć studentów na pierwszy rok, aby z prawdopodobieństwem w przybliżeniu $0,9$, co najmniej 50 osób ukończyło studia w terminie?

14. a) Pewne zdarzenie ma prawdopodobieństwo $0,3$. Za pomocą dystrybuanty rozkładu normalnego oszacować liczbę doświadczeń, które trzeba wykonać, aby z prawdopodobieństwem $0,95$ można było stwierdzić, że częstość występowania interesującego nas zdarzenia będzie się odchylać od prawdziwego prawdopodobieństwa o nie więcej niż $0,02$.

b) Jak zmieni sie odpowiedź, jeśli nie znamy prawdziwego prawdopodobieństwa, ale wiemy, że jest ono pewną liczbą z przedziału $(0, \frac{1}{10})$? A gdy nic nie wiemy o wartości prawdziwego prawdopodobieństwa?

c) Porównać otrzymane wyniki z oszacowaniami wynikającymi z nierówności Czebyszewa.

---

[1] `http://www.who.int/childgrowth/standards/weight_for_age/en/`

15. Przypuśćmy, że aktualne poparcie dla partii politycznych to: 33% PiS, 26% PO, 7% SLD, 7% Kukiz '15, 4% PSL, [...].[2] Ankietujemy 1000 losowo wybranych osób.

a) Jakie jest w przybliżeniu prawdopodobieństwo, że w takiej ankiecie PiS osiągnie mniej niż 30%?

b) Jakie jest w przybliżeniu prawdopodobieństwo, że w takiej ankiecie poparcie dla PiS będzie o co najmniej 10 punktów procentowych wyższe niż poparcie dla PO?

c) Jakie jest prawdopodobieństwo, że poparcie dla SLD przekroczy 10% i jednocześnie poparcie dla Kukiz '15 spadnie poniżej progu wyborczego? Przedstawić wynik w postaci całki podwójnej po odpowiednim obszarze.

16. Zmienne $X_1, \ldots, X_n$ są niezależne i mają ten sam rozkład o średniej μ i wariancji 1. Liczba μ jest ustalona, ale nie znamy jej.

a) Na podstawie CTG zaproponować konstrukcję przedziału $[a, b]$, takiego że w przybliżeniu
$$\mathbb{P}(a \leqslant \mu \leqslant b) \geqslant 0,95.$$
Wartości a, b oczywiście mogą być losowe, tzn. zależeć od wartości $X_1, \ldots, X_n$.

b) Uświadomić sobie, że jeśli $X_1, \ldots, X_n$ są niezależne i mają rozkład $\mathcal{N}(\mu, 1)$, to w naszej konstrukcji nie dokonujemy żadnego przybliżenia oraz
$$\mathbb{P}(a \leqslant \mu \leqslant b) = 0,95.$$

c) Załóżmy, że $X_1, \ldots, X_n, Y_1, \ldots, Y_n$ są niezależne i mają rozkład $\mathcal{N}(\mu, 1)$, zaś $[a, b]$ to przedział skonstruowany w oparciu o próbkę $X_1, \ldots, X_n$. Czy
$$\mathbb{P}\Big(a \leqslant \frac{Y_1 + \cdots + Y_n}{n} \leqslant b\Big) \geqslant 0,95?$$

[2] Dane za badaniem Millward Brown z 20.11.2018.

Dystrybuanta rozkładu normalnego: $\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-x^2/2} dx \overset{z \geq 0}{=} \frac{1}{2}\Big(1 + \mathrm{erf}\big(\frac{z}{\sqrt{2}}\big)\Big)$

z	**+0.00**	**+0.01**	**+0.02**	**+0.03**	**+0.04**	**+0.05**	**+0.06**	**+0.07**	**+0.08**	**+0.09**
0.0	0.50000	0.50399	0.50798	0.51197	0.51595	0.51994	0.52392	0.52790	0.53188	0.53586
0.1	0.53983	0.54380	0.54776	0.55172	0.55567	0.55966	0.56360	0.56749	0.57142	0.57535
0.2	0.57926	0.58317	0.58706	0.59095	0.59483	0.59871	0.60257	0.60642	0.61026	0.61409
0.3	0.61791	0.62172	0.62552	0.62930	0.63307	0.63683	0.64058	0.64431	0.64803	0.65173
0.4	0.65542	0.65910	0.66276	0.66640	0.67003	0.67364	0.67724	0.68082	0.68439	0.68793
0.5	0.69146	0.69497	0.69847	0.70194	0.70540	0.70884	0.71226	0.71566	0.71904	0.72240
0.6	0.72575	0.72907	0.73237	0.73565	0.73891	0.74215	0.74537	0.74857	0.75175	0.75490
0.7	0.75804	0.76115	0.76424	0.76730	0.77035	0.77337	0.77637	0.77935	0.78230	0.78524
0.8	0.78814	0.79103	0.79389	0.79673	0.79955	0.80234	0.80511	0.80785	0.81057	0.81327
0.9	0.81594	0.81859	0.82121	0.82381	0.82639	0.82894	0.83147	0.83398	0.83646	0.83891
1.0	0.84134	0.84375	0.84614	0.84849	0.85083	0.85314	0.85543	0.85769	0.85993	0.86214
1.1	0.86433	0.86650	0.86864	0.87076	0.87286	0.87493	0.87698	0.87900	0.88100	0.88298
1.2	0.88493	0.88686	0.88877	0.89065	0.89251	0.89435	0.89617	0.89796	0.89973	0.90147
1.3	0.90320	0.90490	0.90658	0.90824	0.90988	0.91149	0.91308	0.91466	0.91621	0.91774
1.4	0.91924	0.92073	0.92220	0.92364	0.92507	0.92647	0.92785	0.92922	0.93056	0.93189
1.5	0.93319	0.93448	0.93574	0.93699	0.93822	0.93943	0.94062	0.94179	0.94295	0.94408
1.6	0.94520	0.94630	0.94738	0.94845	0.94950	0.95053	0.95154	0.95254	0.95352	0.95449
1.7	0.95543	0.95637	0.95728	0.95818	0.95907	0.95994	0.96080	0.96164	0.96246	0.96327
1.8	0.96407	0.96485	0.96562	0.96638	0.96712	0.96784	0.96856	0.96926	0.96995	0.97062
1.9	0.97128	0.97193	0.97257	0.97320	0.97381	0.97441	0.97500	0.97558	0.97615	0.97670
2.0	0.97725	0.97778	0.97831	0.97882	0.97932	0.97982	0.98030	0.98077	0.98124	0.98169
2.1	0.98214	0.98257	0.98300	0.98341	0.98382	0.98422	0.98461	0.98500	0.98537	0.98574
2.2	0.98610	0.98645	0.98679	0.98713	0.98745	0.98778	0.98809	0.98840	0.98870	0.98899
2.3	0.98928	0.98956	0.98983	0.99010	0.99036	0.99061	0.99086	0.99111	0.99134	0.99158
2.4	0.99180	0.99202	0.99224	0.99245	0.99266	0.99286	0.99305	0.99324	0.99343	0.99361
2.5	0.99379	0.99396	0.99413	0.99430	0.99446	0.99461	0.99477	0.99492	0.99506	0.99520
2.6	0.99534	0.99547	0.99560	0.99573	0.99585	0.99598	0.99609	0.99621	0.99632	0.99643
2.7	0.99653	0.99664	0.99674	0.99683	0.99693	0.99702	0.99711	0.99720	0.99728	0.99736
2.8	0.99744	0.99752	0.99760	0.99767	0.99774	0.99781	0.99788	0.99795	0.99801	0.99807
2.9	0.99813	0.99819	0.99825	0.99831	0.99836	0.99841	0.99846	0.99851	0.99856	0.99861
3.0	0.99865	0.99869	0.99874	0.99878	0.99882	0.99886	0.99889	0.99893	0.99896	0.99900
3.1	0.99903	0.99906	0.99910	0.99913	0.99916	0.99918	0.99921	0.99924	0.99926	0.99929
3.2	0.99931	0.99934	0.99936	0.99938	0.99940	0.99942	0.99944	0.99946	0.99948	0.99950
3.3	0.99952	0.99953	0.99955	0.99957	0.99958	0.99960	0.99961	0.99962	0.99964	0.99965
3.4	0.99966	0.99968	0.99969	0.99970	0.99971	0.99972	0.99973	0.99974	0.99975	0.99976
3.5	0.99977	0.99978	0.99978	0.99979	0.99980	0.99981	0.99981	0.99982	0.99983	0.99983
3.6	0.99984	0.99985	0.99985	0.99986	0.99986	0.99987	0.99987	0.99988	0.99988	0.99989
3.7	0.99989	0.99990	0.99990	0.99990	0.99991	0.99991	0.99992	0.99992	0.99992	0.99992
3.8	0.99993	0.99993	0.99993	0.99994	0.99994	0.99994	0.99994	0.99995	0.99995	0.99995
3.9	0.99995	0.99995	0.99996	0.99996	0.99996	0.99996	0.99996	0.99996	0.99997	0.99997
4.0	0.99997	0.99997	0.99997	0.99997	0.99997	0.99997	0.99998	0.99998	0.99998	0.99998

Kilka kwantyli: $u_p = \Phi^{-1}(p),\ p \in (0,1)$

p	0.5	0.75	0.9	0.95	0.975	0.99	0.995
u_p	0	0,67449	1.28155	1.64485	1.95996	2.32634	2.57583

RP dla bioinformatyki, seria VIII

1. Bolek i Lolek grają w szachy do dwóch wygranych partii pod rząd.

a) Jakie jest prawdopodobieństwo, że wygra Lolek, jeśli prawdopodobieństwo wygrania przez niego pojedynczej partii wynosi p? Zakładamy, że wyniki poszczególnych partii są niezależne (i że nie ma remisów).

b) Ilu partii średnio potrzeba, aby wyłonić zwycięzcę?

c) Przyjmijmy, że $p = 1/2$. Jakie jest prawdopodobieństwo, że do wyłonienia zwycięzcy potrzeba więcej niż 5 partii?

2. Bolek i Lolek grają w szachy do dwóch wygranych partii pod rząd. Prawdopodobieństwo, że Bolek wygra pierwszą partię wynosi 1/2, ale począwszy od drugiej partii, prawdopodobieństwo wygranej Bolka zależy od wyniku poprzedniej partii. Jeśli dopiero co wygrał, to czuje się pewnie i wygrywa z prawdopodobieństwem 3/4; jeśli przegrał, to zżera go stres i wygrywa z prawdopodobieństwem 1/3.

a) Jakie jest prawdopodobieństwo, że Bolek wygra całą rozgrywkę?

b) Ilu partii średnio potrzeba, aby wyłonić zwycięzcę?

3. Rzucamy monetą aż pojawi się ciąg OOO lub ROO. Jakie jest prawdopodobieństwo, że ciąg ROO pojawi sie wcześniej?

4. a) Wypisać macierz przejścia dla błądzenia losowego po $\mathbb{Z}$ z dwoma ekranami pochłaniającymi. Jak należy ją zmodyfikować, żeby otrzymać macierz przejścia dla błądzenia z ekranami elastycznymi?

b) Wypisać macierz przejścia dla długości nieprzerwanej serii sukcesów w ciągu prób Bernoulliego.

5. Zmienne $r_0, r_1, \ldots$ są niezależne i mają ten sam rozkład: $\mathbb{P}(r_0 = 1) = 1/2 = \mathbb{P}(r_0 = -1)$.

a) Niech $X_n = r_n r_{n+1}$ dla $n = 0, 1, 2, \ldots$ – czy jest to łańcuch Markowa?

b) Niech $Y_n = \frac{r_n + r_{n+1}}{2}$ dla $n = 0, 1, 2, \ldots$ – czy jest to łańcuch Markowa?

6. Rozważmy przestrzeń stanów $S = \{0, 1, 2\}$ i niech $\mathbb{P}(X_0 = 0) = \mathbb{P}(X_0 = 1) = \mathbb{P}(X_0 = 2) = 1/3$ oraz $X_{n+1} = (X_n + 1)(\mathrm{mod}\, 3)$ dla $n = 0, 1, 2, \ldots$; ponadto niech $f(0) = 0$, $f(1) = f(2) = 1$. Czy $(X_n)_{n \geqslant 0}$ to łańcuch Markowa? A $(f(X_n))_{n \geqslant 0}$?

7. Przy okrągłym stole stole siedzi n wikingów i ich wódz. Wódz puszcza w obieg beczkę z piwem. Każdy z wikingów robi to samo: wypija z beczki solidny łyk, a następnie przekazuje beczkę dalej do kolejnego wikinga lub podaje ją z powrotem do tego wikinga, od którego ją otrzymał. Przyjmijmy, że dla i-tego wikinga prawdopodobieństwo podania beczki dalej to $p_i \in (0, 1)$, zaś prawdopodobieństwo „odbicia" i zmienienia kierunku obiegu beczki to $1 - p_i$. Jakie jest prawdopodobieństwo, że beczka wróci do wodza z innej strony niż ta, w którą ją pierwotnie podał?

8. Raz na godzinę ameba albo dzieli się na dwie ameby (to dzieje się z prawdopodobieństwem p), albo umiera (z prawdopodobieństwem $1 - p$). Jakie jest prawdopodobieństwo, że populacja startującą z jednej ameby nigdy nie wyginie?

9. W urnie jest 1000 ponumerowanych kul. Losujemy kulę, po czym usuwamy ją z urny wraz ze wszystkimi kulami o numerach większych od niej. Doświadczenie powtarzamy dopóty, dopóki w urnie są kule. Jaka jest wartość oczekiwana liczby wykonanych losowań?

10. W każdym z dwóch pudełek jest k kul, przy czym łącznie jest k kul białych i k – czarnych. Stan układu jest opisany przez podanie liczby kul białych w pierwszym pudełku. Powtarzamy następującą procedurę: z każdego pudłeka losujemy po jednej kuli, a następnie zamieniamy je miejscami.

a) Wypisać macierz przejścia dla tego łańcucha Markowa.

b) Uzasadnić, że macierz przejścia tego łańcucha Markowa ma dokładnie jeden rozkład stacjonarny: $\tilde{p}_k$ równa się prawdopodobieństwu otrzymania dokładnie k kul białych przy losowaniu N kul spośród N kul białych i N kul czarnych.

11. Jednorodny łańcuch Markowa na przestrzeni stanów $\{1, 2, 3, 4\}$ ma macierz przejścia

$$\begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{4} & \frac{1}{2} & 0 & \frac{1}{4} \\ \frac{2}{3} & 0 & 0 & \frac{1}{3} \\ 0 & \frac{2}{3} & \frac{1}{3} & 0 \end{bmatrix}.$$

a) Jakie jest prawodpoodbieństwo przejścia w dwóch krokach ze stanu 1 do 2? A w trzech? A w 10?

b) Jakie jest prawdopodobieństwo odwiedzenia stanu 2 przed stanem 4, jeśli łańcuch startuje ze stanu 1?

c) Znaleźć wszystkie rozkłady stacjonarne.

d) Obliczyć przybliżone prawdopodobieństwo tego, że $X_{1000} = 1$.

12. a) W wierzchołku sześcianu siedzi mucha. Co minutę mucha przechodzi do losowo wybranego z sąsiednich wierzchołków. Wypisać macierz przejścia i znaleźć wszystkie rozkłady stacjonarne. Co można powiedzieć o zbieżności do nich?

b) W wierzchołku sześcianu siedzi leniwa mucha. Co minutę mucha z prawdodpobieństwem $p \in (0, 1)$ pozostaje w aktualnie zajmowanym wierzchołku, z prawdopodobieństwem $1 - p$ przechodzi do losowo wybranego z sąsiednich wierzchołków. Jakie są rozkłady stacjonarne macierzy przejścia i co można powiedzieć o zbieżności do nich?

13. W pudełku A jest sześć kul ponumerowanych liczbami od 1 do 6, pudełko B jest puste. Rzucono 10000 razy kostką i za każdym razem przekładano kulę z wylosowanym numerem do drugiego pudełka. Jaka jest mniej więcej szansa, że pudełko B jest puste?

14. Wykazać, że w nieprzywiedlnym łańcuchu Markowa wszystkie stany mają ten sam okres.

15. Jednorodny w czasie łańcuch Markowa $(X_t)_{t=0}^{\infty}$ (na skończonej przestrzeni stanów) spełnia $X_0 \sim X_{20}$ oraz $X_3 \sim X_{22}$. Czy $X_0 \sim X_{2019}$?

16. a) Rozważmy błądzenie losowe po $\mathbb{Z}$: z dowolnego stanu j przechodzimy do $j + 1$ z prawdopodobieństwem p, a do $j - 1$ z prawdopodobieństwem $1 - p$. Obliczyć $p_{00}^{(k)}$ – prawdopodobieństwo przejścia z 0 do 0 w k krokach. Czy zero jest stanem chwilowym czy powracającym?

b) Rozważyć analogiczny problem dla (symetrycznego) błądzenia losowego po $\mathbb{Z}^d$.