Kolokwium z Rachunku Prawdopodobieństwa, 1 grudnia 2012r. – rozwiązania

Wersja I

1. Z talii 52 kart ciągniemy 3 karty (bez zwracania). Jakie jest prawdopodobieństwo, że wyciągnięte karty są trzech różnych kolorów, jeśli wiadomo, że wśród wyciągniętych kart są tylko karty numerowane (od 2 do 10)? (*4 pkt*)

   Rozw. Jako zbiór zdarzeń elementarnych przyjmujemy 3-elementowe podzbiory zbioru wszystkich 52 kart (tj. zdarzeniem elementarnym jest zbiór 3 wyciągniętych kart), wobec czego $|\Omega| = \binom{52}{3}$. *Przyjmujemy, że wszystkie zdarzenia elementarne są jednakowo prawdopodobne. Niech* A *będzie zdarzeniem polegającym na wyciągnięciu kart trzech różnych kolorów, natomiast* B *— zdarzeniem polegającym na wyciągnięciu kart numerowanych. W zadaniu mamy obliczyć p-stwo warunkowe* $\mathbb{P}(A|B)$. *Liczymy:*

$$\mathbb{P}(A \cap B) = \frac{\binom{4}{3}9^3}{\binom{52}{3}}, \qquad \mathbb{P}(B) = \frac{\binom{36}{3}}{\binom{52}{3}}, \qquad \mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\binom{4}{3}9^3}{\binom{36}{3}} = \frac{243}{595}.$$

2. W urnie znajduje się 5 kostek do gry, przy czym 3 kostki są prawidłowe, jedna kostka ma dokładnie dwie ściany z sześcioma oczkami (a na pozostałych ścianach od 1 do 4 oczek) i jedna kostka ma sześć oczek na wszystkich ścianach.

   a) Losujemy kostkę z urny, rzucamy nią i wypada szóstka. Jakie jest prawdopodobieństwo, że kostka którą wylosowaliśmy była prawidłowa? (*2 pkt*)

   Rozw. Przez K_1, K_2, K_6 *oznaczmy odpowiednio zdarzenia polegające na wyciągnięciu z urny kostki prawidłowej, kostki o dwóch ścianach z sześcioma oczkami i kostki, której wszystkie ściany mają 6 oczek. Niech* A *oznacza zdarzenie, że na wyciągniętej z urny kostce wypadło 6 oczek. W zadaniu należy policzyć p-stwo warunkowe* $\mathbb{P}(K_1|A)$. *Z danych zadania wiemy, że* $\mathbb{P}(K_1) = 3/5$, $\mathbb{P}(K_2) = 1/5$, $\mathbb{P}(K_3) = 1/5$, *oraz* $\mathbb{P}(A|K_1) = 1/6$, $\mathbb{P}(A|K_2) = 1/3$, $\mathbb{P}(A|K_6) = 1$. *Ze wzoru na p-stwo całkowite,*

$$\mathbb{P}(A) = \mathbb{P}(A|K_1)\mathbb{P}(K_1) + \mathbb{P}(A|K_2)\mathbb{P}(K_2) + \mathbb{P}(A|K_6)\mathbb{P}(K_6) = \frac{11}{30},$$

   zatem (wzór Bayesa) $\mathbb{P}(K_1|A) = \frac{\mathbb{P}(A|K_1)\mathbb{P}(K_1)}{\mathbb{P}(A)} = \frac{3}{11}$.

   b) Losujemy kostkę z urny, rzucamy nią i wypada szóstka. Jakie jest prawdopodobieństwo, że przy kolejnych dwóch rzutach tą samą kostką, raz wypadnie szóstka a raz jakaś inna liczba oczek? (*2 pkt*)

   Rozw. Niech K_1, K_2, K_6 *jak wyżej, podobnie niech* A *oznacza zdarzenie, że po pierwszym rzucie wyciągniętą z urny kostką wypadło 6 oczek, natomiast* B *oznacza zdarzenie, że w dwóch kolejnych rzutach raz wypadła szóstka a raz jakaś inna liczba oczek. Chcemy policzyć* $\mathbb{P}(B|A) = \frac{\mathbb{P}(A\cap B)}{\mathbb{P}(A)}$. *Z danych zadania wiemy, że* $\mathbb{P}(A \cap B|K_1) = 10/6^3$, $\mathbb{P}(A \cap B|K_2) = 4/27$, $\mathbb{P}(A \cap B|K_3) = 0$. *Ze wzoru na p-stwo całkowite*

$$\mathbb{P}(A \cap B) = \mathbb{P}(A \cap B|K_1)\mathbb{P}(K_1) + \mathbb{P}(A \cap B|K_2)\mathbb{P}(K_2) + \mathbb{P}(A \cap B|K_6)\mathbb{P}(K_6) = \frac{31}{540},$$

   zatem $\mathbb{P}(B|A) = \frac{31/540}{11/30} = \frac{31}{198}$. *(Można też rozumować nieco inaczej: korzystając z obliczeń punktu a) i stosując wzór Bayesa dostajemy* $\mathbb{P}(K_2|A) = \frac{2}{11}$ *oraz* $\mathbb{P}(K_3|A) = \frac{6}{11}$. *Ograniczając przestrzeń probabilistyczną do zdarzenia* A *i rozpatrując warunkowe p-stwo* $\mathbb{P}_A$ *zdefiniowane wzorem* $\mathbb{P}_A(C) = \mathbb{P}(C|A)$ *dla dowolnego zdarzenia* $C \subseteq \Omega$, *stosujemy wzór na p-stwo całkowite do szukanego* $\mathbb{P}_A(B)$:

$$\mathbb{P}_A(B) = \mathbb{P}_A(B|K_1)\mathbb{P}_A(K_1) + \mathbb{P}_A(B|K_2)\mathbb{P}_A(K_2) + \mathbb{P}_A(B|K_6)\mathbb{P}_A(K_6) = \frac{10}{36}\frac{3}{11} + \frac{4}{9}\frac{2}{11} + 0\frac{6}{11} = \frac{31}{198}.)$$

3. System bankowy błędnie księguje kwoty średnio 2 na 10^7 przelewów. Korzystając z przybliżenia Poissona, oszacować prawdopodobieństwo, że wśród 10^6 przelewów, które miały miejsce w pewnym miesiącu, błędnie zaksięgowane zostały co najmniej dwa przelewy (trafiły na konto pewnego polityka). (*3 pkt*) Oszacować błąd przybliżenia. (*1 pkt*)

Wsk.: $e^2 \approx 7,39$; $e^{-1/2} = 0,61$; $e^{-1/5} \approx 0,82$; $e^{1/5} \approx 1,22$

Rozw. $n = 10^6, p = 2 \cdot 10^{-7}, \lambda = np = 2/10 = 1/5$. *Z tw. Poissona p-stwo błędnego zaksięgowania dokładnie k przelewów w przybliżeniu wynosi* $\frac{\lambda^k}{k!}e^{-\lambda}$. *P-stwo zdarzenia przeciwnego do zdarzenia opisanego w zadaniu wynosi w przybliżeniu*

$$e^{-\lambda} + \lambda e^{-\lambda} = \frac{6}{5}e^{-1/5} \approx \frac{6}{5} \cdot 0,82 = 0,984,$$

zatem p-stwo błędnego zaksięgowania co najmniej 2 przelewów wynosi w przybliżeniu $0,016$. *Błąd przybliżenia wynikający z zastosowania tw. Poissona (błąd samego przybliżenia* $e^{-\lambda}$ *pomijamy) jest nie większy niż* $\lambda^2/n = 4 \cdot 10^{-8}$.

4. Zmienna losowa X spełnia następujące warunki: $\mathbb{P}(X=1) = \frac{1}{10}$, $\mathbb{P}(X=3) = \frac{2}{10}$, $\mathbb{P}(X \in (3,t)) = \frac{t-3}{10}$ dla $t \in (3,7]$ oraz $\mathbb{P}(X=7) = \frac{3}{10}$.

   a) Wyznacz dystrybuantę X. (*2 pkt*)
   Rozw.

$$F_X(t) = \begin{cases} 0 & \text{dla } t \in (-\infty, 1) \\ \frac{1}{10} & \text{dla } t \in [1,3) \\ \frac{3}{10} + \frac{t-3}{10} = \frac{t}{10} & \text{dla } t \in [3,7) \\ 1 & \text{dla } t \in [7, +\infty). \end{cases}$$

   b) Czy X ma rozkład ciągły? Czy X ma rozkład dyskretny? Odpowiedzi uzasadnij! (*2 pkt*)
   Rozw. Rozkład X nie jest ciągły, ponieważ $\mathbb{P}(X=1) > 0$. *Rozkład X nie jest też dyskretny, jako że jest skupiony na zbiorze* $\{1\} \cup [3,7]$, *które nie jest ani skończony ani nawet przeliczalny.*

   c) Oblicz $\mathbb{E}X$ oraz Var$X$. (*4 pkt*)
   Rozw. Zmienna X jest nieujemna, więc stosujemy wzór

$$\mathbb{E}X = \int_0^\infty \mathbb{P}(X>t)\,dt = \int_0^\infty 1 - F_X(t)\,dt = \int_0^1 1\,dt + \int_1^3 \frac{9}{10}\,dt + \int_3^7 (1 - \frac{t}{10})\,dt = 1 + \frac{9}{5} + 2 = \frac{24}{5}.$$

   Podobnie obliczamy

$$\mathbb{E}X^2 = \int_0^\infty 2t\mathbb{P}(X>t)\,dt = 2\int_0^\infty t(1-F_X(t))\,dt = 2\left(\int_0^1 t\,dt + \int_1^3 \frac{9}{10}t\,dt + \int_3^7 t(1-\frac{t}{10})\,dt\right)$$
$$= 2\left(\frac{1}{2} + \frac{18}{5} + \frac{142}{15}\right) = \frac{407}{15},$$

   zatem $VarX = \mathbb{E}X^2 - (\mathbb{E}X)^2 = \frac{407}{15} - \frac{576}{25} = \frac{307}{75}$.

5. Zmienna losowa X ma rozkład ciągły z gęstością $g(x) = \frac{c}{x^2}\mathbf{1}_{(4,\infty)}(x)$ oraz $Y = \sqrt{X}$.

   a) Oblicz c. (*2 pkt*)
   Rozw. Żeby proponowana funkcja byłą funkcją gęstości, należy sprawdzić, że $\int_{\mathbb{R}} g(x)dx = 1$ *oraz, że* $g(x) \geqslant 0$ *p.w. Pierwszy warunek oznacza, że*

$$1 = \int_{\mathbb{R}} \frac{c}{x^2}\mathbf{1}_{(4,\infty)}(x)dx = c\int_4^\infty \frac{1}{x^2}dx = c\left[-\frac{1}{x}\right|_4^\infty = c(-0 + \frac{1}{4}) = \frac{c}{4},$$

   a zatem $c = 4$. *Dla* $c = 4$ *mamy* $g(x) \geqslant 0$, *więc jest to poprawny wzór na funkcję gęstości.*

   b) Czy Y ma rozkład ciągły? Jeśli tak, to oblicz gęstość Y, w przeciwnym razie uzasadnij, że Y nie ma rozkładu ciągłego. (*3 pkt*)
   Rozw. Są dwie podstawowe metody rozwiązania: jedna wykorzystuje rachunki na dystrybuancie, druga twierdzenie o przekształcaniu gęstości.
   W pierwszym przypadku: wyznaczymy dystrybuantę zmiennej Y.

$$F_Y(t) = \mathbb{P}(Y \leqslant t) = \mathbb{P}(\sqrt{X} \leqslant t) = \mathbb{P}(X \leqslant t^2),$$

dla $t \geqslant 0$ (dla $t < 0$ nie jest możliwe, żeby pierwiastek był od takiego t mniejszy). A zatem, dla $t > 0$:

$$F_Y(t) = \int_{-\infty}^{t^2} \frac{4}{x^2} \mathbf{1}_{(4,\infty)}(x)dx = \begin{cases} \left[-\frac{4}{x}\right|_4^{t^2} & t^2 \geqslant 4 \\ 0 & t^2 < 4 \end{cases} = \begin{cases} 1 - \frac{4}{t^2} & t \geqslant 2 \\ 0 & t < 2 \end{cases},$$

co można połączyć z przypadkiem $t < 0$ bez zmiany wzoru. Dystrybuanta jest ciągła wszędzie (łącznie z punktem 2) i jest różniczkowalna wszędzie poza 2. Można więc wyznaczyć gęstość (a zatem zmienna Y jest ciągła), jako pochodną z dystrybuanty z dodaną zerową wartością w 2:

$$g(t) = \frac{8}{t^3}\mathbf{1}_{(2,\infty)}(t).$$

W drugim przypadku: zauważamy, że $Y = \varphi(X)$, gdzie $\varphi(x) = \sqrt{x}$. $\varphi'(x) = \frac{1}{2\sqrt{x}} \neq 0$ dla $x > 2$, a zatem przekształcenie to spełnia założenia odp. twierdzenia, i mamy:
a) zmienna Y jest ciągła, oraz
b) gęstość Y jest równa

$$g_Y(y) = g_X(\varphi^{-1}(y))|\frac{d\varphi^{-1}(y)}{dy}|\mathbf{1}_{(\varphi(4),\varphi(\infty))}(y) = \frac{4}{(y^2)^2}|2y|\mathbf{1}_{(2,\infty)}(y) = \frac{8}{y^3}\mathbf{1}_{(2,\infty)}(y).$$

c) Rozstrzygnij, czy X jest całkowalna. Jeśli tak, to oblicz $\mathbb{E}X$. (*2 pkt*)
Rozw. X jest całkowalna, jeśli $\mathbb{E}|X| = \int_{\mathbb{R}} |x|g(x)dx < \infty$. W naszym przypadku,

$$\mathbb{E}|X| = \int_{\mathbb{R}} |x|\frac{4}{x^2}\mathbf{1}_{(4,\infty)}(x)dx = \int_4^{\infty} x\frac{4}{x^2}dx = \int_4^{\infty} \frac{4}{x}dx = [4\ln x|_4^{\infty} = \infty,$$

a zatem zmienna X NIE jest całkowalna

d) Rozstrzygnij, czy Y jest całkowalna. Jeśli tak, to oblicz $\mathbb{E}Y$. (*2 pkt*)
Rozw. Y jest całkowalna, jeśli $\mathbb{E}|Y| = \int_{\mathbb{R}} |y|g(y)dy < \infty$ lub równoważnie $\mathbb{E}|Y| = \mathbb{E}|\sqrt{X}| = \mathbb{E}\sqrt{X} = \int_{\mathbb{R}} \sqrt{x}g(x)dx < \infty$. W naszym przypadku,

$$\mathbb{E}\sqrt{X} = \int_{\mathbb{R}} \sqrt{x}\frac{4}{x^2}\mathbf{1}_{(4,\infty)}(x)dx = \int_4^{\infty} x\frac{4}{x^2}dx = \int_4^{\infty} \frac{4}{x^{3/2}}dx = 4\left[-2x^{-1/2}\right|_4^{\infty} = 4,$$

a zatem zmienna Y jest całkowalna oraz $\mathbb{E}Y = \mathbb{E}|Y| = 4$ (bo Y jest nieujemna).

6. Zmienna losowa X ma rozkład jednostajny na odcinku $[-1, 2]$ zaś $Y = X^2 + 2$. Oblicz $\mathbb{E}Y$ (*2 pkt*), $\mathrm{Var}Y$ (*3 pkt*), a także $\mathbb{P}(Y \leqslant 1/4)$ oraz $\mathbb{P}(Y > 1)$. (*1 pkt*)
Rozw. Skoro X ma rozkład jednostajny, to możemy skorzystać ze znanych własności by stwierdzić, iż $g(x) = \frac{1}{3}\mathbf{1}[-1, 2]$, zaś $\mathbb{E}X = \frac{-1+2}{2} = \frac{1}{2}$ oraz $VarX = \frac{(2-(-1))^2}{12} = \frac{9}{12} = \frac{3}{4}$. W związku z tym, $\mathbb{E}X^2 = \frac{3}{4} + \left(\frac{1}{2}\right)^2 = 1$. Mamy wyznaczyć:
(a) $\mathbb{E}Y = \mathbb{E}(X^2 + 5) = \mathbb{E}X^2 + 2 = 1 + 2 = 3$, oraz
(b) $VarY = Var(X^2+2) = VarX^2 = \mathbb{E}(X^2)^2 - (\mathbb{E}X^2)^2 = \mathbb{E}X^4 - 1^2$. Wyznaczymy zatem brakujące

$$\mathbb{E}Y^4 = \int_{\mathbb{R}} x^4 g_X(x)dx = \int_{-1}^{2} x^4\frac{1}{3}dx = \left[\frac{x^5}{15}\right|_{-1}^{2} = \frac{32}{15} + \frac{1}{15} = \frac{33}{15},$$

a zatem $VarY = \frac{33}{15} - 1 = \frac{18}{15}$.
(c) Ponieważ $Y \geqslant 2$, więc $\mathbb{P}(Y \leqslant \frac{7}{4}) = 0$. Mamy również $\mathbb{P}(Y > 3) = \mathbb{P}(X^2 + 2 > 3) = \mathbb{P}(X^2 > 1) = \mathbb{P}(|X| > 1)$, co u nas $= \mathbb{P}(X \in (1, 2]) = \frac{1}{3}$.

7. Przeanalizowano dane pewnego urzędu pracy by stwierdzić, jaki czas upływa od zakończenia pracy do zarejestrowania w urzędzie. W wylosowanej próbce 10 bezrobotnych, czas ten prezentował się następująco (w dniach):

$$1, 1, 2, 2, 2, 3, 3, 4, 5, 7.$$

Wyznaczyć dystrybuantę empiryczną (*2 pkt*), średnią i wariancję z próby (*2 pkt*), oraz pierwszy kwartyl z próby (*1 pkt*).
Rozw. Dystrybuanta empiryczna równa jest

$$F_{10}(t) = \begin{cases} 0 & t < 1 \\ \frac{2}{10} & t \in [1,2) \\ \frac{5}{10} & t \in [2,3) \\ \frac{7}{10} & t \in [3,4) \\ \frac{8}{10} & t \in [4,5) \\ \frac{9}{10} & t \in [5,7) \\ 1 & t \geqslant 7 \end{cases}.$$

Średnia z próby wynosi

$$m = \frac{1 \cdot 2 + 2 \cdot 3 + 3 \cdot 2 + 4 + 5 + 7}{10} = \frac{30}{10} = 3,$$

zaś wariancja z próby

$$\begin{aligned} s^2 &= \frac{(1-3)^2 \cdot 2 + (2-3)^2 \cdot 3 + (3-3)^2 \cdot 2 + (4-3)^2 + (5-3)^2 + (7-3)^2}{10} \\ &= \frac{8+3+0+1+4+16}{10} = 3,2. \end{aligned}$$

Pierwszy kwartyl z próby wynosi zaś 2 (bo 2 jest jedyną wartością spełniającą warunki $\mathbb{P}(X \leqslant 2) \geqslant \frac{1}{4}$ *oraz* $\mathbb{P}(X \geqslant 2) \geqslant \frac{3}{4}$*).*