Rachunek prawdopodobieństwa i statystyka

Lista zadań nr 3. Tydzień rozpoczynający się 23. marca

Zadania

1. A oraz B są zdarzeniami takimi, że: $P(A \cap B) = 1/4$, $P(A^C) = 1/3$, $P(B) = 1/2$. Znaleźć $P(A \cup B)$.

   $P(A) = {}^2\!/_3$ czyli $P(A \cap B) = P(A) + P(B) - P(A \cup B) = {}^2\!/_3 + {}^1\!/_2 - {}^1\!/_4 = {}^{11}\!/_{12}$.

---

2. Czy prawdą jest, że 13. dzień miesiąca powiązany jest z piątkiem? (1 stycznia 1601 – 31 grudnia 2000)
   Założenia: rok numer n jest jest przestępny jeżeli $n \equiv_4 0$, pod warunkiem, że $n \not\equiv_{100} 0$; dodatkowo – jeżeli $n \equiv_{400} 0$ (czyli rok 2000), to wcześniejszy warunek jest nieważny. Ile razy w 400-letnim cyklu 13-tym dniem miesiąca był poniedziałek, wtorek, ..., niedziela?

   Np. w arkuszu kalkulacyjnym

pon	wt	śr	czw	pt	sb	nd
685	685	687	684	**688**	684	687

Mówimy, że zmienne X, Y są niezależne, wtedy gdy – w wypadku dyskretnym – spełniony jest warunek $P(X = x_i, Y = y_k) = P(X = x_i) \cdot P(Y = y_k)$.

3. Zmienna X ma rozkład $B(n_1, p)$ a zmienna Y rozkład $B(n_2, p)$. Zmienne są niezależne. Wykazać, że zmienna $Z = X + Y$ ma rozkład $B(n_1 + n_2, p)$.

   $P(Z = k) = P(X + Y = k) = \sum_{i=0}^{k} P(X = i,\ Y = k - i)$, bo zdarzenia pod sumą są rozłączne. Następnie: $= \sum_{i=0}^{k} \binom{n_1}{i} p^i q^{n_1 - i} \binom{n_2}{k-i} p^{k-i} q^{n_2 - k + i} = p^k q^{n_1 + n_2 - k} \sum_{i=0}^{k} \binom{n_1}{i} \binom{n_2}{k-i}$. Wyrażenie pod sumą to – ze wzoru dwumianowego Newtona – $\binom{n_1 + n_2}{k}$ co daje tezę zadania.

---

4. Niezależne zmienne losowe X, Y mają rozkład Poissona z parametrami λ_1 i λ_2. Wykazać, że zmienna $Z = X + Y$ ma rozkład Poissona z parametrem $\lambda_1 + \lambda_2$.

   $P(Z = k) = P(X + Y = k) = \sum_{i=0}^{k} P(X = i,\ Y = k - i) = (*)$, bo zdarzenia pod sumą są rozłączne. Następnie – z niezależności – $P(X = i,\ Y = k - i) = P(X = i) \cdot P(Y = k - i)$ czyli otrzymujemy wyrażenie $(*) = \sum_{i=0}^{k} \mathrm{e}^{-\lambda_1} \frac{\lambda_1^i}{i!} \cdot \mathrm{e}^{-\lambda_2} \frac{\lambda_2^{k-i}}{(k-i)!} = \mathrm{e}^{-(\lambda_1 + \lambda_2)} \sum_{i=0}^{k} \binom{k}{i} \frac{1}{k!} \lambda_1^i \lambda_2^{k-i} = \mathrm{e}^{-(\lambda_1 + \lambda_2)} \frac{(\lambda_1 + \lambda_2)^k}{k!}$, co daje tezę zadania.

Gęstość 2-wymiarowej zmiennej losowej (X, Y) to $f(x, y) = 3xy$ na obszarze ograniczonym prostymi $y = 0$, $y = x$, $y = 2 - x$.

5. Wyznaczyć gęstości brzegowe $f_1(x), f_2(y)$.

   1. Dla ustalonego $x \in [0, 1]$ y zmienia się od 0 do x, dla $x \in [1, 2]$ jest $y \in [0, 2 - x]$.

$$f_1(x) = \begin{cases} \int_0^x 3xy\,dy = \frac{3}{2} xy^2 \Big|_{y=0}^{x} = \frac{3}{2}\, x^3, & x \in [0, 1], \\ \int_0^{2-x} 3xy\,dy = \frac{3}{2} xy^2 \Big|_{y=0}^{2-x} = \frac{3}{2}\, x(2 - x)^2, & x \in [1, 2]. \end{cases}$$

2. Dla ustalonego $y \in [0,1]$ zmienna x przebiega przedział $[y, 2-y]$.

$$f_2(y) = \int_y^{2-y} 3xy\,dx = \frac{3}{2}x^2y\Big|_{x=y}^{2-y} = 6y(1-y).$$

---

6. Obliczyć wartość oczekiwaną zmiennej brzegowej Y. Czy zmienne X, Y są niezależne? (odpowiedź uzasadnić)

   $\mathrm{E}Y = \int_0^1 y\, f_2(y)\,dy = 6 \cdot \int_0^1 y^2 - y^3\,dy = \frac{1}{2}.$
   Zmienne X, Y są zależne, bo obszar na którym $f(x,y) > 0$ nie jest prostokątem. Na przykład: $x = 0.5$, $y = 0.75$. Jest $0 = f(x,y) \neq f_1(x) \cdot f_2(y) = \frac{3}{16} \cdot \frac{9}{8}$.

---

7. Prawdopodobieństwo sukcesu w pojedynczej próbie jest równe p. Wykonujemy niezależne doświadczenia do momentu uzyskania 3 sukcesów. Zmienna losowa X to liczba przeprowadzonych prób. Wyznaczyć rozkład zmiennej X, tzn. podać jej funkcję gęstości (ppb). Obliczyć wartość oczekiwaną zmiennej X.

   Niech $q = 1 - p$. Rozkład zmiennej X to

X	3	4	5	$\ldots$	k	$\ldots$
p_i	p^3	$\binom{3}{2} p^3\, q$	$\binom{4}{2} p^3\, q^2$	$\ldots$	$\binom{k-1}{2} p^3\, q^{k-3}$	$\ldots$

   Zwięzły dowód jest np. tutaj:
   `https://www.quora.com/How-do-we-derive-the-expected-value-of-a-negative-binomial-distribution-by-directly-evalu`
   How do we derive the expected value of a negative binomial distribution by directly evaluating the sum that arises using the definition of expected value?

---

8. Czytelnie i starannie - bez korzystania z notatek - napisać wielkie i małe greckie litery: alfę α, betę β, (d)zetę ζ, etę η, lambdę λ, chi χ, ksi ξ, fi ϕ, rho ρ.

α	β	ζ	η	λ	χ	ξ	ϕ	ρ
A	B	Z	H	Λ	X	Ξ	Φ	P

---

9. (a) Niech $X \sim U[-2,2]$. Znaleźć rozkład zmiennej $Y = |X|$.
   (b) Dla $X \sim U[-1,1]$ wyznaczyć rozkłady zmiennych $Y = X^3$, $Z = X^2$.

   a) Zbiór wartości zmiennej Y to przedział $[0,2]$. Niech $t \in [0,2]$. Dla dystrybuanty mamy równość $F_Y(t) = P(Y < t) = P(-t < X < 2) = P(X < 2) - P(X \leqslant -2) = P(X < 2) - P(X < -2)$ bo zmienna jest typu ciągłego. Stąd $F_Y(t) = F_X(t) - F_X(-t)$. Różniczkując obustronne to równanie (po zmiennej t, pamiętając o pochodnej funkcji wewnętrznej) otrzymujemy $f_Y(y) = f_X(t) + f_X(-t) = \frac{1}{2}$, czyli $X \sim U[0,2]$.
   b) Zbiór wartości zmiennej Y to przedział $[-1,1]$. Niech $t \in [-1,1]$. Dla dystrybuanty mamy równość $F_Y(t) = P(Y < t) = P(X^3 < t) = P(X < \sqrt[3]{t}) = F_X(\sqrt[3]{t})$. Różniczkując obustronnie to równanie otrzymujemy $f_Y(y) = f_X(t) \cdot \frac{1}{3}\, t^{-2/3} = \frac{1}{6} \cdot \frac{1}{\sqrt[3]{t^2}}$.

---

10. Niech X będzie zmienną o rozkładzie geometrycznym ($X \sim \mathrm{Geom}(p)$). Sprawdzić, że $\mathrm{V}(X) = \dfrac{1-p}{p^2}$.

Prawdopodobieństwa w rozkładzie geometrycznym to $P(X=k) = pq^{k-1}$, gdzie $k = 1, 2, \ldots$ oraz $q = 1-p$. Wykonujemy próby (jak w rozkładzie Bernoulliego) do momentu uzyskania pierwszego sukcesu.

Dla wartości oczekiwanej mamy $\mathrm{E}X = \sum_{k=1}^{\infty} k\,p\,q^{k-1}$. Potraktujmy q jako zmienną a wartość oczekiwaną jako szereg pewnej funkcji. Jest $\mathrm{E}X = \sum_{k=1}^{\infty} k\,p\,q^{k-1} = \sum_{k=1}^{\infty} p\,\dfrac{\partial}{\partial q} q^k$. Zamiast sumy pochodnych zapiszmy pochodną sumy: $\mathrm{E}X = p\,\dfrac{\partial}{\partial q} \sum_{k=0}^{\infty} q^k$ (można rozpocząć sumowanie od 0). Suma (szeregu geometrycznego) jest równa $\dfrac{1}{1-q}$ a pochodna to $\dfrac{1}{(1-q)^2}$. Stąd $\mathrm{E}X = \dfrac{p}{(1-q)^2} = \dfrac{1}{p}$.

Obliczymy teraz $\mathrm{E}(X^2)$. $\mathrm{E}(X^2) = \sum_{k=1}^{\infty} k^2 p\,q^{k-1}$. Zauważmy najpierw, że $k^2 = k\,(k-1) + k$. Powinniśmy obliczyć dwie sumy szeregów. Dla pierwszej sumy postępowanie jest podobne do wcześniejszego: $S_1 = \sum_{k=2}^{\infty} k\,(k-1)\,p\,q^{k-1} = p\,q \sum_{k=0}^{\infty} \dfrac{\partial^2}{\partial q^2} q^k = p\,q \left(\dfrac{1}{1-q}\right)^{(2)} = p\,q\,\dfrac{2}{(1-q)^3} = \dfrac{2q}{p^2}$. Proszę zwrócić uwagę, że początkowo dolny indeks sumowania to 1, zwiększamy go do 2, bo $k-1=0$ dla $k=1$, a potem zmniejszamy do 0 z uwagi na dwukrotne różniczkowanie. Druga suma była wyznaczona wcześniej. Mamy więc $\mathrm{E}(X^2) = \dfrac{2q}{p^2} + \dfrac{1}{p}$.

Korzystając ze wzoru $\mathrm{V}(X) = \mathrm{E}(X^2) - (\mathrm{E}X)^2$ mamy $\mathrm{V}X = \dfrac{2q}{p^2} + \dfrac{1}{p} - \dfrac{1}{p^2} = \dfrac{1-p}{p^2}$.

---

11. Zbiory $A_1, \ldots, A_4$ mają moc – odpowiednio – 40, 32, 20, 50. Losowo wybieramy pewien element (z całości). Wartością zmiennej losowej X jest moc zbioru z którego pochodzi wybrany element. Następnie losowo wybieramy jeden ze zbiorów. Wartością zmiennej losowej Y jest moc wybranego zbioru. Obliczyć $\mathrm{E}(X)$ i $\mathrm{E}(Y)$.

Ppb zmiennych losowych X, Y to

X	40	32	20	50
p_i	$40/142$	$32/142$	$20/142$	$50/142$

Y	40	32	20	50
p_i	$1/4$	$1/4$	$1/4$	$1/4$

Stąd już $\mathrm{E}X = \dfrac{40^2 + 32^2 + 20^2 + 50^2}{142}$ oraz $\mathrm{E}Y = \dfrac{40+32+20+50}{4}$.

Witold Karczewski