Metody probabilistyczne
Rozwiązania zadań

7. Wielowymiarowe zmienne losowe

21.11.2017

Zadanie 1*. *Pokaż, że rozkład* $P_{X|B}(A) = P(X \in A | Y \in B)$ *zdefiniowany jako*

$$P(X \in A|Y \in B) \;=\; \frac{P(X \in A, Y \in B)}{P(Y \in B)} \tag{1}$$

dla $P(Y \in B) > 0$, *spełnia aksjomaty Kołmogorowa*

Odpowiedź:

1. *Nieujemność* $P_{X|B}(A) \geqslant 0$: wynika wprost z definicji, ponieważ wszystkie wyrażenia po prawej stronie (1) są nieujemne.
2. *Normalizacja* $P_{X|B}(\mathbb{R}) = 1$. Wynika to z poniższego:

$$P_{X|B}(\mathbb{R}) \;=\; \frac{P(X \in \mathbb{R}, Y \in B)}{P(Y \in B)} \;\stackrel{(*)}{=}\; \frac{P(Y \in B)}{P(Y \in B)} \;=\; 1,$$

   gdzie w $(*)$ użyliśmy dość oczywistej równości $P(Y \in B) = P(X \in \mathbb{R}, Y \in B)$, wynikającej z tego, że zdarzenia $\{Y \in B\}$ oraz $\{X \in \mathbb{R} \wedge Y \in B\}$ są identyczne.
3. *Addytywność*: mając ciąg $A_1, A_2, \ldots$ zdarzeń rozłącznych, tj. takich, że $A_i \cap A_j = \emptyset$, musimy pokazać, że $P\Big(\bigcup_{j=1}^{\infty} A_j \,\big|\, Y \in B\Big) = \sum_{j=1}^{\infty} P(A_j | Y \in B)$. Mamy:

$$\begin{aligned}
P_{X|B}\Big(\bigcup_{j=1}^{\infty} A_j\Big) \;&=\; \frac{P\Big(\big(X \in \bigcup_{j=1}^{\infty} A_j\big), Y \in B\Big)}{P(Y \in B)} \\
&=\; \frac{P\Big(\bigcup_{j=1}^{\infty}\{X \in A_j \wedge Y \in B\}\Big)}{P(Y \in B)} \\
&\stackrel{(*)}{=}\; \sum_{j=1}^{\infty} \frac{P\big(\{X \in A_j \wedge Y \in B\}\big)}{P(Y \in B)} \\
&=\; \sum_{j=1}^{\infty} \frac{P\big(X \in A_j, Y \in B\big)}{P(Y \in B)} \\
&=\; \sum_{j=1}^{\infty} P(X \in A_j | Y \in B) \;=\; \sum_{j=1}^{\infty} P_{X|B}(A_j),
\end{aligned}$$

   gdzie w $(*)$ wykorzystaliśmy fakt, że skoro zdarzenia $\{X \in A_j\}$ $(j = 1, 2, \ldots)$ są rozłączne (bo zbiory $A_j, j = 1, 2, \ldots$ są rozłączne), to tym bardziej są rozłączne są zdarzenia $\{X \in A_j \wedge Y \in B\}$ $(j = 1, 2, \ldots)$.

Zadanie 2*. *Owad składa Y jajeczek zgodnie z rozkładem Poissona z parametrem λ, a potomek owada wylęga się z jaja z prawdopodobieństwem p niezależnie od innych. Wyznacz rozkład prawdopodobieństwa liczby potomków X*

Odpowiedź: Y ma rozkład Pois(λ), tzn:

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

Zakładając, że $Y = n$ (tzn. owad złożył n jajeczek), liczba potomków X dana jest *rozkładem dwumianowym* (liczba „sukcesów" w n próbach, gdzie „sukces" oznacza wylęgnięcie się potomka z danego jaja, stąd prawdopodobieństwo sukcesu wynosi p):

$$P(X = k|Y = n) = \binom{n}{k}p^k(1-p)^{n-k}, \qquad k = 0, 1, \ldots, n$$

Zauważmy też, że $P(X = k|Y = n) = 0$ dla $n < k$ (nie może się wylęgnąć więcej potomków niż jest złożonych jaj). Celem zadania jest policzenie „bezwarunkowego" rozkładu X, tzn. wyznaczenie prawdopodobieństwa $P(X = k)$, dla $k = 0, 1, 2, \ldots$. Ze wzoru na prawdopodobieństwo całkowite:

$$P(X = k) = \sum_{n=0}^{\infty} P(X = k|Y = n)P(Y = n) = \sum_{n=k}^{\infty}\binom{n}{k}p^k(1-p)^{n-k}\frac{\lambda^n}{n!}e^{-\lambda}, \tag{2}$$

gdzie możemy sumować od $n = k$, ponieważ $P(X = k|Y = n) = 0$ dla $n < k$. Przekształcając ostatnie wyrażenie otrzymujemy:

$$\begin{aligned}\binom{n}{k}p^k(1-p)^{n-k}\frac{\lambda^n}{n!}e^{-\lambda} &= \frac{\cancel{n!}}{k!(n-k)!}p^k(1-p)^{n-k}\frac{\overbrace{\lambda^{n-k}\lambda^k}^{\lambda^n}}{\cancel{n!}}e^{-\lambda} \\ &= \frac{(\lambda p)^k}{k!}e^{-\lambda}\,\frac{\big(\lambda(1-p)\big)^{n-k}}{(n-k)!}.\end{aligned}$$

Podstawiając do (2) dostajemy:

$$\begin{aligned}P(X = k) &= \sum_{n=k}^{\infty}\frac{(\lambda p)^k}{k!}e^{-\lambda}\,\frac{\big(\lambda(1-p)\big)^{n-k}}{(n-k)!} \\ &= \frac{(\lambda p)^k}{k!}e^{-\lambda}\sum_{n=k}^{\infty}\frac{\big(\lambda(1-p)\big)^{n-k}}{(n-k)!} \\ &\overset{(*)}{=} \frac{(\lambda p)^k}{k!}e^{-\lambda}\sum_{n=0}^{\infty}\frac{\big(\lambda(1-p)\big)^{n}}{n!} \\ &\overset{(\dagger)}{=} \frac{(\lambda p)^k}{k!}e^{-\lambda}e^{\lambda(1-p)} = \frac{(\lambda p)^k}{k!}e^{-\lambda p},\end{aligned}$$

gdzie w $(*)$ zmieniliśmy indeks sumowania z n na $n - k$, a w $(\dagger)$ użyliśmy znanego faktu: $\sum_{n=0}^{\infty}\frac{x^n}{n!} = e^x$ (patrz zadanie z poprzedniego wykładu) dla $x = \lambda(1-p)$, a następnie zauważyliśmy, że $e^{-\lambda}e^{\lambda(1-p)} = e^{-\lambda+\lambda(1-p)} = e^{-\lambda p}$.

A więc $P(X = k)$ ma rozkład Poissona z parametrem λp. Możemy z tego natychmiast wyznaczyć wartość oczekiwaną $EX = \lambda p$.

Zadanie 3. *Rozważmy schemat n prób Bernoulliego z prawdopodobieństwem sukcesu p. Jaka jest średnia liczba sukcesów w pierwszej próbie, jeżeli wiemy, że łącznie zaszło k sukcesów?*

Odpowiedź: Niech $Y \in \{0, 1, \ldots, n\}$ będzie zmienną losową o rozkładzie dwumianowym $B(n, p)$ określającą liczbę sukcesów w n próbach. Niech $X \in \{0, 1\}$ będzie binarną zmienną losową określającą czy w pierwszej próbie zaszedł sukces ($X = 1$) czy porażka ($X = 0$). Naszym celem jest policzenie $E(X|Y = k)$. Mamy:

$$E(X|Y = k) \;=\; 0 \cdot P(X = 0|Y = k) + 1 \cdot P(X = 1|Y = k) \;=\; P(X = 1|Y = k).$$

Zgodnie ze wzorem na prawdopodobieństwo warunkowe:

$$P(X = 1|Y = k) \;=\; \frac{P(X = 1, Y = k)}{P(Y = k)}.$$

Mianownik jest po prostu równy $P(Y = k) = \binom{n}{k}p^k(1-p)^{n-k}$, ponieważ jest to prawdopodobieństwo uzyskania k sukcesów w n próbach.

Teraz skupimy się na wyznaczeniu licznika. Od razu zauważmy, że licznik jest równy 0, jeśli $k = 0$ (ponieważ nie ma możliwości aby w pierwszej próbie zaszedł sukces ($X = 1$), a jednocześnie łączna liczba sukcesów k była równa 0). Rozważmy więc tylko $k > 0$. Prawdopodobieństwo w liczniku $P(X = 1, Y = k)$ określa szansę zajścia zdarzenia „k sukcesów w n próbach, przy czym w pierwszej próbie nastąpił sukces”. Dla dowolnego n-elementowego ciągu binarnego z k jedynkami (sukcesami), prawdopodobieństwo tego ciągu wynosi $p^k(1-p)^{n-k}$. Ile jest takich ciągów, w których sukces nastąpił w pierwszej próbie? Jest ich dokładnie $\binom{n-1}{k-1}$, ponieważ na tyle sposobów możemy rozłożyć $k-1$ sukcesów wśród pozostałych $n-1$ prób (jeden sukces jest zarezerwowany dla pierwszej próby). Tym samym:

$$P(X = 1, Y = k) \;=\; \binom{n-1}{k-1}p^k(1-p)^{n-k}, \qquad \text{dla } \; k > 0$$

Otrzymujemy więc:

$$P(X = 1|Y = k) \;=\; \frac{\binom{n-1}{k-1}p^k(1-p)^{n-k}}{\binom{n}{k}p^k(1-p)^{n-k}} \;=\; \frac{\binom{n-1}{k-1}}{\binom{n}{k}} \;=\; \frac{\frac{(n-1)!}{(k-1)!(n-k)!}}{\frac{n!}{k!(n-k)!}} \;=\; \frac{k}{n}.$$

Wynik ten zgadza się również dla $k = 0$ (jak zauważyliśmy wcześniej, $P(X = 1|Y = 0) = 0$). Tym samym $E(X|Y = k) = \frac{k}{n}$. Co zaskakujące, wynik ten w ogóle nie zależy od prawdopodobieństwa sukcesu p!

Zadanie 4. *Rozważmy schemat n prób Bernoulliego z prawdopodobieństwem sukcesu p. Jaka jest średnia liczba sukcesów w pierwszej próbie? Wyznacz wynik na dwa sposoby:*

- *Bezpośrednio (trywialne),*
- *Korzystając ze wzoru $E\big(E(X|Y)\big) = EX$ i wyniku z zadania 3.*

Odpowiedź: Niech zmienna $X \in \{0, 1\}$ określa liczbę sukcesów w pierwszej próbie. Ponieważ X ma rozkład dwupunktowy $B(p)$, więc oczywiście $EX = p$.

Zobaczymy, czy ten sam wynik dostaniemy używając wyniku z poprzedniego zadania. Niech $Y \in \{0, 1, \ldots, n\}$ będzie zmienną losową o rozkładzie dwumianowym $B(n, p)$ określającą liczbę sukcesów w n próbach. Pokazaliśmy, że:

$$E(X|Y = k) = \frac{k}{n}.$$

Traktując warunkową wartość oczekiwaną jako funkcję zmiennej Y dostajemy:

$$E(X|Y) = \frac{Y}{n}$$

Stąd:

$$EX \;=\; E\big(E(X|Y)\big) \;=\; E\left(\frac{Y}{n}\right) \;=\; \frac{np}{n} \;=\; p,$$

czego oczywiście się spodziewaliśmy.