Rachunek prawdopodobieństwa I*

Rozwiazanie zadania domowego nr. 2

Konrad Kaczmarczyk

4 kwietnia 2025

§1 Zadanie

Zadanie. Ponumerujmy genotypy AA, Aa, aa odpowiednio $1, 2, 3$ i niech $p_{ik}, i, k \in \{1, 2, 3\}$ oznacza prawdopodobieństwo warunkowe, że potomek będzie należał do genotypu k jeśli jego ojciec (lub matka) należą do genotypu i. Oblicz p_{ik} zakładając, że prawdopodobieństwo, że drugi z rodziców jest genotypu $1, 2, 3$ wynosi odpowiednio $p^2, 2pq, q^2$. Ogólnie pokaż, że $p_{ik}^{(n)}$-prawdopodobieństwa, że potomek w n-tym pokoleniu będzie należał do genotypu k jeśli ustalony przodek należał do genotypu i wynosi

$$\left\{ \begin{matrix} p^2 + \frac{pq}{2^{n-1}} & 2pq + \frac{q(q-p)}{2^{n-1}} & q^2 - \frac{q^2}{2^{n-1}} \\ p^2 + \frac{p(q-p)}{2^n} & 2pq + \frac{1-4pq}{2^n} & q^2 + \frac{q(p-q)}{2^n} \\ p^2 - \frac{p^2}{2^{n-1}} & 2pq + \frac{p(p-q)}{2^{n-1}} & q^2 + \frac{pq}{2^{n-1}} \end{matrix} \right\}.$$

Oznacza to, że wpływ genotypu przodka maleje z pokolenia na pokolenie o $1/2$.

Skorzystajmy z indukcji
Dla $n = 1$ poprzez ręczne rachunki mamy że powinna być to macierz:

$$\left\{ \begin{matrix} p^2 + \frac{1}{2} \cdot 2pq & q^2 + \frac{1}{2} \cdot 2pq & 0 \\ \frac{1}{2}p^2 + \frac{1}{4} \cdot 2pq & \frac{1}{2} & \frac{1}{2}q^2 + \frac{1}{4}pq \\ 0 & p^2 + pq & q^2 + pq \end{matrix} \right\}.$$

co po podstawieniu do tezy działa.
Zauważmy że zachodzą zależności:

$$p_{k1}^{(n+1)} = p_{k1}^{(n)} \cdot (p^2 + pq) + p_{k2}^{(n)} \cdot (\frac{1}{2}p^2 + \frac{1}{2}pq)$$
$$p_{k2}^{(n+1)} = p_{k1}^{(n)} \cdot (pq + q^2) + p_{k2}^{(n)} \cdot (\frac{1}{2}p^2 + pq + \frac{1}{2}q^2) + p_{k3}^{(n)} \cdot (p^2 + pq)$$
$$p_{k3}^{(n+1)} = p_{k2}^{(n)} \cdot (\frac{1}{2}pq + \frac{1}{2}q^2) + p_{k3}^{(n)} \cdot (pq + q^2)$$

korzystając z oczywistego faktu że $p + q = 1$ mamy że:

$$p_{k1}^{(n+1)} = p_{k1}^{(n)} \cdot p + p_{k2}^{(n)} \cdot \frac{1}{2}p$$
$$p_{k2}^{(n+1)} = p_{k1}^{(n)} \cdot q + p_{k2}^{(n)} \cdot \frac{1}{2} + p_{k3}^{(n)} \cdot p$$
$$p_{k3}^{(n+1)} = p_{k2}^{(n)} \cdot \frac{1}{2}q + p_{k3}^{(n)} \cdot q$$

czyli w postaci macierzy:

$$\begin{bmatrix} p_{11}^{(n+1)} & p_{12}^{(n+1)} & p_{13}^{(n+1)} \\ p_{21}^{(n+1)} & p_{22}^{(n+1)} & p_{23}^{(n+1)} \\ p_{31}^{(n+1)} & p_{32}^{(n+1)} & p_{33}^{(n+1)} \end{bmatrix} = \begin{bmatrix} p_{11}^{(n)} & p_{12}^{(n)} & p_{13}^{(n)} \\ p_{21}^{(n)} & p_{22}^{(n)} & p_{23}^{(n)} \\ p_{31}^{(n)} & p_{32}^{(n)} & p_{33}^{(n} \end{bmatrix} \cdot \begin{bmatrix} p & q & 0 \\ \frac{1}{2}p & \frac{1}{2} & \frac{1}{2}q \\ 0 & p & q \end{bmatrix}$$

wystarczy zatem sprawdzić czy:

$$\begin{bmatrix} p^2 + \frac{pq}{2^{n-1}} & 2pq + \frac{q(q-p)}{2^{n-1}} & q^2 - \frac{q^2}{2^{n-1}} \\ p^2 + \frac{p(q-p)}{2^n} & 2pq + \frac{1-4pq}{2^n} & q^2 + \frac{q(p-q)}{2^n} \\ p^2 - \frac{p^2}{2^{n-1}} & 2pq + \frac{p(p-q)}{2^{n-1}} & q^2 + \frac{pq}{2^{n-1}} \end{bmatrix} \cdot \begin{bmatrix} p & q & 0 \\ \frac{1}{2}p & \frac{1}{2} & \frac{1}{2}q \\ 0 & p & q \end{bmatrix}$$
$$= \begin{bmatrix} p^2 + \frac{pq}{2^n} & 2pq + \frac{q(q-p)}{2^n} & q^2 - \frac{q^2}{2^n} \\ p^2 + \frac{p(q-p)}{2^{n+1}} & 2pq + \frac{1-4pq}{2^{n+1}} & q^2 + \frac{q(p-q)}{2^{n+1}} \\ p^2 - \frac{p^2}{2^n} & 2pq + \frac{p(p-q)}{2^n} & q^2 + \frac{pq}{2^n} \end{bmatrix}$$

co zachodzi i kończy dowód zadania.

Zadanie. Niech $A_1, A_2, \dots$ będą zdarzeniami. Załóżmy, że $\sum_k \mathbf{P}(A_k) = \infty$. Wykaż, że jeśli

$$\limsup_{n\to\infty} \frac{\left(\sum_{k=1}^n \mathbf{P}(A_k)\right)^2}{\sum_{1\le j,k\le n} \mathbf{P}(A_j \cap A_k)} = \alpha > 0,$$

wówczas $\mathbf{P}(\limsup A_n) \ge \alpha$, gdzie $\limsup_{n\to\infty} A_n = \bigcap_n \bigcup_{m\ge n} A_m$.

Rozwiązanie zacznijmy od lematu, wynikającego z tw. Jensen'a:

Lemat 1.1

Niech $Z \ge 0$ będzie zmienną losową, wówczas:

$$(\mathbb{E}Z)^2 \le \mathbb{P}(Z \ge 0) \cdot \mathbb{E}Z^2$$

Dowód:
Skorzystajmy z wzoru na warunkową wartość oczekiwaną dla Z i Z^2:

$$\mathbb{E}Z = \mathbb{P}(Z > 0) \cdot \mathbb{E}\left[Z \mid Z > 0\right]$$
$$\mathbb{E}Z^2 = \mathbb{P}(Z^2 > 0) \cdot \mathbb{E}\left[Z^2 \mid Z > 0\right]$$

zauważmy że $\mathbb{P}(Z^2 > 0) = \mathbb{P}(Z > 0)$, skorzystajmy więc z tw. Jensen'a:

$$\left(\frac{\mathbb{E}[Z]}{\mathbb{P}(Z > 0)}\right)^2 = \left(\mathbb{E}\left[Z \mid Z > 0\right]\right)^2 \overset{\text{Jensen}}{\le} \mathbb{E}\left[Z^2 \mid Z > 0\right] = \frac{\mathbb{E}[Z^2]}{\mathbb{P}(Z^2 > 0)} = \frac{\mathbb{E}[Z^2]}{\mathbb{P}(Z > 0)}$$

co po wymnożeniu daje tezę lematu.
Udowodnijmy teraz więc drugi lemat, który pozwoli rozwiązać zadanie:

Lemat 1.2

Niech $A_1, \dots, A_n$ będą zdarzeniami, wówczas zachodzi:

$$\mathbb{P}\left(\bigcup_{k=1}^{n} A_k\right) \geq \frac{\left(\sum_{k=1}^{n} \mathbf{P}(A_k)\right)^2}{\sum_{1\leq j,k\leq n} \mathbf{P}(A_j \cap A_k)}$$

Dowód:
Ustalmy zmienne losowe:

$$X_k = \chi_{A_k}$$

i korzystając z Lemat 1.1 mamy że:

$$\left(\mathbb{E}\left[\sum_{k=1}^{n} X_k\right]\right)^2 \leq \mathbb{P}\left(\sum_{k=1}^{n} X_k > 0\right) \cdot \mathbb{E}\left[\left(\sum_{k=1}^{n} X_k\right)^2\right]$$

korzystając z faktów że:

$$X_k \cdot X_l = \chi_{A_k} \cdot \chi_{A_l} = \chi_{A_k \cap A_l} \qquad \mathbb{E}[X_k] = \mathbb{P}(A_k)$$

przekształcając mamy:

$$\left(\sum_{k=1}^{n} \mathbb{E}[X_k]\right)^2 \leq \mathbb{P}\left(\sum_{k=1}^{n} \chi_{A_k} > 0\right) \cdot \mathbb{E}\left[\sum_{l,k=1}^{n} X_l X_k\right]$$

$$\left(\sum_{k=1}^{n} \mathbb{P}(A_k)\right)^2 \leq \mathbb{P}\left(\sum_{k=1}^{n} \chi_{A_k} > 0\right) \cdot \mathbb{E}\left[\sum_{l,k=1}^{n} \chi_{A_k \cap A_l}\right]$$

$$\left(\sum_{k=1}^{n} \mathbb{P}(A_k)\right)^2 \leq \mathbb{P}\left(\bigcup_{k=1}^{n} A_k\right) \cdot \sum_{l,k=1}^{n} \mathbb{E}\left[\chi_{A_k \cap A_l}\right]$$

$$\left(\sum_{k=1}^{n} \mathbb{P}(A_k)\right)^2 \leq \mathbb{P}\left(\bigcup_{k=1}^{n} A_k\right) \cdot \sum_{l,k=1}^{n} \mathbb{P}(A_k \cap A_l)$$

$$\mathbb{P}\left(\bigcup_{k=1}^{n} A_k\right) \geq \frac{\left(\sum_{k=1}^{n} \mathbb{P}(A_k)\right)^2}{\sum_{l,k=1}^{n} \mathbb{P}(A_k \cap A_l)}$$

co po dowodzi lematu.

Korzystając z powyższego lematu możemy popełnić dwie obserwacje
Pierwszą, oznaczając:

$$s_n = \sum_{k=1}^{n} A_k \qquad t_n = \sum_{k,l=1}^{n} \mathbb{P}\left(A_k \cap A_l\right)$$

z tezy zadania wiemy że $s_n \to \infty$ czyli:

$$1 \geq \mathbb{P}\left(\bigcup_{k=1}^{n} A_k\right) = \frac{s_n^2}{t_n} \Rightarrow t_n \to \infty$$

Drugą, szacując:

$$\sum_{l,k=m+1}^{n} \mathbb{P}(A_l \cap A_k) =$$

$$\sum_{l,k=1}^{n} \mathbb{P}(A_l \cap A_k) - \sum_{\substack{1 \le l \le m \\ m+1 \le k \le n}} \mathbb{P}(A_l \cap A_k) - \sum_{\substack{1 \le k \le m \\ m+1 \le l \le n}} \mathbb{P}(A_l \cap A_k) - \sum_{l,k=1}^{m} \mathbb{P}(A_l \cap A_k)$$
$$\le t_n - \sum_{\substack{1 \le l \le m \\ m+1 \le k \le n}} \mathbb{P}(A_l \cap A_k) - \sum_{\substack{1 \le k \le m \\ m+1 \le l \le n}} \mathbb{P}(A_l \cap A_k) - t_m \le t_n - t_m$$

i korzystając z Lemat 1.2 mamy:

$$\mathbb{P}\left(\bigcup_{k=m}^{\infty} A_k\right) = \lim_{n\to\infty} \mathbb{P}\left(\bigcup_{k=m}^{n} A_k\right) \ge \lim_{n\to\infty} \frac{\left(\sum_{k=m}^{n} \mathbb{P}(A_k)\right)^2}{\sum_{m \le j,k \le n} \mathbb{P}(A_j \cap A_k)}$$
$$= \lim_{n\to\infty} \frac{(s_n - s_m)^2}{\sum_{l,k=m+1}^{n} \mathbb{P}(A_l \cap A_k)} \ge \lim_{n\to\infty} \frac{(s_n - s_m)^2}{t_n - t_m} =$$
$$\lim_{n\to\infty} \frac{t_n}{t_n - t_m} \cdot \left(\frac{s_n^2}{t_n} + 2\frac{s_n s_m}{t_n} + \frac{s_m^2}{t_n}\right)$$

korzystając zatem z obserwacji (czyli $s_n \to \infty, t_n \to \infty$) oraz warunku zadania (czyli $\frac{s_n^2}{t_n} \to \alpha$), mamy:

$$= 1 \cdot (\alpha + 0 + 0) = \alpha$$

czyli łacząc:

$$\mathbb{P}\left(\bigcup_{k=m}^{\infty} A_k\right) \ge \alpha$$

dla dowolnego m, a wiemy że przy granicy $m \to \infty$:

$$\alpha \le \lim_{m\to\infty} \mathbb{P}\left(\bigcup_{k=m}^{\infty} A_k\right) = \mathbb{P}\left(\limsup_{k} A_k\right)$$

co kończy dowód.