Wstęp do teorii informacji: Wykład 8

I. WSTĘP

Zmierzamy teraz do tzw. asymptotycznej zasady ekwipartycji, będącej informacyjnym odpowiednikiem prawa wielkich liczb z rachunku prawdopodobieństwa. Będzie ona nam później potrzebna do dowodu twierdzenia Shannona o przepustowości kanału informacyjnego. Przypomnijmy zatem podstawowe fakty na temat praw wielkich liczb.

II. NIERÓWNOŚĆ CZEBYSZEWA

Rozpatrzmy zmienną losową X o wartościach należących do zbioru $\{x_1, \ldots, x_N\}$. Wartością średnią jest

$$\overline{X} = \sum_{i=1}^{N} x_i p_i$$

a wariancją

$$\begin{aligned} D^2(X) &= \sum_{i=1}^{N} (x_i - \overline{X})^2 p_i = \sum_{i=1}^{N} x_i^2 p_i - \overline{X}^2 & (1) \\ &= \overline{X^2} - \overline{X}^2 & (2) \end{aligned}$$

Rozpatrzmy pawdopodobieństwo

$$\Pr\Big(|X - \overline{X}| \geqslant \varepsilon\Big) = \sum_{i; |x_i - \overline{X}| \geqslant \varepsilon} p_i \qquad (3)$$

Odpowiada on takiej procedurze: wybieramy losowo i, po czym sprawdzamy, jak bardzo x_i różni się od wartości sredniej $\overline{X}$, a następnie wybiermy tylko te zdarzenia, dla których x_i różni sie od $\overline{X}$ co najmniej o ε.

Twierdzenie 8.1 (*Nierówność Czebyszewa*)

$$\Pr\Big(|X - \overline{X}| \geqslant \varepsilon\Big) \leqslant \frac{D^2(X)}{\varepsilon^2} \qquad (4)$$

Dowód:

$$\begin{aligned} D^2(X) &= \sum_{i=1}^{N} (x_i - \overline{X})^2 p_i & (5) \\ &\geqslant \sum_{i; |x_i - \overline{X}| \geqslant \varepsilon} (x_i - \overline{X})^2 p_i & (6) \\ &\geqslant \sum_{i; |x_i - \overline{X}| \geqslant \varepsilon} \varepsilon^2 p_i & (7) \\ &= \varepsilon^2 \sum_{i; |x_i - \overline{X}| \geqslant \varepsilon} p_i & (8) \\ &= \varepsilon^2 \Pr\Big(|X - \overline{X}| \geqslant \varepsilon\Big), & (9) \end{aligned}$$

co mieliśmy pokazać. □

III. LOSOWE GENEROWANIE BITÓW (PROCES BERNOULLIEGO)

Rozpatrzmy N-krotny rzut monetą (czyli losowe generowanie bitów: orzeł 0, reszka 1). Niech prawdopodobieństwa wynoszą odpowiednio $p_1 = p$, $p_0 = q = 1 - p$. Prawdopodobieństwo k-krotnego wyrzucenia reszki przy N rzutach wynosi

$$p(k, N) = \binom{N}{k} p^k q^{N-k} = \frac{N!}{k!(N-k)!} p^k q^{N-k}. \qquad (10)$$

Niech teraz

$$X_j = \begin{cases} 0 & \text{jeśli wypadł orzeł} \\ 1 & \text{jeśli wypada reszka} \end{cases} \qquad (11)$$

oznacza zmienną losową odpowiadającą j-temu rzutowi. Średnia liczba reszek przy N rutach odpowiada zmiennej losowej

$$Y_N = \frac{X_1 + \cdots + X_N}{N} \qquad (12)$$

przyjmującej wartości ze zbioru

$$\Big\{\frac{0}{N}, \frac{1}{N}, \ldots, \frac{N-1}{N}, \frac{N}{N}\Big\}. \qquad (13)$$

Każda z powyższych wartości pojawia się z prawdopodobieństwem (10).

Obliczmy średnią i wariancję zmiennej losowej Y_N:

$$\begin{aligned} \overline{Y_N} &= \sum_{k=0}^{N} \frac{k}{N} \frac{N!}{k!(N-k)!} p^k q^{N-k} & (14) \\ &= p \sum_{k=1}^{N} \frac{k}{N} \frac{(N-1)!N}{(k-1)!k(N-k)!} p^{k-1} q^{N-1-(k-1)} & (15) \\ &= p \sum_{k=1}^{N} \frac{(N-1)!}{(k-1)!\big(N-1-(k-1)\big)!} p^{k-1} q^{N-1-(k-1)} & (16) \\ &= p \sum_{l=0}^{N-1} \frac{(N-1)!}{l!(N-1-l)!} p^l q^{N-1-l} & (17) \\ &= p(p+q)^{N-1} = p, & (18) \end{aligned}$$

a więc wynik nie zależy od ilości rzutów;

$$
\begin{aligned}
D^2(Y_N) &= \sum_{k=0}^{N}\left(\frac{k}{N}-\overline{Y_N}\right)^2\frac{N!}{k!(N-k)!}p^k q^{N-k} && (19)\\
&= \sum_{k=0}^{N}\left(\frac{k}{N}-p\right)^2\frac{N!}{k!(N-k)!}p^k q^{N-k} && (20)\\
&= \sum_{k=0}^{N}\left(\frac{k^2}{N^2}-2p\frac{k}{N}+p^2\right)\frac{N!}{k!(N-k)!}p^k q^{N-k} && (21)\\
&= \sum_{k=0}^{N}\frac{k^2}{N^2}\frac{N!}{k!(N-k)!}p^k q^{N-k} - \underbrace{\sum_{k=0}^{N}2p\frac{k}{N}\frac{N!}{k!(N-k)!}p^k q^{N-k}}_{2p^2} + \underbrace{\sum_{k=0}^{N}p^2\frac{N!}{k!(N-k)!}p^k q^{N-k}}_{p_2} && (22)\\
&= p\sum_{k=1}^{N}\frac{k^2}{N^2}\frac{(N-1)!N}{(k-1)!k\big(N-1-(k-1)\big)!}p^{k-1}q^{N-1-(k-1)}-p^2 && (23)\\
&= p\sum_{k=1}^{N}\frac{k}{N}\frac{(N-1)!}{(k-1)!\big(N-1-(k-1)\big)!}p^{k-1}q^{N-1-(k-1)}-p^2 && (24)\\
&= p\sum_{l=0}^{N-1}\frac{l+1}{N}\frac{(N-1)!}{l!(N-1-l)!}p^l q^{N-1-l}-p^2 && (25)\\
&= p\frac{N-1}{N}\sum_{l=0}^{N-1}\frac{l}{N-1}\frac{(N-1)!}{l!(N-1-l)!}p^l q^{N-1-l}+\frac{p}{N}\sum_{l=0}^{N-1}\frac{(N-1)!}{l!(N-1-l)!}p^l q^{N-1-l}-p^2 && (26)\\
&= p^2\frac{N-1}{N}+\frac{p}{N}-p^2=\frac{p-p^2}{N}=\frac{p(1-p)}{N}=\frac{pq}{N}. && (27)
\end{aligned}
$$

Twierdzenie 8.2: (*Bernoulliego prawo wielkich liczb*) W procesie Bernoulliego

$$\Pr\Big(\big|Y_N-p\big|\geqslant\varepsilon\Big)\leqslant\frac{pq}{N\varepsilon^2}, \qquad (28)$$

Dowód: Zastosujmy nierówność Czebyszewa do zmiennej losowej Y_N:

$$\Pr\Big(\big|Y_N-\overline{Y_N}\big|\geqslant\varepsilon\Big)\leqslant\frac{D^2(Y_N)}{\varepsilon^2}, \qquad (29)$$

co jest równoważne nierówności

$$\Pr\Big(\big|Y_N-p\big|\geqslant\varepsilon\Big)\leqslant\frac{pq}{N}\frac{1}{\varepsilon^2}, \qquad (30)$$

a to kończy dowód. □

Definicja 8.1: Ciąg zmiennych losowych $\{X_n\}$ jest zbieżny według prawdopodobieństwa do stałej c jeżeli dla każdego $\varepsilon>0$

$$\lim_{n\to\infty}\Pr\Big(\Big|X_n-c\Big|\geqslant\varepsilon\Big)=0. \qquad (31)$$

Uwagi: (a) W procesie Bernoulliego ciąg zmiennych losowych

$$Y_N=\frac{X_1+\cdots+X_N}{N} \qquad (32)$$

jest zbieżny według prawdopodobieństwa do p. (b) Średnia liczba sukcesów przy N próbach może się różnić od p o ε, ale prawdopodobiństwo takiego zdarzenia zmierza do zera przy zwiększającej się liczbie rzutów monetą. Prawa wielkich liczb tak formułowane zwane są też słabymi prawami wielkich liczb.

IV. ASYMPTOTYCZNA ZASADA EKWIPARTYCJI

Rozpatrzmy N rzutów monetą. Wynikiem jest uporządkowany ciąg bitów

$$(x_1,x_2,\ldots,x_N), \qquad (33)$$

który pojawi się z prawdopodobieństwem

$$p(x_1,x_2,\ldots,x_N)=p(x_1)p(x_2)\ldots p(x_N)=p^{N_1}q^{N_0} \qquad (34)$$

gdzie N_0, N_1 to, odpowiednio, liczby wypadnięć orłów i reszek, $p=p_1$, $q=1-p$.

Jeżeli N jest dużą liczbą, to orzeł wypadł w przybliżeniu $N_0\approx Np_0=Nq$ razy, a reszka $N_1\approx Np_1=Np$

razy. Tak więc

$$\begin{aligned}
p(x_1, x_2, \ldots, x_N) &\approx p^{Np} q^{Nq} = (p^p q^q)^N, && (35)\\
\log p(x_1, x_2, \ldots, x_N) &\approx N \log(p^p q^q)\\
&= N\Big(p \log p + (1-p)\log(1-p)\Big)\\
&= -NH(p), && (36)\\
H(p) &\approx -\frac{\log p(x_1, x_2, \ldots, x_N)}{N}. && (37)
\end{aligned}$$

(por. Wykład 2). Ostatnią równość możemy jeszcze przepisać jako

$$p(x_1, x_2, \ldots, x_N) \approx 2^{-NH(p)}. \tag{38}$$

Widać, że te przybliżone równości stają się tym dokładniejsze, im większe jest N. Można to precyzyjniej zapisać jako

Zdefiniujmy zmienną losową X_N o wartościach

$$-\frac{\log\Big(p^{N_1}(1-p)^{N-N_1}\Big)}{N}, \quad N_1 = 0, \ldots, N. \tag{39}$$

Twierdzenie 8.3: (*Asymptotyczna zasada ekwipartycji dla procesu Bernoulliego*) Ciąg zmiennych losowych X_N jest zbieżny według prawdopodobieństwa do $H(p)$.