Chernoffova in Azumova neenakost

Rok Požar

11. januar 2008

1 Vsota enakomerno porazdeljenih neodvisnih ± 1 spremenljivk

Lema 1.1 (Markova neenakost) *Če je X nenegativna slučajna spremenljivka in $a > 0$, potem velja*

$$P[X \geq a] \leq \frac{E[X]}{a}.$$

Dokaz: Če je X nenegativna, potem velja

$$E[X] \geq a \cdot P[X \geq a].$$

■

Izrek 1.2 (Chernoffova neenakost) *Naj bodo $X_1, \ldots, X_n$ neodvisne slučajne spremenljivke, ki zavzamejo vrednost $+1$ in -1, obe z verjetnostjo $\frac{1}{2}$. Naj bo $X = X_1 + \cdots + X_n$. Potem za vsak realen $t \geq 0$ velja*

$$P[X \geq t] < e^{-t^2/2\sigma^2} \quad in \quad P[X \leq -t] < e^{-t^2/2\sigma^2},$$

kjer je $\sigma = \sqrt{Var[X]} = \sqrt{n}$.

Dokaz: Dokažimo le prvo neenakost, druga bo namreč sledila iz simetrije. Definirajmo novo slučajno spremenljivko $Y = e^{uX}$, kjer je $u > 0$ realni parameter (za zdaj še nedoločen). Potem velja $P[X \geq t] = P[Y \geq e^{ut}]$. Po Markovi neenakosti velja $P[Y \geq q] \leq \frac{E[Y]}{q}$. Računajmo

$$E[Y] = E[e^{u \sum_{i=1}^{n} X_i}] = E[\prod_{i=1}^{n} e^{uX_i}] = \prod_{i=1}^{n} E[e^{uX_i}]$$

(zaradi neodvisnosti X_i)

$$= \left(\frac{e^u + e^{-u}}{2}\right)^n \leq e^{nu^2/2}.$$

Zadnja ocena sledi iz neenakosti $\frac{e^x+e^{-x}}{2} = \cosh x \leq e^{x^2/2}$, ki velja za vsa realna števila x (obe strani razvijemo v Taylorjevo vrsto in primerjamo koeficiente). Potem je

$$P[Y \geq e^{ut}] \leq \frac{E[Y]}{e^{ut}} \leq e^{nu^2/2-ut}.$$

Zadnji izraz je minimiziran za $u = \frac{t}{n}$, od koder sledi $e^{-t^2/2n} = e^{-t^2/2\sigma^2}$. ∎

Naj bo X množica z n točkami in naj bo $\mathcal{F}$ družina podmnožic množice X. Radi bi pobarvali točke množice X z rdečo in modro barvo tako, da vsaka množica družine $\mathcal{F}$ vsebuje čim bolj enako število rdečih in modrih točk (uravnoteženo barvanje). Naj ima rdeča barva vrednost $+1$ in modra vrednost -1. Potem lahko barvanje podamo s preslikavo $\psi : X \to \{-1, +1\}$. Za poljubno množico $S \in \mathcal{F}$ nam potem $\psi(S) = \sum_{x \in S} \psi(x)$ ravno pove razliko med številom rdečih in modrih točk. Z $\mathrm{urb}(\mathcal{F},\psi)$ družine $\mathcal{F}$ in barvanjem ψ označimo $\max_{S \in \mathcal{F}} |\psi(S)|$ in z $\mathrm{urb}(\mathcal{F})$ minimum množice $\{\mathrm{urb}(\mathcal{F},\psi); \psi$ neko barvanje množice $X\}$.

Trditev 1.3 *Naj bo* $|X| = n$ *in* $|\mathcal{F}| = m$. *Potem velja* $\mathrm{urb}(\mathcal{F}) \leq \sqrt{2n\ln(2m)}$. *Če je maksimalno število množic v* $\mathcal{F}$ *največ* s, *potem velja* $\mathrm{urb}(\mathcal{F}) \leq \sqrt{2s\ln(2m)}$.

Dokaz: Naj bo $\psi : X \to \{-1, +1\}$ slučajno barvanje množice X, kjer so barve točk izbrane enotno in neodvisno. Za vsako fiksno množico $S \subseteq X$ je $\psi(S) = \sum_{x \in S} \psi(x)$ vsota $|S|$ neodvisnih slučajnih ± 1 spremenljivk. Zaradi zgornjega izreka velja

$$P[|\psi(S)| > t] < 2e^{-t^2/2|S|} \leq 2e^{-t^2/2s}.$$

Za $t = \sqrt{2s\ln(2m)}$, $2e^{-t^2/2s}$ postane $\frac{1}{m}$. Torej s pozitivno verjetnostjo slučajno barvanje zadošča $|\psi(S)| \leq t$ za vse $S \in \mathcal{F}$. ∎

2 Azumova neenakost

Martingala je zaporedje $X_0, \ldots, X_m$ slučajnih spremenljivk, da za $0 \leq i < m$ velja

$$E[X_{i+1}|X_i, X_{i-1}, \ldots, X_0] = X_i.$$

Predstavljajmo si, da gre kockar v igralnico z X_0 denarja. V igralnici je mnogo iger na srečo. Vse igre so poštene, kar pomeni, da je njihovo matematično upanje enako 0. Igralec lahko izbere naslednjo strategijo: podvoji stavo vsakič, ko izgubi, da bo prva zmaga pokrila vse prejšnje vložke in prinesla še dodaten dobiček v vrednosti originalnega vložka. Naj slučajna spremenljivka X_i predstavlja kockarjevo srečo v času i. Če je $X_i = a$, potem mora biti pogojna verjetnost spremenljivke X_{i+1} enaka a, torej je to martingala.

Poglejmo si preprost, a poučen primer. Predpostavimo, da slučajna spremenljivka X_n predstavlja kockarjevo srečo po n metih poštenega kovanca, kjer igralec dobi 1 evro, če pade cifra in izgubi 1 evro, če pade grb. Očitno zaporedje zadošča pogojem zgornje definicije, zato je to martingala.

Lema 2.1 *Naj bo X slučajna spremenljivka in naj velja $E[X] = 0$ in $|X| \leq 1$. Naj bo $\lambda > 0$. Potem velja*

$$E[e^{\lambda X}] \leq e^{\lambda^2/2}.$$

Dokaz: Definirajmo linerano fukcijo

$$h(x) = \frac{e^{\lambda} + e^{-\lambda}}{2} + \frac{e^{\lambda} - e^{-\lambda}}{2} x.$$

Za $x \in [-1, 1]$ velja $e^{\lambda x} \leq h(x)$ ($y = h(x)$ je ravno sekanta skozi točki $x = \pm 1$ konveksne funkcije $y = e^{\lambda x}$). Velja

$$E[e^{\lambda X}] \leq E[h(X)] = h(E[X]) = h(0) = \cosh(\lambda) = e^{\lambda^2/2}.$$

■

Izrek 2.2 (Azumova neenakost) *Naj bo $0 = X_0, \ldots, X_m$ martingala za katero velja*

$$|X_{i+1} - X_i| \leq 1$$

za vse $0 \leq i < m$. Naj bo $\lambda > 0$. Potem velja

$$P[X_m > \lambda\sqrt{m}] < e^{-\lambda^2/2}.$$

Dokaz: Naj bo $\alpha = \frac{\lambda}{\sqrt{m}}$. Pišimo $Y_i = X_i - X_{i-1}$. Iz predpostavk očitno sledi, da je $|Y_i| \leq 1$ in $E[Y_i | X_{i-1}, X_{i-2}, \ldots, X_0] = 0$. Po zgornji lemi sledi

$$E[e^{\alpha Y_i} | X_{i-1}, X_{i-2}, \ldots, X_0] \leq \cosh(\alpha) = e^{\alpha^2/2}.$$

Torej je

$$\begin{aligned} E[e^{\alpha X_m}] &= E[\prod_{i=1}^{m} e^{\alpha Y_i}] \\ &= E[\left(\prod_{i=1}^{m-1} e^{\alpha Y_i}\right) E[e^{\alpha Y_m} | X_{m-1}, X_{m-2}, \ldots, X_0]] \\ &\leq E[\prod_{i=1}^{m-1} e^{\alpha Y_i}] e^{\alpha^2/2} \leq e^{\alpha^2 m/2}. \end{aligned}$$

Od tod sledi

$$\begin{aligned} P[X_m > \lambda\sqrt{m}] &= P[e^{\alpha X_m} > e^{\alpha\lambda\sqrt{m}}] \\ &< E[e^{\alpha X_m}] e^{-\alpha\lambda\sqrt{m}} \\ &\leq e^{\alpha^2 m/2 - \alpha\lambda\sqrt{m}} \\ &= e^{-\lambda^2/2}. \end{aligned}$$

■

Literatura

[1] N. Alon, J. H. Spencer, *The probabilistic method*, Courant Institute, 2000.

[2] J. Matoušek, J. Vondrák, *The probabilistic method*, Institute for Theoretical Computer Science, 2002.