Linearnost matematičnega upanja[1]

Matematično upanje slučajne spremenljivke $X : \Omega \to \mathbb{R}$ končnega prostora (Ω, p) je definirano kot vsota produktov vrednosti slučajne spremenljivke ter verjetnosti, da spremenljivka to vrednost zavzame. Torej

$$\mathbf{E}(X) = \sum_{\omega \in \Omega} p(\omega) X(\omega).$$

Iz definicije sledi, da vedno obstaja elementarni dogodek $\omega_1 \in \Omega$, za katerega velja

$$X(\omega_1) \geq \mathbf{E}(X).$$

Podobno, vedno lahko najdemo dogodek $\omega_2 \in \Omega$ za katerega velja

$$X(\omega_2) \leq \mathbf{E}(X).$$

Tako lahko minimalno oziroma maksimalno vrednost slučajne spremenljivke X ocenimo zgolj s pomočjo matematičnega upanja:

$$\min_{\omega \in \Omega} X(\omega) \leq \mathbf{E}(X) \quad \text{in} \quad \max_{\omega \in \Omega} X(\omega) \geq \mathbf{E}(X).$$

[1] 26. februar 2013

Uporabnost matematičnega upanja temelji na njegovi linearnosti.

Trditev 1 *Upanje je linearen operator, t.j. za vsak par slučajnih spremenljivk X, Y in konstanti α, $\beta \in \mathbb{R}$ velja*

$$\mathbf{E}(\alpha X + \beta Y) = \alpha\, \mathbf{E}(X) + \beta\, \mathbf{E}(Y)\,.$$

Dokaz.

$$\begin{aligned}
\mathbf{E}(\alpha X + \beta Y) &= \sum_{\omega \in \Omega} p(\omega)(\alpha X + \beta Y)(\omega) \\
&= \alpha \sum_{\omega \in \Omega} p(\omega) X(\omega) + \beta \sum_{\omega \in \Omega} p(\omega) Y(\omega) \\
&= \alpha\, \mathbf{E}(X) + \beta\, \mathbf{E}(Y)\,.
\end{aligned}$$

■

Iz zgornje leme sledi, da je matematično upanje vsote slučajnih spremenljivk $X = X_1 + X_2 + \cdots + X_n$ enako

$$\mathbf{E}(X) = \mathbf{E}(X_1) + \mathbf{E}(X_2) + \cdots + \mathbf{E}(X_n)\,.$$

Za slučajen dogodek A definiramo *indikatorsko spremenljivko* I_A na naslednji način:

$$I_A(\omega) = \begin{cases} 1, & \omega \in A \\ 0, & \omega \notin A. \end{cases}$$

Očitno dogodek A enolično določa I_A in obratno. Velja naslednje:

Trditev 2 *Za vsak dogodek A velja $\mathbf{E}(I_A) = \mathbf{P}(A)$.*

Dokaz. Izpeljemo takole:

$$\mathbf{E}(I_A) = \sum_{\omega \in \Omega} p(\omega) I_A(\omega) = \sum_{\omega \in A} p(\omega) = \mathbf{P}(A).$$

■

V veliko primerih obravnavano slučajno spremenljivko lahko zapišemo kot vsoto indikatorskih spremenljivk

$$X = I_{A_1} + I_{A_2} + \cdots + I_{A_n},$$

kjer poznamo verjetnosti dogodkov $A_1, A_2, \ldots, A_n$. Potem velja

$$\mathbf{E}(X) = \mathbf{P}(A_1) + \mathbf{P}(A_2) + \cdots + \mathbf{P}(A_n)\,.$$

Pokazali smo, da je matematično upanje linearen operator. To lastnost s pridom uporabljamo za dokazovanje problemov s pomočjo verjetnosti. V naslednjih razdelkih so prikazani primeri uporabe.

1 Število fiksnih točk permutacije

Permutacija končne množice A je bijektivna preslikava množice A nase. Brez škode za splošnost lahko elemente množice A označimo z naravnimi števili $\{1, 2, \ldots, n\}$. Taki permutaciji rečemo, da je *reda* n. Množico vseh permutacij reda n označimo z S_n. Velja $|S_n| = n!$.

Če za permutacijo σ obstaja $i \in \{1, 2, \ldots, n\}$, pri čemer $\sigma(i) = i$, potem točki i pravimo *fiksna točka* permutacije σ. Število fiksnih točk je očitno med 0 in n. Zanima nas pričakovano število fiksnih točk v naključno izbrani permutaciji.

Izrek 3 *V povprečju ima vsaka permutacija eno fiksno točko.*

Dokaz. Izračunajmo verjetnost pričakovanega števila fiksnih točk slučajne permutacije σ na $\{1, \ldots, n\}$. Če je

$$F(\sigma) = |\{i : \sigma(i) = i\}|,$$

lahko to izrazimo kot vsoto indikatorskih spremenljivk:

$$F(\sigma) = \sum_{i=1}^{n} F_i(\sigma),$$

kjer je $F_i(\sigma) = 1$, če je $\sigma(i) = i$ in $F_i(\sigma) = 0$ sicer. Torej je

$$\mathbf{E}(F_i) = \mathbf{P}(\sigma(i) = i) = \frac{(n-1)!}{n!} = \frac{1}{n}\,.$$

Od tod sledi

$$\mathbf{E}(F) = \frac{1}{n} + \frac{1}{n} + \cdots + \frac{1}{n} = 1\,,$$

kar pomeni, da ima permutacija v povprečju eno fiksno točko. ∎

2 Hamiltonske poti v turnirjih

Turnir je orientiran poln graf, kar pomeni, da med poljubnima točkama u in v nastopa natanko ena usmerjena povezava (u, v) oz. (v, u). *Hamiltonska pot* v turnirju je usmerjena pot skozi vse točke.

Znano je, da ima vsak turnir Hamiltonsko pot. Naslednji izrek (Szele, 1943) pravi, da obstoja turnir z velikim številom Hamiltonskih poti. Temu izreku oziroma njegovemu dokazu se pogosto pripisuje prva uporaba verjetnostne metode.

Izrek 4 *Obstaja turnir na n točkah, ki ima vsaj $\frac{n!}{2^{n-1}}$ Hamiltonskih poti.*

Dokaz. Izračunajmo pričakovano število Hamiltonskih poti v slučajnem turnirju T na n točkah, kjer ima vsaka povezava slučajno orientacijo, izbrano neodvisno z verjetnostjo $\frac{1}{2}$. Označimo vozlišča turnirja T z elementi množice $\{1, 2, \ldots, n\}$.

Za dano permutacijo σ na $\{1, \ldots, n\}$ si oglejmo zaporedje $\sigma(1), \sigma(2), \ldots, \sigma(n)$ in označimo z I_σ indikator dogodka, da vse povezave $(\sigma(i), \sigma(i+1))$ nastopajo v T s to orientacijo. Upoštevajmo, da orientacijo različnih povezav izberemo neodvisno in izračunajmo $\mathbf{E}[I_\sigma]$:

$$\mathbf{E}(I_\sigma) = \mathbf{P}(\sigma) = \frac{1}{2^{n-1}}\,.$$

Naj bo X število vseh Hamiltonskih poti v turnirju. Sledi

$$X = \sum_{\sigma \in S_n} I_\sigma$$

in od tod

$$\mathbf{E}(X) = \sum_{\sigma} \mathbf{E}(I_\sigma) = \sum_{\sigma} \frac{1}{2^{n-1}} = \frac{1}{2^{n-1}} \sum_{\sigma} 1 = \frac{n!}{2^{n-1}},$$

saj je vseh različnih permutacij na n točkah $n!$. Torej obstaja tak turnir T, ki vsebuje vsaj $\frac{n!}{2^{n-1}}$ Hamiltonskih poti. ■

3 Maksimalni prerez grafov

Tukaj obravnavamo problem maksimalnega prereza, ki je predvsem pomemben algoritmičen problem. Imamo graf $G = (V, E)$ in želimo razdeliti množico točk v dva razreda, A in $B = V \setminus A$ tako, da je število povezav med A in B maksimalno. Naslednji izrek nam pove, da je vedno možno doseči, da je število povezav med A in B vsaj polovica vseh povezav v grafu.

Izrek 5 *Vsak graf z m povezavami vsebuje dvodelen podgraf z vsaj $m/2$ povezavami.*

Dokaz. Naj bo $G = (V, E)$. Izberimo slučajno podmnožico $T \subseteq V$ z vstavljanjem vsake točke v T neodvisno z verjetnostjo $\frac{1}{2}$. Za dano povezavo $e = uv$ naj I_e označuje indikatorsko spremenljivko dogodka, da je natanko ena od točk u in v v T. Potem velja

$$\mathbf{E}(I_e) = \mathbf{P}\big((u \in T \ \text{ in } \ v \notin T) \ \text{ ali } \ (u \notin T \ \text{ in } \ v \in T)\big) = \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{2}.$$

Če X označuje število povezav, ki imajo natanko eno točko v T, potem je

$$X = \sum_{e \in E} I_e$$

in zato

$$\mathbf{E}(X) = \sum_{e \in E} \mathbf{E}(I_e) = \sum_{e \in E} \frac{1}{2} = \frac{1}{2} \sum_{e \in E} 1 = \frac{m}{2},$$

saj je m število vseh povezav v grafu. Torej, za nekatere $T \subseteq V$ obstaja vsaj $\frac{m}{2}$ povezav med T in $V \setminus T$. Te povezave inducirajo dvodelen podgraf. ■

4 Dominantne podmnožice v grafih

Dominantna podmnožica grafa $G = (V, E)$ je podmnožica vozlišč $D \subseteq V$, za katero velja, da je vsako vozlišče grafa G bodisi v D bodisi ima v D soseda. Na primer, najmanjša dominantna množica cikla C_n velikosti $\lceil \frac{n}{3} \rceil$. Spodnji izrek navzgor oceni najmanjšo dominantno podmnožico.

Izrek 6 *Naj bo $G = (V, E)$ graf z n vozlišči in z minimalno stopnjo $\delta > 1$. Potem ima graf G dominantno podmnožico z največ $n\frac{1+\ln(\delta+1)}{\delta+1}$ vozlišči.*

Dokaz. Naj bo S slučajna množica točk iz G tako, da vsako točko izberemo neodvisno z verjetnostjo p. Naj bo T množica točk iz G, za katero velja, da ne vsebuje točk iz S niti točk, ki imajo soseda v S. Množica $T \cup S$ tvori dominantno množico grafa G. Velja

$$\mathbf{E}(|S|) = \sum_{v \in V} p = n\,p\,.$$

Naj bo I_v indikatorska spremenljivka dogodka, da je v izbran v T. Torej

$$I_v = \begin{cases} 1, & \text{niti } v \text{ niti sosedje } v\text{-ja niso v } S \\ 0, & \text{sicer.} \end{cases}$$

Potem je $|T| = \sum_{v \in V} I_v$. Verjetnost, da je neko vozlišče v vsebovano v T, je enaka

$$\mathbf{P}(v \in T) = (1-p)^{d(v)+1} \le (1-p)^{\delta+1}\,.$$

Torej

$$\mathbf{E}(|T|) = \sum_{v \in V} \mathbf{E}(I_v) \le n(1-p)^{\delta+1} \le ne^{-p(\delta+1)}$$

in od tod

$$\mathbf{E}(|S \cup T|) = \mathbf{E}(|S|) + \mathbf{E}(|T|) \le np + ne^{-p(\delta+1)}\,.$$

S prijemi matematične analize dobimo najmanjšo zgornjo mejo naše ocene $\mathbf{E}(|S \cup T|)$ za $p = \frac{\ln(\delta+1)}{\delta+1}$. Torej

$$\mathbf{E}(|S \cup T|) \le n\frac{1+\ln(\delta+1)}{\delta+1}\,.$$

■

5 Uravnoteženi vektorji

Izrek 7 *Naj bodo $v_1, v_2, \ldots, v_n \in \mathbb{R}^n$, za katere velja $|v_i| = 1$ za vsak $i \in \{1, 2, \ldots, n\}$. Potem obstajajo taki $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$, ki lahko zavzamejo vrednosti 1 ali -1, da velja*

$$|\varepsilon_1 v_1 + \ldots + \varepsilon_n v_n| \leq \sqrt{n}$$

in taki, da velja

$$|\varepsilon_1 v_1 + \ldots + \varepsilon_n v_n| \geq \sqrt{n}\,.$$

Dokaz. Naj bodo $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ izbrani enakomerno in neodvisno iz množice $\{1, -1\}$. In naj bo

$$X = |\varepsilon_1 v_1 + \ldots + \varepsilon_n v_n|^2\,.$$

Potem je

$$X = \sum_{i=1}^{n} \sum_{j=1}^{n} \varepsilon_i v_i \cdot \varepsilon_j v_j\,,$$

kar zapišemo malo drugače kot

$$X = \sum_{i=1}^{n} \sum_{j=1}^{n} \varepsilon_i \varepsilon_j v_i \cdot v_j\,.$$

Uporabimo linearnost matematičnega upanja

$$\mathbf{E}(X) = \sum_{i=1}^{n} \sum_{j=1}^{n} v_i \cdot v_j \mathbf{E}(\varepsilon_i \varepsilon_j)\,.$$

Če je $i \neq j$, potem je

$$\mathbf{E}(\varepsilon_i \varepsilon_j) = \mathbf{E}(\varepsilon_i) \cdot \mathbf{E}(\varepsilon_j) = 0\,.$$

Če pa je $i = j$, je $\varepsilon_i^2 = 1$ in zato je

$$\mathbf{E}(\varepsilon_i^2) = 1\,.$$

Tako je

$$\mathbf{E}(X) = \sum_{i=1}^{n} v_i v_i = n\,.$$

Od tod sledi, da obstajajo taki $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$, katerih zaloga vrednosti je $\{1, -1\}$ in je $X \geq n$ in taki, da je $X \leq n$. Izrek je s tem dokazan. ■

Izrek 8 *Naj bodo $v_1, v_2, \ldots, v_n \in \mathbb{R}^n$, za katere velja $|v_i| = 1$ za vsak $i \in \{1, 2, \ldots, n\}$. Naj bodo $p_1, p_2, \ldots, p_n \in [0, 1]$ poljubni in naj bo $w = p_1 v_1 + p_2 v_2 + \ldots + p_n v_n$. Potem obstajajo taki $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n \in \{0, 1\}$, da za $v = \varepsilon_1 v_1 + \ldots + \varepsilon_n v_n$ velja*

$$|w - v| \leq \frac{\sqrt{n}}{2} \,.$$

Dokaz. Izberimo vsak ε_i neodvisno z verjetnostjo

$$\mathbf{P}(\varepsilon_i = 1) = p_i \quad \text{in} \quad \mathbf{P}(\varepsilon_i = 0) = 1 - p_i \,.$$

Naključna izbira ε_i nam da slučajno spremenljivko v. Sedaj si oglejmo slučajno spremenljivko

$$X = |w - v|^2 \,,$$

za katero velja:

$$\begin{aligned} X &= |\sum_{i=1}^{n} (p_i - \varepsilon_i) v_i|^2 \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} (p_i - \varepsilon_i) v_i \, (p_j - \varepsilon_j) v_j \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} v_i \cdot v_j (p_i - \varepsilon_i)(p_j - \varepsilon_j) \,. \end{aligned}$$

Zato je

$$\mathbf{E}(X) = \sum_{i=1}^{n} \sum_{j=1}^{n} v_i \cdot v_j \mathbf{E}((p_i - \varepsilon_i)(p_j - \varepsilon_j)) \,.$$

Če je $i \neq j$, potem sta slučajni spremenljivki $p_i - \varepsilon_i$ in $p_j - \varepsilon_j$ neodvisni in zato velja

$$\mathbf{E}((p_i - \varepsilon_i)(p_j - \varepsilon_j)) = \mathbf{E}((p_i - \varepsilon_i)) \mathbf{E}((p_j - \varepsilon_j)) = 0 \,.$$

Za $i = j$ pa je

$$\mathbf{E}((p_i - \varepsilon_i)^2) = p_i(p_i - 1)^2 + (1 - p_i)p_i^2 = p_i(1 - p_i) \leq \frac{1}{4}\,.$$

Torej je

$$\mathbf{E}(X) = \sum_{i=1}^{n} p_i(1 - p_i)|v_i|^2 \leq \frac{1}{4}\sum_{i=1}^{n} |v_i|^2 = \frac{n}{4}\,.$$

■