Václav Dupač
Stochastické početní methody

Časopis pro pěstování matematiky, Vol. 81 (1956), No. 1, 55--68

Persistent URL: http://dml.cz/dmlcz/117172

Časopis pro pěstování matematiky, roč. 81 (1956)

STOCHASTICKÉ POČETNÍ METHODY

VÁCLAV DUPAČ, Praha.

(Došlo dne 14. února 1955.) DT 517:6:519.28

Některé matematické početní úlohy lze přibližně řešit methodami, založenými na theorii náhodného výběru. Článek pojednává o užití těchto method k výpočtu určitých integrálů, objemů vícerozměrných těles, inversní matice a čísla π. Některé výsledky jsou původní.

1. Úvod. Jednou z klasických úloh z okruhu t. zv. geometrických pravděpodobností je Buffonova úloha s jehlou. Na rovinu, rozdělenou na pásy soustavou rovnoběžných přímek o jednotkové vzdálenosti, hoďme náhodným způsobem jehlu délky $l < 1$. Jest určiti pravděpodobnost P, že jehla protne některou přímku soustavy. Snadno se zjistí, že

$$P = \frac{2}{\pi} \int_0^{\pi} \frac{l}{2} \sin \vartheta \, d\vartheta = \frac{2l}{\pi} .$$

Tento výsledek dává pak možnost stanoviti experimentálně přibližnou hodnotu čísla π. Jest

$$\pi = \frac{2l}{P}, \text{ t. j. přibližně } \pi \cong \frac{2ln}{m},$$

kde n je celkový počet provedených hodů a m je počet hodů, v nichž jehla protala přímku; přitom výraz $\frac{2ln}{m}$ představuje náhodnou proměnnou, která s n neomezeně vzrůstajícím konverguje k π s pravděpodobností 1.

Neobvyklost tohoto způsobu určení čísla π spočívá v tom, že numerický početní problém je nahrazen ekvivalentním problémem stochastickým, a ten se pak řeší specificky stochastickou methodou, totiž náhodným výběrem (opakovaným pokusem). Methody tohoto typu budeme nazývat stochastickými početními methodami (v literatuře se někdy nazývají methodami Monte Carlo).

Zajímavá je už sama skutečnost, že úlohy zcela nenáhodového charakteru lze řešit náhodným výběrem. Avšak pozoruhodnější je to, že v určitých úlohách je řešení stochastickou methodou výhodnější (někdy dokonce v jistém smyslu přesnější) než řešení obvyklými methodami. Ukážeme si několik příkladů tohoto typu.

2. Výpočet integrálů. V určitém integrálu $\int\limits_0^1 f(x)\,dx$ považujme integrační proměnnou x za náhodnou proměnnou s rovnoměrným rozložením pravděpodobností v intervalu $\langle 0, 1\rangle$. Potom $f(x)$ je rovněž náhodná proměnná (ovšem obecně s jiným rozložením pravděpodobností) a integrál $\int\limits_0^1 f(x)\,dx$ má význam její střední hodnoty. Ze známých vět počtu pravděpodobnosti následuje pak tvrzení:

Nechť $I = \int\limits_a^b f(x)\,dx$ *je konečný Lebesgueův integrál v konečných mezích. Nechť* $J = \int\limits_a^b f^2(x)\,dx < +\infty$. *Nechť* $x_1, x_2, \ldots, x_n$ *jsou nezávislé náhodné proměnné, vesměs s rovnoměrným rozložením pravděpodobností v* $\langle a, b\rangle$.

Potom náhodná proměnná

$$\hat{I} = \frac{b-a}{n} \sum_{k=1}^{n} f(x_k)$$

je nestranným, asymptoticky normálním odhadem čísla I, *se směrodatnou odchylkou*

$$\sigma(\hat{I}) = [(b-a)\,J - I^2]^{-\frac{1}{2}}\, n^{-\frac{1}{2}} = \sigma n^{-\frac{1}{2}}\,.$$

To znamená

$$E(\hat{I}) = I\,, \tag{1}$$

kde E značí střední hodnotu;

$$P(|\hat{I} - I| < k_\varepsilon \sigma n^{-\frac{1}{2}}) \cong 1 - \varepsilon\,, \tag{2}$$

kde k_ε je dáno rovnicí

$$\sqrt{\frac{2}{\pi}} \int\limits_0^{k_\varepsilon} e^{-\frac{t^2}{2}}\,dt = 1 - \varepsilon.$$

Tato věta dává tedy methodu numerického výpočtu I. Aby však výpočet byl prakticky proveditelný (na př. pomocí tabulek náhodných čísel), je třeba připojit předpoklad (A), že uzávěr množiny bodů nespojitosti funkce $f(x)$ má Lebesgueovu míru 0.

Za tohoto předpokladu, je-li ε libovolné kladné číslo, existuje ke skoro každému $x \in \langle a, b\rangle$ přirozené číslo $s_0(x, \varepsilon)$ tak, že

$$|f(x) - f(\sum_{l=0}^{s} \alpha_l \,.\, 10^{-l})| < \varepsilon \text{ pro všechna } s \geqq s_0\,,$$

kde $\sum\limits_{l=0}^{\infty} \alpha_l \,.\, 10^{-l}$ značí desetinný rozvoj čísla x.

Pro každé $k = 1, \ldots, n$ lze náhodnou proměnnou x_k vyjádřit ve tvaru nekonečného desetinného zlomku

$$x_k = \sum_{l=0}^{\infty} \alpha_{kl} \cdot 10^{-l},$$

kde α_{kl}, $l = 1, 2, \ldots$ jsou náhodné proměnné, které nabývají pouze hodnot $0, 1, 2, \ldots, 9$. Je-li dáno $\varepsilon > 0$, definujme náhodnou proměnnou $s_{k,\varepsilon}$ jako nejmenší přirozené číslo s, pro něž platí nerovnost

$$\operatorname*{Max}_{\alpha_{k,s+1};\,\alpha_{k,s+2};\,\ldots} \left|f\left(\sum_{l=0}^{\infty}\alpha_{kl} \cdot 10^{-l}\right) - f\left(\sum_{l=0}^{s}\alpha_{kl} \cdot 10^{-l}\right)\right| < \frac{\varepsilon}{b-a}. \tag{3}$$

Z předpokladu (A) vyplývá, že náhodná proměnná $s_{k,\varepsilon}$ je konečná s pravděpodobností 1. Definujme posléze náhodnou proměnnou $y_{k,\varepsilon}$ vztahem $y_{k,\varepsilon} = \sum_{l=0}^{s_{k,\varepsilon}} \alpha_{kl} \cdot 10^{-l}$. Označíme-li jako $\hat{I}_\varepsilon$ náhodnou proměnnou $\frac{b-a}{n} \sum_{k=1}^{n} f(y_{k,\varepsilon})$, potom „chyba ze zaokrouhlení" $|\hat{I}_\varepsilon - \hat{I}|$ je menší než ε s pravděpodobností 1.

Lze tedy hodnoty náhodných proměnných x_k nahradit (na příklad) úseky o délce s číslic z tabulek náhodných čísel, považujeme-li tyto úseky za prvých s decimál veličiny x_k. Přitom s volíme tak velké, aby platila nerovnost (3).

Srovnejme přesnost stochastické methody s přesností běžných method, na př. lichoběžníkové a Simpsonovy. Zjistíme, že posledně jmenované methody dávají (při stejném počtu dělících bodů n) chybu řádově značně menší:

$$|I - I_{\text{lich}}| < C_1 n^{-2}; \quad |I - I_{\text{simp}}| < C_2 n^{-4}$$

(nehledě k tomu, že tyto nerovnosti platí jistě, zatím co (2) dává ohraničení chyby jen s určitou pravděpodobností).

V praxi však počítáme s nepříliš velkými hodnotami n a tu je třeba při srovnání různých method přihlédnout i k velikosti konstant C_1, C_2, $k_\varepsilon \sigma$:

$$C_1 = \frac{(b-a)^3}{12} \operatorname*{Max}_{a \leq x \leq b} |f''(x)|,$$

$$C_2 = \frac{(b-a)^5}{180} \operatorname*{Max}_{a \leq x \leq b} |f^{\mathrm{IV}}(x)|,$$

$$k_\varepsilon \sigma = k_\varepsilon [(b-a) J - I^2]^{\frac{1}{2}}.$$

Mohou nastat případy, kdy druhá (resp. čtvrtá) derivace funkce f je značně velká, není ohraničená v $\langle a, b \rangle$, nebo pro některá x vůbec neexistuje. Někdy lze tyto singularity vhodnými úpravami odstranit, někdy však nikoli.

Konstanta σ má naproti tomu tu pozoruhodnou vlastnost, že nezávisí na hodnotě derivací, a může být velmi malá i při složitém průběhu funkce f.

Jako příklad uveďme integrál

$$I = \int_0^{e^{-\pi}} \sqrt[10]{|\sin(\log x)|}\, dx .$$

K bezprostřednímu výpočtu tohoto integrálu nelze spolehlivě užít žádné z běžných numerických method. (V intervalu $\langle 0, e^{-\pi} \rangle$ je nekonečně mnoho bodů x, pro něž $f'(x)$ neexistuje.) Naopak stochastická methoda dává konstantu $\sigma \doteq 46 . 10^{-4}$, t. j.

$$|\hat{I} - I| < 46 k_\varepsilon . 10^{-4} n^{-\frac{1}{2}} \text{ s pravděpodobností } \cong 1 - \varepsilon ,$$

což je ve srovnání s velikostí hodnoty $I \doteq 393 . 10^{-4}$ uspokojivá přesnost. Zvolíme-li $n = 25$ a vezmeme-li jako náhodný výběr tohoto rozsahu prvých 25 řádků tabulek Kendall - Babington Smithových, dostáváme odhad

$$\hat{I} \doteq 3976 . 10^{-5} ,$$

zatím co skutečná hodnota integrálu jest

$$I \doteq 3934 . 10^{-5} ;$$

t. j. relativní chyba je menší než $1{,}1\%$.

3. Výpočet objemů těles v r-rozměrném eukleidovském prostoru. Nejprve připomeňme definice několika pojmů:

Tělesem v E_r budeme rozumět ohraničenou, uzavřenou oblast v E_r; *objemem tělesa* — jeho Jordanovu míru v E_r (pokud existuje).

Jednotkovou krychlí J_r budeme rozumět r-rozměrnou krychli objemu 1, s vrcholem v počátku, která celá leží v nezáporné části prostoru E_r.

Plochou v E_r rozumíme spojitý obraz jednotkové krychle J_{r-1}; plochu L v E_r nazveme *rektifikace schopnou,* jestliže existuje zobrazení φ a konstanta c tak, že platí

1. $L = \varphi(J_{r-1})$,
2. $\varrho_r(\varphi(x), \varphi(y)) \leqq c\, \varrho_{r-1}(x, y)$ pro každé $x \in J_{r-1}$, $y \in J_{r-1}$.

(Pro $r = 2$ — t. j. v rovině — souhlasí tento pojem s pojmem rektifikace schopné křivky.)

Položme si nyní tuto úlohu:

Jest vypočítati přibližně objem V tělesa M v E_r, definovaného nerovnostmi a implicitními vztahy mezi proměnnými, které nedovolují vyjádřit hranici tělesa explicitně (jako funkci parametrů), umožňují však rozhodnout o každém bodu $x \in E_r$, zda náleží nebo nenáleží do M.

Předpokládáme, že těleso M má objem; mimo to předpokládejme (zřejmě bez újmy obecnosti), že M je obsaženo v jednotkové krychli J_r.

Matematický způsob přibližného výpočtu objemu V spočívá v tom, že jednotkovou krychli J_r rozdělíme pravoúhlou sítí na n krychlí o hraně délky h (přitom jest $nh^r = 1$), z každé krychle vezmeme vrchol, který je nejblíže počátku, a zjistíme, zda náleží nebo nenáleží do M.

Nechť m těchto vrcholů náleží do M; potom přibližná hodnota je

$$\tilde{V} = mh^r = \frac{m}{n};$$

přitom platí $\tilde{V} \to V$ pro $n \to \infty$. Abychom získali řádový odhad chyby $|\tilde{V} - V|$, musíme připojit další předpoklad (B):

Hranice H tělesa M se skládá z konečného počtu rektifikace schopných ploch.

Za tohoto předpokladu platí

$$|\tilde{V} - V| = 0(h) = 0(n^{-\frac{1}{r}}). \tag{1}$$

(Obecně nelze tento odhad snížit. Řádově lepší odhad platí pro koule, elipsoidy a jistá speciální konvexní tělesa; objem těchto těles není však třeba počítat přibližnými methodami.)

Důkaz tvrzení (1). Předpokládejme pro jednoduchost, že hranici H tvoří jediná rektifikace schopná plocha $L = \varphi(J_{r-1})$. Rozdělme J_{r-1} na krychle K_i, vesměs o hraně délky $\frac{h}{(c + \varepsilon)\sqrt{r-1}}$, kde c má význam konstanty uvedené v definici rektifikace schopné plochy, ε je kladné číslo. Počet krychlí K_i jest $\left(\frac{(c+\varepsilon)\sqrt{r-1}}{h}\right)^{r-1} = Ah^{1-r}$, jestliže — opět pro jednoduchost — předpokládáme, že $\frac{(c+\varepsilon)\sqrt{r-1}}{h}$ je celé číslo. Tím dostáváme i rozdělení plochy L na Ah^{1-r} ploch $S_i = \varphi(K_i)$ o průměru $d(S_i) < h$; zřejmě každá plocha S_i může mít neprázdný průnik nejvýše se 2^r krychlemi o hraně h, na něž je rozdělena J_r. (Mezi libovolnými $2^r + 1$ krychlemi existují totiž aspoň dvě, jejichž vzdálenost je $\geq h$.) Celkem tedy má H neprázdný průnik nejvýš se $2^r A h^{1-r}$ krychlemi. Tvrzení (1) plyne odtud, že chyba $|\tilde{V} - V|$ je nutně nejvýš rovna objemu těchto krychlí, t. j. $|\tilde{V} - V| \leq 2^r Ah$.

Obraťme se nyní ke *stochastickému řešení* dané úlohy.

Nechť $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ jsou nezávislé r-rozměrné náhodné proměnné, vesměs s rovnoměrným v J_r rozložením. Nechť m' je počet těch $\mathbf{x}_i$, pro něž $\mathbf{x}_i \in M$. Ježto $P(\mathbf{x}_i \in M) = V$, má m' zřejmě binomické rozložení pravděpodobností $b(n; V)$, a tudíž náhodná proměnná

$$\hat{V} = \frac{m'}{n}$$

je nestranným, asymptoticky normálním odhadem čísla V, se směrodatnou odchylkou

$$\sigma(\hat{V}) = [V(1 - V)]^{\frac{1}{2}} n^{-\frac{1}{2}}.$$

Aby výpočet stochastickou methodou byl prakticky proveditelný, musíme opět připojit předpoklad (B). Pro každé $i = 1, 2, \ldots, n$ lze potom náhodnou proměnnou $\mathbf{x}_i = (\xi_{i1}, \xi_{i2}, \dots, \xi_{ir})$ nahradit náhodnou proměnnou $\mathbf{y}_i = (\eta_{i1}, \eta_{i2}, \dots, \eta_{ir})$, kde $\eta_{ik} = \sum_{l=1}^{s} \alpha_l \,.\, 10^{-l}$, je-li $\xi_{ik} = \sum_{l=1}^{\infty} \alpha_l \,.\, 10^{-l}$ (s je pevně zvolené přirozené číslo; α_l píšeme místo α_{ikl}). Pravděpodobnost P_0, že buďto $\mathbf{x}_i \in M$ a současně $\mathbf{y}_i \in' M$, ňebo $\mathbf{y}_i \in M$ a současně $\mathbf{x}_i \in' M$, je totiž nejvýš rovna objemu těch krychlí o hraně 10^{-s} (na něž si mysleme rozdělenu krychli J_r), které mají neprázdný průnik s plochou L; t. j. — jak plyne z důkazu (1) — $P_0 \leqq \text{const } 10^{-s}$.

Označíme-li jako m'' počet těch $\mathbf{y}_i$, pro něž $\mathbf{y}_i \in M$, potom rozdíl $\left|\frac{m''}{n} - \frac{m'}{n}\right|$ — ,,chyba ze zaokrouhlení'' — je náhodná proměnná, jejíž střední hodnota

$$E\left|\frac{m''}{n} - \frac{m'}{n}\right| \leqq \frac{1}{n} \sum_{k=1}^{n} P_0 \leqq \text{const } 10^{-s}$$

a směrodatná odchylka (jak se snadno zjistí) jest nejvýš rovna výrazu const $10^{-\frac{s}{2}} n^{-\frac{1}{2}}$. Volbou dostatečně velkého s lze chybu ze zaokrouhlení prakticky vyloučit.

Můžeme tedy r-rozměrnou veličinu $\mathbf{x}_i$ nahradit (na příklad) r úseky o délce s číslic z tabulek náhodných čísel, považujeme-li tyto úseky za konečné desetinné rozvoje jejích souřadnic $\xi_{i1}, \xi_{i2}, \dots, \xi_{ir}$.

Srovnejme přesnost obou method — matematické a stochastické — při stejném počtu n bodů, jejichž příslušnost k M zjišťujeme:

První methoda dává odhad

$$|\tilde{V} - V| < kn^{-\frac{1}{r}},$$

druhá — jak plyne z Čebyševovy nerovnosti — dává odhad $|\hat{V} - V| < k'_\varepsilon n^{-\frac{1}{2}}$ s pravděpodobností $> 1 - \varepsilon$, kde $k'_\varepsilon = \sqrt{V(1 - V)} \,.\, \varepsilon^{-\frac{1}{2}}$; to znamená

1. přesnost stochastické methody nezávisí na dimensi r,

2. pro $r > 2$ jest $n^{-\frac{1}{2}} < n^{-\frac{1}{r}}$, t. j. stochastická methoda dává v jistém smyslu lepší výsledek než matematická.

Nejlepší výsledek však dostaneme, jestliže obě methody kombinujeme:

Rozdělme opět J_r na n krychlí $I_1, I_2, \dots, I_n$ o hraně h. Označme V_i objem tělesa $M \cap I_i$. V každé krychli I_i zvolme náhodně bod $\tilde{\mathbf{x}}_i$ (t. zn. $\tilde{\mathbf{x}}_i$ je r-rozměrná náhodná proměnná s rovnoměrným v I_i rozložením pravděpodobností, $\tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_2, \dots, \tilde{\mathbf{x}}_n$ jsou nezávislé); budiž $\tilde{m}$ počet těch $\tilde{\mathbf{x}}_i$, pro něž $\tilde{\mathbf{x}}_i \in M$. Ježto $P(\tilde{\mathbf{x}}_i \in M) = nV_i$, má $\tilde{m}$ zřejmě Poissonovo binomické rozložení pravděpodobností, se střední hodnotou $E(\tilde{m}) = nV$ a rozptylem $\sigma^2(\tilde{m}) = \sum_{i=1}^{n} n\, V_i(1 - nV_i)$. Za předpokladu (B) jest počet krychlí I_i, pro něž $n\, V_i(1 - nV_i) \neq 0$ nejvýš řádu $0(h^{1-r}) = 0(n^{1-\frac{1}{r}})$. Poněvadž $n\,V_i(1 - nV_i) \leqq \frac{1}{4}$ $(i = 1, \ldots, n)$ jest $\sigma^2(\tilde{m}) = 0(n^{1-\frac{1}{r}})$.

Odtud následuje, že náhodná proměnná $\hat{\hat{V}} = \frac{\tilde{m}}{n}$ je nestranným odhadem čísla V se směrodatnou odchylkou $\sigma(\hat{\hat{V}}) = 0(n^{-\frac{r+1}{2r}})$.

Pro $r > 1$ jest $n^{-\frac{r+1}{2r}} < n^{-\frac{1}{2}} \leqq n^{-\frac{1}{r}}$, t. j. smíšená methoda je přesnější než methoda předchozí.

Všimněme si ještě jedné souvislosti s otázkou *mřížových bodů.* Z theorie čísel je znám tento výsledek:

Nechť L je uzavřená konvexní plocha v E_r, která obsahuje uvnitř bod $(0, \ldots, 0)$ a která má ve všech bodech totální křivost různou od nuly. Pro $x \geqq 0$ budiž $L(x)$ plocha, jež vznikne z L homothetickou transformací v poměru $\sqrt{x} : 1$ vzhledem k počátku. Označme $V(x)$ objem tělesa omezeného plochou $L(x)$; jako $A(x)$ označme počet bodů s vesměs celočíselnými souřadnicemi, které leží v tělese omezeném plochou $L(x)$.

Potom (za určitých dodatečných předpokladů o ploše L) platí:

$$\frac{A(x) - V(x)}{x^\Theta} \to 0 \text{ pro } x \to \infty, \text{ pro každé } \Theta > \frac{r}{2} - \frac{r}{r+1}.$$

Uvažujme nyní následující stochastickou modifikaci problému.

Nechť L je uzavřená rektifikace schopná plocha v E_r. Nechť $L(x)$ a $V(x)$ mají obdobný význam jako výše. V každé jednotkové krychli s celočíselnými vrcholy zvolme náhodně (a nezávisle na ostatních krychlích) jeden bod. Nechť $\tilde{A}(x)$ značí počet těchto bodů, které leží v tělese omezeném plochou $L(x)$. Potom platí $\frac{\tilde{A}(x) - V(x)}{x^\Theta} \to 0$ (pro $x \to \infty$) podle pravděpodobnosti, pro každé $\Theta > \frac{r-1}{4}$.

(Jest ovšem $\frac{r-1}{4} < \frac{r}{2} - \frac{r}{r+1}$ pro každé $r > 1$.)

4. Výpočet inversní matice. Připomeňme nejprve definici a základní vlastnosti *Markovových řetězců.* Markovovým řetězcem (přesněji: jednoduchým homogenním M-ým řetězcem) nazýváme níže popsaný proces:

Systém S probíhá v nespojitém čase $t = 0, 1, 2, \ldots$ konečnou množinu stavů $A_1, A_2, \ldots, A_r$ podle stochastického zákona, splňujícího podmínku: Podmíněná pravděpodobnost, že systém bude v čase t (> 0) ve stavu A_k, za předpokladu určitého průběhu předcházejícího (A_i v čase $t - 1$, A_h v čase $t - 2$, A_g v čase $t - 3, \ldots$), závisí pouze na A_i a na A_k a je konstantní vzhledem k t.

Označíme-li stavy přirozenými čísly $1, 2, \ldots, r$, lze Markovův řetězec interpretovat jako posloupnost náhodných proměnných, nabývajících celočíselných hodnot $1, 2, \ldots, r$ a splňujících zmíněnou podmínku.

Podmíněné pravděpodobnosti $P(x_t = k \mid x_{t-1} = i)$ se nazývají pravděpodobnosti přechodu a značí se p_{ik}. Hodnoty p_{ik} $(i, k = 1, 2, \ldots, r)$ tvoří nezápornou čtvercovou matici $\boldsymbol{P}$ s řádkovými součty vesměs rovnými jedné. Naopak každá matice těchto vlastností (t. zv. stochastická matice) definuje spolu s vektorem počátečních pravděpodobností určitý Markovův řetězec.

Prvky n-té mocniny matice $\boldsymbol{P}$ mají tento význam: $p_{ik}^{(n)} = P(x_{t+n} = k \mid x_t = i)$. Za určitých předpokladů existují $\lim\limits_{n\to\infty} p_{ik}^{(n)} = p_k^{(\infty)}$ jako čísla nezávislá na i. Stavy A_k, pro něž $p_k^{(\infty)} \doteq 0$, nazývají se přechodné, ostatní návratné. Existuje-li v řetězci jediný stav návratný, nazývá se stavem absorpčním. Název vystihuje tu okolnost, že při libovolných počátečních pravděpodobnostech přejde systém s pravděpodobností 1 v konečném čase do absorpčního stavu a nadále v něm setrvá.

Na základě theorie Markovových řetězců lze odvodit stochastickou methodu výpočtu inversní matice. Mějme regulární matici $\boldsymbol{A} = \|a_{ik}\|$ stupně r-tého takovou, že matice $\boldsymbol{P} = \boldsymbol{E} - \boldsymbol{A}$ je nezáporná, s řádkovými součty vesměs menšími jedné. Hledáme inversní matici $\boldsymbol{A}^{-1} = \|a_{ik}^{(-1)}\|$.

Uvažujme vroubenou matici $\tilde{\boldsymbol{P}}$, která vznikne z matice $\boldsymbol{P}$ připojením $r + 1$-ého řádku $(0, \ldots, 0, 1)$ a $r + 1$-ého sloupce $(p_1, \ldots, p_r, 1)$ kde $p_i = 1 - \sum\limits_{k=1}^{r} p_{ik}$.

Zřejmě $\tilde{\boldsymbol{P}}$ je stochastická matice; z jejího tvaru je patrno, že řetězec, který definuje, obsahuje absorpční stav $(r + 1)$. (To znamená, že skoro každá realisace řetězce skončí po konečně mnoha krocích ve stavu $(r + 1)$.)

Nechť τ je náhodná proměnná, která značí dobu setrvání systému ve třídě přechodných stavů:

$$\tau = \operatorname*{Max}_{x_t \neq r+1} t\,.$$

Definujme náhodné proměnné g_k $(k = 1, 2, \ldots, r)$ *takto:*

$$g_k = \begin{cases} \dfrac{1}{p_k}\,, & \textit{jestliže } x_\tau = k\,, \\ 0, & \textit{jestliže } x_\tau \neq k\,. \end{cases}$$

Potom platí

(1) $E(g_k \mid x_0 = i) = a_{ik}^{(-1)}$ $(i, k = 1, 2, \ldots, r)$,

(2) $\sigma_{ik}^2 = \sigma^2(g_k \mid x_0 = i) = \dfrac{1}{p_k}\, a_{ik}^{(-1)}(1 - p_k a_{ik}^{(-1)})$,

(3) $E(\tau \mid x_0 = i) = \sum\limits_{j=1}^{r} \sum\limits_{k=1}^{r} p_{ij} a_{jk}^{(-2)} p_k$.

Důkaz. 1. $E(g_k \mid x_0 = i) = \dfrac{1}{p_k} \sum\limits_{j=0}^{\infty} p_{ik}^{(j)} p_k = \sum\limits_{j=0}^{\infty} p_{ik}^{(j)}$,

t. j.

$$\|E(g_k \mid x_0 = i)\| = \sum_{j=0}^{\infty} \mathbf{P}^j = (\mathbf{E} - \mathbf{P})^{-1} = \mathbf{A}^{-1}.$$

Poslední rozvoj platí, neboť

$$\operatorname*{Max}_i |\lambda_i(\mathbf{P})| \leqq \operatorname*{Max}_i \sum_{k=1}^{r} p_{ik} < 1 .$$

2. $E(g_k^2 \mid x_0 = i) = \dfrac{1}{p_k^2} \sum_{j=0}^{\infty} p_{ik}^{(j)} p_k = \dfrac{1}{p_k} a_{ik}^{(-1)}$;

3. $E(\tau \mid x_0 = i) = \sum_{\tau=1}^{\infty} \tau \sum_{k=1}^{r} p_{ik}^{(\tau)} p_k$, t. j. $\|E(\tau \mid x_0 = i)\| = \sum_{\tau=1}^{\infty} \tau \mathbf{P}^\tau \mathbf{p} = \mathbf{P}(\mathbf{E} - \mathbf{P})^{-2}\mathbf{p} =$
$= \mathbf{P}\mathbf{A}^{-2}\mathbf{p}$, kde $\mathbf{p} = (p_1, p_2, \ldots, p_r)$.

S praktického hlediska je účelné odvodit odhady pro σ_{ik}^2 a $E(\tau)$, které neobsahují prvky neznámé matice $\mathbf{A}^{-1}$. Položme

$$p = \operatorname*{Max}_k p_k \,, \quad P = \operatorname*{Max}_k \frac{1}{p_k} \,.$$

Jako $N(\,.\,)$ označme normu matice, totiž největší z řádkových součtů prostých hodnot jejích prvků.

Potom platí

(4) $\sigma_{ik} \leqq \dfrac{1}{2p_k} \quad (i, k = 1, 2, \ldots, r)$,

(5) $N(\|\sigma_{ik}^2\|) \leqq P^2 - 1$.

(6) $E(\tau \mid x_0 = i) \leqq P^2\, p(1 - p_i)$.

Důkaz. 4. $p_k a_{ik}^{(-1)} = P(x_\tau = k \mid x_0 = i) \leqq 1$, odtud $\sigma_{ik}^2 \leqq \dfrac{1}{4p_k^2}$.

5. $N(\|\sigma_{ik}^2 + [a_{ik}^{(-1)}]^2\|) \leqq \operatorname*{Max}_k \dfrac{1}{p_k} N(\mathbf{A}^{-1}) \leqq P \,.\, \dfrac{1}{1 - N(\mathbf{P})} = P^2$.

Ježto $\mathbf{A}^{-1} = \sum_{j=0}^{\infty} \mathbf{P}^j \geqq \mathbf{E}$, jest $a_{ii}^{(-1)} \geqq 1$, $a_{ik}^{(-1)} \geqq 0$ pro všechna i, k, a tedy také $\sum_{k=1}^{r} [a_{ik}^{(-1)}]^2 \geqq 1$ pro všechna i.

6. $E(\tau \mid x_0 = i) \leqq \operatorname*{Max}_l p_l \,.\, \sum_{j=1}^{r} p_{ij} \sum_{k=1}^{r} a_{jk}^{(-2)} \leqq p\, N(\mathbf{A}^{-2}) \sum_{j=1}^{r} p_{ij} \leqq (1 - p_i)\, P^2 p$.

Při praktickém užití metody lze postupovat na příklad takto:

Provedeme n nezávislých realisací Markovova řetězce, při čemž počáteční stavy lze volit nenáhodně. Sestavíme čtvercovou tabulku hodnot n_{ik}, totiž absolutních četností realisací, pro něž $x_0 = i$, $x_\tau = k$. Řádky této tabulky dělíme pak řádkovými součty $n_i = \sum_k n_{ik}$, sloupce dělíme čísly p_k (první sloupec číslem p_1 atd.).

Prvky takto získané matice $\hat{A}$ jsou nestrannými odhady jim odpovídajících prvků inversní matice A^{-1}. Řádky matice $\hat{A}$ jsou (stochasticky) nezávislé; jak se snadno zjistí, má i-tý řádek asymptoticky r-rozměrné normální rozložení s kovarianční maticí

$$\frac{1}{n_i}\|{}_i\mu_{jk}\|\,, \quad \text{kde} \quad {}_i\mu_{kk} = \sigma^2_{ik}\,, \quad {}_i\mu_{jk} = -\,a^{(-1)}_{ij}a^{(-1)}_{ik} \quad \text{pro} \quad j \neq k\,.$$

(Závislosti uvnitř řádků matice $\hat{A}$ vyplývají odtud, že při tomto postupu odhadujeme současně celý řádek matice A^{-1}.)

Určitou modifikací popsané methody lze zeslabit předpoklady, omezující její použitelnost; o matici A stačí předpokládat, že $P = \text{mod}\,(E - A)$ je matice s řádkovými součty vesměs menšími jedné. (Symbolem mod (.) značíme matici, která vznikne z dané matice nahrazením každého prvku jeho prostou hodnotou.)

Modifikace spočívá v tom, že náhodné proměnné g_k je třeba definovat takto: (q_{ij} značí prvek matice $E - A$)

$$g_k = \begin{cases} \dfrac{1}{p_k}\,.\,\text{sgn}\,(q_{x_0x_1}q_{x_1x_2}\ldots q_{x_{\tau-1}x_\tau}), & \text{jestliže } x_\tau = k\,, \\ 0, & \text{jestliže } x_\tau \neq k\,. \end{cases}$$

Pro praktický výpočet to znamená, že po každé realisaci zaznamenáváme do čtvercové tabulky $\pm$ 1, podle toho, jakým způsobem přešel systém ze stavu počátečního do stavu absorpčního.

Výrazy pro σ^2_{ik} a $E(\tau)$ je třeba pozměnit:

$$\sigma^2_{ik} = \frac{t^{(-1)}_{ik}}{p_k} - [a^{(-1)}_{ik}]^2\,; \quad \sigma_{ik} < \frac{1}{p_k}\,;$$

$$N(\|\sigma^2_{ik}\|) < P^2\,; \quad E(\tau \mid x_0 = i) = \sum_{j=1}^{r}\sum_{k=1}^{r} p_{ij}t^{(-2)}_{jk}p_k\,,$$

kde $T = E - P$. Ostatní platí beze změny.

Upozorněme ještě na to, že stochastická methoda umožňuje vypočítat určitý prvek (nebo řádek) inversní matice, aniž by bylo třeba počítat celou inversní matici.

Na následujícím — velmi jednoduchém — příkladě ukážeme, jak lze provést výpočet pomocí tabulek *náhodných čísel.*

Mějme matici

$$A = \begin{Vmatrix} 0{,}8 & -0{,}1 \\ -0{,}1 & 0{,}8 \end{Vmatrix}.$$

Potom

$$P = \begin{Vmatrix} 0{,}2 & 0{,}1 \\ 0{,}1 & 0{,}2 \end{Vmatrix}, \quad \tilde{P} = \begin{Vmatrix} 0{,}2 & 0{,}1 & 0{,}7 \\ 0{,}1 & 0{,}2 & 0{,}7 \\ 0 & 0 & 1 \end{Vmatrix}.$$

Vezměme dvojici sousedních sloupců číslic v tabulkách náhodných čísel; tyto sloupce považujme za stavy (1) a (2). Začněme na příklad ve stavu (1) a postupujeme shora dolů podle tohoto předpisu:

číslice 0 značí přechod z (1) do (2), resp. opačně;

číslice 1, 2 značí setrvání v daném stavu;

číslice 3 až 9 značí přechod do stavu (3), t. j. ukončení řetězce.

Tento předpis zřejmě realisuje řetězec, definovaný stochastickou maticí $\tilde{\mathbf{P}}$.

Vzhledem ke speciálnímu tvaru matice $\mathbf{A}$ stačí vypočítat pouze první řádek inversní matice $\mathbf{A}^{-1}$. Použijeme-li tabulek Kendall-Babington Smithových a volíme-li $n = 1000$, dostáváme odhad

$$\hat{\mathbf{A}} \doteq \begin{Vmatrix} 1{,}273 & 0{,}156 \\ \cdot & \cdot \end{Vmatrix}$$

zatím co správný výsledek jest (zaokrouhleno na 3 des. místa)

$$\mathbf{A}^{-1} = \begin{Vmatrix} 1{,}270 & 0{,}159 \\ 0{,}159 & 1{,}270 \end{Vmatrix}.$$

Ostatní konstanty:

$$\sigma_{ik} \doteq 0{,}45 \quad (i, k = 1, 2) \quad (\text{odhad } (4) \ \sigma_{ik} \leqq 0{,}71)\,,$$
$$E(\tau \mid x_0 = i) \doteq 0{,}34 \quad (i = 1, 2) \quad (\text{odhad } (6) \ E(\tau \mid x_0 = i) \leqq 0{,}43)\,.$$

5. Ohraničení čísla π. Někdy vedou stochastické methody k výsledkům, které platí jistě — nikoli jen s určitou pravděpodobností. Jako příklad uvedeme stochastický důkaz tvrzení, že číslo π leží v intervalu $\langle 3{,}1380;\ 3{,}1481 \rangle$.

Mějme v rovině dánu trojúhelníkovou síť, tvořenou třemi soustavami rovnoběžných přímek o jednotkové kolmé vzdálenosti, při čemž přímky těchto soustav svírají s osou X úhly $0, \frac{\pi}{3}, \frac{2\pi}{3}$.

Na rovinu házejme náhodně úsečku délky l. (Slovu „náhodně" je zde rozuměti takto: poloha (x, y) středu úsečky je dvojrozměrná náhodná proměnná s libovolným rozložením; úhel Θ, který svírá úsečka s osou X, je náhodná proměnná s rovnoměrným rozložením v intervalu $\langle 0, \pi \rangle$; (x, y) a Θ jsou nezávislé.) Nechť c značí počet přímek sítě, které úsečka protne, dělený délkou úsečky. Dopadne-li úsečka pod úhlem Θ, nabude náhodná proměnná c hodnoty

$$\frac{1}{l}\left([l|\sin\Theta|] + \left[l\left|\sin\left(\Theta - \frac{\pi}{3}\right)\right|\right] + \left[l\left|\sin\left(\Theta - \frac{2\pi}{3}\right)\right|\right] + \alpha\right)$$

kde hranaté závorky značí největší celé části a α je některé z čísel 0, 1, 2, 3, v závislosti na poloze středu úsečky.

Nahradíme-li c náhodnou proměnnou c', která v tomto případě nabude hodnoty

$$|\sin\Theta| + \left|\sin\left(\Theta - \frac{\pi}{3}\right)\right| + \left|\sin\left(\Theta - \frac{2\pi}{3}\right)\right|$$

nezávisle na poloze středu úsečky, dopustíme se chyby, kterou lze učinit libovolně malou, volíme-li l dostatečně velké.

Pro momenty náhodné proměnné c' dostáváme tyto výrazy:

$$E(c') = \frac{1}{\pi} \int_0^{\pi} \left(|\sin \Theta| + \left|\sin\left(\Theta - \frac{\pi}{3}\right)\right| + \left|\sin\left(\Theta - \frac{2\pi}{3}\right)\right| \right) d\Theta = \frac{6}{\pi} ;$$

$$E(c'^2) = \frac{1}{\pi} \int_0^{\pi} \left(|\sin \Theta| + \left|\sin\left(\Theta - \frac{\pi}{3}\right)\right| + \left|\sin\left(\Theta - \frac{2\pi}{3}\right)\right| \right)^2 d\Theta = 2 + \frac{3\sqrt{3}}{\pi} ;$$

$$\sigma^2 = \sigma^2(c') = 2 + \frac{3\sqrt{3}}{\pi} - \frac{36}{\pi^2} .$$

Úpravou posledního vztahu dostaneme kvadratickou rovnici pro π:

$$(2 - \sigma^2)\, \pi^2 + 3\sqrt{3}\pi - 36 = 0\,, \quad \text{t. j. } \pi = \frac{-3\sqrt{3} + \sqrt{27 + 144(2 - \sigma^2)}}{2(2 - \sigma^2)} .$$

Tento zlomek definuje funkci $\hat{\pi}(\sigma^2)$, považujeme-li σ^2 za proměnnou.

Snadno se zjistí, že σ^2 leží nutně v mezích od 0 do $\frac{7 - 4\sqrt{3}}{4}$. Hodnota 0 odpovídá případu, kdy úsečka protne při každém hodu týž počet přímek; hodnota $\frac{7 - 4\sqrt{3}}{4}$ odpovídá případu, kdy v polovině hodů úsečka protne minimální počet přímek ($c' = \sqrt{3}$; to nastane tehdy, dopadne-li úsečka rovnoběžně s některou soustavou) a v polovině hodů protne maximální počet přímek ($c' = 2$; to nastane tehdy, půlí-li úsečka úhel dvou sousedních soustav).

Ježto $\hat{\pi}(\sigma^2)$ je v daném intervalu rostoucí a $\hat{\pi}(0) = 3{,}1380\ldots$, $\hat{\pi}\left(\frac{7 - 4\sqrt{3}}{4}\right) = 3{,}14803\ldots$, platí $3{,}1380 < \pi < 3{,}1481$.

LITERATURA

K § 1: Řešení Buffonovy úlohy viz na př.:

[1] *B. V. Gněděnko*: Kurs těoriji věrojatnostěj, Moskva 1950, str. 35; experimentální určení čísla π viz tamtéž, str. 347.

K § 2: Jak vyplývá z recense v Math. Rev. 14 (1953), str. 457, zabývá se různými stochastickými methodami výpočtu určitých integrálů $\int_0^1 f(x)\, dx$ ze spojité funkce $f(x)$ T. Kitagawa v článku:

[2] *T. Kitagawa*: Random integrations, Bull. Math. Statist. 4 (Fukuoka, 1950), str. 15—21.

K § 3: Viz článek:

[3] *K. D. Tocher*: Application of automatic computers to sampling experiments, Jrn. Roy. Stat. Soc., Series B 1954/No. 1, str. 49.

Tocher neuvádí předpoklady, za nichž odhady platí. Smíšená methoda a poznámka o „znáhodněných" mřížových bodech jsou nové.
Citovanou větu z theorie mřížových bodů viz:

[4] *S. Krupička*: O mřížových bodech ve vícerozměrných prostorech (disertační práce).
K § 4: Myšlenka pochází od *J. v. Neumanna* a *S. Ulama*. Po prvé byla publikována (v poněkud obecnější formě) v článku:

[5] *G. E. Forsythe & R. A. Leibler*: Matrix inversion by a Monte Carlo method, Math. Tab., 1950, str. 127—129.

Vzorce (3), (4), (5), (6) jsou nové.

K § 5: Uvedený výsledek je obměnou výsledku Mantelova, který užívá čtverečkové sítě a dokazuje, že π leží v intervalu $\langle 3{,}1231;\ 3{,}1752\rangle$ — viz:

[6] *N. Mantel*: An extension of the Buffon needle problem, Ann. Math. Stat. 24 (1953), str. 674—677.

Резюме

СТОХАСТИЧЕСКИЕ ЧИСЛЕННЫЕ МЕТОДЫ

ВАЦЛАВ ДУПАЧ (Václav Dupač), Прага.

(Поступило в редакцию 14/II 1955 г.)

В статье докладывается о применении так называемых методов Монтэ Карло в численном анализе. В § 1-ом природа этого применения объясняется на примере экспериментального определения числа π. В § 2-ом изучаются свойства величины $\hat{I} = (b - a)\, n^{-1} \sum_{k=1}^{n} f(x_k)$, — где x_k независимые, равномерно распределенные в $\langle a, b\rangle$ случайные величины, как статистической оценки интеграла $\int_a^b f(x)\, dx$. В § 3-ем исследуется подробнее мысль Точера [3] о стохастическом вычислении объемов многомерных тел, определенных сложными неявными взаимоотношениями между координатами. Приходится к методу, который является комбинацией математического метода и метода Монтэ Карло, и который дает найлучшую оценку ошибки. Приводится результат, который можно назвать стохастическим видоизменением проблемы целых точек. В § 4-ом описывается обращение матрицы методом Форсейта-Лейблера [5]. Новыми являются формулы (3)—(6), т. е. верхние оценки для σ_{ik}^2 и $E(\tau)$, несодержащие элементов незнакомой обратной матрицы. В § 5-ом доказано только при помощи элементарных средств теории вероятностей, что π лежит в промежутке $\langle 3{,}1380;$

3,1481⟩. Этот результат является небольшим улучшением результата, приведенного в работе Мантела [6]. Выводы §§ 2-ого и 4-ого иллюстрированы двумя численными примерами.

Summary

STOCHASTIC NUMERICAL METHODS

VÁCLAV DUPAČ, Praha.

(Received February 14, 1955.)

The paper reports about the application of the so-called Monte Carlo methods in numerical calculation. The nature of this application is enlightened in § 1 by the example of the empirical determination of π. In § 2 the statistic $I = (b - a)\, n^{-1} \sum_{k=1}^{n} f(x_k)$ — where x_k are independent, uniformly distributed in $\langle a, b\rangle$ — is studied, as an estimator of the integral $\int_a^b f(x)\, dx$. (Cf. Kitagawa [2].) A numerical example is given. In § 3 an idea due to Tocher [3] is developed concerning the Monte Carlo evaluation of volumes of multidimensional bodies which are defined by complicated implicit relations between coordinates. A method is deduced which is a combination of the mathematical method and the Monte Carlo one, and which gives (in certain sense) the best accuracy. A result is given which can be called the stochastic modification of the lattice points problem. In § 4, matrix inversion by a Monte Carlo method is described, according to Forsythe-Leibler [5]. New are the formulae (3)—(6), i. e. the upper estimates of σ^2 and $E(\tau)$ in which the elements of the unknown inverse matrix do not occur. In § 5, the assertion $\pi \in (3{,}1380; 3{,}1481)$ is proved in a purely probabilistic way. This result is a slight improvement of an analogous result contained in a paper of Mantel [6].