

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

Дніпровський національний університет імені Олеся Гончара

Механіко-математичний факультет

Кафедра статистики й теорії ймовірностей

100 років

*Дніпровському національному університету імені Олеся Гончара
(1918–2018)*

Гончаров С. В.

КОНСПЕКТ ЛЕКЦІЙ

ІЗ СТАТИСТИЧНОГО МОДЕЛЮВАННЯ

Дніпро
Видавець Біла К. О.
2018

*Рекомендовано до друку
Вченою радою механіко-математичного факультету
Дніпровського національного університету імені Олеся Гончара
(протокол № 8 від 27.03.2018)*

Рецензенти:

Послайко Н. І., канд. фіз.-мат. наук, доцент кафедри прикладної математики Дніпропетровського національного університету залізничного транспорту імені академіка В. Лазаряна;

Руденко О. О., канд. фіз.-мат. наук, доцент кафедри математичного аналізу і теорії функцій Дніпровського національного університету імені Олеся Гончара.

Гончаров С. В.

Г 65 Конспект лекцій із статистичного моделювання / С. В. Гончаров. – Дніпро : Видавець Біла К. О., 2018. – 44 с.

Посібник містить лекції з курсу статистичного моделювання. Лекції охоплюють наступні теми: рівномірний розподіл з теоретико-ймовірнісної та теоретико-числової точок зору, моделювання рівномірного розподілу, генератори псевдовипадкових чисел, емпіричні критерії рівномірності розподілу, моделювання абсолютно неперервних та дискретних розподілів, обчислення інтегралів методом Монте-Карло, ймовірнісні методи розв'язку СЛАР, застосування методів Монте-Карло.

Для студентів спеціальності «Статистика» механіко-математичного факультету ДНУ, а також для усіх бажаючих ознайомитись з основами статистичного моделювання.

УДК 519.2

Зміст

Вступ	4
1. Моделювання рівномірного розподілу	6
1.1. Рівномірний розподіл (теоретико-ймовірнісне тлумачення)	6
1.2. Рівномірний розподіл (теоретико-числове тлумачення)	9
1.3. Генератори псевдовипадкових чисел	12
1.3.1. Лінійний конгруентний метод	12
1.3.2. Інші ГПВЧ	14
1.4. Емпіричні критерії рівномірності	15
2. Моделювання заданого розподілу	19
2.1. Абсолютно неперервні розподіли	19
2.2. Дискретні розподіли	20
2.3. Метод відбору та бітові алгоритми	22
2.4. Перетворення щільності	23
2.5. Вибірки та перестановки	25
2.6. Випадкові процеси	26
3. Метод Монте-Карло	28
3.1. Обчислення інтегралів	28
3.1.1. Виділення головної частини	28
3.1.2. Зниження порядку інтегрування	29
3.1.3. Заміна міри (метод суттєвої вибірки)	30
3.1.4. Розшарування вибірки	31
3.1.5. Випадкові інтерполяційно-квадратурні формули	32
3.2. Ймовірнісні методи розв'язку СЛАР	35
4. Застосування методів Монте-Карло	38
4.1. Моделювання руху частинок крізь речовину	38
4.2. Моделювання системи масового обслуговування	41
Позначення	43
Список рекомендованої літератури	43

Вступ

Під *статистичним* моделюванням ми розуміємо:

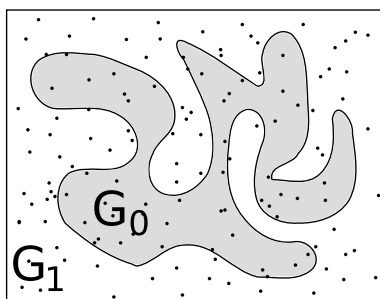
1) Генерування послідовностей, які можна вважати реалізаціями випадкових величин (ВВ) з конкретними розподілами, — моделювання ВВ.

2) Використання цих послідовностей для наближеного обчислення середніх характеристик ймовірнісних моделей деяких об'єктів. До таких характеристик відносяться математичне сподівання, дисперсія, коваріація тощо.

Отже, мова йде про чисельні методи розв'язку математичних задач моделюванням ВВ. Часто вживають назву «Метод Монте-Карло» (ММК).

Початок розвитку цих методів відносять до кінця 1940-х років, зокрема — до статті 1949-го року «Метод Монте-Карло» фон Неймана, Улама та Метрополіса. Їх використовують, коли аналітичні і «звичайні» чисельні методи або не дають розв'язку, або дозволяють лише встановити його грубі оцінки.

Основним прийомом є зведення задачі до обчислення математичного сподівання. Проілюструємо це прикладом.



Необхідно знайти площу області G_0 «складної» форми, — аналітичними або чисельними методами зробити це важко. Помістимо G_0 всередину області G_1 , яка має «просту» (на малюнку — прямокутну) форму, і площа якої відома, нехай це S_{G_1} . На G_1 означимо функцію

$$\xi(x; y) = \begin{cases} S_{G_1}, & (x; y) \in G_0, \\ 0, & (x; y) \notin G_0, \end{cases}$$

де $(x; y)$ — координати точки в області G_1 . Якщо цю точку навмання кидають в G_1 , то вона розподілена рівномірно в G_1 зі щільністю $\rho(x; y) = \frac{1}{S_{G_1}}$, і

$$M\xi = \int_{G_1} \xi(x; y) \rho(x; y) dx dy = \frac{1}{S_{G_1}} \int_{G_0} \xi(x; y) dx dy = \int_{G_0} dx dy = S_{G_0}$$

Тепер кинемо в область G_1 «багато» точок навмання, або, іншими словами, отримаємо багато реалізацій рівномірно розподіленої в G_1 ВВ — $(x_1; y_1)$, $(x_2; y_2)$, ..., $(x_n; y_n)$. Тоді емпірична оцінка $\hat{S}_0 = \frac{1}{n} \sum_{k=1}^n \xi(x_k; y_k)$ є «хорошою» — незміщеною та консистентною — оцінкою параметра $M\xi = S_{G_0}$: $\hat{S}_0$ для достатньо великих n буде з великою ймовірністю близька до шуканої S_{G_0} .

Скільки часу займе такий підрахунок S_{G_0} ? Це залежить від того, як швидко ми для $(x; y) \in G_1$ здатні з'ясувати, чи $(x; y) \in G_0$, аби обчислити $\xi(x; y)$.

Наскільки рівномірно розподілені кинуті точки? Якщо вони зосередяться в/поза G_0 , то $\hat{S}_0$ буде поганою оцінкою S_{G_0} .

Використовуючи ММК, ми повинні вміти:

- 1) Моделювати вибірку з рівномірного розподілу на $[0; 1]$;
- 2) На її основі моделювати вибірку із певного розподілу $F(\cdot)$;
- 3) Обчислювати емпіричну оцінку математичних сподівань.

Є кілька способів здійснення (1):

Метод	Pro	Contra
Таблиці	Одноразова перевірка. Можливість відтворення	Запас чисел обмежений. Займає багато місця у зовнішній пам'яті. Повільний доступ
Датчики	Запас необмежений. Надшвидке отримання. Місця (у пам'яті) не займає	Нестабільна поведінка. Перевірка періодичності. Неможливо відтворити. Потрібний прилад
Псевдовипадкові числа: $x_{n+1} = f(x_n)$	Одноразова перевірка. Можливість відтворення. Швидке отримання. Місця не займає. Не потрібні зовнішні прилади	Запас чисел обмежений. «Випадковість» під питанням (наступні члени послідовності детерміновані попередніми)

Відтворення означає, що розрахунки можна повторити в іншому місці, якщо відомі кілька початкових параметрів, і перевірити результат.

Використовуватимемо псевдовипадкові числа, адже знаючи відповідні алгоритми, їх можна запрограмувати майже на кожному комп'ютері «загального призначення», у той час як таблиці і датчики не завжди доступні.

1. Моделювання рівномірного розподілу

1.1. Рівномірний розподіл (теоретико-ймовірнісне тлумачення)

ОЗН. Рівномірно розподілену на $[0; 1]$ ВВ будемо називати *стандарівно-мірною* (СР).

Відмітимо важливі для нас властивості такої величини.

Розглядатимемо два типи подібних величин — дискретну та неперервну.

Точніше, говоритимемо про випадкову цифру (ВЦ) — $\beta \sim \begin{pmatrix} 0 & 1 & 2 & \dots & 9 \\ \frac{1}{10} & \frac{1}{10} & \frac{1}{10} & \dots & \frac{1}{10} \end{pmatrix}$

(загальніше — випадкова r -ічна цифра, $\beta \sim \begin{pmatrix} 0 & 1 & 2 & \dots & r-1 \\ \frac{1}{r} & \frac{1}{r} & \frac{1}{r} & \dots & \frac{1}{r} \end{pmatrix}$) та випад-

кове число (ВЧ) α зі щільністю розподілу $p(x) = \begin{cases} 1, & x \in [0; 1], \\ 0, & x \notin [0; 1]. \end{cases}$

ВЦ може бути декілька, або нескінченна (зліченна) послідовність $\beta_1, \beta_2, \dots, \beta_n, \dots$. Аналогічно для ВЧ α .

Властивість 1 пов'язує між собою ці два типи ВВ:

В. 1. ВЧ $\alpha = 0.\beta_1\beta_2\dots\beta_m\dots$ СР **тоді** дот ВЦ $\{\beta_i\}_{i=1}^\infty$ Н/З (в сукупності) з однаковим розподілом $\beta_i \sim \begin{pmatrix} 0 & 1 & 2 & \dots & 9 \\ \frac{1}{10} & \frac{1}{10} & \frac{1}{10} & \dots & \frac{1}{10} \end{pmatrix}$.

◀ 1) Якщо α — СР, то

$$P\{\beta_{i_1} = b_1, \beta_{i_2} = b_2, \dots, \beta_{i_k} = b_k\} = \frac{1}{10^k} = \prod_{j=1}^k P\{\beta_{i_j} = b_j\}$$

що означає Н/З величин β_i . Перша рівність випливає із того, що кожна умова виду $\beta_{i_j} = b_j$ відкидає $\frac{9}{10}$ тої частини відрізка $[0; 1]$, що залишилась від попередніх умов, так що міра решти зменшується в 10 разів; друга рівність є наслідком симетрії ($P\{\beta_{i_j} = 0\} = P\{\beta_{i_j} = 1\} = \dots = P\{\beta_{i_j} = 9\} = \frac{1}{10}$).

2) Якщо β_i Н/З і мають розподіл ВЦ, то $\forall x = 0.x_1x_2\dots x_m\dots \in [0; 1]$:

$$\begin{aligned} P\{\alpha < x\} &= P\{0.\beta_1\beta_2\dots\beta_m\dots < 0.x_1x_2\dots x_m\dots\} = \\ &= \sum_{k=1}^{\infty} P\{\beta_1 = x_1, \beta_2 = x_2, \dots, \beta_{k-1} = x_{k-1}, \beta_k < x_k\} = \\ &= \sum_{k=1}^{\infty} P\{\beta_1 = x_1\} \times \dots \times P\{\beta_{k-1} = x_{k-1}\} \cdot P\{\beta_k < x_k\} = \\ &= \sum_{k=1}^{\infty} \frac{1}{10} \times \dots \times \frac{1}{10} \cdot P\{\beta_k = 0, 1, \dots, x_k - 1\} = \sum_{k=1}^{\infty} \frac{1}{10^{k-1}} \cdot \frac{x_k}{10} = \sum_{k=1}^{\infty} \frac{x_k}{10^k} = x \end{aligned}$$

— це ФР рівномірного розподілу на $[0; 1]$, тобто $\alpha \in$ СР. ▶

Наступні властивості відносяться до випадкових векторів (ВипВек) з Н/З рівномірно розподіленими компонентами.

В. 2. Якщо $\alpha_1, \alpha_2, \dots, \alpha_n, \dots$ — Н/З СР ВЧ, а $s \in \mathbb{N}$, то ВипВек $\xi_i = (\alpha_{(i-1)s+1}; \alpha_{(i-1)s+2}; \dots; \alpha_{is})$, де $i \in \mathbb{N}$, є Н/З та рівномірно розподіленими в одиничному гіперкубі $D_s = [0; 1]^s$.

В. 2а. За тих же умов ВипВек $\eta_i = (\alpha_i; \alpha_{i+1}; \dots; \alpha_{i+s-1})$, де $i \in \mathbb{N}$, має рівномірний розподіл в D_s .

В. 3. Якщо змінити нумерацію в будь-якій скінченній підмножині елементів послідовності $\{\alpha_i\}_{i=1}^\infty$ довільним чином (але не залежним від значень α_i) — переставити їх — то В. 2 і тоді матиме місце.

Наступні дві властивості є частинними випадками посиленого закону великих чисел (ПЗВЧ) та центральної граничної теореми (ЦГТ); ВипВек $\xi_i = (\alpha_{i1}; \alpha_{i2}; \dots; \alpha_{is})$ утворені Н/З СР ВЧ.

В. 4. Нехай $f(x) = f(x_1; x_2; \dots; x_s)$ — довільна інтегровна за Лебегом в $D_s = [0; 1]^s$ функція. **Тоді**

$$P\left\{\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n f(\xi_i) = \int_{D_s} f(x) dx\right\} = 1$$

В. 5. В умовах В. 4, якщо додатково $\int_{D_s} |f(x)|^2 dx < \infty$, то

$$\lim_{n \rightarrow \infty} P\left\{\left|\frac{1}{n} \sum_{i=1}^n f(\xi_i) - \int_{D_s} f(x) dx\right| < y\right\} = \sqrt{\frac{2}{\pi}} \int_0^y e^{-\frac{u^2}{2}} du$$

де $\sigma_f^2 = \int_{D_s} f^2(x) dx - \left(\int_{D_s} f(x) dx\right)^2 = Df(\alpha)$.

В. 4 та 5 обґрунтовують використання ВЧ для наближеного обчислення інтегралів виду $I = \int_{D_s} f(x) dx$: із В. 4 випливає, що I можна з ймовірністю 1

наблизити середнім $\frac{1}{n} \sum_{i=1}^n f(\xi_i)$, а якщо $\int_{D_s} f^2(x) dx < \infty$, то — переписавши вираз із В. 5, для достатньо великих n , коли можна вважати, що виконується рівність — отримуємо:

$$P\left\{\left|\frac{1}{n} \sum_{i=1}^n f(\xi_i) - \int_{D_s} f(x) dx\right| < \frac{\sigma_f}{\sqrt{n}} y\right\} \approx \sqrt{\frac{2}{\pi}} \int_0^y e^{-\frac{u^2}{2}} du$$

тоді для $p \approx 1$, взявши $y = y_p$ так, щоб $\sqrt{\frac{2}{\pi}} \int_0^{y_p} e^{-\frac{u^2}{2}} du = p$, матимемо:

$$P\left\{\left|\frac{1}{n}\sum_{i=1}^n f(\xi_i) - \int_{D_s} f(x)dx\right| < \frac{\sigma_f}{\sqrt{n}}y_p\right\} \approx p$$

збільшуючи далі n , можемо зробити похибку $\frac{\sigma_f}{\sqrt{n}}y_p$ як завгодно малою. Маємо швидкість збіжності та оцінку похибки.

Власт. 1–5 — це властивості «справжніх» ВЧ. Але на практиці ми маємо справу лише з конкретними реалізаціями — в принципі нескінченними, $\alpha_1, \alpha_2, \dots, \alpha_n, \dots$ — які можна вважати реалізацією Н/З СР ВЧ, вибіркою із такого розподілу. Вважати чи не вважати — це рішення ми повинні приймати, обґрунтовуючи його; послідовність треба перевірити.

Скористаємось апаратом статистичних гіпотез. Тобто, висуваємо гіпотези, або *тести*, $H_1, H_2, \dots, H_k$ щодо даної послідовності і перевіряємо їх. Вважатимемо послідовність «хорошою», тільки коли вона пройде всі тести.

Якщо послідовність пройшла тести $H_1, H_2, \dots, H_k$, це не означає, що вона пройде тест H_{k+1} . Але, якщо тестів достатньо багато, то послідовність вважається «хорошою», допоки не провалено жодного тесту.

У якості основних критеріїв для перевірки статистичних гіпотез будемо вживати критерій χ^2 та критерій Колмогорова.

Критерій χ^2 . $[0; 1]$ розбивається на m попарно неперетинних вимірних множин (*категорій*) A_i . Позначивши $p_i = \mu A_i = P\{\alpha \in A_i\} > 0$ та $\nu_i = n(\{\alpha_j \in A_i\})$, обчислюємо відхилення Пірсона

$$D = D(\hat{F}_n; G) = \sum_{i=1}^m \frac{(\nu_i - np_i)^2}{np_i} = \frac{1}{n} \sum_{i=1}^m \frac{\nu_i^2}{p_i} - n$$

(де n — загальна кількість реалізацій) і, обравши рівень значущості (РЗ) δ (0.01, 0.05, ...), порівнюємо D з $\chi_{\delta; m-1}^2$ — верхньою δ -межею χ^2 -розподілу з $(m-1)$ ступенями вільності.

Якщо $D > \chi_{\delta; m-1}^2$, відхиляємо гіпотезу про СР розподілу, з якого взято вибірку $\{\alpha_j\}$, інакше — приймаємо. При цьому, як відомо, ймовірність помилково відхилити вірну гіпотезу, $P\{ERR - 1\} \approx \delta$.

Використовуючи той же критерій, можна розбивати не $[0; 1]$, а гіперкуб D_s ; наприклад, розглядаючи пари $(\alpha_{2i-1}; \alpha_{2i})$.

Критерій Колмогорова. Маючи реалізації $\alpha_1, \alpha_2, \dots, \alpha_n$ ВЧ α , будуємо ЕФР $\hat{F}_n(x) = \frac{\mu_x^*}{n}$, де μ_x^* — кількість $\alpha_i < x$, і для гіпотетичного розподілу $G(\cdot)$, СР з ФР $G(x) = x$, обчислюємо відхилення Колмогорова

$$\sup_x |\hat{F}_n(x) - G(x)| = \max_{1 \leq k \leq n} \{ |G(\alpha_k^*) - \hat{F}_n(\alpha_k^*)|; |G(\alpha_k^*) - \hat{F}_n(\alpha_{k+1}^*)| \}$$

де $\{\alpha_k^*\}$ — упорядковані за зростанням $\{\alpha_k\}$, та порівнюємо його з $\varepsilon_{\delta; n}$ — критичним значенням ухилення $D(\hat{F}_n; F)$. Якщо $\sup |\dots| > \varepsilon_{\delta; n}$, відхиляємо гіпотезу про СР ВЧ α , інакше — приймаємо. При цьому $P\{ERR - 1\} \leq \delta$.

Нехай послідовність $\{\alpha_i\}$ отримана за допомогою генератора псевдовипадкових чисел. Очевидно, що самі по собі тести не зроблять конкретний генератор «хорошим», вони можуть лише перевірити його. Для цього алгоритм такого генератора необхідно будувати не аби як, а спираючись на певні теоретичні міркування (можливо, пов'язані із статистичними тестами).

ЗАУВ. Дослідження Колмогорова та його наступників показали, що генератор, який би породжував послідовність $\alpha_1, \alpha_2, \dots$, котра проходить «усі можливі» тести на «випадковість», має «нескінченну складність».

Послідовність $\{\alpha_i\}$, обчислену за допомогою детермінованої програми, називатимемо *псевдовипадковою*, якщо вона «наближено» має властивості 1–5. Але якщо вона детермінована (α_1 однозначно визначає усі наступні α_i), то що ми розумітимемо під розподілами, незалежністю тощо? Звернімося до теорії чисел, теоретико-числових понять та засобів.

1.2. Рівномірний розподіл (теоретико-числове тлумачення)

Коли називати послідовність $\varepsilon = \{\varepsilon_i\}_{i \in \mathbb{N}}$ з $[0; 1]$ рівномірно розподіленою?

Узявши довільну функцію $f \in \mathfrak{L}_{[0;1]}$ (інтегровну за Лебегом на $[0; 1]$), розглянемо $L(f) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n f(\varepsilon_k)$. Якщо послідовність «гарна» — має властивості справжньої вибірки з рівномірного розподілу — то повинно бути $L(f) = \int_{[0;1]} f(x) \mu(dx)$, причому для $\forall f \in \mathfrak{L}_{[0;1]}$, при незмінній $\{\varepsilon_i\}$.

Легко бачити, що такої послідовності не існує. Достатньо узяти f , невідзначену в зліченній множині точок $\{\varepsilon_i\}$ та рівну 0 поза нею: така f інтегровна за Лебегом, але для неї $L(f)$ взагалі не визначене.

Виявляється, коли розглянути замість інтегровності за Лебегом інтегровність за Ріманом ($f \in \mathfrak{R}_{[0;1]}$), послідовності з такою властивістю існують:

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n f(\varepsilon_k) = \int_0^1 f(x) dx \quad (*)$$

ОЗН. 1. Послідовність $\varepsilon = \{\varepsilon_i\}_{i=1}^\infty$, $\varepsilon_i \in [0; 1]$, *рівномірно розподілена (РРП) в теоретико-числовому сенсі*, якщо для $\forall f \in \mathfrak{R}_{[0;1]}$ має місце (*).

ОЗН. 2. ε називається РР, якщо для $\forall$ відрізка $\Delta = [a; b] \subset [0; 1]$:

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n I_\Delta(\varepsilon_k) = |\Delta|, \text{ де } I_\Delta(x) = \begin{cases} 1, & x \in \Delta, \\ 0, & x \notin \Delta. \end{cases}$$

ОЗН. 3. (Критерій Вейля). ε називається РР, якщо

$$\forall k \in \mathbb{Z}, k \neq 0: \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{m=1}^n e^{2\pi i \cdot k \varepsilon_m} = 0$$

ТЕОР. Означення 1, 2, 3 еквівалентні (впливає із того, що ступінчасті і тригонометричні функції породжують системи, всюди щільні в $\mathfrak{R}_{[0;1]}$).

ТЕОР. РРП існують.

◀ Нехай $\beta \in \mathbb{R} \setminus \mathbb{Q}$, $\varepsilon_n := \{n\beta\}$. Узявши $\forall k \in \mathbb{Z}, k \neq 0$, маємо:

$$\begin{aligned} S_n(\varepsilon) &= \sum_{m=1}^n e^{2\pi i \cdot k \varepsilon_m} = \sum_{m=1}^n e^{2\pi i \cdot k \{m\beta\}} = \\ &= \sum_{m=1}^n e^{2\pi i \cdot k (m\beta - [m\beta])} = \sum_{m=1}^n e^{2\pi i \cdot k m \beta} \cdot 1 = e^{2\pi i \cdot k \beta} \cdot \frac{1 - e^{2\pi i \cdot n \beta}}{1 - e^{2\pi i \cdot \beta}} \\ |S_n(\varepsilon)| &= |e^{2\pi i \cdot k \beta}| \cdot \frac{|1 - e^{2\pi i \cdot n \beta}|}{|1 - e^{2\pi i \cdot \beta}|} \leq \frac{2}{|1 - e^{2\pi i \cdot \beta}|} \end{aligned}$$

а $e^{2\pi i \cdot k \beta} \neq 1$, інакше $2\pi i \cdot k \beta = 2\pi i \cdot l \Rightarrow \beta = \frac{l}{k} \in \mathbb{Q} - \otimes$. Значить, $|\frac{1}{n} S_n(\varepsilon)| \leq \frac{2}{|1 - e^{2\pi i \cdot \beta}|} \cdot \frac{1}{n} \xrightarrow{n \rightarrow \infty} 0 \Rightarrow \frac{1}{n} S_n(\varepsilon) \xrightarrow{n \rightarrow \infty} 0$. За критерієм Вейля, ε — РРП. ▶

Виявляється, що РРП достатньо «багато». Точніше,

ТЕОР. (Метрична теорема). Нехай $\{\alpha_i\}_{i=1}^\infty$ — послідовність різних цілих чисел. **Тоді** $\varepsilon_n = \{\alpha_n x\}$, $n \in \mathbb{N}$, є РРП майже для всіх $x \in [0; 1]$.

◀ Візьмемо ціле $k \neq 0$. Нехай $S(N; x) = \frac{1}{N} \sum_{n=1}^N e^{2\pi i \cdot k \varepsilon_n} =$

$$= \frac{1}{N} \sum_{n=1}^N e^{2\pi i \cdot k \{\alpha_n x\}} = \frac{1}{N} \sum_{n=1}^N e^{2\pi i \cdot k \alpha_n x} \cdot e^{-2\pi i \cdot k [\alpha_n x]} = e^{2\pi i \cdot k \alpha_n x}$$

хочемо довести, що $S(N; x) \xrightarrow{n \rightarrow \infty} 0$ майже скрізь на $x \in [0; 1]$.

$$S(N; x) \cdot \overline{S(N; x)} = \frac{1}{N^2} \sum_{n=1}^N e^{2\pi i \cdot k \alpha_n x} \sum_{m=1}^N e^{-2\pi i \cdot k \alpha_m x} = \frac{1}{N^2} \sum_{n,m=1}^N e^{2\pi i \cdot k (\alpha_n - \alpha_m) x}$$

Проінтегруємо цю рівність по x від 0 до 1: якщо $r \in \mathbb{Z}$ і $r \neq 0$, то $\int_0^1 e^{2\pi i \cdot r x} dx = \frac{1}{2\pi i r} (e^{2\pi i \cdot r} - 1) = 0$, тому з $\sum_{n,m=1}^N$ лишаються доданки з $n = m$.

Таких доданків — N , звідки $\int_0^1 |S(N; x)|^2 dx = \frac{1}{N^2} \cdot N = \frac{1}{N}$.

Якщо замість N узяти N^2 : $\int_0^1 |S(N^2; x)|^2 dx = \frac{1}{N^2}$. Значить,

$$\sum_{N=1}^{\infty} \int_0^1 |S(N^2; x)|^2 dx = \sum_{N=1}^{\infty} \frac{1}{N^2} = \frac{\pi^2}{6} < \infty$$

За лемою Фату $\int_0^1 \sum_{N=1}^{\infty} |S(N^2; x)|^2 dx < \infty$. Звідси, ряд під знаком $\int$ збігається майже скрізь на $[0; 1]$, тому майже для всіх $x \in [0; 1]$ загальний член цього ряду прямує до 0: $\lim_{N \rightarrow \infty} S(N^2; x) \rightarrow 0$.

Для $\forall N \in \mathbb{N} \exists m \in \mathbb{N}: m^2 \leq N < (m+1)^2$ (при цьому $m \leq \sqrt{N}$ і $N - m^2 < 2m + 1$). Тоді

$$|S(N; x)| \leq |S(m^2; x)| + \frac{1}{N} \left| \sum_{n=m^2+1}^N e^{2\pi i \cdot k \alpha_n x} \right| \leq$$

$$\leq |S(m^2; x)| + \frac{1}{N} (N - m^2) \leq |S(m^2; x)| + \frac{1}{N} \cdot 2m$$

значить, $|S(N; x)| \leq |S(m^2; x)| + \frac{2}{\sqrt{N}}$. При $N \rightarrow \infty: m^2 \rightarrow \infty \Rightarrow |S(m^2; x)| \rightarrow 0$ МС на $x \in [0; 1]$ та $\frac{2}{\sqrt{N}} \rightarrow 0$ — отже, $S(N; x) \xrightarrow[N \rightarrow \infty]{} 0$ МС на $x \in [0; 1]$.

Прибираючи з $[0; 1]$ усі точки (0-міри для кожного цілого $k \neq 0$), де такої збіжності немає, отримуємо, що вона має місце МС на $[0; 1]$. Критерій Вейля завершує доведення. ►

Аналогічно вводиться поняття РРП векторів $\{e_i\}$:

ОЗН. $\{e_i\}_{i=1}^{\infty} = \{(e_i^{(1)}; \dots; e_i^{(s)})\}_{i=1}^{\infty}$, де $e_i^{(j)} \in [0; 1]$, $j = \overline{1, s}$, називається РР в $D_s = [0; 1]^s$ — одиничному гіперкубі в $\mathbb{R}^s$, якщо для $\forall f \in \mathfrak{R}(D_s)$

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{m=1}^n f(e_m) = \int_{D_s} f(x) dx$$

ТЕОР. (Багатовимірний критерій Вейля). $\{e_i\}_{i \in \mathbb{N}} \in \text{РРП}$ в D_s тоді і тільки тоді, коли $\forall \bar{k} = (k_1; k_2; \dots; k_s) \in \mathbb{Z}^s$, $\bar{k} \neq \theta$ (усі k_i не дорівнюють 0 одночасно):

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{m=1}^n e^{2\pi i \cdot \langle \bar{k}; e_m \rangle} = 0, \text{ де } \langle \bar{k}; e_m \rangle = \sum_{j=1}^s k_j e_m^{(j)}$$

Цей критерій переконує в існуванні РРП у D_s . Достатньо взяти Н/З над полем $\mathbb{Q}$ з різних ірраціональних чисел $\beta_1, \beta_2, \dots, \beta_s$ ($\forall l_1, \dots, l_s \in \mathbb{Z}$, де $\bar{l} \neq \theta$: $l_1 \beta_1 + \dots + l_s \beta_s \neq 0$), тоді $e_n = (\{n \beta_1\}; \dots; \{n \beta_s\})$ буде РРП в D_s .

В основу цих означень покладено теоретико-числовий аналог В. 4 Н/З СР ВЧ та ВипВек в D_s . Якщо вимагати аналоги В. 2 і В. 3, матимемо

ОЗН. $\{\varepsilon_i\}$ називається s -РРП, якщо $e_i = (\varepsilon_i; \varepsilon_{i+1}; \dots; \varepsilon_{i+s-1}) \in \text{РРП}$ в D_s .

ОЗН. $\{\varepsilon_i\}$ називається s -Н/З РРП, якщо $e_{k;s} = (\varepsilon_{(k-1)s+1}; \varepsilon_{(k-1)s+2}; \dots; \varepsilon_{ks}) \in \text{РРП}$ в D_s .

Відомо, що s -РРП можна отримувати за допомогою функції (від k)

$$\rho_k^{(s)} = \{(k^s + P_{s-1}(k))\beta\}, \text{ де } \beta \in \mathbb{R} \setminus \mathbb{Q}, P_{s-1} \in \mathbb{P}_{s-1}[\mathbb{Z}] \text{ і } \deg P_{s-1} = s-1$$

Послідовності, s -РР для $\forall s \in \mathbb{N}$, називаються *цілком РР*. Таку властивість майже напевне має послідовність реалізацій СР ВЧ.

1.3. Генератори псевдовипадкових чисел

Псевдовипадковими (ПВ) називаємо послідовності $\{x_n\}$, отримані детермінованим алгоритмом (або *генератором*) виду $x_n = f(x_{n-1}; x_{n-2}; \dots; x_{n-k})$ (у найпростішому випадку $x_n = f(x_{n-1})$), які наближено мають ВВ. 1–5 реалізацій Н/З СР ВЧ.

З практичної точки зору легше досліджувати алгоритми для цілочисельних послідовностей, не враховуючи особливості округлення нецілих чисел у комп'ютерній арифметиці.

Один з перших ГПВЧ — метод *середин квадратів* фон Неймана:

$$x_n = \overline{a_{2k}^{(n)} a_{2k-1}^{(n)} \dots a_1^{(n)}} \rightarrow x_n^2 = \overline{b_{4k}^{(n)} b_{4k-1}^{(n)} \dots b_1^{(n)}}, \quad x_{n+1} = \overline{b_{3k}^{(n)} b_{3k-1}^{(n)} \dots b_{k+1}^{(n)}}$$

Це досить «бідне» джерело ПВЧ. На початку 1950-х рр. Дж. Форсайт, працюючи із 4-значними числами, перевірів його на 16 початкових значеннях — з них 12 привели до циклу $6100 \rightarrow 2100 \rightarrow 4100 \rightarrow 8100 \rightarrow 6100 \rightarrow \dots$, а ще 2 — до виродженої послідовності з нулів.

Н. К. Метрополіс дослідив його для 20-розрядних двійкових чисел і виявив, що існує всього 13 циклів, з найбільшою довжиною періода 142 (при тому, що загальна кількість 20-бітних чисел — $2^{20} = 1048576$).

1.3.1. Лінійний конгруентний метод

$$x_{n+1} = (ax_n + c) \bmod m$$

де $n \in \mathbb{Z}_+$, а цілі числа m — *модуль*, a — *множник*, c — *приріст*. x_0 , з якого починається послідовність, часто називають *зерном* (seed).

$\{x_n\}_{n=0}^{\infty}$ називається *лінійною конгруентною послідовністю* (ЛКП).

Нехай $m = 10$, $a = c = x_0 = 3$. Тоді $x_1 = (3 \cdot 3 + 3) \bmod 10 = 2$, $x_2 = 9$, $x_3 = 0$, $x_4 = 3$, ... — увійшли до короткого циклу. Очевидно, це «погана» імітація випадковості. Як обирати m , a , c та x_0 , аби вона була «хорошою»?

На комп'ютерах, що використовують бінарну арифметику, як правило вибирають $m = 2^\varepsilon$, де ε — кількість бітів у змінній цілочисельного типу. Одною з головних вимог до ЛКМ є найбільша можлива довжина періоду послідовності — m . Зокрема, маленькі m не годяться, наприклад, для $m = 2$ буде всього 3 типи послідовностей: $0, 0, \dots$ або $1, 1, \dots$ або $0, 1, 0, 1, \dots$.

ТЕОР. ЛКП, визначена параметрами m , a , c та x_0 , має період m (набільший з можливих) **тоді** **дот**

- 1) c та m взаємно прості (ВП): $\text{НСД}(c, m) = 1$;
- 2) $p|b = a - 1$ для кожного простого $p|m$;
- 3) якщо $4|m$, то $4|b$.

Аналогічний результат має місце для *мультиплікативних конгруентних послідовностей* (МКП) з $c = 0$. Введемо кілька означень.

ОЗН. Якщо a і m ВП, то $\min\{\lambda \in \mathbb{N} \mid a^\lambda \equiv 1 \pmod{m}\}$ називається *порядком a за модулем m* (такі λ існують: за теоремою Ейлера $a^{\varphi(m)} \equiv 1 \pmod{m}$).

Нехай $\lambda(m)$ — максимальний з порядків a , котрі ВП з m , за модулем m .

ОЗН. Якщо порядок a за модулем m дорівнює $\lambda(m)$, то a називається *первісним елементом за модулем m* .

ТВ. $\lambda(m)$ визначається так:

$$\lambda(2) = 1, \lambda(4) = 2, \lambda(2^k) = 2^{k-2} \text{ для } k \geq 3.$$

$$\lambda(p^k) = p^{k-1}(p-1), \text{ якщо просте } p > 2.$$

$$\lambda(p_1^{k_1} p_2^{k_2} \times \dots \times p_s^{k_s}) = \text{НСК}(\lambda(p_1^{k_1}); \lambda(p_2^{k_2}); \dots; \lambda(p_s^{k_s})), \text{ де } p_i \text{ — різні прості.}$$

ТЕОР. (Гаусс, 1801). При $c = 0$ найбільшим можливим періодом МКП є $\lambda(m)$, і він досягається, якщо:

- 1) m та x_0 — ВП;
- 2) a — первісний елемент за модулем m .

ТВ. Нехай $m = 2^\varepsilon$, $\varepsilon \geq 4$. a є первісним елементом за модулем m **тоді** і **тільки** тоді $a \equiv 3 \pmod{8}$ або $a \equiv 5 \pmod{8}$. У цьому випадку довжина періоду є $\frac{m}{4}$.

ЗАУВ. Найбільша довжина періоду не є достатньою вимогою; наприклад, такий період має ЛКП $x_{n+1} = (x_n + 1) \pmod{m}$.

Наведемо основні правила вибору параметрів ЛКМ:

- 1) m повинно бути великим, щонайменше 2^{30} .
- 2) Якщо $m = 2^\varepsilon$, то a вибирається так, щоб $a \equiv 5 \pmod{8}$. Якщо $m = 10^\delta$, то $a \equiv 21 \pmod{200}$.
- 3) $\frac{1}{100}m < a < \frac{99}{100}m$, причому двійкові чи десяткові цифри a не повинні мати регулярну структуру на кшталт 12345.
- 4) c та m повинні бути ВП; для $m = 2^\varepsilon$ це означає, що c непарне. При $c = 0$ найбільша довжина періоду зменшується в 4 рази.

5) Якщо $c \neq 0$, то x_0 можна вибирати довільним. Якщо ж $c = 0$ і $m = 2^\varepsilon$, то x_0 повинно бути непарним.

Навіть при виконанні цих вимог

1) Молодші значущі цифри x_n не є достатньо випадковими.

2) Після генерування приблизно $\frac{m}{1000}$ членів поведінка послідовності починає повторюватись. Приміром, для $m = 2^{32}$ необхідно змінювати a через кожні кілька мільйонів кроків.

Якщо процедура передбачає відтворюваність, треба також зберігати всі параметри, що визначають послідовність, включно з моментами їх зміни.

ПР. $x_0 = 5772156649$, $x_{n+1} = (3141592653x_n + 2718281829) \bmod 2^{35}$;
 $x_0 = 1781072418$, $x_{n+1} = (2718281829x_n + 3141592653) \bmod 2^{35}$.

1.3.2. Інші ГПВЧ

ЗАУВ. 1. Обчислення $f(x_n) = x_{n+1}$ для складної f займає багато часу.

ЗАУВ. 2. Чим складніше f , тим важче теоретично обґрунтувати якості відповідної послідовності $\{x_n\}$.

• *Квадратичний конгруентний метод*

$$x_{n+1} = (dx_n^2 + ax_n + c) \bmod m$$

Необхідні та достатні умови найбільшої довжини періоду m :

1) c та m — ВП;

2) $p|d$ і $p|(a-1)$ для всіх непарних простих $p|m$;

3) якщо $4|m$, то $2|d$ і $d \equiv (a-1) \pmod{4}$; якщо $2|m$, то $d \equiv (a-1) \pmod{2}$;

4) якщо $9|m$, то $d \not\equiv 3 \pmod{9}$.

Варіант Р. Ковею: $x_0 = 2 \pmod{4}$, $x_{n+1} = x_n(x_n + 1) \bmod 2^\varepsilon$.

• *Генератор Фібоначчі із запізненням*

$$x_{n+1} = (x_{n-k} + x_{n-l}) \bmod m$$

Один з перших таких запропонували Дж. Мітчел та Д. Мур у 1958 р.:

$$x_n = (x_{n-24} + x_{n-55}) \bmod m, \quad n \geq 55$$

де $2|m$, а $x_0, x_1, \dots, x_{54}$ — довільні, але не всі парні числа.

24 і 55 забезпечують довжину періоду молодшого двійкового розряду, $x_n \bmod 2$, рівну $2^{55} - 1$. При $m = 2^\varepsilon$ період всієї послідовності $\{x_n\}_{n \in \mathbb{N}}$ становить $2^{\varepsilon-1}(2^{55} - 1)$. Числа 24 і 55 в такому ГПВЧ називаються *запізненнями*.

Кілька пар з аналогічною властивістю найбільшої довжини періоду:

$$(30;127), (38;89), (37;100), (83;258), \dots, (7083;19937)$$

Такі генератори використовувались з кінця 1950-х, однак у 1990-ті з'ясувалось, що вони не проходять певний тест на випадковість, т.з. «критерій проміжків між днями народження».

- Розглянемо *генератор випадкових бітів* (0 та 1), запропонований та досліджений Р. Таусвортом.

Нехай $(c_1; c_2; \dots; c_n) \in \{0; 1\}^n$, причому $c_n = 1$ і поліном $f(x) = 1 + c_1x + c_2x^2 + \dots + c_nx^n$ є примітивним над полем $GF(2)$ залишків за модулем 2, тобто не розкладається у добуток поліномів меншого степеня над цим полем. Приміром, для $n = 35$ таким буде $(0; 0; \dots; 0; 1; 0; 1)_2$.

Послідовність бітів $\{\varepsilon_i\}_{i \in \mathbb{N}}$ визначається рекурентним співвідношенням

$$\varepsilon_k = c_1\varepsilon_{k-1} + c_2\varepsilon_{k-2} + \dots + c_n\varepsilon_{k-n} \pmod{2}$$

Примітивність $f(x)$ забезпечує максимальну довжину періоду, — $2^n - 1$.

- *Комбінування кількох ГПВЧ*: наприклад, маючи дві послідовності, побудовані різними генераторами, можна використати одну для перестановки елементів іншої; це т.з. *рандомізація перемішуванням*.

1.4. Емпіричні критерії рівномірності

Опишемо основні тести стандартів рівномірності даної послідовності чисел. Емпіричними вони називаються тому, що працюють не із ГПВЧ і його параметрами, а вже з виданою ним послідовністю.

Деякі з них призначені для перевірки послідовностей чисел з $[0; 1]$, деякі — послідовностей цілих чисел від 0 до $d - 1$. Перші (x_n) та другі (z_n) пов'язані співвідношеннями $x_n = \frac{z_n}{d}$, $z_n = [dx_n]$.

- *Критерій рівномірності/частот*. Перевіряємо, що $\{x_n\}$ рівномірно розподілена на $[0; 1]$, критерієм Колмогорова-Смірнова або, розділивши $[0; 1]$ на d інтервалів довжини $\frac{1}{d}$, застосуємо критерій χ^2 для перевірки, що x_n потрапляють в кожен з однаковою ймовірністю $p_i = \frac{1}{d}$.

- *Критерій серій*. Розглядаємо послідовність цілих $\{z_i\}_{i=1}^n$ від 0 до $d - 1$. Для «хорошої» (у сенсі випадковості) послідовності пари $(k; l) = (z_{2i-1}; z_{2i})$ з'являтимуться приблизно з однаковою частотою. Отже, маємо d^2 категорій, для кожної з яких теоретична ймовірність парі $(z_{2i-1}; z_{2i})$ потрапити в цю категорію — $\frac{1}{d^2}$. Порівнюючи ці ймовірності з емпіричними частотами критерієм χ^2 , робимо висновок щодо гіпотези про рівномірний розподіл.

Аналогічно можна зробити для трійок, ..., m -ок; при цьому треба слідкувати, щоб у кожній категорії було принаймні 5 m -ок з вибірки.

• *Критерій інтервалів.* Розглядаємо $\{x_k\}_{k=1}^n$ з $[0; 1]$. Для $0 \leq \alpha < \beta \leq 1$ порахуємо кількість інтервалів — членів послідовності перед тими x_k , котрі потрапили в $[\alpha; \beta]$. Отже, інтервал — підпослідовність $x_j, x_{j+1}, \dots, x_{j+r}$, де $x_{j+r} \in [\alpha; \beta]$, а всі інші $x_{j+i} \notin [\alpha; \beta]$. r називається довжиною інтервалу; узявши певне ціле t , полічимо інтервали довжин $0, 1, \dots, t-1$ та окремо — «завеликі», довжини не меншої t . Матимемо набір кількостей $\nu_0, \nu_1, \dots, \nu_{t-1}$ та ν_t . Відмітимо, що інтервали можуть накладатись.

Теоретична ймовірність інтервалу довжини r становить $p_r = p(1-p)^r$ при $r = \overline{0, t-1}$ та $p_t = (1-p)^t$ (сумарна для $r \geq t$), де $p = \beta - \alpha$. Близькість емпіричних частот $\frac{\nu_r}{n}$ до ймовірностей $\{p_r\}$ перевіряємо критерієм χ^2 .

• *Покер-критерій/критерій розбиттів.* Ділимо послідовність цілих $\{z_k\}$ на «п'ятірки» ($z_{5j}; z_{5j+1}; \dots; z_{5j+4}$) і рахуємо, скільки разів зустрілись наступні комбінації (порядок всередині групи несуттєвий):

- | | |
|--------------------------------|--------------------------------------|
| 1) abcde — усі числа різні; | 2) aabcd — одна пара; |
| 3) aabbс — дві пари; | 4) aaabc — три числа одного виду; |
| 5) aaabb — повний набір; | 6) aaaab — чотири числа одного виду; |
| 7) aaaaa — усі числа однакові. | |

Або можна рахувати групи з певною кількістю різних чисел — тоді категорій буде 5; у загальному випадку груп з k послідовних значень, теоретична ймовірність групи з r різними числами — $p_r = \frac{d(d-1) \times \dots \times (d-r+1)}{d^k} \cdot \left\{ \begin{matrix} k \\ r \end{matrix} \right\}$, де $\left\{ \begin{matrix} m \\ n \end{matrix} \right\} = S_n^m = \frac{1}{m!} \sum_{s=1}^m (-1)^{m-s} C_m^s s^n$ — числа Стірлінга 2-го роду.

Частоти $\frac{\nu_r}{n}$ порівнюємо з ймовірностями p_r критерієм χ^2 (тут ν_r — кількість груп з r різними числами).

• *Критерій збирання купонів.* Визначаємо у послідовності цілих $\{z_k\}$ відрізки $z_{j+1}, z_{j+2}, \dots, z_{j+r}$, що містять усі числа від 0 до $d-1$. Нехай r — довжина такого відрізка; узявши ціле $t > d$, знаходимо кількості відрізків довжини $d - \nu_d, d+1 - \nu_{d+1}, \dots, t-1 - \nu_{t-1}$, та окремо довжин, не менших $t - \nu_t$. Теоретична ймовірність, з якою відрізок має довжину r —

$$p_r = \frac{d!}{d^r} \cdot \left\{ \begin{matrix} r-1 \\ d-1 \end{matrix} \right\} \text{ для } d \leq r < t, \quad p_t = 1 - \frac{d!}{d^{t-1}} \cdot \left\{ \begin{matrix} t-1 \\ d \end{matrix} \right\}$$

Далі критерієм χ^2 порівнюємо частоти $\frac{\nu_r}{n}$ з ймовірностями p_r .

• *Критерій перестановок.* Розглядаємо $\{x_k\}$ з $[0; 1]$ та розбиваємо їх на послідовні групи по t чисел. Якщо ймовірність співпадіння чисел нульова, то значення в кожній групі можна упорядковувати за зростанням $t!$ різними способами, і теоретична ймовірність кожного з них та сама, що й у інших, звідки $p_r = \frac{1}{t!}$. Порахувавши групи, де числа упорядковуються r -м способом, $r = \overline{1, t!}$, матимемо частоти $\frac{\nu_r}{n}$, які порівнюємо з p_r критерієм χ^2 .

• *Критерій монотонності* перевіряє розподіл довжин монотонних серій в даній послідовності; частоти мають бути близькі до певних ймовірностей.

• *Критерій «максимум- t »*. Розглядаємо $\{x_k\}$ з $[0; 1]$, поділивши їх на групи по t чисел. Перевіряємо розподіл $y_j = \max\{x_{tj}; x_{tj+1}; \dots; x_{tj+t-1}\}$; якщо $\{x_k\}$ Н/З і СР, то $P\{y_j < x\} = \prod_{r=0}^{t-1} P\{x_{tj+r} < x\} = x^t$. Близькість до цього розподілу емпіричного перевіряємо критерієм Колмогорова-Смірнова.

• *Критерій конфліктів*. Недоліком деяких попередніх критеріїв є необхідність мати хоча б одну (а краще більше) реалізацію у кожній категорії.

Нехай є m категорій та n реалізацій, і $n \ll m$. Більшість реалізацій потрапить в категорії, де немає інших; якщо дана реалізація потрапила в категорію з іншими, скажемо, що відбувся «конфлікт». Ймовірність того, що буде рівно r конфліктів, $p_r = \frac{m(m-1) \times \dots \times (m-n+r+1)}{m^n} \cdot \left\{ \frac{n}{n-r} \right\}$, де $r = \overline{0, n}$.

Вважаємо, що ГПВЧ проходить даний тест, якщо конфліктів «не дуже мало» і «не дуже багато». Наприклад, коли $m = 2^{20}$ і $n = 2^{14}$, буде $P\{r \leq 101\} \approx 0.009$, $P\{r \leq 126\} \approx 0.476$, $P\{r \leq 153\} \approx 0.989$. Отже, якщо конфліктів було менше 101 або більше 153, тест не пройдено.

• *Критерій проміжків між днями народження*. Розглядаємо послідовність цілих $\{z_k\}_{k=1}^n$ від 0 до $d-1$, причому $d \gg n$. Назвемо числа від 0 до $d-1$ «днями року», а $\{z_k\}$ — «днями народження». Розташуємо $\{z_k\}$ за неспаданням — $z_{(1)} \leq z_{(2)} \leq \dots \leq z_{(n)}$, визначимо проміжки між ними $s_1 = z_{(2)} - z_{(1)}$, ..., $s_{n-1} = z_{(n)} - z_{(n-1)}$ та $s_n = z_{(1)} + d - z_{(n)}$, і упорядкуємо проміжки за неспаданням — $s_{(1)} \leq s_{(2)} \leq \dots \leq s_{(n)}$. Нехай R — кількість однакових проміжків, тобто таких j , що $s_j = s_{j-1}$. Для $d = 2^{35}$ та $n = 2^9 = 512$

x	0	1	2	≥ 3
$P_R(x)$	0.37	0.37	0.18	0.08

Критерієм χ^2 порівнюємо цей теоретичний розподіл R з емпіричним.

• *Критерій серійної кореляції*. Для $\{x_k\}_{k=0}^{n-1}$ з $[0; 1]$ обчислимо статистику

$$C = \frac{\left[n \sum_{k=0}^{n-1} x_k x_{(k+1) \bmod n} \right] - \left[\sum_{k=0}^{n-1} x_k \right]^2}{\left[n \sum_{k=0}^{n-1} x_k^2 \right] - \left[\sum_{k=0}^{n-1} x_k \right]^2}$$

Якщо $C \approx 0$, можна говорити про відсутність лінійної залежності; якщо $C \approx \pm 1$ — про сильну лінійну залежність. Оскільки $x_k x_{k+1}$ не є Н/З від $x_{k+1} x_{k+2}$ (спільний множник x_{k+1}), то від «хорошої» послідовності очікуємо, щоб C лежало в межах $(\mu_n - 2\sigma_n; \mu_n + 2\sigma_n)$, де $\mu_n = -\frac{1}{n-1}$, а $\sigma_n^2 = \frac{n^2}{(n-1)^2(n-2)}$ (для справжньої вибірки з СР розподілу ймовірність цього близька до 95 %).

• *Критерій підпослідовності.* Послідовні m -ки чисел, виданих ГПВЧ, теж повинні бути випадковими. Перевіримо це, застосувавши до перестановки

$$x_0, x_m, x_{2m}, \dots; x_1, x_{m+1}, x_{2m+1}, \dots; \dots; x_{m-1}, x_{2m-1}, x_{3m-1}, \dots$$

описані вище критерії.

Це приклад «метакритерія», коли вихідну послідовність перетворюємо у нову, для якої відповідний теоретичний розподіл не змінюється. Відтак перевіряємо нову послідовність тими ж критеріями, котрі брали для вихідної.

ЗАУВ. ГПВЧ, які проходять одні тести й не проходять інші:

- Метод середин квадратів задовольняє критерій частот, інтервалів і розбиттів, але провалює критерій серій;
- ЛКМ не проходять критерій монотонності (замало серій довжини 1);
- генератор Фібоначчі із запізненням не проходить однорозрядний критерій частот та критерій проміжків між днями народження.

• *Спектральний критерій* — один з найпотужніших (всі ГПВЧ, які тепер вважаються «поганими», не проходять його, а всі «хороші» — так).

Він базується на перевірці спільних розподілів s послідовних членів $\{x_k\}$ — $(x_k; x_{k+1}; \dots; x_{k+s-1})$ Множина A_s точок з такими координатами в D_s визначає так звані s -вимірні «точності» ν_s , що характеризують нерегулярність структури даної множини; чим вище «точність», тим «кращою» вважається досліджувана послідовність.

Нехай $\mathcal{S}$ — скінченний набір паралельних $(s-1)$ -вимірних гіперплощин, кожна з яких проходить хоча б через одну точку A_s , причому кожна точка A_s належить рівно одній гіперплощині з $\mathcal{S}$. Очевидно, таких наборів «багато». Для кожного з них позначимо через $\rho_s(\mathcal{S})$ найменшу відстань між суміжними гіперплощинами з $\mathcal{S}$. Тоді $\nu_s := 1/\max_{\mathcal{S}} \rho_s(\mathcal{S})$.

Цей критерій застосовується (як правило) при $s \leq 6$, бо із зростанням s час виконання відповідного алгоритму зростає експоненційно. Перевірку слід виконувати для всіх розмірностей s від 1 до найбільшої.

2. Моделювання заданого розподілу

Вважатимемо, що у нас є надійний ГПВЧ, здатний імітувати вибірку зі СР. Виявляється, на його основі можна отримувати, принаймні теоретично, ВВ з іншими розподілами. Підставою для цього є

ТЕОР. (Формула обернення, ФО). Нехай $F: \mathbb{R}^1 \rightarrow [0; 1]$ — неперервна ЙФР, а обернена до неї F^{-1} визначається як

$$F^{-1}(x) = \inf\{t \in \mathbb{R}^1 \mid F(t) = x\} \text{ для } x \in (0; 1)$$

І нехай α — СР ВЧ. Тоді ФР ВВ $\xi = F^{-1}(\alpha) \in F(x)$.

◀ (для зростаючої $F(x)$, тоді звичайна обернена F^{-1} теж зростаюча)

$$P\{F^{-1}(\alpha) < x\} = P\{F(F^{-1}(\alpha)) < F(x)\} = P\{\alpha < F(x)\} = F(x) \blacktriangleright$$

І навпаки, якщо ВВ ξ має ФР $F(x)$, то ВВ $\zeta = F(\xi)$ буде СР ВЧ, адже

$$P\{\zeta < x\} = P\{F(\xi) < x\} = P\{F(F^{-1}(\alpha)) < x\} = P\{\alpha < x\} = x$$

2.1. Абсолютно неперервні розподіли

Нехай α — СР ВЧ, а $\{x_i\}$ — реалізації α , отримані за допомогою ГПВЧ. Безпосереднє застосування ФО дозволяє моделювати такі розподіли:

- *Рівномірний розподіл* на $[a; b]$ моделюємо лінійним перетворенням α : $\xi = a + (b - a)\alpha$, а вибірку із цього розподілу отримуємо так: $y_i = a + (b - a)x_i$.

- *Показниковий розподіл* з параметром λ має ФР $F(x) = \begin{cases} 0, & x \leq 0, \\ 1 - e^{-\lambda x}, & x > 0, \end{cases}$ вона зростає при $x \geq 0$.

Розв'яжемо рівняння $F(x) = t$: $1 - e^{-\lambda x} = t \Leftrightarrow x = -\frac{1}{\lambda} \ln(1 - t)$. Отже, $F^{-1}(t) = -\frac{1}{\lambda} \ln(1 - t)$. Тоді ВВ $\xi = F^{-1}(\alpha) = -\frac{1}{\lambda} \ln(1 - \alpha)$ має показниковий розподіл з параметром λ . Для прискорення обчислень можна взяти $\xi = -\frac{1}{\lambda} \ln(\alpha)$, бо $(1 - \alpha)$ теж СР. Відповідно вибірка $y_i = -\frac{1}{\lambda} \ln x_i$.

- *Розподіл Парето* з ФР $F(x) = \begin{cases} 0, & x \leq b, \\ 1 - (\frac{b}{x})^a, & x > b. \end{cases}$ $F(x) = t \Leftrightarrow x = \frac{b}{\sqrt[a]{1-t}}$; беремо $\xi = \frac{b}{\sqrt[a]{1-\alpha}}$ чи простіше $\xi = \frac{b}{\sqrt[a]{\alpha}}$.

- Аналогічно для *розподілу Вейбулла* з ФР $F(x) = \begin{cases} 0, & x \leq 0, \\ 1 - e^{-(\frac{x}{\theta})^p}, & x > 0. \end{cases}$

Але такий спосіб не годиться для *нормального розподілу* — ФР $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{u^2}{2}} du$ і обернена $\Phi^{-1}(x)$ не виражаються через елементарні функції.

• Метод полярних координат (Бокс, Мюллер, Марсалья, 1958, 1962). Нехай ξ та η — Н/З $N_{0;1}$ -розподілені ВВ, кожна зі щільністю розподілу $p(t) = \frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}$. Їх спільна щільність (тобто щільність $(\xi; \eta)$) є $p(x; y) = \frac{1}{2\pi}e^{-\frac{x^2+y^2}{2}}$.

Перейдемо до полярних координат $(r; \varphi)$, де $x = r \cos \varphi$, $y = r \sin \varphi$; відповідно $\xi = \rho \cos \psi$, $\eta = \rho \sin \psi$. Спільною щільністю розподілу $(\rho; \psi)$ є $p(r; \varphi) = \frac{1}{2\pi}re^{-\frac{r^2}{2}}$ (r — якобіан переходу від $(r; \varphi)$ до $(x; y)$). У $p(r; \varphi)$ немає φ , а це означає, що ρ й ψ Н/З, ψ розподілена рівномірно на $[0; 2\pi]$, у той час як ρ має щільність $f(r) = re^{-\frac{r^2}{2}}$, її ФР $F(r) = \int_0^r f(t)dt = -e^{-\frac{t^2}{2}} \Big|_0^r = 1 - e^{-\frac{r^2}{2}}$ виражається через елементарні функції, і $F^{-1}(t)$ теж:

$$1 - e^{-\frac{r^2}{2}} = t \Leftrightarrow -\frac{r^2}{2} = \ln(1 - t) \Leftrightarrow r = \sqrt{-2 \ln(1 - t)}$$

так що можна моделювати $\rho = \sqrt{-2 \ln(1 - \alpha)}$, а швидше $\rho = \sqrt{-2 \ln \alpha}$.

ψ моделюється як $\psi = 2\pi\beta$, де β — СР ВЧ, причому α та β Н/З. Тоді

$$\xi = \rho \cos \psi = \sqrt{-2 \ln \alpha} \cdot \cos(2\pi\beta), \quad \eta = \rho \sin \psi = \sqrt{-2 \ln \alpha} \cdot \sin(2\pi\beta)$$

Для послідовностей $\{x_i\}$, виданих ГПВЧ, розбиваємо $x_1, x_2, \dots$ на пари $(x_{2k-1}; x_{2k})$ та обчислюємо

$$y_{2k-1} = \sqrt{-2 \ln x_{2k-1}} \cdot \cos(2\pi x_{2k}), \quad y_{2k} = \sqrt{-2 \ln x_{2k-1}} \cdot \sin(2\pi x_{2k})$$

Коли потрібний $N_{a; \sigma^2}$, виконуємо лінійне перетворення $\zeta = \sigma\xi + a$, де $\xi \sim N_{0;1}$ генерується описаним методом.

Для моделювання інших АбсНеп розподілів, як-от гамма- і бета-, використовуються інші методи.

2.2. Дискретні розподіли

Почнемо із загального скінченного випадку, $\xi \sim \begin{pmatrix} x_1 & \dots & x_n \\ p_1 & \dots & p_n \end{pmatrix}$. Вважатимемо всі $p_i > 0$. Цей розподіл розбиває $[0; 1]$ на проміжки довжини p_i .

Застосуємо «дискретизацію»: реалізуємо СР α та визначимо, у який проміжок воно потрапило. Розбиваємо $[0; 1)$ на $\Delta_k = [a_{k-1}; a_k)$, де $a_0 = 0$, $a_k = \sum_{i=1}^k p_i$ для $k = \overline{1, n}$, і якщо α потрапило в Δ_k , — покладаємо $\xi = x_k$.

Найпростіший спосіб визначення номера такого проміжка — послідовно перебирати $\Delta_1, \dots, \Delta_n$, порівнюючи α з кінцями поточного Δ_i , але ефективніше виконувати пошук «діленням навпіл», особливо коли n велике.

Нехай координати a_i обчислені заздалегідь, і ми отримуємо із ГПВЧ чергове α . Покладаємо спочатку номери кінців $S := 0$ та $E := n$. Вони визна-

чають інтервал, у якому знаходиться α : $a_S \leq \alpha < a_E$. Далі запускаємо цикл, котрий виконується, допоки не стане $E = S + 1$, адже тоді α перебуватиме у $[a_S; a_{S+1})$ — проміжок визначено. Якщо $E > S + 1$, ділимо навпіл, тобто беремо серединний (не за координатою, а за номером) кінець $M = \lfloor \frac{S+E}{2} \rfloor$ і дивимось, куди потрапляє α — в $[a_S; a_M)$ чи в $[a_M; a_E)$; в залежності від цього покладаємо $E := M$ чи $S := M$ та переходимо до наступної ітерації циклу.

Кожного разу довжина діапазону — кількість $(E-S)$ проміжків між a_S та a_E — зменшується приблизно удвічі (і хоча б на 1), так що процес скінченний, більш того — кількість ітерацій до завершення циклу не перевищує $C \log_2 n$, а це суттєво швидше, ніж Cn ітерацій при послідовному переборі.

У частинному випадку, коли всі $p_i = \frac{1}{n}$ («дискретний рівномірний розподіл»), номером проміжка, куди потрапляє α , є $1 + \lfloor n\alpha \rfloor$.

- *Біномний розподіл*, $B_{n;p}(k) = C_n^k p^k (1-p)^{n-k}$, $k = \overline{0, n}$. По-перше, спрацює описаний тільки що загальний метод, якщо не буде великих похибок округлення a_i . По-друге, оскільки $\mu \sim B_{n;p} \in \sum_{i=1}^n \mu_i$, де $\mu_i \sim \begin{pmatrix} 0 & 1 \\ 1-p & p \end{pmatrix}$, то для невеликих n можна згенерувати n значень СР α і порахувати, скільки з них потрапить у $[0; p]$. По-третє, для великих n , можна використовувати складніший метод, що спирається на β -розподіл.

- *Геометричний розподіл*, $P_p(k) = p(1-p)^k$, $k \in \mathbb{Z}_+$.

1) Безпосередньо за означенням: генеруємо СР α , поки не відбудеться подія-успіх « $\alpha < p$ », і присвоюємо ξ кількість генерацій перед успіхом.

Але тоді, запускаючи генератор для отримання чергового ξ , ми не знаємо, як довго він працюватиме, скільки генерацій відбудеться до першого успіху. В середньому доведеться чекати $\frac{1}{p} - 1$ генерацій.

2) Покладаємо $\xi = \lfloor \frac{\ln \alpha}{\ln(1-p)} \rfloor$. Тоді для $\forall n \in \mathbb{Z}_+$: $P\{\xi = n\} =$
 $= P\{n \leq \frac{\ln \alpha}{\ln(1-p)} < n+1\} = P\{n \ln(1-p) \geq \ln \alpha > (n+1) \ln(1-p)\} =$
 $= P\{(1-p)^n \geq \alpha > (1-p)^{n+1}\} = (1-p)^n - (1-p)^{n+1} = p(1-p)^n$

- *Пуассонівський розподіл*, $P_\lambda(k) = \frac{\lambda^k}{k!} e^{-\lambda}$, $k \in \mathbb{Z}_+$. Для невеликих λ застосуємо наступний метод. Генеруємо послідовність $\{\xi_i\}$ Н/З ВВ, розподілених показниково з параметром λ , та рахуємо їх разом з сумою $\Xi_m = \xi_1 + \xi_2 + \dots + \xi_m$; зупиняємось, коли $\Xi_m \geq 1$, та покладаємо $\zeta := m - 1$.

Перевіримо, чи має ζ розподіл Пуассона. $P\{\zeta = n\} =$
 $= P\{\Xi_{n+1} \geq 1, \Xi_n < 1\} = P\{\Xi_{n+1} \geq 1 \setminus \Xi_n \geq 1\} = P\{\Xi_{n+1} \geq 1\} - P\{\Xi_n \geq 1\}$
адже $\Xi_n \geq 1 \Rightarrow \Xi_{n+1} \geq 1$. Відомо, що згортка m показникових розподілів з параметром λ є розподілом Ерланга з параметрами $(m; \lambda)$, щільність якого

$$p(x) = \begin{cases} 0, & x \leq 0, \\ \frac{\lambda^m}{(m-1)!} x^{m-1} e^{-\lambda x}, & x > 0. \end{cases}$$

$$\text{Звідси } P\{\Xi_m \geq 1\} = \int_1^{+\infty} p(x) dx = \int_1^{+\infty} \frac{\lambda^{m-1}}{(m-1)!} x^{m-1} e^{-\lambda x} \lambda dx = \int_{\lambda}^{+\infty} \frac{t^{m-1}}{(m-1)!} e^{-t} dt.$$

$$\text{Тому } P\{\zeta = n\} = \int_{\lambda}^{+\infty} \frac{t^n}{n!} e^{-t} dt - \int_{\lambda}^{+\infty} \frac{t^{n-1}}{(n-1)!} e^{-t} dt = \frac{1}{n!} (f_n(\lambda) - n f_{n-1}(\lambda)), \text{ де}$$

$$f_n(\lambda) = \int_{\lambda}^{+\infty} t^n e^{-t} dt. \text{ Так як } f_n(\lambda) = \left\{ \begin{array}{l} u = x^n, du = n x^{n-1} dx, \\ dv = e^{-x} dx, v = -e^{-x} \end{array} \right\} = -x^n e^{-x} \Big|_{\lambda}^{+\infty} +$$

$$n \int_{\lambda}^{+\infty} x^{n-1} e^{-x} dx = \lambda^n e^{-\lambda} + n f_{n-1}(\lambda), \text{ то } P\{\zeta = n\} = \frac{1}{n!} \lambda^n e^{-\lambda} = P_{\lambda}(n).$$

2.3. Метод відбору та бітові алгоритми

• *Метод відбору.* Нехай (гіпер)тіло $A \subseteq D_s = [0; 1]^s$, міра Лебега $\mu(A) > 0$ і необхідно моделювати ВипВек $\bar{e}_i$, рівномірно розподілений в A .

Для цього генеруємо вибірку $\bar{e}_1, \bar{e}_2$ з рівномірного в D_s розподілу, поки не отримаємо точку $\bar{e}_j \in A$; вона стає черговим членом послідовності $\{\bar{e}_i\}$.

Для $\mu(A) > 0$ кількість генерацій $\bar{e}_j$ до першого влучення в A має геометричний розподіл. У загальному випадку, коли $\bar{e}_j$ розподілені рівномірно в $B \subset \mathbb{R}^s$, величина $\frac{\mu(A)}{\mu(B)}$ називається *ефективністю відбору*. Чим вона менше, тим довше у середньому триватиме очікування $\bar{e}_j$, яке потрапить в A .

• *Бітові алгоритми,* застосовувані для моделювання дискретних розподілів, досліджувались, зокрема, Кнутом і Яо. Нехай необхідно моделювати ВВ $\xi \sim \begin{pmatrix} x_1 & \dots & x_n \\ p_1 & \dots & p_n \end{pmatrix}$. Оскільки точність p_i обмежена, то це числа $\frac{m_i}{k_i} \in \mathbb{Q}$ із спільним знаменником $K = \text{НСК}\{k_1; \dots; k_n\}$. Домноживши на нього p_i , матимемо набір $l_i = m_i \frac{K}{k_i}$ цілих чисел від 1 до K , які в сумі дають K . Розіб'ємо діапазон 1, 2, ..., K на послідовні групи по $l_1, l_2, \dots, l_n$ чисел, поставимо у відповідність числам i -ї групи значення x_i , та доповнимо K до найближчого зверху степеня 2, 2^r , вважаючи «нові» $2^r - K$ значень «дірявими».

Наприклад, для $p_1 = \frac{2}{3}, p_2 = \frac{1}{4}, p_3 = \frac{1}{12}$ буде $K = 12$ і $l_1 = 8, l_2 = 3, l_3 = 1$ та 4 «діряві» (13–16), а $r = 4$.

Далі випадкові біти генеруються блоками по r і для кожного блоку визначається група i , куди потрапило число, утворене його бітами; якщо ця група не «дірява», то до вихідної послідовності додається відповідне x_i .

ТЕОР. Якщо $p_1, p_2, \dots, p_n$ — дискретний розподіл, то середня кількість бітів, необхідна для такого моделювання, не перевищує $\log_2(n-1) + 2$ (тобто залежить лише від n , а не від значень $\{p_i\}$).

2.4. Перетворення щільності

Описуючи моделювання нормального розподілу, ми скористались заміною змінних при переході від декартових координат до полярних. У загальному випадку обґрунтуванням такого переходу є

ТЕОР. (Правило перетворення щільності при перетворенні координат.) Нехай області $B, B' \subseteq \mathbb{R}^n$, а $g = (g_1; g_2; \dots; g_n): B \leftrightarrow B'$ — бієктивне диференційоване відображення B на B' . Якщо ВипВек $\xi = (\xi_1; \xi_2; \dots; \xi_n)$, розподілений в B , має щільність $p_\xi(x_1; \dots; x_n)$, то ВипВек $\eta = g(\xi)$ (з координатами $\eta_i = g_i(\xi_1; \dots; \xi_n)$), розподілений в B' , має щільність

$$p_\eta(y_1; \dots; y_n) = p_\xi(x_1; \dots; x_n) \cdot \left| \frac{\partial(x_1; \dots; x_n)}{\partial(y_1; \dots; y_n)} \right|$$

де у правій частині x_i повинні бути виражені через y_j , $x_i = x_i(y_1; \dots; y_n)$, а

$$\left| \frac{\partial(x_1; \dots; x_n)}{\partial(y_1; \dots; y_n)} \right| — \text{якобіан переходу від СК } \{y_j\} \text{ до СК } \{x_i\}, \text{ тобто } \det \begin{pmatrix} \frac{\partial x_1}{\partial y_1} & \dots & \frac{\partial x_1}{\partial y_n} \\ \dots & \dots & \dots \\ \frac{\partial x_n}{\partial y_1} & \dots & \frac{\partial x_n}{\partial y_n} \end{pmatrix}.$$

◀ Нехай D' — довільна область у B' , $D' \subseteq B'$, а $D = g^{-1}(D')$. Тоді $P\{\xi \in D\} = P\{\eta \in D'\}$, і за правилом заміни змінної $\int_{D'} p_\eta(y) dy =$

$$= P\{\eta \in D'\} = P\{\xi \in D\} = \int_D p_\xi(x) \underbrace{dx_1 dx_2 \dots dx_n}_{dx} = \int_{D'} p_\xi(x) \left| \frac{\partial(x_1; \dots; x_n)}{\partial(y_1; \dots; y_n)} \right| dy$$

$$\text{Із } \forall D': \int_{D'} (p_\eta(y) - p_\xi(x) | \dots |) dy = 0 \text{ випливає } p_\eta(y) = p_\xi(x) | \dots |. \blacktriangleright$$

ПР. 1. Моделювання точки, рівномірно розподіленої в одиничній кулі.

Точки одиничної кулі $B_3 \subset \mathbb{R}^3$ визначаються умовою $x^2 + y^2 + z^2 \leq 1$. Нехай ВВ ξ, η, ζ — декартові координати точки Ξ , рівномірно розподіленої в B_3 . Їх спільна щільність (щільність Ξ) є $p_\Xi(x; y; z) = \frac{1}{V_1} = \frac{3}{4\pi}$. Але окремі щільності розподілів ξ, η, ζ складніші, тому перейдемо до сферичних координат:

$$x = r \sin \theta \cos \varphi, y = r \sin \theta \sin \varphi, z = r \cos \theta$$

якобіан переходу від $(r; \theta; \varphi)$ до $(x; y; z) — r^2 \sin \theta$.

У сферичних координатах куля стає паралелепіпедом: $r \in [0; 1]$, $\theta \in [0; \pi]$, $\varphi \in [0; 2\pi]$, а щільність $p_\Xi(r; \theta; \varphi) = \frac{3}{4\pi} \cdot r^2 \sin \theta$ є добутком трьох щільностей: $p_r = 3r^2$, $p_\theta = \frac{1}{2} \sin \theta$ і $p_\varphi = \frac{1}{2\pi}$, так що сферичні координати $\Xi — \text{Н/З}$.

Оскільки ФР $F_r(t) = \int_0^t 3r^2 dr = t^3$, $F_\theta(t) = \int_0^t \frac{1}{2} \sin \theta d\theta = \frac{1 - \cos t}{2}$, $F_\varphi(t) =$

$$\int_0^t \frac{1}{2\pi} d\varphi = \frac{t}{2\pi}, \text{ то, використавши ФО, матимемо для моделювання}$$

$$r = \sqrt[3]{\alpha_1}, \cos \theta = 1 - 2\alpha_2, \varphi = 2\pi\alpha_3, \text{ де } \alpha_1, \alpha_2, \alpha_3 — \text{Н/З СР ВЧ}$$

Якщо радіус кулі — R , домножуємо r на R . Знаючи сферичні координати точки, обчислюємо й декартові: $\xi = r \sin \theta \cos \varphi$, $\eta = r \sin \theta \sin \varphi$, $\zeta = r \cos \theta$.

Інший спосіб отримувати послідовність точок, рівномірно розподілених в одиничній кулі B_n простору $\mathbb{R}^n$, є різновидом метода відбору. Моделюємо $\xi_1 = 1 - 2\alpha_1, \dots, \xi_n = 1 - 2\alpha_n$ — Н/З рівномірно розподілені на $[-1; 1]$ — та перевіряємо умову $d^2 = \sum_{i=1}^n \xi_i^2 \leq 1$. Якщо вона виконується — додаємо $(\xi_1; \dots; \xi_n)$ до послідовності, а якщо ні — повторюємо процедуру знову, поки не з'явиться $\Xi \in B_n$. Відмітимо, що ефективність відбору, $\frac{V(B_n)}{V([-1; 1]^n)} \xrightarrow{n \rightarrow \infty} 0$.

ПР. 2. Вибір випадкового напрямку. Багатовимірний ізотропний вектор.

ОЗН. ВипВек $\xi = (\xi_1; \dots; \xi_n) \in \mathbb{R}^n$ називається *ізотропним*, якщо точка $\frac{\xi}{\|\xi\|}$ розподілена рівномірно на сфері $S_n = \{x \in \mathbb{R}^n : \|x\| = 1\}$.

При $n = 3$ вміємо моделювати точку, рівномірно розподілену в B_3 . Її радіус-вектор ізотропний, адже положення Ξ на S_3 залежить тільки від координат θ й φ , а не від r , так що достатньо моделювати їх — $\cos \theta = 1 - 2\alpha_1$, $\varphi = 2\pi\alpha_2$, і узявши $r = 1$, матимемо ізотропний ВипВек $(\xi; \eta; \zeta)$, де

$$\xi = \sqrt{1 - \cos^2 \theta} \cos \varphi, \eta = \sqrt{1 - \cos^2 \theta} \sin \varphi, \zeta = \cos \theta$$

Аналогічно в $\mathbb{R}^n$ можемо генерувати $(\xi_1; \dots; \xi_n)$, рівномірно розподілений в B_n . Нормуємо його до одиничної довжини, $\Xi = (\frac{\xi_1}{d}; \dots; \frac{\xi_n}{d}) \in S_n$ — ізотропний.

ЛЕМА 1. Якщо η — ізотропний вектор, такий, що $|\eta| = \chi_n$ (де χ_n^2 має χ^2 -розподіл з n ступенями вільності), то $\eta_1, \dots, \eta_n$ Н/З і $\eta_i \sim N_{0;1}$.

◀ Розподіл χ_n^2 має щільність $f_n(x) = c_n x^{\frac{n}{2}-1} e^{-\frac{x}{2}}$ для $x > 0$ (γ -розподіл з параметрами $(\frac{n}{2}; \frac{1}{2})$). Тоді щільність розподілу ВВ $\chi_n = \sqrt{\chi_n^2}$

$$p_{\chi_n}(x) = \frac{d}{dx} P\{\chi_n < x\} = \frac{d}{dx} P\{\chi_n^2 < x^2\} = p_{\chi_n^2}(x^2) \cdot 2x = 2c_n x^{n-1} e^{-\frac{x^2}{2}}$$

У сферичній системі координат η має представлення $(\chi_n; \lambda_n)$, де $\lambda_n \in S_n$ — точка одиничної сфери. Внаслідок ізотропності спільна щільність $h_n(r; s)$ розподілу χ_n та λ_n не залежить від $s \in S_n$ та має вигляд $h_n(r; s) = c_1 r^{n-1} e^{-\frac{r^2}{2}}$.

Якобіан переходу від $(x_1; \dots; x_n)$ до $(r; s)$ дорівнює $r^{-(n-1)}$, бо в декартовій СК нескінченно малий елемент (гіпер)об'єму $dV = r^{n-1} dS dr$, адже $dV = dx_1 dx_2 \dots dx_n$, а в елемента об'єму на відстані r від центра сфери усі виміри (кількістю $n-1$), окрім dr , в r разів менше, ніж у поверхні сфери. За правилом перетворення спільна щільність $\eta_1, \dots, \eta_n$ у декартовій СК

$$p(x_1; \dots; x_n) = c_1 r^{n-1} \cdot e^{-\frac{r^2}{2}} \cdot r^{-(n-1)} = c_1 e^{-\frac{1}{2} \sum_{i=1}^n x_i^2} = c_1 \prod_{i=1}^n e^{-\frac{x_i^2}{2}}$$

звідки $\eta_1, \dots, \eta_n$ — Н/З, і їх щільності $p_i(x_i) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x_i^2}{2}}$, тобто $\eta_i \sim N_{0;1}$. ▶

ЛЕМА 2. Якщо $\eta_1, \dots, \eta_n$ — Н/З ВВ, $\eta_i \sim N_{0,1}$, $i = \overline{1, n}$, то ВипВек $\eta = (\eta_1; \dots; \eta_n)$ є ізотропним.
Доведення аналогічне.

ТВ. Нехай $\xi_1, \xi_2, \dots, \xi_n$ та ξ — Н/З СР ВЧ. Тоді ВВ $\sqrt[n]{\xi}$ і $\max\{\xi_1; \dots; \xi_n\}$ мають однаковий розподіл з ФР $G(x) = x^n$ при $x \in [0; 1]$.

$$\blacktriangleleft F_{\sqrt[n]{\xi}}(x) = P\{\sqrt[n]{\xi} < x\} = P\{\xi < x^n\} = x^n, F_{\max\{\xi_1; \dots; \xi_n\}}(x) = \\ = P\{\max\{\xi_1; \dots; \xi_n\} < x\} = P\{\xi_1 < x, \dots, \xi_n < x\} = \prod_{i=1}^n P\{\xi_i < x\} = x^n \blacktriangleright$$

2.5. Вибірki та перестановки

- Нехай маємо файл з N записів, причому N настільки велике, що цей файл не поміщається в оперативну/швидкодіючу пам'ять. Треба утворити випадкову вибірку об'ємом n з цих записів. Моделлю даного СЕ є класична модель з $\Omega = \{ \{x_1; x_2; \dots; x_n\} \mid x_i = \overline{1, N}, i = \overline{1, n}; x_1 < x_2 < \dots < x_n \}$, $n(\Omega) = C_N^n$. Записи, у свою чергу, можуть теж бути великими, так що навіть n -елементна вибірка не вміщається у пам'ять. Можна тільки записувати вибрані записи (або їх номери) в інший файл.

Отже, програма має послідовно перебирати N записів (чи їх номерів) та для кожного визначати — додавати цей запис до вибірки чи ні. При цьому:

- 1) Наприкінці вихідний файл повинен містити рівно n записів;
- 2) кожний запис вхідного файлу повинен включатися до вибірки з однаковою ймовірністю $1 - \frac{C_{N-1}^n}{C_N^n} = 1 - \frac{N-n}{N} = \frac{n}{N}$.

Аби поєднати ці дві вимоги, діятимемо так (Фан, Мюллер, Резуха, 1962; Джонс, 1962): якщо вже розглянуто t записів, то наступний, $t+1$, включається до вибірки з ймовірністю $p = \frac{n-m}{N-t}$, де m — поточний об'єм вибірки.

Перша вимога виконуватиметься, бо як тільки $m = n$, p стає 0 і процес закінчиться, тобто не буде перебору; з іншого боку, не буде й недобору (у вхідному файлі залишилось менше записів, ніж треба ще додати до вибірки).

Щодо другої вимоги, то ймовірність $\frac{n-m}{N-t}$ можна вивести з неї наступними міркуваннями. Якщо будь-який запис вхідного файлу Н/З від інших вибирається з ймовірністю $\frac{n}{N}$, то за умови, що серед перших t записів вибрано рівно m (подія B), ймовірність того, що й $(t+1)$ -а буде вибрана (подія A) — це частка тих способів вибору n з N , за яких у перших t вибрано m , для котрих вибирається і $(t+1)$ -а, тобто $n(B) = C_t^m \cdot C_{N-t}^{n-m}$, $n(A) = C_t^m \cdot 1 \cdot C_{N-t-1}^{n-m-1}$ і $A \subset B$, звідки $P(A/B) = \frac{P(A)}{P(B)} = \frac{n(A)}{n(B)} = \frac{(N-t-1)!(n-m)!(N-t-n+m)!}{(n-m-1)!(N-t-n+m)!(N-t)!} = \frac{n-m}{N-t}$.

- Для моделювання перестановок n -елементної множини $\{a_1; a_2; \dots; a_n\}$ можна використовувати принаймні два підходи.

Перший полягає в індексації таких перестановок числами від 1 до $P_n = n!$ і випадковому виборі одного з цих чисел з ймовірністю $\frac{1}{n!}$. Цей підхід стає проблематичним, коли $n!$ настільки велике, що точності представлення чисел не вистачає для розрізнення сусідніх проміжків на $[0; 1]$.

Другий підхід (Фішер, Ятс, 1938; Дурстенфелд, 1964) будує перестановку, починаючи з $(a_1; \dots; a_n)$. Вводимо лічильник $m = n$ і щоразу генеруємо СР α , після чого міняємо місцями $a_m \leftrightarrow a_{1+\lfloor m\alpha \rfloor}$ та зменшуємо m на 1, поки не дійдемо до $m = 1$.

2.6. Випадкові процеси

Нагадаємо основні означення: *випадковим процесом* (ВП) на ЙП $\{\Omega; \mathfrak{F}; P\}$ називається довільний клас ВВ $\xi(t; \omega)$, де $t \in \mathbb{R}^1$ тлумачимо як «час». Аргумент ω у позначенні часто опускається.

Якщо у цьому класі $t \in A \subset \mathbb{R}^1$, де A — дискретна множина, то процес називається *дискретним*. Спільні розподіли ВВ $\xi(t_1; \omega)$, $\xi(t_2; \omega)$, ..., $\xi(t_n; \omega)$ при всіх $n \in \mathbb{N}$ та наборах $t_1, t_2, \dots, t_n$ називаються *скінченновимірними розподілами ймовірності* (СВР) ВП.

Наприклад, $\xi(t; \omega) = \omega t$, де $\{\Omega; \mathfrak{F}; P\}$ — СР розподіл, $t \in \mathbb{R}^1$.

При кожному $\omega_0 \in \Omega$ функція $\xi(t; \omega_0)$, де $t \in H \subseteq \mathbb{R}^1$ — *траєкторія* ВП.

ТЕОР. (Колмогорова) у певному сенсі «обертає» це означення: клас ФР $\{F_n(x_1; t_1; \dots; x_n; t_n)\}$, що є $\{P\{\xi(t_1; \omega) < x_1; \dots; \xi(t_n; \omega) < x_n\}\}$, визначає ВП з такими СВР **тоді**дот мають місце *умови узгодження*

$$\lim_{x_n \rightarrow +\infty} F_n(x_1; t_1; \dots; x_n; t_n) = F_{n-1}(x_1; t_1; \dots; x_{n-1}; t_{n-1})$$

Звідси випливає, що (принаймні, теоретично) задача моделювання ВП зводиться до задачі моделювання певного класу розподілів $F_n(x_1; t_1; \dots; x_n; t_n)$.

Фіксуємо t , отримуємо $M\xi(t; \omega)$ — МС/середнє значення ВП.

ВП називається *стаціонарним*, якщо його СВР не залежать від зсуву t_i ; *гауссівським*, якщо всі його СВР — нормальні.

Одним з найпростіших для моделювання є *марківський* ВП з дискретним часом (без зменшення загальності $H = \mathbb{Z}_+$) та скінченною кількістю значень-станів $x_1, \dots, x_n$ (або просто $1, 2, \dots, n$). Такий процес, як відомо, визначається *початковим розподілом* $\begin{pmatrix} 1 & \dots & n \\ p_1 & \dots & p_n \end{pmatrix}$ та *матрицею переходу* $P = \|p_{ij}\|_{1 \leq i, j \leq n}$, де $p_{ij} = P\{\xi(t+1) = j / \xi(t) = i\}$ — ймовірність переходу із стану i у стан j , що залежить тільки від i та j , де $\sum_{j=1}^n p_{ij} = 1$.

У більш загальному випадку, для кожного стану i можливий також *обрив траєкторії* — з ймовірністю g_i процес закінчується / переходить з i у «поглинаючий» стан (0) . Можна вважати $g_i = p_{i0}$, $p_{00} = 1$, а $p_{0j} = 0$ для всіх $j = \overline{1, n}$. Тоді $\sum_{j=1}^n p_{ij} = 1 - g_i$.

Моделювання такого ВП починається з моделювання початкового стану s_0 згідно розподілу $\{p_i\}$. Після цього, на кожному кроці-моменті m , якщо перебуваємо у стані i ($s_m = i$), моделюємо дискретну ВипВел з розподілом $\begin{pmatrix} 0 & 1 & 2 & \dots & n \\ p_{i0} & p_{i1} & p_{i2} & \dots & p_{in} \end{pmatrix}$; її реалізація буде наступним станом s_{m+1} .

Приріст ВП $\xi(t)$ на проміжку часу $[s; t]$, $\xi(t) - \xi(s)$, позначатимемо $\xi[s; t]$.

ОЗН. ВП $\xi(t)$ має *незалежні прирости*, якщо $\forall \{t_i\}_{i=0}^n$, $0 \leq t_0 < t_1 < \dots < t_n$: ВВ $\{\xi[t_{k-1}; t_k]\}_{k=1}^n$ Н/З в сукупності

ОЗН. ВП $\xi(t)$ називається *пуассонівським* з інтенсивністю λ , якщо

- 1) $\xi(0) = 0$ м.н.: $P\{\xi(0) = 0\} = 1$;
- 2) $\xi(t)$ має незалежні прирости;
- 3) $\forall s, t \in [0; +\infty]$, $s < t$: $\xi[s; t]$ має розподіл Пуассона з параметром $\lambda(t-s)$.

Приклади: розпад радіоактивної речовини, виклики аварійних служб.

Алгоритм моделювання пуассонівського процесу такий: моменти τ_n стрибків на 1 моделюються як $\tau_n = \sum_{k=1}^n \eta_k$, де $\{\eta_k\}$ — Н/З ВВ, кожна з яких має показниковий розподіл з параметром λ . Тоді $\xi(t) = \sup\{n \in \mathbb{Z}_+ : \sum_{i=1}^n \eta_i \leq t\}$.

Узагальненням є *складний пуассонівський процес* — коли стрибки в моменти τ_i рівні не +1, а δ_i , де $\{\delta_i\}$ — Н/З однаково розподілені ВВ.

ОЗН. ВП $\xi(t)$ називається *вінерівським*, якщо

- 1) $\xi(0) = 0$ м.н.;
- 2) $\xi(t)$ має Н/З прирости;
- 3) $\forall s, t \in [0; +\infty]$, $s < t$: $\xi[s; t] \sim N_{0; t-s}$.

Приклад: одновимірний броунівський рух.

Алгоритм моделювання залежить від того, що моделюється. Якщо це станни процесу в моменти $t_1, t_2, \dots$, де $t_i < t_{i+1}$, то покладаємо $\xi(0) := 0$, далі послідовно $\xi(t_n) := \xi(t_{n-1}) + \eta_n$, де η_n Н/З і $\eta_n \sim N_{0; t_n - t_{n-1}}$.

3. Метод Монте-Карло

3.1. Обчислення інтегралів

Розглянемо задачу обчислення інтеграла $I = I(f) = \int_{\mathfrak{X}} f(x) \mu(dx)$ на просторі з ймовірнісною мірою $\{\mathfrak{X}; \mathfrak{F}; \mu\}$ методами Монте-Карло (ММК).

РЕС. Згідно ПЗВЧ, якщо ми вміємо моделювати реалізації $x_1, x_2, \dots, x_n$ ВВ ξ , розподіленої на $\mathfrak{X}$ згідно міри $\mu(\cdot)$, то

$$P\left\{\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n f(x_k) = I\right\} = 1$$

Якщо f інтегровна з квадратом на $\mathfrak{X}$ за μ , $f \in \mathcal{L}_{\mathfrak{X}}^2(\mu)$, то за ЦГТ

$$\lim_{n \rightarrow \infty} P\left\{\left|\frac{1}{n} \sum_{k=1}^n f(x_k) - I\right| < \frac{\sigma_f}{\sqrt{n}} y\right\} = \sqrt{\frac{2}{\pi}} \int_0^y e^{-\frac{u^2}{2}} du$$

де $\sigma_f^2 = Df(\xi)$, коли $\xi \sim \mu$: $\sigma_f^2 = \int_{\mathfrak{X}} f^2(x) \mu(dx) - \left(\int_{\mathfrak{X}} f(x) \mu(dx)\right)^2$.

Візьмемо $p \approx 1$ і $y_p > 0$ ($y_p \xrightarrow{p \rightarrow 1} +\infty$), для якого $\sqrt{\frac{2}{\pi}} \int_0^{y_p} e^{-\frac{u^2}{2}} du = p$. Тоді

похибка $\left|\frac{1}{n} \sum_{k=1}^n f(x_k) - I\right|$ не перевищує $\frac{\sigma_f}{\sqrt{n}} y_p$ з ймовірністю $\approx p$ для достатньо великих n ; іншими словами, похибка із зростанням n зменшується як $O(n^{-\frac{1}{2}})$.

Це дуже повільне зростання точності: наприклад, аби вона збільшилась в 1000 разів, треба виконати в 1000000 разів більше обчислень, або витратити в 1000000 разів більше часу.

Опишемо кілька методів, що дозволяють зменшити дисперсію $\sigma_f^2 = Df(\xi)$ і тим самим збільшити точність. Оскільки $\sigma_f^2 = Mf^2(\xi) - I^2$, де I — шуканий інтеграл, то намагаємось зменшити $Mf^2(\xi)$.

3.1.1. Виділення головної частини

Нехай $g(x) \in \mathcal{L}_{\mathfrak{X}}^2(\mu)$ близька до f у метриці $\mathcal{L}_{\mathfrak{X}}^2(\mu)$ ($\int_{\mathfrak{X}} |f(x) - g(x)|^2 \mu(dx)$ малий), причому $J = \int_{\mathfrak{X}} g(x) \mu(dx)$ відоме — знайдене аналітичними або чисельними методами. Тоді $I = \int_{\mathfrak{X}} g(x) \mu(dx) + \int_{\mathfrak{X}} (f(x) - g(x)) \mu(dx) = J + R$, і ММК можна оцінити R : $\hat{R} = \frac{1}{n} \sum_{i=1}^n [f(x_i) - g(x_i)]$, де $\{x_i\}$ — реалізації ВВ ξ з розподілом μ на $\mathfrak{X}$. При цьому

$$\sigma_{f-g}^2 = \int_{\mathfrak{X}} [f(x) - g(x)]^2 \mu(dx) - \left(\int_{\mathfrak{X}} [f(x) - g(x)] \mu(dx)\right)^2 \leq Df(\xi) = \sigma_f^2$$

З іншого боку, доведеться додатково обчислювати $g(x_i)$.

ПР. Нехай $I = \int_0^1 e^x dx (= e - 1)$; тут $f(x) = e^x$, а $\mu(\cdot)$ — СР розподіл.

Значить, $\sigma_f^2 = \int_0^1 e^{2x} dx - I^2 = \frac{1}{2}(e - 1)(3 - e) \approx 0.242$. Представимо $e^x =$

$x + (e^x - x)$. Нехай $g(x) = x$. Тоді $J = \int_0^1 x dx = \frac{1}{2}$, а

$$\sigma_{f-g}^2 = \int_0^1 (e^x - x)^2 dx - (I - J)^2 = \frac{e^2}{2} - \frac{53}{12} \approx 0.0437$$

— дисперсія зменшилась приблизно у 5 разів. Відповідно у $\sqrt{5}$ разів зросте точність та швидкість досягнення бажаної точності.

ЗАУВ. Нехай $g(x) = \sum_{k=1}^m a_k \varphi_k(x)$, де $\{\varphi_k\}_{k=1}^m$ — задана система функцій.

Треба вибрати коефіцієнти a_k так, щоб $\sigma_{f-g}^2 \rightarrow \min$. Відомо, що цей $\min$ досягається, коли $g(x)$ — часткова сума ряду Фур'є функції f за системою, отриманою із $\{\varphi_k\}$ процесом ортогоналізації.

Якщо відразу $\{\varphi_k\}$ ортонормовані в $\mathcal{L}_{\mathfrak{X}}^2(\mu)$, тобто $\int_{\mathfrak{X}} \varphi_k^2(x) \mu(dx) = 1$ та $\int_{\mathfrak{X}} \varphi_i(x) \varphi_j(x) \mu(dx) = 0$ при $i \neq j$, то a_k повинні бути коефіцієнтами Фур'є функції f за системою $\{\varphi_k\}$: $a_k = \int_{\mathfrak{X}} f(x) \varphi_k(x) \mu(dx)$.

3.1.2. Зниження порядку інтегрування

Нехай $\mathfrak{X}$ — s -вимірна область, $\mathfrak{X} \ni x = (x_1; x_2; \dots; x_s)$, і координати x_i розбиті на дві групи, $x = (y; z)$, де $y = (x_1; x_2; \dots; x_t)$, $z = (x_{t+1}; x_{t+2}; \dots; x_s)$, а міра $\mu(\cdot)$ АбсНеп зі щільністю $p(x)$, звідки шуканий інтеграл $I = \int_{\mathfrak{X}} f(x) p(x) dx$.

Нехай також $\mathfrak{X} = Y \times Z$, так що $I = \int_Y \int_Z f(y; z) p(y; z) dz dy$.

Припустимо, що аналітичними/чисельними способами легко обчислювати $p_1(y) = \int_Z p(y; z) dz$ та $f_1(y) = \frac{1}{p_1(y)} \int_Z f(y; z) p(y; z) dz$ (при цьому $\int_Z \frac{p(y; z)}{p_1(y)} dz = 1$).

Тоді $I = \int_Y f_1(y) p_1(y) dy$, і має місце

ТЕОР. При оцінці I сумою $\hat{I}_2 = \frac{1}{n} \sum_{i=1}^n f_1(y_i)$, де $\{y_i\}$ є вибіркою із розподілу

зі щільністю $p_1(y)$, дисперсія буде не більше, ніж при оцінці $\hat{I}_1 = \frac{1}{n} \sum_{i=1}^n f(y_i; z_i)$, де $\{(y_i; z_i)\} = \{x_i\}$ є вибіркою із розподілу зі щільністю $p(y; z) = p(x)$.

◀ Нехай дисперсія при оцінці $\hat{I}_1 \in D_1$, а при оцінці $\hat{I}_2 \in D_2$.

$$\begin{aligned}
D_1 - D_2 &= \int_Y \int_Z f^2(y; z) p(y; z) dz dy - I - \int_Y f_1^2(y) p_1(y) dy + I = \\
&= \int_Y \left[\int_Z f^2(y; z) \frac{p(y; z)}{p_1(y)} dz - f_1^2(y) \right] p_1(y) dy = \\
&= \int_Y \underbrace{\left[\int_Z f^2(y; z) \frac{p(y; z)}{p_1(y)} dz - \left(\int_Z f(y; z) \frac{p(y; z)}{p_1(y)} dz \right)^2 \right]}_{d(y)} p_1(y) dy
\end{aligned}$$

де $d(y) = Df(y; \zeta)$ при фіксованому $y \in Y$ та ζ — ВВ на Z зі щільністю $p_y(z) = \frac{p(y; z)}{p_1(y)}$. Отже, $d(y) \geq 0$, а тому і $D_1 - D_2 = \int_Y d(y) p_1(y) dy \geq 0$. ►

Але якщо $f_1(y)$ буде складнішою за $f(x)$, це нівелює зменшення дисперсії.

3.1.3. Заміна міри (метод суттєвої вибірки)

Нехай $I = \int_{\mathfrak{X}} f(x) p(x) dx$, де $p(x)$ — щільність $\mu(\cdot)$. Візьмемо щільність $q(x)$ на $\mathfrak{X}$: $q(x) > 0$ при $p(x) > 0$. Тоді $I = \int_{\mathfrak{X}} f(x) \frac{p(x)}{q(x)} q(x) dx = \int_{\mathfrak{X}} h(x) q(x) dx$.

Так виникає оцінка $\hat{I}_2 = \frac{1}{n} \sum_{i=1}^n h(y_i)$, де $\{y_i\}$ — реалізації ζ , розподіленої на $\mathfrak{X}$ зі щільністю $q(x)$. $\{\hat{I}_2\} = \{\hat{I}_2(q)\}$ є сукупністю незміщених оцінок I . Для якої $q(x)$ дисперсія σ_h^2 найменша?

ТЕОР. $\sigma_h^2 \rightarrow \min$ для щільності $q_0(x) = \frac{|f(x)|p(x)}{\int_{\mathfrak{X}} |f(x)|p(x) dx}$ (тобто пропорційної $|f(x)|p(x)$), і $\sigma_{h_0}^2 = \left(\int_{\mathfrak{X}} |f(x)|p(x) dx \right)^2 - I^2$.

$$\begin{aligned}
\blacktriangleleft \sigma_{h_0}^2 &= \int_{\mathfrak{X}} h_0^2(x) q_0(x) dx - \left(\int_{\mathfrak{X}} h_0(x) q_0(x) dx \right)^2 = \int_{\mathfrak{X}} f^2(x) \cdot \frac{p^2(x)}{q_0^2(x)} \cdot q_0(x) dx - I^2 = \\
&= \int_{\mathfrak{X}} f^2(x) \cdot \frac{p^2(x)}{|f(x)|p(x)} \cdot \left[\int_{\mathfrak{X}} |f(y)|p(y) dy \right] dx - I^2 = \left(\int_{\mathfrak{X}} |f(x)|p(x) dx \right)^2 - I^2
\end{aligned}$$

$$\begin{aligned}
&\text{Оптимальність випливає із нерівності } \sigma_h^2 - \sigma_{h_0}^2 = \\
&= \left(\int_{\mathfrak{X}} h^2(x) q(x) dx - I^2 \right) - \left(\int_{\mathfrak{X}} h_0^2(x) q_0(x) dx - I^2 \right) = \\
&= \int_{\mathfrak{X}} \left[|f(x)| \cdot \frac{p(x)}{q(x)} \right]^2 q(x) dx - \left(\int_{\mathfrak{X}} |f(x)|p(x) dx \right)^2 = \\
&= \int_{\mathfrak{X}} \left[|f(x)| \cdot \frac{p(x)}{q(x)} \right]^2 q(x) dx - \left(\int_{\mathfrak{X}} \left[|f(x)| \cdot \frac{p(x)}{q(x)} \right] \cdot q(x) dx \right)^2 = \\
&= D \left[|f(\zeta)| \cdot \frac{p(\zeta)}{q(\zeta)} \right] \geq 0, \text{ де } \zeta \text{ розподілена на } \mathfrak{X} \text{ зі щільністю } q(x). \blacktriangleright
\end{aligned}$$

Однак задача моделювання ВВ зі щільністю $q_0(x)$ у загальному випадку настільки ж складна, як обчислення інтеграла I .

ЗАУВ. Цей метод можна використати і коли $I = \int_{\mathfrak{X}} f(x)\mu(dx)$, тобто $\mu(\cdot)$ не обов'язково АбсНеп. При цьому застосовується похідна Радона-Нікодима.

ПР. Покажемо, як у задачах, де σ_f нескінченна, за допомогою даного методу можна перейти до оцінок, у яких вона скінченна. Нехай $I = \int_0^1 \frac{f(x)}{\sqrt{x}} dx$, де $f(x)$ обмежена, $|f(x)| \leq K$, — I збігається, — але у деякому правому околі 0 $|f(x)| \geq C > 0$ (кажуть, що f відділена від 0).

Тоді $\hat{I}_1 = \frac{1}{n} \sum_{i=1}^n \frac{f(x_i)}{\sqrt{x_i}}$, де $\{x_i\}$ — вибірка з СР розподілу, є незміщеною оцінкою I , але $\sigma_f^2 = +\infty$, бо другий момент $\int_0^1 \frac{f^2(x)}{x} dx \geq \int_0^\delta \frac{C^2}{x} dx = +\infty$, і ми не можемо скористатись ЦГТ для з'ясування похибки $\hat{I}_1$ при різних n .

$q(x) = \frac{1}{2\sqrt{x}}$ є щільністю на $[0; 1]$, тож перепишемо $I = 2 \int_0^1 f(x)q(x)dx$.

В оцінки $\hat{I}_2 = \frac{2}{n} \sum_{i=1}^n f(y_i)$, де $\{y_i\}$ є вибіркою із розподілу зі щільністю $q(x)$, дисперсія $\sigma_{f;q}^2 = \int_0^1 (2f(x))^2 q(x)dx - I^2 = 2 \int_0^1 \frac{f^2(x)}{\sqrt{x}} dx - I^2$ скінченна.

3.1.4. Розшарування вибірки

Нехай $\mathfrak{X} = \bigcup_{k=1}^m \mathfrak{X}_k$, де $\mathfrak{X}_i \cap \mathfrak{X}_j = \emptyset$, $i \neq j$ — розбиття області інтегрування.

Тоді $I = \sum_{k=1}^m \int_{\mathfrak{X}_k} f(x)\mu(dx)$. Введемо міру $\mu_k(\cdot) = \frac{\mu(\cdot)}{\mu(\mathfrak{X}_k)}$ на $\mathfrak{X}_k$ та перепишемо

$I = \sum_{k=1}^m \mu(\mathfrak{X}_k) \int_{\mathfrak{X}_k} f(x)\mu_k(dx)$. Оскільки $\mu_k(\cdot)$ є ймовірнісною мірою на $\mathfrak{X}_k$, оцінимо $J_k = \int_{\mathfrak{X}_k} f(x)\mu_k(dx)$ ММК. Позначивши n_k кількість реалізацій $x_i^{(k)} \in \mathfrak{X}_k$

ВВ з розподілом $\mu_k(\cdot)$, матимемо оцінку $\hat{I} = \sum_{k=1}^m \frac{\mu(\mathfrak{X}_k)}{n_k} \sum_{i=1}^{n_k} f(x_i^{(k)})$.

Оскільки $x_i^{(k)}$ як ВВ Н/З, то $D\hat{I} = \sum_{k=1}^m \frac{\mu^2(\mathfrak{X}_k)}{n_k^2} \cdot n_k Df(x_i^{(k)})$; мінімізувати її

означає мінімізувати лінійну комбінацію моментів $\sum_{k=1}^m \frac{\mu^2(\mathfrak{X}_k)}{n_k} \cdot m_2(f(x_i^{(k)}))$.

ТВ. $D\hat{I} \rightarrow \min$, коли пропорції між n_k близькі до пропорцій між величинами $\int_{\mathfrak{X}_k} f(x)\mu(dx)$ (це нагадує метод суттєвої вибірки).

3.1.5. Випадкові інтерполяційно-квадратурні формули

Нехай на $\mathfrak{X}$ задані лінійно незалежні функції $\varphi_1(x), \dots, \varphi_m(x)$ (тотожність $\sum_{i=1}^m a_i \varphi_i(x) \equiv 0$ досягається лише при $a_1 = \dots = a_m = 0$) і $x_1, \dots, x_m$ — різні точки $\mathfrak{X}$. Для функції $f: \mathfrak{X} \rightarrow \mathbb{R}^1$ означимо коефіцієнти $c_1, \dots, c_m$ умовою

$$\sum_{j=1}^m c_j \varphi_j(x_i) = f(x_i), \quad i = \overline{1, m}$$

За формулами Крамера інтерполяційний поліном $L(f; x) = \sum_{j=1}^m c_j \varphi_j(x)$ є

$$L(f; x) = -\frac{1}{\det \|\varphi_i(x_j)\|_{1 \leq i, j \leq m}} \begin{vmatrix} 0 & \varphi_1(x) & \dots & \varphi_m(x) \\ f(x_1) & \varphi_1(x_1) & \dots & \varphi_m(x_1) \\ \dots & \dots & \dots & \dots \\ f(x_m) & \varphi_1(x_m) & \dots & \varphi_m(x_m) \end{vmatrix}$$

де $\det \|\varphi_i(x_j)\|_{1 \leq i, j \leq m} \neq 0$. Дійсно, підставивши $x = x_k$ та віднявши від першого рядка $(k+1)$ -й, отримаємо у першому рядку єдиний ненульовий коефіцієнт у першому стовпчику, $-f(x_k)$. Розклавши цей визначник за першим рядком, $L(f; x_k) = -\frac{1}{\det \|\varphi_i(x_j)\|_{1 \leq i, j \leq m}} \times (-1)^{1+1} (-f(x_k)) \times \det \|\varphi_i(x_j)\|_{1 \leq i, j \leq m} = f(x_k)$

Припустимо, що f та $\{\varphi_i\}$ належать $\mathcal{L}_{\mathfrak{X}}^2(\mu)$, причому $\{\varphi_i\}$ ортонормовані. Якщо вони є достатньо гладкими, то при вдалому виборі точок $\{x_j\}$ поліном $L(f; x)$ є хорошим наближенням $f(x)$. З тих же міркувань $\int_{\mathfrak{X}} \varphi_1(x) L(f; x) \mu(dx)$ буде хорошим наближенням $\int_{\mathfrak{X}} \varphi_1(x) f(x) \mu(dx)$; зокрема, при $\varphi_1(x) \equiv 1$ маємо наближення I . Як воно виглядає? Розкладаючи за першим рядком,

$$\begin{aligned} \int_{\mathfrak{X}} L(f; x) \mu(dx) &= -\frac{1}{\det \|\varphi_i(x_j)\|_{1 \leq i, j \leq m}} \times \\ &\times \sum_{r=1}^m (-1)^{2+r} \det \|f(x_k); \varphi_1(x_k); \dots; \varphi_{r-1}(x_k); \varphi_{r+1}(x_k); \dots; \varphi_m(x_k)\|_{1 \leq k \leq m} \times \\ &\times \int_{\mathfrak{X}} \varphi_1(x) \varphi_r(x) \mu(dx) = \left\{ \text{інтеграл рівний } 0 \text{ при } r \neq 1 \text{ (ортонормованість)} \right\} = \\ &= \frac{1}{\det \|\varphi_i(x_j)\|_{1 \leq i, j \leq m}} \cdot \det \|f(x_k); \varphi_2(x_k); \dots; \varphi_m(x_k)\|_{1 \leq k \leq m} \end{aligned}$$

тепер розкладемо визначник у чисельнику за першим стовпцем $\{f(x_k)\}$ та отримаємо $\sum_{k=1}^m \mathcal{A}_k(x_1; \dots; x_m) f(x_k)$, де

$$\mathcal{A}_k(x_1; \dots; x_m) = \frac{(-1)^{1+k}}{\det \|\varphi_i(x_j)\|_{1 \leq i, j \leq m}} \begin{vmatrix} \varphi_2(x_1) & \dots & \varphi_m(x_1) \\ \dots & \dots & \dots \\ \varphi_2(x_{k-1}) & \dots & \varphi_m(x_{k-1}) \\ \varphi_2(x_{k+1}) & \dots & \varphi_m(x_{k+1}) \\ \dots & \dots & \dots \\ \varphi_2(x_m) & \dots & \varphi_m(x_m) \end{vmatrix}$$

Отже, оцінкою $I \in \hat{I} = \sum_{k=1}^m \mathcal{A}_k(x_1; \dots; x_m) f(x_k) = \mathcal{K}_m(f; x_1; \dots; x_m)$, яка називається *інтерполяційно-квадратурною формулою* (ІКФ).

Нехай набір точок $\xi = (x_1; \dots; x_m) \in \mathfrak{X}^m$ — ВипВек зі щільністю $g(t)$, де $t = (t_1; \dots; t_m)$, відносно міри $\mu^m(dt)$. Позначимо

$$\Delta(t) = \det \|\varphi_i(t_j)\|_{1 \leq i, j \leq m} \neq 0$$

$$\Delta(f; t) = \det \|f(t_k); \varphi_2(t_k); \dots; \varphi_m(t_k)\|_{1 \leq k \leq m}$$

тоді оцінка $\mathcal{K}_m(f; \xi) = \frac{\Delta(f; \xi)}{\Delta(\xi)}$. Як вибрати g ? Повинна мати місце рівність

$$M\mathcal{K}_m(f; \xi) \left\{ = \int_{\mathfrak{X}^m} \mathcal{K}_m(f; t) g(t) \mu^m(dt) \right\} = \int_{\mathfrak{X}} \varphi_1(x) f(x) \mu(dx)$$

ЛЕМА. Для наборів $\{\varphi_i(x)\}_{i=1}^m$ та $\{\psi_i(x)\}_{i=1}^m$ функцій з $\mathcal{L}_{\mathfrak{X}}^2(\mu)$
 $\int_{\mathfrak{X}^m} \det \|\varphi_i(t_j)\|_{1 \leq i, j \leq m} \cdot \det \|\psi_i(t_j)\|_{1 \leq i, j \leq m} \mu^m(dt) = m! \cdot \det \|\langle \varphi_i; \psi_j \rangle\|_{1 \leq i, j \leq m}$
де $\langle \varphi_i; \psi_j \rangle = \int_{\mathfrak{X}} \varphi_i(x) \psi_j(x) \mu(dx)$.

ТЕОР. 1. Нехай $\{\varphi_i(x)\}_{i=1}^m$ — ортонормована в $\mathcal{L}_{\mathfrak{X}}^2(\mu)$ система. Тоді

1) $\frac{1}{m!} \Delta^2(t)$ — щільність за мірою $\mu^m(\cdot)$;

2) Якщо ВипВек ξ має таку щільність, то $M\mathcal{K}_m(f; \xi) = \int_{\mathfrak{X}} \varphi_1(x) f(x) \mu(dx)$.

◀ (1) $\Leftrightarrow \frac{1}{m!} \int_{\mathfrak{X}^m} \Delta^2(t) \mu^m(dt) = 1$ випливає з леми для $\psi_i = \varphi_i$ і того, що

$\|\langle \varphi_i; \varphi_j \rangle\|_{1 \leq i, j \leq m}$ є одинична матриця $\mathcal{E}$ внаслідок ортонормованості $\{\varphi_i\}$.

Відтак $M\mathcal{K}_m(f; \xi) = \int_{\mathfrak{X}^m} \frac{\Delta(f; t)}{\Delta(t)} \cdot \frac{1}{m!} \Delta^2(t) \mu^m(dt) = \frac{1}{m!} \int_{\mathfrak{X}^m} \Delta(f; t) \Delta(t) \mu^m(dt)$.

Візьмемо $\psi_1 = f$ та $\psi_k = \varphi_k$ для $k = \overline{2, m}$. За лемою $M\mathcal{K}_m(f; \xi) =$

$$= \begin{vmatrix} \langle f; \varphi_1 \rangle & \langle f; \varphi_2 \rangle & \dots & \langle f; \varphi_m \rangle \\ 0 & 1 & \dots & 0 \\ \dots & \dots & 1 & \dots \\ 0 & 0 & \dots & 1 \end{vmatrix} = \langle f; \varphi_1 \rangle \cdot \det \mathcal{E} = \int_{\mathfrak{X}} \varphi_1(x) f(x) \mu(dx) \blacktriangleright$$

ОЗН. $\{\varphi_i\}_{i=1}^m$ регулярна відносно міри $\mu(\cdot)$, якщо $\mu^m\{t : \Delta(t) = 0\} = 0$.

ТЕОР. 2. В умовах теореми 1

$$D\mathcal{K}_m(f; \xi) \leq \int_{\mathfrak{X}} |f(x) - \sum_{k=1}^m \langle f; \varphi_k \rangle \cdot \varphi_k(x)|^2 \mu(dx)$$

і для регулярної $\{\varphi_i\}_{i=1}^m$ нерівність перетворюється на рівність.

$$\blacktriangleleft D\mathcal{K}_m(f; \xi) = M\mathcal{K}_m^2(f; \xi) - (M\mathcal{K}_m(f; \xi))^2 = M\mathcal{K}_m^2(f; \xi) - (\langle f; \varphi_1 \rangle)^2,$$

$$M\mathcal{K}_m^2(f; \xi) = \int_{t \in \mathfrak{X}^m : \Delta(t) \neq 0} \frac{\Delta^2(f; t)}{\Delta^2(t)} \cdot \frac{1}{m!} \Delta^2(t) \mu^m(dt) = \frac{1}{m!} \int_{t \in \mathfrak{X}^m : \Delta(t) \neq 0} \Delta^2(f; t) \mu^m(dt)$$

Якщо $\{\varphi_i\}$ — регулярна, то $\mu^m\{t : \Delta(t) = 0\} = 0$, і цією множиною можна знехтувати; у загальному випадку отриманий інтеграл не перевищує аналогічного інтеграла по усій множині $\mathfrak{X}^m$, $\delta_m = \frac{1}{m!} \int_{\mathfrak{X}^m} \Delta^2(f; t) \mu^m(dt)$.

Застосуємо лему для наборів $\{\varphi_i\} = \{\psi_i\} = (f; \varphi_2; \dots; \varphi_m)$:

$$\delta_m = \begin{vmatrix} \langle f; f \rangle & \langle f; \varphi_2 \rangle & \dots & \langle f; \varphi_m \rangle \\ \langle f; \varphi_2 \rangle & 1 & \dots & 0 \\ \dots & \dots & 1 & \dots \\ \langle f; \varphi_m \rangle & 0 & \dots & 1 \end{vmatrix}$$

Розкладемо цей визначник за останнім рядком:

$$\begin{aligned} \delta_m &= \langle f; \varphi_m \rangle \cdot (-1)^{m+1} \cdot \underbrace{\begin{vmatrix} \langle f; \varphi_2 \rangle & \langle f; \varphi_3 \rangle & \dots & \langle f; \varphi_{m-1} \rangle & \langle f; \varphi_m \rangle \\ 1 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & 0 \end{vmatrix}}_{\text{розкладаємо по останньому стовпцю}} + \\ &+ 1 \cdot (-1)^{m+m} \cdot \begin{vmatrix} \langle f; f \rangle & \langle f; \varphi_2 \rangle & \dots & \langle f; \varphi_{m-1} \rangle \\ \langle f; \varphi_2 \rangle & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ \langle f; \varphi_{m-1} \rangle & 0 & \dots & 1 \end{vmatrix} = \\ &= \langle f; \varphi_m \rangle \cdot (-1)^{m+1} \cdot \langle f; \varphi_m \rangle \cdot (-1)^{m-1+1} \cdot \det \mathcal{E} + \delta_{m-1} = \\ &= \delta_{m-1} - (\langle f; \varphi_m \rangle)^2 = \delta_{m-2} - (\langle f; \varphi_{m-1} \rangle)^2 - (\langle f; \varphi_m \rangle)^2 = \dots \\ &= \langle f; f \rangle - \sum_{k=2}^m (\langle f; \varphi_k \rangle)^2, \text{ оскільки } \delta_1 = \langle f; f \rangle. \text{ Значить,} \\ &DK_m(f; \xi) \leq \langle f; f \rangle - \sum_{k=1}^m (\langle f; \varphi_k \rangle)^2 = \\ &= \int_{\mathfrak{X}} f^2(x) \mu(dx) - \sum_{k=1}^m \langle f; \varphi_k \rangle \int_{\mathfrak{X}} f(x) \varphi_k(x) \mu(dx) = \\ &= \int_{\mathfrak{X}} f^2(x) \mu(dx) - \int_{\mathfrak{X}} f(x) \left[\sum_{k=1}^m \langle f; \varphi_k \rangle \varphi_k(x) \right] \mu(dx) = \\ &= \int_{\mathfrak{X}} [f^2(x) - 2f(x) \sum_{k=1}^m \langle f; \varphi_k \rangle \varphi_k(x)] \mu(dx) + \underbrace{\sum_{k=1}^m (\langle f; \varphi_k \rangle)^2 \cdot \int_{\mathfrak{X}} \varphi_k^2(x) \mu(dx)}_{=1} + \\ &+ 2 \sum_{1 \leq i < j \leq m} \langle f; \varphi_i \rangle \cdot \langle f; \varphi_j \rangle \cdot \underbrace{\int_{\mathfrak{X}} \varphi_i(x) \varphi_j(x) \mu(dx)}_{=0, \varphi_i \perp \varphi_j} = \\ &= \int_{\mathfrak{X}} [f^2(x) - 2f(x) \sum_{k=1}^m \langle f; \varphi_k \rangle \varphi_k(x)] \mu(dx) + \int_{\mathfrak{X}} \left[\sum_{k=1}^m (\langle f; \varphi_k \rangle \cdot \varphi_k(x))^2 + \right. \\ &\quad \left. + 2 \sum_{1 \leq i < j \leq m} [\langle f; \varphi_i \rangle \cdot \varphi_i(x)] \cdot [\langle f; \varphi_j \rangle \cdot \varphi_j(x)] \right] \mu(dx) = \\ &= \int_{\mathfrak{X}} [f^2(x) - 2f(x) \sum_{k=1}^m \langle f; \varphi_k \rangle \varphi_k(x)] \mu(dx) + \int_{\mathfrak{X}} \left[\sum_{k=1}^m \langle f; \varphi_k \rangle \cdot \varphi_k(x) \right]^2 \mu(dx) = \\ &= \int_{\mathfrak{X}} \left[f(x) - \sum_{k=1}^m \langle f; \varphi_k \rangle \varphi_k(x) \right]^2 \mu(dx) \end{aligned}$$

і для регулярних $\{\varphi_i\}$ нерівність перетворюється на рівність. ►

Розглядаючи інші методи, ми відмічали, що за зменшення дисперсії, як правило, доводиться «платити» зростанням обчислювальної складності. Для методу випадкових ІКФ це пов'язано з:

- використанням формул виду $\frac{\Delta(f;\xi)}{\Delta(\xi)}$;
- застосуванням методу відбору для отримання елементів вибірки;
- обчисленням визначників великих порядків m .

Він ефективний при невеликих m , коли f близька до лінійної комбінації $\{\varphi_k\}$, — як видно з теореми 2, це зменшує дисперсію та похибку. Також її зменшує додаткова залежність між $\{x_j\}$, коли одні вузли x_j обираються випадково згідно міри $\mu(\cdot)$, а інші є детермінованими функціями від них.

3.2. Ймовірнісні методи розв'язку СЛАР

СЛАР — системи лінійних алгебраїчних рівнянь.

Розглянемо марківський ланцюг (ЛМ) із станами $1, 2, \dots, n$, початковим розподілом $p^0 = (p_1^0; \dots; p_n^0)$ та перехідною матрицею $P = \|p_{ij}\|_{1 \leq i, j \leq n}$, $\sum_{j=1}^n p_{ij} = 1 - g_i$, де $g_i > 0$ — ймовірність обриву траєкторії в стані i .

Будь-якій траєкторії $\omega = i_0 \rightarrow i_1 \rightarrow \dots \rightarrow i_r (\rightarrow \times)$ відповідає її ймовірність $p_{i_0}^0 p_{i_0 i_1} \times \dots \times p_{i_{r-1} i_r} g_{i_r}$, і на дискретній множині усіх скінченних траєкторій виникає ймовірнісна міра:

$$\sum_{r=0}^{\infty} \sum_{i_0=1}^n \sum_{i_1=1}^n \dots \sum_{i_r=1}^n p_{i_0}^0 p_{i_0 i_1} \times \dots \times p_{i_{r-1} i_r} g_{i_r} = 1$$

Доведемо дану рівність, тобто покажемо, що це дійсно ймовірнісна міра.

◀ Розглянемо СЛАР $Z = PZ + G$, де вектор-стовпці $G = (g_1; g_2; \dots; g_n)^T$, а Z — невідоме. Перша норма матриці P , $\lambda_1(P) = \max_{1 \leq i \leq n} \sum_{j=1}^n p_{ij} = \max_i (1 - g_i) < 1$, бо всі $g_i > 0$. Як відомо, така система має єдиний ітераційний розв'язок

$$Z = \lim_{m \rightarrow \infty} Z_m = \lim_{m \rightarrow \infty} (\mathcal{E} + P + P^2 + \dots + P^m)G$$

З іншого боку, легко бачити, що цим розв'язком є $(1; 1; \dots; 1)^T$:

$$P(1; 1; \dots; 1)^T + G = \left\| \sum_{j=1}^n p_{ij} \right\|_{1 \leq i \leq n}^T + \|g_i\|_{1 \leq i \leq n}^T = (1; 1; \dots; 1)^T$$

Тому $(1; 1; \dots; 1)^T = Z = \lim_{m \rightarrow \infty} (\mathcal{E} + P + P^2 + \dots + P^m)G =$

$$= \left\| \sum_{r=0}^{\infty} \sum_{i_0=1}^n \dots \sum_{i_r=1}^n p_{i_0 i_1} \times \dots \times p_{i_{r-1} i_r} g_{i_r} \right\|_{1 \leq i_0 \leq n}^T$$

а оскільки $\langle p^0; Z \rangle = \sum_{j=1}^n p_j^0 = 1$, то, міняючи місцями $\sum_{i_0}$ й $\sum_r$, отримуємо

$$1 = \langle p^0; Z \rangle = \sum_{r=0}^{\infty} \sum_{i_0=1}^n \left[\sum_{i_1=1}^n \dots \sum_{i_r=1}^n p_{i_0 i_1}^0 \times \dots \times p_{i_{r-1} i_r} g_{i_r} \right] \blacktriangleright$$

Якщо на траєкторіях $i_0 \rightarrow i_1 \rightarrow \dots \rightarrow i_r$ задано функціонал $\Phi(i_0; i_1; \dots; i_r)$ із скінченним $M\Phi$, то

$$M\Phi = \sum_{r=0}^{\infty} \sum_{i_0=1}^n \sum_{i_1=1}^n \dots \sum_{i_r=1}^n \Phi(i_0; i_1; \dots; i_r) p_{i_0}^0 p_{i_0 i_1} \times \dots \times p_{i_{r-1} i_r} g_{i_r} \quad (*)$$

Розглянемо тепер СЛАР виду $X = AX + F$, де

$$A = \|a_{ij}\|_{1 \leq i, j \leq n}, \quad X = (x_1; x_2; \dots; x_n)^T, \quad F = (f_1; f_2; \dots; f_n)^T$$

такі, що метод послідовних наближень (ітерацій) розв'язку є збіжним: покладаючи $X_m = AX_{m-1} + F$, де $X_0 = F$ (або $X_0 = (0; 0; \dots; 0)^T$), для $m \in \mathbb{N}$, матимемо $X_m \xrightarrow{m \rightarrow \infty} \tilde{X}$ — розв'язку вихідного рівняння.

Тоді існує єдиний розв'язок $\tilde{X}$, причому він може бути представлений у вигляді *ряду Неймана*: $\tilde{X} = \sum_{k=0}^{\infty} A^k F$.

Якщо додатково $\lambda_1(|A|) < 1$, тобто виконуються умови збіжності *мажорантного* ітераційного процесу

$$\tilde{X}_m = |A| \tilde{X}_{m-1} + |F|, \quad \text{де } |A| = \| \|a_{ij}\| \|_{1 \leq i, j \leq n}, \quad |F| = \| \|f_i\| \|_{1 \leq i \leq n}^T$$

то $\tilde{X}$ може бути представлений у вигляді

$$\tilde{X} = \| \sum_{k=0}^{\infty} \sum_{i_1=1}^n \dots \sum_{i_k=1}^n a_{i_0 i_1} a_{i_1 i_2} \times \dots \times a_{i_{k-1} i_k} f_{i_k} \|_{1 \leq i_0 \leq n}^T$$

і для довільного $H = (h_1; h_2; \dots; h_n)^T$:

$$\langle H; \tilde{X} \rangle = \sum_{k=0}^{\infty} \sum_{i_0=1}^n \sum_{i_1=1}^n \dots \sum_{i_k=1}^n h_{i_0} a_{i_0 i_1} a_{i_1 i_2} \times \dots \times a_{i_{k-1} i_k} f_{i_k} \quad (**)$$

Опишемо ймовірнісний метод обчислення $\langle H; \tilde{X} \rangle$. Візьмемо довільний ЛМ, що задається парою $(p^0; P)$ та задовольняє умови узгодження:

якщо $h_i \neq 0$, то $p_i^0 > 0$; якщо $a_{ij} \neq 0$, то $p_{ij} > 0$; якщо $f_i \neq 0$, то $g_i > 0$

І розглянемо функціонал $\Phi_0(i_0; i_1; \dots; i_r) = \frac{h_{i_0} a_{i_0 i_1} \times \dots \times a_{i_{r-1} i_r} f_{i_r}}{p_{i_0}^0 p_{i_0 i_1} \times \dots \times p_{i_{r-1} i_r} g_{i_r}}$.

Тоді, згідно (*) та (**), $M\Phi_0 = \langle H; \tilde{X} \rangle$. (Існування $M\Phi_0$ впливає зі збіжності мажорантного ітераційного процесу.)

Практично: моделюємо Н/З чином N траєкторій ЛМ, визначеного $(p^0; P)$ — $\{(i_0^j; i_1^j; \dots; i_{r_j}^j)\}_{j=1}^N$ — і обчислюємо на кожній функціонал Φ_0 . Отримаємо числа $\Phi_0^j = \Phi_0(i_0^j; i_1^j; \dots; i_{r_j}^j)$. Тоді на підставі ПЗВЧ

$$P \left\{ \frac{1}{N} \sum_{j=1}^N \Phi_0^j \xrightarrow{N \rightarrow \infty} \langle H; \tilde{X} \rangle \right\} = 1$$

тобто $\frac{1}{N} \sum_{j=1}^N \Phi_0^j$ буде «хорошою» оцінкою для $\langle H; \tilde{X} \rangle$.

ЗАУВ. Такі обчислення можна розпаралелити: j -й «процесор» незалежно від інших моделює траєкторію $\omega_{r_j}^j$ та обчислює Φ_0^j .

Так ми обчислюємо інтеграл (Лебега за дискретним розподілом) по траєкторіях випадкового процесу — скінченних послідовностях $i_0, i_1, \dots, i_r$.

Обчислення $\langle H; \tilde{X} \rangle$ розглядаємо як розв'язок СЛАР $X = AX + F$, бо, узявши $H_k = (\underbrace{0; 0; \dots; 0}_{k-1}; \underbrace{1; 0; 0; \dots; 0}_{n-k})$, маємо $\langle H_k; \tilde{X} \rangle = \tilde{x}_k$; відповідно для

$\Phi_{0;k}$ з $h_{i_0} = \delta_{i_0,k}$ оцінка $\frac{1}{N} \sum_{j=1}^N \Phi_{0;k}^j$ буде наближенням $\tilde{x}_k$. Ці n сум можна обчислювати одночасно, додаючи значення функціонала на траєкторії, яка почалася зі стану k , до одної суми з номером k .

ЗАУВ. Для економії пам'яті достатньо зберігати лише номери початкового, попереднього та поточного станів, адже

$$\Phi_0(i_0; i_1; \dots; i_r) = \left(\frac{h_{i_0}}{p_{i_0}^0} \right) \cdot \left(\frac{a_{i_0 i_1}}{p_{i_0 i_1}} \right) \times \dots \times \left(\frac{a_{i_{r-1} i_r}}{p_{i_{r-1} i_r}} \right) \cdot \left(\frac{f_{i_r}}{g_{i_r}} \right)$$

• Порівняємо *трудоемності* — кількості елементарних обчислювальних операцій — при ітераційному (R_n) та ймовірнісному (S_n) розв'язках СЛАР. Можна показати, що має місце асимптотична по $n \rightarrow \infty$ еквівалентність

$$\frac{S_n}{R_n} \sim \frac{N \log_2 n + \frac{n}{M}}{n^2}$$

де M — кількість ітерацій, необхідних для досягнення бажаної точності ε . Згідно з теорією ітераційних методів, $M \sim \frac{|\ln \varepsilon|}{|\ln \lambda_1|}$, де $\lambda_1 = \lambda_1(|A|)$. У свою чергу, як наслідок із ЦГТ для ММК, $N \sim \frac{\sigma^2}{\varepsilon^2}$. Звідси для фіксованого ε

$$\frac{S_n}{R_n} \sim \frac{1}{n^2} \left(\frac{\sigma^2}{\varepsilon^2} \log_2 n + n \cdot \frac{|\ln \lambda_1|}{|\ln \varepsilon|} \right) = \frac{\sigma^2}{\varepsilon^2} \cdot \frac{\log_2 n}{n^2} + \frac{|\ln \lambda_1|}{|\ln \varepsilon|} \cdot \frac{1}{n} \xrightarrow{n \rightarrow \infty} 0$$

так що ММК асимптотично по вимірності невідомого вектора X виявляється ефективнішим, швидшим за ітераційний метод.

Схожий результат можна отримати, коли матриця A має нульові коефіцієнти, причому кількість ненульових пропорційна qn^2 з $q < 1$.

З іншого боку, ймовірнісний метод внаслідок власне своєї стохастичності може привести до оцінки, далекої від точного $\tilde{X}$, до якого гарантовано збігається ітераційний процес.

4. Застосування методів Монте-Карло

4.1. Моделювання руху частинок крізь речовину

Нехай євклідів простір $\mathbb{R}^3$ з декартовою СК $Oxyz$ являє собою середовище, заповнене неоднорідною речовиною, кожна точка $\bar{r}(x; y; z)$ якого характеризується наступними параметрами:

- $\sigma(\bar{r}) \in [0; +\infty)$ — коефіцієнт послаблення потоку;
- $\sigma_a(\bar{r}) \in [0; \sigma(\bar{r}))$ — коефіцієнт поглинання (absorb),
- $\sigma_s(\bar{r}) \in [0; \sigma(\bar{r}))$ — коефіцієнт розсіювання (scatter), $\sigma_a(\bar{r}) + \sigma_s(\bar{r}) = \sigma(\bar{r})$
- $g(\bar{r}; s)$ — щільність розподілу кута розсіювання на $[0; \pi]$.

Частинка, котру розглядаємо як матеріальну точку, рухається крізь речовину. Внаслідок «зіткнення» з речовиною вона поглинається чи розсіюється (змінює напрямок руху). Вважаємо, що мають місце такі припущення:

- абсолютна величина $|\bar{V}|$ швидкості $\bar{V}$ руху частинки незмінна;
- почавши рух або розсіявшись у точці $\bar{c}$ (collision), частинка має напрям руху, який задається одиничним вектором $\bar{d} = (d_x; d_y; d_z)$, і рухається у цьому напрямку до наступного зіткнення $\bar{c} + l\bar{d}$;

- при цьому довжина вільного пробігу між зіткненнями l — ВВ, розподілена на $[0; +\infty]$ зі щільністю $p_l(t) = \sigma(\bar{z}(t))e^{-\int_0^t \sigma(\bar{z}(u))du} = \frac{d}{dt} \left[1 - e^{-\int_0^t \sigma(\bar{z}(u))du} \right]$, де $\bar{z}(t) = \bar{c} + t\bar{d}$;

- внаслідок зіткнення в т. $\bar{c}$ частинка поглинається з ймовірністю $\frac{\sigma_a(\bar{c})}{\sigma(\bar{c})}$ (тоді рух закінчується), чи розсіюється з ймовірністю $\frac{\sigma_s(\bar{c})}{\sigma(\bar{c})}$.

Таким чином, траєкторія частинки — ламана.

- під час розсіювання відбувається зміна напрямку руху. Вона визначається двома кутами: *азимутальним* φ — повороту навколо осі-напрямку руху до зіткнення, — та *розсіювання* θ — повороту в площині, визначеній азимутальним кутом, що проходить через попередній напрямок, який відраховується від цього напрямку. Азимутальний кут розподілений рівномірно на $[0; 2\pi]$ (ізотропний), кут розсіювання — зі щільністю $g(\bar{c}; s)$ на $[0; \pi]$.

З'ясуємо, як пов'язані декартові координати напрямку руху до $(\bar{d})$ та після $(\bar{d}')$ розсіювання в т. $\bar{c}$, визначеного азимутальним кутом φ та кутом розсіювання θ . Оскільки азимутальний кут розподілений рівномірно, його можна відліковувати від будь-якої вісі, що проходить в площині, перпендикулярній $\bar{d}$, через $\bar{c}$. Якщо перенести початок СК в $\bar{c}$, то в такій СК $O'x'y'z'$ ($O' = \bar{c}$) точки вказаної площини задовольняють рівняння

$$d_x x' + d_y y' + d_z z' = 0 \Leftrightarrow \langle \bar{d}; (x'; y'; z') \rangle = 0$$

Візьмемо в цій СК одиничні вектори $\bar{u}$ та $\bar{v}$, які разом з $\bar{d}$ утворюватимуть репер, тобто праву трійку $(\bar{u}; \bar{v}; \bar{d})$ взаємно перпендикулярних орт (праву — $\overline{uvd} > 0$). Достатньо вибрати $\bar{u} \perp \bar{d}$, тоді можна узяти $\bar{v} = \bar{d} \times \bar{u}$.

В СК з ортами $\bar{u}$, $\bar{v}$, $\bar{d}$ вектор $\bar{d}$ має координати $(0; 0; 1)$. Повернувши його навколо $\bar{u}$ на кут розсіювання θ , матимемо $\bar{d}_1$, який у цій СК має координати $(0; \sin \theta; \cos \theta)$. Повернемо $\bar{d}_1$ на азимутальний кут φ навколо $\bar{d}$, отримаємо (в цій же СК) $\bar{d}_2 = (-\sin \varphi \sin \theta; \cos \varphi \sin \theta; \cos \theta)$. Це і є новий напрямок $\bar{d}'$; аби виразити його в «абсолютній» декартовій СК $Oxyz$, лишається узяти відповідну лінійну комбінацію $\bar{u}$, $\bar{v}$, $\bar{d}$:

$$\bar{d}' = -\sin \varphi \sin \theta \cdot \bar{u} + \cos \varphi \sin \theta \cdot \bar{v} + \cos \theta \cdot \bar{d} \quad (*)$$

Кут φ моделюється як $2\pi\alpha$, де α — СР ВЧ. Моделювання θ залежить від щільності $g(\bar{c}; s)$; у загальному випадку, оскільки функція розподілу θ , $G(t) = \int_0^t g(\bar{c}; s)ds$ для $t \in [0; \pi]$, можна скористатись формулою обернення:

$\theta = G^{-1}(\beta)$, де β — СР, і, моделювавши β , знайти t_β , для якого $\int_0^{t_\beta} g(\bar{c}; s)ds = \beta$ (наприклад, методом ділення навпіл; див. нижче), та покласти $\theta := t_\beta$.

Якщо θ теж рівномірно розподілений — $\theta := \pi\beta$.

Тепер розглянемо моделювання l , довжини вільного пробігу. Згідно означення $p_l(t)$, $F_l(t) = 1 - e^{-\int_0^t \sigma(\bar{z}(u))du}$, де $\bar{z}(u) = \bar{c} + u\bar{d}$.

ОЗН. $\tau(t) = \int_0^t \sigma(\bar{z}(u))du$ — оптична довжина відрізка $[\bar{c}; \bar{c} + t\bar{d}]$.

$F_l(t) = 1 - e^{-\tau(t)}$ і $p_l(t) = \frac{d}{dt}\tau(t) \cdot e^{-\tau(t)}$. У загальному випадку, користуючись ФО, моделюємо $l := F^{-1}(\alpha)$ як розв'язок t рівняння $1 - e^{-\tau(t)} = \alpha$, де α — СР ВЧ. Після заміни $1 - \alpha$ на α матимемо $t = \tau^{-1}(-\ln \alpha)$ (**)

Знайдемо таке t методом ділення навпіл:

0) Покладаємо $S := 0$, $E := 1$;

1) Подвоюємо E (2, 4, ...), поки не стане $\tau(E) = \int_0^E \sigma(\bar{z}(u))du \geq -\ln \alpha$;

2) Покладаємо $M := \frac{S+E}{2}$ та перевіряємо:

а) якщо $\tau(M) < -\ln \alpha$, то $S := M$,

б) якщо $\tau(M) > -\ln \alpha$, то $E := M$;

3) Якщо $|S - E| < \delta$, де δ — бажана точність, беремо $l := M$; інакше повертаємось до кроку (2).

$\tau(M)$ знаходимо чисельними методами. Зокрема, для прискорення обчислень скористаємось адитивністю інтеграла:

$$\int_0^{M_{i+1}} \sigma = \int_0^{M_i} \sigma + \int_{M_i}^{M_{i+1}} \sigma.$$

У випадку, коли $\sigma(\bar{r})$ — кусково-стала, моделювати l простіше: промінь $\bar{c} + t\bar{d}$, $t \in [0; +\infty)$, розбивається на проміжки, на кожному з яких $\sigma(\bar{z})$ стала — $\lambda_1, \lambda_2, \dots$. Тоді $F_l(t)$ — «кусково-показникова» з різними λ_i у $e^{-\lambda_i t}$, точніше $F_l(t) = c_i - e^{-\lambda_i t}$ для t з i -го проміжку, тому шукати $F_l^{-1}(t)$ легше. Зокрема, коли $\sigma(\bar{r}) \equiv \lambda$, маємо $F_l(t) = 1 - e^{-\lambda t}$, звідки $l \leftarrow -\frac{1}{\lambda} \ln \alpha$.

Навіть якщо $\sigma(\bar{r})$ — не кусково-стала, область можна розбити на достатньо маленькі зони, у кожній з яких вона «приблизно» стала, і замість σ узяти відповідне наближення $\tilde{\sigma}$.

Переходимо до моделювання траєкторії частинки. Вона визначається послідовністю точок зіткнення $\bar{c}_0, \bar{c}_1, \dots$, де $\bar{c}_0$ — початкове положення, в якому частинка має початковий напрям $\bar{d}_0$.

Оскільки у даній моделі $\sigma_a(\bar{r})$ та $\sigma_s(\bar{r})$ потрібні тільки для знаходження ймовірностей поглинання $\frac{\sigma_a(\bar{r})}{\sigma(\bar{r})}$ та розсіювання $\frac{\sigma_s(\bar{r})}{\sigma(\bar{r})}$, можна перейти до ймовірності розсіювання — $s(\bar{r}) = \frac{\sigma_s(\bar{r})}{\sigma(\bar{r})}$, і вважати заданою відразу її. Тоді ймовірність поглинання в $\bar{r}$ — $(1 - s(\bar{r}))$.

- 0) Ініціалізація: частинка починає рух з $\bar{c}_0$ у напрямку $\bar{d}_0$;
- 1) Моделюємо довжину вільного пробігу до чергового зіткнення за (**):
 $l := \tau^{-1}(-\ln \alpha_1)$, де α_1 — СР ВЧ, $\tau(t) = \int_0^t \sigma(\bar{c}_i + u\bar{d}_i) du$;
- 2) Наступна точка зіткнення $\bar{c}_{i+1} := \bar{c}_i + l\bar{d}_i$;
- 3) Розсіювання чи поглинання? Моделюємо СР ВЧ α_2 : якщо $\alpha_2 < s(\bar{c}_{i+1})$ — розсіювання, інакше — поглинання і обрив траєкторії;
- 4) Моделюємо новий напрям $\bar{d}_{i+1}$ у т. $\bar{c}_{i+1}$ за попереднім $\bar{d}_i$ згідно (*);
- 5) Повертаємось до кроку (1) з частинкою в $\bar{c}_{i+1}$ і напрямком руху $\bar{d}_{i+1}$.

На модельованих таким чином траєкторіях можна обчислювати різноманітні функціонали на зразок кількості відрізків та загальної довжини. Беручи значення цих функціоналів з ймовірнісними вагами, отримуємо характеристики типу середніх — наближення величин, аналітичне або чисельне знаходження яких складне.

Також можна оцінювати ймовірності подій на зразок «частинка вийде за межі даної області», «частинка увійде в дану область», «частинка буде поглинута у шарі речовини між даними двома областями» тощо.

4.2. Моделювання системи масового обслуговування

Розглянемо просту модель СМО, що складається з n ліній обслуговування. У випадкові моменти часу до системи надходять заявки як пуассонівський потік інтенсивності λ . Чергова заявка обробляється вільною лінією з найменшим номером протягом сталого часу t_3 , у цей час лінія зайнята. Якщо вільних ліній для заявки немає — відмова, ця заявка залишає систему.

Поведінка системи протягом тривалого часу визначає статистичні характеристики, такі як доля відмов, доля зайнятих у середньому ліній тощо.

Скористайтесь тим, що проміжки між послідовними стрибками пуассонівського процесу є Н/З ВВ, показниково розподілені з параметром λ . Знаючи в момент надходження заявки стан кожної лінії, ми можемо вказати вільну лінію (якщо така є), котра прийме заявку, і відповідно змінити її стан. У свою чергу, стан лінії в даний момент описується іншим моментом — коли вона закінчила/закінчить обробляти заявки, які до неї надійшли в минулому.

Алгоритм такий:

0) Ініціалізація: лічильники прийнятих (accepted) та відмінених (cancelled) заявок $A := 0$, $C := 0$; моменти звільнення ліній $T_1 := 0, \dots, T_n := 0$; поточний час системи $t := 0$;

1) Моделюємо інтервал до чергової заявки: $\Delta := -\frac{1}{\lambda} \ln \alpha$, де α — СР ВЧ;

2) Перехід до моменту чергової заявки: $t := t + \Delta$;

3) Кінець? Якщо $t > T$, де T — загальна тривалість моделювання в умовних часових одиницях, воно завершується. Інакше

4) Вибір лінії: перевіряємо по черзі кожну (i -ту) лінію, $i = 1, 2, \dots$, чи закінчила вона обробку своїх заявок — $t \geq T_i$?

4a) якщо $t \geq T_i$, заявку приймає i -та лінія, моментом її звільнення стає $T_i := t + t_3$. Також $A := A + 1$;

4b) якщо вільних ліній немає, $t < T_i$ для усіх i , — відмова, $C := C + 1$;

5) Повертаємось до кроку (1).

Щоб порахувати, скільки часу кожна лінія лишається вільною, введемо лічильники тривалості вільного часу $R_1, \dots, R_n$, ініціалізовані 0, і на кроці (4a) перед зміною T_i на $t + t_3$ додаватимемо до R_i різницю $t - T_i$: це проміжок між звільненням i -ї лінії та поточним моментом прийняття заявки. Наприкінці моделювання додаємо до кожного R_i відповідне $(T - T_i)$, якщо $T > T_i$, — аби врахувати час між останнім звільненням та кінцем періоду моделювання.

Реалізація цього алгоритму на мові програмування **R** наведена далі:

```

modQueue <- function(intens, durat, N, T) {
  nAcc <- 0; nCan <- 0
  fTime <- rep(0, N); relax <- rep(0, N)
  now <- 0
  while (now <= T) {
    delta <- -1 / intens * log(runif(1))
    now <- now + delta
    if (now <= T) {
      acc <- FALSE
      for (i in 1:N) {
        if (now >= fTime[i]) {
          nAcc <- nAcc + 1
          relax[i] <- relax[i] + (now - fTime[i])
          fTime[i] <- now + durat
          acc <- TRUE
          break
        }
      }
      if (acc == FALSE) {
        nCan <- nCan + 1
      }
    }
  }
  for (i in 1:N) {
    relax[i] <- relax[i] + max(0, T - fTime[i])
  }
  return (list(c(nAcc, nCan), relax))
}

```

Результати кількох запусків:

```

>modQueue(2, 0.5, 5, 1000)
[[1]] 1986, 4 [[2]] 495, 705, 863, 954, 991
>modQueue(2, 6, 5, 1000)
[[1]] 743, 1252 [[2]] 80, 94, 101, 138, 145
>modQueue(2, 6, 10, 1000)
[[1]] 1395, 664 [[2]] 79, 92, ..., 253, 314

```

Позначення

ОЗН. — Означення.

РЕС. — Нагадування.

ТВ. — Твердження.

ЗАУВ. — Зауваження.

ПР. — Приклад.

В. — Властивість.

Н. — Наслідок.

ЛЕМА. — Лема.

ТЕОР. — Теорема.

◀ — Початок доведення.

▶ — Кінець доведення.

тоді́дот — тоді й тільки тоді.

⊗ — протиріччя.

Список рекомендованої літератури

- [1] Metropolis N. The Monte Carlo method. [Text] / N. Metropolis, S. Ulam. — J. Amer. Stat. Assoc. — 44 (247). — 1949. — 335–341.
- [2] Ермаков С. М. Статистическое моделирование, 2-е изд., дополн. [Текст] / С. М. Ермаков, Г. А. Михайлов. — М., Наука. — 1982. — 296 с.
- [3] Ермаков С. М. Статистическое моделирование, в 2-х ч. [Текст] / С. М. Ермаков. — СПб., ГУ НИИ ММ. — 2006, 2004. — Ч. I, 63 с.; Ч. II, 46 с.
- [4] Кнут Д. Э. Искусство программирования, том 2. Получисленные методы, 3-е изд. [Текст] : Пер. с англ. / Д. Э. Кнут. — М., ООО «И.Д. Вильямс». — 2007. — 832 с.
- [5] Савелова Т. И. Метод Монте-Карло [Текст] / Т. И. Савелова. — М., НИЯУ МИФИ. — 2011. — 152 с.
- [6] Соболев С. М. Численные методы Монте-Карло [Текст] / С. М. Соболев. — М., Наука. — 1973. — 312 с.

Навчально-методичне видання

Гончаров Сергій Валерійович

КОНСПЕКТ ЛЕКЦІЙ
ІЗ СТАТИСТИЧНОГО МОДЕЛЮВАННЯ

Видання друкується в авторській редакції

Відповідальний редактор *Біла К. О.*
Технічний редактор *Олексенко Н. С.*

Підп. до друку 24.04.18. Формат 60x84 1/16.
Ум. друк. арк. 1,6. Тираж 35 пр. Зам. № 0418-04/3.

Видавець та виготовлювач СПД Біла К. О.

Свідоцтво про внесення до Державного реєстру
суб'єктів видавничої справи ДК № 3618 від 06.11.09

Надруковано на поліграфічній базі видавничо-поліграфічного центру «Імпакт»
Україна, 49000, м. Дніпро, пр. Д. Яворницького, 111, оф. 2
+38 (099) 7805049; +38 (067) 2100256
www.impact.dp.ua
e-mail: impact.dnepr@gmail.com