

BÀI TẬP 2  
(Sinh số ngẫu nhiên)  
THỐNG KÊ MÁY TÍNH VÀ ỨNG DỤNG (CLC)

1. Từ một đồng xu có xác suất ra ngửa là  $p$  ( $0 < p < 1$ ), ta không biết giá trị của  $p$ , tìm cách sinh một biến cố có xác suất là 25%.
2. Từ dãy số giả ngẫu nhiên  $X_1, X_2, \dots, X_N$  được sinh từ bộ sinh đồng dư tuyến tính (LCG) với modulus  $m$ , ta có thể sinh dãy  $U_1, U_2, \dots, U_N$  các số thực thuộc khoảng  $(0, 1)$  bằng cách đặt

$$U_n = \frac{X_n + 1}{m + 1}, n = 1, 2, \dots, N.$$

Xét các dãy số  $U_1, U_2, \dots, U_N$  được sinh từ các phương pháp sau

- (a) LCG:  $m = 2^{35}, a = 2^7, c = 1$ ,
- (b) LCG:  $m = 2^{32}, a = 69069, c = 1327217885$ ,
- (c) LCG:  $m = 2^{31}, a = 2^{16} + 3, c = 0$ ,
- (d) Hàm `numpy.random.uniform` (dãy đối chiếu).

Thực hiện các yêu cầu sau

- (a) Dùng kiểm định Chi-bình phương (Chi-square two-sided test) với mức ý nghĩa 95%, kiểm tra tính phân phối đều của các dãy số trong khoảng  $(0, 1)$ . Lấy  $N = 10^5$  và chia  $(0, 1)$  ra 20 khoảng bằng nhau.
  - (b) Vẽ biểu đồ các cặp  $(U_n, U_{n+1}), n = 1, 2, \dots, N - 1$ . Từ đó đánh giá tính độc lập của các số trong dãy.
3. Cho  $n$  là số nguyên dương, đặt  $U_n = \{1, 2, \dots, n\}$  và  $\mathcal{P}(n) = \{\text{các hoán vị của } U_n\}$ . Tìm cách sinh các **hoán vị ngẫu nhiên** (random permutation) của  $U_n$ , tức là sinh phần tử từ phân phối đều rời rạc trên  $\mathcal{P}(n)$ . Xét các trường hợp
    - (a)  $n = 3$ ,
    - (b)  $n = 10$ ,
    - (c)  $n = 20$ ,
    - (d)  $n = 52$  (một ứng dụng là để "xào bài" tiến lên).
  4. Biến ngẫu nhiên liên tục  $X$  được gọi là có **phân phối Cauchy chuẩn** (standard Cauchy distribution) nếu  $X$  có hàm mật độ xác suất

$$f(x) = \frac{1}{\pi(1 + x^2)}, x \in (-\infty, \infty).$$

Dùng phương pháp biến đổi ngược, sinh các số ngẫu nhiên từ phân phối Cauchy chuẩn.

5. Cho biến ngẫu nhiên liên tục  $X$  với hàm mật độ xác suất

$$f(x) \propto |\sin x|, x \in (-\pi, \pi).$$

Dùng phương pháp lấy mẫu loại bỏ, sinh mẫu cho  $X$ .

6. Đề xuất thuật toán sinh mẫu từ phân phối có hàm mật độ xác suất

$$f(x) = \begin{cases} \frac{x-2}{2} & 2 \leq x < 3, \\ \frac{2-x/3}{2} & 3 \leq x \leq 6, \\ 0 & \text{khác.} \end{cases}$$

*Lưu ý:*

- Các thuật toán cần được trình bày bằng mã giả và cài đặt bằng Python.
- Cần kiểm tra và đánh giá kết quả chạy các thuật toán.
- Ngoại trừ Câu 1 và Câu 2, các thuật toán chỉ được phép dùng nguồn ngẫu nhiên  $\mathcal{U}(0, 1)$  (các số sinh từ hàm `numpy.random.uniform`).

— HẾT —