

BÀI TẬP 3

THỐNG KÊ MÁY TÍNH VÀ ỨNG DỤNG

Câu 1. (4 điểm) Cặp biến ngẫu nhiên X, Y có hàm mật độ xác suất đồng thời

$$f(x, y) \propto e^{-\frac{x^2y^2+x^2+y^2-8x-8y}{2}}.$$

- a) Thiết kế thuật toán lấy mẫu cho (X, Y) bằng các phương pháp:
 - (1) Lấy mẫu độc lập,
 - (2) Metropolis-Hastings (hoặc Metropolis),
 - (3) Lấy mẫu Gibbs.
- b) So sánh các phương pháp ở Câu (a).
- c) Chọn phương pháp tốt nhất ở Câu (a) để ước lượng xác suất $P(X \geq Y)$ và kì vọng $E(XY)$ (có đánh giá sai số).

Câu 2. (3 điểm) Một đồng xu có xác suất ra ngửa là p . Tung đồng xu 10 lần thì thấy 8 lần ra ngửa. Dùng suy diễn Bayes với lần lượt các thông tin tiên nghiệm sau

- a) Không biết thông tin gì cả.
- b) Đồng xu có khả năng cân đối cao (p quanh 0.5).
- c) Đồng xu có khả năng mất cân đối cao (p quanh 0 hoặc 1).
- d) p chỉ có thể nhận 1 trong 3 giá trị: 0.2, 0.7, 0.9.
- e) $0.4 \leq p \leq 0.9$.

để trả lời các câu hỏi:

- (1) Giá trị của p là bao nhiêu?
- (2) p có lớn hơn 0.6 hay không?

Câu 3. (1 điểm) Cho các biến ngẫu nhiên X, Y, Z có phân phối như sau

$$X \sim \mathcal{N}(0, 2^2), Y \sim \mathcal{U}(0, |X|), Z \sim \mathcal{E}(1/Y)$$

trong đó: $\mathcal{N}(\mu, \sigma^2)$ là phân phối chuẩn có kì vọng μ và phương sai σ^2 , $\mathcal{U}(a, b)$ là phân phối đều trên khoảng (a, b) , $\mathcal{E}(\lambda)$ là phân phối mũ có kì vọng $1/\lambda$.

Ước lượng $P(X \geq 3)$ khi quan sát thấy $Z = 4$.

Câu 4. (2 điểm) Tìm hiểu bộ dữ liệu Auto MPG tại

<https://archive.ics.uci.edu/ml/datasets/Auto+MPG>. Download tập tin dữ liệu auto-mpg.data (<https://archive.ics.uci.edu/ml/machine-learning-databases/auto-mpg/auto-mpg.data>). Dùng suy diễn Bayes để so sánh công suất mpg (miles per gallon) của các xe theo vùng sản xuất origin (American, European, Japanese) dùng các mô hình sau

- a) Các nhóm riêng rẽ với phân phối chuẩn.
- b) Các nhóm riêng rẽ với phân phối Student (kháng ngoại lai).
- c) Mô hình phân cấp (“chia sẻ thông tin nhóm”) với phân phối Student.

Lưu ý:

- Trình bày bài làm (lời giải, công thức Toán, mã Python, kết quả, ...) trong tập tin Jupyter Notebook.
- Cần trình bày, cài đặt, kiểm tra và đánh giá các thuật toán ở Câu 1.
- Được phép dùng PyMC để thực hiện các Câu 2, 3, 4.

--- HẾT ---