Một số pp ngẫu nhiên: rời rạc+liên tục

Dùng pp rời rạc để xấp xỉ các pp liên tục.

Định lý luật số lớn.

Cho $X_1, X_2, \cdots, X_n$: độc lập và cùng phân phối với trung bình μ, *phương sai* σ^2

$$\overline{X}_n = \frac{X_1 + X_2 + \cdots + X_n}{n} \sim N\left(\mu, \frac{\sigma^2}{n}\right) : n \text{ đủ } lớn$$

$$S_n = X_1 + X_2 + \cdots + X_n \sim N(n\mu, n\sigma^2)$$

BT: Theo một cuộc khảo sát, thời gian xem tivi trung bình của các bé từ 2 đến 5 tuổi là 25 giờ mỗi tuần. Giả sử thời gian xem tivi có phân phối chuẩn với độ lệch chuẩn là 3 giờ. Chọn ngẫu nhiên 20 em bé từ 2 đến 5 tuổi. Tính xác suất thời gian xem tivi trung bình của 20 bé lớn hơn 26,3 giờ.

Bài làm:

Đặt X_i thời gian xem tv của em bé thứ i có $\mu = 25, \sigma = 3$

$$\overline{X}_{20} = \frac{X_1 + X_2 + \cdots + X_{20}}{20} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

Yêu cầu

$$P(\overline{X}_{20} > 26.3) = P\left(\frac{\overline{X}_{20} - \mu}{\frac{\sigma}{\sqrt{n}}} > \frac{26.3 - \mu}{\frac{\sigma}{\sqrt{n}}}\right) = P\left(\frac{\overline{X}_{20} - \mu}{\frac{\sigma}{\sqrt{n}}} > 1.94\right)$$

$$= 1 - P\left(\frac{\overline{X}_{20} - \mu}{\frac{\sigma}{\sqrt{n}}} \leq 1.94\right) = 1 - \Phi(1.94) = 0.0262$$

BT: Ổ cứng dung lượng 330 megabyte. Cho 300 ảnh: 1 ảnh có dung lượng trung bình 1 megabyte với độ lệch chuẩn là 0.5 megabyte. Xác suất để ổ cứng này có thể lưu trữ được 300 ảnh trên.

Đặt X_i dung lượng ảnh thứ I, có trung bình $\mu = 1, \sigma = 0.5$

Dung lương của 300 ảnh

$$S_{300} = X_1 + X_2 + \cdots + X_{300} \sim N(n\mu, n\sigma^2)$$

Yêu cầu:

$$P(S_{300} \leq 330) = P\left(\frac{S_{300} - 300 * 1}{\frac{0.5}{\sqrt{300}}} \leq \frac{330 - 300 * 1}{\frac{0.5}{\sqrt{300}}}\right)$$

$$= P\left(\frac{S_{300} - 300 * 1}{\frac{0.5}{\sqrt{300}}} \le 3.46\right) = \Phi(3.46) = 0.9997$$

BTQT: Gieo đồng thời 3 con xúc xắc 1000 lần độc lập với nhau. Đặt *Xi* là số số nguyên tố xuất hiện ở mặt trên của 3 con xúc xắc khi gieo lần thứ *i* với $1 \le i \le 1000$. Tính $P(X1 + X2 + \cdots + X1000 > 1550)$.

Các phương pháp xấp xỉ:

K -A	N-k ko A

Chọn ngẫu nhiên n phần từ. X số phần tử có tính chất A trong n phần tử được chọn

$$X \sim H(N, k, n): P(X = x) = \frac{C_k^x C_{N-k}^{n-x}}{C_N^n}$$

Nếu N, k rất lơn hoặc $N \ge 20\, n$

$$X \sim B\left(n, \frac{k}{N}\right): P(X = x) = C_n^x p^x (1-p)^{n-x} \text{ vớip } = \frac{k}{N}$$

Ví dụ : Một vườn lan có 1 000 cây lan sắp nở hoa trong đó có 200 cây hoa màu đỏ. Chọn ngẫu nhiên 20 cây lan. Tính xác suất có 5 cây lan màu đỏ trong 20 cây.

X : số cây lan đỏ được chọn trong 20 cây

$$X \sim H(1000,200,20)$$

$$P(X = 5) = \frac{C_{200}^5 C_{800}^{15}}{C_{1000}^{20}} = 0.1761$$

Nhận thấy $N \ge 20\, n$

$$X \sim B\left(20, \frac{200}{1000}\right)$$

$$P(X = 5) = C_{20}^5 0.2^5 0.8^{15} = 0.1746$$

BT : Một hộp chứa 20 sản phẩm trong đó có 4 sản phẩm có lỗi. Lấy ngẫu nhiên 10 sản phẩm. Tính xác suất có nhiều nhất 2 sản phẩm có lỗi. 10’

24410202, 24410192, 24410178, 24410168, 24410203, 24410146, 24410242, 24410248, 24410157, 24410167, 24410186, 24410152, 24410197, 24410140, 24410238, 24410246, 24410174, 2410247, 24410221, 24410216

X: số sản phẩm lỗi trong 10 sp được chọn

$$X \sim H(20,4,10)$$

$$P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2)$$

$$= \frac{C_{16}^{10}}{C_{20}^{10}} + \frac{C_4^1 C_{16}^5}{C_{20}^{10}} + \frac{C_4^2 C_{16}^4}{C_{20}^{10}} = 0.709$$

Ta thấy N=20, k=4 nhỏ. $N \leq 20\, n$ suy ra ta không dùng được nhị thức để xấp xỉ siêu bội.

$$\begin{aligned} P(X \leq 2) &= P(X=0) + P(X=1) + P(X=2) \\ &= C_{10}^0 0,8^{10} + C_{10}^1 0,2^1 0,8^9 + C_{10}^2 0,2^2 0,8^8 = 0,6778 \end{aligned}$$

Dạng : dùng Poison để xấp xỉ nhị thức

Giả sử $n \geq 30$ và $p \leq 0.05$, khi đó $X \sim B(n,p) \rightarrow X \sim P(np)$

Ví dụ : Trong một dây chuyền sản xuất chip, xác suất một chip bị lỗi là 0,005. Chọn ngẫu nhiên 400 chip để kiểm tra.
a. Tính xác suất có 1 chip bị lỗi.
Đặt X : số chip lỗi trong 400 chip

$$X \sim B(400; 0.05)$$

Yêu cầu

$$P(X = 1) = C_{400}^1 0.05 * 0.95^{399}$$

Cách khác : $n \geq 30, p = 0.05 \leq 0.05$

$$X \sim P(400 * 0.05) = P(2)$$

$$P(X = 1) = e^{-2} 1^1 = 0.271$$

BT : Xác suất một máy tính đã được cài chương trình diệt virus bị nhiễm virus là 0,03. Chọn ngẫu nhiên 200 máy tính đã được cài đặt chương trình diệt virus. Tính xác suất có nhiều nhất 6 máy bị nhiễm virus. 10’

24410247, 24410218, 24410159, 24410246, 24410209, 24410184, 24410163, 24410140, 24410203, 24410216, 24410146, 24410253, 24410191, 24410221, 24410257, 24410212, 24410162

Đặt X: số máy tính bị nhiễm víu trong 200 máy

$$X \sim B(200; 0.03)$$

Nhận thấy $200 \leq 30, p = 0.03 \leq 0.05$

$$X \sim P(200 * 0.03) = P(6)$$

$$P(X \leq 6) = P(X = 0) + P(X = 1) + \cdots + P(X = 6) = 0.606$$

Dạng: xấp xỉ nhị thức bằng pp chuẩn:

Cho $X \sim B(n;p)$. Nếu $pn \geq 5, n(1-p) \geq 5$ thì

$$X \sim N\big(np; np(1-p)\big)$$

Note: hiệu chỉnh

$$P(X = x) = P(x - 0,5 < X < x + 0.5)$$

$$P(X < a) = P(X < a - 0.5)$$

$$P(X > a) = P(X > a + 0.5)$$

$$P(a \leq X \leq b) = P(a - 0,5 \leq X \leq b + 0.5)$$

Ví dụ: Một loại virus máy tính mới tấn công một thư mục bao gồm 1350 tệp. Mỗi tệp bị hỏng với xác suất 0,75 độc lập với các tệp khác

a. Xác suất có 1000 tệp bị hỏng là bao nhiêu?

Đặt X số tệp bị bỏng trong 1350

$$X \sim B(1350; 0.75)$$

Ta thấy $np \geq 5$ $và$ $n(1-p) \geq 5$, như

$$X \sim N\big(np; np(1-p)\big) = N(1012,5; 15.91^2)$$

Yêu cầu :

$$P(X = 1000) = P(1000 - 0.5 < X < 1000 + 0.5)$$

$$= P\left(\frac{999.5 - 1012,5}{15.91} < \frac{X - 1012,5}{15.91} < \frac{1000.5 - 1012,5}{15.91}\right)$$

$$= \Phi\left(\frac{1000.5 - 1012,5}{15.91}\right) - \Phi\left(\frac{999.5 - 1012,5}{15.91}\right) = 0.2266 - 0.2061$$

$$= 0.0205$$

b) Xác suất có từ 1000 đến 1020 tệp bị hỏng là bao nhiêu?

$$P(1000 \leq X \leq 1200) = P(1000 - 0.5 < X < 1200 + 0.5)$$

$$= P\left(\frac{999.5 - 1012,5}{15.91} < \frac{X - 1012,5}{15.91} < \frac{1200.5 - 1012,5}{15.91}\right)$$

$$= \Phi\left(\frac{1200.5 - 1012,5}{15.91}\right) - \Phi\left(\frac{999.5 - 1012,5}{15.91}\right) =?$$

BT : Giả sử xác suất bị lỗi trong một lần giao dịch của một ngân hàng là 0,0003. Nếu có 10000 giao dịch được thực hiện thì xác suất có hơn 6 lỗi là bao nhiêu? (Đáp số: 0,0335)

BT: Theo một cuộc khảo sát, người ta thấy rằng có 25% người lái xe buồn ngủ khi đang lái xe. Chọn ngẫu nhiên 200 người lái xe. Dùng xấp xỉ phân phối chuẩn để tính xác suất có 62 người buồn ngủ khi đang lái xe. (Đáp số: 0,94)

$$10'$$

24410233, 24410174, 24410257, 24410238, 24410216, 24410191, 24410131, 24410162, 24410233, 24410186, 24410257, 24410139, 24410174, 24412018, 24410212, 24410167, 24410203, 24410253, 24410218, 24410163, 24410169, 24410140, 24410241, 24410203

BTQT: Một hộp chứa 10000 viên bi trong đó có 500 bi đỏ. Lấy ngẫu nhiên 10 bi (không thay thế). Xác suất trong 10 bi lấy ra có nhiều nhất 1 bi bỏ là bao nhiêu?

BTQT : Người ta thấy rằng xác suất những người sống đến 65 tuổi ở một thị trấn nọ là 0,8. Chọn ngẫu nhiên 500 người. Dùng xấp xỉ phân phối chuẩn, tính xác suất có từ 375 đến 425 người sống đến 65 tuổi. (Đáp số: 0,9956)

Thường thức : dùng pp chuẩn xấp xỉ pp poisson

Cho biến ngẫu nhiên $X \sim P(\lambda)$ và $\mu = \lambda$, $\sigma^2 = \lambda$.

• Nếu λ đủ lớn thì $X \approx N(\lambda, \lambda)$.

• Nếu $\lambda \geq 20$ thì ta có thể dùng phân phối chuẩn để xấp xỉ cho phân phối Poisson.

Ví dụ : Số lượng người truy cập vào một trang web trên Internet trong một ngày tuân theo phân phối Poisson. Trung bình có 30 người truy cập trong một ngày.

a. Xác suất có ít nhất 25 người truy cập trong một ngày.

$$\begin{aligned} P(X \geq 25) &= P(X > 24,5) \\ &= P\left(\frac{X-30}{\sqrt{30}} > \frac{24,5-30}{\sqrt{30}}\right) \\ &= P(Z > -1) \\ &= 1 - \Phi(-1) = 1 - 0,1587 = 0,8423 \end{aligned}$$

b. Xác suất có nhiều hơn 32 người truy cập trong một ngày.

$$\begin{aligned} P(X > 32) &= P(X > 32,5) \\ &= P\left(\frac{X-30}{\sqrt{30}} > \frac{32,5-30}{\sqrt{30}}\right) \\ &= P(Z > 0,46) \\ &= 1 - \Phi(0,46) = 1 - 0,6772 = 0,3228. \end{aligned}$$