ĐẠI HỌC ĐÀ NẴNG
TRƯỜNG ĐẠI HỌC BÁCH KHOA

TOÁN ỨNG DỤNG CÔNG NGHỆ THÔNG TIN

Khoa Công Nghệ Thông Tin

Chuyên đề xác suất

Biến ngẫu nhiên(1)

- Biến ngẫu nhiên x là một biến số có giá trị ngẫu nhiên thể hiện cho một kết quả của một phép thử.
- Biến ngẫu nhiên gồm 2 loại: rời rạc và liên tục.
- Ví dụ biến ngẫu nhiên rời rạc:
  - mặt đồng xu,
  - mặt chấm của xúc xắc, ...
- Ví dụ biến ngẫu nhiên liên tục:
  - giá một căn nhà ở Đà Nẵng,
  - nhiệt độ ngoài trời, ...

Biến ngẫu nhiên (2)

- **Rời rạc:** hàm phân phối xác suất $p(x)$

$$0 \le p(x) \le 1 \qquad \sum p(x) = 1 \qquad P(X = x) = p(x)$$

- **Liên tục:** hàm mật độ xác suất $f(x)$

$$f(x) \ge 0 \qquad \int f(x)dx = 1 \qquad P(\alpha \le X \le \beta) = \int_{\alpha}^{\beta} f(x)dx$$

- Xác suất ra mặt 6 chấm khi gieo một cục xúc xắc

$$P(X = 6) = p(6) = \frac{1}{6}$$

- Xác suất chiều cao của một nam sinh từ 1m7 đến 1m8 là

$$P(1.7 \le X \le 1.8) = \int_{1.7}^{1.8} f(x)dx$$

Xác suất đồng thời

- Xác suất đồng thời của hai biến ngẫu nhiên x, y:

$$p(x, y)$$

x và y rời rạc: $\sum p(x, y) = 1$

x và y liên tục: $\int p(x, y)dxdy = 1$

x rời rạc, y liên tục: $\sum_x \int p(x, y)dy = 1$

- Điểm thi THPTQG:
  - Biến ngẫu nhiên x $-$ *môn Toán*
  - Biến ngẫu nhiên y – môn Vật lý

$$p(x = 9.4, y = 8.75) > p(x = 9.2, y = 4.5)$$

Xác suất biên

- Xác suất biên là xác suất của từng biến riêng biệt
- Xác suất biên có thể được tính thông qua xác suất đồng thời.
  - Nếu x, y rời rạc: $p(x) = \sum_y p(x, y)$ $\quad p(y) = \sum_x p(x, y)$
  - Nếu x, y liên tục: $p(x) = \int p(x, y)dy$ $\quad p(y) = \int p(x, y)dx$

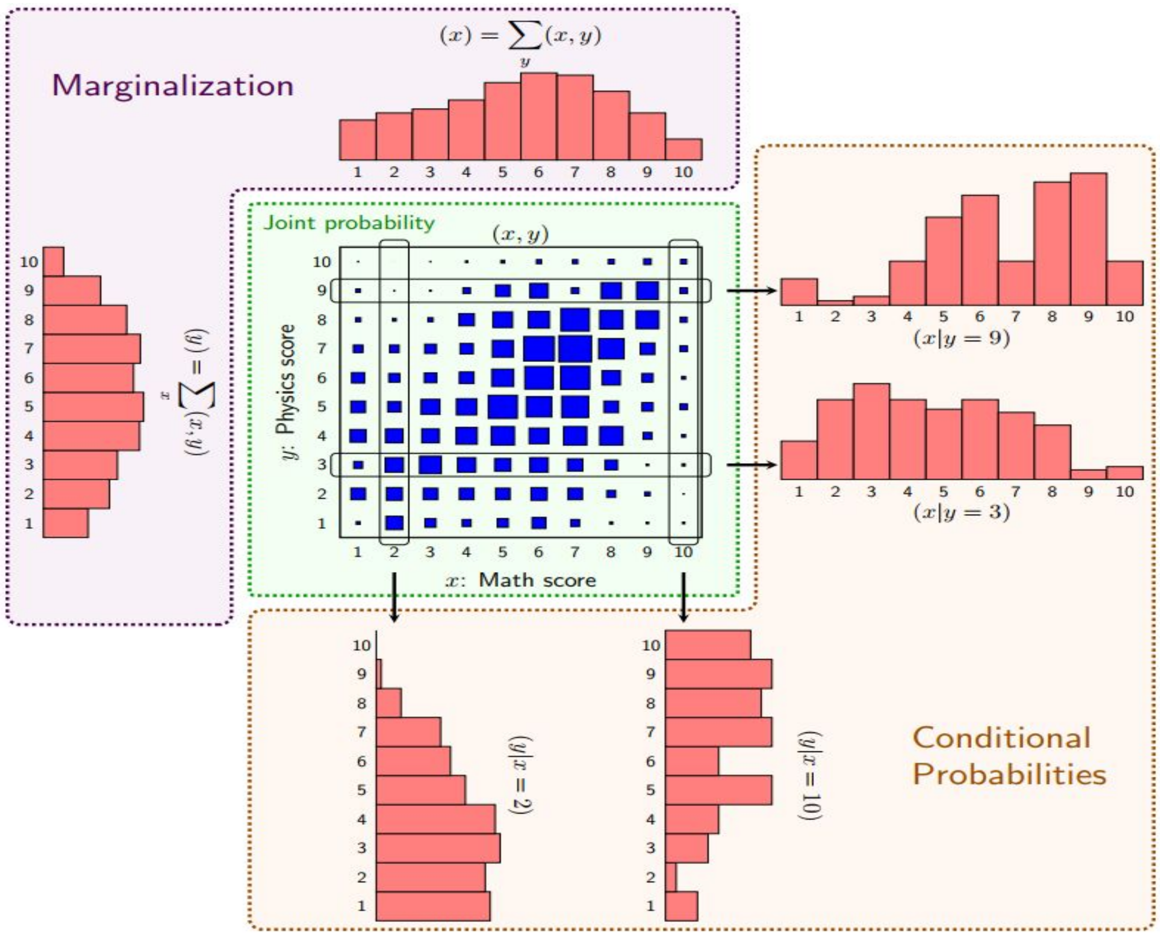
- $p(x, y)$ – xác suất phân phối đồng thời điểm Toán, Lý
  - Xác suất điểm đoánToán: $p(x) = \sum_y p(x, y)$
  - Xác suất điểm Lý: $p(y) = \sum_x p(x, y)$

Xác suất điều kiện (1)

- Xác suất để một biến ngẫu nhiên x nhận giá trị nào đó khi biết biến ngẫu nhiên y có giá trị là y^*: $p(x|y = y^*)$.
  - $p(x|y = y^*) = \frac{p(x,y=y^*)}{\sum_x p(x,y=y^*)} = \frac{p(x,y=y^*)}{p(y=y^*)}$
  - $p(x|y) = \frac{p(x,y)}{p(y)}$
  - $p(x,y) = p(x|y)p(y) = p(y|x)p(x)$

- Xác suất một bạn có điểm thi THPTQG môn toán là 10 với điều kiện điểm Lý là 8.5: $p(10|8.5)$

Xác suất điều kiện (2)



Quy tắc Bayes

$$p(y|x)p(x) = p(x|y)p(y)$$

$$\begin{aligned} p(y|x) &= \frac{p(x|y)p(y)}{p(x)} \\ &= \frac{p(x|y)p(y)}{\sum_y p(x,y)} \\ &= \frac{p(x|y)p(y)}{\sum_y p(x|y)p(y)} \end{aligned}$$

Biến ngẫu nhiên độc lập

- Hai biến ngẫu nhiên x, y độc lập khi chúng không phụ thuộc vào nhau.

$$p(x|y) = p(x)$$
$$p(y|x) = p(y)$$

$$p(x, y) = p(x|y)p(y) = p(x)p(y)$$

- Chiều cao và điểm thi Toán của một học sinh là hai biến ngẫu nhiên độc lập

Kỳ vọng

- Kỳ vọng của một biến ngẫu nhiên x được xác định
  - Nếu x rời rạc: $E[x] = \sum_x xp(x)$
  - Nếu x liên tục: $E[x] = \int xp(x)dx$

- $$\bar{x} = \frac{1}{N}\sum_{n=1}^{N} x_n$$

Phương sai

- Phương sai của một biến ngẫu nhiên X:

$$V(X) = E((X - E(X))^2)$$

$$\sigma^2 = \frac{1}{N}\sum_{n=1}^{N}(x_n - \bar{x})^2$$

Ma trận hiệp phương sai

- Cho vectơ ngẫu nhiên (X, Y). Hiệp phương sai của X, Y được xác định bởi công thức:

$$Cov(X,Y) = E[(X - E(X))(Y - E(Y))]$$

- Ma trận hiệp phương sai:

$$\mathbf{S} = \frac{1}{N}\sum_{n=1}^{N}(\mathbf{x}_n - \bar{\mathbf{x}})(\mathbf{x}_n - \bar{\mathbf{x}})^T = \frac{1}{N}\hat{\mathbf{X}}\hat{\mathbf{X}}^T$$

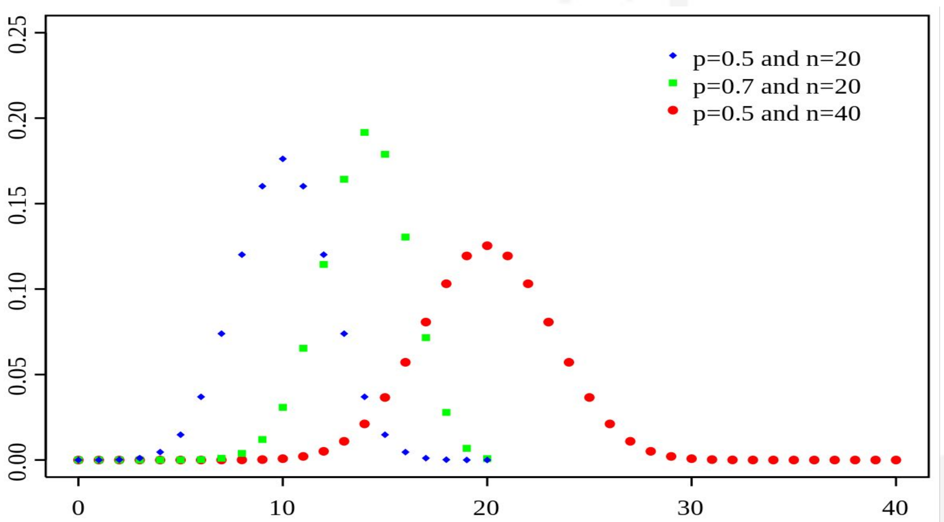
Phân bố rời rạc

- Phân bố Bernoulli
  - Phân bố rời rạc mô tả các biến ngẫu nhiên nhị phân $X \in \{0, 1\}$.

$$X \sim Ber(p), \quad P(X = 1) = p, \quad P(X = 0) = 1 - p$$

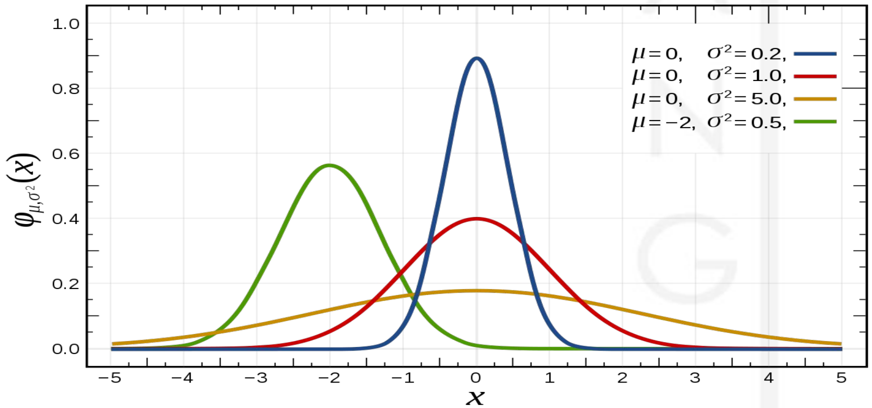
$$E(X) = p, \quad V(X) = p(1 - p)$$

- Minh họa (nhị thức)



Phân bố liên tục

- Phân bố Gaussian
  - Phân bố liên tục với tham số μ và σ, nếu có hàm mật độ xác suất
    - $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, x \in R$
    - $E(X) = \mu, V(X) = \sigma^2$
- Minh họa



Chuỗi Markov

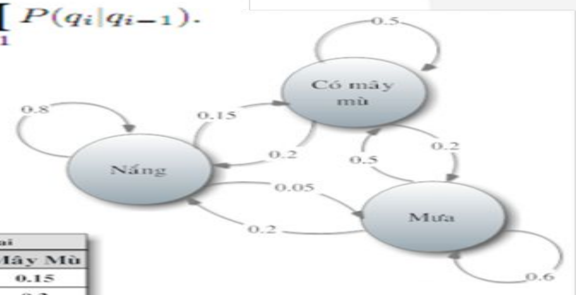
- Dãy các biến cố ngẫu nhiên $X_1, X_2, \ldots, X_n$ thỏa mản tính chất Markov ('tính không gi nhớ')
- Kết quả ở tương lai chỉ dự đoán dựa vào trạng thái hiện tại – quan trọng hơn – dự đoán ấy 'tốt bằng' dự đoán dự trên toàn bộ lịch sử

$$P(X_{n+1} = x \mid X_1 = x_1, \ldots, X_n = x_n) = P(X_{n+1} = x \mid X_n = x_n)$$

- Hôm qua 'mưa', hôm nay 'mây', tính xác suất ngày mai nắng

$$P(q_1, \ldots, q_n) = \prod_{i=1}^{n} P(q_i \mid q_{i-1}).$$

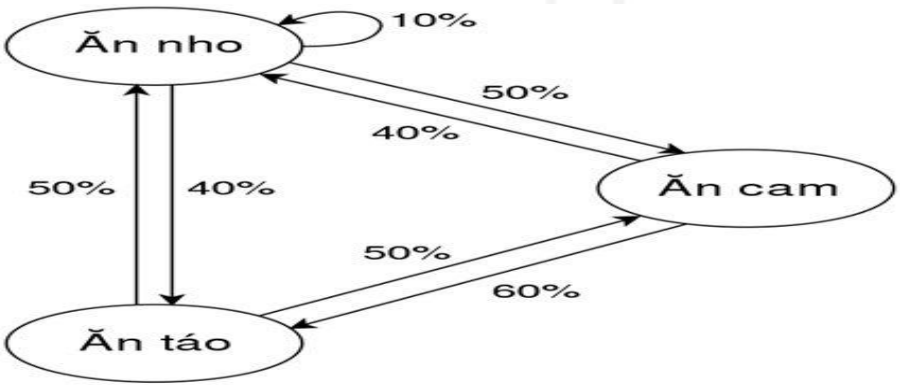
$$P(X_3 = nắng \mid X_2 = mây, X_1 = mưa)$$
$$= P(X_3 = nắng \mid X_2 = mây)$$
$$= 0.2$$



Thời tiết hôm nay	Thời tiết ngày mai		
	Nắng	Mưa	Mây Mù
Nắng	0.8	0.05	0.15
Mưa	0.2	0.6	0.2
Mây Mù	0.2	0.3	0.5

Chuỗi Markov

- Ví dụ chuỗi mô phỏng nho – táo – cam



Chuỗi Markov

- Xác định chính sách thay thế vật tư thiết bị
- Dự báo thất thu cho các hợp đồng thực hiện trước
- cơ sở cho phương pháp mô phỏng ngẫu nhiên xích Markov Monte Carlo,
- Ứng dụng trong thống kê Bayes
- Ứng dụng trong trí tuệ nhân tạo.
- PageRank khởi nguồn cho công cụ tìm kiếm của Google

Ứng dụng dãy Markov

- Mô hình kiểm kê (Inventory Model)
- Mô hình bình Ehrenfest
- Xích Markov trong di truyền
- Mô hình trò chơi hai đấu thủ
- Mô hình phục vụ đám đông (lý thuyết xếp hàng)