Biến cố và xác suất của biến cố (P3)

Giảng viên: Hoàng Thị Điệp
Khoa CNTT – Đại học Công Nghệ

Nội dung

» Phép thử ngẫu nhiên và không gian mẫu
» Biến cố và quan hệ giữa chúng
» Xác suất của một biến cố
» Các qui tắc tính xác suất
» Phép thử lặp – Công thức Becnuli
» Xác suất có điều kiện
» Công thức xác suất đầy đủ
» Công thức Bayes

Công thức Bayes

» Nếu $B_1, B_2, \ldots, B_n$ là một hệ đầy đủ các biến cố và A là 1 biến cố với P(A) > 0 thì với mỗi k=1,2,...,n

$$P(B_k|A) = \frac{P(B_k)P(A|B_k)}{\sum_{i=1}^{n} P(B_i)P(A|B_i)} = \frac{P(B_k)P(A|B_k)}{P(A)} = \frac{P(AB_k)}{P(A)}$$

- Các xác suất $P(B_1),\ldots,P(B_n)$ được gọi là các *xác suất tiên nghiệm* (trước thí nghiệm)
- Các xác suất $P(B_1|A),\ldots,P(B_n|A)$ được gọi là các *xác suất hậu nghiệm* (sau thí nghiệm)

$$P(A) = \sum_{i=1}^{n} P(A|Bi) * P(Bi)$$

Ứng dụng trong thống kê

$$P(H \mid D) = \frac{P(D \mid H) \times P(H)}{P(D)}$$

P(H | D): Xác suất hậu nghiệm (posterior probability)

D: Dữ liệu quan sát được (observed Data)

H: Giả thuyết (Hypothesis)

P(D|H):   Xác suất của dữ liệu với điều kiện giả thuyết H;

hay được gọi là khả năng (likelihood) của giả thuyết H.

P(H): Xác suất tiên nghiệm của giả thuyết, được xác định trước khi quan sát

D

P(D): Xác suất tiên nghiệm của dữ liệu, giống nhau cho mọi giả thuyết

Xét nghiệm covid

- 1% dân cư của thành phố bị covid (99% không bị)
- Nếu bị covid, kiểm tra bằng Kit-V phát hiện 80% (20% không phát hiện được).
- Nếu không bị covid, 9.6% người bị Kit-V trả lời là có covid (90.4% đúng).

	Bị covid (1%)	Không bị (99%)
Kit-V dương tính (positive)	80%	9.6%
Kit-V âm tính (negative)	20%	90.4%

Xét nghiệm covid

	Bị covid (1%)	Không bị (99%)
Kit-V dương tính (positive)	True pos: 1% x 80%	False pos: 99% x 9.6%
Kit-V âm tính (negative)	False neg: 1% x 20%	True neg: 99% x 90.4%

1. Nếu bạn nhận kết quả Kit-V là dương tính, xác suất bạn bị covid là bao nhiêu?
2. Nếu bạn nhận kết quả Kit-V là âm tính, xác suất bạn không bị covid là bao nhiêu?

Bài tập

Nhà máy có 3 phân xưởng X, Y và Z làm ra tương ứng 25%, 35% và 40% tổng sản phẩm. Biết xác suất làm ra sản phẩm hỏng tương ứng của X, Y và Z là 0,01; 0,02 và 0,025. Bạn mua phải 1 sản phẩm hỏng, tính xác suất sản phẩm đó được làm từ:

a) phân xưởng X
b) phân xưởng Y
c) phân xưởng Z

» P(X)=0.25, P(Y)=0.35, P(Z)=0.4
» P(H|X)=0.01, P(H|Y)=0.02, P(H|Z)=0.025
» P(H)?
» P(X|H)=?
» P(Y|H)=?
» P(Z|H)=?

$$P(B_k|A) = \frac{P(B_k)P(A|B_k)}{\sum_{i=1}^{n} P(B_i)P(A|B_i)}$$

Bài tập

- Nhà trường có 3 Khoa ĐT, CH, và CNTT với số sinh viên tương ứng là 20%, 30% và 50%. Xác suất 1 sinh viên không tốt nghiệp đúng hạn từ khoa ĐT, CH, CNTT lần lượt là 25%, 35% và 30%. Biết một sinh viên X không tốt nghiệp đúng hạn, tính xác suất sinh viên đó thuộc khoa CH.

Bài tập

» Có 4 nhóm xạ thủ tập bắn. Nhóm thứ nhất có 5 người, nhóm thứ 2 có 7 người, nhóm thứ 3 có 4 người và nhóm thứ tư có 2 người. Xác suất bắn trúng đích của mỗi người trong nhóm thứ nhất, nhóm thứ 2, nhóm thứ 3 và nhóm thứ tư theo thứ tự là 0.3, 0.4, 0.5 và 0.6. Chọn ngẫu nhiên 1 xạ thủ và xạ thủ này bắn trượt. Hãy xác định xem xạ thủ này có khả năng ở trong nhóm nào nhất.

Bài tập

» Trong số bệnh nhân ở một bệnh viện có 50% điều trị bệnh A, 30% điều trị bệnh B, và 20% điều trị bệnh C. Xác suất để chữa khỏi các bệnh A, B, C trong bệnh viện này tương ứng là 0.7, 0.8 và 0.9. Hãy tính tỉ lệ bệnh nhân được chữa khỏi bệnh A trong tổng số bệnh nhân đã được chữa khỏi bệnh.

- Người ta kiểm tra sức bền của một loại cáp. Nếu cáp bền, máy kiểm tra cho kết quả đúng với xác suất 0.85. Ngược lại, nếu cáp không bền, máy vẫn đánh giá là "bền" với xác suất 0.04. Nếu tỉ lệ cáp bền là 98% và một đoạn cáp chọn ngẫu nhiên bị máy đánh giá là "không bền" thì xác suất đoạn cáp này thực sự không bền là bao nhiêu?

Chuẩn bị bài tới

- Luyện tập các thí dụ/bài tập của Chương 1 giáo trình
- Hoàn thành bài tập gửi qua email