Xác suất thống kê
Không gian Xác suất

Nguyễn Thị Hồng Nhung
nthnhung@hcmus.edu.vn

Khoa Toán-Tin học
Trường Đại học Khoa Học Tự Nhiên Hồ Chí Minh

Ngày 9 tháng 3 năm 2020

Nội dung

Biến cố ngẫu nhiên

Phép thử ngẫu nhiên (Random experiment)

Là sự thực hiện một số điều kiện xác định (một thí nghiệm cụ thể hay quan sát một hiện tượng nào đó), có thể lặp lại nhiều lần và kết quả của phép thử ta không xác định trước được.

Ví dụ 1

Phép thử ngẫu nhiên	*Kết quả*
Tung đồng tiền	*Mặt sấp, mặt ngửa*
Điểm thi kết thúc môn	$\{0, 1, 2, \ldots, 10\}$
Nhóm máu của một người	*A, B, O, AB*

Biến cố ngẫu nhiên

- Tập hợp tất cả các kết quả có thể xảy ra khi thực hiện một phép thử gọi là không gian mẫu hay không gian các biến cố sơ cấp (sample space), ký hiệu Ω.
- Mỗi kết quả của phép thử ngẫu nhiên, $\omega, (\omega \in \Omega)$ gọi là một biến cố/sự kiện sơ cấp (simple event).
- Một tập con của không gian mẫu có nhiều biến cố được gọi là biến cố/sự kiện ngẫu nhiên (event). Dùng các kí tự in hoa để kí hiệu biến cố ngẫu nhiên(sự kiện ngẫu nhiên)$A, B, C, \ldots$
- Biến cố luôn xảy ra khi thực hiện một phép thử gọi là biến cố chắc chắn, ký hiệu Ω.
- Biến cố luôn không xảy ra gọi là biến cố không thể có (empty event), kí hiệu $\emptyset$.

Biến cố ngẫu nhiên

Ví dụ 2

Gieo một lần con xúc xắc. Gọi
$\omega_i =$ *"mặt trên của xúc xắc có i chấm"* $= i$.
Không gian các biến cố sơ cấp

$$\Omega = \{\omega_1, \omega_2, \ldots, \omega_6\} = \{1, 2, \ldots, 6\}$$

$$\begin{array}{llll} A = & \{1,3,5\} = & \text{"chấm lẻ"} & \searrow \\ B = & \{2,4,6\} = & \text{"chấm chẵn"} & \rightarrow \quad \textit{Biến cố ngẫu nhiên} \\ C = & \{5,6\} = & \text{"chấm} > 4\text{"} & \nearrow \end{array}$$

Quan hệ giữa các biến cố I

1. **Sự kéo theo**
   A kéo theo B, ký hiệu $A \subset B$, nếu A xảy ra thì B xảy ra. Ta còn nói A là biến cố thuận lợi cho B.

Ví dụ 3

Tung một con xúc xắc.
Gọi A_i là biến cố được i chấm $(i = \overline{1,6})$,
B là biến cố được số chấm chia hết cho 3,
$C =$ *"số chấm chẵn",*
$\mathbb{P}_2 =$ *"số chấm nguyên tố chẵn".*
Khi đó, ta có $A_2 \subset C, A_3 \subset B, A_2 \subset \mathbb{P}_2, \mathbb{P}_2 \subset A_2$.

Quan hệ giữa các biến cố II

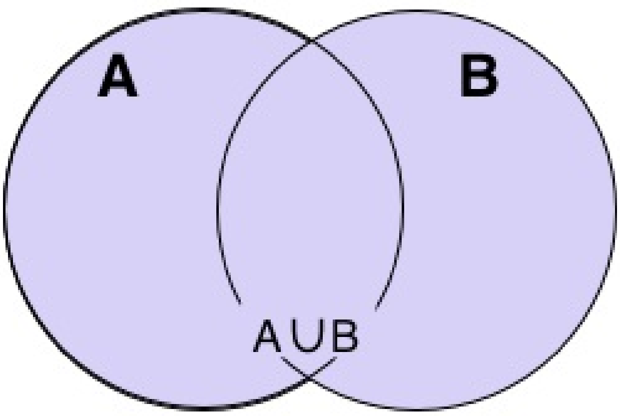
2. **Sự tương đương**
   A tương đương với B, ký hiệu $A = B$, nếu A xảy ra thì B xảy ra và ngược lại.

Ví dụ 4

Trong ví dụ trên $A_2 = \mathbb{P}_2$.

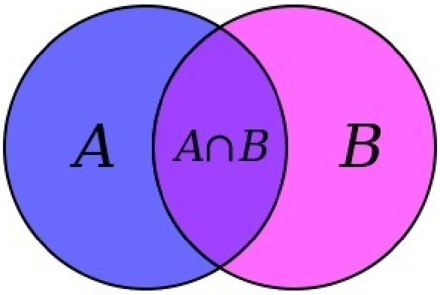
Các phép toán trên biến cố I

1. **Biến cố tổng (union)**
   Biến cố tổng của A và B, ký hiệu $A + B$ hay $A \cup B$ là biến cố xảy ra nếu A hoặc B xảy ra (nghĩa là, có ít nhất một trong hai biến cố xảy ra).



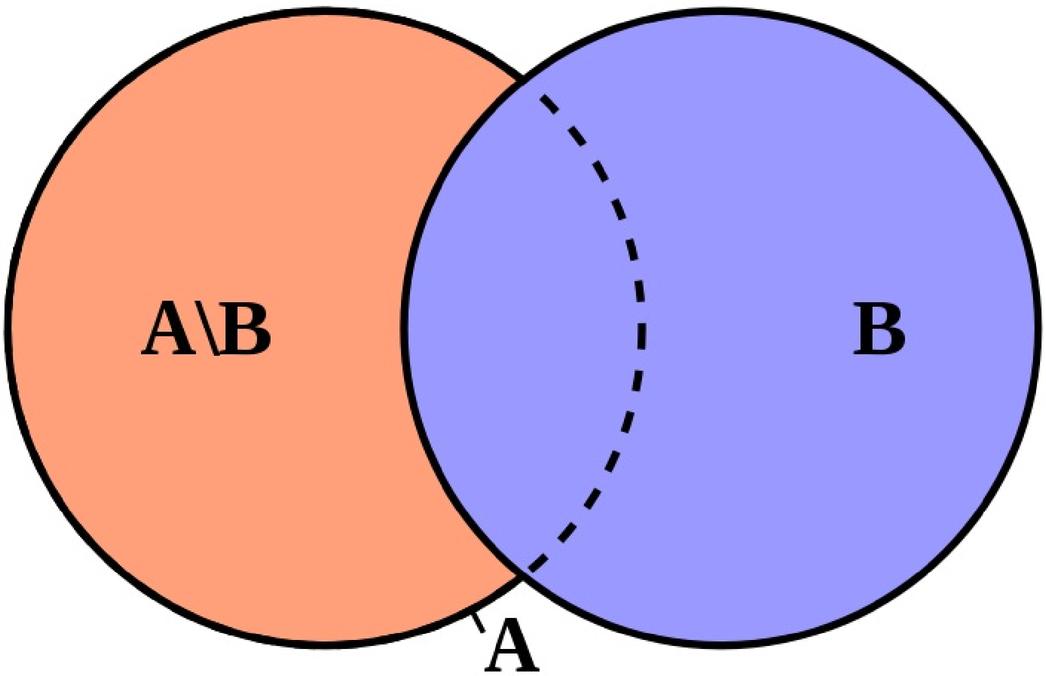
Các phép toán trên biến cố II

2. **Biến cố tích (intersection)**
   Biến cố tích của A và B, ký hiệu AB (hoặc $A \cap B$) là biến cố xảy ra nếu A và B đồng thời xảy ra.



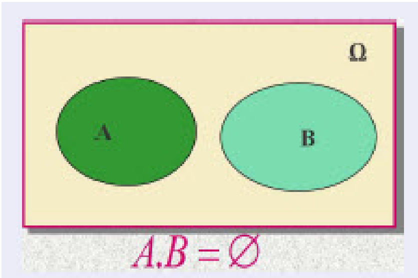
3. **Biến cố hiệu**
   Biến cố hiệu của A và B, ký hiệu $A \setminus B$, là biến cố có được khi biến cố A xảy ra nhưng B không xảy ra.

Các phép toán trên biến cố III



Các phép toán trên biến cố IV

4. **Các biến cố xung khắc (mutually exclusive)**
   A xung khắc với B nếu A và B không đồng thời xảy ra, ký hiệu $AB = \emptyset$.

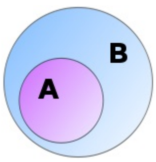


Dãy các biến cố $A_1, A_2, \ldots, A_n$ được gọi là xung khắc từng đôi một nếu $A_iA_j = \emptyset, \ \forall i \neq j$.

Các phép toán trên biến cố V

5. **Biến cố đối lập (Biến cố bù) (complement)**
   Biến cố đối lập của A, ký hiệu $\overline{A}$, là biến cố xảy ra khi A không xảy ra và ngược lại, nghĩa là

$$\begin{cases} A + \overline{A} &= \Omega \\ A\overline{A} &= \emptyset \end{cases} \quad \text{hay } \overline{A} = \Omega \setminus A.$$



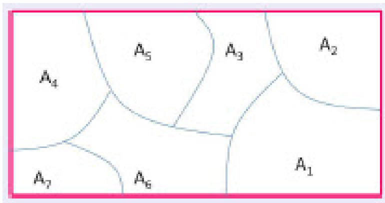
Tính chất

- $\overline{A+B} = \overline{A}.\overline{B}$
- $\overline{AB} = \overline{A} + \overline{B}$

Các phép toán trên biến cố VI

6. **Hệ đầy đủ các biến cố (exhaustive)**
   Dãy n các biến cố $A_1, A_2, \ldots, A_n$ được gọi là một hệ đầy đủ các biến cố nếu:

$$A_iA_j = \emptyset, \ \forall i \neq j, i, j = 1, 2, 3, \ldots, n.$$
$$A_1 + A_2 + \cdots + A_n = \Omega$$



Ví dụ 5

Có 3 bệnh nhân phỏng.
Đặt các biến cố:

$$A_i = \text{"Bệnh nhân } i \text{ tử vong"}, i = 1, 2, 3.$$

Hãy biểu diễn theo A_i các biến cố sau:

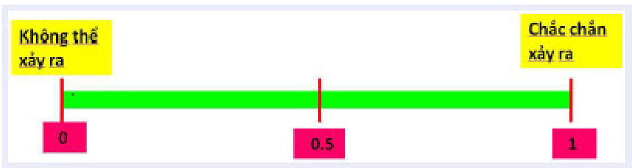
a) B = *"Có không quá hai bệnh nhân tử vong".*
b) C = *"Có ít nhất một bệnh nhân tử vong".*
c) D = *"Có ít nhất hai bệnh nhân tử vong".*
d) E = *"Cả ba bệnh nhân đều sống sót".*

Nội dung

Khái niệm và các định nghĩa về xác suất

Khái niệm về xác suất

Xác suất của biến cố A là môt con số, số đó đặc trưng cho khả năng xuất hiện của biến cố A trong phép thử tương ứng. Ký hiệu là $\mathbb{P}(A)$.



Nhận xét 1

- *$\mathbb{P}(A)$ càng lớn (càng gần 1) thì khả năng xuất hiện A càng nhiều.*
- *$\mathbb{P}(A)$ càng nhỏ (càng gần 0) thì khả năng xuất hiện A càng ít.*

Khái niệm và các định nghĩa về xác suất

Định nghĩa 1 (Định nghĩa xác suất theo quan điểm cổ điển)

Nếu trong một phép thử có tất cả n biến cố sơ cấp đồng khả năng, nghĩa là $\mathbb{P}(\omega_1) = \mathbb{P}(\omega_2) = \cdots = \mathbb{P}(\omega_n) = \frac{1}{n}$*,trong đó có m biến cố thuận lợi cho biến cố A thì xác suất của A, ký hiệu,* $\mathbb{P}(A)$*, là tỉ số* $\frac{m}{n}$.

$$\mathbb{P}(A) = \frac{\text{Số biến cố thuận lợi cho } A}{\text{Số tất cả các biến cố có thể}} = \frac{m}{n}$$

Khái niệm và các định nghĩa về xác suất

Ví dụ 6

Trong một hộp có 3 quả cầu trắng và 5 quả cầu đỏ giống hệt nhau về kích thước. Lấy ngẫu nhiên 3 quả cầu từ hộp đó. Tìm xác suất để được

a) *3 quả cầu đỏ.*

b) *2 quả cầu trắng và 1 quả cầu đỏ.*

Khái niệm và các định nghĩa về xác suất

Ưu và nhược điểm

- Ưu điểm: Tính được chính xác giá trị của xác suất mà không cần tiến hành phép thử.
- Nhược điểm: do đòi hỏi phải có hữu hạn các biến cố và tính đồng khả năng của chúng mà trong thực tế lại có nhiều phép thử không có tính chất đó. Vì vậy, cần đưa ra định nghĩa khác về xác suất để khắc phục những hạn chế trên.

Khái niệm và các định nghĩa về xác suất

Định nghĩa 2 (Định nghĩa xác suất theo quan điểm thống kê)

Thực hiện phép thử n lần. Giả sử biến cố A xuất hiện m lần. Khi đó m là tần số suất hiện biến cố A trong n phép thử, và tỷ số $\frac{m}{n}$ được gọi là tần suất xuất hiện biến cố A trong n phép thử, ký hiệu, $f_n(A) = \frac{m}{n}$. Thực hiện phép thử vô hạn lần, $(n \to \infty)$ tần suất xuất hiện biến cố A tiến về một số xác định gọi là xác suất của biến cố A.

$$\mathbb{P}(A) = \lim_{n\to\infty} f_n(A) = \frac{m}{n}. \tag{1}$$

Khái niệm và các định nghĩa về xác suất

Định nghĩa 3 (Định nghĩa theo quan điểm hình học)

Xét một phép thử đồng khả năng, không gian mẫu có vô hạn phần tử và được biểu diễn thành một miền hình học Ω có độ đo xác định (độ dài, diện tích, thể tích). Biến cố $A \subset \Omega$ được biểu diễn bởi miền hình học A. Khi đó, xác suất xảy ra A được xác định bởi:

$$\mathbb{P}(A) = \frac{\textit{Độ đo của miền } A}{\textit{Độ đo của miền } \Omega} \tag{2}$$

Tính chất của xác suất

1. $0 \leq \mathbb{P}(A) \leq 1$.
2. $\mathbb{P}(\emptyset) = 0, \mathbb{P}(\Omega) = 1$.
3. Nếu $A \subset B$ thì $\mathbb{P}(A) \leq \mathbb{P}(B)$.
4. $\mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A)$.

Nội dung

Công thức cộng xác suất

- A, B là hai sự kiện tùy ý, ta có:

$$\mathbb{P}(A+B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(AB)$$

- $A_1, A_2, \ldots, A_n$ là n sự kiện bất kỳ

$$\begin{aligned}\mathbb{P}\left(\sum_{i=1}^{n} A_i\right) &= \sum_{i=1}^{n}\mathbb{P}(A_i) - \sum_{i<j}^{n}\mathbb{P}(A_iA_j) + \\ &+ \sum_{i<j<k}^{n}\mathbb{P}(A_iA_jA_k) - \cdots + (-1)^{n-1}\mathbb{P}(A_1A_2\ldots A_n)\end{aligned}$$

Công thức cộng xác suất

- Cho A và B là hai biến cố xung khắc ta có :

$$\mathbb{P}(A+B) = \mathbb{P}(A) + \mathbb{P}(B).$$

- $A_1, A_2, \ldots, A_n$ là các sự kiên xung khắc từng đôi một $(A_iA_j = \emptyset \text{ với } i \neq j)$

$$\mathbb{P}\left(\sum_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mathbb{P}(A_i) \tag{3}$$

Công thức cộng xác suất

Ví dụ 7

Trong số 300 sinh viên năm I có 100 sinh viên biết tiếng Anh, 80 sinh viên biết tiếng Pháp, 30 sinh viên biết cả 2 ngoại ngữ Anh-Pháp. Chọn ngẫu nhiên 1 sinh viên năm I. Tính xác suất sinh viên này biết ít nhất 1 ngoại ngữ (Anh hoặc Pháp).

Công thức xác suất có điều kiện

Định nghĩa 4 (Conditional ptobability)

Cho hai biến cố A và B.

- *Xác suất xảy ra biến cố A với điều kiện biến cố B đã xảy ra là*

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(AB)}{\mathbb{P}(B)}, \ \mathbb{P}(B) > 0. \tag{4}$$

- *Tương tự, xác suất xảy ra biến cố B với điều kiện biến cố A đã xảy ra là*

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(AB)}{\mathbb{P}(A)}, \ \mathbb{P}(A) > 0. \tag{5}$$

Công thức xác suất điều kiện

Tính chất của xác suất có điều kiện

1. $0 \leq \mathbb{P}(A|B) \leq 1$.
2. $\mathbb{P}(B|B) = 1$.
3. Nếu $AC = \emptyset$ thì $\mathbb{P}[(A+C)|B] = \mathbb{P}(A|B) + \mathbb{P}(C|B)$.
4. $\mathbb{P}(\bar{A}|B) = 1 - \mathbb{P}(A|B)$.

Công thức xác suất điều kiện

Ví dụ 8

Một bộ bài tây có 52 lá được trộn kỹ. Chọn ngẫu nhiên 1 lá. Biết đã chọn được lá đỏ. Tính xác suất lá đó là lá át cơ.

Ví dụ 9

Một nhóm gồm 300 người trong đó có 200 nam và 100 nữ. Trong 200 nam có 100 người hút thuốc. Trong 100 nữ có 20 người hút thuốc. Chọn ngẫu nhiên một người.

a) *Biết đã chọn được nữ, tính xác suất người đó là người hút thuốc?*

b) *Biết đã chọn được người hút thuốc, tính xác suất người đó là nam?*

Công thức nhân xác suất

Hệ quả 1 (Multiplication rule)

Với các biến cố tùy ý A và B ta có

$$\mathbb{P}(AB) = \mathbb{P}(A|B)\mathbb{P}(B) = \mathbb{P}(B|A)\mathbb{P}(A). \tag{6}$$

Công thức nhân xác suất tổng quát

Cho họ $A_i (i = 1, \dots, n)$ là họ n biến cố, khi đó

$$\mathbb{P}(A_1A_2\dots A_n) = \mathbb{P}(A_1)\mathbb{P}(A_2|A_1)(A_3|A_1A_2)\dots\mathbb{P}(A_n|A_1A_2\dots A_{n-1}). \tag{7}$$

Sự độc lập giữa các biến cố

Ví dụ 10

Có 10 lá thăm, trong đó có 4 lá thăm trúng thưởng. Sinh viên A rút trước, B rút sau.

a) *Hỏi trò chơi có công bằng hay không?*

b) *Nếu B được thưởng, tính xác suất A được thưởng?*

Sự độc lập giữa các biến cố

Hai biến cố độc lập

Hai biến cố A và B được gọi là **độc lập (independent)** với nhau nếu

$$\mathbb{P}(AB) = \mathbb{P}(A).\mathbb{P}(B) \tag{8}$$

Suy ra, nếu A độc lập với B thì

$$\begin{aligned}\mathbb{P}(A|B) = \mathbb{P}(A|\overline{B}) &= \mathbb{P}(A)\\ \mathbb{P}(B|A) = \mathbb{P}(B|\overline{A}) &= \mathbb{P}(B)\end{aligned}$$

Sự độc lập giữa các biến cố

Ví dụ 11

Khảo sát giới tính của những đứa con trong các gia đình có 2 con có độc lập với nhau hay không?

Bài giải 1

Không gian biến cố sơ cấp của phép thử: $\Omega = \{TT, TG, GT, GG\}$
$A =$ *"Con đầu là con trai."* $= \{TT, TG\}$
$B =$ *"Con thứ hai là con gái "* $= \{TG, GG\}$
Ta có:

$$\mathbb{P}(A) = \frac{2}{4} = \frac{1}{2} \text{ và } \mathbb{P}(B) = \frac{2}{4} = \frac{1}{2}$$

và $\mathbb{P}(AB) = \dfrac{1}{4} = \mathbb{P}(A).\mathbb{P}(B)$. *Vậy* A, B *độc lập.*

Sự độc lập giữa các biến cố

n biến cố độc lập

Các biến cố $A_1, A_2, \ldots, A_n$ được gọi là độc lập với nhau nếu chúng thỏa

$$\begin{aligned} \mathbb{P}(A_iA_j) &= \mathbb{P}(A_i)\mathbb{P}(A_j) \\ \mathbb{P}(A_iA_jA_k) &= \mathbb{P}(A_i)\mathbb{P}(A_j)\mathbb{P}(A_k) \\ \mathbb{P}(A_1A_2\ldots A_n) &= \mathbb{P}(A_1)\mathbb{P}(A_2)\ldots\mathbb{P}(A_n) \end{aligned}$$

với mọi tổ hợp chập 2, (i, j), chập 3, $(i, j, k), \ldots$ của n chỉ số.

Chú ý

Sự độc lập từng đôi một không dẫn đến sự độc lập toàn phần.

Sự độc lập giữa các biến cố

Ví dụ 12

Xét phép thử ngẫu nhiên có các kết quả đồng khả năng

ω	ω_1	ω_2	ω_3	ω_4
$\mathbb{P}(\omega)$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$

Đặt $A = \{\omega_1, \omega_4\}$, $B = \{\omega_2, \omega_4\}$, $C = \{\omega_3, \omega_4\}$. *Tính* $\mathbb{P}(AB), \mathbb{P}(AC), \mathbb{P}(BC)$ *và* $\mathbb{P}(ABC)$.

Bài giải 2

Ta có: $\mathbb{P}(A) = \mathbb{P}(B) = \mathbb{P}(C) = \frac{2}{4}$ *và* $\mathbb{P}(AB) = \mathbb{P}(AC) = \mathbb{P}(BC) = \frac{1}{4}$

$$ABC = \{\omega_4\} \Longrightarrow \mathbb{P}(ABC) = \frac{1}{4} \neq \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C).$$

Công thức xác suất đầy đủ

Định nghĩa 5 (Total Probability Rule)

Cho $A_i (i = 1, \ldots, n)$ *là hệ đầy đủ các biến cố và* B *là một biến cố nào đó (trong cùng phép thử) thì*

$$\begin{aligned} \mathbb{P}(B) &= \mathbb{P}(A_1)\mathbb{P}(B|A_1) + \mathbb{P}(A_2)\mathbb{P}(B|A_2) + \cdots + \mathbb{P}(A_n)\mathbb{P}(B|A_n) \\ &= \sum_{i=1}^{n} \mathbb{P}(A_i)\mathbb{P}(B|A_i). \end{aligned} \tag{9}$$

Công thức xác suất đầy đủ

Ví dụ 13

Một nông trường có 4 đội sản xuất. Đội 1 sản xuất 1/3 tổng sản lượng nông sản của nông trường. Đội 2 sản xuất 1/4 tổng sản lượng, đội 3 sản xuất 1/4 tổng sản lượng và đội 4 sản xuất 1/6 tổng sản lượng. Tỉ lệ phế phẩm tương ứng với các đội sản xuất là 0.15; 0.08; 0.05 và 0.01.
Lấy ngẫu nhiên một sản phẩm trong kho của nông trường. Tìm xác suất để lấy phải một phế phẩm.

Công thức xác suất đầy đủ, công thức Bayes

Ví dụ 14

Một đám đông có số đàn ông bằng nửa số phụ nữ. Xác suất để đàn ông bị bệnh tim là 0.06 và phụ nữ là 0.036. Chọn ngẫu nhiên 1 người từ đám đông, tính xác suất để người này bị bệnh tim.

Công thức Bayes

Định nghĩa 6 (Bayes Fomula)

Cho $A_i (i = 1, \ldots, n)$ là hệ đầy đủ các biến cố, B là một biến cố nào đó liên quan đến hệ sao cho $\mathbb{P}(B) > 0$. Khi đó với mọi i

$$\mathbb{P}(A_i|B) = \frac{\mathbb{P}(A_i)\mathbb{P}(B|A_i)}{\mathbb{P}(B)} = \frac{\mathbb{P}(A_i)\mathbb{P}(B|A_i)}{\sum_{i=1}^{n} \mathbb{P}(A_i)\mathbb{P}(B|A_i)} \tag{10}$$

Công thức xác suất đầy đủ, công thức Bayes

Ví dụ 15

Tỷ lệ bệnh B tại một địa phương bằng 0.02. Dùng một phản ứng giúp chuẩn đoán, nếu người bị bệnh thì phản ứng dương tính 95%, nếu người không bị bệnh thì phản ứng dương tính 10%.

a) *Tìm xác suất dương tính của phản ứng.*

b) *Một người làm phản ứng thấy dương tính, tìm xác suất người đó là người bị bệnh.*

c) *Tìm xác suất chuẩn đoán đúng của phản ứng.*