XÁC SUẤT - THỐNG KÊ

CHƯƠNG 2: BIẾN NGẪU NHIÊN, VECTOR NGẪU NHIÊN

TS. Phan Thị Hường

Trường Đại học Bách Khoa TP HCM
Khoa Khoa học ứng dụng, bộ môn Toán ứng dụng
Email: huongphan@hcmut.edu.vn

TP. HCM — 2020.

NỘI DUNG

1. BIẾN NGẪU NHIÊN

NỘI DUNG

1. BIẾN NGẪU NHIÊN
2. BIẾN NGẪU NHIÊN RỜI RẠC

Nội dung

1. Biến ngẫu nhiên
2. Biến ngẫu nhiên rời rạc
3. Biến ngẫu nhiên liên tục

NỘI DUNG

1. BIẾN NGẪU NHIÊN
2. BIẾN NGẪU NHIÊN RỜI RẠC
3. BIẾN NGẪU NHIÊN LIÊN TỤC
4. KỲ VỌNG VÀ PHƯƠNG SAI

NỘI DUNG

1. BIẾN NGẪU NHIÊN
2. BIẾN NGẪU NHIÊN RỜI RẠC
3. BIẾN NGẪU NHIÊN LIÊN TỤC
4. KỲ VỌNG VÀ PHƯƠNG SAI
5. VÉCTO NGẪU NHIÊN RỜI RẠC 2 CHIỀU

Biến ngẫu nhiên

Định nghĩa 1.1 (Biến ngẫu nhiên)

Biến ngẫu nhiên X là một ánh xạ từ không gian các biến cố sơ cấp Ω vào $\mathbb{R}$,

$$\begin{array}{cccccc} X & : & \Omega & \longrightarrow & \mathbb{R} \\ & & \omega & \longmapsto & X = X(\omega) \end{array}$$

Biến ngẫu nhiên

Định nghĩa 1.1 (Biến ngẫu nhiên)

Biến ngẫu nhiên X là một ánh xạ từ không gian các biến cố sơ cấp Ω vào $\mathbb{R}$,

$$\begin{array}{rcccl} X & : & \Omega & \longrightarrow & \mathbb{R} \\ & & \omega & \longmapsto & X = X(\omega) \end{array}$$

Người ta thường dùng các chữ in $X, Y, Z, \ldots$ để ký hiệu các biến ngẫu nhiên và các chữ thường $x, y, z, \ldots$ để chỉ các giá trị của biến ngẫu nhiên.

PHÂN LOẠI ĐẠI LƯỢNG NGẪU NHIÊN

BIẾN NGẪU NHIÊN RỜI RẠC

Biến ngẫu nhiên được gọi là rời rạc nếu tập hợp các giá trị mà nó có thể nhận là một tập hữu hạn hoặc vô hạn đếm được.

Phân loại đại lượng ngẫu nhiên

Biến ngẫu nhiên rời rạc

Biến ngẫu nhiên được gọi là rời rạc nếu tập hợp các giá trị mà nó có thể nhận là một tập hữu hạn hoặc vô hạn đếm được.

Ví dụ 1.1

Các biến ngẫu nhiên sau là rời rạc: điểm số sinh viên, số linh kiện bị lỗi, số ca nhiễm Covid-19, ...

Phân loại đại lượng ngẫu nhiên

Biến ngẫu nhiên liên tục

Biến ngẫu nhiên được gọi là liên tục nếu tập hợp các giá trị mà nó nhận được là một khoảng dạng (a, b)($[a, b)$ hoặc $[a, b]$) hoặc toàn bộ $\mathbb{R}$.

PHÂN LOẠI ĐẠI LƯỢNG NGẪU NHIÊN

BIẾN NGẪU NHIÊN LIÊN TỤC

Biến ngẫu nhiên được gọi là liên tục nếu tập hợp các giá trị mà nó nhận được là một khoảng dạng (a,b)($[a,b)$ hoặc $[a,b]$) hoặc toàn bộ $\mathbb{R}$.

VÍ DỤ 1.2

Các biến ngẫu nhiên sau là biến ngẫu nhiên liên tục: Nhiệt độ không khí ở mỗi thời điểm nào đó, thời gian hoạt động bình thường của một bóng đèn điện, Độ pH của một chất hóa học nào đó.

Quy luật phân phối xác suất

Định nghĩa 2.1

Một hệ thức cho phép biểu diễn mối quan hệ giữa các giá trị có thể có của biến ngẫu nhiên với xác suất nhận các giá trị tương ứng gọi là quy luật phân phối xác suất của biến ngẫu nhiên.

HÀM XÁC SUẤT (PROBABILITY MASS FUNCTION)

ĐỊNH NGHĨA 2.2

Biến ngẫu nhiên rời rạc X *nhận các giá trị có thể* $x_1, x_2, \ldots, x_n, \ldots$ *với xác suất tương ứng là* $\mathbb{P}(X = x_i)$, *ta đặt*

$$f(x) = \begin{cases} \mathbb{P}(X = x) & \textit{khi } x \in \{x_1, \ldots, x_n, \ldots\} \\ 0 & \textit{khi } x \notin \{x_1, \ldots, x_n, \ldots\} \end{cases}$$

gọi là hàm giá trị xác suất biến ngẫu nhiên rời rạc X *nhận giá trị* x, *để đơn giản ta gọi là hàm xác suất.*

HÀM XÁC SUẤT (PROBABILITY MASS FUNCTION)

ĐỊNH NGHĨA 2.2

Biến ngẫu nhiên rời rạc X *nhận các giá trị có thể* $x_1, x_2, \ldots, x_n, \ldots$ *với xác suất tương ứng là* $\mathbb{P}(X = x_i)$, *ta đặt*

$$f(x) = \begin{cases} \mathbb{P}(X = x) & \text{khi } x \in \{x_1, \ldots, x_n, \ldots\} \\ 0 & \text{khi } x \notin \{x_1, \ldots, x_n, \ldots\} \end{cases}$$

gọi là hàm giá trị xác suất biến ngẫu nhiên rời rạc X *nhận giá trị* x, *để đơn giản ta gọi là hàm xác suất.*
Hàm xác suất thường được biểu diễn dư?i dạng bảng:

X	x_1	x_2	$\cdots$	x_n	$\cdots$
$\mathbb{P}$	$f(x_1)$	$f(x_2)$	$\cdots$	$f(x_n)$	$\cdots$

VÍ DỤ 2.1

Giả sử rằng hệ thống máy tính cần truyền đi một tính hiệu gồm 4 *bit. Gọi* X *là số số bit bị lỗi trong* 4 *bit đã truyền đi. Vậy* X *có thể nhận các giá trị* $\{0, 1, 2, 3, 4\}$ *với các xác suất tương ứng:*

X	0	1	2	3	4
$\mathbb{P}$	0.6561	0.2916	0.0486	0.0036	0.0001

TÍNH CHẤT 2.1 (TÍNH CHẤT:)

Hàm số $f(x)$ *là hàm giá trị xác suất của biến ngẫu nhiên rời rạc* X *nếu và chỉ nếu*

1. $f(x_i) \geq 0$ *for all* x_i.
2. $\sum_i f(x_i) = 1$.

Ví dụ 2.2

Giả sử rằng BNN X có hàm xác suất như sau:

$$f(u) = \begin{cases} cu, & u = 1, \cdots, 5 \\ 0, & u \notin \{1,2,3,4,5\} \end{cases}$$

- (A) *Xác định hằng số c.*
- (B) *Tính* $\mathbb{P}(X \in [0,3])$.

Hàm phân phối xác suất (cumulative distribution function)

Định nghĩa 2.3 (Hàm phân phối xác suất)

Hàm phân phối xác suất của biến ngẫu nhiên X (xác định trên không gian các biến cố sơ cấp Ω) là hàm $F(x)$ được định nghĩa

$$F(x) = \mathbb{P}(X \leq x) \tag{1}$$

với mọi $x \in (-\infty, +\infty)$.

HÀM PHÂN PHỐI XÁC SUẤT (CUMULATIVE DISTRIBUTION FUNCTION)

TÍNH CHẤT 2.2

Hàm phân phối xác suất $F(x)$ có các tính chất cơ bản sau

1. *Hàm phân phối là hàm không giảm.*
2. *Liên tục phải, có giới hạn bên trái tại mọi điểm.*
3. $F(-\infty) = \lim\limits_{x\to-\infty} F(x) = 0, \quad F(+\infty) = \lim\limits_{x\to+\infty} F(x) = 1.$
4. $\mathbb{P}(a < X \leq b) = F(b) - F(a)$ *với mọi* $a, b \in \mathbb{R}$ *và* $a \leq b$.
5. *Cho* $I \subset \mathbb{R}$ *thì*

$$\mathbb{P}(X \in I) = \sum_{x_i \in I} \mathbb{P}(X = x_i) = \sum_{x_i \in I} f(x_i)$$

HÀM PHÂN PHỐI XÁC SUẤT (CUMULATIVE DISTRIBUTION FUNCTION)

VÍ DỤ 2.3

Cho biến ngẫu nhiên rời rạc X có hàm phân phối xác suất sau:

$$F(x) = \begin{cases} 0, & x < -2 \\ 0.2, & -2 \leq x < 0 \\ 0.7, & 0 \leq x < 2 \\ 1, & 2 \leq x \end{cases}$$

- (A) *Tính* $F(2)$.
- (B) *Tinhs* $\mathbb{P}(-2 < X \leq 2)$.
- (C) *Tính* $\mathbb{P}(1.5 \leq X \leq 3)$.

HÀM CỦA BIẾN NGẪU NHIÊN RỜI RẠC

VÍ DỤ 2.4

Cho biến ngẫu nhiên rời rạc X có bảng phân phối xác suất như sau:

X	-1	0	1	2
$\mathbb{P}$	0.1	0.3	0.4	0.2

- (A) *Lập bảng phân phối xác suất cho biến ngẫu nhiên* $Y = 2X + 3$.
- (B) *Lập bảng phân phối xác suất cho biến ngẫu nhiên* $Z = X^2$.

HÀM MẬT ĐỘ XÁC SUẤT (PROBABILITY DENSISTY FUNCTION)

ĐỊNH NGHĨA 3.1

Cho biến ngẫu nhiên liên tục X, hàm mật độ xác suất của biến X là hàm số $f(x)$ thoả các tính chất sau:

I)

$$f(x) \geq 0$$

II)

$$\int_{-\infty}^{\infty} f(x)dx = 1$$

Chú ý:

Chú ý:

1. Mọi hàm $f(x)$ không âm, và thỏa điều kiện $\int\limits_{-\infty}^{\infty} f(x)dx = 1$ đều là hàm mật độ của 1 biến ngẫu nhiên X nào đó.
2. $P(a \le X \le b) = \int_a^b f(x)dx =$
diện tích miền dưới hàm số f(x) từ a đến b
3. $P(a \le X \le b) = P(a < X \le b) = P(a \le X < b) = P(a < X < b) = \int_a^b f(x)dx$

HÀM MẬT ĐỘ XÁC SUẤT

VÍ DỤ 3.1

Cho một biến ngẫu nhiên liên tục X là độ đường kính của một lỗ khoan trên một miếng kim loại. Đường kính mục tiêu là $12.5(mm)$, nhưng thực tế đường kính này thường lớn hơn đường kính mục tiêu. Và người ta ghi nhận được rằng biến ngẫu nhiên X sẽ có hàm mật độ xác suất như sau:

$$f(x) = 20e^{-20(x-12.5)},\ cho\ \ x \geq 12.5$$

Những lỗ khoan mà có đường kính > 12.6 là không đạt yêu cầu và sẽ bị loại bỏ.

- (A) *Hãy tính tỉ lệ lỗ khoan bị loại.*
- (B) *Hãy tính tỉ lệ các lỗ khoan có đường kính trong khoảng $12.5 < X < 12.6$.*

Hàm phân phân phối xác suất (Cumulative Distribution Function)

Định nghĩa 3.2

Hàm phân phối xác suất của biến ngẫu nhiên liên tục X được định nghĩa bởi:

$$F(x) = P(X \le x) = \int_{-\infty}^{x} f(u)\,du$$

với mọi $x \in \mathbb{R}$.

Hàm phân phân phối xác suất (Cumulative Distribution Function)

Định nghĩa 3.2

Hàm phân phối xác suất của biến ngẫu nhiên liên tục X được định nghĩa bởi:

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(u)\,du$$

với mọi $x \in \mathbb{R}$.

Chú ý:

Hàm phân phân phối xác suất (Cumulative Distribution Function)

Định nghĩa 3.2

Hàm phân phối xác suất của biến ngẫu nhiên liên tục X được định nghĩa bởi:

$$F(x) = P(X \le x) = \int_{-\infty}^{x} f(u)\,du$$

với mọi $x \in \mathbb{R}$.

Chú ý:

1. $P(a \le X \le b) = P(a < X \le b) = P(a \le X < b) = P(a < X < b) = F(b) - F(a)$
2. $F'(x) = \frac{d}{dx}F(x) = f(x)$

HÀM PHÂN PHÂN PHỐI XÁC SUẤT

VÍ DỤ 3.2

Tuổi thọ Y của một thiết bị (đơn vị: giờ) có hàm phân phối xác suất có dạng

$$F(x) = \begin{cases} \frac{a}{100} - \frac{a}{x} & \text{nếu } x \geq 100 \\ 0 & \text{nếu } x < 100 \end{cases}$$

với $a \in \mathbb{R}$.

- (A) *Hãy xác đinh giá trị của hằng số a để hàm $F(x)$ xác định.*
- (B) *Thiết bị được gọi là loại A nếu tuổi thọ của nó kéo dài ít nhất 400 giờ. Tính tỉ lệ loại A.*

KỲ VỌNG

ĐỊNH NGHĨA 4.1 (KỲ VỌNG CỦA BIẾN NGẪU NHIÊN RỜI RẠC)

Giả sử BNN rời rạc X có hàm xác suất $f(x)$ hoặc có bảng phân phối xác suất như sau:

X	x_1	x_2	x_3	x_4	...	x_n	...
$\mathbb{P}$	$f(x_1)$	$f(x_2)$	$f(x_3)$	$f(x_4)$	...	$f(x_n)$	...

Kỳ vọng

Định nghĩa 4.1 (Kỳ vọng của biến ngẫu nhiên rời rạc)

Giả sử BNN rời rạc X có hàm xác suất $f(x)$ hoặc có bảng phân phối xác suất như sau:

X	x_1	x_2	x_3	x_4	...	x_n	...
$\mathbb{P}$	$f(x_1)$	$f(x_2)$	$f(x_3)$	$f(x_4)$	...	$f(x_n)$	...

Kỳ vọng (Expectation) của X được định nghĩa như sau

$$\mathbb{E}(X) = \sum_{i=1}^{\infty} x_i \mathbb{P}(X = x_i) = \sum_{i=1}^{\infty} x_i f(x_i)$$

Kỳ vọng (Expectation) của $h(X)$ được định nghĩa như sau

$$\mathbb{E}(h(X)) = \sum_{i=1}^{\infty} h(x_i)\mathbb{P}(X = x_i) = \sum_{i=1}^{\infty} h(x_i) f(x_i)$$

KỲ VỌNG

ĐỊNH NGHĨA 4.2 (KỲ VỌNG CỦA BIẾN NGẪU NHIÊN LIÊN TỤC)

Giả sử BNN liên tục X có hàm mật độ xác suất $f(x)$, kỳ vọng của X là

$$\mathbb{E}(X) = \int_{-\infty}^{+\infty} x f(x) dx \tag{2}$$

Kỳ vọng

Định nghĩa 4.2 (Kỳ vọng của biến ngẫu nhiên liên tục)

Giả sử BNN liên tục X có hàm mật độ xác suất $f(x)$, kỳ vọng của X là

$$\mathbb{E}(X) = \int_{-\infty}^{+\infty} x f(x) dx \tag{2}$$

Kì vọng của biến ngẫu nhiên liên tục $h(X)$ được định nghĩa bởi

$$\mathbb{E}(h(X)) = \int_{-\infty}^{+\infty} h(x) f(x) dx \tag{3}$$

Kỳ vọng

Tính chất của kỳ vọng

Cho X, Y là hai biến ngẫu nhiên bất kỳ và $C \in \mathbb{R}$ thì kỳ vọng của biến ngẫu nhiên có các tính chất sau

i) $\mathbb{E}(C) = C$.

ii) $\mathbb{E}(CX) = C\mathbb{E}(X)$.

iii) $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$.

iv) Nếu hai biến ngẫu nhiên X và Y độc lập thì $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$.

KỲ VỌNG

TÍNH CHẤT CỦA KỲ VỌNG

Cho X, Y là hai biến ngẫu nhiên bất kỳ và $C \in \mathbb{R}$ thì kỳ vọng của biến ngẫu nhiên có các tính chất sau

I) $\mathbb{E}(C) = C$.

II) $\mathbb{E}(CX) = C\mathbb{E}(X)$.

III) $\mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y)$.

IV) Nếu hai biến ngẫu nhiên X và Y độc lập thì $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$.

Ý NGHĨA CỦA KỲ VỌNG

- Là giá trị trung bình theo xác suất của tất cả các giá trị có thể có của biến ngẫu nhiên.
- Kỳ vọng phản ánh giá trị trung tâm của phân phối xác suất.

Ví dụ 4.1

Khảo sát khả năng tính toán của trẻ em ở một trường mầm non với 5 *câu hỏi, người ta ghi nhận số câu trả lời đúng* (X) *và tỷ lệ tương ứng* $(\mathbb{P}(X))$ *như trong bảng số liệu sau*

X	0	1	2	3	4	5
$\mathbb{P}(X)$	0.05	0.05	0.1	0.2	0.2	0.4

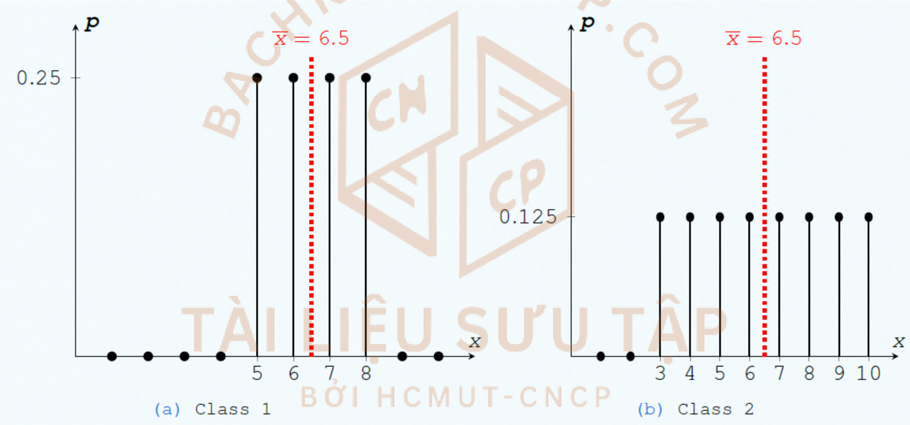
Hãy tính trung bình $\mathbb{E}(X)$ *và phương sai* $\mathbb{V}(X)$ *cho số câu trả lời đúng của các bé.*

Ví dụ 4.2

Người ta nhận thấy cân nặng (kg) của những bưu kiện được gửi tại một bưu điện là một đại lượng ngẫu nhiên có hàm mật độ xác suất $f(x) = \dfrac{80}{79x^2}$ *với* $1 < x < 80$ *và bằng 0 với các giá trị khác của* x*. Biết rằng chi phí gởi một bưu kiện bao gồm phí dịch vụ và phí bưu kiện. Biết rằng phí dịch vụ là* 1.5 *đô la và phí bưu kiện là* 2 *đô la/kg.*

- (A) *Tính trung bình cân nặng của các gói bưu kiện được gửi ở bưu điện này.*
- (B) *Tính trung bình chi phí để gửi một bưu kiện ở bưu điện này.*
- (C) *Tính độ lệch chuẩn trong cân nặng của các gói bưu kiện được gửi ở bưu điện này.*
- (D) *Tính độ lệch chuẩn trong chi phí gửi một bưu kiện ở bưu điện này.*

PHƯƠNG SAI


p
$\overline{x} = 6.5$
0.25
x
5
6
7
8
(a) Class 1
p
$\overline{x} = 6.5$
0.125
x
3
4
5
6
7
8
9
10
(b) Class 2

PHƯƠNG SAI

ĐỊNH NGHĨA 4.3 (PHƯƠNG SAI)

Nếu biến ngẫu nhiên X có kỳ vọng $\mathbb{E}(X)$ thì phương sai, ký hiệu $\mathbb{V}(X)$, được định nghĩa

$$\mathbb{V}(X) = \mathbb{E}(X - \mathbb{E}(X))^2 = \mathbb{E}(X^2) - \mathbb{E}(X)^2 \tag{4}$$

Phương sai

Định nghĩa 4.3 (Phương sai)

Nếu biến ngẫu nhiên X có kỳ vọng $\mathbb{E}(X)$ thì phương sai, ký hiệu $\mathbb{V}(X)$, được định nghĩa

$$\mathbb{V}(X) = \mathbb{E}(X - \mathbb{E}(X))^2 = \mathbb{E}(X^2) - \mathbb{E}(X)^2 \tag{4}$$

Định nghĩa 4.4 (Độ lệch chuẩn)

Độ lệch chuẩn của biến ngẫu nhiên X, ký hiệu $\sigma(X)$, là căn bậc hai của $\mathbb{V}(X)$.

$$\sigma(X) = \sqrt{\mathbb{V}(X)}$$

PHƯƠNG SAI

NHẬN XÉT 4.1

Nếu X là BNN rời rạc thì phương sai của X là:

$$\mathbb{V}(X) = \sum_{i=1}^{\infty} (x_i - \mathbb{E}(X))^2 f(x_i) = \left(\sum_{i=1}^{\infty} x_i^2 f(x_i) \right) - \mathbb{E}(X)^2$$

Nếu X là biến ngẫu nhiên liên tục thì phương sai của X là:

$$\mathbb{V}(X) = \int_{-\infty}^{+\infty} (x - \mathbb{E}(X))^2 f(x)dx = \int_{-\infty}^{+\infty} x^2 f(x)dx - \mathbb{E}(X)^2$$

PHƯƠNG SAI

TÍNH CHẤT PHƯƠNG SAI

Cho hai biến ngẫu nhiên X, Y và hằng số thực $C \in \mathbb{R}$, phương sai có các tính chất sau

- I) $\mathbb{V}(C) = 0$.
- II) $\mathbb{V}(CX) = C^2\mathbb{V}(X)$.
- III) Nếu X và Y độc lập thì $\mathbb{V}(X+Y) = \mathbb{V}(X) + \mathbb{V}(Y)$.

PHƯƠNG SAI

Ý NGHĨA CỦA PHƯƠNG SAI

- Phương sai là kỳ vọng của bình phương các sai lệch giữa X và $\mathbb{E}(X)$, nói cách khác phương sai là trung bình bình phương sai lệch, nó phản ánh mức độ phân tán các giá trị của biến ngẫu nhiên xung quanh giá trị trung bình.
- Trong công nghiệp phương sai biểu thị độ chính xác trong sản xuất. Trong canh tác, phương sai biểu thị mức độ ổn định của năng suất...

VÉCTƠ NGẪU NHIÊN RỜI RẠC 2 CHIỀU

ĐỊNH NGHĨA 5.1 (JOINT PROBABILITY MASS FUNCTION)

Hàm xác suất đồng thời của hai biến ngẫu nhiên X và Y, kí hiệu bởi $f_{XY}(x,y)$ là hàm thoả mãn,

1. $f_{XY}(x,y) \geq 0$
2. $\sum_X \sum_Y f_{XY}(x,y) = 1$
3. $f_{XY}(x,y) = P(X = x, Y = y)$

VÉCTƠ NGẪU NHIÊN RỜI RẠC 2 CHIỀU

NHẬN XÉT 5.1

Bảng phân phối xác suất của hai biến ngẫu nhiên X và Y được biểu diễn bởi

Y/X	x_1	x_2	$\ldots$	x_n	*Tổng dòng*
y_1	$\mathbb{P}(X = x_1, Y = y_1)$	$\mathbb{P}(X = x_2, Y = y_1)$	$\ldots$	$\mathbb{P}(X = x_n, Y = y_1)$	$\mathbb{P}(X = x., Y = y_1)$
y_2	$\mathbb{P}(X = x_1, Y = y_2)$	$\mathbb{P}(X = x_2, Y = y_2)$	$\ldots$	$\mathbb{P}(X = x_n, Y = y_n)$	$\mathbb{P}(X = x., Y = y_2)$
$\ldots$					
y_n	$\mathbb{P}(X = x_1, Y = y_n)$	$\mathbb{P}(X = x_2, Y = y_n)$	$\ldots$	$\mathbb{P}(X = x_n, Y = y_n)$	$\mathbb{P}(X = x., Y = y_n)$
Tổng cột	$\mathbb{P}(X = x_1, Y = y.)$	$\mathbb{P}(X = x_2, Y = y.)$	$\ldots$	$\mathbb{P}(X = xn, Y = y.)$	*1*

VÉCTƠ NGẪU NHIÊN RỜI RẠC 2 CHIỀU

ĐỊNH NGHĨA 5.2 (HÀM XÁC SUẤT LỀ (MARGINAL PROBABILITY DISTRIBUTION))

Hai BNN X và Y có hàm xác suất đồng thời $f_{XY}(x,y)$, thì hàm xác suất lề của X và Y lần lượt là

$$f_X(x) = P(X = x) = \sum_Y f_{XY}(x,y)$$

và

$$f_Y(y) = P(Y = y) = \sum_X f_{XY}(x,y)$$

ĐỊNH NGHĨA 5.3

Xét vector ngẫu nhiên (X, Y), *nếu* X *có hàm xác suất lề* $f_X(x)$ *thì*

$$\mathbb{E}(X) = \sum_{i=1}^{\infty} x_i f_X(x_i)$$

Và

$$\mathbb{V}(X) = \sum_{i=1}^{\infty} (x_i - \mathbb{E}(X))^2 f_X(x_i) = \left(\sum_{i=1}^{\infty} x_i^2 f_X(x_i)\right) - \mathbb{E}(X)^2$$

Tương tự cho Y.

ĐỊNH NGHĨA 5.4

Xét vector ngẫu nhiên (X,Y)*, có hàm xác suất đồng thời* $f_{XY}(x,y)$ *thì*

$$\mathbb{E}(XY) = \sum_i \sum_j x_i y_j f_{XY}(x_i, y_j)$$

và

$$\mathbb{E}(h(X,Y)) = \sum_i \sum_j h(x_i, y_j) f_{XY}(x_i, y_j)$$

Ví dụ 5.1

Cho hai biến ngẫu nhiên X và Y có bảng phân phối xác suất đồng thời như sau:

	$X = x$		
$Y = y$	1	2	3
0	0.1	0.15	a
1	0.2	b	0.4

(A) *Tìm a và b, biết rằng* $\mathbb{E}(Y) = 0.7$.

(B) *Tính* $\mathbb{E}(XY)$.

VÉCTƠ NGẪU NHIÊN RỜI RẠC 2 CHIỀU

ĐỊNH NGHĨA 5.5 (PHÂN PHỐI XÁC SUẤT CÓ ĐIỀU KIỆN (CONDITIONAL PROBABILITY DISTRIBUTION))

Hai biến ngẫu nhiên X và Y có hàm xác suất đồng thời $f_{XY}(x,y)$, các hàm xác xuất có điều kiện được định nghĩa bởi

$$f_{X|y_i} = P(X = x|Y = y_i) = \frac{P(X = x, Y = y_i)}{P(Y = y_i)} = \frac{f_{XY}(x, y_i)}{f_Y(y_i)}$$

và

$$f_{Y|x_i} = P(Y = y|X = x_i) = \frac{P(X = x_i, Y = y)}{P(X = x_i)} = \frac{f_{XY}(x_i, y)}{f_X(x_i)}$$

VÉCTƠ NGẪU NHIÊN RỜI RẠC 2 CHIỀU

ĐỊNH NGHĨA 5.6 (SỰ ĐỘC LẬP GIỮA HAI BIẾN NGẪU NHIÊN)

Cho hai biến ngẫu nhiên X và Y, X và Y là độc lập nếu thoả mãn một trong các điều kiện sau:

1. $f_{XY}(x,y) = f_X(x)f_Y(y)$ *cho mọi x và y.*
2. $f_{Y|x}(y) = f_Y(y)$ *cho mọi giá trị x và y với* $f_Y(y) > 0$.
3. $f_{X|y}(x) = f_X(x)$ *cho mọi giá trị x và y với* $f_X(x) > 0$.

VÉCTƠ NGẪU NHIÊN RỜI RẠC 2 CHIỀU

VÍ DỤ 5.2

Thời gian phản hồi là tốc độ tải trang và nó rất quan trọng đối với một trang web di động. Khi thời gian phản hồi tăng lên, khách hàng trở sẽ cảm thấy khó chịu và có khả năng từ bỏ trang web. Gọi X là số thanh tín hiệu (độ mạnh của đường truyền) và gọi Y là thời gian tải trang (giây) cho một người dùng và trang web cụ thể. Bảng phân phối xác suất đồng thời của X và Y như sau:

	x = Number of Bars of Signal Strength		
y = Response time (nearest second)	1	2	3
4	0.15	0.1	0.05
3	0.02	0.1	0.05
2	0.02	0.03	0.2
1	0.01	0.02	0.25

- A. *Tìm các hàm xác suất lề của X và Y.*
- B. *Tìm các hàm xác suất có điều kiện của X và Y.*
- C. *Hỏi X và Y có độc lập.*

BÀI TẬP 5.1

Cho biến ngẫu nhiên rời rạc X có bảng phân phối xác suất

X	a	0.1	0.3	0.4	2
$\mathbb{P}$	0.3	0.2	0.2	b	0.1

- (A) *Tìm a và b biết* $\mathbb{E}(X) = 0.3$.
- (B) *Tại giá trị a, b vừa tìm được, hãy tính phương sai và độ lệch chuẩn của X.*
- (C) *Tìm hàm phân phối xác suất của X.*

BÀI TẬP 5.2

Tuổi thọ X (năm) của người dân ở một địa phương là biến ngẫu nhiên có hàm phân phối xác suất như sau

$$F(x) = \begin{cases} 1 - e^{-\lambda x}, & khi\ x > 0, \\ 0, & khi\ x \leq 0, \end{cases}$$

với $\lambda = 0.013$.

- (A) *Tính tỷ lệ người dân thọ từ* 60 *đến* 70 *tuổi.*
- (B) *Hàm mật độ của X.*
- (C) *Tính tuổi thọ trung bình và tính độ lệch chuẩn của X.*