BIẾN NGẪU NHIÊN

Nguyễn Văn Thìn

BỘ MÔN THỐNG KÊ TOÁN HỌC
KHOA TOÁN - TIN HỌC
ĐẠI HỌC KHOA HỌC TỰ NHIÊN TP.HCM

Tháng 2 năm 2016

BIẾN NGẪU NHIÊN
Nguyễn Văn Thìn
Định nghĩa
Phân loại
Phân phối xác suất
Bảng phân phối xác suất
Hàm phân phối xác suất
Hàm mật độ xác suất
Các tham số đặc trưng

Outline

BIẾN NGẪU NHIÊN
Nguyễn Văn Thìn
Định nghĩa
Phân loại
Phân phối xác suất
Bảng phân phối xác suất
Hàm phân phối xác suất
Hàm mật độ xác suất
Các tham số đặc trưng

Outline

1 Định nghĩa
2 Phân loại
3 Phân phối xác suất
- Bảng phân phối xác suất
- Hàm phân phối xác suất
- Hàm mật độ xác suất

4 Các tham số đặc trưng

BIẾN NGẪU NHIÊN
Nguyễn Văn Thìn
Định nghĩa
Phân loại
Phân phối xác suất
Bảng phân phối xác suất
Hàm phân phối xác suất
Hàm mật độ xác suất
Các tham số đặc trưng

Biến ngẫu nhiên: Định nghĩa

Định nghĩa 1

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

Biến ngẫu nhiên: Ví dụ

Ví dụ 2

Xét phép thử tung hai đồng xu. Không gian mẫu của phép thử này là

$$\Omega = \{SS, SN, NS, NN\}$$

Gọi X là số mặt ngửa xuất hiện. Khi đó, X là một ánh xạ từ không gian mẫu Ω vào $\mathbb{R}$ như sau:

ω	SS	NS	SN	NN
$X(\omega)$	0	1	1	2

Outline

BIẾN NGẪU NHIÊN
Nguyễn Văn Thìn
Định nghĩa
Phân loại
Phân phối xác suất
Bảng phân phối xác suất
Hàm phân phối xác suất
Hàm mật độ xác suất
Các tham số đặc trưng

Biến ngẫu nhiên: Phân loại

BIẾN NGẪU NHIÊN
Nguyễn Văn Thìn
Định nghĩa
Phân loại
Phân phối xác suất
Bảng phân phối xác suất
Hàm phân phối xác suất
Hàm mật độ xác suất
Các tham số đặc trưng

Dựa vào miền giá trị của biến ngẫu nhiên mà người ta phân thành 2 loại chính như sau

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.

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) hoặc toàn bộ $\mathbb{R}$

Biến ngẫu nhiên: Phân loại

BIẾN NGẪU NHIÊN
Nguyễn Văn Thìn
Định nghĩa
Phân loại
Phân phối xác suất
Bảng phân phối xác suất
Hàm phân phối xác suất
Hàm mật độ xác suất
Các tham số đặc trưng

Ví dụ 3

Tung 1 con xúc sắc cân đối. Gọi X là số chấm xuất hiện thì X là một biến ngẫu nhiên rời rạc vì tập các giá trị mà nó có thể nhận là $\{1, 2, 3, 4, 5, 6\}$.

Ví dụ 4

Các biến ngẫu nhiên X sau là biến ngẫu nhiên liên tục:

(a) Chọn ngẫu nhiên một thời điểm trong ngày và đo nhiệt độ không khí (X).

(b) Chọn ngẫu nhiên một bóng đèn điện tử và đo thời gian hoạt động bình thường của nó (X).

(c) Chọn ngẫu nhiên một hợp chất hóa học và đo độ pH của nó (X).

Outline

Phân phối xác suất của biến ngẫu nhiên

BIẾN NGẪU NHIÊN
Nguyễn Văn Thìn
Định nghĩa
Phân loại
Phân phối xác suất
Bảng phân phối xác suất
Hàm phân phối xác suất
Hàm mật độ xác suất
Các tham số đặc trưng

Kí hiệu

Cho $A \subset \mathbb{R}$. Ta kí hiệu

$$(X \in A) = \{\omega \in \Omega \colon X(\omega) \in A\}$$

Chẳng hạn, ta viết

$$(X = a) = \{\omega \in \Omega \colon X(\omega) = a\}$$
$$(X \le a) = \{\omega \in \Omega \colon X(\omega) \le a\}$$

Phân phối xác suất của biến ngẫu nhiên
Bảng phân phối xác suất

BIẾN NGẪU NHIÊN
Nguyễn Văn Thìn
Định nghĩa
Phân loại
Phân phối xác suất
Bảng phân phối xác suất
Hàm phân phối xác suất
Hàm mật độ xác suất
Các tham số đặc trưng

Để mô tả biến ngẫu nhiên rời rạc X nhận giá trị nào đó với xác suất tương ứng là bao nhiêu thì người ta dùng bảng phân phối xác suất. Bảng này có hai dòng như sau

- Dòng thứ nhất là các giá trị mà biến ngẫu nhiên X nhận được.
- Dòng thứ hai là xác suất biến ngẫu nhiên X nhận các giá trị tương ứng.

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

Hàm phân phối xác suất: Định nghĩa và tính chất

BIẾN NGẪU NHIÊN
Nguyễn Văn Thìn
Định nghĩa
Phân loại
Phân phối xác suất
Bảng phân phối xác suất
Hàm phân phối xác suất
Hàm mật độ xác suất
Các tham số đặc trưng

Định nghĩa 5

Cho biến ngẫu nhiên X, hàm thực

$$\begin{aligned} F_X : \mathbb{R} &\longrightarrow [0,1] \\ x &\longmapsto P(X \le x) \end{aligned}$$

được gọi là hàm phân phối xác suất của X.

Mệnh đề 6

Hàm phân phối xác suất $F(x) \equiv F_X(x)$ có các tính chất sau:

(i) *không giảm: $x \le y \Rightarrow F(x) \le F(y)$,*

(ii) *liên tục phải: $\lim_{x \to x_0^+} F(x) = F(x_0)$ với mọi số thực x_0,*

(iii) $F(-\infty) = \lim_{x \to -\infty} F(x) = 0, \quad F(+\infty) = \lim_{x \to +\infty} F(x) = 1,$

(iv) $\mathbb{P}(a < X \le b) = F(b) - F(a)$ *với $a \le b$ bất kì.*

Hàm phân phối xác suất của biến ngẫu nhiên rời rạc

Giả sử biến ngẫu nhiên X có bảng phân phối xác suất

X	x_1	x_2	x_3	$\dots$	x_k	x_{k+1}	$\dots$
$\mathbb{P}$	p_1	p_2	p_3	$\dots$	p_k	p_{k+1}	$\dots$

- Nếu $x < x_1$ thì $(X \le x) = \emptyset$ và ta có
$$F(x) = P(X \le x) = P(\emptyset) = 0$$
- Nếu $x_k \le x < x_{k+1}$ thì,
$$(X \le x) = (X \in \{x_1, x_2, \dots, x_k\}) = (X = x_1) \cup (X = x_2) \cup \dots \cup (X = x_k)$$
Mặt khác, do các biến cố $(X = x_i)$ xung khắc nhau từng đôi một nên
$$F(x) = P(X \le x) = P(X = x_1) + P(X = x_2) + \dots + P(X = x_k)$$

Hàm phân phối xác suất của biến ngẫu nhiên rời rạc

BIẾN NGẪU NHIÊN
Nguyễn Văn Thìn
Định nghĩa
Phân loại
Phân phối xác suất
Bảng phân phối xác suất
Hàm phân phối xác suất
Hàm mật độ xác suất
Các tham số đặc trưng

$$\begin{aligned} F(x) &= P(X \le x) = P(X = x_1) + P(X = x_2) + \dots + P(X = x_k) \\ &= p_1 + p_2 + \dots + p_k \end{aligned}$$

Vậy hàm phân phối xác suất của X là

$$F_X(x) = \begin{cases} 0 & \text{nếu } x < x_1 \\ p_1 & \text{nếu } x_1 \le x < x_2 \\ p_1 + p_2 & \text{nếu } x_2 \le x < x_3 \\ \dots & \dots \\ p_1 + p_2 + \dots + p_k & \text{nếu } x_k \le x < x_{k+1} \\ \dots & \dots \end{cases}$$

BIẾN NGẪU NHIÊN
Nguyễn Văn Thìn
Định nghĩa
Phân loại
Phân phối xác suất
Bảng phân phối xác suất
Hàm phân phối xác suất
Hàm mật độ xác suất
Các tham số đặc trưng

Ví dụ 7

Tung một đồng xu cân đối đồng chất. Gọi X là số mặt sấp xuất hiện. Hãy lập bảng phân phối xác suất của X và xác định hàm phân phối của nó.

Gợi ý 8

Bảng phân phối xác suất của X là

X	0	1
$\mathbb{P}$	0.5	0.5

Hàm phân phối xác suất của X là

$$F(x) = \begin{cases} 0 & \text{nếu } x < 0 \\ 0.5 & \text{nếu } 0 \le x < 1 \\ 1 & \text{nếu } x \ge 1 \end{cases}$$

BIẾN NGẪU NHIÊN
Nguyễn Văn Thìn
Định nghĩa
Phân loại
Phân phối xác suất
Bảng phân phối xác suất
Hàm phân phối xác suất
Hàm mật độ xác suất
Các tham số đặc trưng

Ví dụ 9

Gọi X là số nút xuất hiện khi tung một con xúc sắc. Hãy lập bảng phân phối và xác định hàm phân phối xác suất của X.

Ví dụ 10

Tung đồng thời hai đồng xu cân đối đồng chất. Gọi Y là số mặt sấp xuất hiện khi thực hiện phép thử, hãy lập bảng phân phối xác suất và xác định hàm phân phối xác suất của Y.

Ví dụ 11

Một người đi thi bằng lái xe, xác suất đậu của anh ta ở mỗi lần thi là 0.3. Anh ta sẽ thi đến khi đạt được bằng lái xe thì thôi. Gọi Z là số lần người đó dự thi. Lập bảng phân phối xác suất của Z.

Hàm mật độ xác suất: Định nghĩa

Định nghĩa 12

Cho biến ngẫu nhiên liên tục X. Hàm số $f(x)$ không âm, xác định trên $\mathbb{R}$ và thỏa các tính chất

i)
$$\mathbb{P}(X \in I) = \int_I f(x)dx, \quad \forall I \subset \mathbb{R}$$

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

được gọi là hàm mật độ xác suất của biến ngẫu nhiên X.

Hàm mật độ xác suất: Tính chất

BIẾN NGẪU NHIÊN
Nguyễn Văn Thìn
Định nghĩa
Phân loại
Phân phối xác suất
Bảng phân phối xác suất
Hàm phân phối xác suất
Hàm mật độ xác suất
Các tham số đặc trưng

Nhận xét 13

1) Mọi hàm $f(x)$ không âm và thỏa điều kiện $\int\limits_{-\infty}^{\infty} f(x)dx = 1$ đều là hàm mật độ xác suất của 1 biến ngẫu nhiên X nào đó.
2) Từ định nghĩa về hàm mật độ ta có hàm phân phối xác suất của biến ngẫu nhiên X có hàm mật độ $f(x)$ là
$$F(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^{x} f(u)du$$
3)
$$F'(x) = \frac{dF}{dx}(x) = f(x)$$
4) Trong trường hợp liên tục, $P(X = a) = \int_a^a f(x)dx = 0$.

BIẾN NGẪU NHIÊN
Nguyễn Văn Thìn
Định nghĩa
Phân loại
Phân phối xác suất
Bảng phân phối xác suất
Hàm phân phối xác suất
Hàm mật độ xác suất
Các tham số đặc trưng

Ví dụ 14

Cho hàm
$$f(x) = \begin{cases} 2x & \text{nếu } x \in [0,1] \\ 0 & \text{nếu } x \notin [0,1] \end{cases}$$

a) Chứng tỏ rằng $f(x)$ là hàm mật độ xác suất của biến ngẫu nhiên X nào đó.
b) Tìm hàm phân phối xác suất $F(x)$ của X.
c) Tính xác suất $\mathbb{P}(0 < X \leq \frac{1}{2})$.

BIẾN NGẪU NHIÊN
Nguyễn Văn Thìn
Định nghĩa
Phân loại
Phân phối xác suất
Bảng phân phối xác suất
Hàm phân phối xác suất
Hàm mật độ xác suất
Các tham số đặc trưng

Ví dụ 15

Tuổi thọ Y của một thiết bị (đơn vị: giờ) có hàm mật độ xác suất có dạng
$$f(x) = \begin{cases} \frac{a}{x^2} & \text{nếu } x \geq 100 \\ 0 & \text{nếu } x < 100 \end{cases}$$
với $a \in \mathbb{R}$.

a) Hãy xác định hàm phân phối của Y.
b) Thiết bị được gọi là loại A nếu tuổi thọ của nó kéo dài ít nhất 400 giờ. Tính tỉ lệ loại A.

Outline

Đặc trưng của biến ngẫu nhiên: Kỳ vọng

Định nghĩa 16

Kỳ vọng của X, ký hiệu $\mathbb{E}(X)$, là một số được định nghĩa

$$\mathbb{E}(X) = \begin{cases} \sum_x^{+\infty} x\mathbb{P}(X = x) & \text{nếu } X \text{ là BNN rời rạc} \\ \int_{-\infty}^{+\infty} xf(x)dx & \text{nếu } X \text{ là BNN liên tục} \end{cases}$$

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.

Kỳ vọng của biến ngẫu nhiên rời rạc

Ví dụ 17

Một hộp chứa 10 viên bi, trong đó có 3 viên bi nặng 10g, 5 viên nặng 50g, 2 viên nặng 20g. Chọn ngẫu nhiên ra 1 viên bi và gọi X là khối lượng của viên bi đó. Tính $\mathbb{E}(X)$.

Ví dụ 18

Một chùm chìa khóa có 6 chìa, trong đó có 2 chìa mở được cửa. Thử từng chìa (thử xong bỏ ra ngoài) cho đến khi mở được cửa. Tìm số lần thử trung bình để mở được cửa.

Kỳ vọng của biến ngẫu nhiên liên tục

Ví dụ 19

Cho X là một biến ngẫu nhiên có hàm mật độ

$$f(x) = \begin{cases} 2x & \text{nếu } x \in [0,1] \\ 0 & \text{nếu } x \notin [0,1] \end{cases}$$

Tìm kỳ vọng của X.

Ví dụ 20

Cho biến ngẫu nhiên Y có hàm mật độ xác suất

$$g(y) = \begin{cases} \frac{2}{y^2} & \text{nếu } y \in [1,2] \\ 0 & \text{nếu } y \notin [1,2] \end{cases}$$

Tìm $\mathbb{E}(Y)$.

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

Đặc trưng của biến ngẫu nhiên: Phương sai

BIẾN NGẪU NHIÊN
Nguyễn Văn Thìn
Định nghĩa
Phân loại
Phân phối xác suất
Bảng phân phối xác suất
Hàm phân phối xác suất
Hàm mật độ xác suất
Các tham số đặc trưng

Định nghĩa 21

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

$$\mathbb{V}ar(X) = \mathbb{E}(X - \mathbb{E}(X))^2 \tag{1}$$

Lưu ý

- Trong tính toán, để tính phương sai của biến ngẫu nhiên X ta thường sử dụng công thức $\mathbb{V}ar(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$.
- Đôi khi người ta còn kí hiệu $D(X)$ để chỉ phương sai của X trong một số giáo trình.

BIẾN NGẪU NHIÊN
Nguyễn Văn Thìn
Định nghĩa
Phân loại
Phân phối xác suất
Bảng phân phối xác suất
Hàm phân phối xác suất
Hàm mật độ xác suất
Các tham số đặc trưng

Định nghĩa 22 (Độ lệch chuẩn)

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

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

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) $\mathbb{V}ar(C) = 0$.
ii) $\mathbb{V}ar(CX) = C^2\mathbb{V}ar(X)$.
iii) Nếu X và Y độc lập thì $\mathbb{V}ar(X+Y) = \mathbb{V}ar(X) + \mathbb{V}ar(Y)$.

BIẾN NGẪU NHIÊN
Nguyễn Văn Thìn
Định nghĩa
Phân loại
Phân phối xác suất
Bảng phân phối xác suất
Hàm phân phối xác suất
Hàm mật độ xác suất
Các tham số đặc trưng

Ví dụ 23

Một hộp chứa 10 viên bi, trong đó có 3 viên bi nặng 10g, 5 viên nặng 50g, 2 viên nặng 20g. Chọn ngẫu nhiên ra 1 viên bi và gọi X là khối lượng của viên bi đó. Tính $\mathbb{E}(X)$, $\mathbb{V}ar(X)$.

Ví dụ 24

Cho biến ngẫu nhiên Y có hàm mật độ xác suất

$$g(y) = \begin{cases} \frac{2}{y^2} & \text{nếu } y \in [1,2] \\ 0 & \text{nếu } y \notin [1,2] \end{cases}$$

Tìm $\mathbb{E}(Y)$, $\mathbb{V}ar(Y)$.

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...

Đặc trưng của biến ngẫu nhiên: Mod

Định nghĩa 25 (Mod)

Mod của biến ngẫu nhiên X, ký hiệu $Mod(X)$, là giá trị mà biến ngẫu nhiên X nhận được với xác suất lớn nhất.

Từ định nghĩa, nếu biến ngẫu nhiên rời rạc X có bảng phân phối xác suất

X	x_1	x_2	$\cdots$	x_n	$\cdots$
$\mathbb{P}$	p_1	p_2	$\cdots$	p_n	$\cdots$

thì

$$Mod(X) = x_i \Leftrightarrow p_i = \mathbb{P}(X = x_i) = \max\{p_1, p_2 \dots\}$$

còn nếu X có phân phối liên tục với hàm mật độ xác suất $f(x)$ thì

$$Mod(X) = x_0 \Leftrightarrow f(x_0) = \max_{x\in\mathbb{R}} f(x)$$

Mod

Ví dụ 26 (Trường hợp rời rạc)

Tìm *Mod* của biến ngẫu nhiên X có phân phối rời rạc với bảng phân phối xác suất

X	1	2	3	4	5
$\mathbb{P}$	0,3	0,25	0,18	0,14	0,13

Ví dụ 27 (Trường hợp liên tục)

Tìm *Mod* của biến ngẫu nhiên X có hàm mật độ xác suất sau

$$f(x) = \begin{cases} \frac{3}{4}x(2-x) & \text{khi } 0 \le x \le 2 \\ 0 & \text{nơi khác} \end{cases}$$

Đặc trưng của biến ngẫu nhiên: Trung vị

Định nghĩa 28 (Trung vị)

Cho biến ngẫu nhiên X bất kỳ, trung vị của X, ký hiệu $Med(X)$, là giá trị m của biến ngẫu nhiên X sao cho

$$\begin{cases} \mathbb{P}(X \le m) \ge \frac{1}{2} \\ \mathbb{P}(X \ge m) \ge \frac{1}{2} \end{cases}$$

ta viết $Med(X) = m$.

Đặc trưng của biến ngẫu nhiên: Trung vị

Nhận xét 29

Khi X là biến ngẫu nhiên liên tục thì trung vị của X chính là điểm chia phân phối xác suất thành hai phần bằng nhau. Nghĩa là

$$\mathbb{P}(X \geq m) = \mathbb{P}(X \leq m) = 1/2$$

tương đương với $\mathbb{P}(X \geq m) = 1/2$ hoặc $\mathbb{P}(X \leq m) = 1/2$.

Chứng minh

Thật vậy, từ điều kiện $P(X \geq m) \geq 1/2$ suy ra $P(X \leq m) = P(X < m) \leq 1/2$. Kết hợp với điều kiện $P(X \leq m) \geq 1/2$ ta phải có $P(X \leq m) = 1/2$.

Đặc trưng của biến ngẫu nhiên: Trung vị

Ví dụ 30 (Trung vị phân phối rời rạc)

Giả sử biến ngẫu nhiên rời rạc X có bảng phân phối xác suất như sau

X	1	2	3	4
$\mathbb{P}$	0.1	0.2	0.3	0.4

Tìm $Med(X)$.

Ví dụ 31 (Trung vị phân phối rời rạc cho trường hợp không duy nhất)

Giả sử biến ngẫu nhiên rời rạc X có bảng phân phối xác suất như sau

X	1	2	3	4
$\mathbb{P}$	0.1	0.4	0.3	0.2

Tìm $Med(X)$.

Đặc trưng của biến ngẫu nhiên: Trung vị

Ví dụ 32 (Trung vị phân phối liên tục)

Giả sử biến ngẫu nhiên liên tục X có hàm mật độ xác suất cho bởi

$$f(x) = \begin{cases} 4x^3 & \text{khi } 0 < x < 1 \\ 0 & \text{nơi khác} \end{cases}$$

Tìm $Med(X)$.

Ví dụ 33 (Trung vị phân phối liên tục cho trường hợp không duy nhất)

Giả sử biến ngẫu nhiên liên tục X có hàm mật độ xác suất cho bởi

$$f(x) = \begin{cases} \frac{1}{2} & \text{khi } 0 \leq x \leq 1 \\ 1 & \text{khi } 2.5 \leq x \leq 3 \\ 0 & \text{nơi khác} \end{cases}$$

Tìm $Med(X)$.