



VLU | ĐẠI HỌC VĂN LANG

KHOA KHOA HỌC CƠ BẢN



GIÁO TRÌNH

**THỐNG KÊ ỨNG DỤNG
TRONG KINH TẾ VÀ KINH
DOANH**

THÀNH PHỐ HỒ CHÍ MINH - 2019

MỤC LỤC

CHƯƠNG 1 : BIẾN NGẪU NHIÊN

Mục lục chương 1	1
1.1 Khái niệm biến ngẫu nhiên	1
1.2 Biểu diễn biến ngẫu nhiên.....	3
1.2.1 Biến ngẫu nhiên rời rạc.	3
1.2.2 Biến ngẫu nhiên liên tục.....	5
1.3 Hàm phân phối biến ngẫu nhiên.....	7
1.3.1. Định nghĩa hàm phân phối xác suất.	7
1.3.2. Tính chất hàm phân phối xác suất.....	9
1.4 Hai biến ngẫu nhiên rời rạc độc lập.	9
1.4.1 Hai biến ngẫu nhiên rời rạc độc lập.....	9
1.4.2 Kết hợp hai biến ngẫu nhiên rời rạc độc lập.	9
1.5 Hàm của biến ngẫu nhiên.....	10
1.5.1 Hàm của biến ngẫu nhiên rời rạc.....	10
1.5.2 Hàm của biến ngẫu nhiên liên tục.....	11
1.6 Các đặc trưng của biến ngẫu nhiên.	12
1.6.1 Kỳ vọng	12
1.6.2 Phương sai	15
1.6.3 Giá trị tin chắc nhất (Mode).	18
1.6.4 Trung vị	20

CHƯƠNG 2 : MỘT SỐ PHÂN PHỐI XÁC SUẤT THÔNG DỤNG

Mục lục chương 2	22
2.1 Phân phối nhị thức.....	22
2.2 Phân phối siêu bội.....	26
2.3 Phân phối Poisson.....	30
2.4 Phân phối chuẩn	31



2.5	Phân phối Chi bình phương.....	35
2.6	Phân phối Student	37

CHƯƠNG 3 : NGUYÊN LÝ THỐNG KÊ VÀ CÁC KHÁI NIỆM CƠ BẢN

Mục lục chương 3.....	41
3.1. Một số khái niệm dùng trong thống kê.	41
3.1.1 Tổng thể thống kê và đơn vị tổng thể, và mẫu.....	43
3.1.2 Tiêu thức.....	44
3.1.3 Lượng biến.....	44
3.1.4 Tham số.	45
3.1.5 Thang đo.....	45
3.1.6 Thiết kế thang đo.	47
3.2. Thu thập và trình bày dữ liệu thống kê.	47
3.2.1 Xác định dữ liệu và phương pháp thu thập dữ liệu sơ cấp.....	47
3.2.2 Các kỹ thuật lấy mẫu dữ liệu.....	48
3.2.3 Xác định quy mô mẫu.....	49
3.2.4 Phân tổ	50
3.2.5 Trình bày dữ liệu thống kê	52
3.2.6 Đồ thị biểu đồ thống kê.....	54

CHƯƠNG 4 : TÓM TẮT DỮ LIỆU BẰNG ĐẠI LƯỢNG SỐ

Mục lục chương 4.....	55
4.1. Các đại lượng đo lường mức độ tập trung của dữ liệu.....	55
4.1.1 Số trung bình số học.	55
4.1.2 Số trung bình điều hòa.	57
4.1.3 Số trung bình nhân.....	58
4.1.4 Yếu vị (Mod).	59
4.1.5 Số trung vị (Median).....	60
4.2. Các khuynh hướng đo độ phân tán.	63

4.2.1	Khoảng biến thiên.	65
4.2.2	Độ lệch tuyệt đối trung bình.	64
4.2.3	Phương sai , độ lệch chuẩn.	64
4.2.4	Hệ số biến thiên.....	65
4.3.	Các khuynh hướng đo vị trí tương đối.	67
4.3.1	Phân vị.....	67
4.3.2	Tứ phân vị	67
4.3.3	Giá trị	68
4.4.	Hệ số tương quan của các bộ dữ liệu	70
4.4.1	Hiệp phương sai.....	70
4.4.2	Hệ số tương quan.....	72
4.5.	Hệ số đo hình dạng của quy luật phân phối.	74
4.2.5	Hệ số Kurtosis (độ nhọn).....	74
4.2.6	Độ lệch – Skewness	75

CHƯƠNG 5 : ƯỚC LƯỢNG THAM SỐ

Mục lục chương 5	77
5.1. CÁC TIÊU CHUẨN ƯỚC LƯỢNG.....	77
5.1.1 Ước lượng không chệch.	77
5.1.2 Khoảng tin cậy.	78
5.2. Khoảng tin cậy cho giá trị trung bình.	79
5.3. Khoảng tin cậy cho độ lệch hai giá trị trung bình.....	81
5.4. Khoảng tin cậy cho giá trị tỷ lệ.	83
5.5. Khoảng tin cậy cho độ lệch hai giá trị tỷ lệ.	85
5.6. Khoảng tin cậy cho giá trị phương sai.....	85
5.7. Khoảng tin cậy cho dự đoán giá trị quan sát	87

CHƯƠNG 6 : KIỂM ĐỊNH GIẢ THIẾT THỐNG KÊ

Mục lục chương 7	89
------------------------	----



6.1.	KHÁI NIỆM	89
6.1.1	Giả thiết và đối thuyết	89
6.1.2	Sai lầm loại I và sai lầm loại II.....	90
6.2.	KIỂM ĐỊNH GIẢ THIẾT CHO MỘT GIÁ TRỊ TỶ LỆ TỔNG THỂ.	91
6.2.1	Phân tích.....	91
6.2.2	Mô hình kiểm định.....	92
6.3.	KIỂM ĐỊNH GIẢ THIẾT CHO MỘT TRUNG BÌNH TỔNG THỂ.	93
6.3.1	Phân tích.....	93
6.3.2	So sánh trung bình tổng thể với một số khi biết phương sai.....	84
6.3.3	So sánh trung bình tổng thể với một số khi không biết phương sai.....	94
6.4.	KIỂM ĐỊNH GIẢ THIẾT CHO PHƯƠNG SAI TỔNG THỂ.	96
6.4.1	Phân tích	96
6.4.2	So sánh phương sai tổng thể với một số khi biết trung bình μ	96
6.4.3	So sánh phương sai tổng thể với một số khi chưa biết trung bình μ	97
6.5.	KIỂM ĐỊNH GIẢ THIẾT CHO HAI GIÁ TRỊ TỶ LỆ TỔNG THỂ.	98
6.5.1	Kiểm định giả thiết so sánh 2 tỷ lệ tổng thể sử dụng phân phối chuẩn.	98
i.	Phân tích.....	98
ii.	Mô hình kiểm định.....	99
6.5.2	Kiểm định giả thiết so sánh 2 tỷ lệ tổng thể sử dụng phân phối chi bình phương	100
i.	Phân tích.....	100
ii.	Mô hình kiểm định	100
6.6.	KIỂM ĐỊNH GIẢ THIẾT CHO HAI TRUNG BÌNH TỔNG THỂ.....	102
6.6.1	Phân tích.....	102
6.6.2	So sánh hai trung bình tổng thể khi biết phương sai.	102
6.6.3	So sánh hai trung bình tổng thể khi không biết phương sai và cỡ mẫu lớn.....	103
6.6.4	So sánh hai trung bình tổng thể khi không biết phương sai, phương sai bằng nhau và cỡ mẫu nhỏ.	103

6.7.	KIỂM ĐỊNH GIẢ THIẾT CHO HAI PHƯƠNG SAI TỔNG THỂ.....	105
6.8.	KIỂM TRA GIẢ THIẾT VỀ SỰ ĐỘC LẬP	106
6.8.1	Phân tích.	106
6.8.2	Kiểm định độc lập của hai bộ dữ liệu định tính.	107

CHƯƠNG 7 : DỰ BÁO CHUỖI THỜI GIAN

Mục lục chương 7	109
7.1 CHUỖI THỜI GIAN, CÁC KHÁI NIỆM CƠ BẢN	109
7.1.1 Khái niệm chuỗi thời gian.....	109
7.1.2 Các thành phần chuỗi thời gian.	109
i. Thành phần xu hướng.	109
ii. Thành phần chu kỳ.	110
iii. Thành phần mùa.....	110
iv. Thành phần bất thường.	110
7.1.3 Các đại lượng mô tả chuỗi thời gian.....	111
i. Mức độ trung bình theo thời gian	111
ii. Lượng tăng giảm tuyệt đối.....	112
iii. Tốc độ phát triển.....	112
iv. Tốc độ tăng giảm.....	113
7.2 CÁC PHƯƠNG PHÁP DỰ BÁO ĐƠN GIẢN.....	113
7.3 CÁC PHƯƠNG PHÁP LÀM TRƠN.....	114
7.3.1 Dự báo bằng phương pháp trung bình trượt	114
7.3.2 Dự báo bằng san bằng hàm mũ	116
7.3.3 Dự báo bằng hàm xu thế tuyến tính	117

PHỤ LỤC.BẢNG PHÂN PHỐI XÁC SUẤT

B1. BẢNG GIÁ TRỊ HÀM MẬT ĐỘ $N(0;1)$	121
B2. BẢNG TÍCH PHÂN LAPLACE	122
B3. BẢNG PHÂN PHỐI STUDENT	123

B4. BẢNG PHÂN PHỐI CHI BÌNH PHƯƠNG.....124

Tailieu.vn

CHƯƠNG 1

Biến ngẫu nhiên

Mục lục chương 1

1.1	Khái niệm biến ngẫu nhiên	1
1.2	Biểu diễn biến ngẫu nhiên	3
1.3	Hàm phân phối biến ngẫu nhiên	7
1.4	Hai biến ngẫu nhiên rời rạc độc lập	9
1.5	Hàm của biến ngẫu nhiên	11
1.6	Các đặc trưng số của biến ngẫu nhiên	13

Ở bước ban đầu khi tiếp cận về lý thuyết xác suất, sinh viên đã nghiên cứu về khái niệm biến cố, phân loại và phương pháp tính xác suất xảy ra của các biến cố. Trong chương một này, mục tiêu là hệ thống và quản lý khả năng xảy ra của các kết quả có thể có trong một phép thử. Khái niệm mới được đưa vào trong chương này là thuật ngữ **biến ngẫu nhiên**, là một khái niệm quan trọng trong lý thuyết xác suất, giúp chúng ta hiểu rõ quy luật, bản chất của các hiện tượng và phép thử.

1.1 KHÁI NIỆM BIẾN NGẪU NHIÊN

Trong nhiều trường hợp, chúng ta không quan tâm chi tiết đến mọi kết quả trong không gian mẫu của phép thử mà thay vào đó ta quan tâm đến phân nhóm cho các kết quả đó. Ví dụ thực hiện phép thử tung 3 đồng xu lần lượt, ta có không gian mẫu của phép thử là:

$$S = \{NNN; NNS; NSN; NSS; SNN; SNS; SSN; SSS\}$$

Trong đó ký hiệu S : tung đồng xu được sắp và N : tung đồng xu được ngửa. Như vậy ta có thể phân loại kết quả của phép thử thành 4 trường hợp: không được mặt sấp nào có xác suất là $\frac{1}{8}$,

được một mặt sấp có xác suất là $\frac{3}{8}$, được hai mặt sấp có xác suất là $\frac{3}{8}$ và được ba mặt sấp có xác suất là $\frac{1}{8}$. Như vậy nếu ta đặt một biến ngẫu nhiên chỉ số mặt sấp có được sau 3 lần tung, kí hiệu

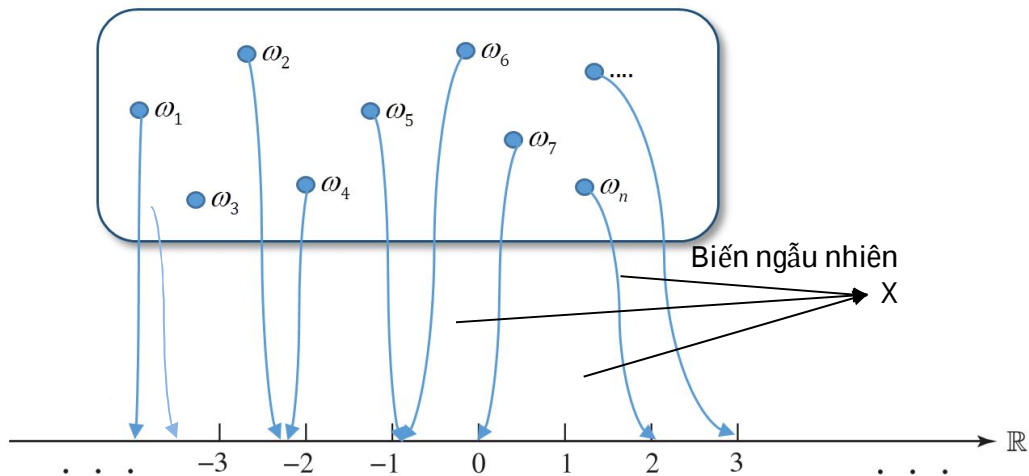
là X , thì $X \in \{0, 1, 2, 3\}$ sẽ đại diện cho 8 trường hợp trong không gian mẫu của phép thử.

Như vậy khái niệm biến ngẫu nhiên được mô hình hóa như sau:

Định nghĩa Biến ngẫu nhiên X của một phép thử là một hàm số đi từ không gian các biến cố sơ cấp Ω vào $\mathbb{R}$:

$$X : \Omega \rightarrow \mathbb{R}$$

$$\omega \mapsto X = X(\omega)$$



Hình 1.1: Biến ngẫu nhiên X.

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

Ta ký hiệu biến ngẫu nhiên X nhận giá trị x là $X = x$ và xác suất để X nhận giá trị x là $P(X = x)$.

Ví dụ 1.1.

Thực hiện phép thử tung đồng xu 3 lần, gọi X là biến ngẫu nhiên chỉ số mặt sấp có được trong 3 lần tung.

Ta có không gian mẫu của phép thử $S = \{NNN; NNS; NSN; NSS; SNN; SNS; SSN; SSS\}$

Và biến ngẫu nhiên $X : S \rightarrow \mathbb{R}$ có các giá trị như sau

$$X(NNN) = 0; X(NNS) = 1; X(NSN) = 1; X(NSS) = 2$$

$$X(SNN) = 1; X(SNS) = 2; X(SSN) = 2; X(SSS) = 3$$

Như vậy về mặt xác suất của biến ngẫu nhiên ta có:

$$P(X=0) = \frac{1}{8}; P(X=1) = \frac{3}{8}; P(X=2) = \frac{3}{8}; P(X=3) = \frac{1}{8}$$

Lưu ý. Ký hiệu $P(X=2) = \frac{3}{8}$ có thể hiểu là xác suất tung đồng xu 3 lần 2 lần được sấp là bằng $\frac{3}{8}$.

Ví dụ 1.2.

Một quyển tập Starbook có kích thước chuẩn $175 \times 255 (\pm 2mm)$ đang được lưu hành ngoài thị trường. Chọn một quyển tập bất kỳ và đo chiều dài quyển tập. Gọi X là biến ngẫu nhiên chỉ số đo chiều dài quyển tập.

Trong trường hợp này thì tập giá trị của biến ngẫu nhiên X là tất cả các giá trị nằm trong khoảng $(253; 257)mm$.

Dựa trên tập giá trị của biến ngẫu nhiên có thể nhận được, người ta phân biến ngẫu nhiên ra làm hai loại.

Biến ngẫu nhiên được gọi là rời rạc: nếu tập giá trị của biến ngẫu nhiên chỉ nhận hữu hạn hoặc vô hạn đếm được các giá trị. Ta có thể liệt kê các giá trị của biến ngẫu nhiên rời rạc $x_1; x_2; \dots; x_n; \dots$

Biến ngẫu nhiên được gọi là liên tục: nếu tập giá trị của biến ngẫu nhiên có thể lấy bất kỳ trên một khoảng của trục số thực.

Ví dụ 1.3.

Quan sát kết quả bài thi lấy chứng chỉ kiểm toán viên (CPA) của một nhân viên kế toán. Bài kiểm tra gồm 4 phần. Gọi X là số phần của bài kiểm tra mà nhân viên đó đã vượt qua. Khi đó, X là biến ngẫu nhiên rời rạc vì tập các giá trị mà nó có thể nhận là hữu hạn gồm các giá trị 0, 1, 2, 3, 4.

Ví dụ 1.4.

Quan sát xe ô tô đi qua một trạm thu phí. Biến ngẫu nhiên X là số xe hơi đi qua trạm thu phí trong 1 ngày. Khi đó, X là biến ngẫu nhiên rời rạc có thể nhận một trong các giá trị của dãy vô hạn (0, 1, 2, ...).

Ví dụ 1.5.

Chiều cao của thanh niên Việt Nam thường nằm trong khoảng từ 150 cm đến 180 cm. Chiều cao đo được cụ thể của một thanh niên nào đó có thể nhận bất kỳ giá trị nào nằm trong khoảng này, tùy thuộc vào độ chính xác của phép đo.

Ví dụ 1.6.

Quan sát các cuộc gọi đến phòng tiếp nhận thông tin của một công ty bảo hiểm. Gọi X là thời gian giữa hai cuộc gọi liên tiếp. X có thể nhận bất kỳ giá trị nào trong khoảng $[0; +\infty)$. X có thể nhận vô số các giá trị, chẳng hạn 1,26 phút, 2,755 phút, ...

1.2 BIỂU DIỄN BIẾN NGẪU NHIÊN.**1.2.1 Biến ngẫu nhiên rời rạc.****I. Bảng phân phối xác suất.**

Với X là biến ngẫu nhiên rời rạc, tập giá trị của X gồm các giá trị $x_1; x_2; \dots; x_n; \dots$ với $x_1 < x_2 < \dots < x_n < \dots$. Và xác suất tương ứng với các giá trị của biến ngẫu nhiên là $P(X = x_i) = p_i$ với $i = 1, 2, \dots, n, \dots$.

Để biểu diễn biến ngẫu nhiên X ta dùng bảng phân phối xác suất có cấu trúc như sau:

X	x_1	x_2	...	x_n	...
P	p_1	p_2	...	p_n	...

Nhận xét. Trong kết quả phép thử ngẫu nhiên, biến ngẫu nhiên rời rạc chỉ nhận một trong các giá trị $x_1, \dots, x_n, \dots$ nên các biến cố $\{X = x_j\}$ và $\{X = x_i\}$ xung khắc với mọi $i \neq j$.

Tính chất. Bảng phân phối xác suất của biến ngẫu nhiên có tính chất sau:

$$\begin{aligned} \text{i} \quad & \sum_{i=1}^{+\infty} P(X = x_i) = \sum_{i=1}^{+\infty} p_i = 1 \\ \text{ii} \quad & P(a \leq X \leq b) = \sum_{a \leq x_i \leq b} P(X = x_i) = \sum_{a \leq x_i \leq b} p_i \end{aligned}$$

II. Hàm mật độ xác suất.

Cho biến ngẫu nhiên rời rạc X nhận các giá trị $x_1; x_2; \dots; x_n; \dots$ với xác suất tương ứng là $P(X = x_i) = p_i$. **Hàm mật độ xác suất** biến ngẫu nhiên X dạng $f: \mathbb{R} \rightarrow \mathbb{R}$ thỏa

$$f(x) = \begin{cases} p_i & ; x = x_i, i = 1, 2, \dots, n, \dots \\ 0 & ; x \neq x_i, \forall i \end{cases}$$

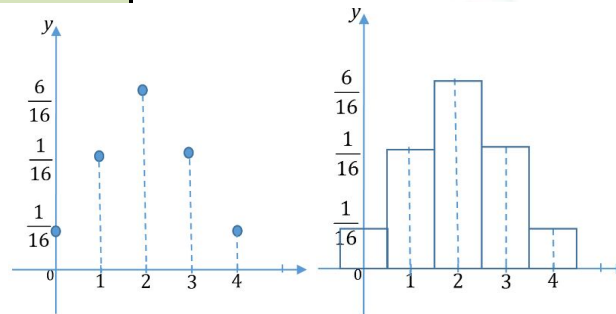
Tính chất. Tương tự bảng phân phối, hàm mật độ xác suất có các tính chất sau:

- i $f(x) \geq 0; \forall x \in \mathbb{R}$
- ii $\sum_x f(x) = 1$
- iii $P(a \leq X \leq b) = \sum_{a \leq x \leq b} f(x)$

Ví dụ 1.7.

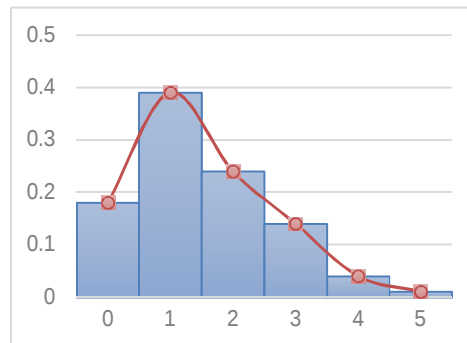
Với phép thử gieo 4 đồng xu lần lượt, và đặt X là biến ngẫu nhiên chỉ số mặt sấp có được sau 4 lần tung. Ta có bảng phân phối xác suất và hàm mật độ xác suất cho X .

X	0	1	2	3	4
P	$\frac{1}{16}$	$\frac{4}{16}$	$\frac{6}{16}$	$\frac{4}{16}$	$\frac{1}{16}$

**Hình 1.2:** Hàm mật độ xác suất biến ngẫu nhiên rời rạc.**Ví dụ 1.8.**

Xem xét doanh thu bán xe ô tô tại cửa hàng Dicalo Motors ở Saratoga, New York. Quan sát 300 ngày, thấy rằng có 54 ngày không bán được chiếc ô tô nào, 117 ngày bán được một chiếc, 72 ngày bán được 2 chiếc, 42 ngày bán được 3 chiếc, 12 ngày bán được 4 chiếc, 42 ngày bán được 3 chiếc, 12 ngày bán được 4 chiếc và 3 ngày bán được 5 chiếc. Giả sử phép thử là chọn một ngày bất kỳ của DiCarlo Motors và định nghĩa biến ngẫu nhiên X là số chiếc ô tô bán được trong ngày đó. Từ dữ liệu quá khứ, ta biết X là biến ngẫu nhiên rời rạc có thể nhận một trong các giá trị 0, 1, 2, 3, 4, 5. Ta có bảng phân phối xác suất của biến ngẫu nhiên X như sau:

X	0	1	2	3	4	5
P	0,18	0,39	0,24	0,14	0,04	0,01

**Ví dụ 1.9.**

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

X	1	2	3	4
P	$\frac{1}{10}$	$\frac{2}{10}$	$\frac{3}{10}$	$\frac{4}{10}$

Phân phối xác suất của biến ngẫu nhiên X có thể được biểu diễn bằng công thức:

$$f(x) = \frac{x}{10}, \text{ với } x = 1, 2, 3, \text{ hoặc } 4$$

Ứng với từng giá trị có thể có của X , ta có thể xác định phân phối xác suất $f(x)$ tương ứng.

Chẳng hạn, ta có thể xác định $f(2) = \frac{2}{10}$ chính là xác suất để X nhận giá trị 2.

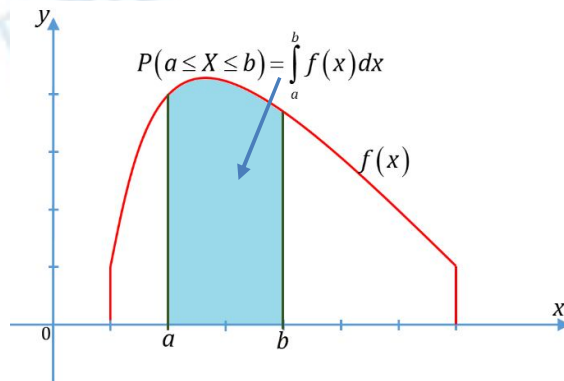
1.2.2 Biến ngẫu nhiên liên tục.

I. Hàm mật độ xác suất.

Cho biến ngẫu nhiên liên tục X , có tập giá trị D , **hàm mật độ xác suất** của biến ngẫu nhiên X là hàm $f(x)$ thỏa với mọi $a, b \in D$ thì:

$$P(a \leq X \leq b) = \int_a^b f(x) dx$$

Ý nghĩa. Hàm mật độ xác suất của biến ngẫu nhiên liên tục miêu tả xác suất biến ngẫu nhiên thuộc một khoảng có giá trị bằng vùng diện tích của hàm mật độ trong khoảng đó.



Hình 1.3: Hàm mật độ xác suất biến ngẫu nhiên liên tục.

II. Tính chất hàm mật độ xác suất.

Theo định nghĩa của hàm mật độ xác suất của biến ngẫu nhiên liên tục ta có hai tính chất cơ bản của hàm mật độ:

i $f(x) \geq 0$ với mọi $x \in (-\infty, +\infty)$

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

Nhận xét. Tính chất ii) giúp chỉ ra mối quan hệ giữa định nghĩa hàm mật độ xác suất và công thức tính xác suất trong chương 1.

$$P(X \in A) = \frac{|A|}{|S|} = \frac{\int_{x \in A} f(x) dx}{\int_{-\infty}^{+\infty} f(x) dx} = \int_{x \in A} f(x) dx$$

Hệ quả.

iii Đối với biến ngẫu nhiên liên tục thì mật độ xác suất tại một điểm thì bằng 0,

$$P(X = x_0) = 0, \forall x_0 \in \mathbb{R}.$$

iv Từ đó ta có

$$P(a \leq X \leq b) = P(a < X < b) = P(a \leq X < b) = P(a < X \leq b) = \int_a^b f(x) dx$$

Ví dụ 1.10.

Cho biến ngẫu nhiên X liên tục có hàm mật độ dạng

$$f(x) = \begin{cases} kx^3 & \text{khi } 0 < x < 1 \\ 0 & \text{khi } x \leq 0 \vee x \geq 1 \end{cases}$$

- Xác định hằng số k .
- Tính xác suất $X \in [0,4; 0,6]$

Giải**a. Xác định hằng số k .**

Theo tính chất ii) ta có

$$\int_{-\infty}^{+\infty} f(x) dx = 1 \Leftrightarrow \int_{-\infty}^0 0 dx + \int_0^1 kx^3 dx + \int_1^{+\infty} 0 dx = 1 \Leftrightarrow \left(\frac{k}{4} x^4 \right) \Big|_0^1 = 1 \Leftrightarrow \frac{k}{4} = 1$$

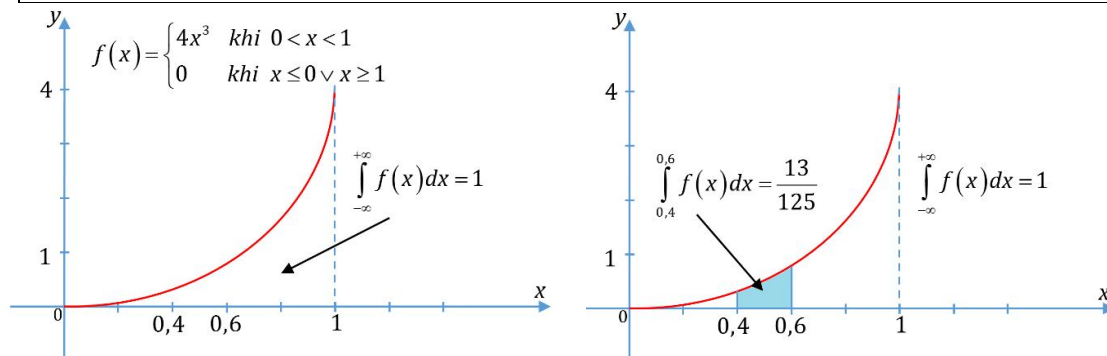
Vậy để $f(x)$ là hàm mật độ xác suất của biến ngẫu nhiên X thì $k = 4$. Và

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

b. Xác suất biến ngẫu nhiên $X \in [0,4; 0,6]$

$$P(0,4 \leq X \leq 0,6) = \int_{0,4}^{0,6} 4x^3 dx = \frac{13}{125}$$

Xác suất $X \in [0,4; 0,6]$ bằng $\frac{13}{125}$ so với 1 là xác suất X chắc chắn thuộc $[0;1]$.



Hình 1.4: Xác suất biến ngẫu nhiên liên tục

Ví dụ 1.11.

Nhãn trên chai nước giặt cho biết mỗi chai chứa 12 ounces. Giả sử dung tích trên các chai sản xuất được phân phối đều theo hàm mật độ xác suất sau:

$$f(x) = \begin{cases} 8 & \text{khi } 11,975 \leq x \leq 12,1 \\ 0 & \text{khi } x < 11,975 \vee x > 12,1 \end{cases}$$

Gọi X là dung tích trên một chai nước giặt.

a. Xác suất để một chai chứa từ 12 đến 12,05 ounces là bao nhiêu?

Tức là ta cần tính $P(12 \leq X \leq 12,05)$

$$P(12 \leq X \leq 12,05) = \int_{12}^{12,05} 8dx = 8x \Big|_{12}^{12,05} = 0,4.$$

b. Xác suất để một chai chứa từ 12,02 ounces trở lên là bao nhiêu?

$$P(X \geq 12,05) = P(12,02 \leq X \leq 12,1) = \int_{12,02}^{12,1} 8dx = 8x \Big|_{12,02}^{12,1} = 0,64.$$

c. Những chai có dung tích sai lệch không quá 0,02 ounces so với số in trên nhãn được chấp nhận là đạt tiêu chuẩn. Xác suất để một chai không đạt tiêu chuẩn là bao nhiêu?
Xác suất để một chai đạt tiêu chuẩn là

$$P(|X - 12| \leq 0,02) = P(11,98 \leq X \leq 12,02) = \int_{11,98}^{12,02} 8dx = 8x \Big|_{11,98}^{12,02} = 0,32.$$

Vậy, xác suất để một chai không đạt tiêu chuẩn là $1 - 0,32 = 0,68$.

1.3 HÀM PHÂN PHỐI BIẾN NGẪU NHIÊN.

1.3.1. Định nghĩa hàm phân phối xác suất.

Hàm phân phối xác suất của biến ngẫu nhiên X là hàm $F(x)$ được định nghĩa: $F: \mathbb{R} \rightarrow \mathbb{R}$ với

$$F(x) = P(X < x)$$

Hàm phân phối xác suất hay còn gọi là hàm phân phối tích lũy.

Nhận xét. Khai triển công thức hàm phân phối trong hai trường hợp:

I. Biến ngẫu nhiên rời rạc

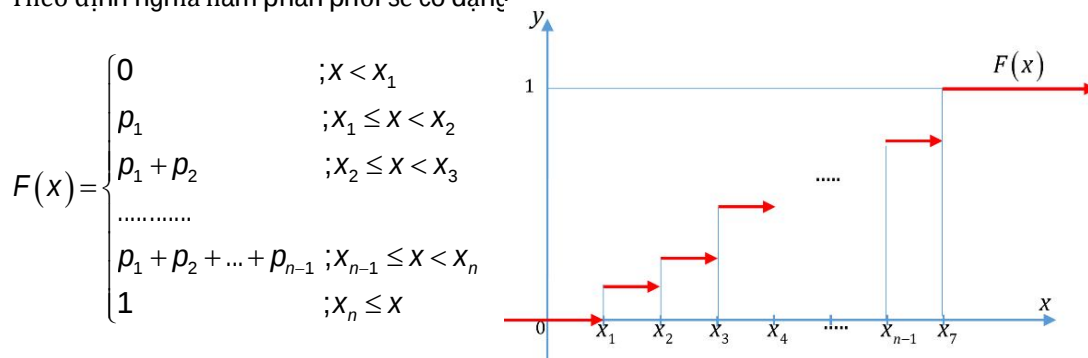
$$F(x) = P(X < x) = \sum_{x_i < x} P(X = x_i) = \sum_{x_i < x} p_i$$

Ví dụ 1.12.

Ta có bảng phân phối xác suất của X

X	x_1	x_2	...	x_n
P	p_1	p_2	...	p_n

Theo định nghĩa hàm phân phối sẽ có dạng:



Hình 1.5: Hàm phân phối rời rạc.

II. Biến ngẫu nhiên liên tục

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

Ví dụ 1.13.

Cho biến ngẫu nhiên X liên tục có hàm mật độ xác suất

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

Lập hàm phân phối xác suất cho biến ngẫu nhiên X

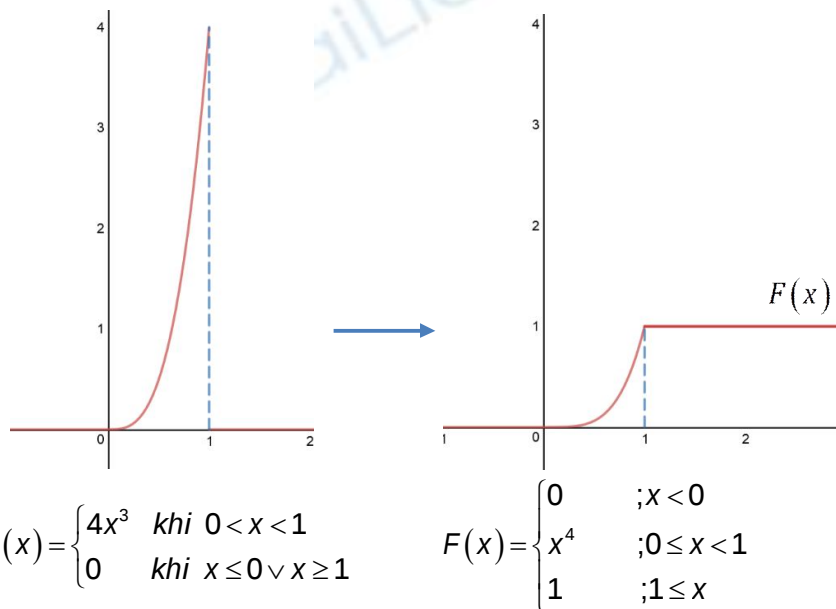
Nếu $x < 0$ ta có $F(x) = \int_{-\infty}^x f(t)dt = \int_{-\infty}^x 0dt = 0$

Nếu $0 \leq x < 1$ ta có $F(x) = \int_{-\infty}^x f(t)dt = \int_0^x f(t)dt = \int_0^x 4t^3dt = t^4 \Big|_0^x = x^4$

Nếu $x \geq 1$ ta có $F(x) = \int_{-\infty}^x f(t)dt = \int_0^1 f(t)dt + \int_1^x 0dt = \int_0^1 4t^3dt = 1$

Vậy hàm phân phối của biến ngẫu nhiên X có dạng

$$F(x) = \begin{cases} 0 & ; x < 0 \\ x^4 & ; 0 \leq x < 1 \\ 1 & ; 1 \leq x \end{cases}$$



Nhận xét. Nếu một biến ngẫu nhiên X liên tục có

hàm mật độ dạng $f(x) = \begin{cases} h(x) & ; x \in [a; b] \\ 0 & ; x \notin [a; b] \end{cases}$; với $h(x)$ liên tục trên $[a; b]$.

thì hàm phân phối xác suất dạng $F(x) = \begin{cases} 0 & ; x < a \\ H(x) & ; a \leq x < b; \text{ với } H'(x) = h(x). \\ 1 & ; b \leq x \end{cases}$

1.3.2. Tính chất hàm phân phối xác suất.

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

i. $0 \leq F(x) \leq 1, \forall x \in \mathbb{R}$

ii. Hàm $F(x)$ là hàm không giảm, nghĩa là với $x_1 < x_2 \Leftrightarrow F(x_1) \leq F(x_2)$

iii. Với $F(x)$ là hàm phân phối của biến ngẫu nhiên liên tục ta có $P(a < X < b) = F(b) - F(a)$

với mọi $a, b \in \mathbb{R}$, $a \leq b$

iv. **Quan hệ giữa $f(x)$ và $F(x)$:** Nếu hàm phân phối xác suất $F(x)$ của biến ngẫu nhiên X khả vi tại mọi x , với $f(x)$ là hàm mật độ xác suất, thì ta có:

$$F(x) = \int_{-\infty}^x f(t) dt \quad \longrightarrow \quad F'(x) = f(x)$$

1.4 HAI BIẾN NGẪU NHIÊN RỜI RẠC ĐỘC LẬP.

Cho biến ngẫu nhiên X và Y rời rạc có bảng phân phối xác suất lần lượt:

X	x_1	x_2	...	x_n	Y	y_1	y_2	...	y_m
P	p_1	p_2	...	p_n	P	q_1	q_2	...	q_m

1.4.1 Hai biến ngẫu nhiên rời rạc độc lập.

Hai biến ngẫu nhiên $X; Y$ được gọi là độc lập với nhau khi và chỉ khi xác suất biến ngẫu nhiên này nhận giá trị không ảnh hưởng đến xác suất biến ngẫu nhiên kia nhận giá trị. Và theo công thức nhân xác suất trong chương 1 ta có:

$$P((X = x_i) \cdot (Y = y_j)) = P(X = x_i) \cdot P(Y = y_j) = p_i \cdot q_j \quad \forall i, j$$

Ví dụ 1.14.

Tung 2 viên xúc sắc riêng biệt. Gọi $X; Y$ là biến ngẫu nhiên chỉ số nút của xúc sắc thứ nhất và thứ hai. Ta có ví dụ

$$P((X = 2) \cdot (Y = 3)) = \frac{1}{36} = \frac{1}{6} \cdot \frac{1}{6} = P(X = 2) \cdot P(Y = 3)$$

Là một trường hợp về xác suất của cặp giá trị $X; Y$ nhận được lần lượt là 2 ; 3.

Tương tự ta có $P((X = i) \cdot (Y = j)) = \frac{1}{36} = \frac{1}{6} \cdot \frac{1}{6} = P(X = i) \cdot P(Y = j) \quad \forall i, j = \overline{1, 6}$

Nghĩa là hai biến ngẫu nhiên $X; Y$ độc lập với nhau.

1.4.2 Kết hợp hai biến ngẫu nhiên rời rạc độc lập.

Cho biến ngẫu nhiên X và Y rời rạc, độc lập có bảng phân phối như ban đầu

Ta có biến ngẫu nhiên $(X + Y)$ có bảng phân phối xác suất dạng:

$X + Y$	z_1	z_2	...	z_k
P	P_1	P_2	...	P_k

Trong đó $\{z_1; z_2; \dots; z_k\} = \{x_i + y_j \mid i = \overline{1, n}; j = \overline{1, m}\}$

$$\text{Và } P_i = P(X + Y = z_i) = \sum_{x_i; y_j : x_i + y_j = z_i} P(X = x_i) \cdot P(Y = y_j) = \sum_{x_i; y_j : x_i + y_j = z_i} p_i q_j$$

Tương tự trong một trường hợp khác nếu ta kết hợp $(X \cdot Y)$ thì bảng phân phối xác suất có cấu trúc tương tự:

$X.Y$	Z_1	Z_2	...	Z_k
P	P_1	P_2	...	P_k

Trong đó $\{z_1; z_2; \dots; z_k\} = \{x_i \cdot y_j / i = \overline{1, n}; j = \overline{1, m}\}$

$$\text{Và } P_i = P(X.Y = z_i) = \sum_{x_i; y_j : x_i \cdot y_j = z_i} P(X = x_i) \cdot P(Y = y_j) = \sum_{x_i; y_j : x_i \cdot y_j = z_i} p_i q_j$$

Ví dụ 1.15.

Cho hai biến ngẫu nhiên rời rạc $X; Y$ độc lập có bảng phân phối xác suất lần lượt:

X	-1	1	2	3
P	0,3	0,4	0,2	0,1

Y	1	3	5
P	0,3	0,5	0,2

Ta có bảng phân phối xác suất của $(X + Y)$ dạng:

$X + Y$	0	2	3	4	5	6	7	8
P	0,09	0,27	0,06	0,29	0,1	0,13	0,04	0,02

Trong đó:

$$P(X + Y = 0) = P(X = -1) \cdot P(Y = 1) = 0,3 \cdot 0,3 = 0,09$$

$$P(X + Y = 2) = P(X = -1) \cdot P(Y = 3) + P(X = 1) \cdot P(Y = 1) = 0,3 \cdot 0,5 + 0,4 \cdot 0,3 = 0,27$$

$$P(X + Y = 3) = P(X = 2) \cdot P(Y = 1) = 0,2 \cdot 0,3 = 0,06$$

$$P(X + Y = 4) = P(X = -1) \cdot P(Y = 5) + P(X = 1) \cdot P(Y = 3) + P(X = 3) \cdot P(Y = 1) \\ = 0,3 \cdot 0,2 + 0,4 \cdot 0,5 + 0,1 \cdot 0,3 = 0,29$$

$$P(X + Y = 5) = P(X = 2) \cdot P(Y = 3) = 0,2 \cdot 0,5 = 0,1$$

$$P(X + Y = 6) = P(X = 1) \cdot P(Y = 5) + P(X = 3) \cdot P(Y = 3) = 0,4 \cdot 0,2 + 0,1 \cdot 0,5 = 0,13$$

$$P(X + Y = 7) = P(X = 2) \cdot P(Y = 5) = 0,2 \cdot 0,2 = 0,04$$

$$P(X + Y = 8) = P(X = 3) \cdot P(Y = 5) = 0,1 \cdot 0,2 = 0,02$$

1.5 HÀM CỦA BIẾN NGẪU NHIÊN.

Cho biến ngẫu nhiên X và $f(x)$ là một hàm số xác định tại mọi giá trị trong tập giá trị của biến ngẫu nhiên X , thì $Y = f(X)$ là một biến ngẫu nhiên mới và là hàm theo biến ngẫu nhiên X .

1.5.1 Hàm của biến ngẫu nhiên rời rạc.

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

X	x_1	x_2	...	x_n
P	p_1	p_2	...	p_n

Và $Y = f(X)$ là hàm theo biến ngẫu nhiên X .

Bảng phân phối xác suất của biến ngẫu nhiên Y có dạng

$Y = f(X)$	y_1	y_2	...	y_k
P	P_1	P_2	...	P_k

Theo nguyên tắc:

Tập giá trị của biến ngẫu nhiên $Y : \{y_1, y_2, \dots, y_k\} \equiv \{f(x_1), f(x_2), \dots, f(x_n)\}$.

Giá trị xác suất : $P_i = P(Y = y_i) = \sum_{x_j: f(x_j)=y_i} p_j$.

Ví dụ 1.16.

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

X	1	2	3	4
P	0,4	0,3	0,2	0,1

Bảng phân phối xác suất của biến ngẫu nhiên $Y = X^2$ có dạng:

$Y = X^2$	1	4	9	16
P	0,4	0,3	0,2	0,1

Ví dụ 1.17.

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

X	-1	1	2	3
P	0,4	0,3	0,2	0,1

Bảng phân phối xác suất của biến ngẫu nhiên $Y = X^2$ có dạng

$Y = X^2$	1	4	9
P	0,7	0,2	0,1

Trong đó $P(Y=1) = P(X^2=1) = P(X=1) + P(X=-1) = 0,4 + 0,3 = 0,7$.

Cũng như $P(Y=4) = P(X^2=4) = P(X=2) + P(X=-2) = 0,2 + 0 = 0,2$.

$P(Y=9) = P(X^2=9) = P(X=3) + P(X=-3) = 0,1 + 0 = 0,1$.

1.5.2 Hàm của biến ngẫu nhiên liên tục.

Cho biến ngẫu nhiên X rời rạc có hàm mật độ xác suất $f(x)$. Và $Y = h(X)$ là hàm theo biến ngẫu nhiên X .

Gọi $G(y)$ là hàm phân phối xác suất của biến ngẫu nhiên Y , ta có:

$$G(y) = P(Y \leq y) = P(h(X) \leq y) = \int_{x: h(x) \leq y} f(x) dx$$

Và $g(y)$ là hàm mật độ xác suất của biến ngẫu nhiên, ta có:

$$g(y) = G'(y)$$

Ví dụ 1.18.

Cho biến ngẫu nhiên liên tục X có hàm mật độ xác suất

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

Và hàm biến ngẫu nhiên $Y = X^3$, lập hàm mật độ xác suất cho biến ngẫu nhiên Y .

Gọi $G(y)$ là hàm phân phối xác suất cho biến ngẫu nhiên Y .

$$G(y) = P(Y < y) = P(X^3 < y) = P(X < \sqrt[3]{y}) = \int_{-\infty}^{\sqrt[3]{y}} f(x) dx$$

Nếu $\sqrt[3]{y} < 0 \Leftrightarrow y < 0$ ta có $G(y) = \int_{-\infty}^{\sqrt[3]{y}} 0 dx = 0$

Nếu $0 \leq \sqrt[3]{y} < 1 \Leftrightarrow 0 \leq y < 1$ ta có $G(y) = \int_{-\infty}^{\sqrt[3]{y}} f(x) dx = \int_0^{\sqrt[3]{y}} 4x^3 dx = x^4 \Big|_0^{\sqrt[3]{y}} = y^{\frac{4}{3}}$

Nếu $1 \leq \sqrt[3]{y} \Leftrightarrow 1 \leq y$ ta có $G(y) = \int_{-\infty}^{\sqrt[3]{y}} f(x) dx = \int_0^1 f(x) dx + \int_1^{\sqrt[3]{y}} 0 dx = 1$

$$\text{Vậy } G(y) = \begin{cases} 0 & ; y < 0 \\ y^{\frac{4}{3}} & ; 0 \leq y < 1 \\ 1 & ; 1 \leq y \end{cases}$$

Ta có hàm mật độ của biến ngẫu nhiên $Y = X^3$ có dạng:

$$g(y) = \begin{cases} \frac{4}{3} \sqrt[3]{y} & ; y \in [0; 1] \\ 0 & ; y \notin [0; 1] \end{cases}$$

1.6 CÁC ĐẶC TRƯNG CỦA BIẾN NGẪU NHIÊN.

1.6.1 Kỳ vọng

I. Kỳ vọng biến ngẫu nhiên rời rạc.

Định nghĩa. Giả sử biến ngẫu nhiên rời rạc X có bảng phân phối xác suất

X	x_1	x_2	...	x_n	...
P	p_1	p_2	...	p_n	...

Kỳ vọng của X , ký hiệu EX , là một số được định nghĩa:

$$EX = \sum_{i=1}^{+\infty} x_i p_i$$

Ví dụ 1.19.

Gọi X là biến ngẫu nhiên chỉ số nút nhận được khi tung xúc sắc. Ta có bảng phân phối xác suất của X

X	1	2	3	4	5	6
P	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

thì kỳ vọng của biến ngẫu nhiên X

$$EX = 1 \times \frac{1}{6} + 2 \times \frac{1}{6} + 3 \times \frac{1}{6} + 4 \times \frac{1}{6} + 5 \times \frac{1}{6} + 6 \times \frac{1}{6} = \frac{21}{6} = 3,5$$

Nhận xét. Giá trị 3,5 cũng chính là số nút trung bình nhận được khi tung một xúc sắc.

Ví dụ 1.20.

Một người tham gia vào một trò chơi vòng xoay roulette, (Một bánh xe xoay xung quanh tâm, trên bánh xe có 38 ô tương ứng với các ô 00, 0, 1, 2, 3, ..., 35, 36), tương ứng với ô 00 là màu trắng, ô 0 là màu xanh, các ô chẵn (2, 4, ..., 36) màu đỏ, và các ô lẻ (1, 3, ..., 35) màu đen. Có nhiều hình thức đặt cược trong trò chơi, và người này đặt cược vào chẵn lẻ. Giả sử người này đặt cược vào ô chẵn \$1, trung bình trong một lượt chơi thì người này thắng hay thua bao nhiêu tiền.