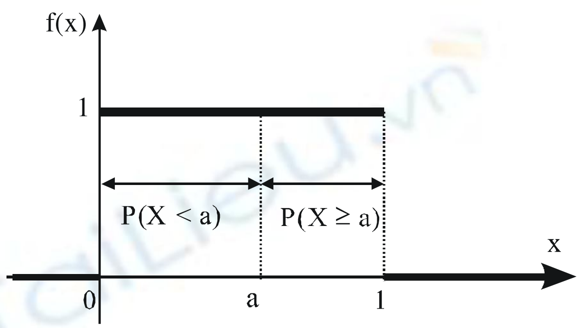
Cần chú ý rằng: $\sum_{i=1}^{3} p_i =$ 0,3 + 0,4 +0,3 = 1.

Một số phân phối xác suất thường dùng của biến ngẫu nhiên liên tục và rời rạc được liệt kê dưới đây.

Phân phối đều trong [0,1): X nhận các giá trị thuộc nửa khoảng [0,1) với khả năng "như nhau". Hàm mật độ xác suất f(x) của nó được biển diễn trên hình IV.1.



Hình IV.1. Đồ thị hàm mật độ phân phối đều

Phân phối Poát−xông: Với một hệ thống hàng chờ một kênh (xem mục 3), số lượng X tín hiệu đến trong một khoảng thời gian là một biến ngẫu nhiên, X có thể nhận các giá trị nguyên không âm 0, 1,..., k,...

Giả sử số tín hiệu đến trung bình trong một khoảng thời gian đã biết được (kí hiệu số đó là λ) thì với một số điều kiện nhất định có thể coi X tuân theo luật phân phối xác suất Poát−xông (*Poisson*) như sau:

Các giá trị của X: x_i	0	1	...	k	... $+\infty$
Xác suất p_i tương ứng	$P(X = 0)$	$P(X = 0)$	...	$P(X = k) = \dfrac{\lambda^k e^{-\lambda}}{k!}$	...

Dễ thấy: $\sum_{i=0}^{+\infty} p_i = e^{-\lambda}\left[\frac{\lambda^0}{0!} + \frac{\lambda^1}{1!} + \frac{\lambda^2}{2!} + \ldots + \frac{\lambda^k}{k!} + \ldots\right] = e^{-\lambda} \times e^{\lambda} = 1.$

Chú ý rằng số đặc trưng cho giá trị trung bình của biến ngẫu nhiên X được gọi là kì vọng. Trong phân phối Poát−xông, kì vọng của X là λ. Số đặc trưng cho độ phân tán các giá trị của X xung quanh giá trị kì vọng của nó được gọi là độ lệch chuẩn σ. Với phân phối Poát−xông thì $\sigma^2 = \lambda$.

Phân phối mũ: Trên đây ta đã xét phân phối Poát–xông của số các tín hiệu đến trong một đơn vị thời gian. Một kiểu biến ngẫu nhiên thường xét là khoảng thời gian giữa hai tín hiệu liên tiếp sẽ tuân theo phân phối mũ. Đây là biến ngẫu nhiên liên tục chỉ nhận các giá trị không âm với hàm mật độ xác suất là $f(\tau) = \lambda e^{-\lambda\tau}$. Kí hiệu biến ngẫu nhiên đang xét là τ thì xác suất $P(\tau \le t) = \int_0^t \lambda e^{-\lambda\tau} d\tau$ có thể hiểu là xác suất cộng dồn cho tới t. Do đó hàm phân phối xác suất của τ là: $F(t) = \int_0^t f(\tau)d\tau = \int_0^t \lambda e^{-\lambda\tau} d\tau = -e^{-\lambda\tau} \Big|_0^t$ = $1 - e^{-\lambda t}$.

Phân phối chuẩn tắc N(0, 1): Giả sử X là biến ngẫu nhiên có phân phối chuẩn tắc N(0,1). Lúc đó nó có kì vọng m = 0 và độ lệch chuẩn σ = 1. Hàm phân phối xác suất của X có dạng:

$$F(x) = P(X \le x) = \int_{-\infty}^{x} f(x)dx = \int_{-\infty}^{x} (1/\sqrt{2\pi})\exp(-x^2/2)dx.$$

Cho X là biến ngẫu nhiên tuân theo luật phân phối chuẩn $N(m, \sigma^2)$ có kì vọng m, độ lệch chuẩn σ. Lúc đó, thực hiện phép đổi biến $Z = \dfrac{X - m}{\sigma}$ thì Z là một biến ngẫu nhiên tuân theo luật phân phối chuẩn tắc N(0,1).

Mô phỏng các phân phối xác suất

Ví dụ 1: Mô phỏng phân phối đều trên [0, 1)

Cách 1: Dùng bảng số ngẫu nhiên (xem phụ lục 2A và 2B). Đây là các bảng số ghi lại các số (giả) ngẫu nhiên được phát sinh nhờ các hàm sinh số ngẫu nhiên trong máy tính. Chẳng hạn, sử dụng phụ lục 2B chúng ta nhận được một dãy số ngẫu nhiên: 0,10; 0,09; 0,73; 0,25...

Cách 2: Sử dụng các hàm sinh số ngẫu nhiên (*Random number generator*) đã được cài đặt trên máy tính.

Dù dùng bảng số ngẫu nhiên hay sử dụng các hàm sinh số ngẫu nhiên trong máy tính, ta cũng lấy ra hoặc tính được liên tiếp các số ngẫu nhiên $\mathbf{x_i}$ trong [0, 1) với i = 1, 2,..., n. Tần số các giá trị này rơi vào k khoảng nhỏ với độ dài bằng nhau 1/k được chia ra từ [0, 1) là gần như nhau (≈ n/k). Với n lớn thì các tần số đó càng sát gần n/k. Vì vậy ta coi các giá trị phát sinh được là các thể hiện của biến ngẫu nhiên X tuân theo phân phối đều trên [0, 1).

Trong trường hợp cần mô phỏng biến Y phân phối đều trên [a, b), ta chỉ việc tính $y_i = a + (b - a)x_i$. Chú ý rằng để phát sinh các số ngẫu nhiên nhận giá trị nguyên 0, 1, 2,..., N, chỉ cần áp dụng công thức $y_i = [(N + 1)x_i]$, trong đó vế phải là phần nguyên của $(N + 1)x_i$. Một số bảng số ngẫu nhiên nguyên hay hàm sinh số ngẫu nhiên nguyên cài đặt sẵn trong các hệ máy tính cũng giúp giải quyết vấn đề này.

Ví dụ 2: Mô phỏng phân phối rời rạc với luật phân phối xác suất sau

Các giá trị của X: x_i	6	9	12
Xác suất p_i	0,3	0,4	0,3

Muốn mô phỏng phân phối trên, trước hết cần tạo ra một dãy các chữ số ngẫu nhiên bằng cách tra bảng số ngẫu nhiên hay dùng hàm sinh số ngẫu nhiên đã được cài đặt trong máy tính. Chẳng hạn ta có thể chọn dãy sau 1009732533 7652013586 3467354876... lấy từ hàng đầu bảng số ngẫu nhiên trong phụ lục 2B. Ta quy định nếu các chữ số 0, 1, 2 xuất hiện thì coi X = 6, nếu 3, 4, 5, 6 xuất hiện thì coi X = 9, còn nếu có 7, 8, 9 xuất hiện thì coi X = 12. Lúc đó ứng với 10 chữ số đầu tiên của dãy trên $a_1a_2...a_{10}$ = 1009732533 ta có bảng sau đây cho biết các giá trị của X có thể lấy:

a_i	1	0	0	9	7	3	2	5	3	3
Các giá trị của X: x_i	6	6	6	12	12	9	6	9	9	9

Như vậy, đã có 10 giá trị (thể hiện) của X được tạo ra. Tương tự, có thể tạo ra các thể hiện khác của X. Do tần suất (hay xác suất thực nghiệm) của mỗi chữ số ngẫu nhiên từ 0 tới 9 trong bảng số ngẫu nhiên là khoảng 10% nên tần suất (xác suất thực nghiệm) X nhận giá trị 6, 9 và 12 theo thứ tự là 30%, 40% và 30%. Do đó có thể coi P(X = 6) = 30%, P(X = 9) = 40%, P(X = 12) = 30%.

Vậy muốn mô phỏng phân phối của X phải phát sinh ra một loạt các giá trị (các thể hiện) $\mathbf{x_i}$ của biến ngẫu nhiên X tuân theo quy luật phân phối đã cho.

Ví dụ 3: Mô phỏng phân phối mũ.

Giả sử biến ngẫu nhiên $\boldsymbol{\tau}$ tuân theo phân phối mũ với hàm phân phối xác suất là F(t) = $P(\boldsymbol{\tau} \le t) = 1 - e^{-\lambda t}$, với λ là tham số đã cho của phân phối mũ. F(t) chính là xác suất để $\boldsymbol{\tau}$ nhận giá trị không lớn hơn một số t cho trước.

Nếu $\mathbf{r}$ là biến ngẫu nhiên có phân phối đều trên [0, 1) thì $P(\mathbf{r} \ge e^{-\lambda t}) = 1 - e^{-\lambda t} = P(\boldsymbol{\tau} \le t)$ (xem hình III.1). Do đó, $P(\ln\mathbf{r} \ge -\lambda t) = P(-\frac{1}{\lambda}\ln\mathbf{r} \le t) = P(\boldsymbol{\tau} \le t)$. Vậy để phát sinh ra giá trị ngẫu nhiên τ của $\boldsymbol{\tau}$ thì trước hết cần phát sinh ra giá trị ngẫu nhiên r của $\mathbf{r}$ và tính $\tau = -\frac{1}{\lambda}\ln r$. Chẳng hạn, từ bảng số ngẫu nhiên (phụ lục 2B), nếu lấy r = 0,10 và λ = 5 thì $\tau = -0{,}2\times\ln r = -0{,}2\times\ln 0{,}1 = 0{,}46$. Tiếp theo, nếu lấy r = 0,09 thì $\tau = -0{,}2\times\ln 0{,}09 = 0{,}482$. Cứ như vậy ta thu được một dãy các thể hiện của $\boldsymbol{\tau}$.

2. ÁP DỤNG MÔ PHỎNG NGẪU NHIÊN

2.1. Vai trò của phương pháp mô phỏng

Nhiều bài toán thực tế chứa các yếu tố ngẫu nhiên, bất ổn định không giải được bằng các phương pháp giải tích. Nếu chúng ta áp dụng các phương pháp giải tích thì trong nhiều trường hợp buộc phải công nhận những giả thiết chặt chẽ không được thoả mãn trên thực tế và do đó lời giải tìm được cũng ít có giá trị thực tiễn. Phương pháp mô phỏng được dùng rộng rãi để giải các bài toán loại đó, nhất là những bài toán liên quan đến hệ thống lớn, bất ổn định, hàm chứa nhiều yếu tố ngẫu nhiên.

Chúng ta cần áp dụng phương pháp mô phỏng trong các tình huống sau đây:

– Khi không tìm được mô hình giải tích nào thích hợp.

– Các hoạt động của hệ thống thường bị ngắt quãng, đứt đoạn không theo quy luật nào cả.

– Mô phỏng là phương pháp duy nhất cho chi phí tiết kiệm và tốn ít thời gian.

Tuy nhiên phương pháp mô phỏng có một số điểm hạn chế sau:

– Không đưa ra được lời giải chính xác.

– Khó xác định được sai số.

– Mô phỏng chỉ sử dụng khi môi trường có tính bất ổn định.

– Mô phỏng chỉ tạo ra các phương án đánh giá chứ không đưa ra được kĩ thuật tìm lời giải tối ưu.

– Mô phỏng đôi khi rất đắt tiền.

2.2. Các bước cần tiến hành khi áp dụng mô phỏng

– Xác định vấn đề hay hệ thống cần mô phỏng.

– Xác định mô hình mô phỏng.

– Đo và thu thập số liệu cần thiết cho mô hình.

– Chạy mô phỏng.

– Phân tích kết quả mô phỏng, nếu cần thì phải sửa lại phương án đã được đánh giá qua chạy mô phỏng.

– Chạy mô phỏng để kiểm chứng phương án mới.

– Kiểm tra tính đúng đắn của mọi kết luận về hệ thống thực tế được rút ra sau khi chạy mô phỏng.

Trên đây là các bước cần làm khi áp dụng mô phỏng ngẫu nhiên để tìm ra các phương án hợp lí cho các bài toán thực tế. Ngoài ra, mô phỏng còn được áp dụng để giải quyết nhiều vấn đề khác.

2.3. Một số ví dụ về áp dụng phương pháp mô phỏng

Ví dụ 1: Cần lựa chọn một trong hai chiến lược để phát triển sản phẩm, với các số liệu thu thập được cho trong ba bảng IV.1, IV.2 và IV.3.

Bảng IV.1. Xác suất thời gian phát triển sản phẩm

Thời gian phát triển sản phẩm	Xác suất	
	Chiến lược I	Chiến lược II
6	0,2	0,4
9	0,3	0,4
12	0,5	0,2

Bảng IV.2. Chi phí lợi nhuận

Chi phí/giá bán	Chiến lược I	Chiến lược II
Chi phí cố định	600.000	1.500.000
Chi phí biến thiên/đơn vị	7,5	6,75
Giá bán/đơn vị sản phẩm	10	10

Bảng IV.3. Doanh số phụ thuộc thời gian phát triển sản phẩm

Doanh số	Xác suất		
	6 tháng	9 tháng	12 tháng
1.000.000	0,2	0,4	0,5
1.500.000	0,8	0,6	0,5

Vấn đề đặt ra là áp dụng phương pháp mô phỏng để tính lợi nhuận trung bình của từng chiến lược, sau đó kiểm tra kết quả (so sánh với kết quả lí thuyết).

Như vậy có năm phân phối xác suất cần mô phỏng ứng với năm biến ngẫu nhiên: X_1 – thời gian phát triển sản phẩm (theo chiến lược) I, X_2 – thời gian phát triển sản phẩm II, X_3 – doanh số cho thời gian 6 tháng, X_4 – doanh số cho thời gian 9 tháng và X_5 – doanh số cho thời gian 12 tháng. Trong ví dụ này, để trình bày đơn giản về vấn đề mô phỏng các phân phối xác suất của các biến trên, ta dùng mười số ngẫu nhiên, mỗi số gồm mười chữ số ngẫu nhiên rút ra từ bảng số ngẫu nhiên – phụ lục 2A (vì vậy các chữ số 0, 1, 2,..., 9 mỗi số chiếm khoảng 10%).

a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}
1	5	8	1	9	2	2	3	9	6
2	0	6	8	5	7	7	9	8	4
8	2	6	2	1	3	0	8	9	2
8	3	7	4	8	3	6	0	4	9
4	6	3	7	5	6	7	4	8	8

0	9	2	8	1	0	5	5	8	2
7	2	9	5	0	8	8	5	7	9
9	5	8	6	1	1	1	6	3	2
7	0	5	5	5	0	8	7	6	7
6	4	7	2	3	8	2	9	3	4

Ta quy định a_1 ứng với X_1, a_2 ứng với X_2, a_6 ứng với X_3, a_8 ứng với X_4 và a_{10} ứng với X_5. Ngoài ra cũng quy định:

$$a_1 = \left[\begin{array}{ll} 0,1 & \text{thì } X_1 = 6 \text{ tháng (thời gian phát triển sản phẩm I)} \\ 2,3,4 & \text{thì } X_1 = 9 \text{ tháng} \\ 5,6,7,8,9 & \text{thì } X_1 = 12 \text{ tháng} \end{array}\right.$$

$$a_2 = \left[\begin{array}{ll} 0,1,2,3 & \text{thì } X_2 = 6 \text{ tháng (thời gian phát triển sản phẩm II)} \\ 4,5,6,7 & \text{thì } X_2 = 9 \text{ tháng} \\ 8,9 & \text{thì } X_2 = 12 \text{ tháng} \end{array}\right.$$

$$a_6 = \left[\begin{array}{ll} 0,1 & \text{thì } X_3 = 10^6 \text{ (doanh số 6 tháng phát triển sản phẩm)} \\ 2,3,\ldots,9 & \text{thì } X_3 = 1,5.10^6 \end{array}\right.$$

$$a_8 = \left[\begin{array}{ll} 0,1,2,3 & \text{thì } X_4 = 10^6 \text{ (doanh số 9 tháng phát triển sản phẩm)} \\ 4,5,\ldots,9 & \text{thì } X_4 = 1,5.10^6 \end{array}\right.$$

$$a_{10} = \left[\begin{array}{ll} 0,1,2,3,4 & \text{thì } X_5 = 10^6 \text{ (doanh số 12 tháng phát triển sản phẩm)} \\ 5,6,\ldots,9 & \text{thì } X_5 = 1,5.10^6 \end{array}\right.$$

Cần nhắc lại một số công thức trong lĩnh vực quản trị kinh doanh như sau:

+ Lợi nhuận = (Doanh số − Điểm hoà vốn) × (Lợi nhuận/đơn vị sản phẩm)

+ Điểm hoà vốn = (Chi phí cố định)/(Lợi nhuận/đơn vị sản phẩm)

+ Lợi nhuận/đơn vị sản phẩm = (Giá bán/đơn vị sản phẩm) - (chi phí/đơn vị sản phẩm)

Các tính toán mô phỏng được tổng hợp trong bảng IV4.

Bảng IV.4. Kết quả tính toán mô phỏng

Số ngẫu nhiên										Thời gian		Doanh số		Lợi nhuận	
a_1	a_2				a_6		a_8		a_{10}	I	II	I	II	I	II
1	5	8	1	9	2	2	3	9	6	6	9	$1,5.10^6$	10^6	$3,15.10^6$	$1,75.10^6$
2	0	6	8	5	7	7	9	8	4	9	6	$1,5.10^6$	$1,5.10^6$	$3,15.10^6$	$3,38.10^6$
8	2	6	2	1	3	0	8	9	2	12	6	10^6	$1,5.10^6$	$1,9.10^6$	$3,38.10^6$
8	3	7	4	8	3	6	0	4	9	12	6	$1,5.10^6$	$1,5.10^6$	$3,15.10^6$	$3,38.10^6$
4	6	3	7	5	6	7	4	8	8	9	9	$1,5.10^6$	$1,5.10^6$	$3,15.10^6$	$3,38.10^6$

0	9	2	8	1	0	5	5	8	2	6	12	10^6	10^6	$1{,}9.10^6$	$1{,}75.10^6$
7	2	9	5	0	8	8	5	7	9	12	6	$1{,}5.10^6$	$1{,}5.10^6$	$3{,}15.10^6$	$3{,}38.10^6$
9	5	8	6	1	1	1	6	3	2	12	9	10^6	$1{,}5.10^6$	$1{,}9.10^6$	$3{,}38.10^6$
7	0	5	5	5	0	8	7	6	7	12	6	$1{,}5.10^6$	10^6	$3{,}15.10^6$	$1{,}75.10^6$
6	4	7	2	3	8	2	9	3	4	12	9	10^6	$1{,}5.10^6$	$1{,}9.10^6$	$3{,}38.10^6$

Điểm hoà vốn của chiến lược I $= \dfrac{600.000}{10 - 7{,}5} = 240.000$

Điểm hoà vốn của chiến lược II $= \dfrac{1.500.000}{10 - 6{,}75} = 461.538$

Bảng IV.5. So sánh lợi nhuận giữa chiến lược I và II

	Chiến lược I	Chiến lược II
Tổng lợi nhuận	$26{,}5 \times 10^6$	$28{,}91 \times 10^6$
Lợi nhuận trung bình (Σ lợi nhuận/10)	$2{,}65 \times 10^6$	$2{,}891 \times 10^6$

Cần chú ý rằng trong bảng IV.5 là kết quả tính toán khi chạy mô phỏng 10 lượt ứng với 10 số đã chọn ra. Nếu ta lấy càng nhiều số ngẫu nhiên thì độ chính xác đạt được càng cao. Vì vậy, nếu việc tính toán trên đây được lập trình và chạy trên máy tính với hàng trăm, hàng ngàn lượt thì độ chính xác sẽ rất cao.

Qua các phân tích trên ta thấy, để tiến hành mô phỏng cần phải có:

– Cơ sở dữ liệu *(DataBase)*

– Cơ sở tri thức *(KnowledgeBase)*

Kiểm tra kết quả mô phỏng trên bằng cách so sánh với kết quả lí thuyết được thực hiện như sau:

Doanh số chiến lược I = $0{,}2\times(0{,}2\times10^6 + 0{,}8\times1{,}5\times10^6) + 0{,}3\times(0{,}4\times10^6 + 0{,}6\times1{,}5\times10^6) + 0{,}5\times(0{,}5\times10^6 + 0{,}5\times1{,}5\times10^6) = 1{,}295\times10^6$. Lợi nhuận trung bình chiến lược I = $(1{,}295 - 0{,}24)\times2{,}5\times10^6 = 2{,}637\times10^6$. Kết quả tính toán mô phỏng là $2{,}65\times10^6$ rất sát với kết quả này.

Tương tự ta tính được doanh số và lợi nhuận trung bình cho chiến lược II ($2{,}84\times10^6$) và rút ra được kết luận về độ chính xác của tính toán mô phỏng.

Ví dụ 2: Tìm xác suất p để bao lồi của 4 điểm lấy bất kì trong vòng tròn đơn vị là một hình tam giác *(bài toán Sylvester)*.

Có lẽ cách đơn giản nhất để giải bài toán này là áp dụng mô phỏng ngẫu nhiên theo các bước sau đây: