Đại lượng ngẫu nhiên rời rạc

Giảng viên: PGS.TS. Lê Sỹ Vinh
Khoa CNTT – Đại học Công Nghệ

Nội dung

- Đại lượng ngẫu nhiên
- Phân bố xác suất
- Kì vọng, Phương sai
- Phân bố nhị thức
- Phân bố poisson
- Phân bố đồng thời

Đại lượng ngẫu nhiên rời rạc

Đại lượng (biến) ngẫu nhiên (ĐLNN) X biểu diễn định lượng kết quả của một phép thử **C**. X ánh xạ mỗi kết quả của phép thử **C** sang một giá trị thực

$$X : \Omega \rightarrow R$$

- X(Ω): tập hợp các giá trị có thể của ĐLNN X

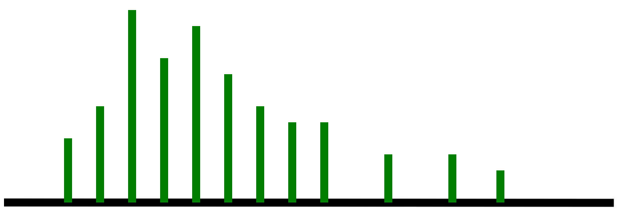
Ví dụ:

- Gieo một con xúc xắc. Gọi X là số nốt xuất hiện trên con xúc xắc, X là một ĐLNN, kí hiệu X(Ω) = {1, 2, 3, 4, 5, 6}.
- Tung đồng xu. Không gian mẫu Ω = {H, T}. ĐLNN X(Ω) = {0, 1}.

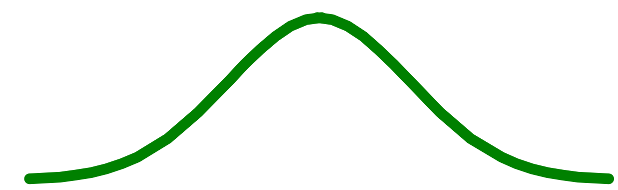
Đại lượng ngẫu nhiên

Đại lượng ngẫu nhiên

ĐLNN rời rạc



ĐLNN liên tục



Đại lượng ngẫu nhiên rời rạc

- Có miền giá trị là tập hữu hạn hoặc vô hạn đếm được
- Ví dụ:
  - Tung một con xúc xắc 2 lần. Đặt X là số lần mặt 6 điểm xuất hiện. X có thể nhận các giá trị 0, 1, hoặc 2.
  - Tung đồng xu 5 lần. Đặt Y là số lần xuất hiện mặt hình. Y có thể nhận các giá trị 0, 1, 2, 3, 4, hoặc 5.

Phân bố xác suất

- Phân bố xác suất (probability mass distribution) của một ĐLNN rời rạc X là một bảng bao gồm tất cả các giá trị mà ĐLNN X có thể nhận và kèm theo xác suất để nhận giá trị đó.

X	x_1	x_2	…	x_n
P	p_1	p_2	…	p_n

ở đó $p_i = P(X = xi)$. Lưu ý $p_1 + p_2 + \cdots + p_n = 1$.

- Hàm phân bố tích lũy (cumulative distribution function)

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

Ví dụ 1

- Một túi chứa ba tấm thẻ đánh số 1, 2, 3 và 1 túi chứa 3 tấm thẻ đánh số 4, 5, 6. Chọn ngẫu nhiên một tấm thẻ từ mỗi túi và tính tổng 2 tấm thẻ chọn được. Gọi X là kết quả, hãy lập bảng phân bố xác suất, và hàm phân bố tích lũy của X.
- Chọn ngẫu nhiên ba đứa trẻ từ một nhóm gồm 6 bé trai và 4 bé gái. Gọi X là số bé gái trong nhóm được chọn. Lập bảng phân bố xác suất của X.

Mode của X, kí hiệu mod(X) , là giá trị x_i có xác suất lớn nhất.

Kì vọng

Cho X là ĐLNN rời rạc có bảng phân bố xác suất

X	x_1	x_2	…	x_n
P	p_1	p_2	…	p_n

Kì vọng (hay gọi là giá trị trung bình) của X, kí hiệu là EX được tính như sau:

$$EX = \sum x_i . p_i$$

Ví dụ 2

Bảng phân bố xác suất của độ tuổi vào đại học ở Việt Nam được cho như sau

X	<17	17	18	19	20	21	>21
P	0	0.08	0.65	0.2	0.05	0.02	0.0

Tính kì vọng của tuổi vào đại học tại Việt Nam.

EX = ?

Ví dụ 3

Bảng phân bố xác suất của lương sinh viên CNTT sau khi ra trường

X	<3	4	5	6	7	8	9	10	11	12	>12
P	0	0.01	0.01	0.05	0.05	0.1	0.4	0.22	0.11	0.05	0

Tính kì vọng của lương sinh viên CNTT sau khi ra trường.

Tính chất của kỳ vọng

1) $EC = C$ nếu C là hằng số
2) $E(CX) = C.EX$ nếu C là hằng số
3) $E(X + Y) = EX + EY$
4) $E(XY) = EX.EY$ nếu X và Y độc lập

Phương sai và độ lệch chuẩn

Cho X là ĐLNN rời rạc có bảng phân bố xác suất

X	x_1	x_2	…	x_n
P	p_1	p_2	…	p_n

và kì vọng EX = μ. Độ lệch khỏi giá trị trung bình là X-μ.

- Phương sai của X, kí hiệu là DX:

  $$DX = E(X-\mu)^2 = \Sigma\, (x_i-\mu)^2 * p_i$$

- Độ lệch chuẩn của X, kí hiệu σ_X là căn bậc hai của phương sai DX.

Ví dụ 4

Lương của nhân viên 1 công ty TechQ

X	<3	4	5	6	7	8	9	10	200	>200
P	0	0.03	0.1	0.1	0.2	0.5	0.02	0.02	0.03	0

Tính kì vọng, phương sai, độ lệch chuẩn của lương nhân viên công ty TechQ.

Ví dụ 5

Bảng phân bố xác suất của lương sinh viên CNTT sau khi ra trường

X	<3	4	5	6	7	8	9	10	>10
P	0	0.02	0.05	0.1	0.2	0.5	0.02	0.11	0

Tính phương sai, độ lệch chuẩn của lương sinh viên CNTT sau khi ra trường.

Ví dụ 6

Bảng phân bố xác suất của độ tuổi vào đại học ở Việt Nam được cho như sau

X	<17	17	18	19	20	21	>21
P	0	0.03	0.65	0.2	0.05	0.07	0.0

Tính phương sai, độ lệch chuẩn của tuổi vào đại học tại Việt Nam.

Tính chất của phương sai

1) $D(c)=0$ nếu c là hằng số

2) $D(cX)=c^2DX$ nếu c là hằng số

   $D(X+c)=DX$

3) $D(X + Y) = DX + DY$ nếu X và Y độc lập.

Nội dung

- Đại lượng ngẫu nhiên
- Phân bố xác suất
- Kì vọng, Phương sai
- Phân bố nhị thức
- Phân bố poisson
- Phân bố đồng thời

Phân bố nhị thức

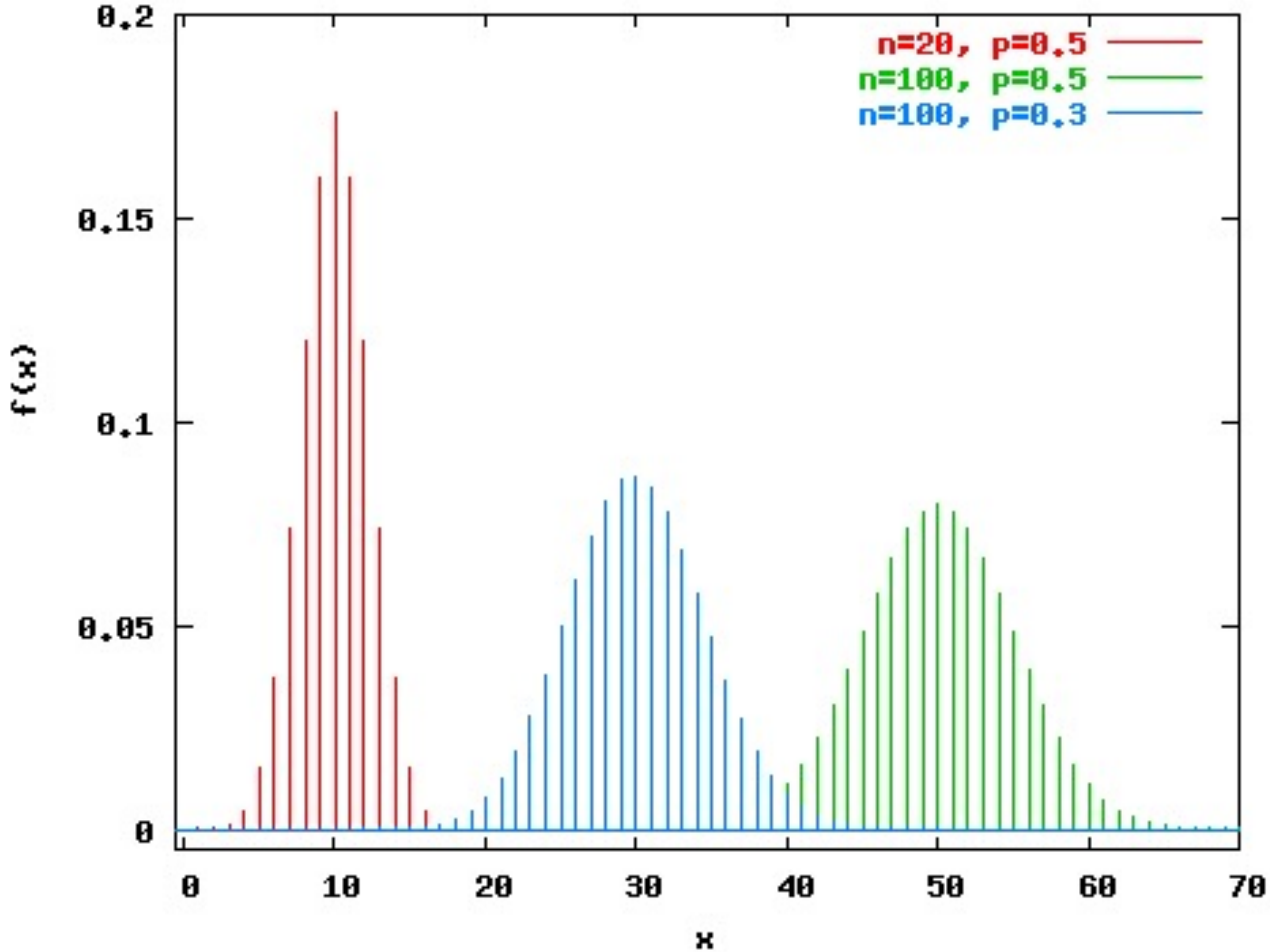
Xét phép thử ngẫu nhiên **C** chỉ có 2 kết quả là thành công hay thất bại. Xét biến cố A là phép thử thành công với $P(A) = p$. Phép thử **C** được tiến hành lặp đi lặp lại n lần. Gọi X là số lần biến cố A xuất hiện.

- X là một ĐLNN với $X(\Omega) = \{0,1,\ldots n\}$.
- Theo công thức Becnuli:

$$P\{X=k\} = C^k_n \, p^k \, (1-p)^{n-k}$$

- ĐLNN X được gọi là có phân bố nhị thức với tham số n và p và kí hiệu là $X \sim B(n, p)$.
- Kì vọng $EX = np$; Phương sai $DX = np(1-p)$

Phân bố nhị thức



Tính chất phân bố nhị thức

- Phép thử **C** chỉ có 2 kết quả là thành công hay thất bại
- Phép thử **C** được tiến hành đúng n lần.
- Xác suất của thành công hay thất bại là cố định trong cả n lần thử.
- Kết quả của phép thử là độc lập trong các lần thử khác nhau.
- Xác suất tích lũy:

$$P\{X \leq k\} = \sum_{i=0}^{k} C_n^i p^i (1-p)^{n-i}$$

- Với các giá trị k khác nhau được tính sẵn ở Bảng 1 (Phụ lục 2).

Ví dụ 7

Tỉ lệ hộ gia đình sử dụng internet ở VN là 80%. Chọn ngẫu nhiên 10 gia đình và gọi X là số gia đình có sử dụng internet.

a) Tính xác suất P{X=1}

b) Tính xác suất P{X <= 5}

c) Tính xác suất P{ X >=9}

d) Tính kì vọng của X (EX=?)

e) Tính phương sai của X (DX=?)

Ví dụ 8

Tỉ lệ một động cơ ô tô bị hỏng trong thời gian bảo hành 1 năm là 1%. Theo dõi 12 xe ô tô trong thời gian bảo hành. Gọi X là số xe hỏng trong thời gian bảo hành. Tính

- Tính $P\{X = 1\}$
- Tính $P\{X = 2\}$
- Tính $P\{X > 10\}$
- Tính kì vọng của X
- Tính phương sai của X

Phân bố Poisson

- Phân bố nhị thức quan tâm đến xác suất của số lần thử thành công sau n lần thử.
- Phân bố Poisson quan tâm đến số lần xuất hiện của một biến cố trong một khoảng thời gian (không gian, khoảng cách, hay một độ đo nào đó) xác định trước.

Ví dụ:

- Số bệnh nhân xuất hiện trong 1 đêm tại bệnh viện để bố trí số bác sỹ trực.
- Số khách hàng vào cửa hàng trong 1 tiếng để bố trí nhân viên bán hàng.
- Số lượng sinh viên vắng mặt trong buổi học XSTK

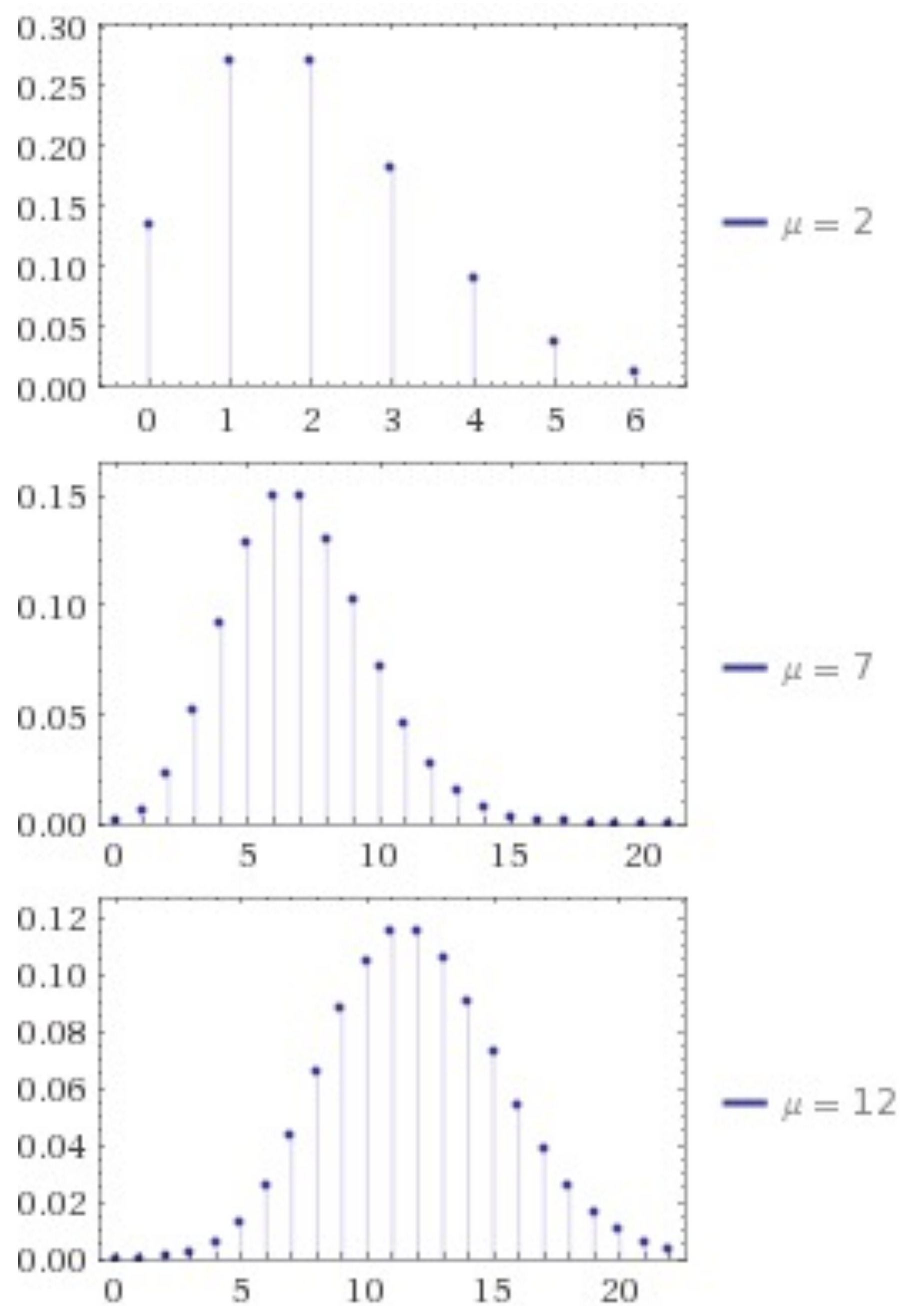
Phân bố Poisson

- Kì vọng của phân bố Poisson chính là số lượng xuất hiện trung bình của một biến cố trong một khoảng thời gian.
- Số lượng xuất hiện biến cố trong một khoảng thời gian khác nhau là độc lập với nhau.
- Gọi X là ĐLNN biểu diễn số lần xuất hiện của biến cố trong một khoảng thời gian xác định. Xác suất của ĐLNN X theo phân bố Poisson:

$$P\{X = k\} = \frac{e^{-\mu}\mu^{k}}{k!}$$

Trong đó μ là kì vọng của X.

Phân bố Poisson



Phân bố Poisson

- Xác suất tích lũy

$$P\{X \leq k\} = \sum_{i=0}^{k} \frac{e^{-\mu}\mu^{i}}{i!}$$

- Các giá trị μ và k khác nhau được tính sẵn ở Bảng 2 (Phụ lục 2).

Ví dụ 9

Một gara cho thuê ô tô thấy rằng số lượng người đến thuê xe là một ĐLNN theo phân bố Poisson. Trung bình có 6 người đến thuê vào thứ Bảy. Giả sử gara có 10 chiếc xe, hãy tính các xác suất sau đây:

a) Tất cả 10 xe đều được thuê

b) Không xe nào được thuê

c) Ít nhất 5 xe được thuê

d) Gara cần có bao nhiêu xe để xác suất không đáp ứng được nhu cầu thuê xe bé hơn 1%.

$$P\{X = k\} = \frac{e^{-\mu}\mu^k}{k!}$$

Ví dụ 10

Ở một tổng đài chăm sóc khách hàng, các cuộc điện thoại xuất hiện ngẫu nhiên với tần suất trung bình khoảng 6 cuộc trong 1 phút. Hãy tính xác suất sau:

a) Có đúng 10 cú điện thoại trong 1 phút

b) Không có cú điện thoại nào trong 1 phút

c) Có ít nhất 1 cú điện thoại trong thời gian 1 phút

$$P\{X = k\} = \frac{e^{-\mu}\mu^{k}}{k!}$$

Nội dung

- Đại lượng ngẫu nhiên
- Phân bố xác suất
- Kì vọng, Phương sai
- Phân bố nhị thức
- Phân bố poisson
- Phân bố đồng thời

Phân bố đồng thời

Gọi X và Y là hai ĐLNN rời rạc với

X(Ω) = $\{x_1, \dots, x_m\}$ và

Y(Ω) = $\{y_1, \dots, y_n)$.

Kí hiệu $P_{ij} = P\{X = x_i, Y = y_j\}$ là xác suất đồng thời của $X = x_i$ và $Y = y_j$.

Bảng phân bố xác suất đồng thời của X và Y:

		Y					
		y_1	**y_2**	**…**	**y_j**	**…**	**y_n**
X	**x_1**	p_{11}	p_{12}	…	p_{1j}	…	p_{1n}
	x_2	p_{21}	p_{22}	…	p_{2j}	…	p_{2n}
	…	…	…	…	…	…	…
	x_i	p_{i1}	p_{i2}	…	p_{ij}	…	p_{in}
	…			…		…	
	x_m	p_{m1}	p_{m2}	…	p_{mj}	…	p_{mn}

Ví dụ 11

Ba đồng tiền cân đối A, B, C được tung. Gọi X, Y là các ĐLNN được xác định như sau:

X: Số mặt ngửa trên đồng tiền A và B

Y: Số mặt ngửa trên cả 3 đồng tiền A, B và C

Hãy lập bảng phân bố đồng thời của X và Y.

Lưu ý: Nếu X và Y độc lập thì $(X = x_i)$ độc lập với $(Y = y_j)$ với mọi cặp (i,j). Cách khác, $P(X = x_i, Y = y_j) = P(X = x_i) * P(Y = y_j)$.