Đại lượng ngẫu nhiên rời rạc

Giảng viên: Hoàng Thị Điệp
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

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

$$X : W \to R$$

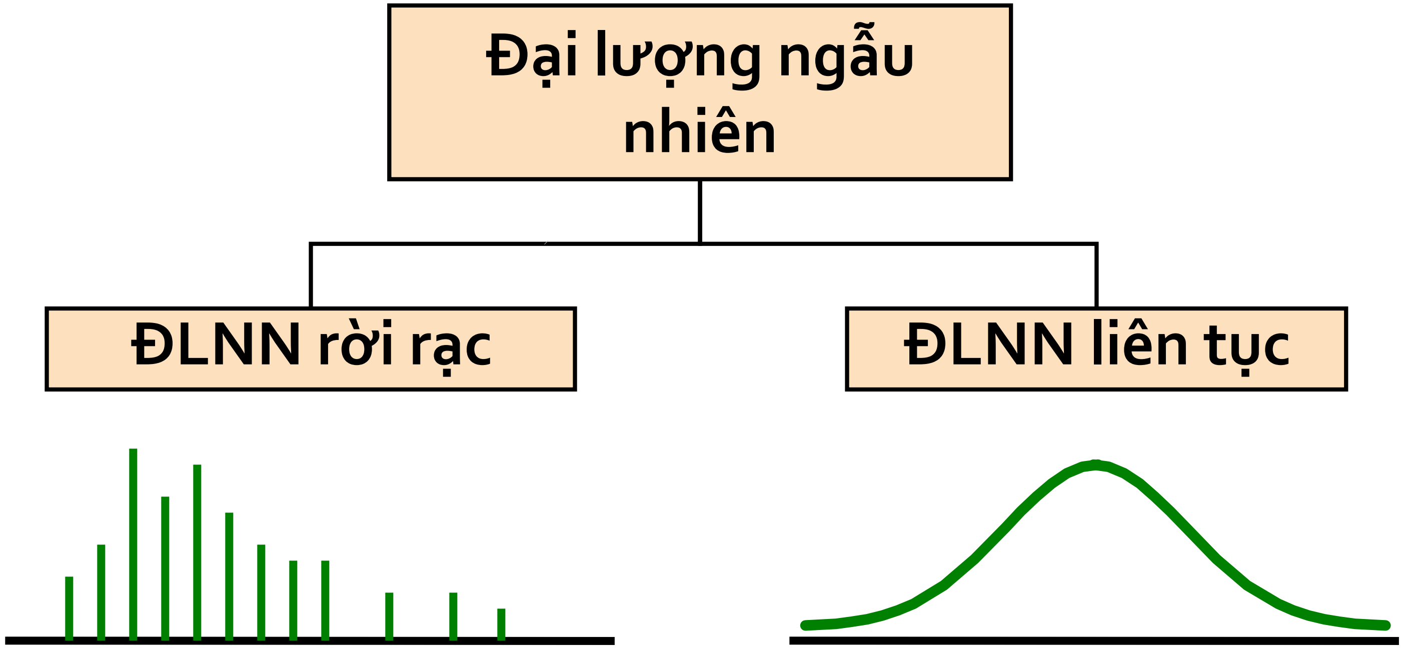
» X(Ω): tập hợp các giá trị có thể của ĐLNN X

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



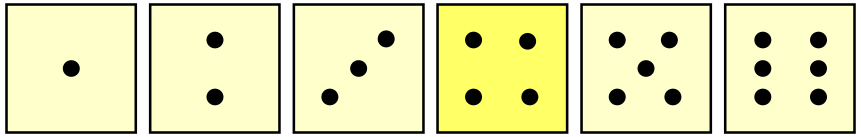
Đại lượng ngẫu nhiên rời rạc

» 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 4 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.

  Thì Y = 0, 1, 2, 3, 4, hoặc 5

Đại lượng ngẫu nhiên rời rạc

» Ví dụ

Tung một con xúc xắc cân đối và đồng chất

Đặt X = Số lần tung cho đến khi mặt 6 điểm xuất hiện.

X = 1, 2, …

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 $p_i = P(X = x_i)$. Lưu ý $p_1+p_2+\ldots+p_n = 1$.

» Hàm phân bố tích lũy (cumulative distribution function)

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

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 của X.

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

» C: Tung 1 con xúc sắc cân đối 2 lần

» X: Số lần ra mặt 4 chấm

» Lập bảng phân bố xác suất của X

X	?	?	?
P	?	?	?

» Xác định hàm F của X

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

- C: Tung 1 con xúc sắc cân đối 2 lần
- X: Số lần ra mặt 4 chấm
- Lập bảng phân bố xác suất của X

X	0	1	2
P	25/36	10/36	1/36

- Xác định hàm F của X

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

- Bảng phân bố xác suất
- Hàm phân bố xác suất

Ví dụ (tiếp)

Ví dụ 2. 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.

Ví dụ 3. Khi một người đi thi lấy bằng lái xe nếu không đạt anh ta đăng kí thi lại cho đến khi đạt mới thôi. Gọi X là số lần anh ta dự thi. Lập phân bố xác suất của X biết rằng xác suất thi đỗ của anh ta là 1/4.

Hãy dự đoán xem trong 1024 người (mỗi người đều có xác suất thi đỗ là ¼) có bao nhiêu người thi đạt ngay lần đầu, thi đạt ở lần thứ hai, phải thi ít nhất 4 lần.

Problem 1

Let X be a discrete random variable with the following PMF:

$$P_X(x) = \begin{cases} \frac{1}{2} & \text{for } x = 0 \\ \frac{1}{3} & \text{for } x = 1 \\ \frac{1}{6} & \text{for } x = 2 \\ 0 & \text{otherwise} \end{cases}$$

a. Find R_X, the range of the random variable X.
b. Find $P(X \geq 1.5)$.
c. Find $P(0 < X < 2)$.
d. Find $P(X = 0 | X < 2)$

Problem 2

Let X be the number of the cars being repaired at a repair shop. We have the following information:

- At any time, there are at most 3 cars being repaired.
- The probability of having 2 cars at the shop is the same as the probability of having one car.
- The probability of having no car at the shop is the same as the probability of having 3 cars.
- The probability of having 1 or 2 cars is half of the probability of having 0 or 3 cars.

Find the PMF of X.

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

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	0.7	0.2	0.05	0.02	0

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

EX = 17 x 0.03 + 18 x 0.7 + 19 x 0.2 + 20 x 0.05 + 21 x 0.02

Ví dụ

Bảng phân bố xác suất của lương SV Cơ-Kĩ thuật sau khi ra trường

X	<3	4	5	6	7	8	9	10	>10
P	0	0.06	0.1	0.5	0.2	0.1	0.02	0.02	0

Tính kì vọng của lương sinh viên Cơ-Kĩ thuật sau khi ra trường.

$EX = ?, EX^2 = ?$

» C: Tung 1 con xúc sắc cân đối 2 lần

» X: Số lần ra mặt 4 chấm

» Lập bảng phân bố xác suất của X

X	0	1	2
P	25/36	10/36	1/36

Ví dụ

Tính chất của kỳ vọng

1) EC = C, C: hằng số
2) E(CX) = C.EX, C: hằng số
3) $E(X \pm Y) = EX \pm EY$
4) E(XY) = EX.EY nếu X và Y độc lập
5) $Ef(X) = \sum_i f(x_i)p_i$ nếu $P(X = x_i) = p_i$

X	x_1	x_2	…	x_n
P	p_1	p_2	…	p_n

f(X)	$f(x_1)$	$f(x_2)$	…	$f(x_n)$
P	p_1	p_2	…	p_n

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:

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

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

Ví dụ

Lương của nhân viên 1 công ty ABC

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

Tính kì vọng, phương sai của lương nhân viên công ty ABC.

Công ty XYZ

Y	<3	4	5	6	7	8	9	10	>10
P	0	0.06	0.1	0.5	0.2	0.1	0.02	0.02	0

Ví dụ

Bảng phân bố xác suất của lương SV Cơ-Kĩ thuật sau khi ra trường

X	<3	4	5	6	7	8	9	10	>10
P	0	0.06	0.1	0.5	0.2	0.1	0.02	0.02	0

Tính phương sai, độ lệch chuẩn của lương sinh viên Cơ-Kĩ thuật sau khi ra trường.

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

Ví dụ

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	0.7	0.2	0.05	0.02	0

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

Tính chất của phương sai

1) $D(c)=0$, c:hằng số
2) $D(cX)=c^2DX$, c:hằng số

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

Thế nào là 2 biến ngẫu nhiên độc lập?

Luyện tập

1. Không đặt bút tính, hãy so sánh kì vọng và phương sai của 4 biến ngẫu nhiên X,Y,Z,W.

X	1	2	3	4	5
P	1/5	1/5	1/5	1/5	1/5

Y	1	2	3	4	5
P	1/10	2/10	4/10	2/10	1/10

Z	1	2	3	4	5
P	5/10	0	0	0	5/10

W	1	2	3	4	5
P	0	0	1	0	0

2. Tung 1 đồng xu cân đối đến khi thu được mặt ngửa (head). Tính kì vọng số lần tung phải mặt xấp (tail).

3. Tung 2 con xúc xắc cân đối. Bạn được 1000$ nếu tổng 2 con bằng 2 và mất 100$ nếu tổng khác. Bạn kì vọng mình sẽ thắng trung bình bao nhiêu $/lần nếu chơi rất nhiều lần.

4. Tung 2 con xúc xắc cân đối. Gọi X là tổng 2 mặt. Nếu hàm phần thưởng $Y=X^2-6X+1$ thì game này có lợi cho người chơi hay không?

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(Ω) = {0,1,…n}.

» Theo công thức Becnuli thì

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

» 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)

$$
\begin{aligned}
EX &= \sum_{k=0}^{n} k C_n^k p^k q^{n-k} = \\
&= \sum_{k=1}^{n} \frac{kn!}{k!(n-k)!} p^k q^{n-k} = n \sum_{k=1}^{n} \frac{(n-1)! p^k q^{n-k}}{(k-1)!(n-k)!} \\
&= np \sum_{k=1}^{n} C_{n-1}^{k-1} p^{k-1} q^{n-k} = np \sum_{i=0}^{n-1} C_{n-1}^{i} p^{i} q^{n-1-i} \\
&= np(p+q)^{n-1} = np.
\end{aligned}
$$

$$EX^2 = \sum_{k=0}^{n} k^2 C_n^k p^k q^{n-k} = \sum_{k=1}^{n} \frac{k^2 n!}{k!(n-k)!} p^k q^{n-k} = \sum_{k=1}^{n} \frac{k(k-1)n!}{k!(n-k)!} p^k q^{n-k} +$$

$$+ \sum_{k=1}^{n} \frac{kn!}{k!(n-k)!} p^k q^{n-k} = S_1 + S_2.$$

Ta có

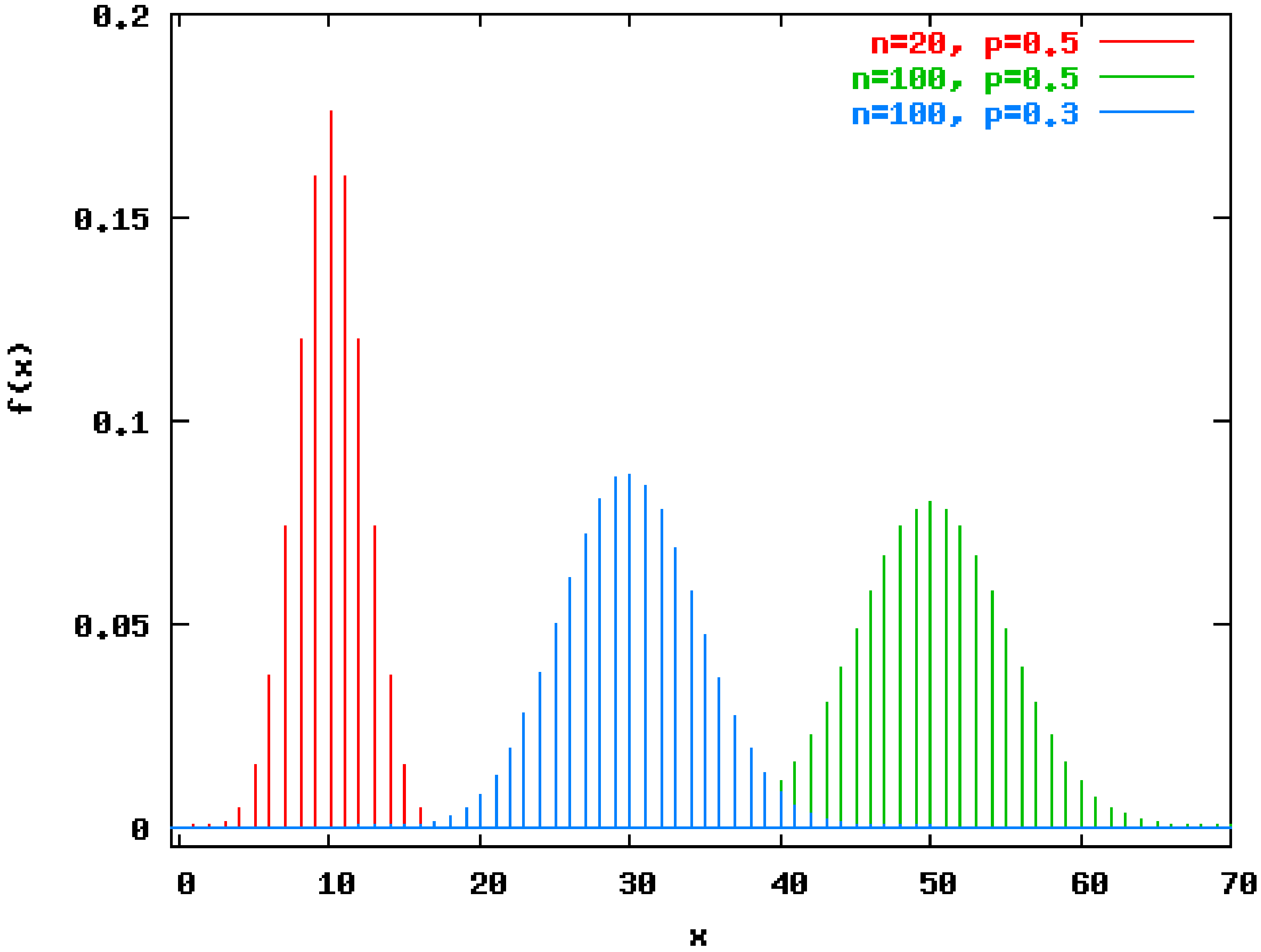
$$S_1 = \sum_{k=2}^{n} \frac{n!}{(k-2)!(n-k)!} p^k q^{n-k} = n(n-1)p^2$$

$S_2 = np$ (đã tính ở trên).

Thành thử

$$DX = EX^2 - (EX)^2 = n^2p^2 - np^2 + np - n^2p^2 = np(1-p) = npq.$$

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 P{X≤k}

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

- Với các giá trị k khác nhau được tính sẵn ở Bảng 1 (Phụ lục 2).
- Xác suất P(X) theo phân bố nhị thức có thể được tính bằng Excel
  - BINOMDIST(k, n, p, cumulative)
  - trong đó
  - cumulative = FALSE, thì tính P(X=k)
  - cumulative = TRUE, thì tính P(X <=k)

Ví dụ

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

Tỉ lệ sinh con thứ ba của 1 tỉnh M là 15%. Chọn ngẫu nhiên 10 gia đình và gọi X là số gia đình có con thứ 3.

a) Tính xác suất P{X=5}

b) Tính xác suất P{X <= 3}

c) Tính xác suất P{ X >= 5}

d) Tính kì vọng của X. (EX=?)

e) Tính phương sai của X. (DX=?)

Ví dụ

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

a) Tính P{X = 1}

b) Tính P{X = 2}

c) Tính P{X > 10}

d) Tính kì vọng của X

e) Tính phương sai của X

C: Chọn ngẫu nhiên 1 xe ô tô và xem nó có bị hỏng trong thời gian bảo hành hay không

A: ?

p: ?

n: ?

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.

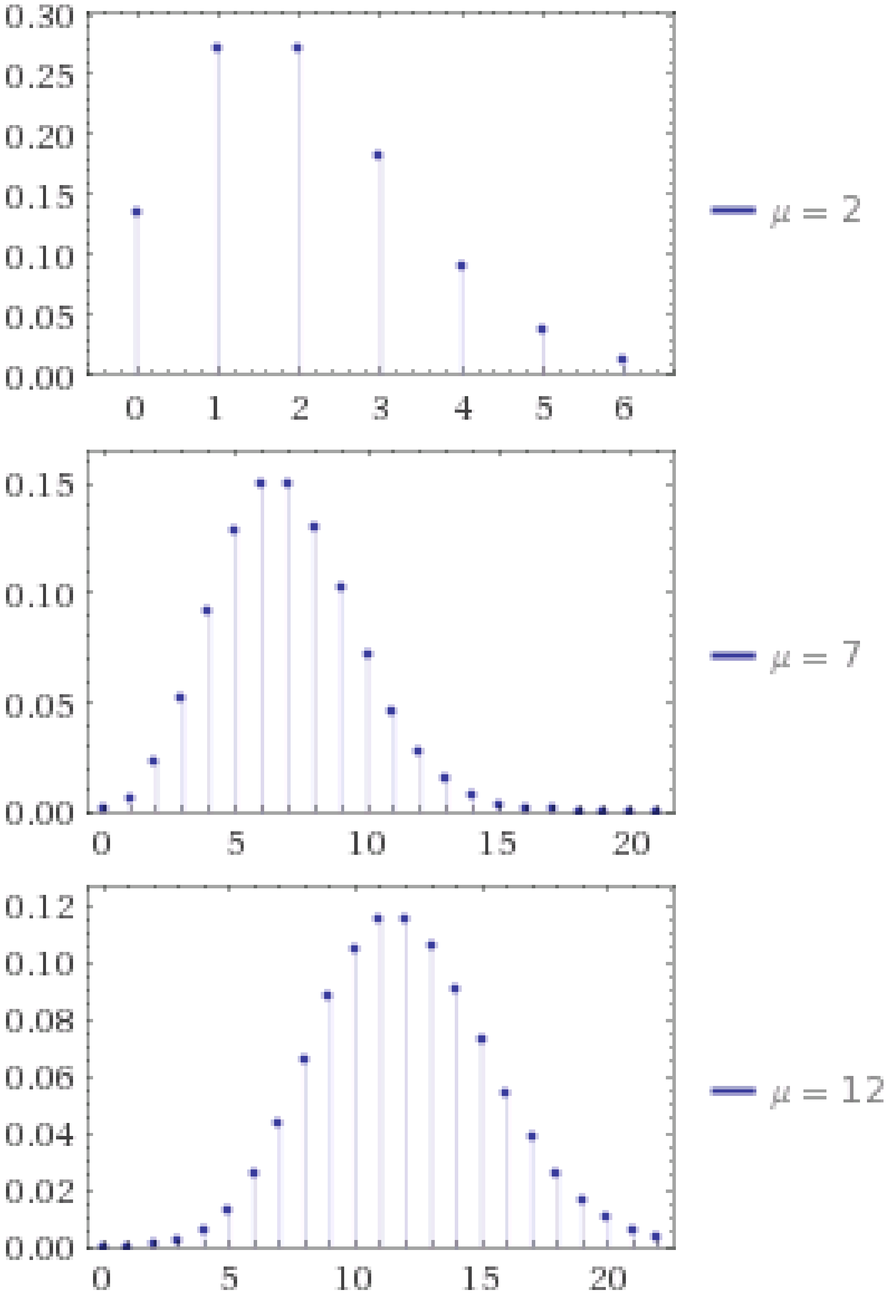
» Số lượng xuất hiện biến cố trong một khoảng thời gian khác nhau là độc lập với nhau.

» 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:

Trong đó μ là kì vọng của X.

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

Phân bố Poisson



Phân bố Poisson

» Xác suất tích lũy P{X <= k}

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

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

» Xác suất P(X) theo phân bố Poisson có thể được tính bằng Excel

- POISSON (k, μ, cumulative)
- trong đó
  - cumulative = FALSE, thì tính P(X=k)
  - cumulative = TRUE, thì tính P(X <=k)

Tra bảng ở trang 208-

» EX=6

a) $P(X=3)=$
b) $P(X=0)=$
c) $P(X<=4)=$
d) $P(X<3)=$
e) $P(X>1)=1- P(X<=1)$

Ví dụ

Một gara cho thuê ô tô thấy rằng số người đến thuê xe là một ĐLNN theo phân bố Poisson. Trung bình số người đến thuê xe vào thứ bảy là 3. Giả sử gara có 5 chiếc xe, hãy tính các xác suất sau đây:

a) Tất cả 5 xe đều được thuê

b) Không xe nào được thuê

c) Ít nhất 3 xe được thuê

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

a) P(X>=5)

b) P(X=0)

c) P(X>=3)

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

Ví dụ

Ở một tổng đài chăm sóc khách hàng, các cú 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 30 giây

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

Bài tập cuối chương (giáo trình)

Bài 14: Tại một trạm kiểm soát giao thông trung bình một phút có 2 xe ô tô đi qua.

a) Tính xác suất để có đúng 6 xe đi qua trong vòng 3 phút.

b) Tính xác suất để trong khoảng thời gian ***t*** phút có ít nhất 1 xe ô tô đi qua. Xác định ***t*** để xác suất này bằng 0.99.

Bài 16: Một trạm cho thuê xe taxi có 3 chiếc xe. Hàng ngày trạm phải nộp thuế 8 USD cho 1 chiếc xe (dù xe đó có được thuê hay không). Mỗi chiếc xe được cho thuê với giá 20 USD. Một xe chỉ được thuê tối đa 1 lần trong ngày.

Giả sử số yêu cầu thuê xe của trạm trong 1 ngày là ĐLNN X có phân bố Poisson với $\mu = 2.8$.

a) Gọi Y là số tiền thu được trong 1 ngày của trạm (nếu không có ai thuê thì số tiền thu được là -24 USD). Tìm phân bố xác suất của Y. Từ đó, tính số tiền trung bình thu được của trạm trong 1 ngày.

b) Giải bài toán trên trong trường hợp trạm có 5 chiếc xe.

c) Trạm nên có 3 hay 5 chiếc xe?

Luyện tập

Bài 67 (BT): Trong một cuộc sổ số người ta phát hành 10 vạn vé trong đó có 1 vạn vé trúng giải. Cần phải mua ít nhất bao nhiêu vé để với xác suất không nhỏ hơn 0.95 ta sẽ trúng ít nhất 1 vé.

Đề thi cũ

Theo thống kê của một hãng hàng không, trung bình 3% khách đã mua vé sẽ không có mặt check-in. Do đó hãng này bán vé vượt số ghế trên loại máy bay 100 ghế. Tính số vé vượt mà hãng có thể bán (cho mỗi chuyến bay) để xác suất thiếu ghế không quá 5%.

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

(Tiết 3, trang 51 giáo trình)

X và Y là hai ĐLNN rời rạc với

$X(\Omega) = \{x_1,\ldots,x_m\}$ và

$Y(\Omega) = \{y_1,\ldots,y_n\}$.

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

Ba đồng tiền cân đối A, B, C được gieo. 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

X và Y có độc lập hay không?

Ví dụ

Ba đồng tiền cân đối A, B, C được gieo. 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

a) Hãy lập bảng phân bố đồng thời của X và Y.
b) Lập bảng phân bố xác suất của X. (Tính P(X=0), P(X=1), P(X=2).)
c) Lập bảng phân bố xác suất của Y

Nếu X và Y độc lập thì (X=xi) độc lập với (Y=yj) với mọi cặp (i,j)

hay P(X=xi và Y=yj) = P(X=xi).P(Y=yj)

vế trái = ô giao của hàng i cột j

vế phải = (tổng của hàng i) x (tổng của cột j)

Chuẩn bị bài tới

- Đọc Chương 3 giáo trình
- Hoàn thành bài tập gửi qua email