Đại lượng ngẫu nhiên liên tục

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

Nội dung

- Đại lượng ngẫu nhiên liên tục
- Hàm mật độ xác suất và hàm phân bố tích lũy
- Kì vọng, Phương sai
- Phân bố đều
- Phân bố chuẩn
- Phân bố mũ
- Đại lượng ngẫu nhiên liên tục nhiều chiều (tự đọc)

Biến ngẫu nhiên liên tục

Tập các giá trị có thể lấp đầy một hay một số khoảng của trục số, thậm chí lấp đầy toàn bộ trục số.

Ví dụ

- Chiều cao, cân nặng.
- Thời gian để hoàn thành 1 công việc.

Hàm mật độ xác suất

f(x) gọi là hàm mật độ xác suất của biến ngẫu nhiên liên tục X nếu

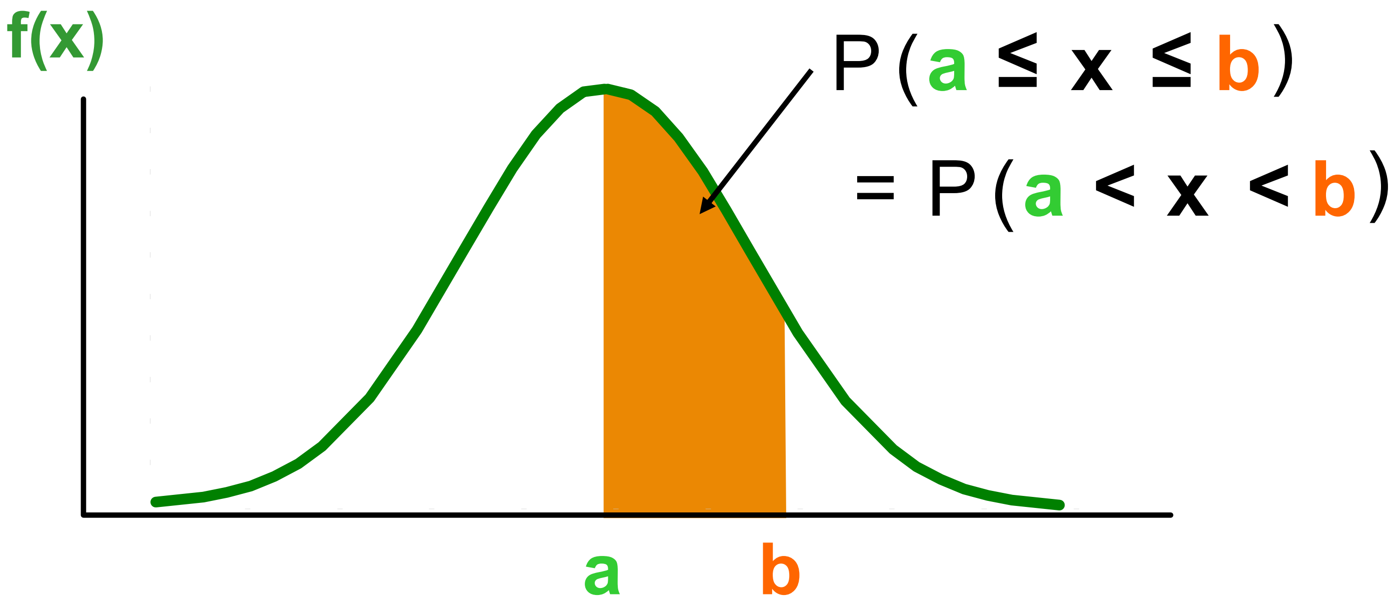
$$\boxed{\begin{aligned} &i)\, f(x) \geq 0 \quad \forall x \\ &ii) \int_{-\infty}^{+\infty} f(x)dx = 1 \end{aligned}}$$

Ví dụ: Biến ngẫu nhiên liên tục X với hàm mật độ xác suất

$$f(x) = \begin{cases} 2x & 0 \leq x \leq 1 \\ 0 & , \neq \end{cases}$$

Biến ngẫu nhiên liên tục

Tìm P(a<X<b)?



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

Lưu ý: P(a <= x <= b) = P(a < x < b)

Hàm phân phối tích lũy

- Xét biến ngẫu nhiên X, hàm phân phối tích lũy của X, ký hiệu F(x), được định nghĩa như sau

$$F(x) = P\left(X \le x\right)$$

- Xác suất X thuộc [a,b]

$$P(a \le X \le b) = F(b) - F(a)$$

Tính chất hàm phân phối tích lũy

- 1) $0 \leq F(x) \leq 1$
- 2) F(x) là hàm không giảm: nếu a<b thì F(a) ≤ F(b).
- 3) $$F(-\infty) = \lim_{x \to -\infty} F(x) = 0$$
  $$F(+\infty) = \lim_{x \to +\infty} F(x) = 1$$

Nếu X là biến ngẫu nhiên liên tục có hàm phân phối tích lũy F(x) thì hàm mật độ f(x) = F' (x) tại những điểm liên tục của X.

Ví dụ 1

Giả sử X có giá trị trong đoạn [0,2] và hàm mật độ xác suất $f(x) = cx^2$.

a) Tính giá trị của c

b) Tính hàm phân bố tích lũy F(x)

c) Tính $P(1 \leq X \leq 2)$

Ví dụ 2

Giả sử X có giá trị trong đoạn [0,b] và hàm phân phối tích lũy $F(x) = x^2/9$.

a) Tính giá trị của b
b) Tính hàm mật độ xác suất f(x)

Kỳ vọng của biến ngẫu nhiên liên tục

Xét biến ngẫu nhiên liên tục X có hàm mật độ xác suất f(x).

Kỳ vọng của X:

$$EX = \int_{-\infty}^{+\infty} xf(x)dx$$

Ví dụ: Biến ngẫu nhiên liên tục X với hàm mật độ xác suất

$$f(x) = \begin{cases} x/2 & 0 \le x \le 2 \\ 0 & ,\neq \end{cases}$$

Tính kỳ vọng EX.

Tính chất của kỳ vọng

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

Phương sai của biến ngẫu nhiên liên tục

Xét X là biến NNLT có hàm mật độ xác suất f(x), kỳ vọng μ = EX.

Phương sai, kí hiệu DX hay VarX

$$VarX = E(X - EX)^2 = \int_{-\infty}^{+\infty} (x-\mu)^2 f(x)\,dx$$

hoặc

$$VarX = EX^2 - (EX)^2 = \int_{-\infty}^{+\infty} x^2 f(x)dx - \mu^2$$

Tính chất của phương sai

1. $Var(c)=0$, c:hằng số
2. $Var(cX)=c^2VarX$;
3. $Var(X+c)=VarX$
4. $Var(X + Y) = VarX + VarY$ nếu X và Y độc lập.

Ví dụ 3

Giả sử X có giá trị trong đoạn [0,2] và hàm mật độ xác suất $f(x) = cx^2$.

a) Tính kì vọng EX

b) Tính phương sai DX

Ví dụ 4

Giả sử X nằm trong đoạn [0,3] với hàm mật độ $f(x) = cx^3$. Hãy tìm:

a) Hằng số c

b) Kì vọng

c) Phương sai và độ lệch chuẩn

d) *Median:* Giá trị m được gọi là median của ĐLNN X nếu

$$P\{X<m\} = P\{X > m\} \quad \text{hay } F(m) = 1/2$$

Nội dung

- Đại lượng ngẫu nhiên liên tục
- Hàm mật độ xác suất và hàm phân bố tích lũy
- Kì vọng, Phương sai
- Phân bố đều
- Phân bố chuẩn
- Phân bố mũ
- Đại lượng ngẫu nhiên liên tục nhiều chiều (tự đọc)

Phân phối đều

Một ĐLNN liên tục X có phân phối đều (uniform distribution) trong đoạn [a,b] nếu và chỉ nếu hàm mật độ xác suất f(x) có dạng sau

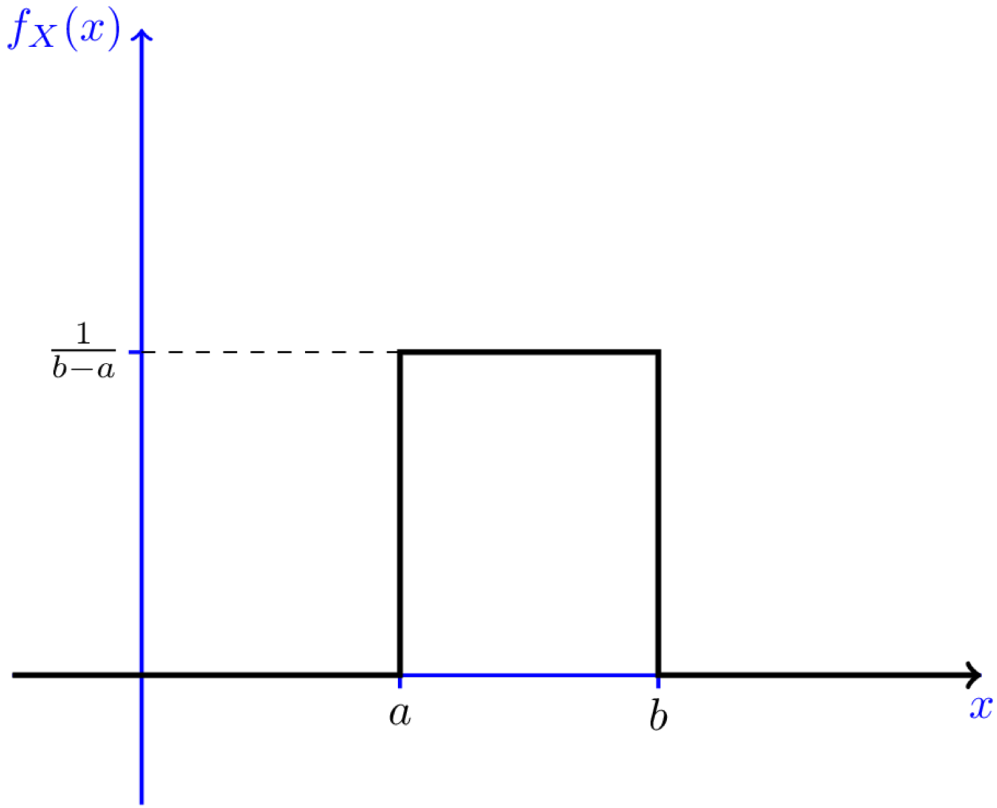
f(x, a, b) = 1/(b-a); nếu a<= x <= b

= 0; ngược lại

Ví dụ: RAND () là phân phối đều trong đoạn [0,1].

- Tính kì vọng EX?
- Tính phương sai DX?

Hàm mật độ của phân phối đều



Ví dụ 6

ĐLNN X có phân bố đều trên đoạn [2,5]. Hãy tính

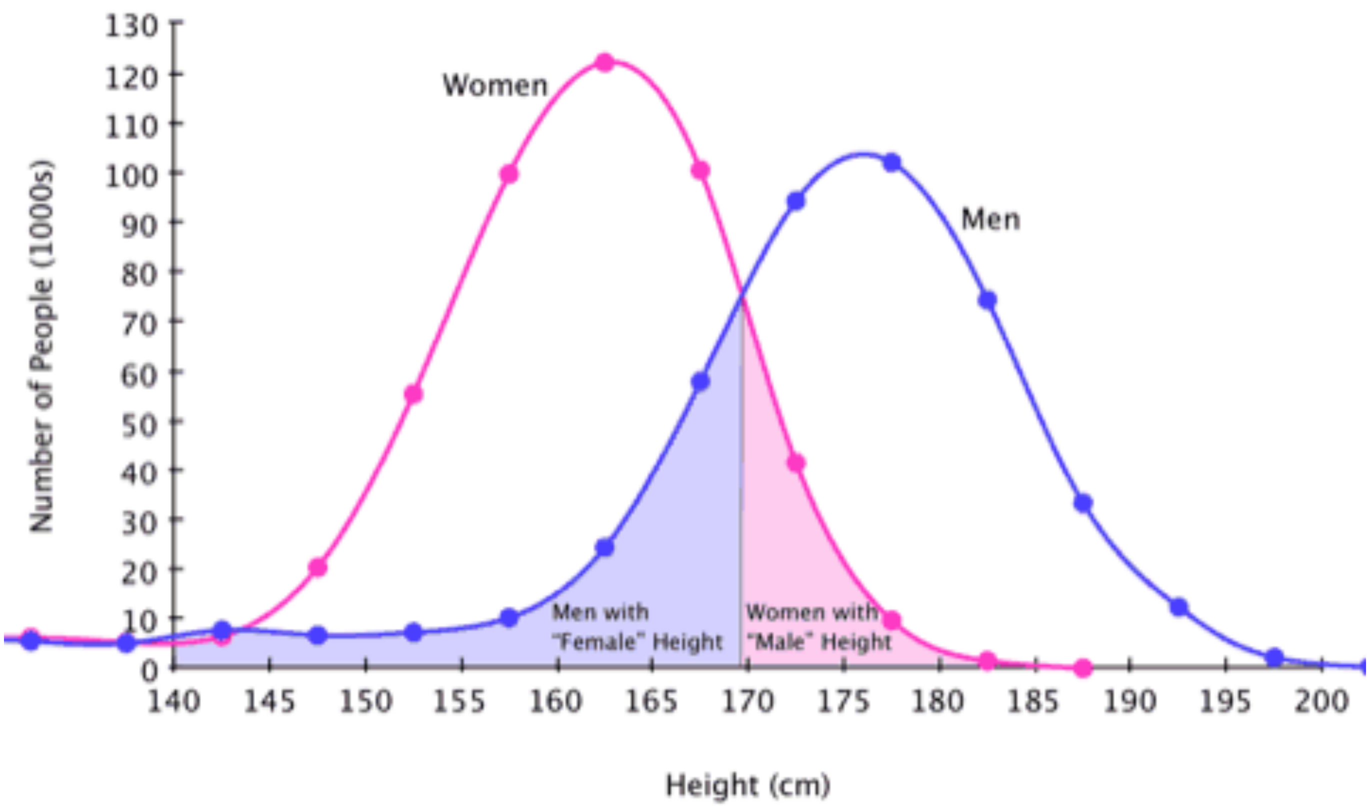
a) $P(X < 3)$

b) $P(X > 4)$

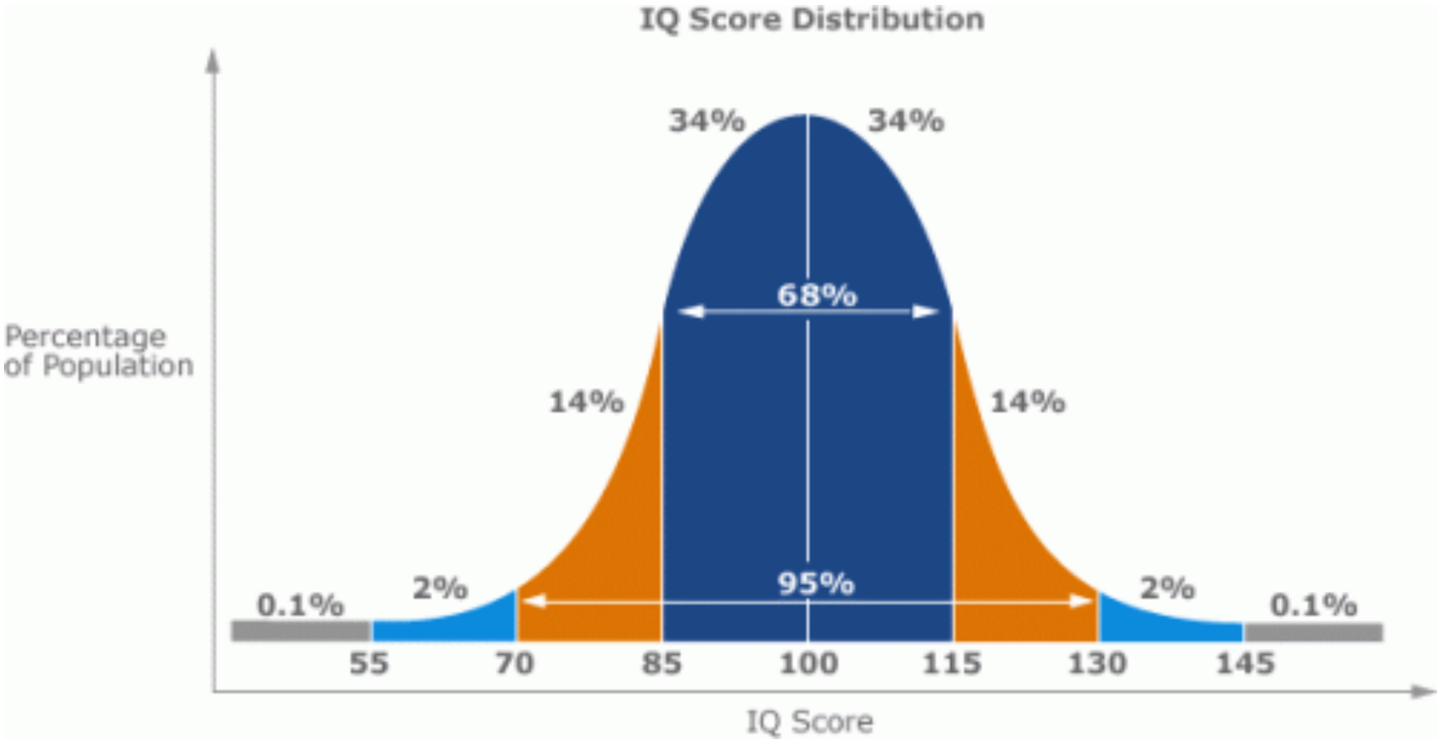
c) $P(3.5 < X \leq 7)$

d) Tính kì vọng, phương sai của X.

Phân bố chuẩn
normal/Gaussian distribution



Phân bố chuẩn normal/Gaussian distribution



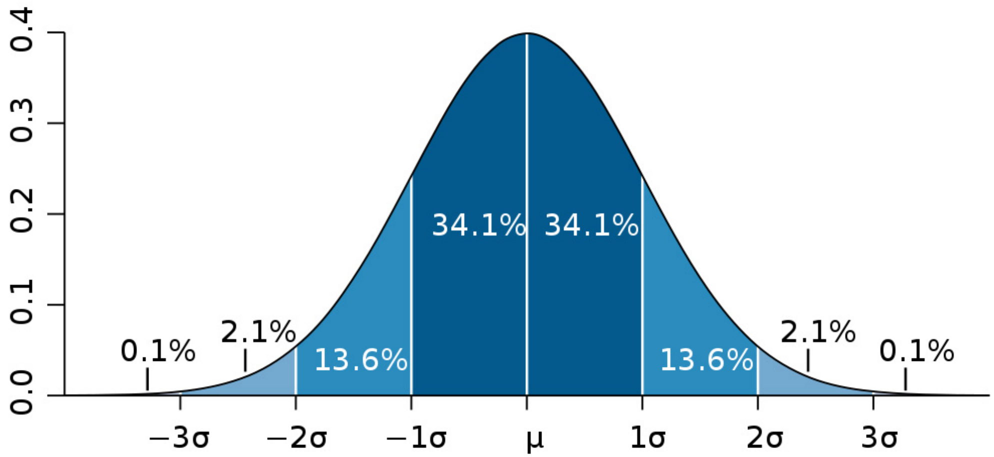
Phân bố chuẩn normal/Gaussian distribution

Hàm mật độ f(x)

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-[(x-\mu)/\sigma]^2/2}$$

Trong đó:

- μ là kì vọng
- σ là độ lệch chuẩn
- Kí hiệu X ~ N(μ, σ^2)
- Kì vọng, median, và mode cùng một giá trị
- Phân bố là đường cong đối xứng qua giá trị kì vọng
- Hai đuôi của phân bố kéo dài đến vô cùng

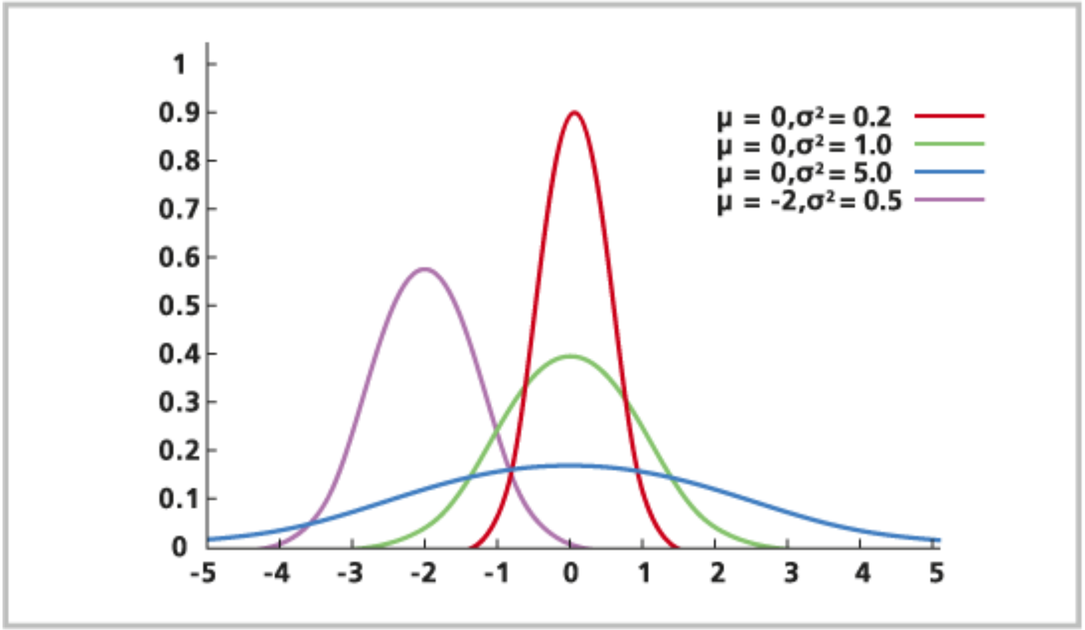


Phân bố chuẩn tắc
standard normal distribution

ĐLNN X *có phân bố chuẩn tắc* nếu

X phân bố chuẩn với μ = 0, σ = 1

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$



Tính xác suất theo phân bố chuẩn normal/Gaussian distribution

- Gọi X có phân bố chuẩn $N(\mu, \sigma^2)$
- $Z = (X-\mu) / \sigma$: Số lần độ lệch chuẩn giữa X và μ.

  Z có phân bổ chuẩn tắc hay $Z \sim N(0, 1)$

- $P(X < x) = P(Z < z)$. Giá trị $P(Z < z)$ đã được tính sẵn trong bảng.

Ví dụ 7: Giả sử X là ĐLNN có phân bố chuẩn với kì vọng 2100 và độ lệch chuẩn 200. Hãy tính

1. $P\{X > 2400\}$
2. $P\{1700 < X < 2200\}$
3. Xác định a để $P\{X > a\} = 0.03$

Ví dụ 8

Tốc độ của xe ô tô qua 1 điểm kiểm tra tốc độ là một phân phối chuẩn với kì vọng 60km/giờ và độ lệch chuẩn là 5km/giờ. Tính xác suất để tốc độ một chiếc xe sẽ đi qua điểm kiểm tra:

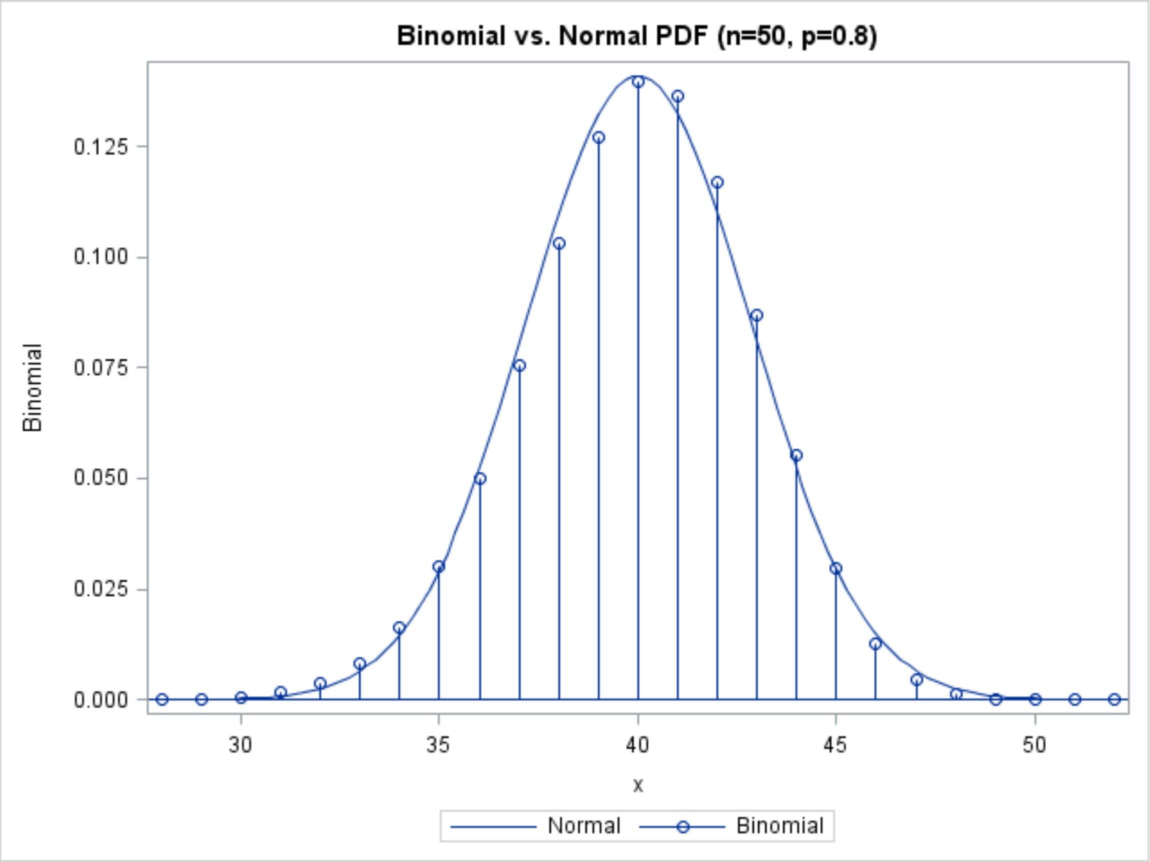
1. Nhỏ hơn 60km/giờ
2. Lớn hơn 70km/giờ
3. Từ 60-65km/giờ

Ví dụ 9

Lương một sinh viên CNTT ra trường có phân bố chuẩn với kì vọng 9 triệu và độ lệch chuẩn là 4 triệu. Tính xác suất lương một sinh viên

1. <6 triệu
2. 6-12 triệu
3. >12 triệu

Xấp xỉ phân bố nhị thức bằng phân bố chuẩn



Xấp xỉ phân bố nhị thức bằng phân bố chuẩn

- ĐLNN X ~ B (p, n) thì $P(X = k) = C^k_n\, p^k q^{n-k}$
- X có phân bố xấp xỉ X' ~ N(np, npq) khi np và nq lớn hơn 5 hoặc khi npq lớn hơn 20.

  Lưu ý: EX'=np, DX'=npq
- Hiệu chỉnh để giảm sai số:

  P{k1 <= X <=k2) được xấp xỉ bởi P(k1-0.5 < X' < k2 + 0.5)

Ví dụ 10: Một kí túc xá có 650 sinh viên. Xác suất 1 sinh viên đi xem phim vào tối thứ bảy là 0.7.

a) Tính xác suất để số sinh viên đi xem vào tối thứ bảy ít hơn 470

b) Cần phải chuẩn bị bao nhiêu ghế để có thể đảm bảo với xác suất 95% đủ ghế cho người xem.

Nhắc lại phân bố Poisson

- Xác suất tích lũy

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

- Ví dụ:
  - Số lượng khách vào một cửa hàng trong 1 giờ theo phân bố Poisson
  - Số lượng bệnh nhân cấp cứu một đêm tại một bệnh viện theo phân bố Poisson.

Phân bố mũ

Với một giả thiết nào đó, khoảng thời gian giữa hai lần xuất hiện một biến cố E sẽ theo phân bố mũ.

Ví dụ:

- Khoảng thời gian giữa hai ca cấp cứu
- Khoảng thời gian giữa hai lần hỏng một máy tính
- Tuổi thọ trung bình của một máy tính

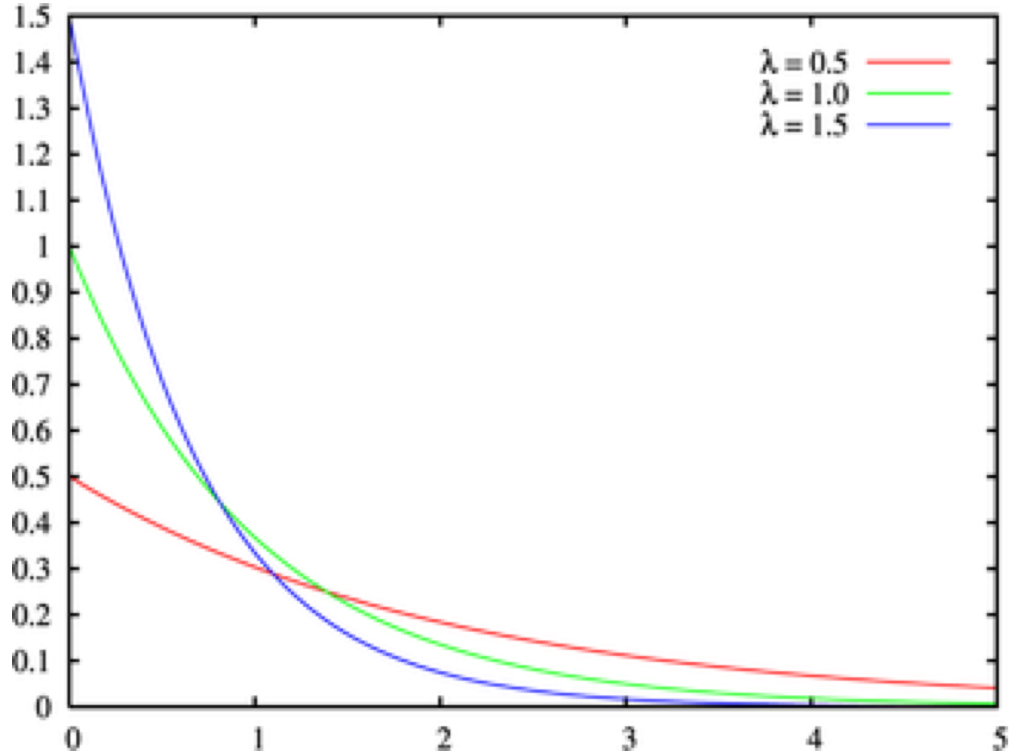
Phân bố mũ

Hàm mật độ

$$f(x) = \begin{cases} \lambda e^{-\lambda x}; x > 0 \\ 0; x < 0 \end{cases}$$

Hàm phân bố tích lũy

$$F(x) = \begin{cases} 1 - e^{-\lambda x}; x > 0 \\ 0; x \le 0 \end{cases}$$



Kì vọng và độ lệch chuẩn đều bằng 1/λ

Ví dụ 11

- Tuổi thọ trung bình của một mạch điện có phân bố mũ là 6.5 năm. Trong thời gian 5 năm bảo hành, có bao nhiêu % mạch điện tử bị hỏng?

- Trung bình có 5 bệnh nhân xuất hiện trong 1 tiếng tại bệnh viện theo phân bố mũ. Một bệnh nhân vừa xuất hiện, tính xác suất bệnh nhân tiếp theo xuất hiện:
  - Trong vòng 10 phút
  - Trong vòng 20 phút
  - Không có bệnh nhân nào xuất hiện trong vòng 15 phút
  - Không có bệnh nhân nào xuất hiện trong vòng 30 phút

Ví dụ 12

Trung bình 1 năm có 12 trận mưa to tại Hà Nội và theo phân bố mũ. Một trận mưa to vừa diễn ra cách đây 2 tuần. Tính xác suất:

- Trận mưa tiếp theo diễn ra hôm nay
- Trận mưa tiếp theo diễn ra trong vòng 1 tuần
- Trận mưa tiếp theo diễn ra trong vòng 1 tháng
- Không có trận mưa nào diễn ra trong vòng 2 tháng