MÔN HỌC

THỐNG KÊ ỨNG DỤNG - XD (KC107)

GIÁO VIÊN BIÊN SOẠN

ĐẶNG THẾ GIA

Bộ môn Kỹ Thuật Xây Dựng
Khoa Công Nghệ, Trường Đại Học Cần Thơ

Chương 6 & 7:

PHÂN PHỐI XÁC SUẤT

PROBABILITY DISTRIBUTION

Nội dung chương

1. **Luật phân phối xác suất**
2. **Đặc trưng của phân phối xác suất**
3. **Phân loại các phân phối xác suất**
4. **Phân phối rời rạc điển hình**
5. **Phân phối liên tục điển hình**
6. **Các bảng tra**

1. Luật phân phối xác suất

Hàm phân phối xác suất

- Một **phân phối xác suất** hay thường gọi hơn là một **hàm phân phối xác suất** là một mô tả toán học của một hiện tượng ngẫu nhiên thông qua khái niệm xác suất.
- Luật phân phối xác suất của biến X có thể được mô tả một cách duy nhất bởi hàm phân phối lũy tích F(x) (cumulative distribution function, CDF) được định nghĩa như sau:

  $F(x) = P(X \leq x)$ với mọi x là số thực (R)

$$Biên \quad Rời \quad Rac: F(x) = \sum_{x_i < x} p_i$$

$$Biên \quad Liên \quad Tuc: F(x) = \int_{-\infty}^{x} fx(t)dt$$

Ý nghĩa & Tính chất

- **Hàm phân phối xác suất** là quy luật cho biết cách gán mỗi xác suất cho mỗi khoảng giá trị của tập số thực, sao cho các tiên đề xác suất (Probability axioms) được thỏa mãn.
- Hàm phân phối xác suất phản ánh mức độ tập trung xác suất về phía trái điểm X.
  - $0 \leq F(x) \leq 1$, với mọi x
  - $F(-\infty) = 0$ và $F(+\infty) = 1$
  - F(x) là hàm số không giảm
  - $P(a \leq X < b) = F(b) – F(a)$
  - Nếu X là biến liên tục thì $F'(x) = f(x)$

Hàm mật độ xác suất

- **Hàm mật độ xác suất** của đại lượng ngẫu nhiên liên tục X ký hiệu là f(x) là đạo hàm bậc nhất của hàm phân phối xác suất của đại lượng ngẫu nhiên đó: $f(x) = F'(x)$.

$$\text{Biên rời rac}: f(x) = \begin{cases} p_i & \text{khi } x = x_i, i = 1,2,..,n \\ 0 & \text{khi } x \neq x_i \end{cases}$$

$$\text{Biên liên tuc}: f(x) = F'(x)$$

Phân phối rời rạc & Phân phối liên tục

- Một phân phối được gọi là **rời rạc** nếu hàm phân phối tích lũy của nó bao gồm một dãy các bước nhảy hữu hạn, hoặc vô hạn đếm được, cách quảng nhau.
- Do vậy phân phối rời rạc được sinh ra từ một biến ngẫu nhiên rời rạc X (một biến chỉ có thể nhận giá trị trong một tập hợp hữu hạn hoặc đếm được nhất định).
- Một phân phối được gọi là **liên tục** nếu hàm phân phối tích lũy của nó là hàm liên tục, tức là tập giá trị của biến ngẫu nhiên lắp đầy một khoảng hay toàn bộ trục số thực.
- Khi đó nó sinh ra từ một biến ngẫu nhiên X mà $P(X=x_0) = 0$ với mọi x thuộc R.

2. Đặc trưng của một phân phối xác suất

- Hàm mật độ xác suất
- Hàm phân phối xác suất
- Giá trị kỳ vọng (giá trị trung bình)
- Trung vị
- Giá trị thường gặp
- Phương sai
- Độ xiên
- Độ nhọn
- Entropy
- Hàm sinh moment
- Hàm đặc trưng

Kỳ vọng toán

- Cho một biến ngẫu nhiên X, kỳ vọng toán của X là:

$$Biên \quad Rời \quad Rac: E(X) = \sum_{all\ x_i} x_i \cdot p(x_i)$$

$$Biên \quad Liên \quad Tuc: E(X) = \int_{-\infty}^{+\infty} x.f(x)dx$$

với $p(x_i)$ là xác suất của giá trị x_i

- Kỳ vọng toán của biến ngẫu nhiên X là bình quân gia quyền (weighted average) của các giá trị khả dĩ của X, khi đó trọng số (gia quyền) tương ứng với xác suất của mỗi x_i.
- Kỳ vọng toán của biến ngẫu nhiên là con số đăc trưng cho giá trị bình quân của biến ngẫu nhiên đó.

Kỳ vọng toán – Tính chất

- E(c) = c
- E(c*X) = c*E(X)
- E(X + Y) = E(X) + E(Y)
- E(X - Y) = E(X) - E(Y)
- E(X*Y) = E(X)*E(Y) nếu X và Y là các biến ngẫu nhiên độc lập

Giá trị thường gặp

- Biến rời rạc: Là giá trị của biến ngẫu nhiên mà tại đó nó có **xác suất lớn nhất**
- Biến liên tục: Là giá trị của biến ngẫu nhiên mà tại đó **hàm mật độ đạt giá trị cực đại**

Phương sai

- Gọi X là một biến ngẫu nhiên rời rạc, phương sai của X là:

$$Tông\quad quát: V(X) = E[X - E(X)]^2$$

$$Biên\quad Rời\quad Rac: V(X) = \sum_{\forall\, x_i} x^2 p_i - [E(X)]^2 = \sum_{\forall\, x_i} (x_i - \mu)^2 p(x_i)$$

$$Biên\quad Liên\quad Tuc: V(X) = \int_{-\infty}^{+\infty} x^2 . f(x) dx - [E(X)]^2$$

với giá trị x_i có xác suất $p(x_i)$, và $E(x_i)=\mu$

Phương sai – Tính chất & Ý nghĩa

- V(C) = 0
- V(C*X) = C^2*V(X)
- V(X$\pm$Y) = V(X) + V(Y) nếu X và Y là các biến ngẫu nhiên độc lập
- Phương sai của biến ngẫu nhiên X là bình quân gia quyền (weighted average) của bình phương các độ lệch của các biến x_i so với giá trị bình quân μ, khi đó trọng số (gia quyền) tương ứng với xác suất của mỗi x_i.

Độ lệch chuẩn

- Độ lệch chuẩn của biến ngẫu nhiên rời rạc, ký hiệu σ(X), là căn (dương) bậc hai của phương sai: $\sigma(X) = \sqrt{V(X)}$

Ví Dụ

- Tổng số lô vật liệu sẽ được bán trong tuần tới với xác suất như sau:

x	0	1	2	3	4
p(x)	.05	.15	.35	.25	.20

- Xác định giá trị kỳ vọng và độ lệch chuẩn?

$$E(X) = \mu = \sum_{i=1}^{5} x_i p(x_i)$$

x	0	1	2	3	4
p(x)	.05	.15	.35	.25	.20

$$= 0(0.05) + 1(0.15) + 2(0.35) + 3(0.25) + 4(0.20)$$
$$= 2.40$$

$$V(X) = \sigma^2 = \sum_{i=1}^{5} (x_i - 2.4)^2 p(x_i)$$
$$= (0-2.4)(.05) + (1-2.4)(.15) + (2-2.4)(.35)$$
$$+ (3-2.4)(.25) + (4-2.4)(.20) = 1.24$$
$$\sigma = \sqrt{1.24} = 1.11$$

Ví dụ

- Giả sử xác suất số lô vật liệu bán trong tuần tới như trong ví dụ trước. Tiền lương tuần của nhân viên là 150 ngàn VNĐ cộng thêm 200 ngàn VNĐ tiền thưởng cho mỗi lô vật liệu bán được.
- Tính giá trị kỳ vọng và phương sai cho số tiền mà nhân viên có thể nhận?

Giải:

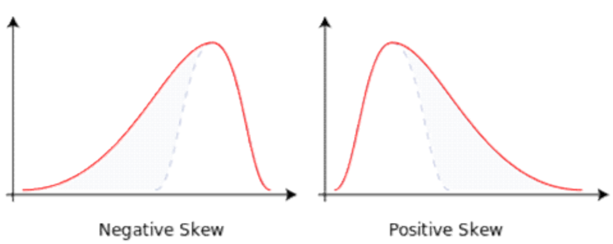
- Số tiền nhận được trong tuần: Y = 200X + 150

E(Y) = E(200X+150) = 200E(X)+150= 200(2.4)+150=630 $

V(Y) = V(200X+150) = 200^2V(X) = 200^2 (1.24) = 49,600 $^{2}

Độ xiên (*Skewness*) – Định nghĩa

- **Độ xiên** là một đại lượng đo lường mức độ **mức độ bất đối xứng** của phân phối xác suất của một biến ngẫu nhiên. Nó còn tên gọi nữa là **hệ số bất đối xứng**.



Độ xiên (*Skewness*) – Công thức

Để tính ra giá trị độ xiên, có thể sử dụng công thức:

$$\gamma_1 = \frac{\mu_3}{\sigma^3}$$

trong đó γ_1 là độ xiên hay moment chuẩn hóa

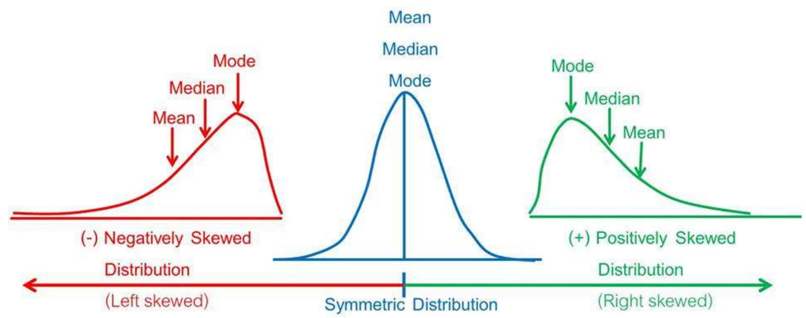
μ_3 là tâm moment thứ ba

σ là độ lệch chuẩn

Độ xiên (*Skewness*) – Tính chất

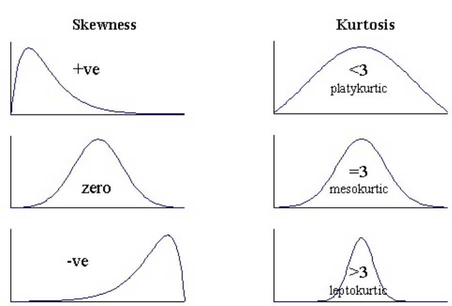
- Nếu hệ số này bằng 0, thì phân phối là cân xứng. Các số bình quân, trung vị và giá trị thường gặp (mode) bằng nhau.
- Nếu hệ số này lớn hơn 0, thì phân phối nghiêng dương. Số giá trị thường gặp (mode) nhỏ hơn số trung vị, và số trung vị lại nhỏ hơn số bình quân.
- Nếu hệ số này nhỏ hơn 0, thì phân phối nghiêng âm. Số bình quân nhỏ hơn số trung vị, và số trung vị nhỏ hơn số giá trị thường gặp (mode).

Độ xiên (*Skewness*) – Tính chất



Độ nhọn (*Kurtosis*) – Định nghĩa

- Độ nhọn là một đại lượng thống kê mô tả đo **mức độ tập trung** của phân phối xác suất của một biến ngẫu nhiên, cụ thể là mức độ tập trung của các quan sát quanh trung tâm của phân phối trong mối quan hệ với hai đuôi.



Platy: Rộng, phẳng

Meso: Trung

Lepto: Nhỏ, hẹp

Độ nhọn (*Kurtosis*) – Công thức

Cách tính độ nhọn có công thức:

$$\gamma_2 = \frac{\mu_4}{\sigma^4}$$

trong đó γ_2 là giá trị độ nhọn,

μ_4 là tâm moment thứ tư

δ là độ lệch chuẩn

Độ ngọn (*Kurtosis*) – Tính chất

- Khi γ_2 nhỏ hơn 3, phân phối tập trung kém mức bình thường; đỉnh của đồ thị hình chuông của phân phối thấp và tù hơn, với 2 đuôi dài hơn.
- Khi γ_2 bằng 3, phân phối tập trung ở mức độ bình thường.
- Khi γ_2 lớn hơn 3, phân phối tập trung hơn mức bình thường; đỉnh của đồ thị hình chuông của phân phối cao và nhọn trong khi 2 đuôi ngắn hơn.

Tâm moment thứ n – Định nghĩa

$$\mu_n = \mathrm{E}[(X - \mathrm{E}[X])^n] = \int_{-\infty}^{+\infty} (x - \mu)^n f(x)\,\mathrm{d}x$$

- Tâm moment thứ zero (n=0), $\mu_0 = 1$
- Tâm moment thứ nhất (n=1), $\mu_1 = 0$ (không phải mean, μ)
- Tâm moment thứ hai (n=2), $\mu_2 = \sigma^2$ (phương sai)
- Tâm moment thứ ba (μ_3) và thứ tư (μ_4) dùng để tính độ xiên và độ nhọn.

Tâm moment thứ n – Tính chất

Với mọi n và C ta có:

$$\mu_n(X + c) = \mu_n(X)$$
$$\mu_n(cX) = c^n \mu_n(X)$$

Với n = {1,2,3} ta có:

$$\mu_n(X + Y) = \mu_n(X) + \mu_n(Y)$$

3. Phân loại các phân phối xác suất

Phân phối xác suất rời rạc

- **Biến có giá trị hữu hạn:**
  - **Phân phối Bernoulli**
  - Phân phối Rademacher
  - **Phân phối nhị thức (binomial distribution)**
  - Phân phối suy biến (degenerate distribution)
  - Phân phối đều rời rạc (discrete uniform distribution)
  - Phân phối siêu bội (hypergeometric distribution)
  - Phân phối Zipf
  - Phân phối Zipf-Mandelbrot

Phân phối xác suất rời rạc

- **Biến có giá trị vô hạn:**
  - Phân phối Boltzmann (các trường hợp đặc biệt gồm có: Phân phối Gibbs, Phân phối Maxwell-Boltzmann, Phân phối Bose-Einstein, Phân phối Fermi-Dirac)
  - Phân phối hình học
  - Phân phối lôga
  - Phân phối nhị thức âm (một suy rộng của phân phối hình học)
  - Phân phối bật hai phân dạng
  - **Phân phối Poisson**
  - Phân phối Skellam
  - Phân phối Yule-Simon
  - Phân phối zeta

Phân phối xác suất liên tục

- **Biến có giá trị trên một khoảng bị chặn:**
  - Phân phối Beta trên đoạn [0,1]
  - **Phân phối đều liên tục trên đoạn [a,b]** (Continuous Uniform distribution)
    - Phân phối chữ nhật trên đoạn [-1/2,1/2]
  - Hàm delta Dirac
  - Phân phối Kumaraswamy
  - Phân phối lôga (liên tục)
  - Phân phối tam giác trên đoạn [a, b]
  - Phân phối Von Mises
  - Phân phối nửa hình tròn Wigner (Wigner semicircle distribution)

Phân phối xác suất liên tục

- **Biến có giá trị trên một khoảng nửa hữu hạn (thường là [0,∞):**
  - Phân phối Khi
  - Phân phối Khi không trung tâm (noncentral chi distribution)
  - **Phân phối Khi-bình phương**
    - Phân phối Khi-bình phương nghịch đảo (inverse-chi-square distribution)
    - Phân phối Khi-bình phương nghịch đảo không trung tâm (noncentral chi-square distribution)
    - Phân phối Khi-bình phương nghịch đảo tỉ lệ (scale-inverse-chi-square distribution)

Phân phối xác suất liên tục

- **Biến có giá trị trên một khoảng nửa hữu hạn (thường là [0,∞):**
  - **Phân phối mũ**
  - Phân phối F
    - Phân phối F không trung tâm (noncentral F-distribution)
  - Phân phối Gamma
    - Phân phối Erlang
    - Phân phối gamma đảo (inverse-gamma distribution)
  - Phân phối z của Fisher (Fisher's z-distribution)
  - Phân phối nửa chuẩn (half-normal distribution)
  - Phân phối Lévy

Phân phối xác suất liên tục

- **Biến có giá trị trên một khoảng nửa hữu hạn (thường là [0,∞):**
  - Phân phối logarit-lý luận (log-logistic distribution)
  - Phân phối logarit chuẩn (log-normal distribution)
  - Phân phối Pareto
  - Phân phối Rayleigh
    - Phân phối Rayleigh hỗn hợp (Rayleigh mixture distribution)
  - Phân phối Rice
  - Phân phối Gumgel loại 2 (type-2 Gumbel distribution)
  - Phân phối Wald
  - Phân phối Weibull

Phân phối xác suất liên tục

- **Biến có giá trị trên toàn tập số thực:**
  - Phân phối nguyên tố Beta
  - Phân phối Cauchy
  - Phân phối Fisher-Tippett
    - Phân phối Gumbel
  - Phân phối giá trị cực tổng quát (generalized extreme value distribution)
  - Phân cát tuyến hyperbolic (Hyperbolic secant distribution)
  - Phân phối Landau
  - Phân phối Laplace
  - Phân phối Lévy nghiêng alpha ổn định (Lévy skew alpha-stable distribution)

Phân phối xác suất liên tục

- **Biến có giá trị trên toàn tập số thực:**
  - Phân phối bản đồ Airy (map-Airy distribution)
  - **Phân phối chuẩn** (normal distribution) còn gọi là phân phối theo đường cong Gauss
  - **Phân phối Student**, là phân phối của biến ngẫu nhiên biểu diễn giá trị trung bình chưa biết của phân phối Gauss.
    - Phân phối Student không tâm
  - Phân phối Gumbel loại 1

Phân phối điều kiện

- **Phân phối đồng thời của các biến ngẫu nhiên trên cùng một không gian mẫu:**
  - Phân phối Dirichlet
  - Công thức mẫu Ewen (Ewens's sampling formula)
  - Phân phối bội, là tổng quát hóa của phân phối nhị thức.
  - Phân phối chuẩn bội, là tổng quát hóa của phân phối chuẩn
- **Các phân phối của các ma trận ngẫu nhiên:**
  - Phân phối Wishart
  - Phân phối ma trận chuẩn
  - Phân phối ma trận Student
  - Phân phối T-bình phương Hotelling (Hotelling's T-square distribution)

4. Phân phối rời rạc điển hình

Phân phối nhị thức
(*Binomial Distribution*)

Phân phối rời rạc
Biến có giá trị hữu hạn

Khái niệm

- Phân phối nhị thức là một hàm phân phối xác suất của số lượng thành công trong ***n*** lượt thử độc lập. Tìm kết quả CÓ hay KHÔNG thành công.

Hàm số phân phối nhị thức:

$$P[X=k]=\binom{n}{k}p^k(1-p)^{n-k} \quad \text{for } k=0,1,2,\ldots,n$$

ở đây $$\binom{n}{k}=\frac{n!}{k!(n-k)!}$$

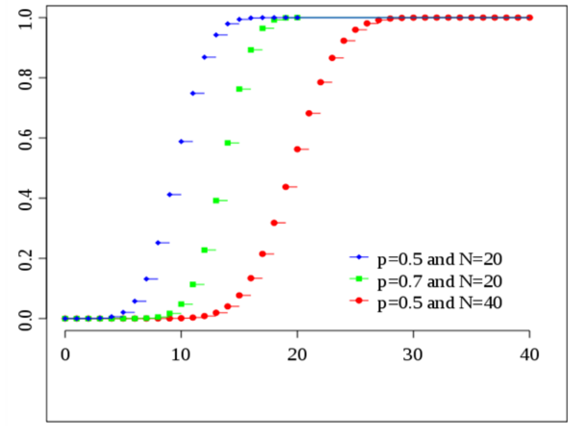
với k là số lần thu được kết quả CÓ thành công trong n lượt thử

Khái niệm

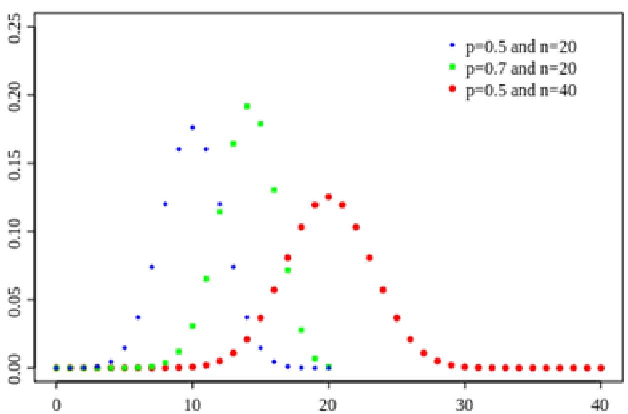
- Thực nghiệm nhị thức chỉ cho ra một trong hai kết quả, CÓ hoặc KHÔNG.
- Các trường hợp điển hình ứng dụng thực nghiệm nhị thức:
  - Một đồng xu lật xấp hay ngửa
  - Ứng cử viên trong cuộc bầu cử thắng hay thua
  - Một sinh viên nam hay nữ
  - Một chiếc xe dung xăng chỉ số Octane 95 hay dùng xăng khác.

41

Hàm phân phối xác suất lũy tích



Hàm mật độ xác suất



Đặc trưng

Tham số	$n \geq 0$ số lượt thử (số nguyên) $0 \leq p \leq 1$ xác suất CÓ thành công (số thực)
Giá	$k \in \{0, \dots, n\}$
Hàm mật độ	$\binom{n}{k} p^k (1-p)^{n-k}$
Hàm phân phối tích lũy	$I_{1-p}(n - \lfloor k \rfloor, 1 + \lfloor k \rfloor)$
Giá trị kỳ vọng	np
Trung vị	one of $\{\lfloor np \rfloor, \lceil np \rceil\}$ [1]
Mode	$\lfloor (n+1)\,p \rfloor$
Phương sai	$np(1-p)$

Đặc trưng

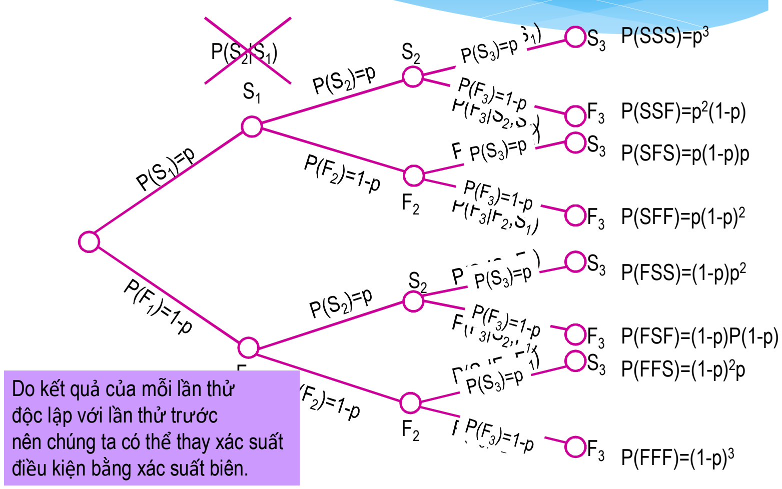
Độ xiên	$\frac{1-2p}{\sqrt{np(1-p)}}$
Độ nhọn	$\frac{1-6p(1-p)}{np(1-p)}$
Entropy	$\frac{1}{2}\ln(2\pi nep(1-p)) + O\left(\frac{1}{n}\right)$
Hàm sinh moment	$(1-p+pe^{t})^{n}$
Hàm đặc trưng	$(1-p+pe^{it})^{n}$

- Ghi chú: Phân phối Bernoulli là trường hợp đặc biệt của phân phối nhị thức với n=1

Thực nghiệm nhị phân

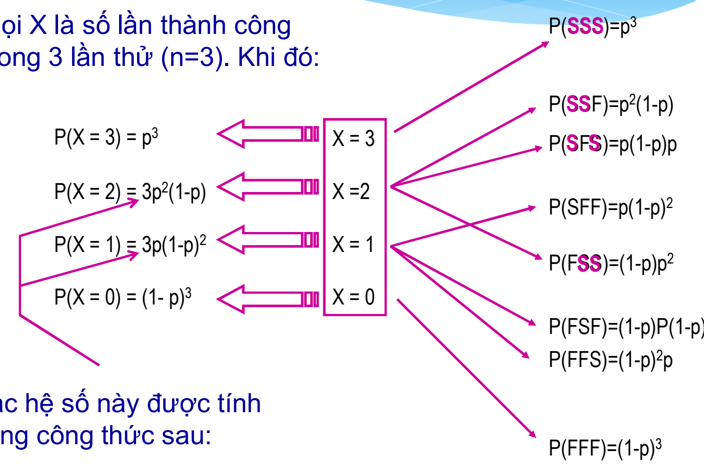
- **Điều kiện cho phép thử nhị phân**
  - Có ***n*** phép/lần thử (***n*** được xác định và không đổi).
  - Mỗi phép/lần thử sẽ cho kết quả CÓ hoặc KHÔNG.
  - Xác suất thành công (CÓ) ***p*** là như nhau cho mọi phép/lần thử.
  - Tất cả các phép thử độc lập nhau.
- **Biến ngẫu nhiên nhị phân**
  - Biến ngẫu nhiên nhị phân đếm số lần thành công (CÓ) trong n phép/lần thử.
  - Theo định nghĩa, đây là biến rời rạc.

Lập công thức xác suất



Lập công thức xác suất

Gọi X là số lần thành công trong 3 lần thử (n=3). Khi đó:



Các hệ số này được tính bằng công thức sau:

Công thức tính xác suất nhị thức

- Tổng quát, xác suất nhị thức được tính theo công thức:

$$P(X=x)=p(x)=C_x^n p^x(1-p)^{n-x}$$

$$với\ C_x^n=\frac{n!}{x!(n-x)!}$$

Ví dụ: Cho n = 3

$$x=0:\ C_0^3=\frac{3!}{0!(3-0)!}=\frac{1\cdot 2\cdot 3}{(1)(1\cdot 2\cdot 3)}=1$$

$$x=1:\ C_1^3=\frac{3!}{1!(3-1)!}=\frac{1\cdot 2\cdot 3}{(1)(1\cdot 2)}=3$$

Ví dụ 1

- 5% số gạch sản xuất ra bị lỗi.
- Một mẫu gồm 3 viên gạch được lấy ra. Tìm phân phối xác suất của số viên gạch bị lỗi.

Giải:

- Một viên gạch có thể tốt hoặc có thể bị lỗi.
- Số lần thử là cố định (n=3)
- Mỗi viên gạch là độc lập với các viên gạch khác.
- Xác suất của một viên gạch bị phát hiện lỗi là không đổi trong các lần thử (p=.05).

Các điều kiện của thực nghiệm nhị thức đều thỏa

- Gọi X là biến ngẫu nhiên nhị thức.
- Xác định xác suất số lần viên gạch bị phát hiện "có lỗi".

$$P(X=0)=p(0)=\frac{3!}{0!(3-0)!}(.05)^0(.95)^{3-0}=.8574$$

$$P(x=1)=p(1)=\frac{3!}{1!(3-1)!}(.05)^1(.95)^{3-1}=.1354$$

$$P(x=2)=p(2)=\frac{3!}{2!(3-2)!}(.05)^2(.95)^{3-2}=.0071$$

$$P(x=3)=p(3)=\frac{3!}{3!(3-3)!}(.05)^3(.95)^{3-3}=.0001$$

X	P(X)
0	.8574
1	.1354
2	.0071
3	..0001

Giá trị bình quân & Phương sai

E(X) = μ = np

V(X) = σ^2 = np(1-p)

Ví dụ:

- Các ghi chép cho thấy 30% số khách hàng mua gạch trong cửa hàng dùng thẻ tín dụng để thanh toán.
- Sáng nay có 20 khách đến mua hàng.
- Trả lời các câu hỏi sau:

- Xác suất để có tối thiểu 12 khách hàng dùng thẻ tín dụng?
  - Đây là bài toán nhị phân với n=20 and p=.30

k	p	
	.01..........	30
0		
.		
11		.995

P(Tối thiểu 12 khách dùng thẻ tín dụng)
= P(X>=12)=1-P(X<=11)
= 1-.995 = .005

- Xác suất để tối thiểu 3 nhưng không quá 6 khách hàng dùng thẻ tín dụng?

k	p	
	.01..........	30
0		
2		.035
.		
6		.608

P(3<=X<=6)=
P(X=3 or 4 or 5 or 6)
=P(X<=6) -P(X<=2)
=.608 - .035 = .573

- Bao nhiêu khách hàng được kỳ vọng dùng thẻ tín dụng?
  - E(X) = np = 20(.30) = 6

- Xác suất sẽ có chính xác 14 khách hàng không dùng thẻ tín dụng?
  - Gọi Y là số khách hàng không dùng thẻ tín dụng.
    P(Y=14) = P(X=6) = P(X<=6) - P(x<=5) = .608 - .416 = .192

- Xác suất có tối thiểu 9 khách hàng không dùng thẻ tín dụng.
  - Gọi Y là số khách hàng không dùng thẻ tín dụng.
    P(Y>=9) = P(X<=11) = .995

Phân phối Poisson
(*Poisson Distribution*)

Phân phối rời rạc
Biến có giá trị vô hạn

Khái niệm

- Phân phối Poisson là một phân phối xác suất rời rạc.
- Nó khác với các phân phối xác suất rời rạc khác ở chỗ thông tin cho biết không phải là xác suất để một sự kiện (event) xảy ra (thành công) trong một lần thử như trong phân phối Bernoulli, hay là số lần mà sự kiện đó xảy ra trong **n** lần thử như trong phân phối nhị thức, mà chính là **trung bình số lần xảy ra thành công của một sự kiện trong một khoảng thời gian hay một phạm vi nhất định**.
- Giá trị trung bình này được gọi là **lamda**, kí hiệu là **λ**

(Trong nhiều tài liệu giá trị này cũng được ký hiệu là μ).

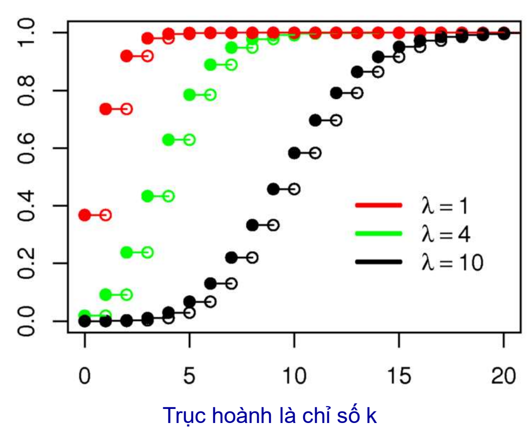
Khái niệm

- Thí nghiệm Poisson thường phù hợp với trường hợp của các sự kiện hiếm xảy ra trong một khoảng thời gian nhất định hoặc trong một phạm vi xác định.
- Trường hợp điển hình:
  - Số lỗi người đánh máy mắc trong một trang
  - Số khách hàng bước vào một quầy dịch vụ trong một khoảng thời gian xác định (giờ, ngày,…)
  - Số cuộc gọi tới trong thời gian một giờ.

Tính chất của thực nghiệm Poisson

- Số dữ kiện thành công xảy ra trong một khoảng thời gian là độc lập với số dữ kiện thành công xảy ra trong một khoảng thời gian khác.
- Xác suất thành công trong một khoảng thời gian xác định:
  - Bằng nhau cho bất kỳ khoảng thời gian nào của cùng kích thức mẫu
  - Tỉ lệ với chiều dài của khoảng thời gian
- Xác suất của hai hay nhiều lần thành công sẽ gần với zero khi khoảng thời gian nhỏ dần.

Hàm phân phối xác suất lũy tích



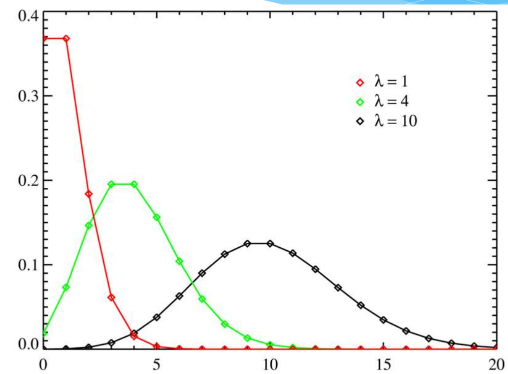
Hàm mật độ xác suất

$$f(k;\lambda) = \frac{\lambda^k e^{-\lambda}}{k!}$$

với

- e là cơ số của logarit tự nhiên (e = 2.71828...)
- k là số lần xuất hiện của một sự kiện - mà xác suất của nó là cho bởi công thức trên
- $k!$ là giai thừa của k
- λ là số thực dương, bằng với giá trị kì vọng xuất hiện của sự kiện trong một khoảng cho sẵn. Ví dụ, nếu một sự kiện trung bình xảy ra 1 lần trong 4 phút, giờ ta quan tâm số lần sự kiện xảy ra trong khoảng thời gian 10 phút, ta dùng mô hình phân phối Poisson với λ = 10/4 = 2.5.

Hàm mật độ (khối) xác suất



Trục hoành là chỉ số k.

Hàm khối xác suất được định nghĩa dựa trên duy nhất biến nguyên k.

Đường nối dùng để minh họa chứ không có nghĩa là liên tục.

Đặc trưng

Tham số	$\lambda \in (0, \infty)$
Giá	$k \in \{0, 1, 2, \ldots\}$
Hàm mật độ	$\frac{e^{-\lambda}\lambda^k}{k!}$
Hàm phân phối tích lũy	$\frac{\Gamma(\lfloor k+1 \rfloor, \lambda)}{\lfloor k \rfloor!}$ for $k \ge 0$
	(với $\Gamma(x, y)$ là hàm gamma không đầy đủ)
Giá trị kỳ vọng	λ
Trung vị	usually about $\lfloor \lambda + 1/3 - 0.02/\lambda \rfloor$

Đặc trưng

Mode	$\lceil \lambda \rceil - 1$
Phương sai	λ
Độ xiên	$\lambda^{-1/2}$
Độ nhọn	λ^{-1}
Entropy	$\lambda[1 - \ln(\lambda)] + e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k \ln(k!)}{k!}$
Hàm sinh moment	$\exp(\lambda(e^t - 1))$
Hàm đặc trưng	$\exp(\lambda(e^{it} - 1))$

Biến ngẫu nhiên & Phân phối xác suất

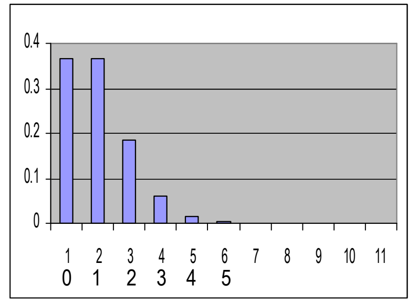
- Biến ngẫu nhiên Poisson
  - Biến Poisson chỉ số lần thành công xảy ra trong một khoảng thời gian cho trước hoặc trong một miền xác định trong thực nghiệm Poisson.
- Phân phối xác suất của biến ngẫu nhiên Poisson

$$P(X=x)=p(x)=\frac{e^{-\mu}\mu^{x}}{x!} \qquad x=0,1,2...$$

$$E(X)=V(X)=\mu$$

Phân phối xác suất Poisson

Phân phối xác suất Poisson với μ**=1**



$$P(X=0)=p(0)=\frac{e^{-1}1^{0}}{0!}=e^{-1}=.3678$$

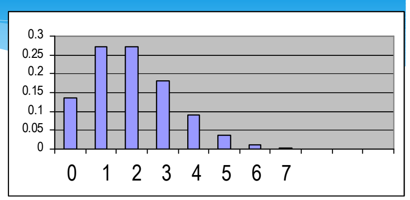
$$P(X=1)=p(1)=\frac{e^{-1}1^{1}}{1!}=e^{-1}=.3678$$

$$P(X=2)=p(2)=\frac{e^{-1}1^{2}}{2!}=\frac{e^{-1}}{2}=.1839$$

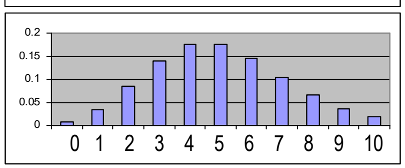
$$P(X=3)=p(3)=\frac{e^{-1}1^{3}}{3!}=\frac{e^{-1}}{6}=.0613$$

Trục X trong Excel bắt đầu với x=1!!

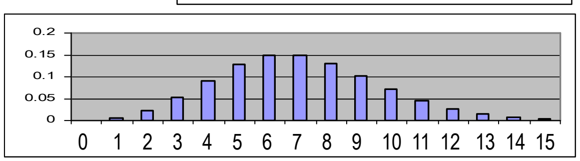
Phân phối xác suất Poisson với μ=2

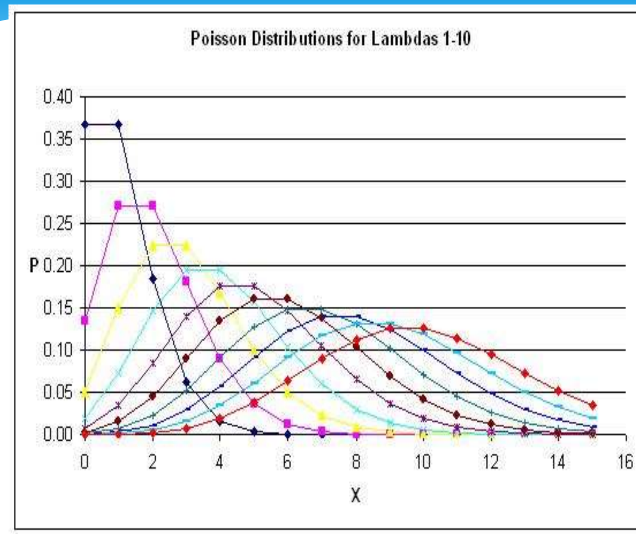


Phân phối xác suất Poisson với μ=5



Phân phối xác suất Poisson với μ=7





Ví dụ 2

- Nghiên cứu giao thông cho thấy số xe qua quầy thu phí giao thông là 360 xe/giờ.
- Tìm xác suất để chỉ có 2 xe/phút?
  - Dùng công thức:
    - μ = 360/60 = 6 xe/phút
    - Gọi X là số xe qua trạm trong thời gian 1 phút.

$$P(X=2)=\frac{e^{-6}6^2}{2!}=.0446$$

  - Dùng bảng tra:
    - P(X = 2) = P(X<=2) - P(X<=1) = .062 - .017 = .045

Tìm xác suất để chỉ có 2 xe đến trong thời gian 1 phút? (Dùng bảng tra)

- P(X = 2) = P(X<=2) - P(X<=1) = .062 - .017 = .045

	μ					
K	0.1	....	3	....	$\boxed{6}$	....
0	0.905		0.05		0.002	
1	0.995		0.199		$\boxed{0.017}$	
2	1		0.423		$\boxed{0.062}$	

Xác suất để có tối thiểu 4 xe/phút? (Dùng bảng tra)

- P(X>=4) = 1 - P(X<=3) = 1 - .151 = .849

Xấp xỉ Poisson của phân phối Nhị thức

- Khi **n** quá lớn, xác suất nhị thức sẽ tính toán với khối lượng lớn.
- Nếu **p** khá nhỏ (ví dụ: p< .05), chúng ta có thể tính gần đúng phân phối nhị thức bằng cách sử dụng phân phối Poisson.
- Cho **μ = np** và thực hiện như sau:

$$P(\underset{\substack{\uparrow\\ \text{Với tham số n \& p}}}{X_{Binomial}}=x)\cong P(\underset{\substack{\uparrow\\ \text{Với } \mu = np}}{X_{Poisson}}=x)$$

Ví dụ 3

- Một kho hàng thường kiểm tra 50 viên gạch khi có lô hàng mới đến, và sẽ chỉ chấp nhận lô hàng nếu không quá 2 viên bị phát hiện có lỗi.
- Một lô hàng trong thực tế có 2% số gạch lỗi. Tìm xác suất để lô hàng được chấp nhận?

Giải

- Đây là thực nghiệm nhị thức với n=50, p=.02.
- Giá trị n khá lớn, nếu dùng bảng tra cũng không có giá trị, p=0.02<.05, do vậy sử dụng xấp xỉ Poisson [μ=(50)(.02)=1]
- → P(Xpoisson<=2) = .920

  Giá trị này gần với xác suất nhị thức (=.922)

5. Phân phối liên tục điển hình

Phân phối đều liên tục
(*Continuous Uniform Distribution*)

Phân phối liên tục
Biến có giá trị trên một khoảng bị chặn

Khái niệm

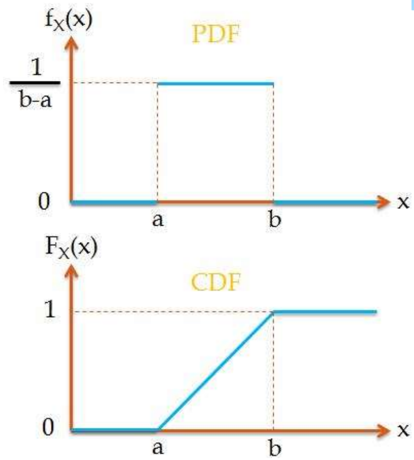
- **Phân phối đều liên tục**, đôi khi còn được gọi là **phân phối hình chữ nhật**, là một phân phối mà xác suất xảy ra như nhau cho mọi kết cục của biến ngẫu nhiên liên tục.
- Hàm mật độ xác suất của phân phối đều như sau:

$$f(x) = \frac{1}{b-a} \qquad a \le x \le b.$$

- Kỳ vọng toán và phương sai:

$$\mathrm{E(X)} = \frac{a+b}{2} \qquad V(\mathbf{X}) = \frac{(b-a)^2}{12}$$

Hàm phân phối & Hàm mật độ



Đặc trưng

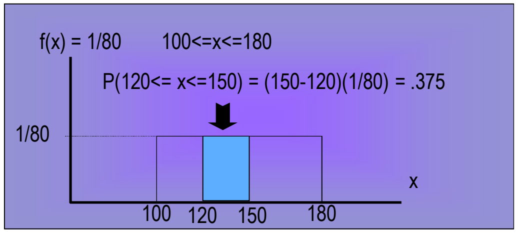
Tham số	$a, b \in (-\infty, \infty)$
Giá	$a \le x \le b$
Hàm mật độ xác suất	$\frac{1}{b-a}$ for $a \le x \le b$ 0 for $x < a$ or $x > b$
Hàm phân phối tích lũy	0 for $x < a$ $\frac{x-a}{b-a}$ for $a \le x < b$ 1 for $x \ge b$
Giá trị kỳ vọng	$\frac{a+b}{2}$
Trung vị	$\frac{a+b}{2}$

Đặc trưng

Mode	mọi giá trị thuộc $[a, b]$
Phương sai	$\frac{(b-a)^2}{12}$
Độ xiên	0
Độ nhọn	$-\frac{6}{5}$
Entropy	$\ln(b-a)$
Hàm sinh moment	$\frac{e^{tb}-e^{ta}}{t(b-a)}$
Hàm đặc trưng	$\frac{e^{itb}-e^{ita}}{it(b-a)}$

Ví dụ 4

- Thời gian chờ kể từ khi đặt hàng và nhận hàng, nằm trong khoảng 100 đến 180 phút, là phân phối đều.
- Vẽ sơ đồ và viết hàm mật độ?
- Tìm xác suất để khoảng thời gian chờ nằm trong khoảng từ 2 giờ đến 2,5 giờ?



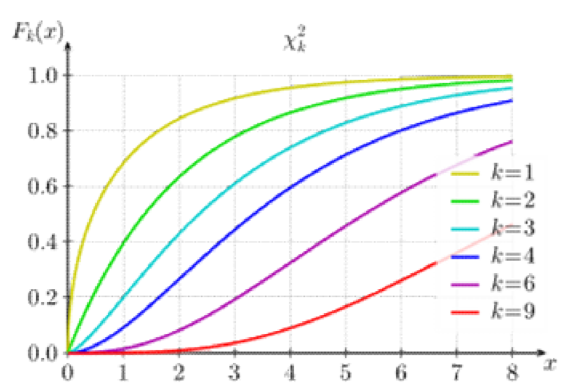
Phân phối Khi-Bình Phương κ^2 (*Chi-squared Distribution*)

Phân phối liên tục
Biến có giá trị trên một nửa hữu hạn

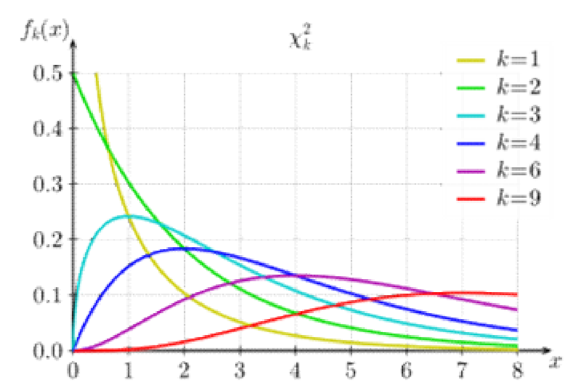
Khái niệm

- Phân phối Khi-Bình phương với ***k*** bậc tự do là sự phân bố của một tổng các bình phương của ***k*** biến ngẫu nhiên độc lập tiêu chuẩn bình thường.
- Đó là một trường hợp đặc biệt của phân phối gamma và là một trong những bản phân phối xác suất được sử dụng rộng rãi nhất trong thống kê suy luận.

Hàm phân phối xác suất lũy tích



Hàm mật độ xác suất



Hàm mật độ xác suất

- Đồ thị của hàm mật độ xác suất là đường cong không đối xứng, lệch mạnh về phía trái.
- Khi bậc tự do **k≥30**, đồ thị hàm mật độ dịch chuyển về bên phải và gần đối xứng.
- Khi đó đồ thị hàm mật độ của phân phối Khi-Bình Phương tiệm cận **phân phối chuẩn**.

Đặc trưng

Notation	$\chi^2(k)$ or χ^2_k
Parameters	$k \in \mathbb{N}_{>0}$ (known as "degrees of freedom")
Support	$x \in [0, +\infty)$
PDF	$\frac{1}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)} x^{\frac{k}{2}-1} e^{-\frac{x}{2}}$
CDF	$\frac{1}{\Gamma\left(\frac{k}{2}\right)} \gamma\left(\frac{k}{2}, \frac{x}{2}\right)$
Mean	k
Median	$\approx k\left(1-\frac{2}{9k}\right)^3$

Đặc trưng

Mode	$\max\{k-2, 0\}$
Variance	$2k$
Skewness	$\sqrt{8/k}$
Ex. kurtosis	$12/k$
Entropy	$\frac{k}{2} + \ln(2\Gamma(\frac{k}{2})) + (1-\frac{k}{2})\psi(\frac{k}{2})$ (nats)
MGF	$(1-2t)^{-k/2}$ for $t < ½$
CF	$(1-2it)^{-k/2}$ [1]

Phân phối mũ
(*Exponential Distribution*)

Phân phối liên tục

Biến có giá trị trên một khoảng nửa hữu hạn

Khái niệm

- Phân phối mũ có thể dùng để mô phỏng:
  - Thời gian giữa các cuộc gọi
  - Thời gian giữa các lần xe qua trạm thu phí
  - Tuổi thọ của thiết bị điện tử
- Khi số lần xuất hiện của một sự kiện tuân theo phân phối Poisson, thì **thời gian giữa các lần xuất hiện** của sự kiện đó tuân theo phân phối mũ.

Tính chất

- Một biến ngẫu nhiên theo phân phối mũ nếu hàm mật độ xác suất là:

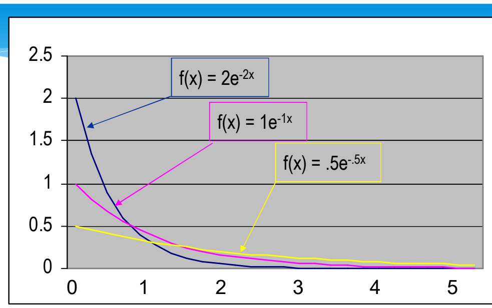
$$f(x) = \lambda e^{-\lambda x}, \quad x>=0$$

- λ là một tham số của phân phối (λ>0)

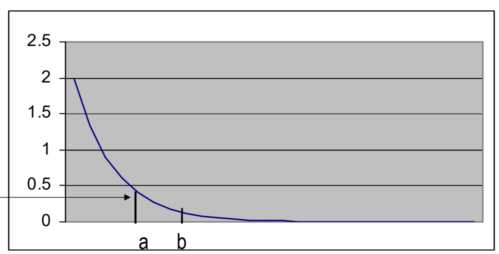
$$E(X) = 1/\lambda \qquad V(X) = 1/\lambda$$

$$P(X>a) = e^{-\lambda a}$$

Phân phối mũ cho λ = 0.5, 1, 2



$P(a<x<b) = e^{-\lambda a} - e^{-\lambda b}$



Ví dụ 5

- Lượt xe chạy qua trạm thu phí (độc lập và ngẫu nhiên) đạt mức trung bình 360 xe/giờ. Dùng phân phối mũ tính xác suất để thời gian giữa hai lượt xe cách nhau một phút.

Giải

- Gọi X là thời gian (phút) trôi qua giữa hai lượt xe. X là biến phân phối mũ với λ = 360/60 = 6 xe/phút.

$$P(X>.5) = e^{-6(.5)} = .0498$$

- Tính xác suất để không có xe nào chạy đến trong khoảng thời gian ½ phút?

Giải

- Nếu Y là số lượt xe đến trong vòng ½ phút, khi đó Y là biến theo phân phối Poisson với μ = (.5)(6) = 3 xe / ½ phút

$$P(Y=0) = e^{-3}(3^0)/0! = .0498.$$

Bình luận: Nếu chiếc xe đầu tiên không đến trạm trong vòng ½ phút thì chiếc xe kế tiếp cũng không đến trong vòng ½ phút. Do vậy, không ngạc nhiên khi xác suất ở câu hỏi này bằng với xác suất của câu hỏi trước.

Ví dụ 6

- Tuổi thọ của transitor (bán dẫn) tuân theo phân phối mũ, với giá trị bình quân là 1000 giờ.
- Tính xác suất cái transistor hoạt động được từ 1000 đến 1500 giờ?

Giải

- Gọi X là tuổi thọ của transistor (giờ)

$$E(X) = 1000 = 1/\lambda, \quad \rightarrow \lambda = 1/1000 = .001$$

$$P(1000<X<1500) = e^{-(.001)(1000)} - e^{-(.001)(1500)} = .1448$$

Phân phối chuẩn
(*Normal Distribution*)

Phân phối liên tục
Biến có giá trị trên toàn tập số thực

Khái niệm

- **Phân phối chuẩn**, còn gọi là **phân phối Gauss**, là một phân phối xác suất cực kì quan trọng trong nhiều lĩnh vực.
  - Nhiều biến ngẫu nhiên có thể được mô phỏng thành phân phối chuẩn,
  - Nhiều phân phối khác có thể được tính gần đúng từ phân phối chuẩn.
  - Phân phối chuẩn là phân phối nền tảng của thống kê suy luận.
- Nó là họ phân phối có dạng tổng quát giống nhau, chỉ khác tham số giá trị trung bình μ (***vị trí***) và phương sai σ^2.

Khái niệm

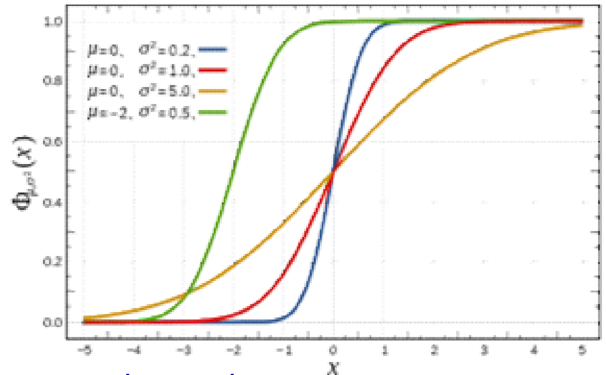
- Phân phối chuẩn còn được gọi là **đường cong chuông** (bell curve) hay đường cong gò (mount) do hình dạng của đồ thị của hàm mật độ xác suất.

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

$$\text{với } \pi = 3.14159... \text{ và } e = 2.71828...$$

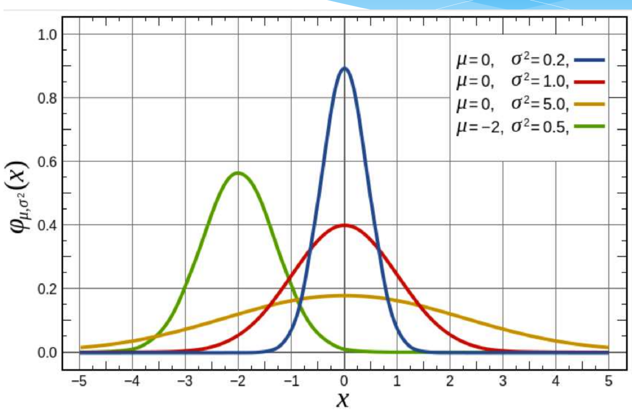
- **Phân phối chuẩn chuẩn hóa** (standard normal distribution), còn gọi là **phân phối chuẩn tắc**, là phân phối chuẩn có giá trị trung bình μ=0 phương sai σ^2=1 (đường cong màu đỏ trong hình bên dưới).

Hàm phân phối xác suất lũy tích



Hàm phân phối xác suất là đường cong **phản xứng**. Đường màu đỏ là **phân phối chuẩn chuẩn hóa**.

Hàm mật độ xác suất



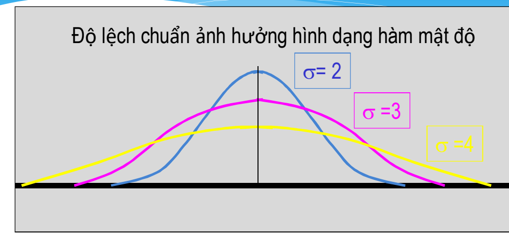
Hàm mật độ xác suất **đối xứng** qua trục tại μ. Đường màu đỏ là **phân phối chuẩn chuẩn hóa**.

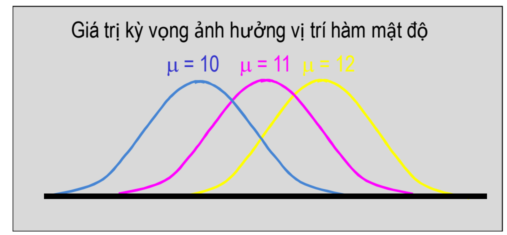
Đặc trưng

Tham số	μ cho biết vị trí (thực) $\sigma^2 > 0$ bình phương tỉ lệ (thực)
Giá	$x \in (-\infty; +\infty)$
Hàm mật độ xác suất	$\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$
Hàm phân phối tích lũy	$\frac{1}{2}\left(1 + \operatorname{erf} \frac{x-\mu}{\sigma\sqrt{2}}\right)$
Giá trị kỳ vọng	μ
Trung vị	μ

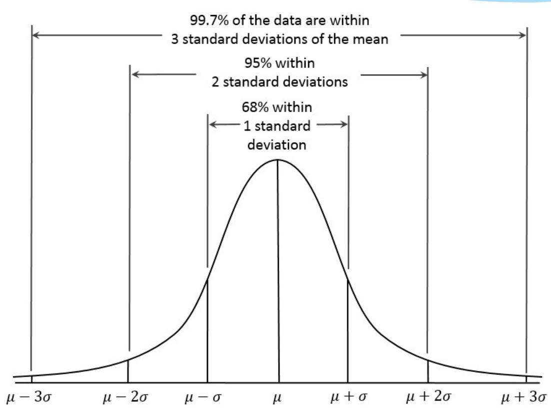
Đặc trưng

Mode	μ
Phương sai	σ^2
Độ xiên	0
Độ nhọn	0
Entropy	$\ln\left(\sigma\sqrt{2\,\pi\,e}\right)$
Hàm sinh moment	$M_X(t) = \exp\left(\mu\,t + \frac{\sigma^2 t^2}{2}\right)$
Hàm đặc trưng	$\phi_X(t) = \exp\left(\mu\,i\,t - \frac{\sigma^2 t^2}{2}\right)$





Quy tắc thực nghiệm



Các bài toán xác suất phân phối chuẩn

- Lưu ý 2 đặc điểm giúp tính xác suất phân phối chuẩn:
  - Phân phối chuẩn là hàm đối xứng
  - Bất kỳ một phân phối chuẩn nào cũng có thể được chuyển về dạng phân phối chuẩn chuẩn hóa

Ví dụ 7

- Thời gian hoàn thành (và đạt điểm tuyệt đối) bài kiểm tra giữa kỳ môn TKUD của các sinh viên trong lớp tuân theo phân phối chuẩn với thời gian bình quân là 60 phút và độ lệch là 8 phút.
- Tính xác suất để một sinh viên hoàn thành trong khoảng thời gian 60 đến 70 phút?

Giải

- Gọi X là thời gian cần để hoàn thành bài kiểm tra, chúng ta cần tìm xác suất P(60<X<70).
- Xác suất này có thể được tính bằng cách chuyển biến X (phân phối chuẩn) thành một biến Z tuân theo phân phối chuẩn hóa.

Mọi biến pp chuẩn X với μ & σ, đều có thể chuyển sang biến Z.

$$Z = \frac{X - \mu_x}{\sigma_x}$$

Vì vậy, sau khi xác suất của biến (chuẩn hóa) Z đã được tính thì có thể suy ra xác suất của biến (chuẩn) X.

E(Z) = 0 V(Z) = 1

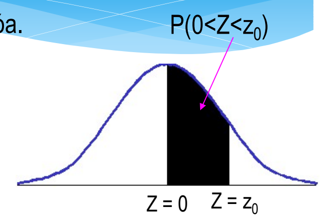
$$P(60<X<70) = P\left(\frac{60-60}{8} < \frac{X-\mu}{\sigma} < \frac{70-60}{8}\right)$$

$$= P(0<Z<1.25)$$

Để thực hiện bài toán,
cần tính xác suất theo phân phối chuẩn hóa

Dùng công thức hoặc sử dụng bảng tra để tìm xác suất chuẩn hóa.

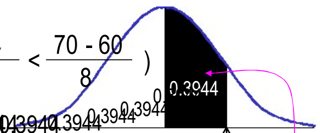
Giá trị xác suất tương ứng với phần diện tích giữa Z=0 và $Z=z_0 > 0$



z	0	0.1	....	0.05	0.06
0.0	0.0000	0.0040		0.0199	0.0239
0.1	0.0398	0.0438		0.0596	0.636
.	.	.		.	.
.	.	.		.	.
1.0	0.3413	0.3438		0.3531	0.3554
.	.	.		.	.
.	.	.		.	.
1.2	0.3849	0.3869	....	0.3944	0.3962
.	.	.	.	.	.
.	.	.	.	.	.

$$P(60<X<70) = P\left(\frac{60-60}{8} < \frac{X-\mu}{\sigma} < \frac{70-60}{8}\right)$$

$$= P(0<Z<1.25) = 0.3944$$

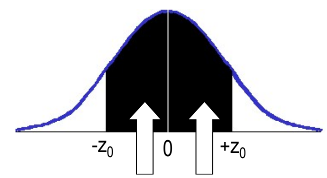


Trong ví dụ này z_0 = 1.25

z	0	0.1	....	0.05	0.06
0.0	0.0000	0.0040		0.0199	0.0239
0.1	0.0398	0.0438		0.0596	0.636
.	.	.		.	.
.	.	.		.	.
1.0	0.3413	0.3438		0.3531	0.3554
.	.	.		.	.
.	.	.		.	.
1.2	0.3849	0.3869	....	0.3944	0.3962
.	.	.	.	.	.
.	.	.	.	.	.

Vận dụng tính đối xứng

- Do tính đối xứng của phân phối chuẩn, có thể tính xác suất cho giá trị âm của Z_0 ($-Z_0$) như sau:

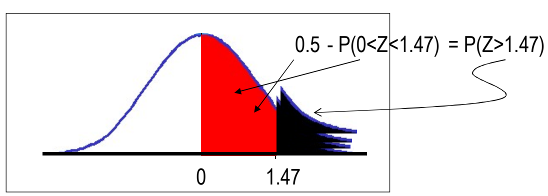


$P(-z_0<Z<0) = P(0<Z<z_0)$

Ví dụ 8

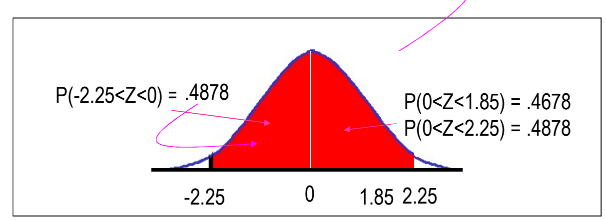
- Tính các xác suất sau:

P(Z>1.47) = ?



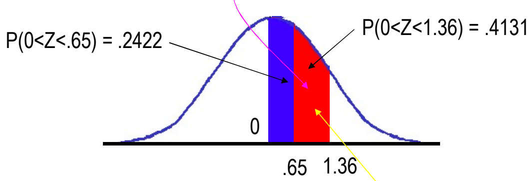
P(Z>1.47) = 0.5 - 0.4292 = 0.0708

P(-2.25<Z<1.85) = ?



P(-2.25<Z<1.85) = 0.4878 + 0.4678 = 0.9556

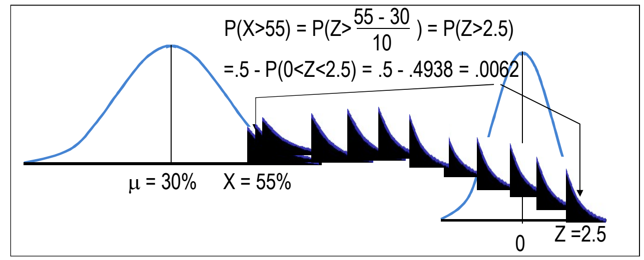
P(.65<Z<1.36) = ?



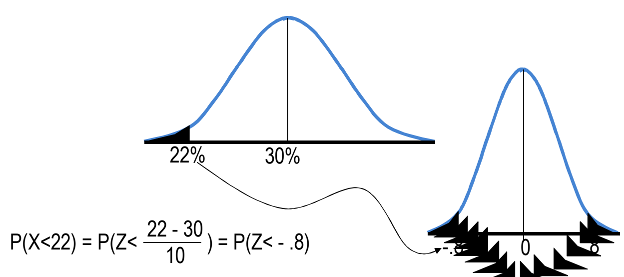
P(.65<Z<1.36) = .4131 - .2422 = .1709

Ví dụ 9

- Suất thu lợi (gọi là X) tuân theo luật phân phối chuẩn với giá trị bình quân 30% và độ lệch chuẩn 10%
- Tính xác suất suất thu lợi (X) có thể vượt 55%?



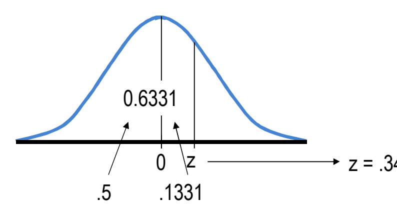
- Tính xác suất suất thu lợi (X) có thể dưới 22%?



Ví dụ 10

- Nếu Z là biến chuẩn hóa, xác định giá trị của Z để:

P(Z<z)=0.6331

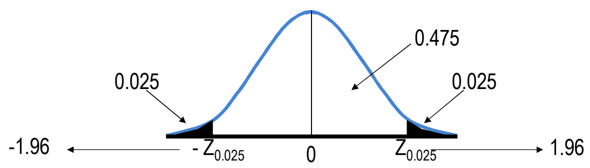


Ví dụ 11

- Xác định $Z_{0.025}$?

Giải

- Z_A là giá trị mà tại đó diện tích (đường chuẩn hóa) về phía phải bằng A.



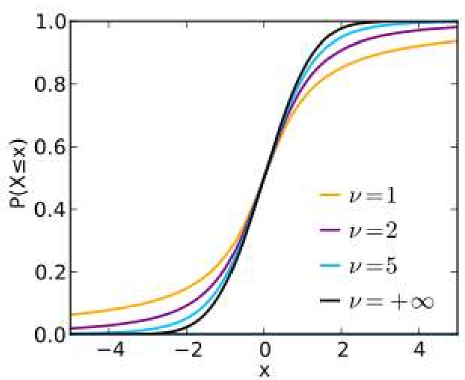
Phân phối Student (*Student Distribution*)

Phân phối liên tục
Biến có giá trị trên toàn tập số thực

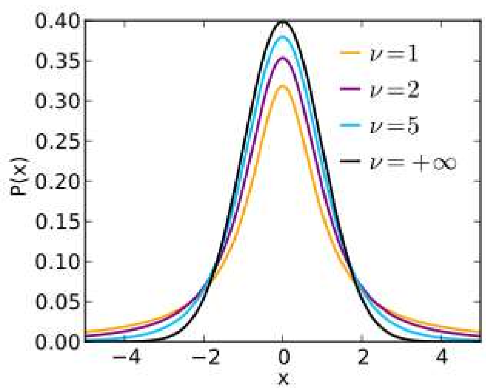
Khái niệm

- Phân phối student (phân phối t) là phân phối xác suất liên tục phát sinh khi ước tính bình quân của một tổng thể phân bố chuẩn khi kích thước mẫu nhỏ và độ lệch tiêu chuẩn của tổng thể chưa được biết rõ.
- Phân phối chuẩn mô tả một tổng thể đầy đủ, còn phân phối student mô tả một mẫu được rút ra từ tổng thể đầy đủ; do vậy, phân phối t sẽ khác nhau cho các kích thước mẫu khác nhau.
- Khi mẫu càng lớn (n lớn), phân phối t càng gần **phân phối chuẩn**.

Hàm phân phối xác suất lũy tích



Hàm mật độ xác suất



Đặc trưng

Parameters	$\nu > 0$ degrees of freedom (real)
Support	$x \in (-\infty; +\infty)$
PDF	$\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma\left(\frac{\nu}{2}\right)}\left(1+\frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}$
CDF	$\frac{1}{2} + x\Gamma\left(\frac{\nu+1}{2}\right) \times$ $\frac{{}_2F_1\left(\frac{1}{2},\frac{\nu+1}{2};\frac{3}{2};-\frac{x^2}{\nu}\right)}{\sqrt{\pi\nu}\,\Gamma\left(\frac{\nu}{2}\right)}$ where ${}_2F_1$ is the hypergeometric function
Mean	0 for $\nu > 1$, otherwise undefined
Median	0
Mode	0
Variance	$\frac{\nu}{\nu-2}$ for $\nu > 2$, ∞ for $1 < \nu \le 2$, otherwise undefined

Đặc trưng

Skewness	0 for $\nu > 3$, otherwise undefined
Ex. kurtosis	$\frac{6}{\nu-4}$ for $\nu > 4$, ∞ for $2 < \nu \le 4$, otherwise undefined
Entropy	$\frac{\nu+1}{2}\left[\psi\left(\frac{1+\nu}{2}\right) - \psi\left(\frac{\nu}{2}\right)\right]$ $+\ln\left[\sqrt{\nu}B\left(\frac{\nu}{2},\frac{1}{2}\right)\right]$ (nats) • ψ: digamma function, • *B*: beta function
MGF	undefined
CF	$\frac{K_{\nu/2}(\sqrt{\nu}\|t\|)\cdot(\sqrt{\nu}\|t\|)^{\nu/2}}{\Gamma(\nu/2)2^{\nu/2-1}}$ for $\nu > 0$ • $K_\nu(x)$: Modified Bessel function of the second kind[1]

6. Các bảng tra

Statistical Tables

Bảng tra của phân phối chuẩn tắc

566 APPENDIX A

NEGATIVE z Scores

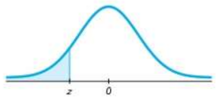
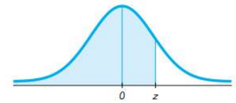


TABLE A-2 Standard Normal (z) Distribution: Cumulative Area from the LEFT

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.50 and lower	.0001									
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
−2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051 *	.0049	.0048
−2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
−2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
−2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
−2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
−2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
−1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
−1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
−1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367

Bảng tra của phân phối chuẩn tắc



Appendix A

POSITIVE z Scores

TABLE A-2 (*continued*) Cumulative Area from the LEFT

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495 *	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633

Bảng tra của phân phối Student

t Table

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$	$t_{.999}$	$t_{.9995}$
one-tail	0.50	0.25	0.20	0.15	0.10	0.05	0.025	0.01	0.005	0.001	0.0005
two-tails	1.00	0.50	0.40	0.30	0.20	0.10	0.05	0.02	0.01	0.002	0.001
df											
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66	318.31	636.62
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.000	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.000	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.000	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.000	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.000	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.000	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.000	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.000	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.000	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.000	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.000	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.579	3.883

Bảng tra của phân phối Student

$$P(|X|>t_{k,\alpha}) = \alpha \Leftrightarrow P(X<-t_{k,\alpha}) + P(X>t_{k,\alpha}) = \alpha$$

Do phân phối là đối xứng nên:

$P(X<-t_{k,\alpha}) = P(X>t_{k,\alpha}) = \alpha/2$

Giả sử: n = 3

df = n - 1 = 2

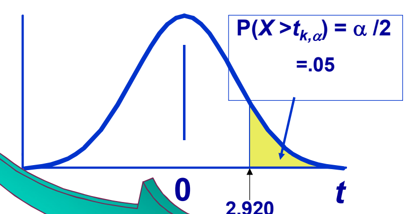
α = .10

α/2 =.05

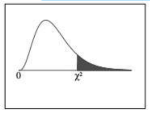
	Đuôi trên		
df \ α	.5 .25	.20 .10	.10 .05
1	1.000	3.078	6.314
2	0.817	1.886	2.920
3	0.765	1.638	2.353

α / 2

t Values



Bảng tra của phân phối Khi-Bình phương



The shaded area is equal to α for $\chi^2 = \chi^2_\alpha$.

df	$\chi^2_{.995}$	$\chi^2_{.990}$	$\chi^2_{.975}$	$\chi^2_{.950}$	$\chi^2_{.900}$	$\chi^2_{.100}$	$\chi^2_{.050}$	$\chi^2_{.025}$	$\chi^2_{.010}$	$\chi^2_{.005}$
1	0.000	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.070	12.833	15.086	16.750
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725	26.757
12	3.074	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.300
13	3.565	4.107	5.009	5.892	7.042	19.812	22.362	24.736	27.688	29.819
14	4.075	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319
15	4.601	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578	32.801
16	5.142	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000	34.267
17	5.697	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409	35.718
18	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.156
19	6.844	7.633	8.907	10.117	11.651	27.204	30.144	32.852	36.191	38.582
20	7.434	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566	39.997
21	8.034	8.897	10.283	11.591	13.240	29.615	32.671	35.479	38.932	41.401
22	8.643	9.542	10.982	12.338	14.041	30.813	33.924	36.781	40.289	42.796
23	9.260	10.196	11.689	13.091	14.848	32.007	35.172	38.076	41.638	44.181
24	9.886	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980	45.559
25	10.520	11.524	13.120	14.611	16.473	34.382	37.652	40.646	44.314	46.928
26	11.160	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642	48.290
27	11.808	12.879	14.573	16.151	18.114	36.741	40.113	43.195	46.963	49.645
28	12.461	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278	50.993

XIN CẢM ƠN!