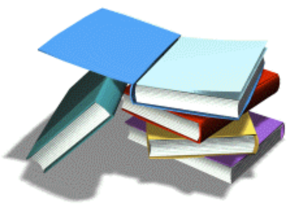
TRƯỜNG ĐẠI HỌC TÀI CHÍNH-MARKETING
KHOA CƠ BẢN
BỘ MÔN TOÁN – THỐNG KÊ

BÀI GIẢNG

LÝ THUYẾT XÁC SUẤT VÀ THỐNG KÊ TOÁN

Giảng viên
ThS. Lê Trường Giang



Chương 2

BIẾN NGẪU NHIÊN VÀ PHÂN PHỐI XÁC SUẤT

Bài 3. Một số luật phân phối xác suất thường dùng

1. Phân phối Bernoulli B(1, *p*)
2. Phân phối nhị thức B(*n*, *p*)
3. Phân phối siêu bội
4. Phân phối Poisson
5. Phân phối đều (SV tự đọc)
6. Phân phối chuẩn
7. Phân phối chi bình phương (SV tự đọc)
8. Phân phối Student (SV tự đọc)
9. Phân phối Fisher (SV tự đọc)

1. Phân phối Bernoulli B(1, *p*)

a. Định nghĩa

BNN X có tập giá trị $\mathrm{Im}(X)=\{0,1\}$ và bảng phân phối xác suất

X	0	1
B	1-p	p

gọi là có phân phối Bernoulli tham số $p\in(0,1)$ kí hiệu $X\sim B(1,p)$

b. Tham số đặc trưng

i. Kỳ vọng $E(X)=p$.

ii. Phương sai $Var(X)=p(1-p)$.

James BERNOULLI

1. Phân phối Bernoulli B(1, *p*)

c. Mô hình ứng dụng

Thực hiện phép thử Bernoulli, chỉ xảy ra hai kết quả. Một kết quả là sự kiện T xảy ra gọi là thành công với xác suất $p > 0$ và kết quả còn lại là $\overline{T}$ gọi là thất bại với xác suất $1 - p$. Xây dựng biến ngẫu nhiên X cho mô hình như sau $X\left(T\right)=1,\ X\left(\overline{T}\right)=0$.

Khi đó X tuân theo phân phối Bernoulli.

d. Ví dụ 1.

Phép thử tung đồng tiên cân đối đồng chất, kết quả xuất hiện mặt sấp S, $X\left(S\right)=1$ hoặc xảy ra mặt ngửa N, $X\left(N\right)=0$. Xác suất $P\left(X=1\right)=\dfrac{1}{2}$, vậy $X \sim B\left(1,\dfrac{1}{2}\right)$.

2. Phân phối nhị thức B(*n*, *p*)

a. Định nghĩa

Biến ngẫu nhiên X nhận các giá trị là số tự nhiên $\mathrm{Im}(X)=\{0,\ 1,\ 2,\dots,n\}$ với xác suất sau

$$P(X=k)=C_n^k p^k(1-p)^{n-k}$$

được gọi là có phân phối Nhị thức, kí hiệu là $X\sim B(n,p)$

b. Tham số đặc trưng

i. Kỳ vọng $E(X)=np$.

ii. Phương sai $Var(X)=np(1-p)$.

iii. Mode : $np-q\le Mod(X)\le np+p$

2. Phân phối nhị thức B(*n*, *p*)

Ví dụ 2. Một máy sản xuất lần lượt từng sản phẩm với xác suất có một phế phẩm là 3%. Cho máy sản suất ra 20 sản phẩm.

a) Tính xác suất trong 20 sản phẩm sản xuất ra có 4 phế phẩm?

b) Tìm số sản phẩm tốt trung bình trong 20 sản phẩm được sản xuất ra?

2. Phân phối nhị thức B(*n*, *p*)

Ví dụ 3 (BTN). Một nhà máy có 50 máy giống nhau hoạt động độc lập. Xác suất xảy ra hư hỏng của mỗi máy trong ngày là 0,05. Tính xác suất mỗi ngày có nhiều nhất là 2 máy hư hỏng? Số máy hư hỏng trung bình trong ngày?Nhiều khả năng nhất là bao nhiêu máy hỏng?

Ví dụ 4(BTN). Một người nuôi 200 con gà mái đẻ. Xác suất để 1 con gà đẻ trứng trong ngày là 80%.

a) Tìm số trứng gà trung bình thu được trong ngày?

b) Nếu muốn mỗi ngày thu được 300 trứng gà thì cần phải nuôi thêm bao nhiên con gà?

3. Phân phối siêu bội

Định nghĩa. Biến ngẫu nhiên X thuộc loại rời rạc, phụ thuộc vào ba tham số nguyên dương N, M và n. Xác suất của biến ngẫu nhiên X tại các giá trị k được cho như sau

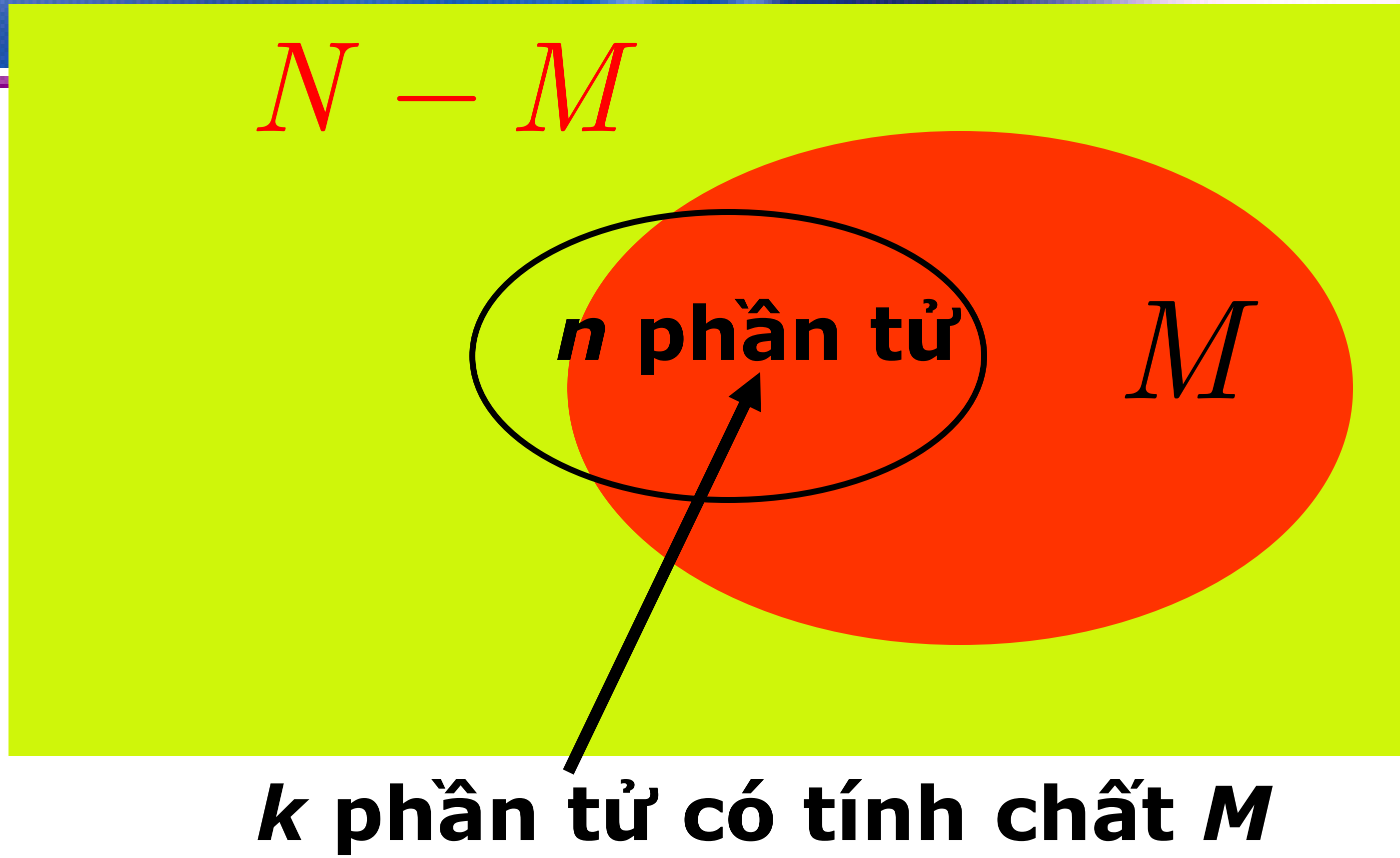
$$P(X=k)=\frac{C_M^k C_{N-M}^{n-k}}{C_N^n}, \quad k\in\mathbb{N}\cap\left[\max\{0,n-(N-M)\},\min\{n,M\}\right].$$

Biến ngẫu nhiên X có phân phối siêu bội kí hiệu là $X\sim H(N,M,n)$.

Đặc trưng số. $p=\dfrac{M}{N}$

i. Kỳ vọng $E(X)=np$.

ii. Phương sai $Var(X)=np(1-p)\dfrac{N-n}{N-1}$.



PHÂN PHỐI SIÊU Bội

3. Phân phối siêu bội

Xấp xỉ phân phối siêu bội bằng phân phối nhị thức

Cho $X \sim H(N,M,n)$. Trong thực tế nếu N khá lớn, n rất nhỏ so với N $(n < 0,05N)$, đặt $p = \frac{M}{N}$ thì

$$P(X = k) = \frac{C_M^k . C_{N-M}^{n-k}}{C_N^n} \approx C_n^k p^k q^{n-k}$$

3. Phân phối siêu bội

Ví dụ 5. Một ông chủ vườn lan đã để nhầm 20 chậu lan có hoa màu đỏ vào cùng với 100 chậu lan có hoa màu tím (lan chưa nở hoa). Một khách hàng chọn mua ngẫu nhiên đồng thời 15 chậu từ 120 chậu lan này.

a. Tính xác suất khách hàng mua được từ 5 đến 6 chậu lan có hoa màu đỏ?

b. Gọi X là số chậu lan có hoa màu đỏ mà khách hàng chọn được. Tính trung bình và phương sai của X.

3. Phân phối siêu bội

Ví dụ 6(BTN). Một lô hàng có 30 sản phẩm được đóng gói giống nhau, trong có có 5 sản phẩm lỗi kỹ thuật và 25 sản phẩm đạt tiêu chuẩn. Lấy ngẫu nhiên 10 sản phẩm từ lô hàng trên.

a. Xác định phân phối xác suất của biến ngẫu nhiên chỉ số sản phẩm đạt tiêu chuẩn?

b. Tính xác suất để có ít nhất 8 sản phẩm đạt tiêu chuẩn?

c. Tính trung bình số sản phẩm đạt tiêu chuẩn?

3. Phân phối siêu bội

Ví dụ 7 (BTN).

Một lô hàng có 5000 sản phẩm, trong đó có 250 phế phẩm. Lấy ngẫu nhiên 10 sản phẩm từ lô hàng này. Tính xác suất để trong 10 sản phẩm lấy ra có 2 phế phẩm? (tính theo hai cách)

4. Phân phối Piosson

Định nghĩa. BNN X được gọi là có phân phối Poisson với tham số dương λ, kí hiệu $X \sim Poisson(\lambda)$, nếu X nhận các giá trị $k = 0,1,2,\ldots$ với xác suất tương ứng là

$$P(X = k) = e^{-\lambda}\frac{\lambda^k}{k!}, \quad k = 0,1,2,\ldots$$

Đặc trưng số.

i. Kỳ vọng $E(X) = \lambda$.

ii. Phương sai $Var(X) = \lambda$.

4. Phân phối Piosson

Người đầu tiên mô tả phân phối Poisson là Simeon Denis Poisson vào năm 1837. Phân phối này có nhiều ứng dụng đối với các quá trình liên quan đến số quan sát trên một đơn vị thời gian hay không gian.

Thông tin cho biết của phân phối Poisson là trung bình số lần xảy ra thành công của một sự kiện trong một khoảng thời gian hay không gian nhất định. Giá trị này gọi là lambda, ký hiệu là λ.

Siméon Denis Poisson

(1781 - 1840)

4. Phân phối Piosson

BNN X chỉ số lần sự kiện A xảy ra trong một **khoảng** thời gian hay không gian liên tục nhất định.

Điều kiện đặt ra là số lần sự kiện A xảy ra trong những khoảng khác nhau là độc lập và số lần tỉ lệ thuận với chiều dài của khoảng. Trung bình số lần xảy ra của sự kiện A chính là λ được xác định

$$\lambda = c.\Delta_t .$$

Trong đó, c được gọi là cường độ (số lần trên một đơn vị thời gian), Δ_t là khoảng thời gian.

4. Phân phối Piosson

Ví dụ 8. Ở một tổng đài điện thoại, các cú điện thoại gọi đến xuất hiện ngẫu nhiên, độc lập với nhau và trung bình có 2 cuộc gọi đến trong một phút. Tìm xác suất để:

a. Có đúng 5 cú điện thoại gọi đến trong 2 phút.

b. Không có cú điện thoại nào gọi đến trong 30 giây.

c. Có ít nhất một cú điện thoại gọi đến trong 10 giây.

4. Phân phối Piosson

Ví dụ 9 (VN).

Một bệnh viện tiếp nhận trung bình 5 ca bệnh trong 30 phút. Tính xác suất

a. Bệnh viện tiếp nhận 17 ca bệnh trong một giờ?

b. Xác định khoảng thời gian Δ_t mà bệnh viên không tiếp nhận ca bệnh nào với xác suất 0,80?

Xấp xỉ phân phối nhị thức bằng phân phối Piosson

Thực hiện n lần phép thử Bernoulli độc lập, xác suất sự kiện A xảy ra trong mỗi phép thử là p_n. Nếu $\lim\limits_{n\to\infty} p_n = 0$ và $n.p_n = \lambda$, với λ là hằng số dương. Khi đó với $k \in \{0,1,\ldots,n\}$ ta có

$$\lim_{n\to\infty} C_n^k p_n^k \left(1-p_n\right)^{n-k} = \frac{\lambda^k}{k!}.e^{-\lambda}.$$

Trong thực tế ta có xấp xỉ, nếu $X \sim B\left(n,p\right)$ với $n > 30$ và $np \le 5$ hoặc $(1-p)n \le 5$ thì $X \sim P\left(np\right)$.

4. Phân phối Piosson

Ví dụ 10.

Cho biết xác suất trúng đích trong mỗi lần bắn của xạ thủ là 0,99. Xạ thủ bắn 500 phát, tính xác suất để có ít nhất 2 phát bắn trượt? (tính theo hai cách)

Ví dụ 11 (BTN). Một công ty sản xuất dược phẩm với tỉ lệ phế phẩm của những viên thuốc là 0,002. Nếu công ty sản xuất 1000 viên thuốc. Tính xác suất để có không quá 2 viên thuốc là phế phẩm.

4. Phân phối Piosson

Ví dụ 12 (VN). Một nhà máy sản xuất với tỉ lệ phế phẩm là 0,001. Nhập vào một lô hàng với số lượng là 1000 sản phẩm.

a. Tính xác suất để có một đến hai phế phẩm trong lô hàng này?

b. Tính xác suất để lô hàng nhập có phế phẩm?

Bài 3. Một số luật phân phối xác suất thường dùng
5. Phân phối đều

Định nghĩa. BNN X được gọi là có phân phối đều trên đoạn $\left[a,b\right]\left(a<b\right)$, kí hiệu là $X\sim U\left[a,b\right]$ nếu X có hàm mật độ là

$$f\left(x\right)=\begin{cases}\dfrac{1}{b-a} & x\in\left[a,b\right]\\ 0 & x\notin\left[a,b\right]\end{cases}$$

Đặc trưng số.

i. Kỳ vọng $E\left(X\right)=\dfrac{a+b}{2}$.

ii. Phương sai $Var\left(X\right)=\dfrac{\left(b-a\right)^2}{12}$.

iii. Mod(X) là bất kì giá trị nào trên [a,b].

6. Phân phối chuẩn

Định nghĩa. BNN X có hàm mật độ xác suất f phụ thuộc hai tham số μ và σ^2 $(\sigma > 0)$

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}.e^{-\frac{1}{2}\cdot\left(\frac{x-\mu}{\sigma}\right)^2},\ x \in \mathbb{R}$$

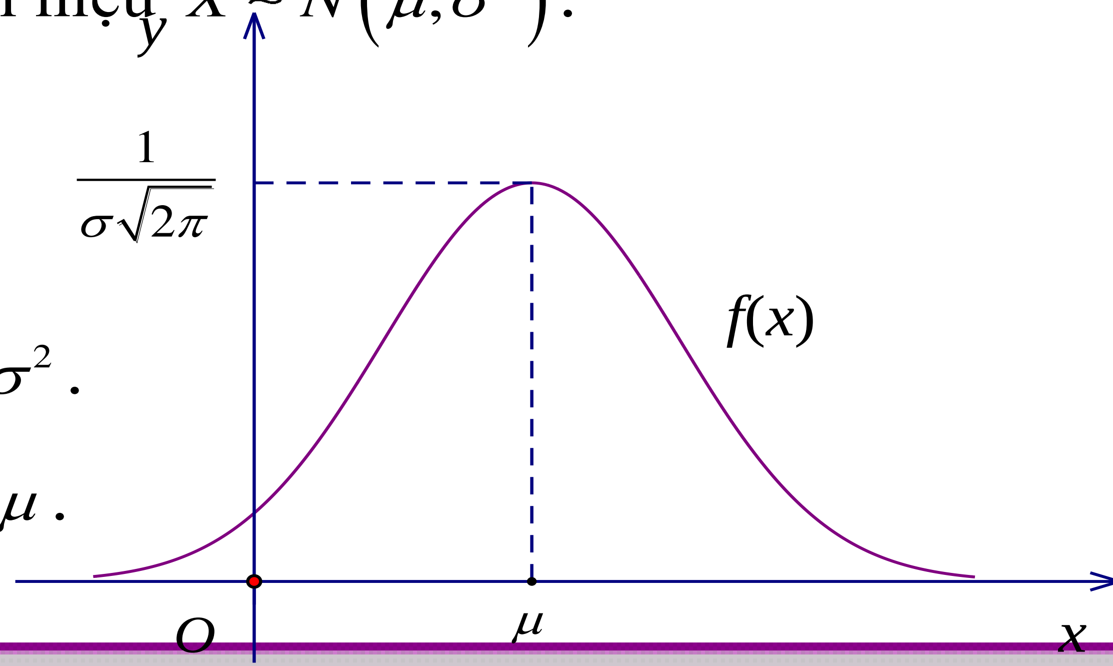
gọi là có phân phối chuẩn, kí hiệu $X \sim N(\mu, \sigma^2)$.

Đặc trưng số.

i. Kỳ vọng $E(X) = \mu$.

ii. Phương sai $Var(X) = \sigma^2$.

iii. Mod(X) = Med(X) = μ.



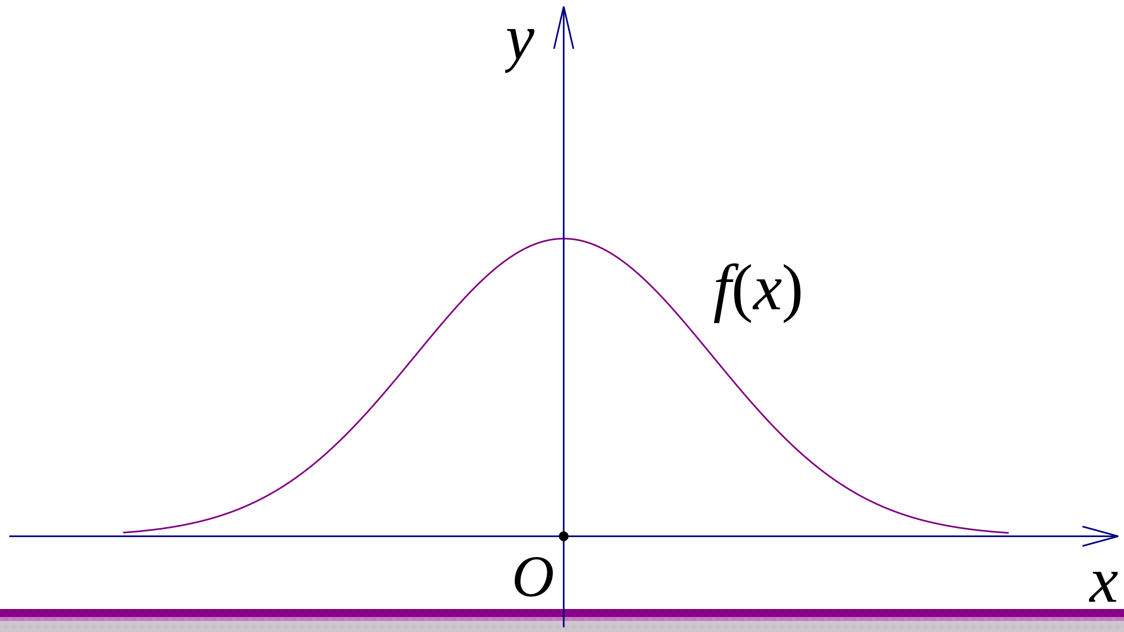
Phân phối chuẩn do Carl F. Gauss đưa ra năm 1795.

Carl Friedrich Gauss
(1777 - 1855)

6. Phân phối chuẩn

Dạng chuẩn tắc kí hiệu $X \sim N(0;1)$ có hàm mật độ

$$f(x) = \frac{1}{\sqrt{2\pi}}.e^{-\frac{x^2}{2}},\ x \in \mathbb{R}$$

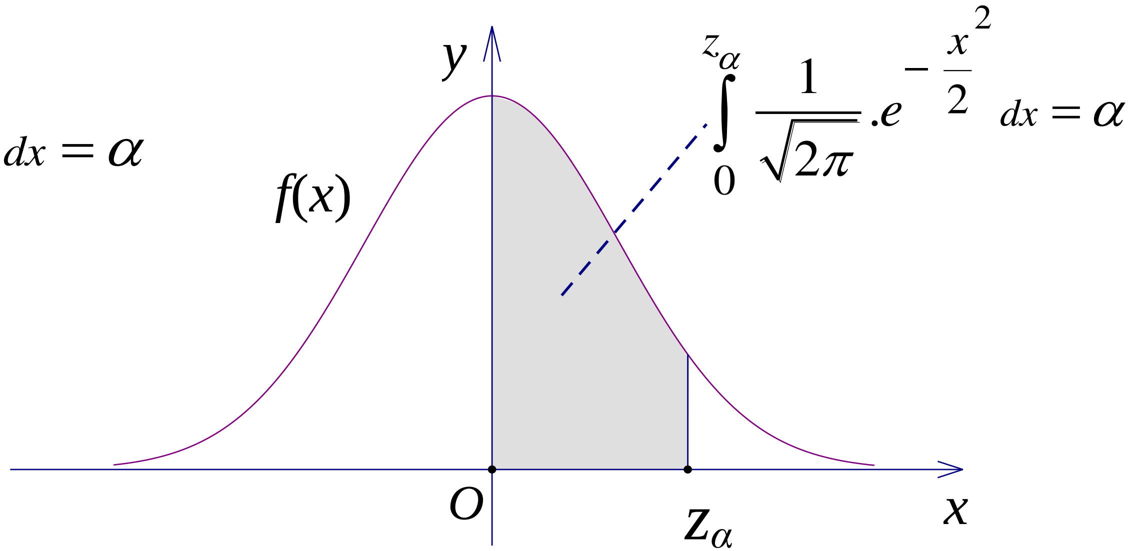


6. Phân phối chuẩn

Bảng tích phân Laplace.

Các giá trị của hàm φ được tính trong bảng Laplace

$$\varphi(z_\alpha) = \int_0^{z_\alpha} \frac{1}{\sqrt{2\pi}}.e^{-\frac{x^2}{2}}dx = \alpha$$



Hàm φ có tính chất $\varphi(-z) = -\varphi(z)$. Nếu gọi $Z \sim N(0;1)$ thì

$$P(0 \le Z \le z_\alpha) = \varphi(z_\alpha) = \alpha.$$

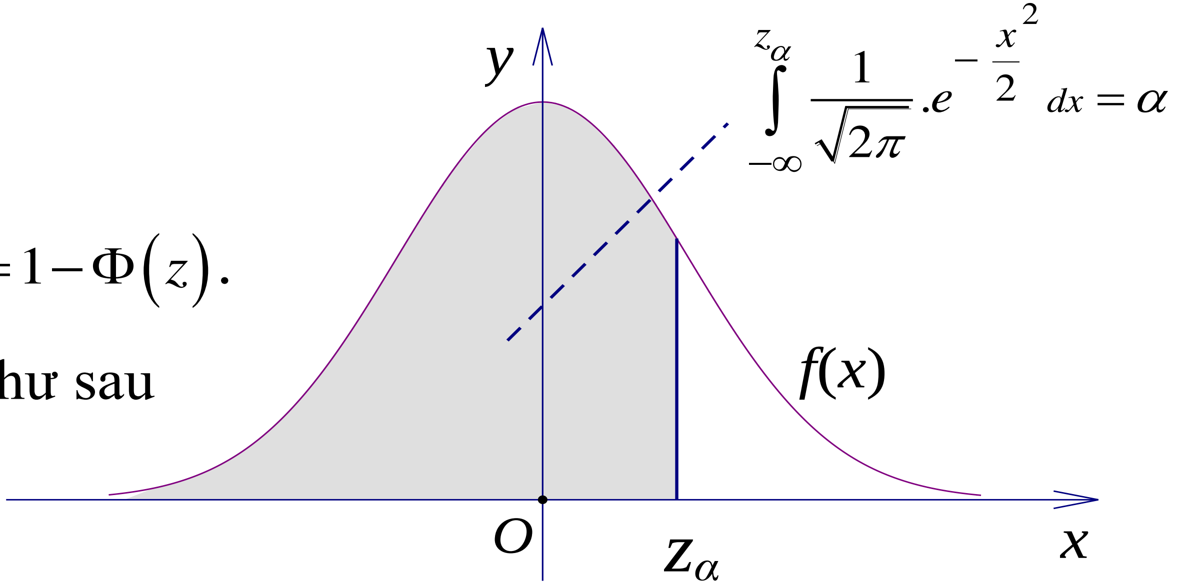
Ta có $\varphi(+\infty) = 0,5$ và $\varphi(-\infty) = -0,5$. Nếu $a > 5$ thì $\varphi(a) \approx 0,5$.

6. Phân phối chuẩn

Bảng phân vị chuẩn mức α.

Các giá trị của Φ cho trong bảng phân vị chuẩn

$$\Phi\left(z_\alpha\right)=\int_{-\infty}^{z_\alpha}\frac{1}{\sqrt{2\pi}}.e^{-\frac{x^2}{2}}dx=\alpha$$



Hàm Φ có tính chất $\Phi\left(-z\right)=1-\Phi\left(z\right)$.

Mối liên hệ giữa φ và Φ như sau

$$\varphi\left(z\right)=\Phi\left(z\right)-\frac{1}{2}.$$

Nếu gọi $Z\sim N\left(0;1\right)$ thì $P\left(Z\le z_\alpha\right)=\Phi\left(z_\alpha\right)=\alpha$.

Tính chất của hàm phân phối ta có ngay $\Phi\left(+\infty\right)=1$ và $\Phi\left(-\infty\right)=0$.

6. Phân phối chuẩn

Công thức tính xác suất

Với $Z \sim N(0;1)$, dùng bảng tích phân Laplace của hàm φ hoặc phân vị chuẩn Φ cho trong bảng tính các xác suất sau.

i. $P(a \le Z \le b) = \varphi(b) - \varphi(a) = \Phi(b) - \Phi(a)$

ii. $P(Z \le b) = \varphi(b) + 0,5 = \Phi(b).$

iii. $P(a \le Z) = 0,5 - \varphi(a) = 1 - \Phi(a).$

Với $X \sim N(\mu, \sigma^2)$, ta chuẩn hóa X là $Z = \dfrac{X - \mu}{\sigma^2}$ và sử dụng các kết quả trên

Bài 3. Một số luật phân phối xác suất thường dùng
6. Phân phối chuẩn

Công thức ứng với $\varphi(x)=\int_0^x \frac{1}{\sqrt{2\pi}}.e^{-\frac{t^2}{2}}dt$

$$P(\alpha \le X \le \beta)=\varphi\left(\frac{\beta-\mu}{\sigma}\right)-\varphi\left(\frac{\alpha-\mu}{\sigma}\right);$$

$$P(X \le \beta)=\varphi\left(\frac{\beta-\mu}{\sigma}\right)+\frac{1}{2};\ P(X \ge \alpha)=\frac{1}{2}-\varphi\left(\frac{\alpha-\mu}{\sigma}\right);$$

$$\varphi(+\infty)=\frac{1}{2};\ \varphi(-\infty)=-\frac{1}{2};\ \varphi(0)=0;\varphi(-z)=-\varphi(z).$$

$$P(|X-\mu|<\varepsilon)=2\varphi\left(\frac{\varepsilon}{\sigma}\right).$$

6. Phân phối chuẩn

Công thức ứng với $\Phi(x) = \int\limits_{-\infty}^{x} \frac{1}{\sqrt{2\pi}}.e^{-\frac{t^2}{2}} dt$

$$P(\alpha \le X \le \beta) = \Phi\left(\frac{\beta - \mu}{\sigma}\right) - \Phi\left(\frac{\alpha - \mu}{\sigma}\right);$$

$$P(X \le \beta) = \Phi\left(\frac{\beta - \mu}{\sigma}\right) + 1;\ P(X \ge \alpha) = 1 - \Phi\left(\frac{\alpha - \mu}{\sigma}\right);$$

$$\Phi(+\infty) = 1;\ \Phi(-\infty) = -1;\ \ \Phi(0) = 0.5;\ \ \Phi(-z) = 1 - \Phi(z).$$

$$P(|X - \mu| < \varepsilon) = 2\Phi\left(\frac{\varepsilon}{\sigma}\right) - 1.$$

6. Phân phối chuẩn

Quy tắc $k\sigma$. Cho biến ngẫu nhiên $X \sim N\left(\mu,\sigma^2\right)$.

Ta có

$$P\left(|X-\mu| \le \sigma\right) = 2\varphi(1) = 0,68268$$

$$P\left(|X-\mu| \le 2\sigma\right) = 2\varphi(2) = 0,9545$$

$$P\left(|X-\mu| \le 3\sigma\right) = 2\varphi(3) = 0,9973$$

Ta nhận thấy hầu như chắc chắn X sẽ lấy giá trị trong khoảng $[\mu-3\sigma;\mu+3\sigma]$

Định lý. Cho X_1, X_2 độc lập và $X_1 \sim N\left(\mu_1,\sigma_1^2\right)$,

$X_2 \sim N\left(\mu_2,\sigma_2^2\right)$.

Với $a,b \in \mathbb{R}$ ta có $aX_1 + bX_2 \sim N\left(\mu_1+\mu_2, a^2\sigma_1^2 + b^2\sigma_2^2\right)$.

Xấp xỉ phân phối nhị thức bằng phân phối chuẩn

Cho $X \sim B(n,p)$ thỏa $\begin{cases} n > 30 \\ np \geq 5 \\ n(1-p) \geq 5 \end{cases}$ thì $X \sim N(np, npq)$.

Ta có công thức xấp xỉ

$$P(k_1 \leq X \leq k_2) \approx \varphi\left(\frac{k_2 - np}{\sqrt{npq}}\right) - \varphi\left(\frac{k_1 - np}{\sqrt{npq}}\right)$$

$$P(X = k) \approx \frac{1}{\sqrt{npq}} f\left(\frac{k - np}{\sqrt{npq}}\right)$$

6. Phân phối chuẩn

Ví dụ 13. Điểm thi Toeic của sinh viên năm cuối ở một trường đại học là BNN X có phân phối chuẩn với giá trị trung bình là 560 điểm và độ lệch chuẩn là 78 điểm. Tính:

a. Tỷ lệ sinh viên có điểm từ 600 đến 700 điểm?

b. Tỷ lệ sinh viên có điểm thi trên 500 điểm?

6. Phân phối chuẩn

Ví dụ 14 (BTN). Thời gian để sản xuất một sản phẩm loại A là biến ngẫu nhiên tuân theo luật phân phối chuẩn với trung bình 10 và độ lệch chuẩn 1 (đơn vị là phút).

a. Tính xác suất để một sản phẩm loại A nào đó được sản xuất trong khoảng thời gian từ 9 phút đến 12 phút?

b. Tính thời gian cần thiết để sản xuất một sản phẩm loại A bất kỳ?

6. Phân phối chuẩn

Ví dụ 15 (VN). Trọng lượng của một loại cá một năm tuổi được nuôi trong hồ là biến ngẫu có phân phối chuẩn với trung bình 1,2kg và độ lệch chuẩn 0,2kg.

a. Tính tỉ lệ cá từ 0,9kg đến 1,1kg?

b. Trọng lượng cá từ x gam trở lên đạt tỉ lệ 90%, hãy tìm giá trị của x ?

6. Phân phối chuẩn

Ví dụ 16. Người ta lấy ngẫu nhiên 400 hạt lúa từ một kho lúa rất lớn, có tỉ lệ hạt lép là 0,2 để kiểm tra. Tính xác suất để trong đó có từ 80 đến 100 hạt lép.

Ví dụ 17 (BTN). Theo thống kê của một công ty sản xuất xe máy thì có 60% khách mua của họ là phụ nữ. Chọn ngẫu nhiên 200 người mua xe của công ty này. Tính xác suất để trong 200 người này có ít nhất 140 phụ nữ.

7. Phân phối chi bình phương

Định nghĩa. Biến ngẫu nhiên X có hàm mật độ xác suất f

$$f(x)=\begin{cases}\dfrac{1}{2^{\frac{n}{2}}.\Gamma\left(\dfrac{n}{2}\right)}x^{\frac{n}{2}-1}e^{-\frac{x}{2}} & \text{nếu } x>0\\ 0 & \text{nếu } x\le 0.\end{cases}$$

Trong đó, Γ là hàm Gamma xác định như sau

$$\Gamma(\alpha)=\int_0^{+\infty} y^{\alpha-1}e^{-y}dy\,.$$

Biến ngẫu nhiên X được xác định như trên gọi là tuân theo luật phân phối chi bình với n bậc tự do, kí hiệu $X\sim\chi^2(n)$.

7. Phân phối chi bình phương

Đặc trưng số.

i. Kỳ vọng $E(X) = n$.

ii. Phương sai $Var(X) = 2n$.

Định lý. Biến ngẫu nhiên X có phân phối chuẩn tắc $X \sim N(0;1)$ thì biến ngẫu nhiên $X^2 \sim \chi^2(1)$.

Định lý. Cho X_1, X_2 độc lập và $X_1 \sim \chi^2(k_1), X_2 \sim \chi^2(k_2)$.

Khi đó tổng $X_1 + X_2 \sim \chi^2(k_1 + k_2)$.

- Phân phối $\chi^2(n)$ do Karl Pearson đưa ra năm 1900.

Karl Pearson (1857 - 1936)

8. Phân phối Student

Định nghĩa. Cho Z và V độc lập, $Z \sim N(0;1)$, $V \sim \chi^2(n)$. Khi đó, BNN T được xác định như sau

$T = \dfrac{Z}{\sqrt{\dfrac{V}{n}}}$ được gọi là luật Student với *n* bậc tự do, kí hiệu là $T \sim t(n)$.

BNN $T \sim t(n)$ có hàm mật độ xác suất f

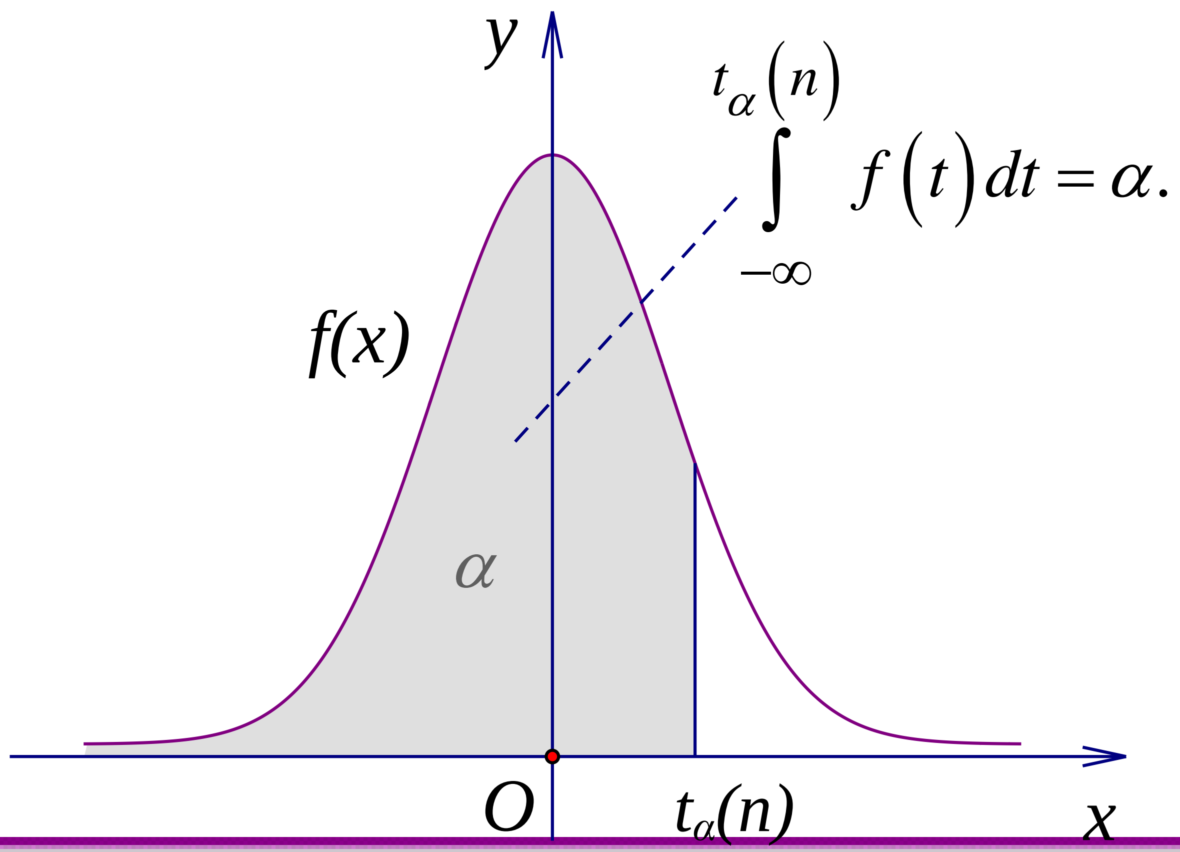
$$f(x) = \frac{\Gamma\left(\dfrac{n+1}{2}\right)}{\sqrt{n\pi}.\Gamma\left(\dfrac{n}{2}\right)}\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}} \qquad \forall x \in \mathbb{R}, \quad n > 0.$$

Đặc trưng số.

i. Kỳ vọng $E(X) = 0$. *ii.* Phương sai $Var(X) = \dfrac{n}{n-2}$.

8. Phân phối Student

Bảng phân vị. BNN $T \sim t(n)$, giá trị phân vị $t_\alpha(n)$ và xác suất $P\left(T \le t_\alpha(n)\right)$ được cho trong bảng Student.



- Phân phối $St(n)$ do Willam.S.Gosset đưa ra năm 1908.

Willam. Sealy .Gosset 1876 - 1937

XIN CHÂN THÀNH CẢM ƠN!