西北工業大學
NORTHWESTERN POLYTECHNICAL UNIVERSITY

概率论与数理统计

徐爽

西北工业大学理学院概率统计教研室

第一节 一维随机变量及其分布(3)

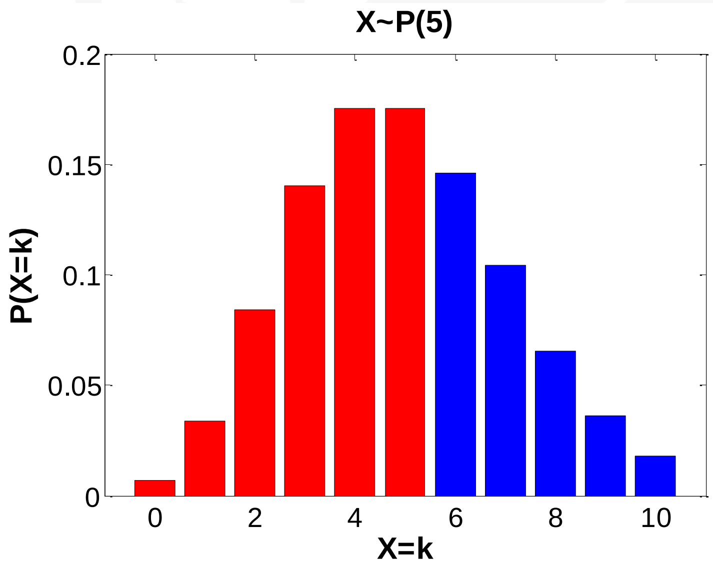
- 五、连续型随机变量
- 六、典型的连续型随机变量及其分布

离散型随机变量：X的取值**有限或无限可列个**。如人数、产品数、事件发生的次数等。

例如$X \sim P(5)$

$$P\{X \le a\} = F(a) = \sum_{x_i \le a} P(X = x_i); \quad P\{X = a\} = F(a) - F(a^-)$$

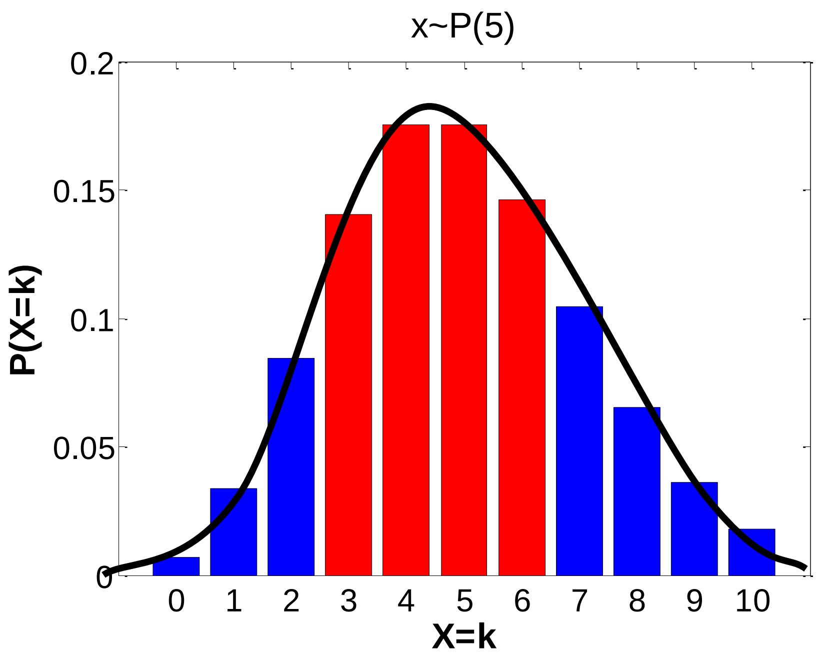
$$P\{a < X \le b\} = F(b) - F(a) = \sum_{a < x_i \le b} P(X = x_i)$$



$$P\{X \le 5\} = F(5)$$

$$= \sum_{x_k \le 5} P\{X = x_k\}$$

$$= \sum_{k=0}^{5} \frac{5^k}{k!} e^{-5} \approx 0.616$$



$$P\{3 \le X < 6\} =$$

$$= \sum_{3 \le x_k < 6} P\{X = x_k\}$$

$$\approx 0.6375$$

连续型随机变量：X的取值为**无限不可列个**。

如速度、候车时间、寿命、降水量等。

$P\{X \le x\} = ?$ $P\{X = a\} = ?$ $P\{a < X \le b\} = ?$

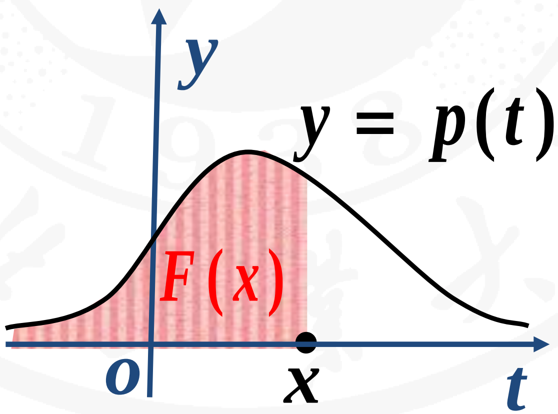
五、连续型随机变量

1. 连续型随机变量及其密度函数

定义 对于随机变量X，若存在**非负可积**函数 $p(x)$ ($x\in R$), 使得X 的分布函数

$$F(x)=P(X\leq x)=\int_{-\infty}^{x}p(y)\mathrm{d}y$$

则称X为**连续型随机变量**,且称$p(x)$ 为**密度函数**,或**概率密度**.

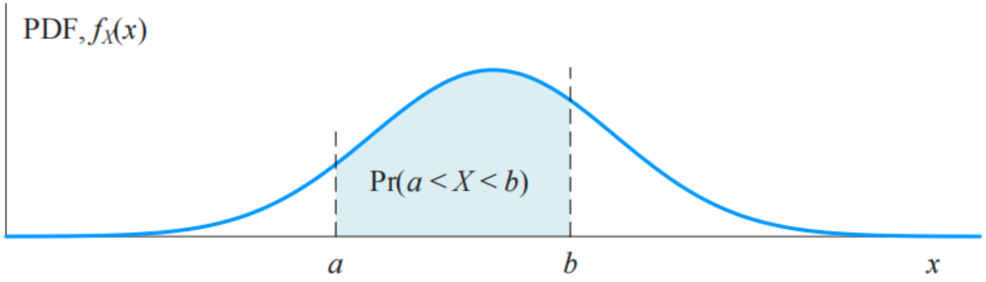


2. 密度函数的性质

设X为连续型随机变量, $p(x)$ 为X的密度函数,

$F(x)$为X的分布函数 ,则

(1) $p(x) \geq 0, \quad x \in R;$



(2) $\int_{-\infty}^{+\infty} p(x)\mathrm{d}x = 1;$

(3) $P\{a < X \leq b\} = F(b) - F(a) = \int_a^b p(x)\mathrm{d}x;$

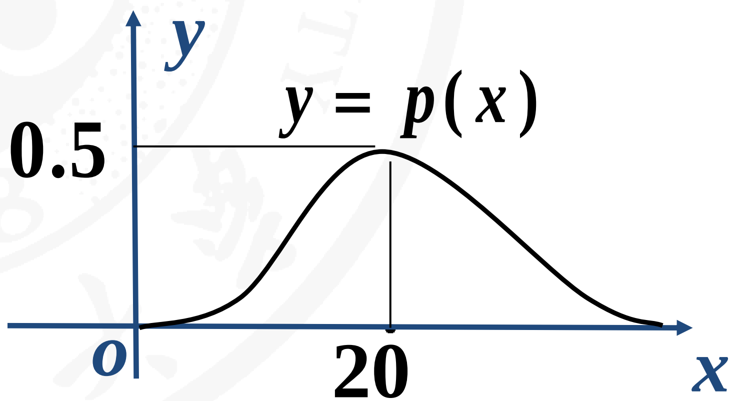
(4) $P\{X = c\} = F(c) - F(c^-) = 0.$

注 1° 性质4说明对于任意可能值c,连续型随机变量取c的概率等于零.

例 X: 西安地区某一天的降雨量 $X \geq 0$

$P\{X=20\}=\cancel{0.5}$?

$$P\{X=20\}=\int_{20^-}^{20} 0.5dx = 0$$



2° 若X为连续型随机变量，则

$$P\{a < X \le b\} = P\{a < X < b\}$$

$$= P\{a \le X < b\} = P\{a \le X \le b\}$$

连续型随机变量的概率与区间的开闭无关

3° $$P(A)=0 \;\overset{\nRightarrow}{\Leftarrow}\; A=\Phi \qquad \because A=\{X=a\}$$

$$P(A)=1 \;\overset{\nRightarrow}{\Leftarrow}\; A=\Omega \qquad \because A=\Omega-\{X=a\}$$

4. （连续型随机变量）描述随机事件的概率

$$P\{X\in D\}=\int_{x\in D} p(x)dx$$

例 1　某仪器电子元件的寿命(单位:h) 分布具有如下密度函数

$$p(x)=\begin{cases}\frac{1}{600}e^{-\frac{x}{600}}, & x>0,\\ 0, & x\le 0.\end{cases}$$

试求在仪器使用的最初200h内,该电子元件损坏的概率。

解　A:{在仪器使用最初200h内电子元件损坏}

设X表示元件的使用寿命.

$$P(A)=P\{X\le 200\}$$

$$p(x)=\begin{cases}\frac{1}{600}\mathrm{e}^{-\frac{x}{600}}, & x>0,\\ 0, & x\le 0.\end{cases}$$

$$=P\{0<X\le 200\}=\int_0^{200}\frac{1}{600}\mathrm{e}^{-\frac{x}{600}}\mathrm{d}x=1-\mathrm{e}^{-\frac{1}{3}}$$

法 2

$$F(x)=\int_0^x\frac{1}{600}\mathrm{e}^{-\frac{y}{600}}dy$$

$$=-\mathrm{e}^{-\frac{y}{600}}\Big|_0^x=1-\mathrm{e}^{-\frac{x}{600}}$$

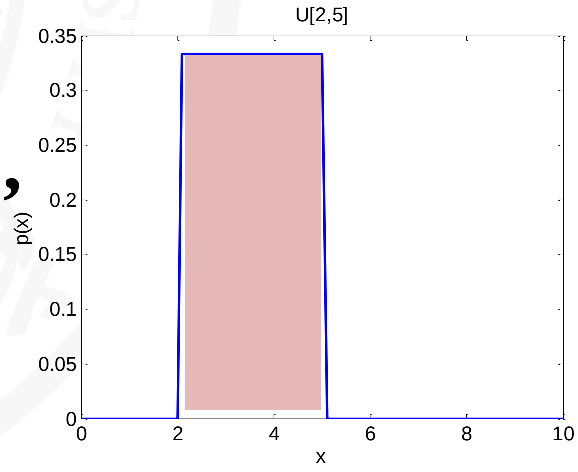
$$\therefore P(A)=P\{X\le 200\}=F(200)=1-\mathrm{e}^{-\frac{1}{3}}$$

六、典型的连续型随机变量的分布

1.均匀分布

(1) 定义 设连续型随机变量 X 具有概率密度:

$$p(x)=\begin{cases}\dfrac{1}{b-a}, & a\le x\le b,\\ 0, & \text{其它}.\end{cases}$$

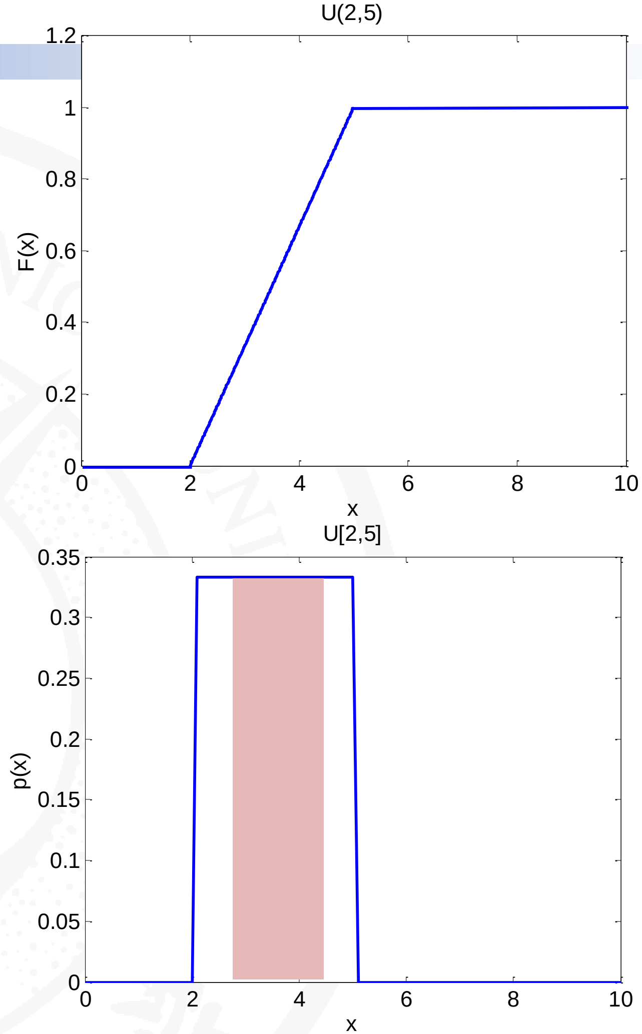


则称 X在区间$[a,b]$上服从均匀分布,

记为 $X\sim U[a,b]$.

分布函数为:

$$F(x)=\begin{cases}0, & x\le a,\\ \dfrac{x-a}{b-a}, & a<x\le b,\\ 1, & x>b.\end{cases}$$



(2) 均匀分布的性质

如果$X\sim U[a,b]$,则

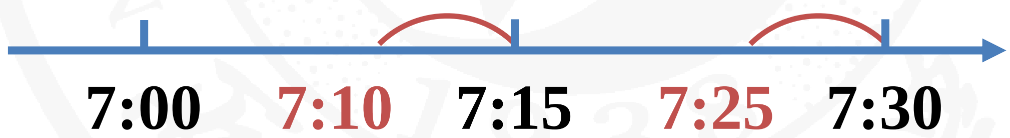
1° $P\{X<a\}=P\{X>b\}=0;$

2° 当$a\le c<d\le b$时，有$P\{c\le X<d\}=\dfrac{d-c}{b-a}.$

例1：

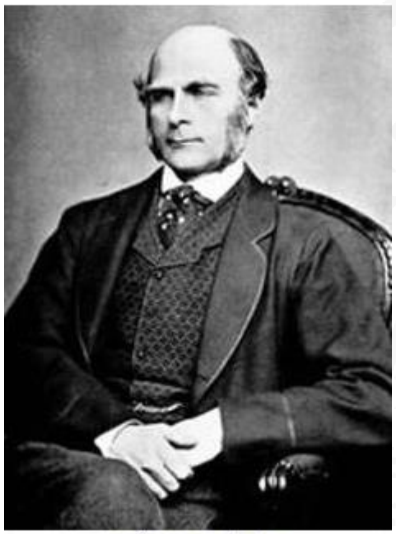
- 某城际列车从早上六点整开始每15 分钟发出一趟列车。
- 假设某乘客达到车站的时间服从七点到七点半的均匀分布。
- 若忽略买票等其它时间,试求该乘客等车少于5分钟的概率。

解： 只要到达的时刻在发车前5分钟即可。



$$P(\{10 < X < 15\} \cup \{25 < X < 30\}) = \frac{10}{30}$$

2.正态分布(Gauss分布)

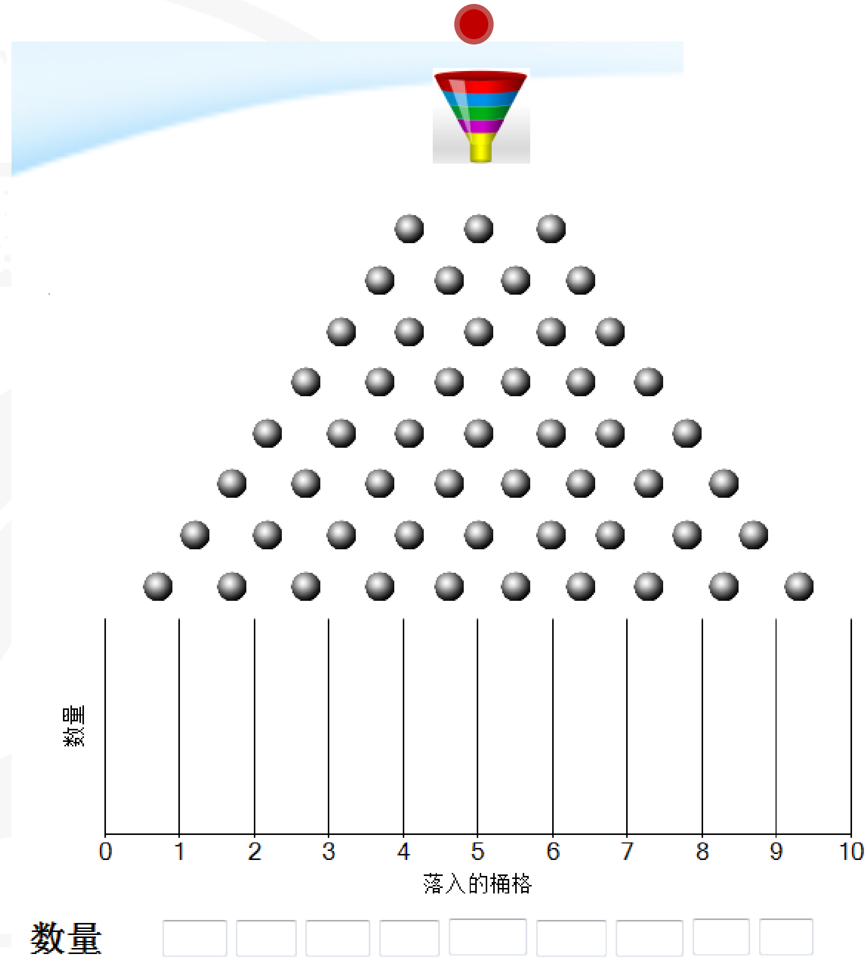


Francis Galton
(1822-1911)

$$P_{左}=P_{右}=1/2$$

每个格子中小球的个数

统计特性

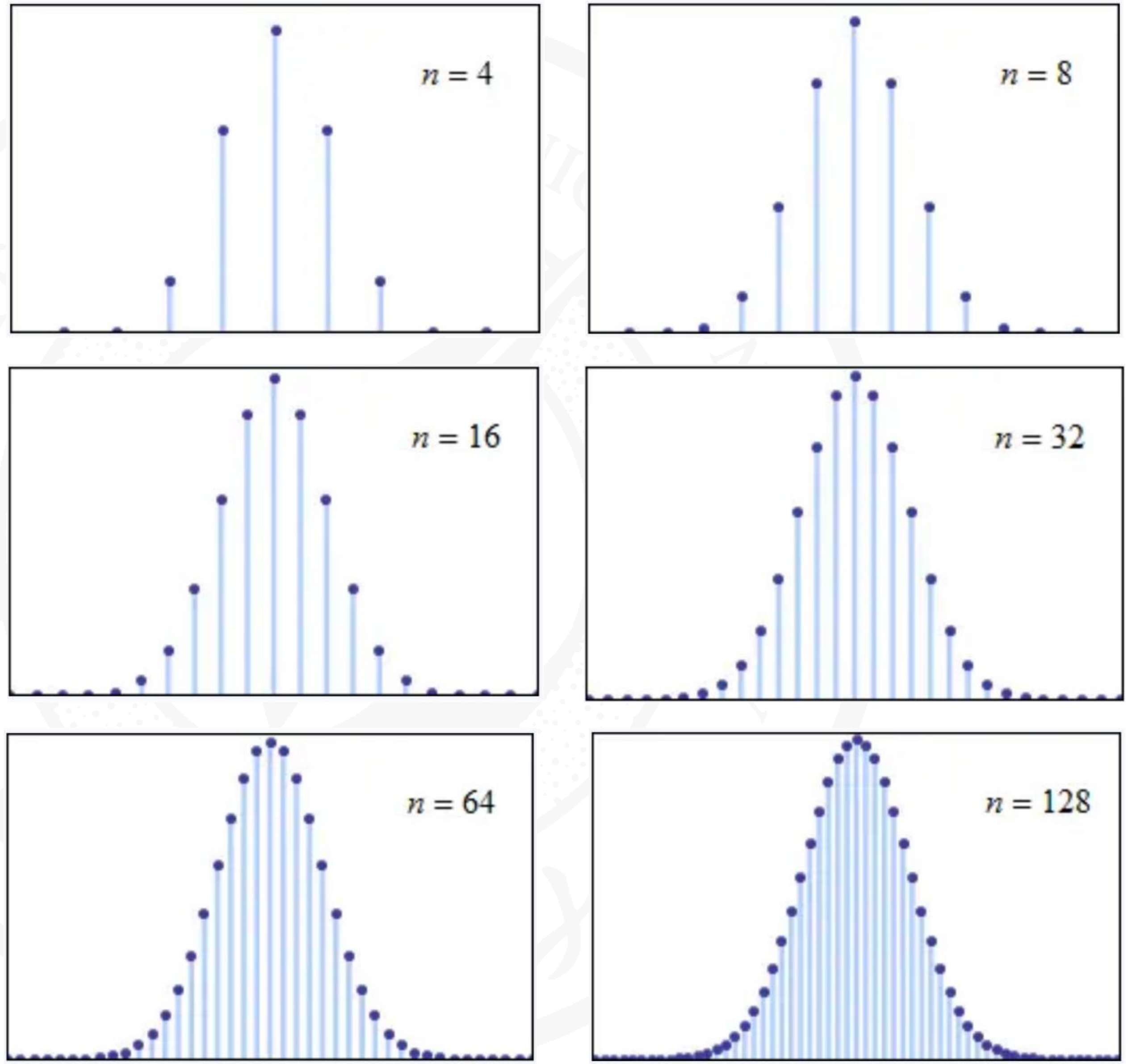


高尔顿(Galton)板实验

“钟型” 曲线

数学上描述统计特性？

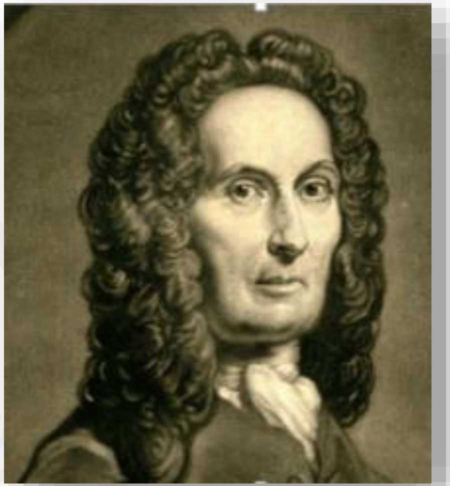
$$C_n^k\left(\frac{1}{2}\right)^k\left(\frac{1}{2}\right)^{n-k}$$



(1) 定义

若随机变量 X 的密度函数为

$$\boxed{p(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad -\infty<x<+\infty}$$



法国数学家 棣莫佛

A. De Moivre

B. (1667-1754)

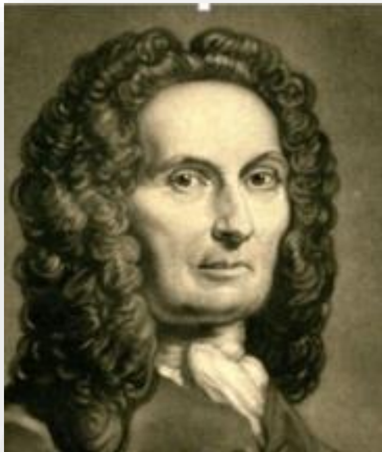
其中 $-\infty<\mu<+\infty,\sigma>0$ 为常数，则称 X

为服从参数 μ 和 σ 的正态分布[1]，记作 $X\sim N(\mu,\sigma^2)$.

分布函数： $F(x)=\dfrac{1}{\sqrt{2\pi}\sigma}\displaystyle\int_{-\infty}^{x}e^{-\frac{(t-\mu)^2}{2\sigma^2}}\,\mathrm{d}t$

THE
NORMAL
LAW OF ERROR
STANDS OUT IN THE
EXPERIENCE OF MANKIND
AS ONE OF THE BROADEST
GENERALIZATIONS OF NATURAL
PHILOSOPHY · IT SERVES AS THE
GUIDING INSTRUMENT IN RESEARCHES
IN THE PHYSICAL AND SOCIAL SCIENCES AND
IN MEDICINE AGRICULTURAL,AND ENGINEERING
IT ISAN INDISPENSABLE TOOL FOR THE ANALYSIS AND THE
INTERPRETATION OF BASIC DATA OBTAINED BY OBSERVATION AND EXPERIMENT

拓展阅读： 靳志辉，正态分布的前世今生

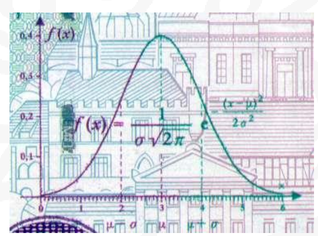


法国数学家 棣莫佛
(1667-1754)

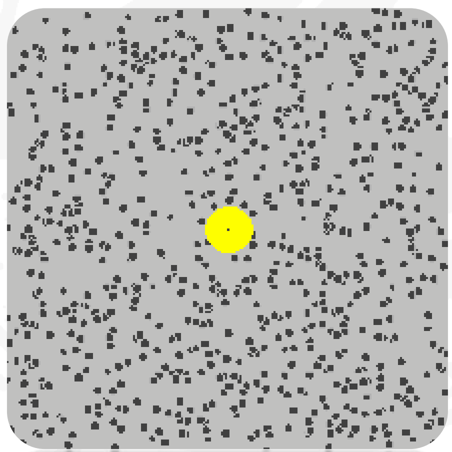
德国数学家、物理学家、天文学家 高斯 (1777-1855)

正态分布

高斯分布



正态分布是自然界及工程技术中最常见的分布之一，**大量的随机现象**例如测量误差、分子的运动、同龄人的身高等都可以用正态分布描述。



正态分布 $\mu=0,\ \sigma^2=1$ 标准正态分布

$X \sim N(\mu,\sigma^2)$ → $X \sim N(0,1)$

密度函数

$$\varphi(x)=\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}} \quad -\infty<x<\infty$$

分布函数

$$\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\mathrm{e}^{-\frac{t^2}{2}}\mathrm{d}t$$

(2) 正态分布密度函数的性质

$$p(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

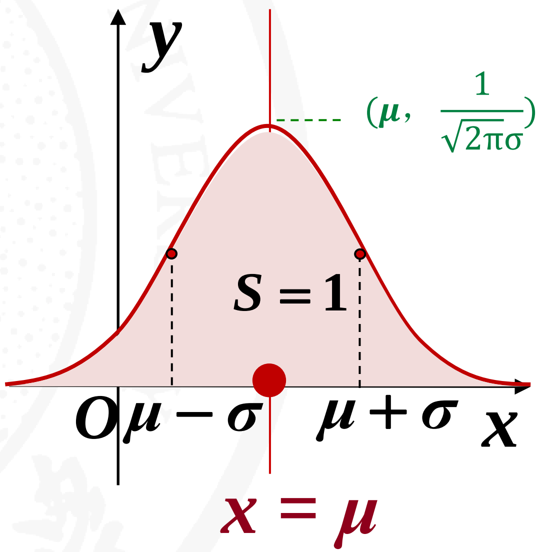
1) 曲线关于 $x=\mu$ 对称;

2) 当$x=\mu$时, $p(x)$取得最大值$\dfrac{1}{\sqrt{2\pi}\sigma}$;

3) $p(x)$以 x 轴为渐近线,

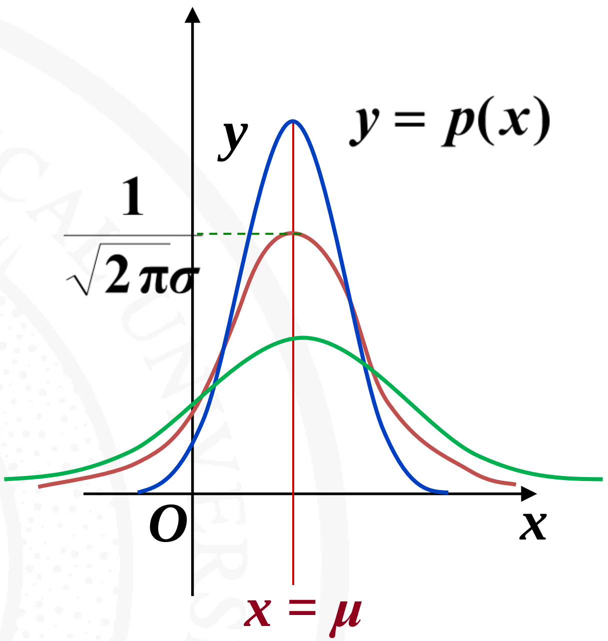
即当$x\to\pm\infty$时，$p(x)\to 0$

4) 曲线在 $x=\mu\pm\sigma$ 处有拐点;



5) $\int_{-\infty}^{+\infty}p(x)dx=1$, 即$p(x)$与 x轴围城的面积为1。

$$p(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$



6) 当固定 σ, 改变 μ 的大小时, $p(x)$ 的形状不变,只是沿着 x 轴作平移;

7) 当固定 μ, 改变 σ 的大小时, $p(x)$ 的对称轴不变, 而形状在改变.

σ 越小, $p(x)$越高越瘦。

σ 越大，$p(x)$越矮越胖.

(3) 标准正态分布的性质

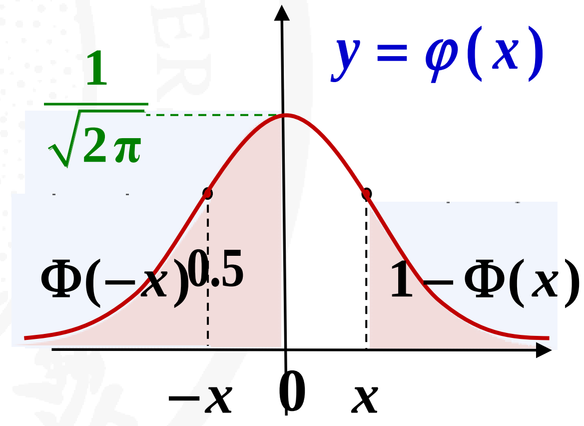
$$\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\mathrm{e}^{-\frac{t^2}{2}}\mathrm{d}t$$

1) $\varphi(x)$为偶函数；

2) $M=\max\varphi(x)=\varphi(0)=\frac{1}{\sqrt{2\pi}}$;

3) $\int_{-\infty}^{+\infty}\varphi(x)dx=1$ $\Rightarrow\int_{-\infty}^{+\infty}\mathrm{e}^{-\frac{x^2}{2}}dx=\sqrt{2\pi}$

$\Rightarrow\int_{-\infty}^{+\infty}\mathrm{e}^{-t^2}dt=\sqrt{\pi}$



4) $\Phi(0)=\int_{-\infty}^{0}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}dx=0.5$;

5) $\Phi(-x)=1-\Phi(x)$

(4) 正态分布概率的计算

$X \sim N(\mu, \sigma^2)$ $\qquad$ $P(X \le b)$

统计规律性 ⇨ **随机事件的概率**

$$P(X \le b) = F(b) = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{b} \boxed{e^{-\frac{(t-\mu)^2}{2\sigma^2}}}\,dt$$

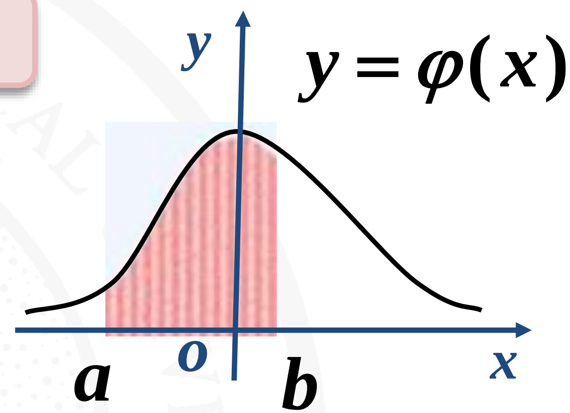
?

标准正态分布：

当$X \sim N(0,1)$时，

$$P\{a < X \le b\} = \frac{1}{\sqrt{2\pi}}\int_a^b \boxed{e^{-\frac{x^2}{2}}}dx$$

非初等函数



查标准正态分布表

$$\Phi(z) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z} e^{-\frac{x^2}{2}}dx \qquad \Rightarrow P\{a < X \le b\} = \Phi(b) - \Phi(a)$$

实例 若$X \sim N(0,1)$，求$P\{1 < X \le 2\}$.

解 $P\{1 < X \le 2\} = \Phi(2) - \Phi(1) \quad = 0.9773 - 0.8413 = 0.136$

查附表2，$\Phi(2) = 0.9773, \Phi(1) = 0.8413$

非标准正态分布：

标准化 如果 $X \sim N(\mu,\sigma^2)$, $Y=\dfrac{X-\mu}{\sigma}$,那么 $Y \sim N(0,1)$.

证明：

$$F_Y(y)=P\{Y\le y\}=P\{\frac{X-\mu}{\sigma}\le y\}$$

$$=P\{X\le \sigma y+\mu\}$$

$$=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\sigma y+\mu}\exp\{-\frac{(x-\mu)^2}{2\sigma^2}\}\mathrm{d}x$$

令 $t=\dfrac{x-\mu}{\sigma}$,则

$$F_Y(y)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{y}\exp\{-\frac{t^2}{2}\}\mathrm{d}t$$

所以 $Y \sim N(0,1)$.

标准化 如果$X \sim N(\mu,\sigma^2)$, $Y=\dfrac{X-\mu}{\sigma}$,那么$Y \sim N(0,1)$.

则 $P\{a < X \le b\}$ $=P\{\dfrac{a-\mu}{\sigma} < \boxed{\dfrac{X-\mu}{\sigma}} \le \dfrac{b-\mu}{\sigma}\}$

查标准正态分布表 $=\Phi(\dfrac{b-\mu}{\sigma})-\Phi(\dfrac{a-\mu}{\sigma})$

例 2 设随机变量X服从正态分布$N(\mu,\sigma^2)$

则随着σ的增大,概率$P\left\{|X-\mu|<\sigma\right\}$是( )

A 单调增加　　*B* 单调减少

C 保持不变　　*D* 非单调变化

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

可见概率不随σ的增大而变化，故选3

例 3 若中国人的身高X（单位：厘米）服从$\mu = 170$, $\sigma = 6$的正态分布，那么公交车门的高度最低设计为多少可以使上下车需要低头的比例不超过0.5%？

解 记高度为h，则$P\{X < h\} = 99.5\%$

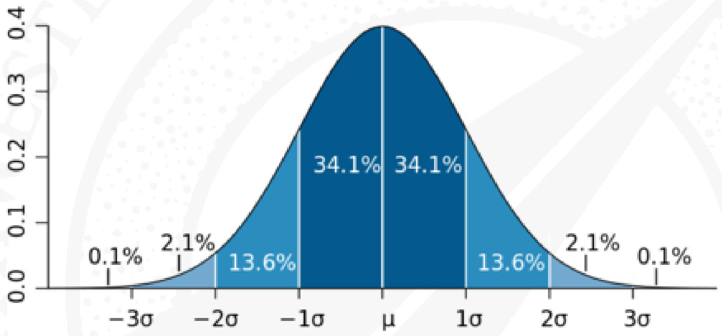
由$Y = \dfrac{X-\mu}{\sigma} \sim N(0,1)$可得

$$P\left\{\frac{X-170}{6} \le \frac{h-170}{6}\right\} = 99.5\%$$

查附表2得$\Phi(2.58) \approx 0.995$, 则

$$\frac{h-170}{6} = 2.58 \qquad \Rightarrow h = 170 + 6 \times 2.58 = 185.48$$

思考：查询资料，解读



六西格玛法则

例 4

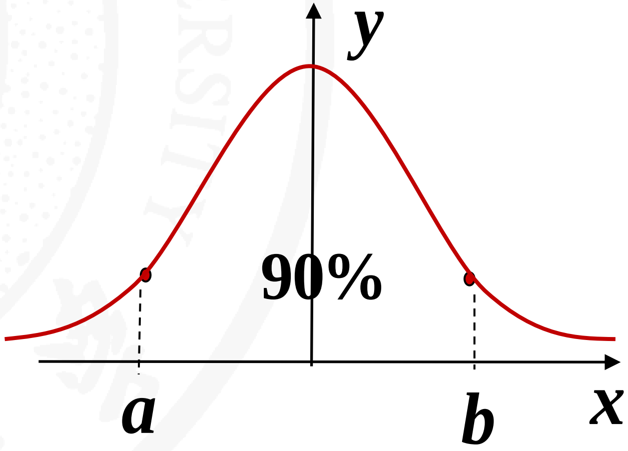
广西玉林市皮肤病防治站检验报告单

姓名： 病人类型： 门诊 床号： 标 本 号： 32
性别： 住 院 号： 费别： 标本类型： 全血
年龄： 岁 科 室： 门诊 诊断： 备 注：

代号	项 目	结 果	参 考 值	代号	项 目	结 果	参 考 值
WBC	白细胞	6.90	4--10 10^9/L	HCT	红细胞压值	0.463	0.35--0.51
LYMPH%	淋巴细胞百分比	31.40	20--40 %	MCV	平均红细胞体积	85.9	85--96 fL
MONO%	单核细胞百分比	10.00	3--10 %	MCH	平均红细胞血红蛋白量	28.9	26--34 pg
NEUT%	中性粒细胞百分比	56.60	46--64 %	MCHC	平均红细胞血红蛋白浓度	337	300--380 g/L
EO%	嗜酸细胞百分比	1.60	0.5--5 %	RDW-CV	红细胞分布宽度(CV)	11.8	10.9--15.4 %
BASO%	嗜碱细胞百分比	0.40	0--1 %	RDW-SD	红细胞分布宽度(SD)	36.7	35.1--46.3 fL
LYMPH#	淋巴细胞绝对值	2.17	1.5--4 10^9/L	PLT	血小板	277	100--300 10^9/L
MONO#	单核细胞绝对值	0.69	0.12--4 10^9/L	PDW	血小板分布宽度	8.8↓	10.5--18.1 fL
NEUT#	中性粒细胞绝对值	3.90	2--7 10^9/L	MPV	血小板平均体积	8.3	7.6--13.2 fL
EO#	嗜酸细胞绝对值	0.11	0--0.45 10^9/L	P-LCR	大血小板比率	13.1	13--43 %
BASO#	嗜碱细胞绝对值	0.03	0--0.2 10^9/L	PCT	血小板比积	0.23	0.11--0.28 %
RBC	红细胞	5.39	3.5--5.5 10^12/L				
HGB	血红蛋白	156↑	110--150 g/L				

例 4

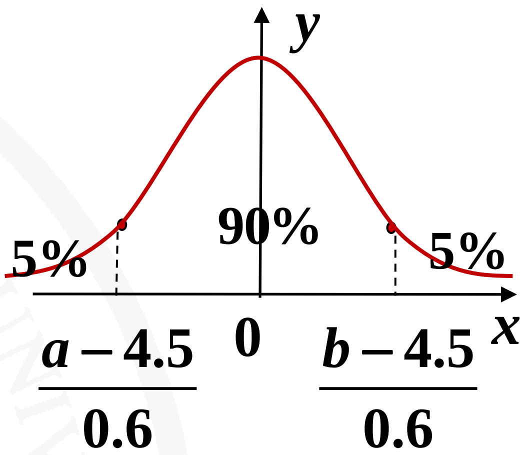
某地区抽样调查结果表明成年男子的红细胞数服从正态分布 $N(4.5, 0.6^2)$，试求该地区成年男子红细胞数90%的医学参考范围。



解: X：成年男子红细胞数

$$X \sim N(4.5, 0.6^2)$$

设参考值的上、下限分别为b和a



则 $P\{a \le X \le b\}$

$$= P\{\frac{a-4.5}{0.6} \le \frac{X-4.5}{0.6} \le \frac{b-4.5}{0.6}\}$$

$$= \Phi(\frac{b-4.5}{0.6}) - \Phi(\frac{a-4.5}{0.6}) = 0.9 \qquad (1)$$

假设参考范围为对称区间， 则 $\frac{b-4.5}{0.6} = -\frac{a-4.5}{0.6}$

$$\because \Phi(-x) = 1 - \Phi(x)$$

$$\therefore \Phi(\frac{b-4.5}{0.6}) = 1 - \Phi(\frac{a-4.5}{0.6}) \qquad (2)$$

广西玉林市皮肤病防治站检验报告单

姓名：	病人类型：门诊	床号：	标 本 号：32
性别：	住 院 号：	费别：	标本类型：全血
年龄：　　岁	科　　室：门诊	诊断：	备　　注：

代号	项　　目	结 果	参 考 值	代号	项　　目	结 果	参 考 值
WBC	白细胞	6.90	4--10 10^9/L	HCT	红细胞压值	0.463	0.35--0.51
LYMPH%	淋巴细胞百分比	31.40	20--40 %	MCV	平均红细胞体积	85.9	85--96 fL
MONO%	单核细胞百分比	10.00	3--10 %	MCH	平均红细胞血红蛋白量	28.9	26--34 pg
NEUT%	中性粒细胞百分比	56.60	46--64 %	MCHC	平均红细胞血红蛋白浓度	337	300--380 g/L
EO%	嗜酸细胞百分比	1.60	0.5--5 %	RDW-CV	红细胞分布宽度(CV)	11.8	10.9--15.4 %
BASO%	嗜碱细胞百分比	0.40	0--1 %	RDW-SD	红细胞分布宽度(SD)	36.7	35.1--46.3 fL
LYMPH#	淋巴细绝对值	2.17	1.5--4 10^9/L	PLT	血小板	277	100--300 10^9/L
MONO#	单核细胞绝对值	0.69	0.12--4 10^9/L	PDW	血小板分布宽度	8.8↓	10.5--18.1 fL
NEUT#	中性粒细胞绝对值	3.90	2--7 10^9/L	MPV	血小板平均体积	8.3	7.6--13.2 fL
EO#	嗜酸细胞绝对值	0.11	0--0.45 10^9/L	P-LCR	大血小板比率	13.1	13--43 %
BASO#	嗜碱细胞绝对值	0.03	0--0.2 10^9/L	PCT	血小板比积	0.23	0.11--0.28 %
RBC	红细胞	5.39	3.5--5.5 10^12/L				
HGB	血红蛋白	156↑	110--150 g/L				

.05

该地区成年男子红细胞数的90%的医学参考范围为[3.51，5.48]