# 第四章 正态分布

刘 春 光

暨南大学数学系

2018年4月

# 目录

# 目录

# 正态分布

**Definition (正态分布(Normal distribution)、标准正态分布(standard normal distribution))**

如果随机变量$X$有以下概率密度

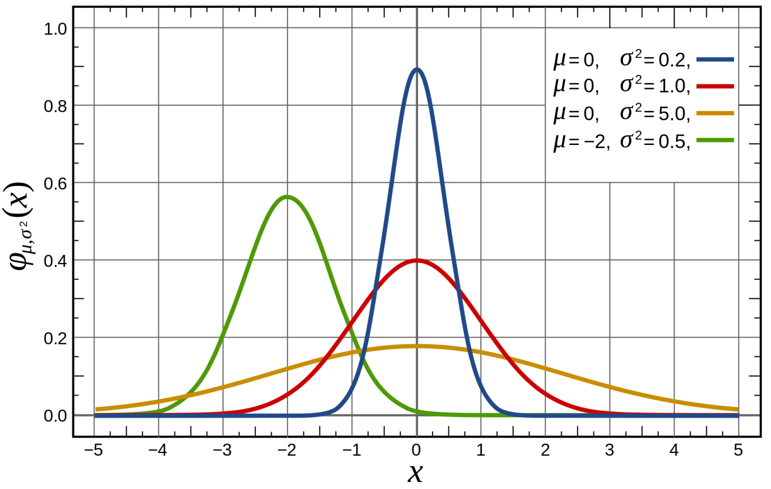
$$\varphi_{\mu,\sigma^2}(x) = \frac{1}{\sqrt{2\pi}\cdot\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}},$$

其中$\mu,\sigma$为常数且$\sigma > 0$，则称$X$服从正态分布。简记为

$$X \sim N(\mu,\sigma^2).$$

称$N(0,1)$为标准正态分布，并简写$\varphi_{0,1}(x)$为$\varphi(x)$。

# 正态分布的密度函数


$\mu=0,\quad \sigma^2=0.2,$
$\mu=0,\quad \sigma^2=1.0,$
$\mu=0,\quad \sigma^2=5.0,$
$\mu=-2,\quad \sigma^2=0.5,$
1.0
0.8
0.6
0.4
0.2
0.0
$\varphi_{\mu,\sigma^2}(x)$
−5
−4
−3
−2
−1
0
1
2
3
4
5
$x$

# 正态分布的密度函数

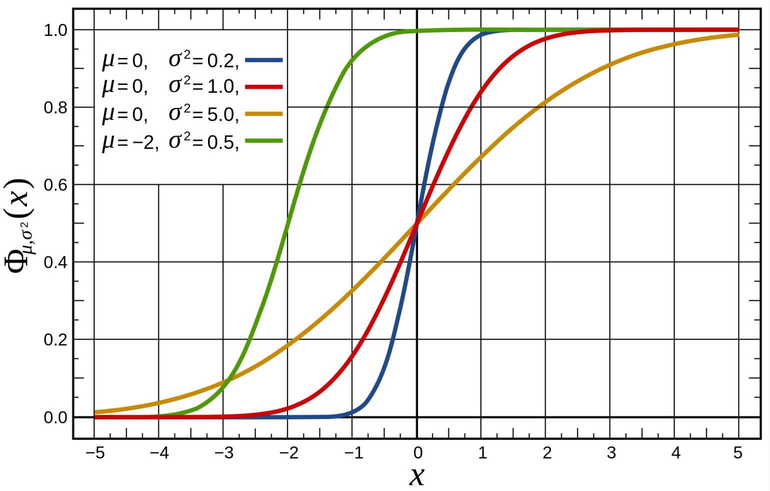

GN4480100S8
DEUTSCHE BUNDESBANK
Banknote
10
Deutsche Bundesbank
Frankfurt am Main
1. September 1999
1777–1855 Carl Friedr. Gauß
ZEHN DEUTSCHE MARK
GN4480100S8

# 标准正态分布的密度函数

标准正态分布的概率密度函数$\varphi(\cdot)$有以下性质：

1. 无穷次可微；
2. 偶函数；
3. 在零点取得最大值；
4. 有拐点$\pm 1$；
5. 有水平渐近线（$x$轴）。

# 正态分布的分布函数


1.0
0.8
0.6
0.4
0.2
0.0
$\mu=0,\quad \sigma^2=0.2,$
$\mu=0,\quad \sigma^2=1.0,$
$\mu=0,\quad \sigma^2=5.0,$
$\mu=-2,\quad \sigma^2=0.5,$
$\Phi_{\mu,\sigma^2}(x)$
−5
−4
−3
−2
−1
0
1
2
3
4
5
$x$

# 常用连续型随机变量

正态分布的分布函数为

$$\Phi_{\mu,\sigma^2}(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\cdot\sigma} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt,$$

该函数不是初等函数。标准正态分布的分布函数简记为$\Phi(x)$。

## 标准正态分布的分布函数

标准正态分布的分布函数满足性质

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

例：已知$\xi \sim N(0,1)$，求

$$P(\xi \leq 1.5), \quad P(\xi \leq -1.5), \quad P(|\xi| \leq 1.5),$$
$$P(\xi \leq 3.2), \quad P(-1 < \xi \leq 2).$$

# 标准正态分布的分布函数

总结：若$X \sim N(0,1)$，则对任意$x > 0$,

$$P\{|X| < x\} = 2\Phi(x) - 1.$$

# 正态分布的分布函数

一般正态分布的分布函数$\Phi_{\mu,\sigma^2}(x)$可以表示为

$$\Phi_{\mu,\sigma^2}(x) = \Phi\left(\frac{x-\mu}{\sigma}\right).$$

事实上，若$X \sim (\mu, \sigma^2)$，则

$$\frac{X-\mu}{\sigma} \sim N(0,1).$$

## 正态分布的概率

例1 设$X \sim N(\mu, \sigma^2)$，求

$P(|X-\mu| < \sigma)$，　　$P(|X-\mu| < 2\sigma)$，

$P(|X-\mu| < 3\sigma)$.

$3\sigma$原则：在应用中，对$X \sim N(\mu, \sigma^2)$，通常认为$X$只取$(\mu - 3\sigma, \mu + 3\sigma)$中的值。计算可知，使用该原则犯错误的概率不到千分之三。

## 正态分布的概率

练习：设$\xi \sim N(8, 0.25)$，求

$$P(|\xi - 8| < 1), \quad P(\xi \leq 10).$$

练习：设$\xi \sim N(\mu, \sigma^2)$，且

$$P(\xi \leq -5) = 0.0446, \quad P(\xi \leq 3) = 0.6179.$$

求$\mu, \sigma$。

# 正态分布

例2 设$X \sim N(0,1)$，求$Y = X^2$的概率密度函数。

# 目录

# 正态分布的数字特征

设随机变量$X \sim N(\mu, \sigma^2)$，则

$$EX = \mu, \qquad DX = \sigma^2.$$

进一步，$X$的$k$阶中心矩为

$$\mu_k(X) = E[(X - EX)^k] = \begin{cases} 0, & k = 1, 3, 5, \cdots, \\ (k-1)!!\sigma^k, & k = 2, 4, 6, \cdots. \end{cases}$$

## 正态分布

例1 设$X \sim N(0,1)$，求$Y = X^2$的期望与方差。

例2 某次考试的成绩$X$近似服从正态分布，平均分为75分。已知95分以上的考生比例为$2.3\%$，求这次考试的不及格率（60分及以上为及格）。

# 目录

# 二维正态分布

**Definition (二元正态分布(Bivariate normal distribution))**

以以下函数为密度的分布称为二元正态分布，

简记为$N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$

$$f(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \cdot \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right]\right\},$$

其中$\mu_1, \mu_2$为实数，$\sigma_1, \sigma_2 > 0$，$|\rho| < 1$。

# 二维正态分布

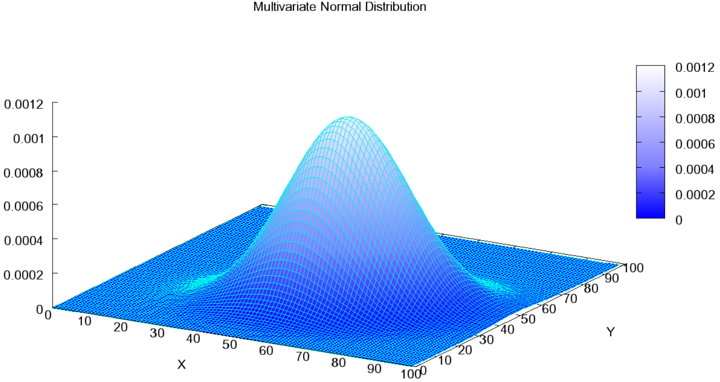



Figure: 二维正态分布的密度函数

## 二维正态分布

Theorem (二维正态分布的数字特征)

设随机向量$(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$，则

$$X\sim N(\mu_1,\sigma_1^2),\quad Y\sim N(\mu_2,\sigma_2^2),$$

且

$$R(X,Y)=\rho.$$

Corollary

对服从二维正态分布的随机向量$(X,Y)$，$X$与$Y$相互独立当且仅当它们的相关系数为零。

# 二维正态分布

例1-2 设随机变量$X$与$Y$相互独立，均服从标准正态分布，其它们的平方和$Z = X^2 + Y^2$的概率密度、数学期望及方差。

## 二维正态分布

例3 设随机向量$(X,Y)$的联合概率密度为

$$f(x,y)=\frac{1}{2\pi\sigma^2}\cdot\exp\left\{-\frac{1}{2\sigma^2}\left[(x-\mu_1)^2+(y-\mu_2)^2\right]\right\},$$

其中$\sigma>0$，求随机变量$U$与$V$的相关系数，其中$U=aX+bY$，$V=aX-bY$，$a,b$为不全为零的常数。

## $n$维正态分布

例4 设随机向量$(X,Y)$的联合概率密度为

$$f(x,y)=\frac{1}{2\pi\sigma^2}\cdot\exp\left\{-\frac{1}{2\sigma^2}(x^2+y^2)\right\},$$

其中$\sigma>0$，求$X$与$Y$的最大值$Z=\max\{X,Y\}$的数学期望。

## $n$维正态分布

二元正态分布的密度可写为以下形式：

$$\varphi(\vec{x}) = \frac{1}{2\pi|B|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(\vec{x}-\vec{\mu})^T B^{-1}(\vec{x}-\vec{\mu})\right\},$$

其中

$$\vec{x} = \begin{pmatrix} x \\ y \end{pmatrix}, \vec{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, B = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$$

此时该分布也可简记为$N(\vec{\mu}, B)$。

# $n$维正态分布

**Definition (多元正态分布(Multivariate normal distribution))**

给定$n$维向量$\vec{\mu} = (\mu_1, \mu_2, \cdots, \mu_n)^T$及$n$阶对称正定矩阵$B$，以

$$\varphi(\vec{x}) = \frac{1}{(2\pi)^{\frac{n}{2}}|B|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(\vec{x}-\vec{\mu})^T B^{-1}(\vec{x}-\vec{\mu})\right\}$$

为密度函数的连续型分布称为$n$元正态分布，记为$N(\vec{\mu}, B)$。

# $n$维正态分布

Theorem ($n$维正态分布的数字特征)

设$n$维随机向量$\vec{X}=(X_1,X_2,\cdots,X_n)^T$服从正态分布$N(\vec{\mu},B)$，则

$$E\vec{X}=\vec{\mu},\quad \mathrm{cov}\vec{X}=B.$$

# 目录

# 一维正态随机变量的线性函数

**Theorem**

设$X \sim N(\mu,\sigma^2)$，则$Y=a+bX(b\neq 0)$也服从正态分布，且有

$$Y=a+bX\sim N(a+b\mu,b^2\sigma^2).$$

**Corollary**

设$X \sim N(\mu,\sigma^2)$，则

$$\frac{X-\mu}{\sigma}\sim N(0,1).$$

# 一维正态随机变量的线性函数

例1 假设测量的误差$X \sim N(0, 10^2)$，试求在100次独立重复测量中，至少有三次测量误差的绝对值大于19.6的概率$\alpha$，并利用泊松分布求概率$\alpha$的近似值。

# 相互独立正态随机变量的线性组合

**Theorem**

设随机变量$X$与$Y$相互独立，且都服从正态分布：

$$X \sim N(\mu_1, \sigma_1^2), \quad Y \sim N(\mu_2, \sigma_2^2),$$

则它们的和也服从正态分布，且有

$$X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2).$$

# 相互独立正态随机变量的线性组合

## Theorem

设随机变量$X_1, X_2, \cdots, X_n$相互独立，且都服从正态分布：

$$X_k \sim N(\mu_k, \sigma_k^2), \quad k = 1, 2, \cdots, n,$$

则它们的线性组合也服从正态分布，且有

$$\sum_{i=1}^{n} c_k X_k \sim N\left(\sum_{k=1}^{n} c_k \mu_k, \sum_{k=1}^{n} c_k^2 \sigma_k^2\right),$$

其中$c_1, c_2, \cdots, c_n$为常数。

# 相互独立正态随机变量的线性组合

例2 设随机变量$X$与$Y$相互独立，且

$$X \sim N(0, 4^2),\ Y \sim N(0, 3^2),$$

求$Z = |X - Y|$的数学期望与方差。

## ★补充：特征函数

**Definition (特征函数(Characteristic function))**

设$F(x)$为一个分布函数，称

$$h(t) = \int_{-\infty}^{+\infty} e^{itx} dF(x),$$

为$F(x)$的特征函数，其中$i = \sqrt{-1}$为虚数单位。如果$F(x)$为随机变量$X$的分布函数，则$h(t)$也称为$X$的特征函数，此时有

$$h(t) = E(e^{itX}).$$

## ★补充：特征函数

若离散随机变量$X$的分布为

$$P\{X = x_k\} = p_k, \qquad k = 1, 2, \cdots,$$

则其特征函数为

$$h(t) = \sum_k e^{itx_k} p_k.$$

若连续随机变量$X$的密度为$f(x)$，则其特征函数为

$$h(t) = \int_{-\infty}^{+\infty} e^{itx} f(x) dx.$$

## ★补充：特征函数

**Example (常用离散分布的特征函数)**

1. 几何分布$G(p)$的特征函数为

$$h(t)=\sum_{k=1}^{\infty}e^{itk}pq^{k-1}=\frac{pe^{it}}{1-qe^{it}}.$$

2. 二项分布$B(n,p)$的特征函数为

$$h(t)=\sum_{k=0}^{n}e^{itk}C_n^kp^kq^{n-k}=(q+pe^{it})^n.$$

## ★补充：特征函数

**Example (常用离散分布的特征函数)**

3. 泊松分布$P(\lambda)$的特征函数为

$$h(t) = \sum_{k=0}^{n} e^{itk} \cdot \frac{\lambda^k}{k!} e^{-\lambda} = e^{\lambda(e^{it}-1)}.$$

## ★补充：特征函数

Example (常用连续分布的特征函数)

1. 均匀分布$U[a,b]$的特征函数为
$$h(t)=\int_a^b \frac{e^{itx}}{b-a}dx=\frac{e^{itb}-e^{ita}}{it(b-a)}.$$
2. 指数分布$e(\lambda)$的特征函数为
$$h(t)=\int_0^{+\infty} e^{itx}\cdot\lambda e^{-\lambda x}dx=\frac{\lambda}{\lambda-it}.$$

## ★补充：特征函数

**Theorem**

若随机变量$X$的特征函数为$h_X(t)$，则$Y=a+bX$ $(b\neq 0)$的特征函数为

$$h_Y(t)=e^{iat}h_X(bt).$$

**Example**

标准正态分布$N(0,1)$的特征函数为$e^{-\frac{t^2}{2}}$，从而一般正态分布$N(\mu,\sigma^2)$的特征函数为

$$\exp\{i\mu t-\frac{1}{2}\sigma^2t^2\}.$$

## ★补充：特征函数

**Theorem**

设随机变量$X$与$Y$相互独立，则它们和的特征函数等于各自特征函数的乘积，即有

$$h_{X+Y}(t) = h_X(t) \cdot h_Y(t).$$

**Theorem (唯一性定理)**

分布函数由其特征函数唯一确定。

# ★补充：多元特征函数

**Definition (特征函数(Characteristic function))**

设$F(\vec{x})$为一个$n$元分布函数，称

$$h(\vec{t}) = \int_{\mathbb{R}^n} e^{i\vec{t}^{\,T}\vec{x}} dF(\vec{x}),$$

为$F(\vec{x})$的特征函数。如果$F(\vec{x})$为随机变量$\vec{X}$的分布函数，则$h(\vec{t})$也称为$\vec{X}$的特征函数，此时有

$$h(\vec{t}) = E(e^{i\vec{t}^{\,T}\vec{X}}).$$

# 补充：特征函数

**Theorem**

$n$元正态分布$N(\vec{\mu}, B)$的特征函数为

$$h(\vec{t}) = \exp\left\{ i\vec{\mu}^T\vec{t} - \frac{1}{2}\vec{t}^{\,T}B\vec{t} \right\}.$$

**Theorem (正态随机向量的线性变换)**

设$n$维随机向量$\vec{X} \sim N(\vec{\mu}, B)$。则对任意的满秩矩阵$C \in \mathbb{R}^{m\times n}$ $(m \leq n)$，$m$维随机向量$\vec{Y} = C\vec{X}$服从正态分布

$$N(C\vec{\mu}, CBC^T).$$

# 正态随机变量的线性组合

**Corollary**

设$n$维随机向量$\vec{X} \sim N(\vec{\mu}, B)$。则对任意非零向量$\vec{a} \in \mathbb{R}^n$，有

$$\vec{a}^T \vec{X} \sim N(\vec{a}^T \vec{\mu}, \vec{a}^T B \vec{a}).$$

# 目录

# 中心极限定理

中心极限定理(Central limit theorem, CLT)是描述以下思想的定理的统称：如果一个随机现象由众多的随机因素所引起，且每一因素在总的变化里所起的作用不显著，则描述这个随机现象的随机变量近似地服从正态分布。

## 中心极限定理

符号说明：对独立随机变量序列$\{X_n\}_{n\geq 1}$，记$Z_n$为序列前$n$项和$\sum_{i=1}^n X_i$的标准化的随机变量，即

$$Z_n := \frac{\sum_{k=1}^n [X_k - E(X_k)]}{\sqrt{D\left(\sum_{k=1}^n X_k\right)}}, \qquad \forall n \geq 1.$$

另外，记

$$s_n^2 := D\left(\sum_{k=1}^n X_k\right) = \sum_{k=1}^n DX_k.$$

## 中心极限定理

**Theorem (林德伯格定理(Lindeberg's theorem))**

设独立随机变量序列$\{X_n\}_{n\geq 1}$满足林德伯格条件：对任意$\varepsilon > 0$，有

$$\lim_{n\to\infty}\frac{1}{s_n^2}\sum_{k=1}^{n}E\big[(X_k-EX_k)^2\cdot\mathbf{1}_{\{|X_k-EX_k|>\varepsilon s_n\}}\big]=0,$$

则当$n\to\infty$时，$Z_n$的分布函数逐点收敛于标准正态分布函数，即对任意$x\in\mathbb{R}$,

$$\lim_{n\to\infty}P\{Z_n\leq x\}=\Phi(x)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-t^2/2}dt.$$

# 中心极限定理

说明：若林德伯格条件成立，则对任意$\varepsilon > 0$，有

$$\lim_{n\to\infty} P\left\{\max_{1\leq k\leq n} \frac{|X_k - EX_k|}{s_n} > \varepsilon\right\} = 0.$$

上式说明$\{X_n\}_{n\geq 1}$的各项“依概率均匀地小”。

## 中心极限定理

**Theorem (列维定理(Lindeberg – Lévy CLT))**

设独立随机变量序列$\{X_n\}_{n\geq 1}$服从相同的分布，有共同的期望$\mu$及方差$\sigma^2 > 0$，则当$n \to \infty$时，$Z_n$的分布函数逐点收敛于标准正态分布函数，即对任意$x \in \mathbb{R}$，

$$\lim_{n\to\infty} P\left\{\frac{\sum_{k=1}^{n} X_k - n\mu}{\sqrt{n}\sigma} \leq x\right\} = \Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt.$$

## ★中心极限定理的证明思路

各种中心极限定理的证明均利用了以下等价关系：

**Lemma**

$Z_n$的分布函数逐点收敛于标准正态分布函数当且仅当$Z_n$的特征函数逐点收敛于标准正态分布的特征函数，即对任意$t \in \mathbb{R}$,

$$\lim_{n\to\infty} E(e^{itZ_n}) = e^{-t^2/2}.$$

# 中心极限定理

常用结论：大量的同分布随机变量的和、平均值近似地服从正态分布。

例1 任取一个实数，对其小数点后第一位四舍五入，将该数变为整数。此时舍入误差服从区间$[-0.5, 0.5)$上的均匀分布。若对300个实数进行这样的操作，求所有舍入误差的总和的绝对值小于10的概率。

## 中心极限定理

练习：一个螺丝钉的重量是一个随机变量，期望值是10克，标准差是1克，求一盒（100个）同型号螺丝钉的重量超过1.02千克的概率。

# 中心极限定理

**Theorem (De Moivre-Laplace定理)**

设在独立试验序列中，事件$A$在每次试验中发生的概率为$p$ $(0<p<1)$，记$Y_n$为前$n$次试验中事件$A$发生的总次数，则对任意$x\in\mathbb{R}$，有

$$\lim_{n\to\infty} P\left\{\frac{Y_n-np}{\sqrt{npq}}<x\right\}=\Phi(x)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-t^2/2}dt,$$

其中$p+q=1$。

## 中心极限定理

De Moivre-Laplace定理的意义：若$X \sim B(n,p)$，则当$n$充分大时，$X$近似服从正态分布，即可以近似认为

$$X \sim N(np, npq),$$

其中$p+q=1$。

# 中心极限定理

例2 某工厂有200台同类型的机器，每台机器工作时需要的电功率为$Q$千瓦。由于工艺等原因，每台机器的实际工作时间只占全部工作时间的75%，各台机器是否工作是相互独立的。求：

1. 任一时刻有144-160台机器正在工作的概率；
2. 需要供应多少电功率可以保证所有机器正常工作的概率不小于0.99？

# 中心极限定理

练习：某公司有200名员工参加一种资格证书考试。按往年经验，该考试的通过率为0.8。试计算这200名员工至少有150人考试通过的概率。