概率论与数理统计

附录 · 常见的重要分布

2020 年 2 月 12 日

■ 暨南大学数学系 ■ 吕荐瑞

第一节 离散型分布

第二节 连续型分布

第一节	离散型分布
A	两点分布
B	二项分布
C	泊松分布

两点分布

若随机变量 X 只能取 0 或 1，其概率分布为：

$$P\{X=1\}=p,\ P\{X=0\}=1-p \quad (0<p<1)$$

则称 X 服从参数为 p 的两点分布或 0-1 分布，记为

$$X \sim B(1,p).$$

对应分布表为

X	0	1
P	$1-p$	p

两点分布

两点分布的数字特征：若

$$X \sim B(1, p),$$

则

$$EX = p, \qquad \mathrm{Var}(X) = p(1-p).$$

第一节	离散型分布
A	两点分布
B	二项分布
C	泊松分布

二项分布

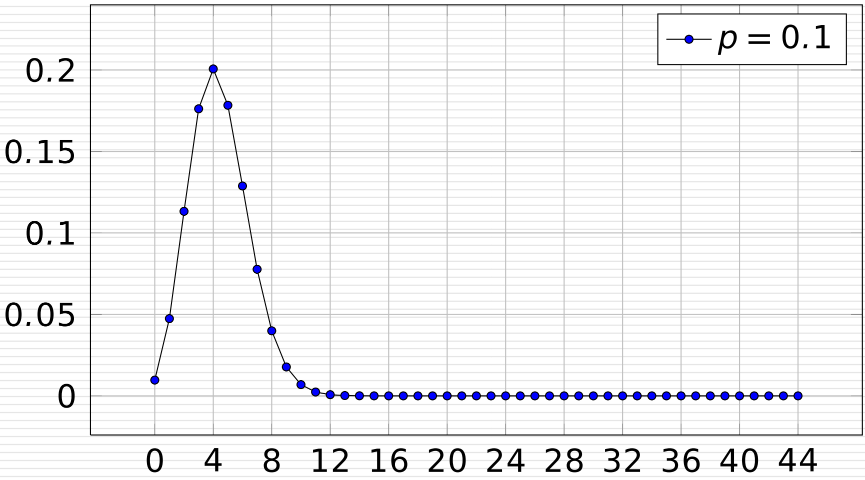
如果随机变量 X 服从以下分布律

$$P\{X=k\}=b(k;n,p)=C_n^k p^k(1-p)^{n-k},$$

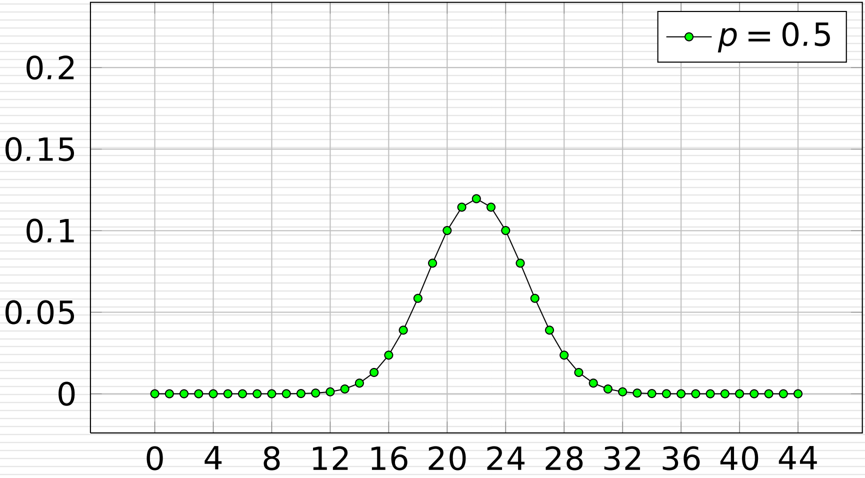
其中 $p\in(0,1)$，$0\leqslant k\leqslant n$，则称 X 服从参数为 n,p 的二项分布．简记为

$$X\sim B(n,p).$$

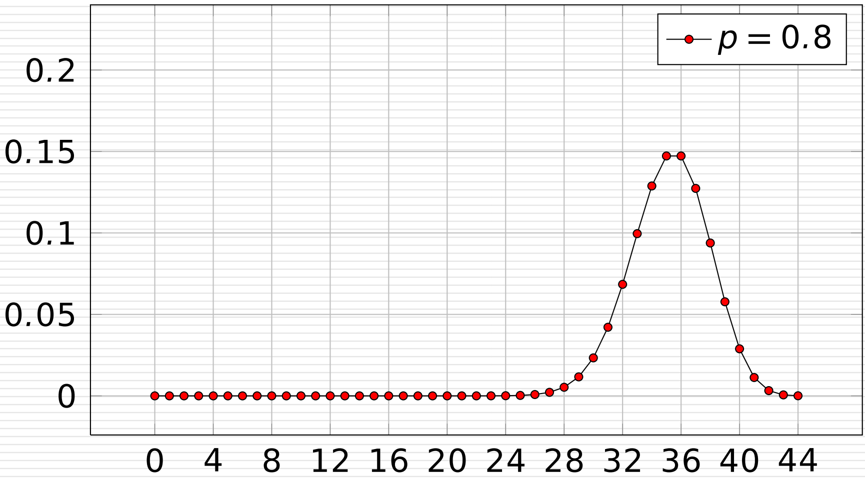
二项分布：$n = 44$



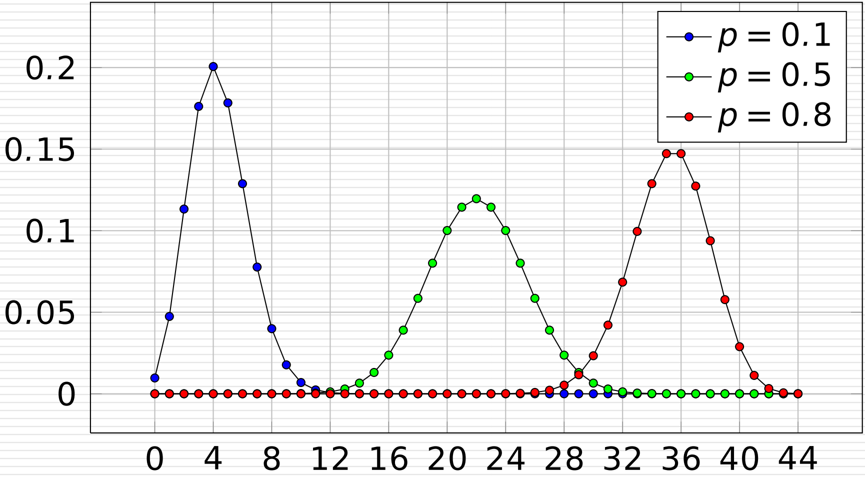
二项分布：$n = 44$



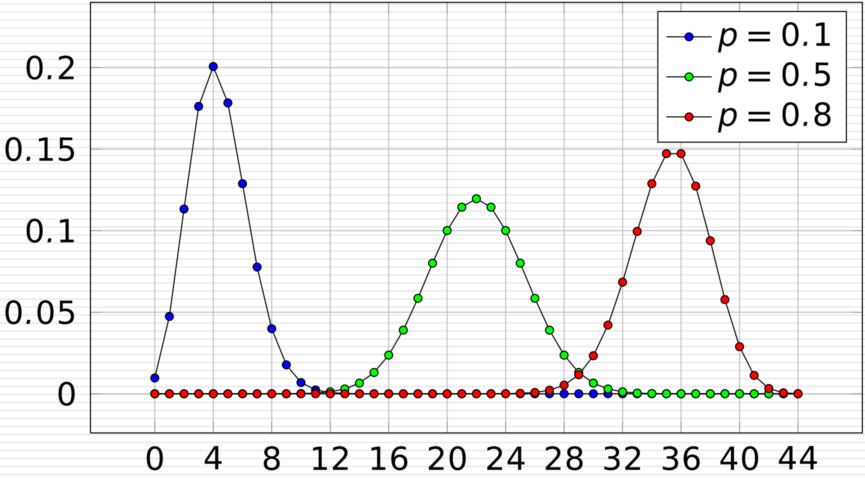
二项分布：$n = 44$



二项分布：$n = 44$



二项分布：$n = 44$



注记　当 $k = [np + p]$ 时概率最大.

二项分布

设在某试验中事件 A 的概率为 p，将该试验独立地进行 n 次．记 X 为 n 次试验中事件 A 发生的总次数，X_i 为第 i 次试验中事件 A 发生的次数，则

$$X \sim B(n,p),\ \ X_i \sim B(1,p), i = 1,2,\cdots,n$$

$X_1,X_2,\cdots,X_n$ 相互独立，且

$$X = X_1 + X_2 + \cdots + X_n$$

二项分布

二项分布的数字特征：若

$$X \sim B(n,p),$$

则

$$EX = np, \qquad \mathrm{Var}(X) = np(1-p).$$

第一节	离散型分布
A	两点分布
B	二项分布
C	泊松分布

泊松分布

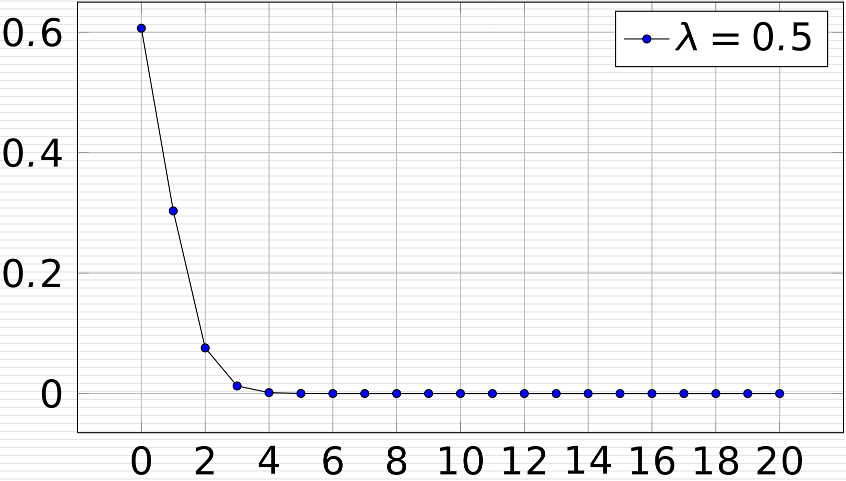
如果随机变量 X 服从以下分布律

$$P\{X=m\}=\frac{\lambda^m}{m!}e^{-\lambda},\quad m=0,1,\cdots$$

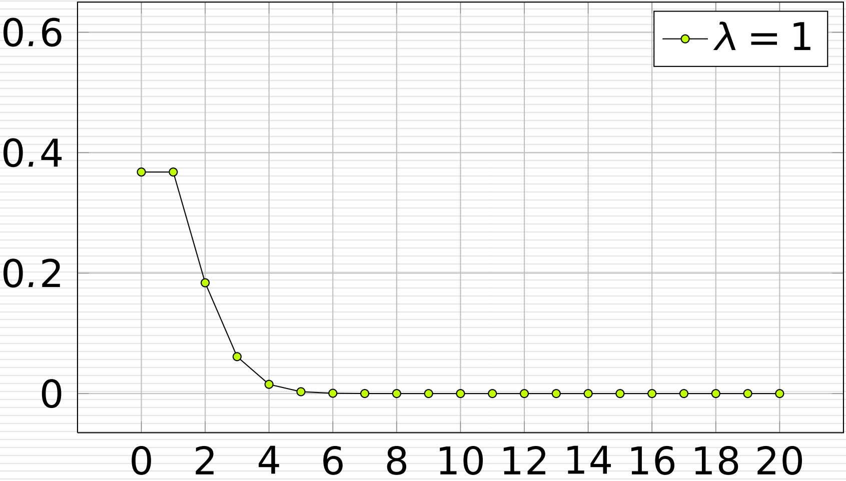
其中 $\lambda>0$，则称 X 服从参数为 λ 的泊松分布，简记为

$$X\sim P(\lambda).$$

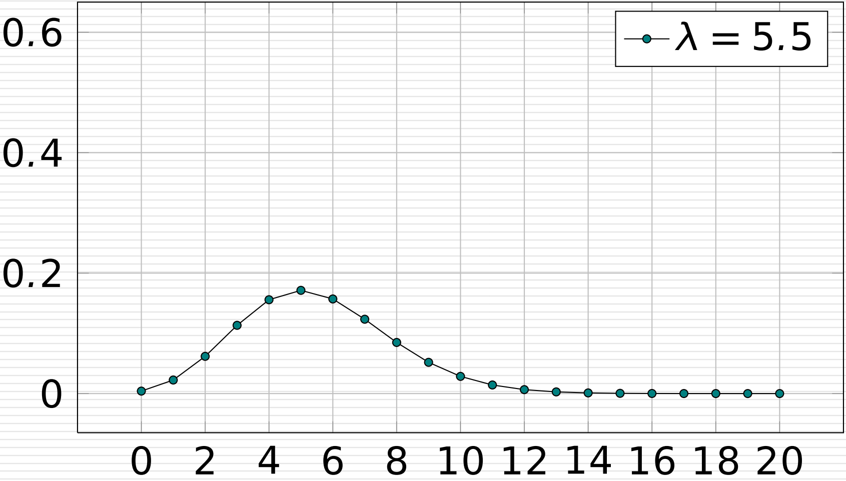
泊松分布



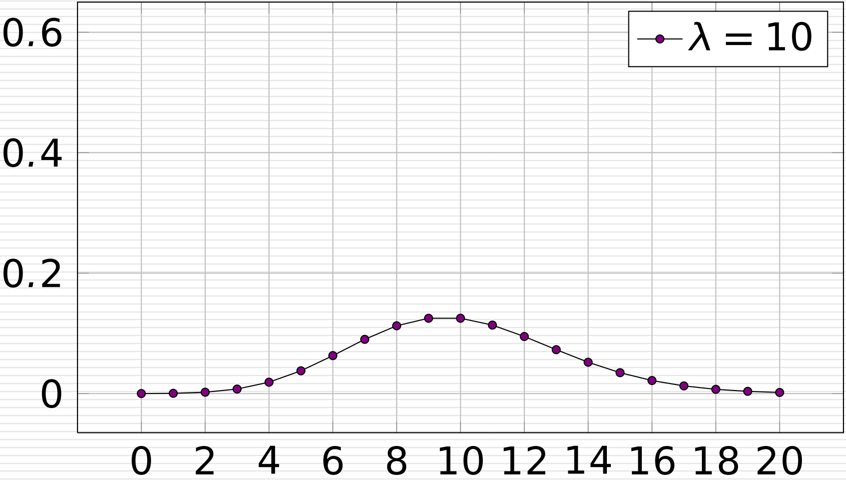
泊松分布



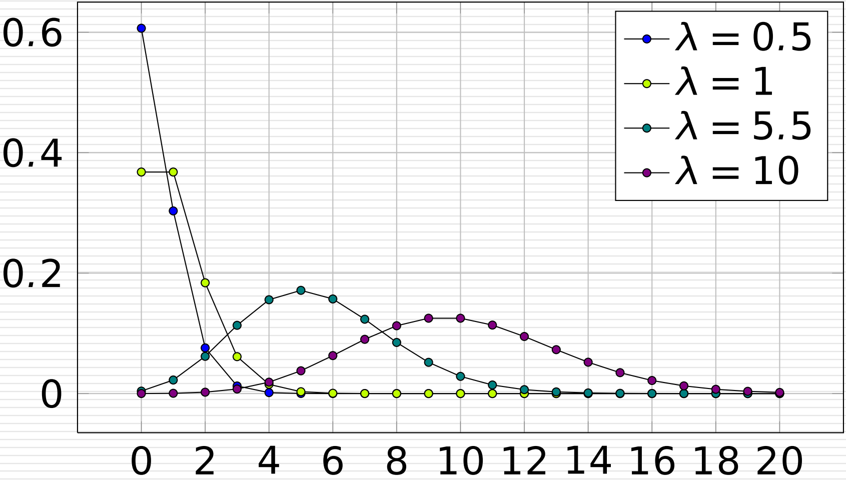
泊松分布



泊松分布

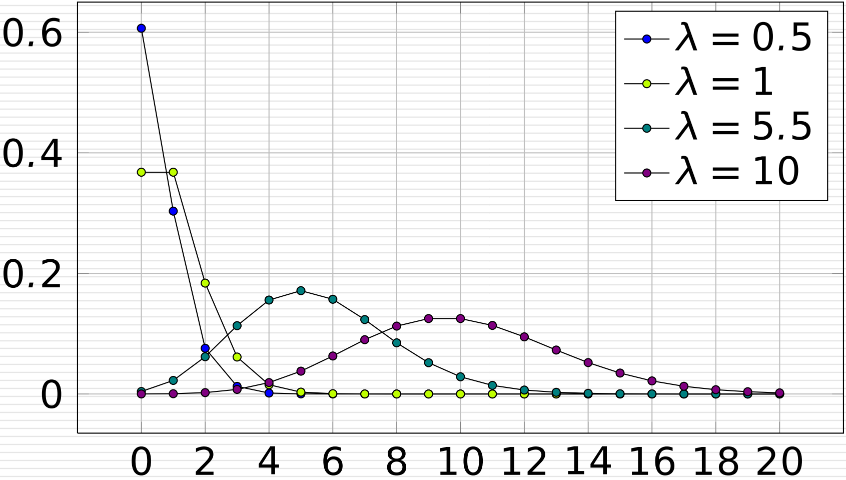


泊松分布



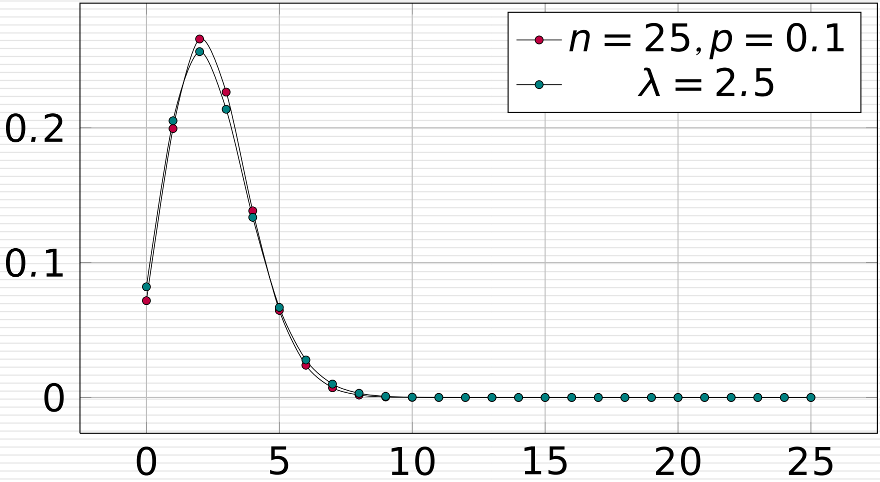
0.6
0.4
0.2
0
0
2
4
6
8
10
12
14
16
18
20
$\lambda = 0.5$
$\lambda = 1$
$\lambda = 5.5$
$\lambda = 10$

泊松分布

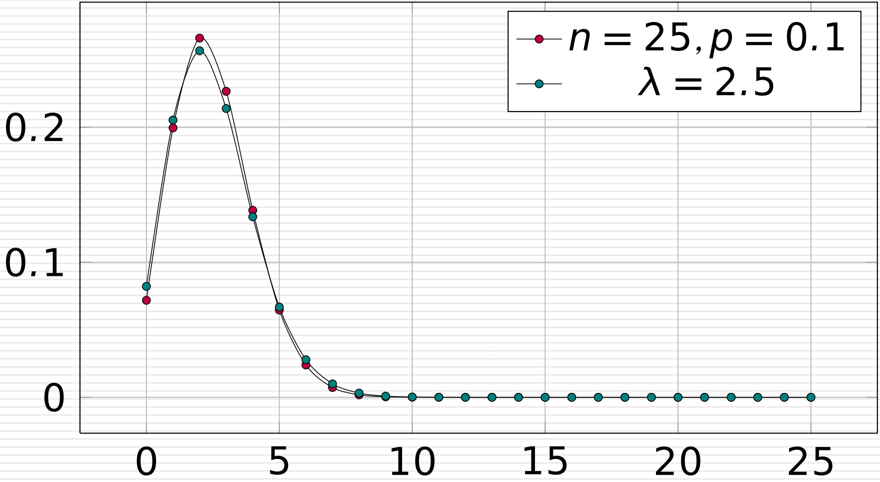


注记　当 $k = [\lambda]$ 时概率最大.

二项分布与泊松分布



二项分布与泊松分布



注记 $B(n,p)$ 与 $P(\lambda)$ 近似（n 大、p 小，$\lambda = np$）.

泊松分布

泊松分布的数字特征：如果

$$X \sim P(\lambda),$$

则

$$EX = \lambda, \qquad \mathrm{Var}(X) = \lambda.$$

第一节 离散型分布

第二节 连续型分布

第二节	连续型分布
A	均匀分布
B	指数分布
C	正态分布

均匀分布

若随机变量 X 有概率密度

$$f(x)=\begin{cases}\frac{1}{b-a}, & x\in[a,b]\\ 0, & \text{otherwise}\end{cases}, \qquad (a<b)$$

则称 X 服从区间 $[a,b]$ 上的均匀分布，记为

$$X\sim U[a,b].$$

均匀分布

上述均匀分布的分布函数为

$$F(x)=\begin{cases} 0, & x\leqslant a \\ \frac{x-a}{b-a}, & x\in[a,b] \\ 1, & x\geqslant b \end{cases}.$$

若 $X\sim U[a,b]$，则对 $[c,d]\subset[a,b]$，有

$$P\{c\leqslant X\leqslant d\}=\frac{d-c}{b-a}.$$

均匀分布

均匀分布的数字特征：如果

$$X \sim U[a,b],$$

则

$$EX = \frac{a+b}{2}, \qquad \mathrm{Var}(X) = \frac{(b-a)^2}{12}.$$

二维均匀分布

设 D 是平面上的有界区域，其面积为 d，若二维随机向量 (X, Y) 的概率密度为

$$f(x,y) = \begin{cases} \frac{1}{d} & (x,y) \in D \\ 0 & \text{else} \end{cases}$$

则称 (X, Y) 服从 D 上的均匀分布.

二维均匀分布

若 (X,Y) 服从 D 上的均匀分布，则 (X,Y) 落在某一区域 A 内的概率

$$
\begin{aligned}
P\{(X,Y)\in A\} &= \iint_A f(x,y)\,\mathrm{d}x\,\mathrm{d}y \\
&= \iint_{A\cap D} \frac{1}{d}\,\mathrm{d}x\,\mathrm{d}y \\
&= \frac{S}{d}
\end{aligned}
$$

其中 S 为 $A\cap D$ 的面积.

第二节	连续型分布
A	均匀分布
B	指数分布
C	正态分布

指数分布

如果随机变量 X 有以下概率密度

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geqslant 0 \\ 0, & \text{otherwise} \end{cases},$$

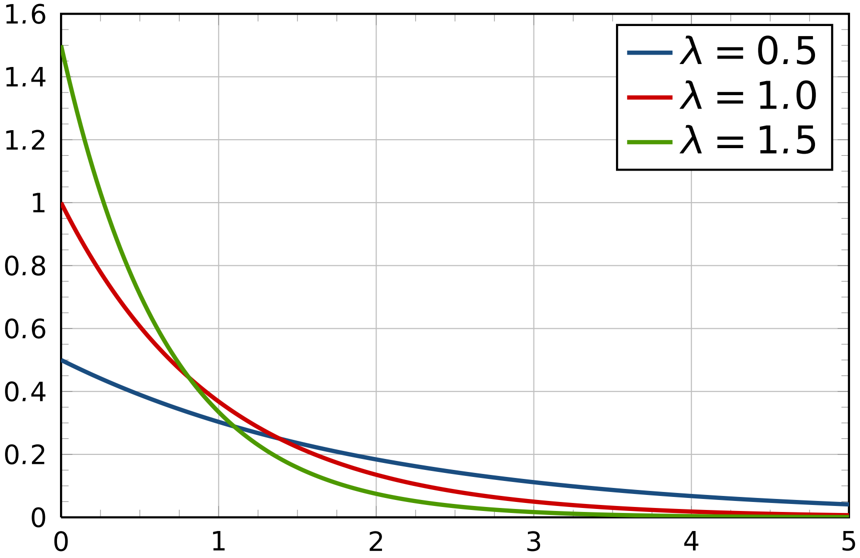
其中 $\lambda > 0$，则称 X 服从参数为 λ 的指数分布，记为

$$X \sim EP(\lambda).$$

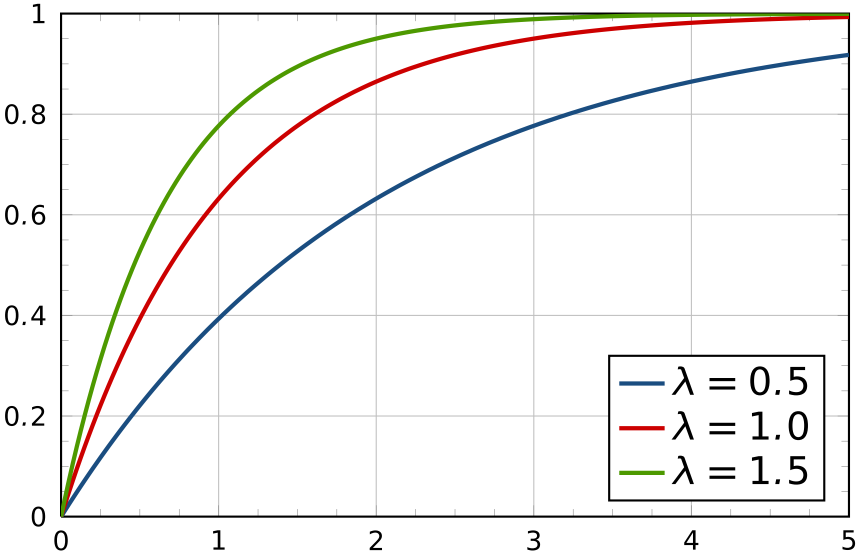
其分布函数为

$$F(x) = \begin{cases} 1 - e^{-\lambda x}, & x \geqslant 0 \\ 0, & x < 0 \end{cases}.$$

指数分布的密度函数



指数分布的分布函数



指数分布

指数分布的数字特征：设随机变量 X 服从参数为 λ 的指数分布，则

$$EX = \frac{1}{\lambda}, \qquad \mathrm{Var}X = \frac{1}{\lambda^2}.$$

第二节	连续型分布
A	均匀分布
B	指数分布
C	正态分布

正态分布

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 μ, σ 为常数且 $\sigma > 0$，则称 X 服从正态分布．简记为

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

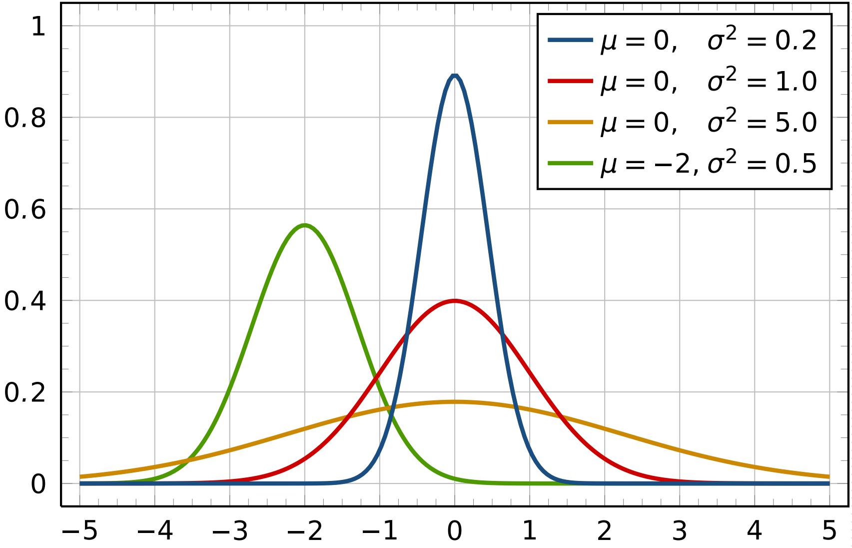
称 $N(0,1)$ 为标准正态分布，并简写 $\varphi_{0,1}(x)$ 为 $\varphi(x)$.

正态分布的数字特征

设随机变量 $X \sim N(\mu, \sigma^2)$，则

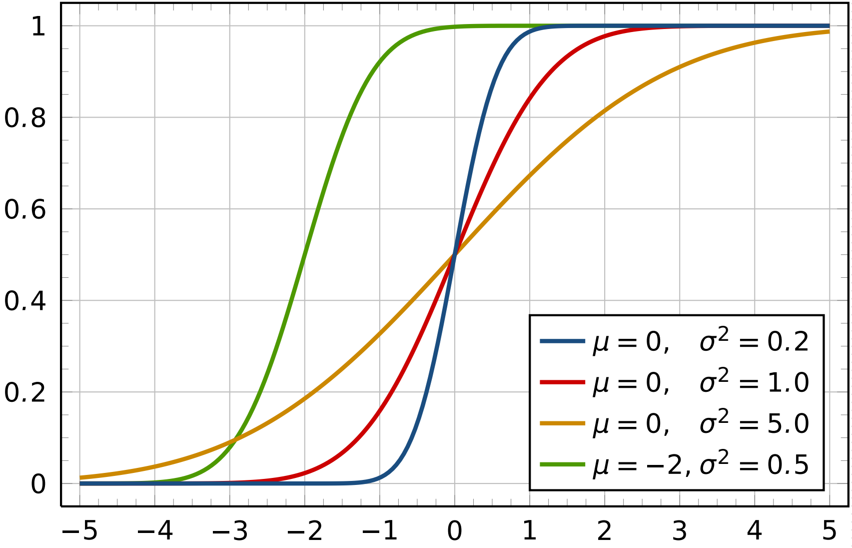
$$EX = \mu, \qquad \mathrm{Var}(X) = \sigma^2.$$

正态分布的密度函数



$\mu = 0, \quad \sigma^2 = 0.2$
$\mu = 0, \quad \sigma^2 = 1.0$
$\mu = 0, \quad \sigma^2 = 5.0$
$\mu = -2, \sigma^2 = 0.5$
1
0.8
0.6
0.4
0.2
0
−5
−4
−3
−2
−1
0
1
2
3
4
5

正态分布的分布函数



1
0.8
0.6
0.4
0.2
0
−5
−4
−3
−2
−1
0
1
2
3
4
5
$\mu = 0, \quad \sigma^2 = 0.2$
$\mu = 0, \quad \sigma^2 = 1.0$
$\mu = 0, \quad \sigma^2 = 5.0$
$\mu = -2, \sigma^2 = 0.5$

正态分布

正态分布的分布函数为

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

该函数不是初等函数.

正态分布

正态分布的分布函数为

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

该函数不是初等函数.

标准正态分布的分布函数简记为 $\Phi(x)$.

正态分布

标准正态分布的分布函数的性质

1. $\Phi(-x) = 1 - \Phi(x)$.
2. 若 $X \sim N(0,1)$，则对任意 $x \geqslant 0$，有
$$P\{|X| \leqslant x\} = 2\Phi(x) - 1.$$

正态分布

若 X 服从正态分布，则其线性函数 $aX+b(a\neq 0)$ 仍然服从正态分布．且若 $X\sim N(\mu,\sigma^2)$，则有

$$aX+b\sim N(a\mu+b,a^2\sigma^2).$$

正态分布

若 X 服从正态分布，则其线性函数 $aX+b(a\neq 0)$ 仍然服从正态分布．且若 $X\sim N(\mu,\sigma^2)$，则有

$$aX+b\sim N(a\mu+b,a^2\sigma^2).$$

服从正态分布随机变量的标准化：若 $X\sim N(\mu,\sigma^2)$，则有

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

正态分布

定理 若 $X_1, X_2, \cdots, X_n$ 相互独立，且

$$X_i \sim N(\mu_i, \sigma_i^2),\ i = 1, 2, \cdots, n$$

则对不全为零的数 $a_1, a_2, \cdots, a_n$，

$$a_1 X_1 + \cdots + a_n X_n \sim N\left(\sum_{i=1}^{n} a_i \mu_i, \sum_{i=1}^{n} a_i^2 \sigma_i^2\right).$$

常用连续型随机变量

设随机变量 $X \sim N(0,1)$，对给定的 $\alpha \in (0,1)$，称满足条件

$$P\{X > Z_\alpha\} = \alpha$$

的点 Z_α 为标准正态分布的上 α 分位点.

常用连续型随机变量

设随机变量 $X \sim N(0,1)$，对给定的 $\alpha \in (0,1)$，称满足条件

$$P\{X > Z_\alpha\} = \alpha$$

的点 Z_α 为标准正态分布的上 α 分位点.

由性质 $\Phi(-x) = 1 - \Phi(x)$ 可得

$$Z_{1-\alpha} = -Z_\alpha.$$

二元正态分布

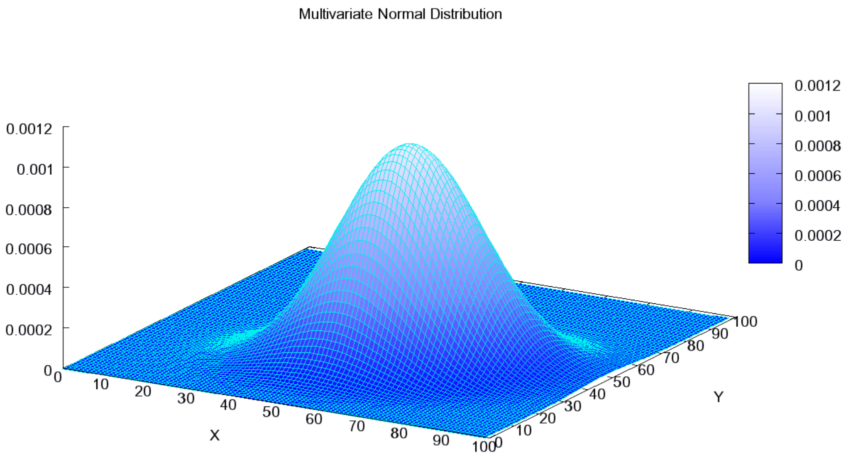
定义　以下面函数为密度的分布称为**二元正态分布**，简记为 $N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$：

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}$$

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

其中 μ_1,μ_2 为实数，$\sigma_1,\sigma_2>0$，$|\rho|<1$.

二元正态分布


Multivariate Normal Distribution
0.0012
0.001
0.0008
0.0006
0.0004
0.0002
0
X
10
20
30
40
50
60
70
80
90
100
Y
0
10
20
30
40
50
60
70
80
90
100
0.0012
0.001
0.0008
0.0006
0.0004
0.0002
0

二元正态分布

二元正态分布的边缘分布：若

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

则

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

二元正态分布

二元正态分布的独立性：若

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

则 X 与 Y 相互独立的充要条件为

$$\text{相关系数 } \rho = 0.$$