中国科学院数学与系统科学研究院
Academy of Mathematics and Systems Science
Chinese Academy of Sciences

第四章 重要概率分布

洪永淼
中国科学院数学与系统科学研究院
中国科学院大学经济与管理学院

目　　录

第一节 引言

第二节 离散概率分布

第三节 连续概率分布

第四节 小结

引言

- 概率空间三元组合 $(S, \mathbb{B}, P)$ 完整刻画了一个随机试验。在实际应用中，往往无法知晓某随机试验的真实概率分布。
- 统计学家与计量经济学家通常从某一类概率测度如 PMF/PDF $f(x, \theta)$ 中选择真实的概率分布，其中函数形式 $f(\cdot,\cdot)$ 已知，而参数值 θ 未知。通常假设存在某个唯一参数值 θ_0 使得 $f(x, \theta_0)$ 与随机试验的真实概率分布 $f_X(x)$ 一致。

引言

- 统计学与计量经济学的主要目的之一就是利用经济观测数据估计这一真实参数值。
- 在估计真实参数值 θ_0 之前，首先介绍几类重要的参数概率分布模型，并讨论其在经济学、金融学与管理学领域的应用。
- 特别地，将重点强调各类参数概率分布模型中参数的含义及其作用。

目　录

第一节 介绍

第二节 离散概率分布

第三节 连续概率分布

第四节 小结

4.2.1 伯努利分布 (Bernoulli Distribution)

- 若随机变量 X 的 PMF 为

$$f_X(x) = \begin{cases} p, & x = 1 \\ 1-p, & x = 0 \end{cases}$$
$$= p^x(1-p)^{1-x}, \qquad x = 0,1$$

其中 $0 < p < 1$，则称其服从**伯努利分布** Bernoulli(p)。

- 伯努利随机变量为**二元变量**，取 1 的概率为 p，取 0 的概率为 $1-p$。

伯努利分布的均值、方差与 MGF

- 对 Bernoulli(p) 随机变量 X，有

  ✓ 均值：$E(X) = p$

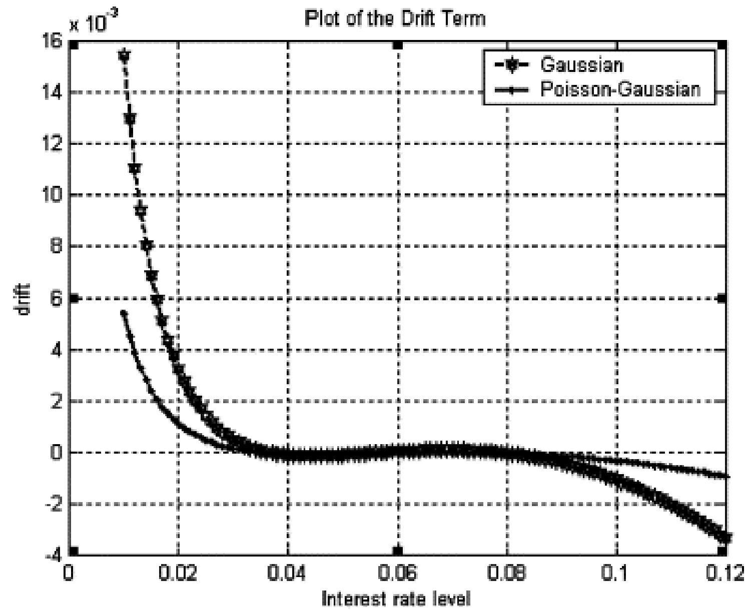
  ✓ 方差：$\mathrm{var}(X) = p(1-p)$

  ✓ MGF：$M_X(t) = pe^t + 1 - p,\ -\infty < t < \infty$

- 注意 X 的各阶矩均为 p 的函数，而 p 是伯努利分布唯一的参数。
- 由于 $E(X) = P(X = 1)$ 且 X 仅有两个可能取值，均值参数 $E(X) = p$ 完全刻画伯努利随机变量的概率分布。

伯努利分布的应用

- 该分布在经济学与金融学有广泛的应用。
  - ✓ 定义一个随机变量 X，若 IBM 股票价格上涨，其取值为 1；若 IBM 股价下跌则取值为 0。则随机变量 X 作为 IBM 股价变化方向的指示变量，服从伯努利分布。
  - ✓ Das (2002) 在研究美国联邦基准利率跳跃行为时，采用伯努利分布近似利率跳跃所服从的概率分布。



Source: Das (2002)

4.2.2 二项分布 (Binomial Distribution)

- 若离散随机变量 X 的 PMF 为

$$f_X(x) = \binom{n}{x} p^x (1-p)^{n-x}, \qquad x = 0,1,\cdots,n$$

其中 $n \geq 1$ 和 $0 < p < 1$，则称其服从**二项分布**，记作 $B(n,p)$。

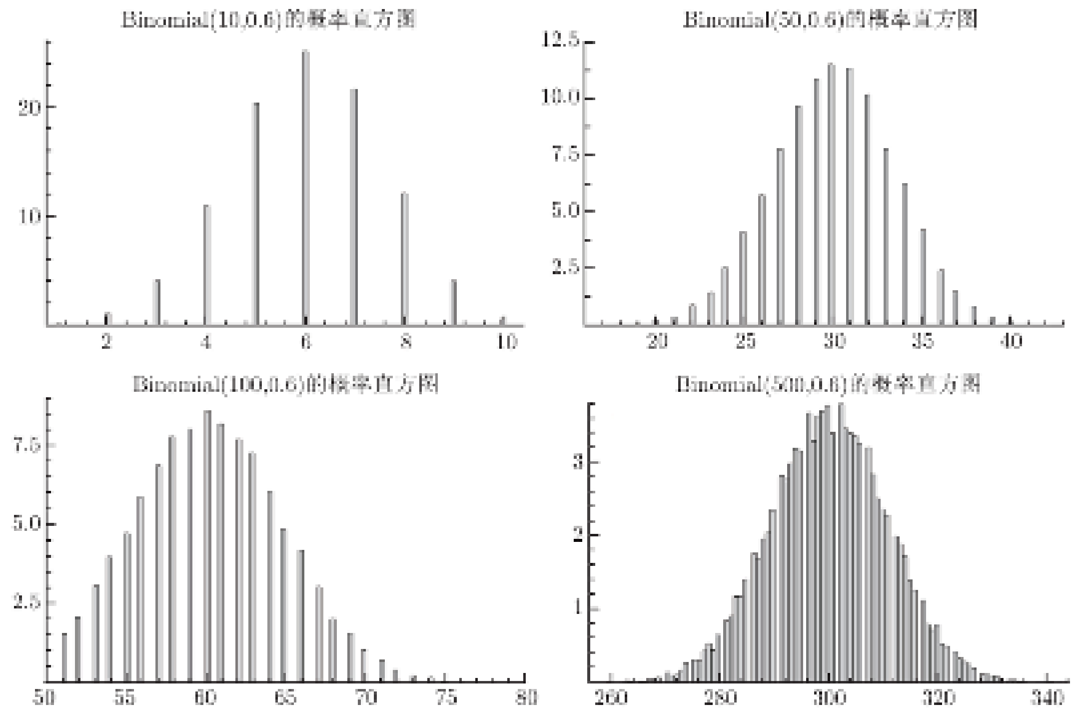


图 4.1：n 取不同值时 $B(n,p)$ 的概率直方图

何时会出现二项分布呢?

- 若独立抛掷 n 次硬币，每次出现正面朝上的概率为 p。假设共进行 n 次试验，则正面朝上的次数是多少?
  - 令 X_i 表示第 i 次试验的结果，正面朝上时 X_i 取值为 1，否则取值为 0；令 X 表示 n 次试验正面朝上的总次数，则

$$X = \sum_{i=1}^{n} X_i$$

其中 X_i 为服从 Bernoulli(p) 的独立同分布 (IID) 随机变量。

可以证明 $X = \sum_{i=1}^{n} X_i$ 服从 $B(n, p)$ 分布。

◆ 问题 4.1

如何证明对所有 $n \geq 1$ 和所有 $p \in (0,1)$，有

$$\sum_{x=0}^{n} f_X(x) = \sum_{x=0}^{n} \binom{n}{x} p^x (1-p)^{n-x} = 1?$$

- 由二项式定理，对任意实数 x 和 y 以及整数 $n \geq 0$，有

$$(x+y)^n = \sum_{i=0}^{n} \binom{n}{i} x^i y^{n-i}$$

令 $x = p$ 和 $y = 1-p$ 即得上式。该二项展开式是该分布称为二项分布的原因。

二项分布的均值、方差与 MGF

- $B(n,p)$ 的**均值:**

$$
\begin{aligned}
E(X) &= \sum_{x=0}^{n} x f_X(x) = \sum_{x=0}^{n} x\binom{n}{x}p^x(1-p)^{n-x} \\
&= \sum_{x=1}^{n} x\binom{n}{x}p^x(1-p)^{n-x} + 0\cdot\binom{n}{0}p^0(1-p)^{n-0} \\
&= \sum_{x=1}^{n} n\binom{n-1}{x-1}p^x(1-p)^{n-x} \quad (\text{令} y = x-1) \\
&= np\sum_{y=0}^{n-1} \binom{n-1}{y} p^y(1-p)^{(n-1)-y} \\
&= np\sum_{y=0}^{n-1} f_Y(y) \\
&= np
\end{aligned}
$$

其中，最后一个求和运算可视为对一个服从二项分布 $B(n-1,p)$ 的随机变量 Y 的 PMF $f_Y(\cdot)$ 求和，故有 $\sum_{y=0}^{n-1} f_Y(y)=1$。

二项分布的均值、方差与 MGF

- 计算 $B(n,p)$ 的方差。其中，**二阶矩**：

$$
\begin{aligned}
E\left(X^2\right) &= \textstyle\sum_{x=0}^{n} x^2 f_X(x) = \sum_{x=0}^{n} x^2 \binom{n}{x} p^x (1-p)^{n-x} \\
&= \textstyle\sum_{x=1}^{n} x^2 \binom{n}{x} p^x (1-p)^{n-x} \\
&= \textstyle\sum_{x=1}^{n} xn \binom{n-1}{x-1} p^x (1-p)^{n-x} \qquad (\text{令 } y = x-1) \\
&= n \textstyle\sum_{y=0}^{n-1} (y+1) \binom{n-1}{y} p^{y+1} (1-p)^{(n-1)-y} \\
&= np \textstyle\sum_{y=0}^{n-1} y \binom{n-1}{y} p^{y} (1-p)^{(n-1)-y} \\
&\quad + np \textstyle\sum_{y=0}^{n-1} \binom{n-1}{y} p^{y} (1-p)^{(n-1)-y} \\
&= npE(Y) + np \textstyle\sum_{y=0}^{n-1} f_Y(y) = np(n-1)p + np \\
&= np[(n-1)p + 1]
\end{aligned}
$$

二项分布的均值、方差与 MGF

- $B(n,p)$ 的**方差**

$$\begin{aligned}\sigma_X^2 &= E\left(X^2\right) - \mu_X^2 \\ &= \{np\cdot[(n-1)p] + np\} - (np)^2 \\ &= np(1-p)\end{aligned}$$

- $B(n,p)$ 的 **MGF**

$$\begin{aligned}M_X(t) &= E\left(e^{tX}\right) \\ &= \sum_{x=0}^{n} e^{tx}\binom{n}{x}p^x(1-p)^{n-x} \\ &= \sum_{x=0}^{n} \binom{n}{x}(pe^t)^x(1-p)^{n-x} \\ &= (pe^t+1-p)^n, \qquad -\infty < t < \infty\end{aligned}$$

伯努利分布与二项分布的关系

- 许多试验都可视为伯努利试验序列，而多次伯努利试验的结果的总和即服从二项分布。
  - ✓ 例如，二项分布可用于近似 n 个产品中次品数量的分布，其中每个产品为次品的概率均为 p。

4.2.3 负二项分布 (Negative Binomial Distribution)

- 二项分布 $B(n,p)$ 描述了在试验次数 n 固定的情况下，成功次数的概率分布。
- 为了获得某一给定成功次数所需试验总次数的概率分布。该分布称为负二项分布，记作 $NB(n,p)$。
  - ✓ 具体而言，在一系列独立 Bernoulli(p) 试验中，令随机变量 X 表示获得 r 次成功所需的试验次数，即在第 X 次试验时获得第 r 次成功，其中 r 为固定整数。换言之，$X-1$ 是获得第 r 次成功之前所进行的试验次数。

负二项分布的 PMF

- 由于前 $X-1$ 次试验中共有 $r-1$ 次成功，X 的 **PMF** 为

$$\begin{aligned} f_X(x) &= \left[\binom{x-1}{r-1}p^{r-1}(1-p)^{(x-1)-(r-1)}\right]p \\ &= \binom{x-1}{r-1}p^r(1-p)^{x-r},\ x=r,\ r+1,\cdots \end{aligned}$$

其中 $\binom{x-1}{r-1}p^{r-1}(1-p)^{(x-1)-(r-1)}$ 表示前 $x-1$ 次试验中共获得 $r-1$ 次成功的概率，而 p 则为第 x 次试验成功的概率。

负二项分布的应用

- 负二项分布在经济学有广泛应用。
  - ✓ 例如，若一个家庭希望生育一定数量的男孩或女孩，可用负二项分布描述该家庭的人数规模 (Rao et al., 1973)。

4.2.4 几何分布 (Geometric Distribution)

- 几何分布是进行一系列独立伯努利试验时获得第一次成功所需试验次数的概率分布，即负二项分布在 $r = 1$ 时的特殊情形。当 $r = 1$ 时，负二项分布变成

$$f_X(x) = p(1-p)^{x-1}, \qquad x = 1, 2, \cdots$$

- 随机变量 X 可解释为获得第一次成功所需的试验次数，即“等待第一次成功”所需的次数或时间。

几何分布的无记忆性

- 几何分布具有以下“无记忆（memoryless）”或“无时效(nonaging)”性质：对整数 $s > t$，有

$$P(X > s \mid X > t) = P(X > s - t)$$

- 在某一时期失败的概率仅依赖于这段时期的长度，与开始的时间点无关。

4.2.5 泊松分布 (Poisson Distribution)

- 若离散随机变量 X 的 **PMF** 为

$$f_X(x) = e^{-\lambda}\frac{\lambda^x}{x!}, x = 0,1,\cdots$$

其中 $\lambda > 0$，则称其服从泊松分布 Poisson(λ)。参数 λ 称为强度参数 (intensity parameter)。

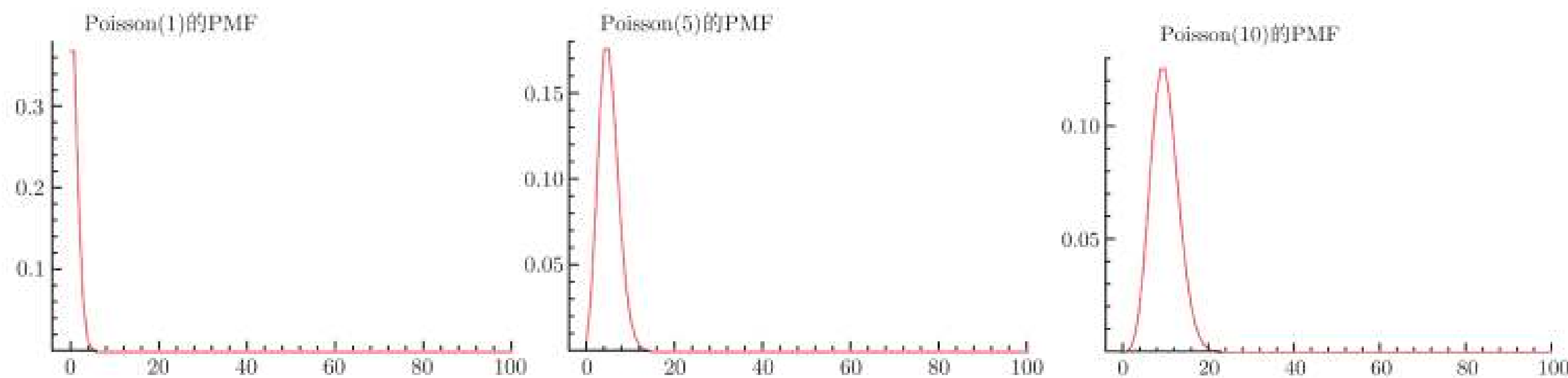


图 4.2 : 不同参数值 λ 对应的 PMF

泊松分布

对任意给定 $\lambda > 0$

$$\sum_{x=0}^{\infty} f_X(x) = 1$$

■ **证明：** 根据麦克劳林级数展开

$$e^{\lambda} = \sum_{x=0}^{\infty} \frac{\lambda^x}{x!}$$

有

$$\begin{aligned}\sum_{x=0}^{\infty} f_X(x) &= e^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} \\ &= e^{-\lambda} e^{\lambda} \\ &= 1\end{aligned}$$

泊松分布的均值

计算 $X \sim \text{Poisson}(\lambda)$ 的**均值**:

$$E(X) = \sum_{x=0}^{\infty} xe^{-\lambda}\frac{\lambda^x}{x!}$$

$$= \sum_{x=1}^{\infty} xe^{-\lambda}\frac{\lambda^x}{x!} \quad [\text{使用}x! = x(x-1)!]$$

$$= \lambda\sum_{x=1}^{\infty} e^{-\lambda}\frac{\lambda^{x-1}}{(x-1)!}$$

$$= \lambda\sum_{y=0}^{\infty} e^{-\lambda}\frac{\lambda^y}{y!} \quad (\text{令 } y = x-1)$$

$$= \lambda$$

泊松分布的方差

- Poisson(λ) 分布的**二阶矩**:

$$E(X^2) = \sum_{x=0}^{\infty} x^2 e^{-\lambda} \frac{\lambda^x}{x!}$$
$$= \sum_{x=1}^{\infty} x e^{-\lambda} \frac{\lambda^x}{(x-1)!}$$
$$= \lambda \sum_{x=1}^{\infty} x e^{-\lambda} \frac{\lambda^{x-1}}{(x-1)!} \quad (\text{令 } y = x-1)$$
$$= \lambda \sum_{y=0}^{\infty} (y+1) e^{-\lambda} \frac{\lambda^y}{y!}$$
$$= \lambda \sum_{y=0}^{\infty} y e^{-\lambda} \frac{\lambda^y}{y!} + \lambda \sum_{y=0}^{\infty} e^{-\lambda} \frac{\lambda^y}{y!}$$
$$= \lambda E(Y) + \lambda \sum_{y=0}^{\infty} f_Y(y)$$
$$= \lambda^2 + \lambda$$

注意：Poisson(λ) 分布的均值和方差均等于 λ。

- 因此，Poisson(λ) 分布的**方差**为

$$\sigma_X^2 = E(X^2) - \mu_X^2 = (\lambda^2 + \lambda) - \lambda^2 = \lambda$$

泊松分布的 MGF

$$
\begin{aligned}
M_X(t) &= E\left(e^{tX}\right) \\
&= \sum_{x=0}^{\infty} e^{tx} f_X(x) \\
&= \sum_{x=0}^{\infty} e^{tx} e^{-\lambda} \frac{\lambda^x}{x!} \\
&= e^{-\lambda} \sum_{x=0}^{\infty} \frac{\left(\lambda e^t\right)^x}{x!} \\
&= e^{-\lambda} e^{\lambda e^t} \quad [\text{用到 } e^a = \sum_{x=0}^{\infty} \frac{a^x}{x!}] \\
&= e^{\lambda\left(e^t - 1\right)}, -\infty < t < \infty
\end{aligned}
$$

小数定律 (Law of Small Numbers, LSN)

- 当$n \to \infty$ 但 $np \to \lambda$，可用 Poisson(λ) 分布近似**二项分布** $B(n,p)$，因为

$$
\begin{aligned}
M_B(t) &= (pe^t + 1 - p)^n \\
&= \left[1 + \frac{np(e^t-1)}{n}\right]^n \\
&\to e^{\lambda(e^t-1)} = M_P(t)
\end{aligned}
$$

- 上式推导中使用了当 $n \to \infty$ 时，$\left(1+\frac{a}{n}\right)^n \to e^a$ 这一极限性质。

Bortkiewicz (1898) 将这一性质称为小数定律。

小数定律 (Cont.)

- 概率论中，小数定律也称为**稀有事件定律** (the law of rare events) 或**泊松极限定理** (Poisson limit theorem)。

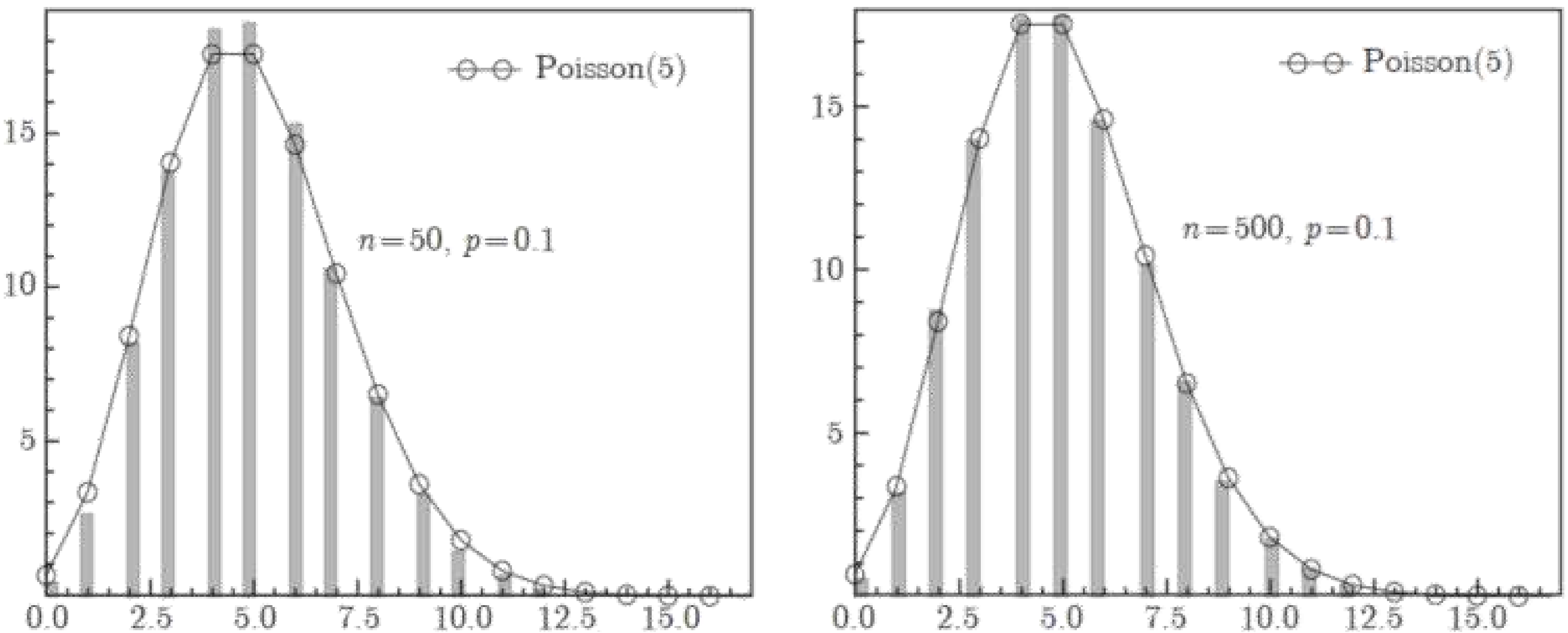


图 4.3：Poisson(λ) 分布对二项分布 $B(n,p)$ 的近似

泊松分布与负二项分布的联系

- 当 $r \to \infty$ 而 $r(1-p) \to \lambda$ 时，负二项分布 $NB(r,p)$ 同样收敛于 $\text{Poisson}(\lambda)$ 分布，因为有

$$f_Y(y) = \binom{y+r-1}{r-1} p^r (1-p)^y \to \frac{e^{-\lambda}\lambda^y}{y!}, \qquad y = 0,1,\cdots$$

- 因此，$NB(r,p)$ 同样可用 $\text{Poisson}(\lambda)$ 分布近似。这可用当 $r \to \infty$，$r(1-p) \to \lambda$ 时，$NB(n,p)$ 的 MGF 收敛于 $\text{Poisson}(\lambda)$ 的 MGF 的方法加以证明。

泊松分布的应用

- 泊松分布广泛应用于刻画某事件在特定时间段或者特定空间内发生的次数的分布

收银台的顾客人数

地震的次数

电话响起的次数

行业内违约或破产的次数

事故发生的次数

资产价格跳跃的次数

泊松分布的应用 (Cont.)

- Bortkiewicz（1898）考察了可能产生泊松分布的各种情形。特别地，他用泊松分布刻画普鲁士军队每年在行军中被马踢致死的士兵人数，其中每个士兵因该原因而死亡的概率很小但面对这种风险的士兵人数却很庞大。

泊松过程

- 泊松分布通常用于研究某个时期内或某个空间内发生的随机事件序列。假设从 $t = 0$ 时刻开始计算某事件发生的次数，则在每个时刻 t 都可定义一个整数，记作 $N(t)$。

$N(t)$ 是在时间段 $[0, t]$ 内该事件发生的次数。显然，对任意 t，$N(t)$ 是一个**离散随机变量**，其可能取值的集合为 $\{0,1,2,\cdots\}$。

泊松过程

假设 $N(0)=0$，时间 t 固定，计算事件在区间 $[0,t]$ 上发生的次数。

- 首先，将区间 $[0,t]$ 等分为 n 个子区间 t/n。$N(t)$ 是 n 个独立**伯努利 Bernoulli(p) 随机变量的总和**。
- 根据小数定律，二项分布收敛于泊松分布，满足

$$P[N(t)=m]=\frac{(\lambda t)^m e^{-\lambda t}}{m!}, \qquad m=0,1,2,\cdots$$

- 对任意 $t>0$，$N(t)$ 是一个参数为 λt 的**泊松随机变量**，故有 $E[N(t)]=\lambda t$，以及 $\lambda=E[N(1)]$。
- 因此，参数 λ 表示事件在单位时间内发生的平均次数。具有这一性质的随机过程称为发生率为 λ 的**泊松过程**，记 $\{N(t),t>0\}$。

泊松分布的应用

- 在金融学，泊松分布广泛应用于刻画资产价格在一定时期内**跳跃次数**的概率分布。
  - ✓ 例如，Merton（1976）假设资产价格除了包含布朗运动成分之外还包含泊松跳跃成分。这种用于金融衍生品定价的所谓跳跃扩散模型是一种复合模型，包含泊松跳跃过程和扩散过程。

  - ✓ 此外，泊松回归，即关于某一事件发生的次数和一组经济解释变量 (如 Z) 之间关系（假设 $\lambda = \alpha + \beta Z$）的分析。

目　录

第一节 介绍

第二节 离散概率分布

第三节 连续概率分布

第四节 小结

4.3.1 均匀分布 (Uniform Distribution)

- 若连续随机变量 X 的 PDF 为

$$f_X(x)=\begin{cases}\frac{1}{b-a}, & a\le x\le b\\ 0, & \text{其他}\end{cases}$$

则称其在区间 $[a,b]$ 上服从均匀分布，记作 $X\sim U[a,b]$。

- 当 $a=0$，$b=1$ 时，$U[0,1]$ 称为**标准均匀分布** (standard uniform distribution)。

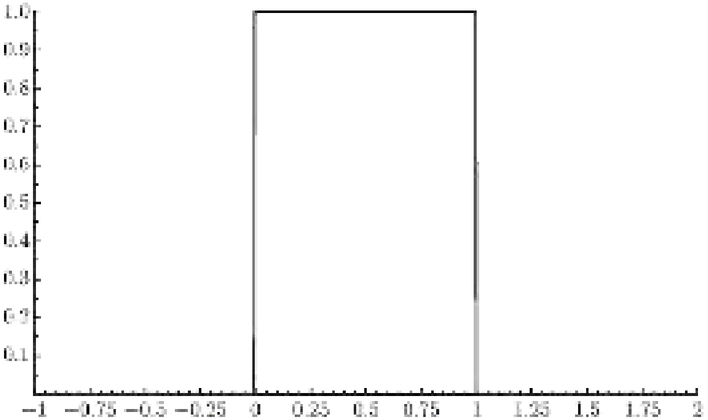


图 4.4：$U[0,1]$ 分布的 PDF

均匀分布的第 k 阶矩

- 因为 X 是**有界**随机变量，其各阶矩均存在。第 k 阶矩为

$$E\left(X^k\right)=\int_{-\infty}^{\infty}x^k f_X(x)dx=\frac{1}{b-a}\int_a^b x^k dx$$

$$=\frac{1}{b-a}\frac{x^{k+1}}{k+1}\bigg|_a^b=\frac{1}{b-a}\frac{b^{k+1}-a^{k+1}}{k+1}$$

> 令 $k=1$，可得 X 的**均值**： $\mu_X=\frac{1}{2}(a+b)$
>
> 令 $k=2$，可得**二阶矩**： $E\left(X^2\right)=\frac{1}{3}\left(b^2+a^2+ab\right)$
>
> X 的**方差**： $\sigma_X^2=E\left(X^2\right)-\mu_X^2=\frac{1}{12}(b-a)^2$

均匀分布的矩生成函数

- **矩生成函数**

$$
\begin{aligned}
M_X(t) &= \int_{-\infty}^{\infty} e^{tx} f_X(x)dx \\
&= \int_a^b e^{tx} \frac{1}{b-a} dx \\
&= \frac{1}{t(b-a)} e^{tx} \Big|_a^b \\
&= \frac{1}{t(b-a)} \left(e^{tb} - e^{ta}\right), -\infty < t < \infty
\end{aligned}
$$

标准均匀分布

- 对标准均匀分布 $U[0,1]$，

> ✓ **均值**为：$\frac{1}{2}$，
>
> ✓ **方差**为：$\frac{1}{12}$。

- 第三章已经证明，概率积分变换 $Y = F_X(X)$ 服从 $U[0,1]$ 分布。
- 均匀分布在统计学和计量经济学中有十分重要的应用。

4.3.2 贝塔分布 (Beta Distribution)

- 若连续随机变量 X 的 PDF 为

$$f_X(x)=\frac{1}{B(\alpha,\beta)}x^{\alpha-1}(1-x)^{\beta-1},0\le x\le 1$$

则称其服从贝塔分布 $BETA(\alpha,\beta)$，其中参数 $\alpha>0$， $\beta>0$，而 $B(\alpha,\beta)$ 称为**贝塔函数** (Beta function)，其定义为

$$B(\alpha,\beta)=\int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx$$
$$=\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

$\Gamma(\alpha)$ 称为**伽玛函数** (Gamma function)，其定义为

$$\Gamma(\alpha)=\int_0^\infty t^{\alpha-1}e^{-t}dt$$

引理 4.1: [$\Gamma(\alpha)$ 的性质]

(1) $\Gamma(\alpha+1)=\alpha\Gamma(\alpha)$;

(2) 若 k 是正整数，则 $\Gamma(k)=(k-1)!$;

(3) $\Gamma\left(\frac{1}{2}\right)=\sqrt{\pi}$。

贝塔分布的均值

- 贝塔分布 $BETA(\alpha,\beta)$ 的**均值**为

$$
\begin{aligned}
E(X) &= \int_0^1 \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{(\alpha+1)-1}(1-x)^{\beta-1}dx \\
&= \frac{\alpha}{\alpha+\beta}\int_0^1 \frac{(\alpha+\beta)\Gamma(\alpha+\beta)}{\alpha\Gamma(\alpha)\Gamma(\beta)} x^{(\alpha+1)-1}(1-x)^{\beta-1}dx \\
&= \frac{\alpha}{\alpha+\beta}\int_0^1 \frac{\Gamma(\alpha+1+\beta)}{\Gamma(\alpha+1)\Gamma(\beta)} x^{(\alpha+1)-1}(1-x)^{\beta-1}dx \\
&= \frac{\alpha}{\alpha+\beta}
\end{aligned}
$$

贝塔分布的二阶矩

- 为了计算方差，首先求**二阶矩**：

$$
\begin{aligned}
E\left(X^2\right) &= \int_0^1 \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{(\alpha+2)-1}(1-x)^{\beta-1}dx \\
&= \frac{\alpha}{(\alpha+\beta)} \int_0^1 \frac{\Gamma(\alpha+1+\beta)}{\Gamma(\alpha+1)\Gamma(\beta)} x^{(\alpha+2)-1}(1-x)^{\beta-1}dx \\
&= \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)} \int_0^1 \frac{\Gamma(\alpha+2+\beta)}{\Gamma(\alpha+2)\Gamma(\beta)} x^{(\alpha+2)-1}(1-x)^{\beta-1}dx \\
&= \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)}
\end{aligned}
$$

贝塔分布的方差

- 由此可得**方差**：

$$
\begin{aligned}
\operatorname{var}(X) &= E\left(X^2\right) - E^2(X) \\
&= \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)} - \left(\frac{\alpha}{\alpha+\beta}\right)^2 \\
&= \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}
\end{aligned}
$$

贝塔分布的 MGF

- 而 **MGF** 为

$$
\begin{aligned}
M_X(t) &= E\left(e^{tX}\right) \\
&= \int_0^1 \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} e^{tx} x^{\alpha-1}(1-x)^{\beta-1} dx \\
&= 1 + \sum_{j=1}^{\infty} \left(\prod_{i=0}^{j-1} \frac{\alpha+i}{\alpha+\beta+i}\right) \frac{t^j}{j!}
\end{aligned}
$$

其中，利用了麦克劳林级数展开

$$
e^{tx} = \sum_{j=0}^{\infty} \frac{t^j x^j}{j!}
$$

贝塔分布与均匀分布的关系

- 与均匀分布类似，$BETA(\alpha,\beta)$ 也是一类常用的有界支撑分布。
- 标准均匀分布 $U[0,1]$ 是 $BETA(\alpha,\beta)$ 分布当 $\alpha=\beta=1$ 时的特例。BETA(α,β) 分布的形状取决于参数 (α,β) 的取值，α 和 β 称为形状参数 (shape parameters)。

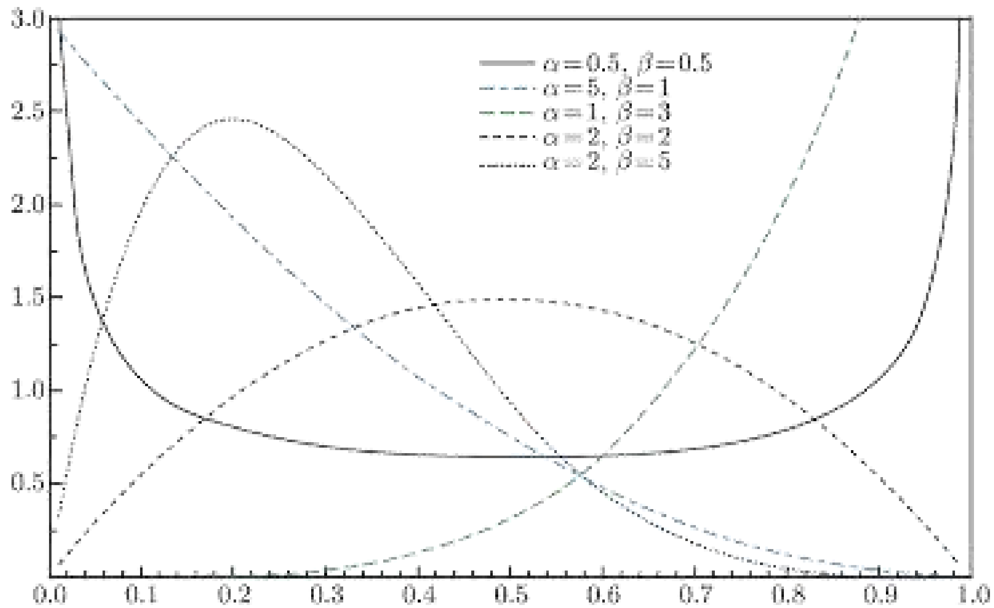


图 4.5：$BETA(\alpha,\beta)$ 分布的 PDF

贝塔分布的应用

- 由于 $BETA(\alpha,\beta)$ 的支撑为 $[0,1]$，贝塔分布可用于对取值落在区间 $[0,1]$ 的比率或者数量的**概率分布建模**。
  - ✓ 例如，Granger (1980) 用贝塔分布对个体消费者的**边际消费倾向**建模，发现具有“短期记忆 (short memory)”特征的个体消费加总后将显现“长期记忆 (long memory)”时间序列性质，即很久以前的消费与当期消费仍然高度相关。

4.3.3 正态分布 (Normal distribution)

- 若连续随机变量 X 的 PDF 为

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2}, -\infty < x < \infty$$

其中 $-\infty < \mu < \infty$ 和 $\sigma^2 > 0$，则称其服从正态分布，记作 $X \sim N(\mu, \sigma^2)$。

- 参数 μ 和 σ^2 分别为**位置** (location) 和**尺度** (scale) 参数。当 $\mu = 0$，$\sigma^2 = 1$ 时，$X \sim N(0,1)$ 称为**标准正态分布** (unit normal distribution)。

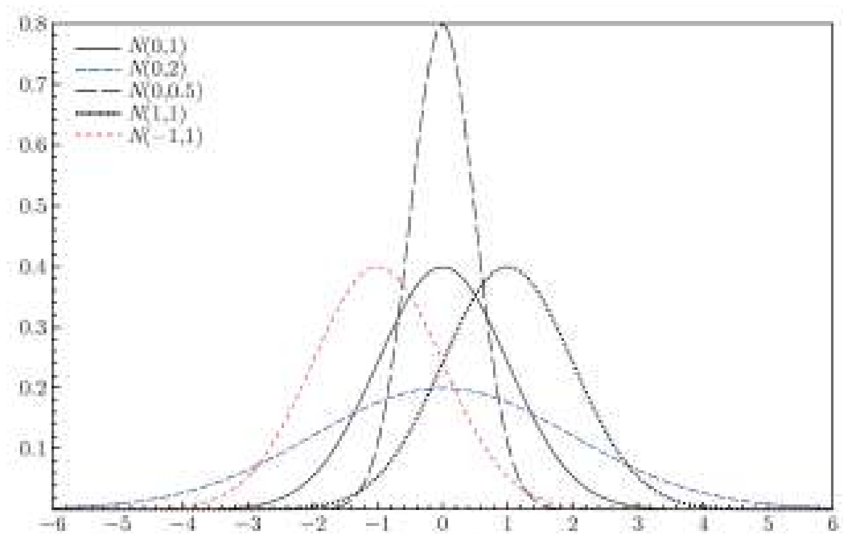


图 4.6：若干正态分布的 PDF

正态分布的发现与应用

- 亚伯拉罕·棣莫弗 (Abraham de Moivre) 1733 年在研究抛掷硬币概率（即**伯努利试验**，$p=\frac{1}{2}$）的近似计算时发现了正态分布。他将所发现的 PDF 命名为指数钟形曲线。
- 1809 年，约翰·高斯（Johann C. F. Gauss）用正态分布**预测星体**的位置，并因此确立了正态分布的重要地位。正态分布也因此常称为**高斯分布**。

正态分布与中心极限定理

- 作为概率论最重要的分布，正态分布在统计学和计量经济学中长期占据着核心位置。
- **中心极限定理** (central limit theorem, CLT) 为正态分布的广泛应用及其合理性提供了保障。其指出，在若干正则条件下，n 个独立同分布 (IID) 随机变量 $\{X_1,\cdots,X_n\}$ 的样本均值经标准化后，将随样本容量 n 的增加依分布收敛于标准正态分布，即当 $n\to\infty$ 时

$$\sqrt{n}\frac{\bar{X}_n-\mu_X}{\sigma_X}\xrightarrow{d}N(0,1),$$

其中 $\bar{X}_n=n^{-1}\sum_{i=1}^{n}X_i$ 为样本均值，且 $\xrightarrow{d}$ 表示依分布收敛。

正态分布与中心极限定理 (Cont.)

- 不论 X_i 是离散或连续随机变量，也无论 X_i 的支撑是有界或无界，中心极限定理均成立。
  - ✓ 例，假设 X_i 是伯努利 $\text{Bernoulli}(p)$ 随机变量，其中 $0 < p < 1$。$X = \sum_{i=1}^{n} X_i$ 服从二项分布 $B(n, p)$，当 $n \to \infty$ 时，$\frac{x-np}{\sqrt{np(1-p)}} \xrightarrow{d} N(0,1)$，故二项分布可任意用正态分布进行良好近似。

◆ 问题 4.4

如何证明对任意给定的 $-\infty < \mu < \infty$ 和 $\sigma^2 > 0$，

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = 1?$$

令 $y = (x-\mu)/\sigma$，可将上述问题转化为验证下式是否成立：

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy = 1$$

continue

■ **证明：** 由于

$$\left(\int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx\right)^2 = \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} dy$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{x^2+y^2}{2}} dxdy$$

令$x = r\cos(\theta)$, $y = r\sin(\theta)$

$$= \int_{0}^{\infty} \int_{0}^{2\pi} e^{-\frac{r^2}{2}} r dr d\theta$$

$$= 2\pi \int_{0}^{\infty} e^{-\frac{r^2}{2}} r dr$$

$$= 2\pi$$

故有

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx = 1$$

正态分布的矩

- 现在，计算正态分布的矩。
  - 首先，**均值**为

$$
\begin{aligned}
E(X) &= \int_{-\infty}^{\infty} x f_X(x)dx \quad (\text{令 } x = (x-\mu)+\mu) \\
&= \int_{-\infty}^{\infty} (x-\mu) f_X(x)dx + \mu \int_{-\infty}^{\infty} f_X(x)dx \\
&= \int_{-\infty}^{\infty} (x-\mu) \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{\sqrt{2\sigma^2}}dx+\mu} \quad (\text{令 } y = x-\mu) \\
&= \int_{-\infty}^{\infty} y \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{\pi^2}{2\sigma^2}} dy + \mu \\
&= \mu
\end{aligned}
$$

正态分布的矩 (Cont.)

- 其次，使用分部积分求**方差**：

$$\sigma_X^2 = \int_{-\infty}^{\infty} (x-\mu)^2 f_X(x)dx$$

$$= \int_{-\infty}^{\infty} (x-\mu)^2 \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \quad (令\ y = x-\mu)$$

$$= \int_{-\infty}^{\infty} y^2 \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{y^2}{2\sigma^2}} dy$$

$$= -\sigma^2 \int_{-\infty}^{\infty} y d\left(\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{y^2}{2\sigma^2}}\right) \quad \left(令\ u = y, v = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-y^2/2\sigma^2}\right)$$

$$= -\sigma^2 \left(y \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{y^2}{2\sigma^2}} \Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{y^2}{2\sigma^2}} dy \right)$$

$$= \sigma^2$$

因此，$X \sim N(\mu, \sigma^2)$ 的矩生成函数。

定理 4.1

假设 $X \sim N\left(\mu, \sigma^2\right)$，则

$$M_X(t) = e^{\mu t + \frac{\sigma^2}{2} t^2}, -\infty < t < \infty$$

正态分布的高阶矩

- 由于正态分布关于 μ 对称，故对任意整数 $k \geq 0$，中心化的奇数阶矩满足 $E(X-\mu)^{2k+1}=0$。
- 假设对任意正整数 k，需要计算偶数阶矩 $E(X-\mu)^{2k}$。这可通过对 $M_X(t)$ 求 $2k$ 次导数获得，但是，当 k 很大时，计算十分繁琐。
- 以下首先利用微分与积分运算的**对偶性**简化运算，推导出正态随机变量 X 的高阶矩 $E(X-\mu)^{2k}$ 的表达式。

正态分布的高阶矩 $E(X-\mu)^{2k}$

- 令 $\beta=\frac{1}{2\sigma^2}$ 或等价地 $\sigma=\frac{1}{\sqrt{2\beta}}$，得

$$E(X-\mu)^{2k}=\int_{-\infty}^{\infty}(x-\mu)^{2k}\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{1}{2\sigma^2}(x-\mu)^2}dx$$

$$=\int_{-\infty}^{\infty}y^{2k}\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{y^2}{2\sigma^2}}dy$$

$$=\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma}y^{2k}e^{-\beta y^2}dy$$

$$=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}(-1)^k\frac{d^k}{d\beta^k}e^{-\beta y^2}dy$$

$$=\frac{1}{\sqrt{2\pi}\sigma}(-1)^k\frac{d^k}{d\beta^k}\int_{-\infty}^{\infty}\sqrt{2\pi\sigma^2}\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{1}{2\sigma^2}y^2}dy$$

$$=\frac{1}{\sqrt{2\pi}\sigma}(-1)^k\frac{d^k}{d\beta^k}(\sqrt{2\pi}\sigma)$$

continue

正态分布的高阶矩 $E(X-\mu)^{2k}$

$$=\frac{1}{\sqrt{2\pi}\sigma}(-1)^k\sqrt{2\pi}\frac{d^k}{d\beta^k}\left(\frac{1}{\sqrt{2\beta}}\right) \qquad (注意\ \sigma=\frac{1}{\sqrt{2\beta}})$$

$$=\frac{1}{\sqrt{2}\sigma}(-1)^k\frac{d^k}{d\beta^k}\left(\beta^{-\frac{1}{2}}\right)$$

$$=\frac{1}{\sqrt{2}\sigma}(-1)^k\left(-\frac{1}{2}\right)\left(-\frac{3}{2}\right)\cdots\left(\frac{1}{2}-k\right)\beta^{-\frac{1}{2}-k}$$

$$=\frac{1}{\sqrt{\pi}}\Gamma\left(k+\frac{1}{2}\right)2^k\sigma^{2k}$$

对 $k=2$ 的特殊情形，有

$$E(X-\mu)^4=(-1)^2\left(-\frac{1}{2}\right)\left(-\frac{3}{2}\right)4\sigma^4=3\sigma^4$$

则 $N(\mu,\sigma^2)$ 的峰度为: $K=\frac{E(X-\mu)^4}{\sigma^4}=3$

引理 4.2: [斯特恩引理 (Stein's Lemma)]

假设 $X \sim N(\mu, \sigma^2)$，并且 $g(\cdot)$ 为满足 $E|g'(X)| < \infty$ 的可导函数，则 $E[g(X)(X-\mu)] = \sigma^2 E[g'(X)]$

■ **证明：** 根据分部积分公式，得

$$
\begin{aligned}
E[g(X)(X-\mu)] &= \int_{-\infty}^{\infty} g(x)(x-\mu)\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx \\
&= -\int_{-\infty}^{\infty} g(x)d\left[\frac{\sigma^2}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}\right] \\
&= -g(x)\frac{\sigma^2}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}\Bigg|_{-\infty}^{\infty} + \int_{-\infty}^{\infty}\frac{\sigma^2}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dg(x) \\
&= \sigma^2\int_{-\infty}^{\infty} g'(x)\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx \\
&= \sigma^2 E[g'(X)]
\end{aligned}
$$

4.3.4 柯西与稳态分布

(Cauchy Distribution & Stationary Distribution)

- 若连续随机变量 X 的 PDF 为

$$f_X(x) = \frac{1}{\pi\sigma}\frac{1}{1+\left(\frac{x-\mu}{\sigma}\right)^2}, -\infty < x < \infty$$

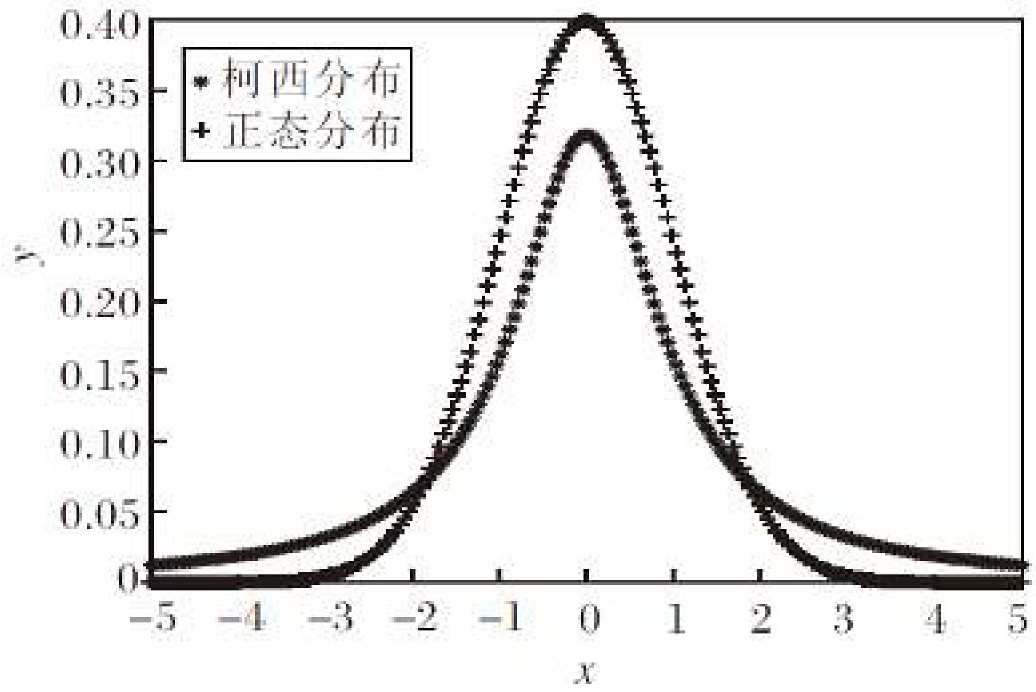
  其中参数 $\sigma > 0$，则称其服从柯西分布 $\text{Cauchy}(\mu, \sigma)$。

- 参数 μ 和 σ 分别为**位置**参数和**尺度**参数。该分布关于 μ 对称，其支撑无界。

- 当 $\mu = 0$，$\sigma = 1$ 时，该分布称为**标准柯西分布**，记作 $\text{Cauchy}(0, 1)$。

柯西分布的意义

- 柯西分布的实际应用价值不大，但具有特殊的理论重要性。
- 对 Cauchy(μ,σ) 分布，其 PDF 的尾部以非常慢的双曲速度衰减至 0。因此，柯西分布的所有大于等于 1 阶的各阶矩均不存在，因而其 MGF 也不存在。



柯西分布的均值

- 对于 Cauchy(0, 1) 分布，有

$$
\begin{aligned}
E|X| &= \int_{-\infty}^{\infty} |x| f_X(x) dx \\
&= \int_{-\infty}^{\infty} |\mathrm{x}| \frac{1}{\pi} \frac{1}{1+\mathrm{x}^2} \mathrm{dx} \\
&= \frac{2}{\pi} \int_0^{\infty} \frac{\mathrm{x}}{1+\mathrm{x}^2} \mathrm{dx} \\
&= \frac{1}{\pi} \ln\left(1+\mathrm{x}^2\right)\Big|_0^{\infty} \\
&= \infty
\end{aligned}
$$

- 这表明柯西分布的均值以及所有高阶矩均不存在。因此，位置参数 μ 不能解释为均值，而尺度参数 σ 也不能解释为标准差。

柯西分布的特征函数

- Cauchy(μ, σ) 的**特征函数**为

$$\begin{aligned}\varphi_X(t) &= E\left(e^{\mathbf{i}tX}\right) \\ &= e^{\mathbf{i}\mu t-\sigma|t|}\end{aligned}$$

该特征函数关于 t 在原点处不可导，这与其所有大于等于 1 阶的各阶矩均不存在的性质一致。

◆ **问题 4.5**

何种情况会出现柯西分布?

- 两个独立正态随机变量的比率服从柯西分布。

- 实际上，柯西分布和正态分布同属于所谓的**稳态分布**（stable distributions），稳态分布是概率论非常重要的一类分布。

◆ 问题 4.6

什么是稳态分布?

- 稳态分布的 PDF 通常没有解析形式，但它具有解析形式的特征函数:

$$\varphi_X(t)=e^{\mathbf{i}\mu t-\sigma|t|^c[1+\mathbf{i}\lambda \operatorname{sgn}(t)\omega(|t|,c)]}$$

其中 $\mathbf{i}=\sqrt{-1}$, $0<c\le 2$, $-1\le\lambda\le 1$, $\sigma>0$，并且

$$\omega(|t|,c)=\begin{cases}\tan\left(\frac{1}{2}\pi c\right), & c\neq 1\\ -\frac{2}{\pi}\ln(|t|), & c=1\end{cases}\qquad sgn(t)=\begin{cases}1, & t>0\\ 0, & t=0\\ -1, & t<0\end{cases}$$

continue

- 直观上，μ 是**位置** (location) 参数，σ 是**尺度** (scale) 参数，c 是**尾部** (tail) 参数，λ 是**偏度** (skew) 参数。PDF 的形状由 c 和 λ 共同决定。
  - ✓ 当 $\lambda = 0$ 时，则为**对称分布**；
  - ✓ 当 $\lambda = 0$ 时，令 $c = 2$ 可得**正态分布**；
  - ✓ 当 $\lambda = 0$ 时，令 $c = 1$ 可得**柯西分布**。

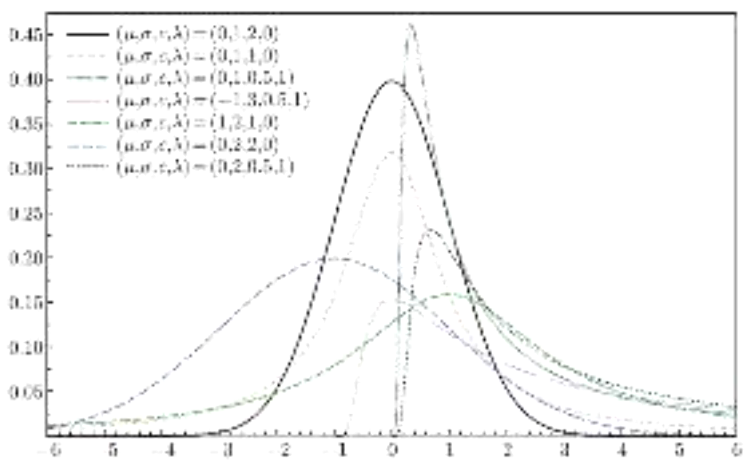


图 4.7：参数 (μ,σ,c,λ) 取不同值时对应的稳态分布的 PDF

- 稳态分布的矩仅当 $c > 1$ 时存在。若 $c < 2$，则稳态分布的方差不存在。
- 稳态分布与金融计量经济学的近期研究热点之一 —— **列维过程** (Levy process) 密切相关。列维过程可更为恰当地刻画金融数据中通常出现的**厚尾现象** (heavy tails)。

✓ Mandelbrot (1963) 和 Fama (1965) 曾用稳态分布对股票收益率建模。

4.3.5 对数正态分布 (Logarithmic Normal Distribution)

- 若连续随机变量 X 的 PDF 为

$$f_X(x) = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma}\frac{1}{x}e^{\frac{1}{2\sigma^2}(\ln x-\mu)^2}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

- 则称其服从对数正态分布，记作 $LN(\mu,\sigma^2)$。

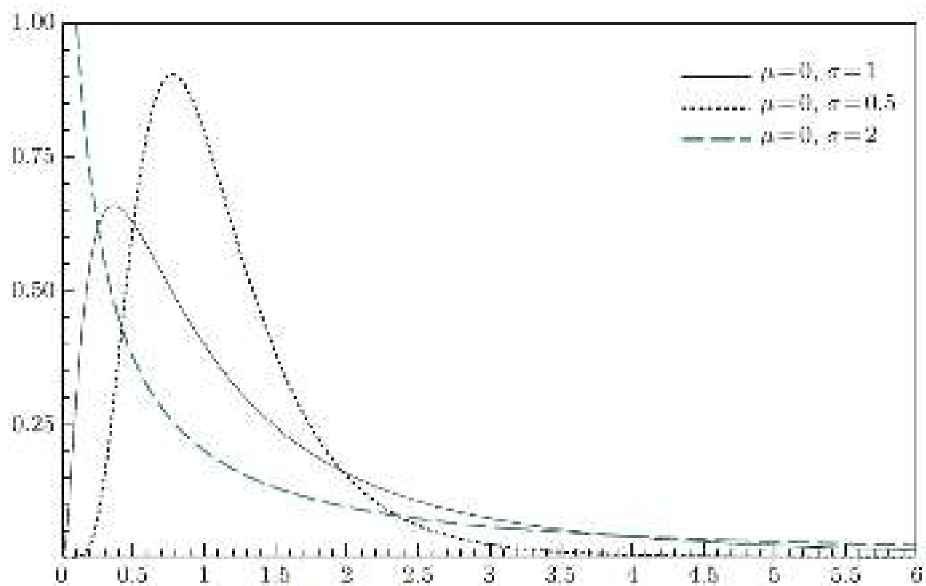


图 4.8：对数正态分布的 PDF

对数正态分布与正态分布的关联

- 若 $X \sim LN(\mu, \sigma^2)$，则 $Y = \ln(X) \sim N(\mu, \sigma^2)$。
- 事实上，正是由于随机变量 X 的对数形式服从正态分布，该随机变量被称为对数正态随机变量。
- **对数正态分布**有时也称为**反对数正态分布**。
  - 这一命名的逻辑基础在于该分布并非对正态变量取对数的分布，而是对其取指数的分布（$X = e^Y$），即随机变量的反对数函数服从正态分布。

对数正态分布的各阶矩

- **正态随机变量** $Y \sim N(\mu,\sigma^2)$ 的 **MGF** 为

$$M_Y(t)=E\left(e^{tY}\right)=e^{\mu t+\frac{\sigma^2}{t^2}t^2}$$

- 因此，**对数正态随机变量** $X \sim LN(\mu,\sigma^2)$ 的**各阶矩**均存在且

$$\begin{aligned} E\left(X^k\right) &= E\left(e^{kY}\right) \\ &= M_Y(k) \\ &= e^{k\mu+\frac{\varepsilon^2}{2}k^2}, k=1,2,\cdots \end{aligned}$$

✓ **均值**：$\mu_X=e^{\mu+\frac{\alpha^2}{2}}$

✓ **方差**：$\sigma_X^2=E\left(X^2\right)-\mu_X^2=e^{2\mu+\sigma^2}\left(e^{\sigma^2}-1\right)$

对数正态分布的 MGF

尽管对数正态分布的各阶矩均存在，但其**矩生成函数**并不存在。

$$
\begin{aligned}
M_X(t) &= E(e^{tx}) \\
&= \int_0^{\infty} e^{tx} \frac{1}{\sqrt{2\pi}\sigma} \frac{1}{x} e^{-\frac{1}{2\sigma^z}(\ln x-\mu)^2} dx \\
&= \int_0^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{te^{\ln x}-\frac{1}{2\sigma^2}(\ln x-\mu)^2} d\ln x \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{te^{y}-\frac{1}{2\sigma^2}(y-\mu)^2} dy \quad (\text{令 } y = \ln x) \\
&\geq \int_{M}^{M+1} \frac{1}{\sqrt{2\pi}\sigma} e^{te^{y}-\frac{1}{2\sigma^2}(y-\mu)^2} dy \quad (\text{对任意 } M > 0) \\
&\geq \frac{1}{\sqrt{2\pi}\sigma} e^{te^{M}-\frac{1}{2\sigma^2}(M-\mu)^2} (M+1-M) \quad (\text{对足够大的 } M) \\
&= \frac{1}{\sqrt{2\pi}\sigma} e^{te^{M}-\frac{1}{2\sigma^2}(M-\mu)^2} \\
&\to \infty, \text{ 当 } M \to \infty
\end{aligned}
$$

- 对数正态分布广泛应用于对非负且右偏随机变量的建模。特别地，该分布被广泛地应用于刻画资产价格、商品价格、收入以及人口等变量的概率分布。
- 为了说明这一点，考察**非负经济变量**

$$X_t = X_{t-1}(1 + Y_t)$$

其中 $\{Y_t\}$ 为 IID 随机变量序列。这里，X_{t-1} 可理解为时期 $t-1$ 的基数，而 Y_t 是从时期 $t-1$ 到时期 t 的随机增长率。

continue

- 根据该递归关系，有

$$X_n = X_0 \prod_{t=1}^{n} (1 + Y_t)$$

- 因此

$$\ln X_n = \ln X_0 + \sum_{t=1}^{n} \ln(1 + Y_t)$$

- 根据中心极限定理，当 n 很大时，$\sum_{t=1}^{n} \ln(1 + Y_t)$ 经过标准化后近似服从正态分布。因此，其指数形式 $\prod_{t=1}^{n} (1 + Y_t)$ 近似服从对数正态分布。

对数正态分布的应用

- 假设某股票价格 $P_t \sim LN(\mu t, \sigma^2 t)$，其中时间 t 可连续变化。则 $\ln P_t \sim N(\mu t, \sigma^2 t)$，其**对数收益率**为

$$
\begin{aligned}
R_t &= \ln(P_t / P_{t-1}) \\
&= \ln(P_t) - \ln(P_{t-1})
\end{aligned}
$$

  该式近似等于股票价格从时期 $t-1$ 到时期 t 的相对价格变化率，且服从正态分布。因此，从时期 $t=1$ 到时期 $t=m$ 的 m 期累积收益率 $\sum_{t=1}^{m} R_t$ 同样服从正态分布。
- Black & Scholes (1973) 在推导欧式期权价格时，假设股票价格服从对数正态分布。

对数正态分布的应用 (Cont.)

- 对数正态分布也可用于描述各种经济单位规模。
  - ✓ Eeckhout (2004) 基于 2000 年美国人口普查数据，指出美国城市规模服从对数正态分布，并提出一个具有局部外部性的经济活动均衡模型来解释其对数正态分布。

Gibrat's Law for (All) Cities

By Jan Eeckhout*

Two empirical regularities concerning the size distribution of cities have repeatedly been established: Zipf's law holds (the upper tail is Pareto), and city growth is proportionate. Census 2000 data are used covering the entire size distribution, not just the upper tail. The nontruncated distribution is shown to be lognormal, rather than Pareto. This provides a simple justification for the coexistence of proportionate growth and the resulting lognormal distribution. An equilibrium theory of local externalities that can explain the empirical size distribution of cities is proposed. The driving force is a random productivity process of local economies and the perfect mobility of workers. (*JEL* D30, D51, J61, R12)

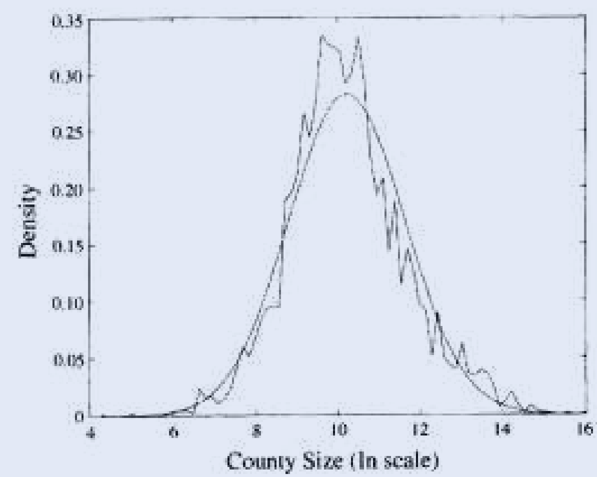


Figure A-1. Empirical and Theoretical Density Functions of All Counties

对数正态分布的应用 (Cont.)

- ✓ 该分布还在描述员工任职年限与其离职之间的关系上取得了巨大成功 (Young, 1971; McClean, 1976)。
- ✓ O’Neil & Wells (1972) 指出，对数正态分布可有效地对个人保险理赔分布建模。
- ✓ Alizadeh et al. (2002) 指出，资产价格范围（即最高价减最低价）观测数据近似服从对数正态分布，并通过高斯拟极大似然估计法，利用该典型经验特征事实预估随机波动模型。

4.3.6 伽玛分布与广义伽玛分布

(Gamma Distribution and Generalized gamma distribution)

- 若非负连续随机变量 X 的 PDF 为

$$f_X(x)=\begin{cases}\frac{1}{\Gamma(\alpha)\beta^{\alpha}}x^{\alpha-1}e^{-x/\beta}, & x>0\\ 0, & x\le 0\end{cases}$$

其中参数 $\alpha,\beta>0$，$\Gamma(\alpha)=\int_0^{\infty}t^{\alpha-1}e^{-t}dt$ 为伽玛函数，则称其服从伽玛分布 $G(\alpha,\beta)$。

- 伽玛分布 $G(\alpha,\beta)$ 是刻画区间 $[0,\infty)$ 上非负随机变量的一类非常灵活的分布族，其中 α 为**形状参数**，β 为控制分布分散程度的**尺度参数**。
- 当 $\beta=1$ 时，伽玛分布 $G(\alpha,1)$ 称为**标准伽玛分布**。

伽玛分布的均值、方差以及 MGF

- **均值**为

$$\mu_X = \int_{-\infty}^{\infty} x f_X(x)dx = \int_0^{\infty} \frac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{\alpha} e^{-x/\beta} dx$$

$$= \int_0^{\infty} \frac{\alpha\beta}{\alpha\Gamma(\alpha)\beta^{\alpha+1}} x^{(\alpha+1)-1} e^{-x/\beta} dx$$

$$= \alpha\beta \int_0^{\infty} \frac{1}{\Gamma(\alpha^*)\beta^{\alpha^*}} x^{\alpha^*-1} e^{-x/\beta} dx$$

$$= \alpha\beta$$

令 $\alpha^* = \alpha + 1$
且 $\alpha\Gamma(\alpha) = \Gamma(\alpha + 1)$

伽玛分布的均值、方差以及 MGF

- 其**二阶矩**为

$$E(X^2) = \int_0^\infty \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha+2-1} e^{-x/\beta} dx$$

$$= \int_0^\infty \frac{\alpha(\alpha+1)\beta^2}{\Gamma(\alpha+2)\beta^{\alpha+2}} x^{\alpha+2-1} e^{-x/\beta} dx$$

$$= \alpha(\alpha+1)\beta^2 \int_0^\infty \frac{1}{\Gamma(\alpha+2)\beta^{\alpha+2}} x^{\alpha+2-1} e^{-x/\beta} dx$$

$$= \alpha(\alpha+1)\beta^2$$

- 由此可得**方差**

$$\sigma_X^2 = E(X^2) - \mu_X^2 = \alpha\beta^2$$

伽玛分布的均值、方差以及 MGF

- 其 **MGF** 为

$$
\begin{aligned}
M_X(t) &= \int_0^\infty e^{tx}\frac{1}{\Gamma(\alpha)\beta^\alpha}x^{\alpha-1}e^{-x/\beta}dx \\
&= \int_0^\infty \frac{1}{\Gamma(\alpha)\beta^\alpha}x^{\alpha-1}e^{-x(1/\beta-t)}dx \ (\text{令 } \beta^* = \tfrac{1}{1/\beta-t}) \\
&= \int_0^\infty \frac{1}{\Gamma(\alpha)\beta^\alpha}x^{\alpha-1}e^{-x/\beta^*}dx \\
&= \frac{(\beta^*)^\alpha}{\beta^\alpha}\int_0^\infty \frac{1}{\Gamma(\alpha)(\beta^*)^\alpha}x^{\alpha-1}e^{-x/\beta^*}dx = \frac{(\beta^*)^\alpha}{\beta^\alpha} \\
&= (1-\beta t)^{-\alpha}, t < 1/\beta
\end{aligned}
$$

伽玛分布的应用

- 伽玛分布的形状与对数正态分布类似，可用于对经济事件持续等待时间的分布建模（如失业持续期、价格久期、贫困持续期、企业存活期等），也可用于对非负随机变量，如收入、人口和距离等建模。
- Cox et al. (1985) 提出了一个即期利率期限结构的**连续时间均衡模型**。假设即期利率服从一个平方根过程，即

$$dr_t = k(\theta - r_t)dt + \sigma\sqrt{r_t}dB_t$$

其中 B_t 为布朗运动。给定参数 $k, \theta > 0$，利率将趋于中心位置或长期均衡值 θ，而参数 k 决定其调整速度。

continue

伽玛分布的应用 (Cont.)

- 关于即期利率的上述设定避免了类似像 Vasicek (1977) 模型中可能出现的负利率情况。在 Vasicek (1977) 模型中，

$$dr_t = k(\theta - r_t)dt + \sigma dB_t$$

可以证明 Cox et al. (1985) 模型的即期利率 r_t 服从具有如下 PDF 的伽玛分布

$$f(r) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} r^{\alpha-1} e^{-r/\beta}$$

其中 $\alpha = \frac{2k\theta}{\sigma^2}$ 和 $\beta = \frac{\sigma^2}{2k}$。

- 因此，短期利率 r_t 的稳态均值和方差分别为 θ 和 $\sigma^2\theta/2k$。

广义伽玛分布 (Generalized Gamma Distributions)

- 假设随机变量 Y

$$Y = \left(\frac{X-\gamma}{\beta}\right)^{c}$$

服从**标准伽玛分布** (standard Gamma distribution)，即 $Y \sim G(\alpha, 1)$ 或者等价地有 PDF

$$f_Y(y) = \frac{y^{\alpha-1}e^{-y}}{\Gamma(\alpha)}, y \geq 0$$

- 则称随机变量 X 服从**形状参数**为 α 和 c，**尺度参数**为 β，**位置参数**为 γ 的**广义伽玛分布**。

广义伽玛分布的 PDF 与矩表达式

- 根据单变量变换方法 (定理 3.12)，可证 X 的 **PDF** 为

$$f_X(x)=\frac{c}{\Gamma(\alpha)\beta^{c\alpha}}(x-\gamma)^{c\alpha-1}e^{-\left(\frac{x-\gamma}{\beta}\right)^{\varepsilon}},x\geq\gamma$$

- **广义伽玛分布的矩**可通过以下关系从标准伽玛分布 $G(\alpha,1)$ 的矩表达式求得：

$$E\left[\left(\frac{X-\gamma}{\beta}\right)^{k}\right]=E\left[\left(\frac{X-\gamma}{\beta}\right)^{c(k/c)}\right]$$

$$=E\left(Y^{(k/c)}\right)$$

$$=\frac{\Gamma(\alpha+k/c)}{\Gamma(\alpha)}$$

4.3.7 卡方分布 (Chi-Square Distribution)

- 若非负连续随机变量 X 的 PDF 为

$$f_X(x)=\begin{cases}\frac{1}{\Gamma\left(\frac{v}{2}\right)\sqrt{2^v}}x^{\frac{v}{2}-1}e^{-\frac{x}{2}}, & x>0\\ 0, & x\le 0\end{cases}$$

- 则称其服从自由度为 ν 的卡方分布，记作 χ_ν^2。
- 卡方分布 χ_ν^2 是伽玛分布 $G(\alpha,\beta)$ 当 $\alpha=\frac{\nu}{2}$，$\beta=2$ 时的特例，其 ***k* 阶矩**为

$$E\left(X^k\right)=\frac{2^k\Gamma\left(\frac{\nu}{2}+k\right)}{\Gamma\left(\frac{\nu}{2}\right)}$$

卡方分布的均值、方差及 MGF

- ✓ **均值**: $E(X) = \nu$
- ✓ **方差**: $\mathrm{var}(X) = 2\nu$
- ✓ **MGF**: $M_X(t) = (1-2t)^{-\frac{\nu}{2}}, t < \frac{1}{2}$

卡方分布与正态分布的关系

- 卡方分布 χ_ν^2 的上述定义允许自由度参数 ν 取**非整数值**。第五章将指出，当 ν 为整数时，χ_ν^2 分布等价于 ν 个独立 $N(0,1)$ 随机变量的平方和。
- 卡方分布 χ_ν^2 是右偏态分布。当自由度 $\nu \to \infty$ 时，χ_ν^2 分布近似于正态分布 $N(\nu, 2\nu)$。
- 与正态分布类似，χ_ν^2 分布也是概率论与统计学最为重要的分布之一，并在计量经济学占据核心地位。计量经济学的许多常用检验统计量都是二次项形式的，而且渐近服从于 χ_ν^2 分布。

4.3.8 指数分布与韦伯分布

(Exponential Distribution and Weibull distribution)

- 伽玛分布的另一种特例称为**指数分布**。若非负连续随机变量 X 的 PDF 为

$$f_X(x) = \begin{cases} \frac{1}{\beta} e^{-x/\beta}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

其中 $\beta > 0$ 为**尺度参数**，则称其服从指 $EXP(\beta)$ 分布。

- 当 $\beta = 1$ 时，则称 X 服从**标准指数分布**，记作 $EXP(1)$。

指数分布的MGF、均值和方差

- 指数分布 $EXP(\beta)$ 是伽玛分布 $G(\alpha,\beta)$ 当 $\alpha=1$ 时的一个特例。

$EXP(\beta)$ 的**MGF**为：$M_X(t)=E\left(e^{tX}\right)=\frac{1}{1-\beta t},t<\frac{1}{\beta}$

$EXP(\beta)$ 的**均值**为：$E(X)=\beta$

$EXP(\beta)$ 的**方差**为：$\mathrm{var}(X)=\beta^2$

指数分布的应用：劳动经济学案例

- 以下是劳动经济学的一个例子：令 X 表示具有 PDF $f_X(x)$ 的工人失业持续期，其相应的风险率或风险函数定义如下

$$\begin{aligned}\lambda(x) &= \lim_{\Delta x \to 0^+} \frac{P(X \le x + \Delta x \mid X > x)}{\Delta x} \\ &= \lim_{\Delta x \to 0^+} \frac{P(x < X \le x + \Delta x)}{P(X > x)\Delta x} \\ &= \left[\lim_{\Delta x \to 0^+} \frac{\int_x^{x+\Delta x} f_X(u)du}{\Delta x}\right] \frac{1}{P(X > x)} \\ &= \frac{f_X(x)}{P(X > x)} \\ &= \frac{f_X(x)}{1 - F_X(x)} \\ &= -\frac{d}{dx} \ln[1 - F_X(x)]\end{aligned}$$

continue

指数分布的应用：劳动经济学案例 (Cont.)

- 直观上，风险函数 $\lambda(x)$ 是失业工人在持续失业 x 时期后找到工作的瞬时概率。
- 久期分析试图使用一些经济变量解释 $\lambda(x)$。最简单的参照例子是假设对任意失业持续期 x，风险率均为常数，即

$$\lambda(x) = \lambda_0, \text{对所有 } x$$

- 此时，可以证明失业持续期 X 服从指数分布 $EXP(1/\lambda_0)$：

$$f_X(x) = \lambda_0 e^{-\lambda_0 x}, x > 0$$

指数分布的应用

- 在金融计量经济学中，高频股票收益率 $\{X_t\}$ 的一个经验典型特征事实是股票收益的绝对值 $|X_t|$ 近似地服从于标准指数分布 (Ding et al., 1993)，这里 X_t 为第 t 期的标准化资本收益率。

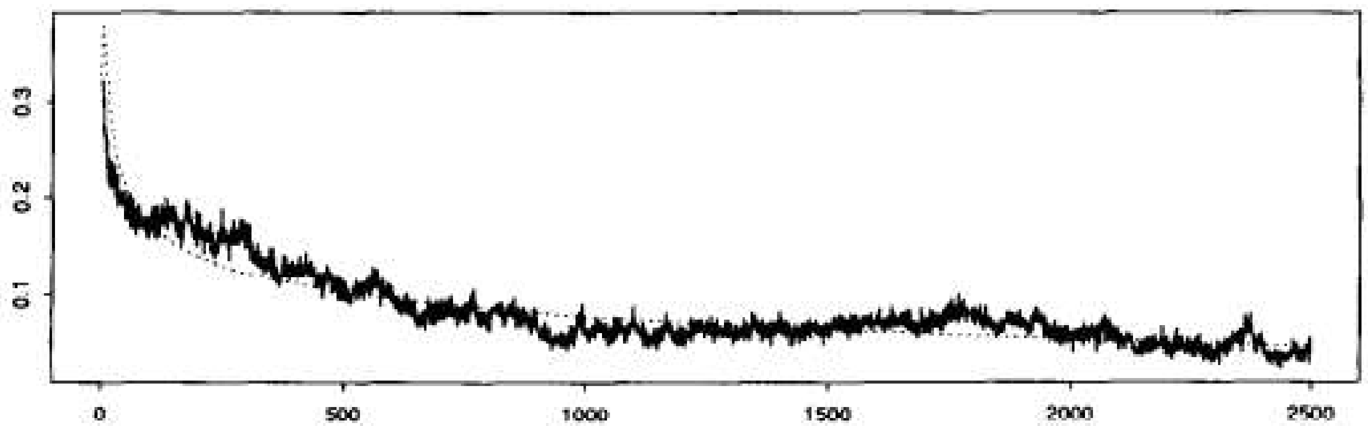


Fig. 3.9. Autocorrelation of $|r|$ (solid line) and its fitted value (dotted line).

韦伯分布 (Weibull distribution)

- 许多重要分布都与指数分布紧密相关。例如，若 $Y=(X-\alpha)^c$ 服从参数为 β 的**指数分布**，则称 X 服从**韦伯 (Weibull) 分布**，其 PDF 为

$$f_X(x)=\frac{c}{\beta}(x-\alpha)^{c-1}e^{-\frac{(x-\alpha)^c}{\beta}}, x>\alpha$$

其中 α 为**位置参数**，β 为**尺度参数**，c 为**形状参数**，且 c 必须大于 1。实际应用中通常设 $\alpha=0$。

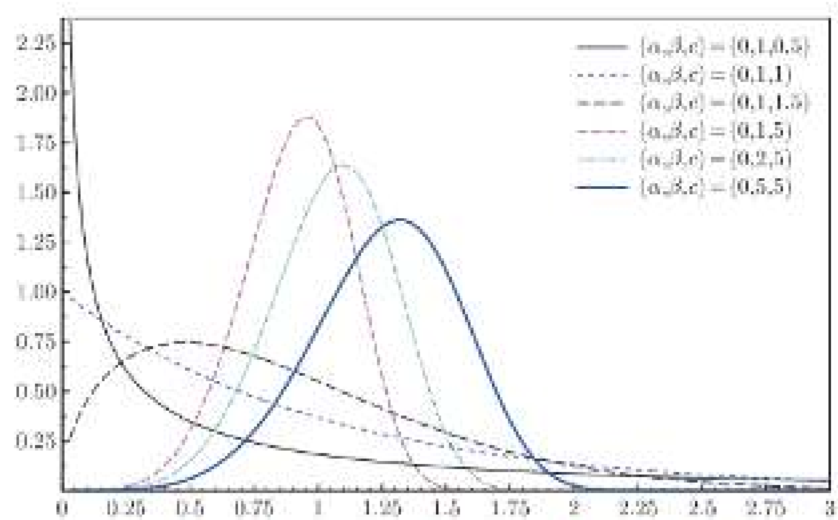


图 4.11：韦伯分布在各参数 (α,β,c) 取值下的 PDF

4.3.9 双指数分布 (Laplace distribution)

- 若连续随机变量 X 的 PDF 为

$$f_X(x) = \frac{1}{2\beta} e^{-\frac{|x-\alpha|}{\beta}}, -\infty < x < \infty$$

其中参数 $\beta > 0$，则称其服从双指数 $DEXP(\alpha, \beta)$ 分布。

- $DEXP(\alpha, \beta)$ 分布是关于 α **对称**的分布，其尾部相对于正态分布更厚。该分布在 $x = \alpha$ 处取得峰值且其导数在该点不存在。

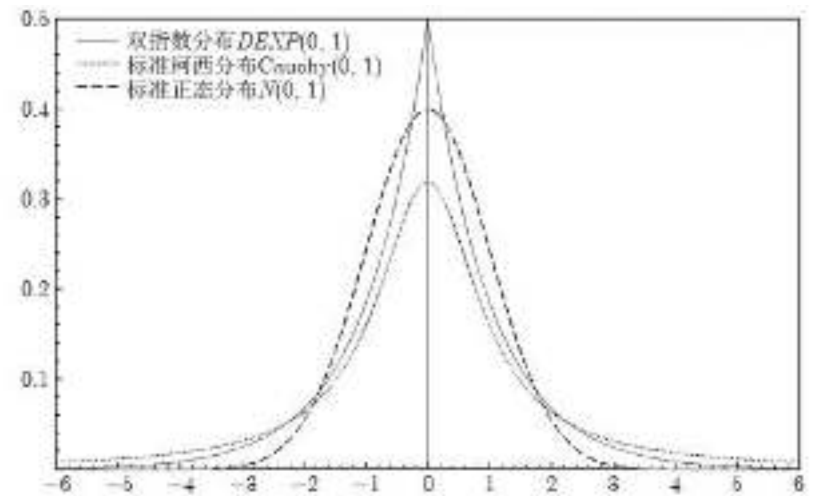


图 4.12：Cauchy(0, 1) 分布、$N(0,1)$ 分布以及 $DEXP(0,1)$ 分布的 PDF

双指数分布的均值、方差与 MGF

- ✓ **均值**：$E(X) = \alpha$
- ✓ **方差**：$\mathrm{var}(X) = 2\beta^2$
- ✓ **MGF**：$M_X(t) = \frac{e^{\alpha t}}{1-\beta^2 t^2}, |t| < \frac{1}{\beta}$

双指数分布也称为**拉普拉斯分布**。当 $\alpha = 0$ 时，X 的绝对值，$Y = |X|$，服从指数分布 $EXP(\beta)$。

目 录

第一节 介绍

第二节 离散概率分布

第三节 连续概率分布

第四节 小结

小结

- 本章介绍了一些重要的离散概率分布和连续概率分布，并探讨了各分布的性质及其相互关系。
- 离散分布包括：
  - ✓ 伯努利分布
  - ✓ 二项分布
  - ✓ 负二项分布
  - ✓ 几何分布
  - ✓ 泊松分布
  - ✓ …

小结

- 连续分布包括：
  - ✓ 均匀分布
  - ✓ 贝塔分布
  - ✓ 正态分布
  - ✓ 柯西分布 & 稳态分布
  - ✓ 对数正态分布
  - ✓ 伽玛分布 & 广义伽玛分布
  - ✓ 卡方分布
  - ✓ 指数分布 &韦伯分布
  - ✓ 双指数分布
  - ✓ …

中国科学院数学与系统科学研究院
Academy of Mathematics and Systems Science
Chinese Academy of Sciences

Thank You！