第二章　随机变量及其分布

§2.1　随机变量及其分布

1. 随机变量　定义在样本空间 Ω 上的实值函数 $X=X(\omega)$ 称为**随机变量**.

（1）仅取有限个或可列个值的随机变量称为**离散随机变量**；

（2）取值充满某个区间 (a,b) 的随机变量称为**连续随机变量**，这里 a 可为 $-\infty$，b 可为 $+\infty$.

2. 分布函数　设 X 是一个随机变量，对任意实数 x，称 $F(x)=P(X\leqslant x)$ 为 X 的**分布函数**，记为 $X\sim F(x)$. 分布函数具有如下三条基本性质：

（1）**单调性**　$F(x)$ 是单调非减函数，即对任意的 $x_1<x_2$，有 $F(x_1)\leqslant F(x_2)$.

（2）**有界性**　对任意的 x，有 $0\leqslant F(x)\leqslant 1$，且

$$F(-\infty)=\lim_{x\to-\infty}F(x)=0,\qquad F(+\infty)=\lim_{x\to+\infty}F(x)=1.$$

（3）**右连续性**　$F(x)$ 是 x 的右连续函数，即对任意的 x_0，有

$$\lim_{x\to x_0+0}F(x)=F(x_0),\qquad 即\quad F(x_0+0)=F(x_0).$$

反之，可以证明：具有上述三条性质的函数 $F(x)$ 一定是某一个随机变量的分布函数.

3. 离散随机变量的概率分布列　若离散随机变量 X 的可能取值是 x_1，x_2，…，x_n，…，则称 X 取 x_i 的概率

$$p_i=p(x_i)=P(X=x_i),i=1,2,\cdots,n,\cdots$$

为 X 的**概率分布列**，简称**分布列**. 分布列也可用列表方式来表示：

X	x_1	x_2	…	x_n	…
P	$p(x_1)$	$p(x_2)$	…	$p(x_n)$	…

分布列 $p(x_i)$ 具有如下两条基本性质：

(1) **非负性**：$p(x_i) \geqslant 0, i=1,2,\cdots$；　　(2) **正则性**：$\sum_{i=1}^{+\infty} p(x_i) = 1$.

离散随机变量 X 的分布函数为 $F(x) = \sum_{x_i \leqslant x} p(x_i)$，它是有限级或可列无限级阶梯函数.离散随机变量 X 取值于区间 $(a,b]$ 上的概率为 $P(a<X\leqslant b)=F(b)-F(a)$.常数 c 可看作仅取一个值的随机变量 X，即 $P(X=c)=1$，它的分布常称为**单点分布**或**退化分布**.

4. 连续随机变量的概率密度函数 记连续随机变量 X 的分布函数为 $F(x)$，若存在一个非负可积函数 $p(x)$，使得对任意实数 x，有 $F(x) = \int_{-\infty}^{x} p(t)\,\mathrm{d}t$，则称 $p(x)$ 为 X 的**概率密度函数**，简称为**密度函数**.密度函数 $p(x)$ 具有如下两条基本性质：

(1) **非负性**：$p(x) \geqslant 0$；　　(2) **正则性**：$\int_{-\infty}^{+\infty} p(x)\,\mathrm{d}x = 1$.

连续随机变量 X 的分布函数 $F(x)$ 是 $(-\infty,+\infty)$ 上的连续函数，在 $F(x)$ 导数存在的点上有 $F'(x)=p(x)$.

连续随机变量 X 仅取一点值的概率恒为零，从而有

$$P(a \leqslant X \leqslant b) = P(a < X \leqslant b) = P(a \leqslant X < b) = P(a < X < b) = \int_a^b p(x)\,\mathrm{d}x.$$

这给计算带来很大方便.

连续随机变量 X 的密度函数不唯一.

5. 设随机变量 X 的分布函数为 $F(x)$，则可用 $F(x)$ 表示下列概率：

(1) $P(X\leqslant a)=F(a)$；

(2) $P(X<a)=F(a-0)$；

(3) $P(X>a)=1-P(X\leqslant a)=1-F(a)$；

(4) $P(X=a)=P(X\leqslant a)-P(X<a)=F(a)-F(a-0)$；

(5) $P(X\geqslant a)=1-P(X<a)=1-F(a-0)$；

(6) $P(|X|<a)=P(-a<X<a)=P(X<a)-P(X\leqslant -a) = F(a-0)-F(-a)$.

§2.2 随机变量的数学期望

1. 数学期望 设随机变量 X 的分布用分布列 $p(x_i)$ 或用密度函数 $p(x)$ 表示,若

$$\begin{cases} \sum_i |x_i| p(x_i) < +\infty, & \text{当 } X \text{ 为离散随机变量}, \\ \int_{-\infty}^{+\infty} |x| p(x) \mathrm{d}x < +\infty, & \text{当 } X \text{ 为连续随机变量}, \end{cases}$$

则称

$$E(X) = \begin{cases} \sum_i x_i p(x_i), & \text{当 } X \text{ 为离散随机变量}, \\ \int_{-\infty}^{+\infty} x p(x) \mathrm{d}x, & \text{当 } X \text{ 为连续随机变量} \end{cases}$$

为 X 的数学期望,简称期望或均值,且称 X 的数学期望存在.否则称数学期望不存在.

数学期望是由分布决定的,它是分布的位置特征.如果两个随机变量同分布,则其数学期望(存在的话)总是相等的.

2. 数学期望的性质 以下所涉及的数学期望均假定其存在.

(1) X 的某一函数 $g(X)$ 的数学期望为

$$E[g(X)] = \begin{cases} \sum_i g(x_i) p(x_i), & \text{在离散场合}, \\ \int_{-\infty}^{+\infty} g(x) p(x) \mathrm{d}x, & \text{在连续场合}; \end{cases}$$

(2) 若 c 是常数,则 $E(c)=c$;

(3) 对任意常数 a,有 $E(aX)=aE(X)$;

(4) 对任意的两个函数 $g_1(x)$ 和 $g_2(x)$,有

$$E[g_1(X) \pm g_2(X)] = E[g_1(X)] \pm E[g_2(X)].$$

§2.3 随机变量的方差与标准差

1. 方差 随机变量 X 对其期望 $E(X)$ 的偏差平方的数学期望(设其存在)

$$\mathrm{Var}(X)=E[X-E(X)]^2$$

称为 X 的方差,方差的正平方根 $\sigma(X)=\sigma_X=\sqrt{\mathrm{Var}(X)}$ 称为 X 的**标准差**.

方差是由分布决定的,它是分布的散布特征,方差愈大,分布愈分散;方差愈小,分布愈集中.标准差与方差的功能相似,只是量纲不同.

2. 方差的性质 以下所涉及的方差均假定其存在.

(1) $\mathrm{Var}(X)=E(X^2)-[E(X)]^2$;

(2) 若 c 是常数,则 $\mathrm{Var}(c)=0$;

(3) 若 a,b 是常数,则 $\mathrm{Var}(aX+b)=a^2\mathrm{Var}(X)$;

(4) 若随机变量 X 的方差存在,则 $\mathrm{Var}(X)=0$ 的充要条件是 X 几乎处处为某个常数 a,即 $P(X=a)=1$.

3. 切比雪夫不等式 设 X 的数学期望和方差都存在,则对任意常数 $\varepsilon>0$,有

$$P\left(|X-E(X)|\geqslant\varepsilon\right)\leqslant\frac{\mathrm{Var}(X)}{\varepsilon^2},$$

或

$$P\left(|X-E(X)|<\varepsilon\right)\geqslant1-\frac{\mathrm{Var}(X)}{\varepsilon^2}.$$

切比雪夫不等式给出随机变量取值的大偏差(指事件$\{|X-E(X)|\geqslant\varepsilon\}$)发生概率的上限,该上限与分布的方差成正比.

4. 随机变量的标准化 对任意随机变量 X,如果 X 的数学期望和方差存在,则称

$$X^*=\frac{X-E(X)}{\sqrt{\mathrm{Var}(X)}}$$

为 X 的**标准化随机变量**,此时有 $E(X^*)=0,\mathrm{Var}(X^*)=1$.

§2.4 常用离散分布

1. 二项分布

(1) 若 X 的概率分布列为

$$P(X=k)=\binom{n}{k}p^k(1-p)^{n-k},\qquad k=0,1,\cdots,n.$$

则称 X 服从二项分布,记为 $X\sim b(n,p)$,其中 $0<p<1$.

(2) 背景:n 重伯努利试验中成功的次数 X 服从二项分布 $b(n,p)$,其中 p 为一次伯努利试验中成功的概率.

(3) $n=1$ 时的二项分布 $b(1,p)$ 称为**二点分布**,或称 **0-1 分布**.因为当 $X\sim b(1,p)$ 时,X 可表示一次伯努利试验中成功的次数,它只能取值 0 与 1.

(4) 二项分布 $b(n,p)$ 的数学期望和方差分别是

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

(5) 若 $X\sim b(n,p)$,则 $Y=n-X\sim b(n,1-p)$,其中 $Y=n-X$ 是 n 重伯努利试验中失败的次数.

2. 泊松分布

(1) 若 X 的概率分布列为

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

则称 X 服从泊松分布,记为 $X\sim P(\lambda)$,其中参数 $\lambda>0$.

(2) 背景:与单位时间(或单位面积、单位产品等)上的计数过程相联系.

(3) 泊松分布 $P(\lambda)$ 的数学期望和方差分别是

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

(4) **二项分布的泊松近似(泊松定理)** 在 n 重伯努利试验中,记事件 A 在一次试验中发生的概率为 p_n(与试验次数 n 有关),如果当 $n\to+\infty$ 时,有 $np_n\to\lambda$,则

$$\lim_{n\to+\infty}\binom{n}{k}p_n^k(1-p_n)^{n-k}=\frac{\lambda^k}{k!}\mathrm{e}^{-\lambda}.$$

3. 超几何分布

（1）若 X 的概率分布列为

$$P(X=k)=\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}},\qquad k=0,1,\cdots,r.$$

则称 X 服从超几何分布，记为 $X\sim h(n,N,M)$，其中 $r=\min\{M,n\}$，且 $M\leqslant N,n\leqslant N.n,N,M$ 均为正整数.

（2）背景：设有 N 个产品，其中有 M 个不合格品.若从中不放回地随机抽取 n 个，则其中含有的不合格品的个数 X 服从超几何分布 $h(n,N,M)$.

（3）超几何分布 $h(n,N,M)$ 的数学期望和方差分别是

$$E(X)=n\frac{M}{N},\qquad \operatorname{Var}(X)=\frac{nM(N-M)(N-n)}{N^2(N-1)}.$$

（4）**超几何分布的二项近似**　当 $n\ll N$ 时，超几何分布 $h(n,N,M)$ 可用二项分布 $b(n,M/N)$ 近似，即

$$\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}\approx\binom{n}{k}p^k(1-p)^{n-k},\qquad 其中\ p=\frac{M}{N}.$$

（5）实际应用中，在不返回抽样时，常用超几何分布描述抽出样品中不合格品数的分布；在返回抽样时，常用二项分布 $b(n,p)$ 描述抽出样品中不合格品数的分布；当批量 N 较大，而抽出样品数 n 较小时，不返回抽样可近似看作返回抽样.

4. 几何分布

（1）若 X 的概率分布列为

$$P(X=k)=(1-p)^{k-1}p,\qquad k=1,2,\cdots,$$

则称 X 服从几何分布，记为 $X\sim Ge(p)$，其中 $0<p<1$.

（2）背景：在伯努利试验序列中，成功事件 A 首次出现时的试验次数 X 服从几何分布 $Ge(p)$，其中 p 为每次试验中事件 A 发生的概率.

(3) 几何分布 $Ge(p)$ 的数学期望和方差分别是

$$E(X)=\frac{1}{p},\qquad \mathrm{Var}(X)=\frac{1-p}{p^2}.$$

(4) 几何分布的无记忆性：若 $X\sim Ge(p)$，则对任意正整数 m 与 n 有

$$P(X>m+n\mid X>m)=P(X>n).$$

5. 负二项分布

(1) 若 X 的概率分布列为

$$P(X=k)=\binom{k-1}{r-1}p^r(1-p)^{k-r},\qquad k=r,r+1,\cdots.$$

则称 X 服从负二项分布或**帕斯卡分布**，记为 $X\sim Nb(r,p)$，其中 r 为正整数，$0<p<1$.

(2) 背景：在伯努利试验序列中，成功事件 A 第 r 次出现时的试验次数 X 服从负二项分布 $Nb(r,p)$，其中 p 为每次试验中事件 A 发生的概率.

(3) $r=1$ 时的负二项分布为几何分布，即 $Nb(1,p)=Ge(p)$.

(4) 负二项分布 $Nb(r,p)$ 的数学期望和方差分别是

$$E(X)=r/p,\qquad \mathrm{Var}(X)=r(1-p)/p^2.$$

(5) 负二项分布的随机变量可以表示成 r 个独立同分布的几何分布随机变量之和，即若 $X\sim Nb(r,p)$，则 $X=X_1+X_2+\cdots+X_r$，其中 $X_1,X_2,\cdots,X_r$ 是相互独立、服从几何分布 $Ge(p)$ 的随机变量. 用图形表示如图 2.1，图中"1"表示 A，"0"表示 $\overline{A}$：

$$\underbrace{\overbrace{0\ 0\ \cdots\ 1}^{X_1}\ \overbrace{0\ 0\ \cdots\ 1}^{X_2}\ \cdots\ \overbrace{0\ 0\ \cdots\ 1}^{X_r}}_{X}$$

图 2.1

这里随机变量间的相互独立是指一个变量的取值不影响其他变量的取值，详见 §3.2.

6. 常用离散分布表

名称与记号	分布列	期望	方差
0-1分布 $b(1,p)$	$p_k=p^k(1-p)^{1-k},\quad k=0,1$	p	$p(1-p)$
二项分布 $b(n,p)$	$p_k=\binom{n}{k}p^k(1-p)^{n-k},k=0,1,\cdots,n$	np	$np(1-p)$
泊松分布 $P(\lambda)$	$p_k=\dfrac{\lambda^k}{k!}e^{-\lambda},k=0,1,\cdots$	λ	λ
超几何分布 $h(n,N,M)$	$p_k=\dfrac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}},\quad k=0,1,\cdots,r,\ r=\min\{M,n\}$	$n\dfrac{M}{N}$	$\dfrac{nM(N-M)(N-n)}{N^2(N-1)}$
几何分布 $Ge(p)$	$p_k=(1-p)^{k-1}p,k=1,2,\cdots$	$\dfrac{1}{p}$	$\dfrac{1-p}{p^2}$
负二项分布 $Nb(r,p)$	$p_k=\binom{k-1}{r-1}(1-p)^{k-r}p^r,k=r,r+1,\cdots$	$\dfrac{r}{p}$	$\dfrac{r(1-p)}{p^2}$

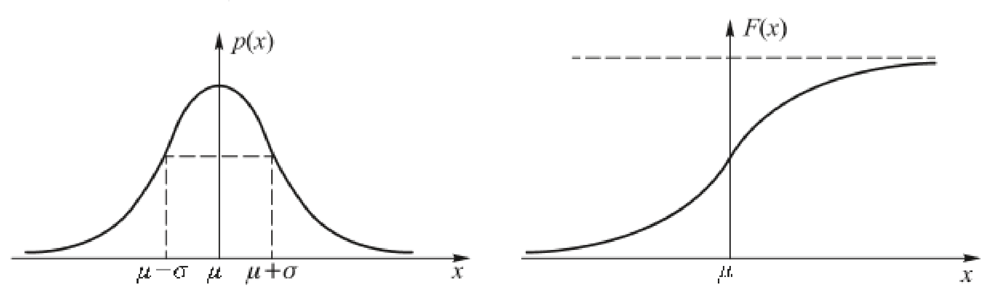
§2.5 常用连续分布

1. 正态分布

(1) 若 X 的密度函数和分布函数(如图 2.2)分别为

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

$$F(x)=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{x}e^{-\frac{(t-\mu)^2}{2\sigma^2}}dt,\qquad -\infty<x<+\infty,$$



(a) 密度函数 $p(x)$　(b) 分布函数 $F(x)$

图 2.2

则称 X 服从正态分布,记作 $X\sim N(\mu,\sigma^2)$,其中参数 $-\infty<\mu<+\infty$,$\sigma>0$.

(2) 背景:一个变量若是由大量微小的、独立的随机因素的叠加的结果,则此变量一定是正态变量(服从正态分布的变量).测量误差就是由量具偏差、测量环境的影响、测量技术的影响、测量人员的心理影响等随机因素叠加而成的,所以测量误差常认为服从正态分布.

(3) 关于参数 μ:

• μ 是正态分布的数学期望,即 $E(X)=\mu$,称 μ 为正态分布的位置参数.

• μ 是正态分布的对称中心,在 μ 的左侧和 $p(x)$ 下的面积为 0.5;在 μ 的右侧和 $p(x)$ 下的面积也为 0.5,所以 μ 也是正态分布的中位数(见 §2.7).

• 若 $X\sim N(\mu,\sigma^2)$,则 X 离 μ 愈近取值的可能性愈大,离 μ 愈远取值的可能性愈小.

关于参数 σ:

• σ^2 是正态分布的方差,即 $\mathrm{Var}(X)=\sigma^2$.

• σ 是正态分布的标准差,σ 愈小,正态分布愈集中;σ 愈大,正态分布愈分散.σ 又称为正态分布的尺度参数.

• 若 $X\sim N(\mu,\sigma^2)$,则其密度函数 $p(x)$ 在 $\mu\pm\sigma$ 处有两个拐点.

(4) 称 $\mu=0,\sigma=1$ 时的正态分布 $N(0,1)$ 为**标准正态分布**.记 U 为标准正态变量,$\varphi(u)$ 和 $\Phi(u)$ 为标准正态分布的密度函数和分布函数.$\varphi(u)$ 和 $\Phi(u)$ 满足:

• $\varphi(-u)=\varphi(u)$;

• $\Phi(-u)=1-\Phi(u)$.对 $u>0$,$\Phi(u)$ 的值有表可查.

(5) 标准化变换:若 $X\sim N(\mu,\sigma^2)$,则 $U=(X-\mu)/\sigma\sim N(0,1)$,其中 $U=(X-\mu)/\sigma$ 称为 X 的标准化变换.

(6) 若 $X\sim N(\mu,\sigma^2)$,则对任意实数 a 与 b,有

$$P(X\leqslant b)=\Phi\left(\frac{b-\mu}{\sigma}\right),$$

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

$$P(a<X\leqslant b)=\Phi\left(\frac{b-\mu}{\sigma}\right)-\Phi\left(\frac{a-\mu}{\sigma}\right),$$

可见,涉及正态变量的概率计算,一般是化为标准正态变量再通过查表获得.

(7) 正态分布的 3σ 原则:设 $X\sim N(\mu,\sigma^2)$,则

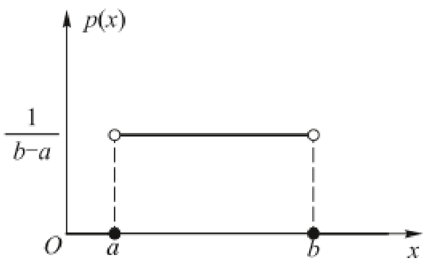
$$P(|X-\mu|<k\sigma)=\Phi(k)-\Phi(-k)=\begin{cases}0.682\,6, & k=1,\\ 0.954\,5, & k=2,\\ 0.997\,3, & k=3.\end{cases}$$

2. 均匀分布

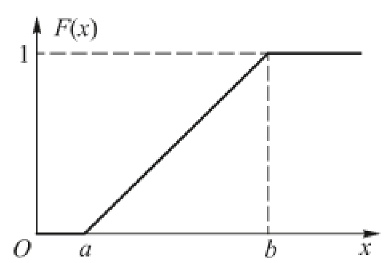
(1) 若 X 的密度函数和分布函数(如图 2.3)分别为

$$p(x)=\begin{cases}\dfrac{1}{b-a}, & a<x<b,\\ 0, & \text{其他},\end{cases}\qquad F(x)=\begin{cases}0, & x<a,\\ \dfrac{x-a}{b-a}, & a\leqslant x<b,\\ 1, & x\geqslant b.\end{cases}$$

则称 X 服从区间 (a,b) 上的均匀分布,记作 $X\sim U(a,b)$.



(a) 密度函数 $p(x)$



(b) 分布函数 $F(x)$

图 2.3

(2) 背景:向区间 (a,b) 随机投点,落点坐标 X 一定服从均匀分布 $U(a,b)$.这里“随机投点”是指:点落在任意相等长度的小区间上的可能性是相等的.

(3) 均匀分布 $U(a,b)$ 的数学期望和方差分别是

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

(4) 称区间 $(0,1)$ 上的均匀分布 $U(0,1)$ 为标准均匀分布,它是导出其他分布随机数的桥梁.

3. 指数分布

(1) 若 X 的密度函数(如图 2.4)和分布函数分别为

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

则称 X 服从指数分布，记作 $X \sim Exp(\lambda)$，其中参数 $\lambda>0$.

(2) 背景：若一个元器件(或一台设备、或一个系统)遇到外来冲击时即告失效，则首次冲击来到的时间 X(寿命)服从指数分布.很多产品的寿命可认为服从或近似服从指数分布.

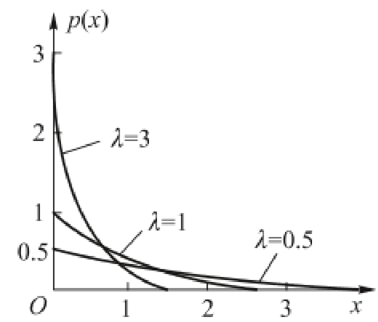


图 2.4

(3) 指数分布 $Exp(\lambda)$ 的数学期望和方差分别为

$$E(X)=\frac{1}{\lambda}, \qquad \operatorname{Var}(X)=\frac{1}{\lambda^{2}}.$$

(4) 指数分布的无记忆性： 若 $X \sim Exp(\lambda)$，则对任意 $s>0, t>0$，有

$$P(X>s+t \mid X>s)=P(X>t).$$

4. 伽马分布

(1) **伽马函数** 称 $\Gamma(\alpha)=\int_{0}^{+\infty} x^{\alpha-1} \mathrm{e}^{-x} \mathrm{d} x$ 为伽马函数，其中参数 $\alpha>0$. 伽马函数具有如下性质：

(a) $\Gamma(1)=1$;

(b) $\Gamma(1/2)=\sqrt{\pi}$;

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

(d) $\Gamma(n+1)=n\Gamma(n)=n!$ (n 为自然数).

(2) **伽马分布** 若 X 的密度函数(如图 2.5)为

$$p(x)=\begin{cases}\dfrac{\lambda^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} \mathrm{e}^{-\lambda x}, & x \geqslant 0, \\ 0, & x<0,\end{cases}$$

则称 X 服从伽马分布，记作 $X \sim Ga(\alpha, \lambda)$，其中 $\alpha>0$ 为形状参数，$\lambda>0$ 为尺度参数.

(3) 背景：若一个元器件(或一台设备、或一个系统)能抵挡一些外来冲击，但遇到第 k 次冲击时即告失效，则第 k 次冲击来到的时间 X(寿命)服从形状参数为 k 的伽马分布 $Ga(k, \lambda)$.

(4) 伽马分布 $Ga(\alpha, \lambda)$ 的数学期望和方差分别为

$$E(X)=\frac{\alpha}{\lambda}, \qquad \operatorname{Var}(X)=\frac{\alpha}{\lambda^{2}}.$$

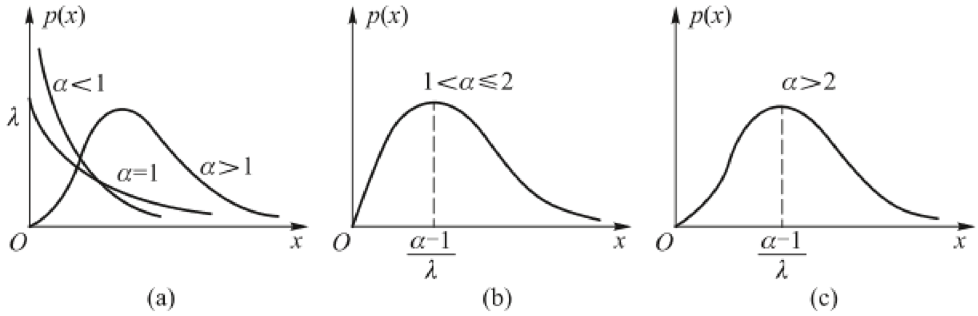


图 2.5

（5）伽马分布的两个特例：

（a）$\alpha=1$ 时的伽马分布就是指数分布，即 $Ga(1,\lambda)=Exp(\lambda)$.

（b）称 $\alpha=n/2,\lambda=1/2$ 时的伽马分布为自由度为 n 的 **χ^2（卡方）分布**，记为 $\chi^2(n)$，其密度函数为

$$p(x)=\begin{cases}\dfrac{1}{2^{\frac{n}{2}}\Gamma\left(\dfrac{n}{2}\right)}\mathrm{e}^{-\frac{x}{2}}x^{\frac{n}{2}-1}, & x>0,\\ 0, & x\leqslant 0.\end{cases}$$

$\chi^2(n)$ 分布的期望和方差分别为　$E(X)=n$，$\mathrm{Var}(X)=2n$.

（6）若形状参数为整数 k，则伽马变量可以表示成 k 个独立同分布的指数变量之和，即若 $X\sim Ga(k,\lambda)$，则 $X=X_1+X_2+\cdots+X_k$，其中 $X_1,X_2,\cdots,X_k$ 是相互独立且都服从指数分布 $Exp(\lambda)$ 的随机变量，见图 2.6，其中“×”表示冲击来到的时间.

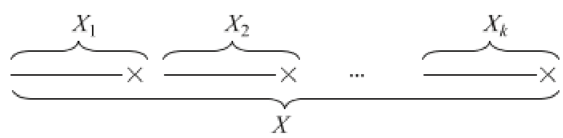


图 2.6

5. 贝塔分布

（1）**贝塔函数**　称 $\mathrm{B}(a,b)=\int_0^1 x^{a-1}(1-x)^{b-1}\mathrm{d}x$ 为贝塔函数，其中参数 $a>0,b>0$. 贝塔函数具有如下性质：

（a）$\mathrm{B}(a,b)=\mathrm{B}(b,a)$；　　（b）$\mathrm{B}(a,b)=\dfrac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$.

（2）**贝塔分布**　若 X 的密度函数（如图 2.7）为

$$p(x)=\begin{cases}\dfrac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}x^{a-1}(1-x)^{b-1}, & 0<x<1,\\ 0, & \text{其他},\end{cases}$$

则称 X 服从贝塔分布,记作 $X\sim Be(a,b)$,其中 $a>0,b>0$ 都是形状参数.

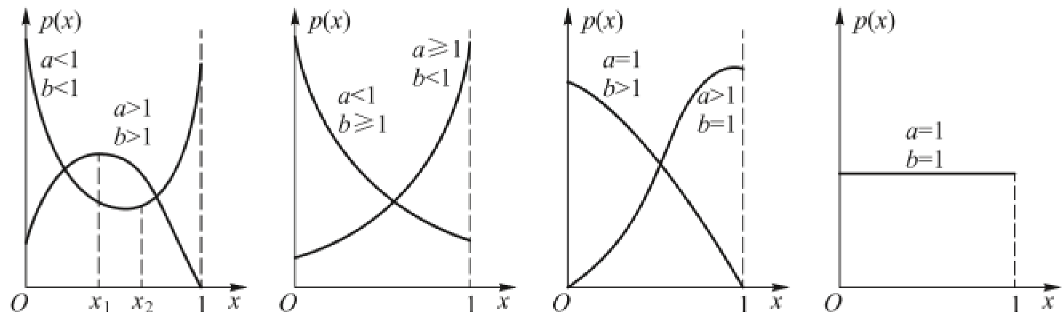


图 2.7

(3) 背景:很多比率,如产品的不合格品率、机器的维修率、某商品的市场占有率、射击的命中率等都是在区间(0,1)上取值的随机变量,贝塔分布 $Be(a,b)$ 可供描述这些随机变量之用,而在应用中,可调节 a 与 b 以适应实际中的要求.

(4) 贝塔分布 $Be(a,b)$ 的数学期望和方差分别为

$$E(X)=\frac{a}{a+b},\qquad \operatorname{Var}(X)=\frac{ab}{(a+b)^2(a+b+1)}.$$

(5) $a=b=1$ 时的贝塔分布就是区间(0,1)上的均匀分布,即 $Be(1,1)=U(0,1)$.

6. 常用连续分布表

名称与记号	密度函数 $p(x)$	期望	方差
正态分布 $N(\mu,\sigma^2)$	$\frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\},-\infty<x<+\infty$	μ	σ^2
均匀分布 $U(a,b)$	$\frac{1}{b-a},a<x<b$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
指数分布 $Exp(\lambda)$	$\lambda e^{-\lambda x},x\geqslant 0$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$

续表

名称与记号	密度函数 $p(x)$	期望	方差
伽马分布 $Ga(\alpha,\lambda)$	$\frac{\lambda^{\alpha}}{\Gamma(\alpha)}x^{\alpha-1}e^{-\lambda x},x\geqslant 0$	$\frac{\alpha}{\lambda}$	$\frac{\alpha}{\lambda^2}$
$\chi^2(n)$分布	$\frac{x^{n/2-1}e^{-x/2}}{\Gamma(n/2)2^{n/2}},x\geqslant 0$	n	$2n$
贝塔分布 $Be(a,b)$	$\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}x^{a-1}(1-x)^{b-1},0<x<1$	$\frac{a}{a+b}$	$\frac{ab}{(a+b)^2(a+b+1)}$
对数正态分布 $LN(\mu,\sigma^2)$	$\frac{1}{\sqrt{2\pi}\sigma x}\exp\left\{-\frac{(\ln x-\mu)^2}{2\sigma^2}\right\},x>0$	$e^{\mu+\sigma^2/2}$	$e^{2\mu+\sigma^2}(e^{\sigma^2}-1)$
柯西分布 $Cau(\mu,\lambda)$	$\frac{1}{\pi}\frac{\lambda}{\lambda^2+(x-\mu)^2},-\infty<x<+\infty$	不存在	不存在
韦布尔分布 $Wei(m,\eta)$	$F(x)=1-\exp\left\{-\left(\frac{x}{\eta}\right)^m\right\},x>0$	$\eta\Gamma\left(1+\frac{1}{m}\right)$	$\eta^2\left[\Gamma\left(1+\frac{2}{m}\right)-\Gamma^2\left(1+\frac{1}{m}\right)\right]$

§2.6 随机变量函数的分布

1. 设连续随机变量 X 的密度函数为 $p_X(x)$，$Y=g(X)$.

（1）若 $y=g(x)$ 严格单调，其反函数 $h(y)$ 有连续导函数，则 $Y=g(X)$ 的密度函数为

$$p_Y(y)=\begin{cases}p_X[h(y)]\,|h'(y)|, & a<y<b,\\ 0, & \text{其他},\end{cases}$$

其中 $a=\min\{g(-\infty),g(+\infty)\}$，$b=\max\{g(-\infty),g(+\infty)\}$.

（2）若 $y=g(x)$ 在不相重叠的区间 $I_1,I_2,\cdots$ 上逐段严格单调，其反函数 $h_1(y),h_2(y),\cdots$ 有连续导函数，则 $Y=g(X)$ 的密度函数为

$$p_Y(y)=\sum_i p_X(h_i(y))\,|h_i'(y)|.$$

2. 正态变量的线性变换仍为正态变量 若 X 服从正态分布 $N(\mu,$

$\sigma^2)$，则当 $a \neq 0$ 时，有 $Y=aX+b \sim N(a\mu+b, a^2\sigma^2)$.

3. 对数正态分布

（1）若 X 的密度函数为

$$p_X(x)=\begin{cases}\dfrac{1}{\sqrt{2\pi}x\sigma}\exp\left\{-\dfrac{(\ln x-\mu)^2}{2\sigma^2}\right\}, & x>0, \\ 0, & x \leqslant 0.\end{cases}$$

则称 X 服从对数正态分布，记为 $X \sim LN(\mu,\sigma^2)$，其中 $-\infty<\mu<+\infty$，$\sigma>0$. 对数正态分布的密度函数如图 2.8.

（2）若 $X \sim LN(\mu,\sigma^2)$，则 $E(X)=e^{\mu+\sigma^2/2}$，$\mathrm{Var}(X)=(e^{\sigma^2}-1)\cdot e^{2\mu+\sigma^2}$.

（3）若 $X \sim LN(\mu,\sigma^2)$，则 $Y=\ln X \sim N(\mu, \sigma^2)$.

图 2.8

4. 若 $X \sim Ga(\alpha,\lambda)$，则当 $k>0$ 时，有 $Y=kX \sim Ga(\alpha,\lambda/k)$.

5. 若 X 的分布函数 $F_X(x)$ 为严格单调增的连续函数，其反函数 $F_X^{-1}(y)$ 存在，则 $Y=F_X(X)$ 服从 $(0,1)$ 上的均匀分布 $U(0,1)$.

§2.7　分布的其他特征数

1. k 阶矩

（1）称 $\mu_k=E(X^k)$ 为 X 的 k 阶**原点矩**. 一阶原点矩就是数学期望；

（2）称 $\nu_k=E(X-E(X))^k$ 为 X 的 k 阶**中心矩**. 二阶中心矩就是方差；

（3）前 k 阶中心矩可用原点矩表示，如

$$\nu_1=0;$$

$$\nu_2=\mu_2-\mu_1^2;$$

$$\nu_3=\mu_3-3\mu_2\mu_1+2\mu_1^3;$$

$$\nu_4=\mu_4-4\mu_3\mu_1+6\mu_2\mu_1^2-3\mu_1^4.$$

2. 变异系数　称比值 $C_v(X)=\dfrac{\sqrt{\mathrm{Var}(X)}}{E(X)}$ 为 X 的变异系数. 变异系数是

一个无量纲的量.

3. 分位数 设连续随机变量 X 的分布函数为 $F(x)$,密度函数为$p(x)$.对任意 $p\in(0,1)$,

(1) 称满足条件

$$F(x_p)=\int_{-\infty}^{x_p}p(x)\,\mathrm{d}x=p$$

的 x_p 为此分布的 p **分位数**,又称**下侧** p **分位数**,它把密度函数下的面积一分为二,左侧面积恰好为 p;

(2) 称满足条件

$$1-F(x'_p)=\int_{x'_p}^{+\infty}p(x)\,\mathrm{d}x=p$$

的 x'_p 为此分布的**上侧** p **分位数**;

(3) 分位数与上侧分位数的转换公式: $x'_p=x_{1-p}$; $x_p=x'_{1-p}$;

(4) 称 $p=0.5$ 时的 p 分位数 $x_{0.5}$为此分布的**中位数**,即 $x_{0.5}$满足

$$F(x_{0.5})=\int_{-\infty}^{x_{0.5}}p(x)\,\mathrm{d}x=0.5;$$

(5) 若随机变量 X 的密度函数 $p(x)$是偶函数,则此分布的 p 分位数 x_p 满足 $x_p=-x_{1-p}$.中位数为分布对称中心;

(6) 记标准正态分布的 p 分位数为 u_p.因为标准正态密度函数是偶函数,所以 $u_p=-u_{1-p}$;

(7) 一般正态分布 $N(\mu,\sigma^2)$的 p 分位数 x_p 满足 $x_p=\mu+\sigma u_p$;

(8) 分布的矩有可能不存在,但连续分布的分位数总存在.p 分位数 x_p 总是 p 的增函数.

4. 偏度系数

(1) 称比值

$$\beta_s=\frac{E(X-E(X))^3}{[E(X-E(X))^2]^{3/2}}$$

为 X 的分布的偏度系数,简称**偏度**;

(2) 偏度系数刻画的是分布的不对称程度,$|\beta_s|$愈大,分布的对称性愈差;

（3）任一对称分布的偏度 $\beta_s=0$. 当 $\beta_s>0$ 时，分布为正偏（又称右偏）；当$\beta_s<0$时，分布为负偏（又称左偏）.

5. 峰度系数

（1）称

$$\beta_k=\frac{E(X-E(X))^4}{[E(X-E(X))^2]^2}-3$$

为 X 的分布的峰度系数，简称**峰度**；

（2）峰度系数是刻画分布的尖峭性和尾部粗细的一个特征数；

（3）任一正态分布的峰度 $\beta_k=0$. 当 $\beta_k<0$ 时，分布比正态分布平坦；当$\beta_k>0$时，分布比正态分布更尖峭；

6. 偏度与峰度都是描述分布（密度）形状的参数.