第三章 随机变量与分布函数

§3.1 随机变量及其分布

一、随机变量及分布函数的定义

二、分布函数的性质

三、离散型随机变量

四、连续型随机变量

一、随机变量的定义

在随机试验中,有些试验结果本身与数值有关(本身就是一个数).

例如，掷一颗骰子面上出现的点数；

每天进入地铁站的人数；

昆虫的产卵数；

六月份南京市的最高温度；

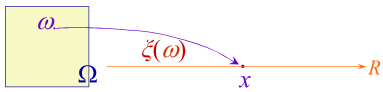
在有些试验中，试验结果看来与数值无关，但我们可以引进一个变量来表示它的各种结果. 也就是说，把试验结果数值化.

例如，在抛硬币的试验中，我们可以如下方法将试验结果与数值联系起来，当出现正面时对应数“1”，当出现反面时对应数“0”.

一般地，对任意的随机事件 A，都可以引进一个函数($\Omega \to \{0,1\}$)来表示 A 是否发生:

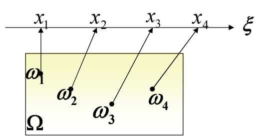
$$I_A = \begin{cases} 1, & \text{如果 } A \text{ 发生} \\ 0, & \text{如果 } A \text{ 不发生} \end{cases}$$

这种对应关系在数学上理解为定义了一种单值实值函数.



◆ 随机变量(random variable, r.v.)的概念

若对于随机试验 E 的每一个可能结果 $\omega \in \Omega$，都有唯一的一个实数值 $\xi(\omega)$ 相对应，则称实值函数 $\xi(\omega)$ 为随机变量，简记为 ξ.



随机变量通常用大写字母 $X, Y, Z, W, N, \cdots$，或希腊字母 $\xi, \eta, \zeta, \cdots$，等表示

而表示随机变量所取的值时，一般采用小写字母 x, y, z, w, n 等.

如何把握一个随机变量ξ？

❶ ξ 的取值情况;　❷ 它取值的概率的分布情况.

(1) 随机变量ξ是定义在样本空间上的实值函数，

它的取值与试验结果形成对应，

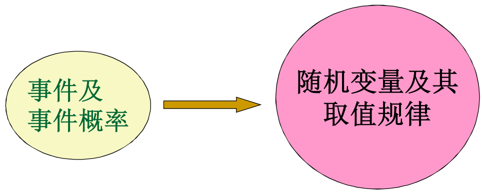
随着实验结果的不同而取不同的值，

在试验之前只知道ξ可能取值的范围，而不能预先肯定它将取哪个值.

(2) 由于试验结果的出现具有一定的概率，所以随机变量取每个值和每个确定范围内的值也有一定的概率.

随机变量的取值既具有可变性，也有随机性。这种双重性正是随机变量与普通变量(函数)的本质区别。

随机变量概念的产生是概率论发展史上的重大事件. 引入随机变量后，随机试验中的任一随机事件就可以通过随机变量的取值关系式表达出来，对随机现象统计规律的研究，就由对事件及事件概率的研究扩大为对随机变量及其取值规律的研究.



Example:

1. 生化检验结果分阳性和阴性，

我们用变量X表示，X=0表阴性；X＝1表阳性。

2. 抛一枚硬币，结果分为“正面”、“反面”

X=1表正面；X＝0表反面。

---

3 抛掷骰子,观察出现的点数.

则有　Ω={1，2，3，4，5，6}

样本点本身就是数量

$\xi(\omega)=\omega$ 恒等变换

$\xi(1)=1, \xi(2)=2, \xi(3)=3, \quad \xi(4)=4, \xi(5)=5, \xi(6)=6,$

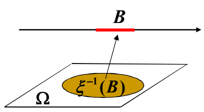
且有：　$P\{\xi=i\}=\frac{1}{6}, \quad (i=1,2,3,4,5,6).$

下面给出随机变量的严格数学定义.

定义 3.1.1 设ξ(ω) 是定义于概率空间(Ω, $\mathcal{F}$, P)上的单值实函数，如果对于直线上任一Borel点集B，有

$$\{\omega:\xi(\omega)\in B\}\in\mathcal{F},$$

则称ξ(ω)为随机变量，而 P(ξ(ω)∈B) 称为随机变量ξ(ω)的概率分布.



例3.1.3　对于Ω ={$\omega_1,\omega_2,\omega_3$}，构造σ代数

$\mathcal{F}$={Ω, Φ,{ω_1,ω_2},{ω_3}}，再定义概率：

P(Ω)=1, P(Φ)=0, P({ω_1,ω_2})=2/3,P({ω_3})=1/3.

这里(Ω,$\mathcal{F}$, P)显然是一个概率空间，但实值函数 ξ(ω_k) =k, 1≤k≤3 不是随机变量，例如取Borel点集 B=(0, 1.5)，则因为有

$$\{\omega: \xi(\omega)\in B\}=\{\omega_1\}\notin\mathcal{F}.$$

特别地，若取 $B=(-\infty,x)$，则有 $\{\omega:\xi(\omega)<x\}\in\mathcal{F}$.

另外，注意到，

$$P\{a\le\xi(\omega)<b\}=P\{\xi(\omega)<b\}-P\{\xi(\omega)<a\}.$$

所以只要对一切实数 x 给出概率 $P\{\xi(\omega)<x\}$，就能算出 $\xi(\omega)$ 落入某个区间的概率，再利用概率的性质还可以算出 $\xi(\omega)$ 属于某些相当复杂的直线点集的概率。

定义 3.1.2 称

$$F(x)=P\{\xi(\omega)<x\},-\infty<x<+\infty$$

为随机变量 ξ(ω)的(累积)分布函数(*cumulative distribution function, c.d.f.*)。记作

$$\xi(\omega)\sim F(x).$$

二、分布函数的性质

(1) 单调性：若 $a<b$，则 $F(a)\le F(b)$；

(2) 有界性： $0\le F(x)\le 1$，且

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

(3) 左连续性：$F(x-)=\lim_{t\to x-}F(t)=F(x)$.

注：可证：满足上述三条性质的函数必为某个随机变量的分布函数。因此，以后我们把满足这三条性质的函数称为分布函数。

如果分布函数定义中的“ $<$ ”改为“ $\le$ ”，则为右连续.

证明：(1). $F(b)-F(a)=P\{a\le\xi<b\}\ge 0$

(2) $1=P\{-\infty<\xi<+\infty\}=P\{\sum_{n=-\infty}^{+\infty}(n\le\xi<n+1)\}=\sum_{n=-\infty}^{+\infty}P\{n\le\xi<n+1\}$

$$=\sum_{n=-\infty}^{+\infty}[F(n+1)-F(n)]=\lim_{n\to+\infty}F(n)-\lim_{m\to-\infty}F(m).$$

由于 $F(x)$ 的单调有界性，

$$\lim_{x\to-\infty}F(x)=\lim_{m\to-\infty}F(m),\ \lim_{x\to+\infty}F(x)=\lim_{n\to+\infty}F(n)\text{存在。}$$

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

（3）由于 $F(x)$ 是单调函数，只须证明对于任意一列单调上升的数列 $x_n\uparrow x$，成立 $\lim_{n\to\infty}F(x_n)=F(x)$ 即可。

根据概率的可列可加性可知

$$F(x)-F(x_0)=P\{x_0\le\xi<x\}=P\{\sum_{n=1}^{\infty}(x_{n-1}\le\xi<x_n)\}$$

$$=\sum_{n=1}^{\infty}[F(x_n)-F(x_{n-1})]=\lim_{n\to\infty}F(x_n)-F(x_0)$$

因而，$F(x-)=\lim_{n\to\infty}F(x_n)=F(x)$.

另证：利用概率的下连续性

$$F(x)=P\{\omega:\xi(\omega)<x\}=P\{\bigcup_{n=1}^{\infty}(\omega:\xi(\omega)<x_n)\}$$

$$=\lim_{n\to\infty}P\{\omega:\xi(\omega)<x_n\}=\lim_{n\to\infty}F(x_n)=F(x-).$$

有了分布函数，关于随机变量的许多概率都能方便算出：

$$1.P\{\xi(\omega)\le a\}=P\{\bigcap_{n=1}^{+\infty}(\xi(\omega)<a+\frac{1}{n})\}$$

$$=\lim_{n\to+\infty}P\{\xi(\omega)<a+\frac{1}{n}\}=\lim_{n\to+\infty}F(a+\frac{1}{n})=F(a+).$$

$$2.P\{\xi(\omega)=a\}=P\{\xi(\omega)\le a\}-P\{\xi(\omega)<a\}=F(a+)-F(a).$$

$$3.P\{\xi(\omega)\ge a\}=1-P\{\xi(\omega)<a\}=1-F(a).$$

$$4.P\{\xi(\omega)>a\}=P\{\xi(\omega)\ge a\}-P\{\xi(\omega)=a\}=1-F(a+).$$

$$5.P\{a<\xi(\omega)\le b\}=P\{\xi(\omega)\le b\}-P\{\xi(\omega)\le a\}=F(b+)-F(a+).$$

……

两类重要随机变量：

- 离散型随机变量；
- 连续型随机变量。

三、离散型随机变量

定义1. 若随机变量 ξ 的全部可能取值是有限个或可列无限多个，则称 ξ 是离散型随机变量。

定义2. 设离散型随机变量 ξ 的所有可能取值为 x_k，其中 $k=1,2,\cdots$，事件 $\{\xi=x_k\}$ 的概率：$P\{\xi=x_k\}=p_k, k=1,2,\cdots$

称为 ξ 的概率分布或分布律。

性质：(1) $p_k \ge 0, k=1,2,\cdots$ (2) $\sum_{k=1}^{\infty} p_k = 1$

分布律也可用如下表格的形式表示：

ξ	x_1	x_2	$\cdots$	x_k	$\cdots$
p_k	p_1	p_2	$\cdots$	p_k	$\cdots$

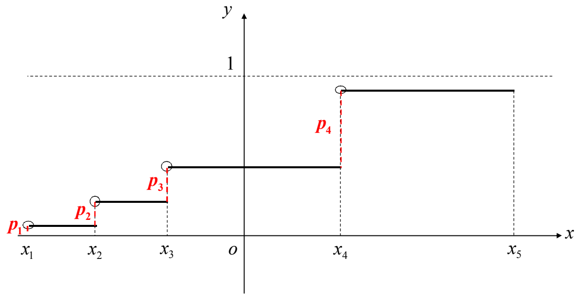
称之为随机变量的分布列.

有了分布列，可以通过下式求得分布函数

$$F(x) = P\{\xi < x\} = \sum p(x_k) \qquad (3.1.12)$$

显然这时 $F(x)$ 是一个跳跃函数，它在每个 x_k 处有跳跃度 $p(x_k)$. 当然，由分布函数 $F(x)$ 也可唯一决定 x_k 及 $p(x_k)$，因此用分布列或分布函数都能描述离散型随机变量.

离散型分布函数的图像



可见，ξ为离散型随机变量当且仅当ξ的分布函数$F(x)$为阶梯型函数. 进一步，ξ的取值点恰为$F(x)$的间断点，且

$$p_i = P\{X = x_i\} = F(x_i + 0) - F(x_i).$$

➢ 常见的离散型随机变量及其分布

1. 退化(degenerate)分布：单点分布

随机变量α只取常数值c，即 P{α=c}=1.

这时分布函数为

$$F(x) = I_c(x) = \begin{cases} 0, & x \le c, \\ 1, & x > c. \end{cases}$$

2. 伯努利(Bernoulli)分布

若以β记一次伯努利试验中事件A出现的次数，则

$$P\{\beta = k\} = b(k;1,p) = p^k q^{1-k}, k = 0,1,$$

称为伯努利分布，亦称两点分布或(0－1)分布。

3. 二项(binomial)分布

若以μ记n重伯努利试验中事件A出现的次数，则

$$P\{\mu = k\} = b(k;n,p) = \binom{n}{k} p^k q^{n-k}, k = 0,1,\cdots,n$$

简记作$\mu \sim B(n,p)$.

伯努利分布可以看成为n=1的二项分布。

4. 超几何(hyper-geometric)分布

N件产品中共有M件次品，若以ν记从这批产品中（不放回地）抽取n件产品中出现的次品数，则

$$P\{\nu = k\} = h_k = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}, \quad \begin{matrix} 0 \le k \le n \le N \\ k \le M \end{matrix}$$

如果 n << N 以及 n << M，可用二项分布近似.

5. 泊松(Poisson)分布

若随机变量ξ可取一切非负整数值，且

$$P\{\xi = k\} = p(k,\lambda) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0,1,2,\cdots$$

其中λ>0，则称ξ服从泊松分布. 记为$\xi \sim P(\lambda)$.

6. 几何(geometric)分布

在事件A发生的概率为p 的伯努利试验中,若以 η 记A首次出现时的试验次数，则

$$P\{\eta=k\}=g(k;p)=q^{k-1}p,\quad k=1,2,\cdots$$

称为几何分布，它是一种等待时间分布.

➢ 几何分布的无记忆性

定理. 取值于自然数的随机变量η服从几何分布，当且仅当η有无记忆性:

$$P\{\eta>m+n\mid\eta>m\}=P\{\eta>n\},\quad \forall m,n\geq 1.$$

或者

$$P\{\eta=m+n\mid\eta>m\}=P\{\eta=n\},\quad \forall m,n\geq 1.$$

Remark: 几何分布是唯一无记忆性的离散型分布.

直观解释:

在伯努利试验中，等待首次成功的时间服从几何分布。无记忆性表明：已知试验了m次未获得成功，再加做n 次试验仍不成功的条件概率，等于从开始算起做n次试验都不成功的概率。就是说，已做过的m次失败的试验被忘记了。

产生几何分布这种无记忆性的根本原因在于，我们进行的是独立重复试验。

如果真的在做“独立重复试验”，那么不管已经失败过多少次，也不会为今后的试验留下可借鉴的东西。

证明:

【必要性】 因为η 服从几何分布，故

$$P\{\eta>n\}=\sum_{k=n+1}^{\infty}q^{k-1}p=\frac{pq^n}{1-q}=q^n,\quad n=1,2,\cdots$$

于是，对任意的 $m,n\geq 1$ 有

$$P\{\eta>m+n\mid\eta>m\}=\frac{P\{\eta>m+n\}}{P\{\eta>m\}}=\frac{q^{m+n}}{q^m}=q^n$$

因而 η 具有无记忆性。

【充分性】 设

$$Q_n=P\{\eta>n\}>0,\ \forall n\geq 1.$$

由乘法公式知:

$$P\{\eta>m+n\}=P\{\eta>m\}\cdot P\{\eta>m+n\mid\eta>m\},$$

因为 η 具有无记忆性，故有 $P(\eta>m+n)=P(\eta>m)P(\eta>n)$,

$$Q_{m+n}=Q_mQ_n,\ \forall m,n\geq 1.$$

因而， $Q_m=Q_1^{\,m}$.

令 $Q_1=q$, 则 $Q_m=q^m$, 且对任意的 $k\geq 1$ 有

$$P\{\eta=k\}=Q_{k-1}-Q_k=q^{k-1}-q^k=q^{k-1}p$$

至此得证 η 服从几何分布。 证毕.

7. 帕斯卡(Pascal)分布

在事件A发生的概率为p的伯努利试验中，若以 ζ 记 第r 次成功出现时的试验次数，则

$$P\{\zeta=k\}=\binom{k-1}{r-1}p^rq^{k-r},\quad k=r,r+1,\cdots$$

称为帕斯卡分布. 显然, 当 r=1时, 即为几何分布.

若以 η_i 记从第 $i-1$次成功之后的第一次试验算起至第i次成功为止共进行的试验次数，则 η_i 服从几何分布:

$$P\{\eta_i=k\}=q^{k-1}p,\quad k=1,2,3,\cdots.$$

η_i 相互独立，并且 $\zeta=\eta_1+\cdots+\eta_r$.

因而，参数为 r的帕斯卡分布可以分解为r个独立同分布的几何分布的随机变量的和。

另一方面，若记 $\tilde{\zeta}=\zeta-r$, 则 $\tilde{\zeta}$表示为等待第 r次成功所经历过的失败次数，那么

$$P\{\tilde{\zeta}=l\}=\binom{r+l-1}{l}p^r q^l=\binom{-r}{l}p^r(-q)^l, l=0,1,2,\cdots$$

注：不少书上把这个分布作为帕斯卡分布的定义。

显然，ζ 与 $\tilde{\zeta}$ 只是计数方式的不同，一个记全部试验数，另一个则只记失败的次数，它们描述的是同样的随机模型，因而它们本质上是相同的。

特别地，取 $r=1$，考察为等待第1次成功所经过的失败次数 $\tilde{\eta}$，那么，

$$P\{\tilde{\eta}=l\}=q^l p,\quad l=0,1,2,\cdots.$$

这也是一个分布，表示等待首次成功所经历过的失败数，也称为几何分布。

8、负二项分布

对于任意的实数 $r>0$，称

$$Nb(l;r,p)=\binom{-r}{l}p^r(-q)^l,\quad l=0,1,2,\cdots$$

为负二项分布.

四、连续型随机变量

定义. 设F(x)是随机变量 ξ 的分布函数，若存在非负可积函数p(x), 使得对任意实数x有

$$F(x)=\int_{-\infty}^{x}p(t)dt,$$

则称 ξ 为连续型随机变量，称 p(x)为ξ的概率密度函数，简称概率密度或密度函数(probability density function, pdf) .

Reamrk: 连续型r.v.的分布函数$F(x)$是绝对连续的, 且$F(x)$几乎处处可导, 即 $p(x)=F'(x)$, for $a.e.\ x\in R$.

概率密度的性质:

(1) 非负性: $p(x)\geq 0$,

(2) 规范性: $\int_{-\infty}^{+\infty}p(x)dx=1$,

注：概率密度函数的充要条件：性质（1）+（2）。

注意到, $P\{a\leq\xi<b\}=F(b)-F(a)=\int_a^b p(x)dx,$

故给定密度函数p(x), 便可以算出r.v. 落入某个区间的概率. 进一步可以证明，r.v. 落入任何Borel点集B的概率为:

$$P\{\xi\in B\}=\int_B p(x)dx.$$

特别地, 在p(x)的连续点处, ξ 落在小区间 $(x,x+\Delta x)$上的概率近似为: $p(x)\Delta x$.

另外, $P\{\xi=c\}=0, \forall c\in(-\infty,\infty).$

因为 $P\{\xi=c\}\leq P\{c\leq\xi<c+h\}=\int_c^{c+h}p(x)dx$

故 $0\leq P\{\xi=c\}\leq\lim\limits_{h\to 0}\int_c^{c+h}p(x)dx=0$

Remark:

- 一个事件的概率为零，它并不一定是不可能事件，一个事件的概率为一，它也不一定是必然事件。
- 在零测集上可以改变密度函数的值，而不影响分布函数的值。
- 若两个连续性随机变量的概率密度几乎处处相等，我们即可说二者同分布。
- 分布函数、概率密度函数，都是对随机变量的概率性质的完整刻划，它们含有相同的信息量，可以相互唯一确定（几乎处处意义下），但在图形上，密度函数对各种分布特征的显示要优胜得多，因此它比分布函数更常用。

例：设ξ的密度函数为

$$p(x)=\begin{cases} ax, & 0<x<1, \\ 2-x, & 1\le x<2, \\ 0 & \text{其它}, \end{cases}$$

(1)求常数a的值； (2) 写出ξ的分布函数；

解：(1) $1=\int_{-\infty}^{+\infty}p(x)dx=\int_{-\infty}^{0}p(x)dx+\int_{0}^{1}p(x)dx+\int_{1}^{2}p(x)dx+\int_{2}^{+\infty}p(x)dx$

$$=0+\frac{a}{2}+\frac{1}{2}+0=\frac{a+1}{2} \quad \Rightarrow a=1.$$

(2) $F(x)=P\{\xi<x\}$

$$=\begin{cases} 0, & x\le 0, \\ \int_0^x t\,dt, & 0<x<1, \\ \int_0^1 t\,dt+\int_1^x (2-t)dt, & 1\le x<2, \\ 1, & x\ge 2, \end{cases} = \begin{cases} 0, & x\le 0, \\ \frac{x^2}{2}, & 0<x<1, \\ 2x-\frac{x^2}{2}-1, & 1\le x<2, \\ 1, & x\ge 2. \end{cases}$$

◆ 常见的连续型随机变量及其分布

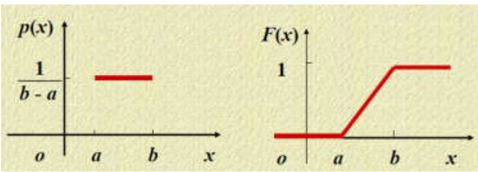
1. 均匀(uniform)分布

定义. 若连续型随机变量ξ的密度函数为:

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

则称ξ服从 $[a, b]$上的均匀分布，记作 $\xi\sim U[a, b]$.

分布函数为: $F(x)=\int_{-\infty}^{x}p(t)dt=\begin{cases} 0, & x\le a, \\ \frac{x-a}{b-a}, & a<x\le b, \\ 1, & x>b. \end{cases}$



服从区间$[a,b]$上均匀分布的随机变量ξ，具有下述意义的等可能性：对任意子区间$(c, c+l)$,

$$P\{c\le\xi<c+l\}=\int_c^{c+l}p(x)dx=\int_c^{c+l}\frac{1}{b-a}dx=\frac{l}{b-a},$$

即随机变量ξ在区间$[a,b]$上的任一子区间上取值的概率与该子区间的长度成正比，而与该子区间的位置无关.

2. 正态 (normal) 分布

定义. 若随机变量ξ的概率密度函数为

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

其中，μ 与σ为常数，且σ>0，则称ξ服从参数为(μ, σ)的正态分布或高斯分布，记作 $\xi\sim N(\mu, \sigma^2)$.

定义. 若μ=0, σ=1，称ξ服从标准正态分布，记作 $\xi\sim N(0, 1)$.

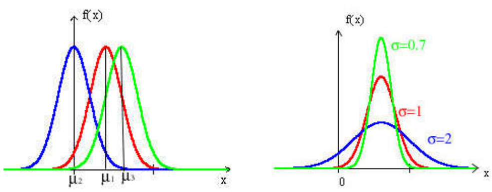
pdf : $\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$, cdf : $\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{t^2}{2}}dt.$

性 质： $\varphi(-x)=\varphi(x), \Phi(-x)=1-\Phi(x).$

特点：(1) $p(x)$关于$x=\mu$对称；

(2) $p(x)$在$x=\mu$处取最大值 $p_{\max}=\frac{1}{\sqrt{2\pi}\sigma}$;

(3) μ决定位置, σ 决定形状.



思考：怎样验证 $\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx=1$?

定理：若$X\sim N(\mu,\sigma^2)$，则$Y=\frac{X-\mu}{\sigma}\sim N(0,1)$.

证明：$F_Y(y)=P\{Y<y\}=P\{\frac{X-\mu}{\sigma}<y\}=P\{X<\mu+\sigma y\}$

$$=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\mu+\sigma y}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx \quad \text{作变量代换}\ \frac{x-\mu}{\sigma}=v$$

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

$$\therefore p_Y(y)=F_Y'(y)=\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}},-\infty<y<+\infty.$$

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

上述定理沟通了一般正态分布与标准状态分布的关系，从而一般正态分布的概率计算可以转化为标准状态分布的概率计算。事实上，若$X \sim N(\mu,\sigma^2)$，则

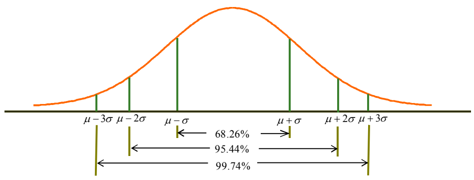
$$P\{a \le X < b\} = P\{\frac{a-\mu}{\sigma} \le \frac{X-\mu}{\sigma} < \frac{b-\mu}{\sigma}\}$$
$$=\Phi(\frac{b-\mu}{\sigma}) - \Phi(\frac{a-\mu}{\sigma}).$$

正态分布的3σ原则：

$$1.P\{|\xi-\mu|<\sigma\} = P\{-1 < \frac{\xi-\mu}{\sigma} < 1\}$$
$$=\Phi(1)-\Phi(-1) = 2\Phi(1)-1 = 0.6826.$$

$$2.P\{|\xi-\mu|<2\sigma\} = 2\Phi(2)-1 = 0.9544.$$

$$3.P\{|\xi-\mu|<3\sigma\} = 2\Phi(3)-1 = 0.9974.$$



◆ 正态分布是概率论中最重要的分布。

一方面，正态分布是自然界最常见的一种分布，例如测量的误差;炮弹弹落点的分布;人的生理特征的尺寸:身高，体重等；农作物的收获量；工厂产品的尺寸:直径，长度，宽度，高度等等都近似服从正态分布。

一般来说,若影响某一数量指标的随机因素很多，而每个因素所起的作用不太大,则这个指标服从正态分布。这一点可以利用概率论的极限定理来证明。

另一方面，正态分布具有许多良好的性质，许多分布可用正态分布来近似,另外一些分布又可以通过正态分布来导出,因此在理论研究中,正态分布十分重要。

下例说明正态分布与实测数据符合得很好.

［例 2］ 上海手表厂曾对其生产的某个零件的重量收集了大量资料,对测量得的 3 805 个数据,按不同重量加以分组,并记录了不同范围内零件的个数(频数),计算了它们的频率,结果如下表所示,它们与 $\mu = 56.94$, $\sigma = 8.2$ 的正态分布符合得相当好,图 3.1.7表示了两者符合的情况.

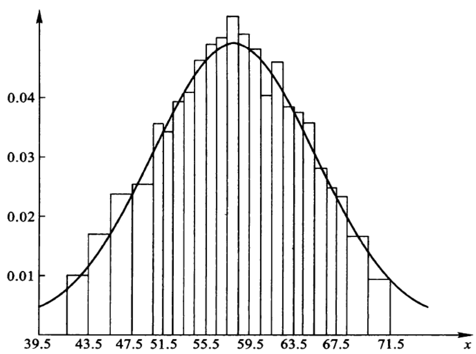


图 3.1.7 实测数据频率直方图与正态密度曲线

例4. 公共汽车车门的高度是按男子与车门顶头碰头机会在0.01以下来设计的. 设男子身高 $X \sim N(170, 6^2)$，问车门高度应如何确定?

解: 设车门高度为h cm, 按设计要求，

$$P(X \ge h) \le 0.01, \text{即} P(X < h) \ge 0.99,$$

故：$P(X<h) = \Phi(\frac{h-170}{6}) \ge 0.99,$

查表得：$\Phi(2.33)=0.99$， 因为分布函数非减,

$$\therefore \frac{h-170}{6} \ge 2.33,$$

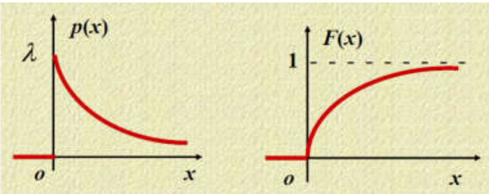
$$h \ge 2.33 \times 6 + 170 \approx 184cm.$$

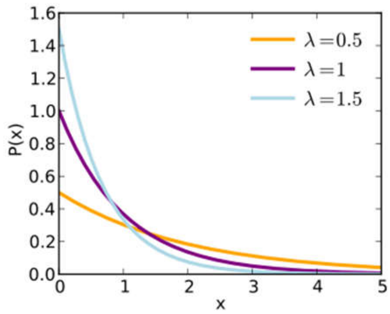
3. 指数(exponential)分布

定义 分布密度函数为： $p(x)=\begin{cases}\lambda e^{-\lambda x}, & x\ge 0\\ 0, & x<0\end{cases}$

分布函数为： $F(x)=\begin{cases}1-e^{-\lambda x}, & x\ge 0\\ 0, & x<0\end{cases}$

其中 λ>0为参数,这个分布称为指数分布,简记为 Exp(λ).





◆ 指数分布的重要性还表现在它具有类似于几何分布的“无记忆性”.

命题. 取非负实值的随机变量ξ服从指数分布，当且仅当它是无记忆性的：

$$P\{\xi>s+t\mid\xi>s\}=P\{\xi>t\},\quad \forall s,t>0$$

证明：必要性。因为ξ服从指数分布，故

$$P\{\xi>t\}=1-F(t)=e^{-\lambda t},\ \forall\, t\ge 0$$

于是, 对任意的 s,t≥0, 有

$$P\{\xi>s+t\mid\xi>s\}=\frac{P\{\xi>s+t\}}{P\{\xi>s\}}=\frac{e^{-\lambda(s+t)}}{e^{-\lambda s}}=e^{-\lambda t},$$

因而ξ具有无记忆性.

◆ 指数分布的应用

指数分布有重要应用，常用它来作为各种“寿命”分布的近似，例如无线电元件的寿命，动物的寿命，电话问题中的通话时间，随机服务系统中的服务时间等都常假定服从指数分布。

假定某种产品的使用寿命服从指数分布，那么，“无记忆性” 表明,无论它已经被使用了多长一段时间,只要还没有损坏,它能再使用一段时间 t 的概率与一件新产品能使用到时间 t 的概率一样，就是说，这种产品将“永远年轻”.

这一点正好说明以指数分布作为寿命分布是有缺陷的.

尽管如此，在有些场合人们还是愿意采用这种易于计算的分布作为产品使用寿命的模型．当然，应用中也可以选用正态分布或威利分布作为寿命分布.

例3.1.8 假设乘客在公交车站等车的时间ξ(分钟)服从参数为0.2的指数分布，密度是：

$$p(x)=0.2\,e^{-0.2x},\quad x>0$$

某人每天上班时如果等车的时间超过10分钟他就决定打车。求他在一个月(30天)里至少有3天是坐出租车上班的概率。

解. 首先需求出他每次等车超过10分钟的概率p，那么在30天里坐公交还是打车，相当于一个 30重 Bernoulli 试验。因此坐出租车上班的次数$\eta\sim B(30,p)$

根据指数分布的分布函数这个人每次等车时间超过 10 分钟的概率是:

$$p=P(\xi>10)=1-F(10)=e^{-2};$$

每个月等车超过10分钟次数$\eta\sim B(30,e^{-2})$; 他至少有三天坐出租车上班的概率就是:

$$P(\eta\ge 3)=\sum_{k=3}^{30}C_{30}^{k}e^{-2k}(1-e^{-2})^{30-k}$$

$$=1-P(\eta\le 2)=0.791549。\quad\square$$

4. 埃尔朗(Erlang)分布

对任意的正整数 r 及实数 $\lambda>0$，密度函数为

$$p(x)=\frac{\lambda^r}{(r-1)!}x^{r-1}e^{-\lambda x},\ x\ge 0.$$

的分布，称为埃尔兰分布。

可以证明：参数为λt 的泊松过程中第 r 个跳跃发生的时刻W_r 服从埃尔兰分布。

证明：$F(t)=P\{W_r<t\}=P(N(t)\ge r)$

$$=1-P(N(t)<r)=1-\sum_{k=0}^{r-1}\frac{(\lambda t)^k}{k!}e^{-\lambda t}$$

$$\therefore\ p(t)=F'(t)=\sum_{k=0}^{r-1}\frac{e^{-\lambda t}}{k!}[\lambda(\lambda t)^k-k\lambda(\lambda t)^{k-1}]$$

$$=\lambda e^{-\lambda t}\left[\sum_{k=0}^{r-1}\frac{(\lambda t)^k}{k!}-\sum_{k=1}^{r-1}\frac{(\lambda t)^{k-1}}{(k-1)!}\right]$$

$$=\lambda e^{-\lambda t}\frac{(\lambda t)^{r-1}}{(r-1)!}=\frac{\lambda^r}{\Gamma(r)}t^{r-1}e^{-\lambda t},$$

其中：$\Gamma(r)=\int_0^\infty x^{r-1}e^{-x}dx$. 当$r$为正整数时，有$\Gamma(r)=(r-1)!$.

可以验证： $\int_0^\infty\frac{\lambda^r}{\Gamma(r)}x^{r-1}e^{-\lambda x}dx=1.$

即：上述$p(x)$的确为一密度函数。

5. Γ 分布

对任意的实数 $r>0,\ \lambda>0$，密度函数为

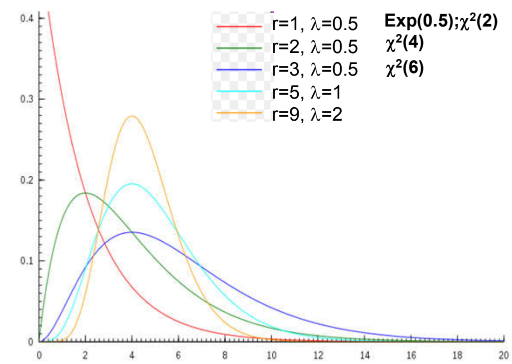
$$p(x)=\begin{cases}\dfrac{\lambda^r}{\Gamma(r)}x^{r-1}e^{-\lambda x}, & x>0\\ 0, & x\le 0\end{cases}$$

的分布称为伽玛Γ分布. 简记为 $\Gamma(r,\lambda)$.

特别地，$\Gamma(1,\lambda)=\mathrm{Exp}(\lambda)$；

当r为正整数数时，$\Gamma(r,\lambda)$ 就是Erlang 分布;

当n为正整数数时，$\Gamma(n/2,1/2)$ 就是自由度为n的$\chi^2(n)$ 分布.



埃尔朗分布是丹麦数学家埃尔朗在研究电话问题时引进的，这些研究开创了排队论这一学科。

埃尔朗分布与帕斯卡分布有着类似的性质。

当r=1时，埃尔兰分布化作指数分布，另外，若记

$$\tau_1=W_1,\ \tau_r=W_r-W_{r-1},\ r=2,3,\cdots$$

则 τ_r 表示泊松过程的第r个来到间隔，用它们可以给出来到时刻W_r的如下表达式：

$$W_r=\tau_1+\tau_2+\cdots+\tau_r.$$

可以证明，$\tau_1,\tau_2,\cdots\tau_r$ 均服从参数为 λ 的指数分布，并且相互独立。

Γ分布包含埃尔朗分布作为它的特例，此外，它在概率论和数理统计中有许多应用，是重要分布之一.

§ 3.2 随机向量, 随机变量的独立性

一、随机向量及其分布

二、边际分布

三、条件分布

四、随机变量的独立性

一、随机向量及其分布

➢ **问题的提出:**

- 例1：研究某一地区学龄儿童的发育情况。仅研究身高H的分布或仅研究体重W的分布是不够的。需要同时考察每个儿童的身高和体重值，研究身高和体重之间的关系，这就要引入定义在同一样本空间的两个随机变量。
- 例2：研究某种型号炮弹的弹着点分布。每枚炮弹的弹着点位置需要由横坐标X和纵坐标Y来确定，而它们是定义在同一样本空间的两个随机变量。

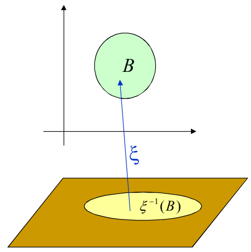
定义. 若随机变量$\xi_1(\omega),\xi_2(\omega),\cdots\xi_n(\omega)$定义在同一概率空间$(\Omega,\mathcal{F},P)$上，则称

$$\xi(\omega)=(\xi_1(\omega),\xi_2(\omega),\cdots\xi_n(\omega))$$

构成一个n维随机变量 或 n维随机向量。

若B_n为R^n上任一博雷尔点集, 有

$$\{\xi(\omega)\in B_n\}\in\mathcal{F}$$



二维随机向量定义示意图

特别地, 对于任意的n个实数 $x_1,x_2,\cdots x_n$,

$$\{\xi_1(\omega)<x_1,\xi_2(\omega)<x_2,\cdots,\xi_n(\omega)<x_n\}=\bigcap_{i=1}^{n}\{\xi_i(\omega)<x_i\}\in\mathcal{F},$$

亦即对于R^n中的 n 维矩形 $C_n=\prod_{i=1}^{n}(-\infty,x_i)$，有

$$\{\xi(\omega)\in C_n\}\in\mathcal{F}$$

定义. 称 n 元函数

$$F(x_1,x_2,\cdots x_n)=P\{\xi_1(\omega)<x_1,\xi_2(\omega)<x_2,\cdots,\xi_n(\omega)<x_n\}$$

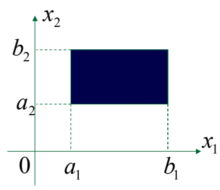
为随机向量$(\xi_1(\omega),\xi_2(\omega),\cdots\xi_n(\omega))$的(联合)分布函数。

给出了分布函数以后，我们可以计算事件

$$\{a_1\le\xi_1(\omega)<b_1,a_2\le\xi_2(\omega)<b_2,\cdots,a_n\le\xi_n(\omega)<b_n\}$$

的概率，例如当$n=2$ 时，有

$$P\{a_1\le\xi_1<b_1,a_2\le\xi_2<b_2\}$$
$$=F(b_1,b_2)-F(a_1,b_2)-F(b_1,a_2)+F(a_1,a_2).$$



类似于一元的场合，可以证明多元分布函数的一些性质：

(1) 单调性：关于每个变元是单调不减函数。

(2) 规范性：

$$F(x_1,x_2,\cdots,-\infty,\cdots x_n)=0,F(+\infty,+\infty,\cdots,+\infty)=1.$$

(3) 关于每个变元左连续。

在二元场合，还应该有：对任意 $a_1\le b_1,a_2\le b_2$, 都有

(4) $F(b_1,b_2)-F(a_1,b_2)-F(b_1,a_2)+F(a_1,a_2)\ge 0$

性质(4)能推出单调性(1)，但(1)不能推出(4)(习题12). 这是多元场合与一元场合的不同之处。

可以证明：满足(2)，(3)，(4)这三条性质的二元函数必为某二元随机变量的分布函数。因此，以后我们称满足以上三条性质的函数$F(x,y)$为二元联合分布函数。

随机向量也有离散型与连续型。

- 离散型：多项分布、多元超几何分布等;
- 连续型：多元均匀分布、多元正态分布等.

◆ 常见的多元离散型分布

1. 多项(multinomial)分布

在实验中，若每次实验的可能结果为 $A_1, A_2, \cdots A_r$ ，而 $P(A_i) = p_i, i = 1,2,\cdots,r,$ 且 $p_1 + p_2 + \cdots + p_r = 1,$

重复这种实验n次，并假定这些实验是相互独立的，若以 $\xi_1, \xi_2, \cdots \xi_r$ 分别记 $A_1, A_2, \cdots A_r$ 出现的次数，则

$$P\{\xi_1 = k_1, \xi_2 = k_2, \cdots \xi_r = k_r\} = \frac{n!}{k_1! k_2! \cdots k_r!} p_1^{k_1} p_2^{k_2} \cdots p_r^{k_r}.$$

这里整数 $k_i \ge 0$，且仅当 $k_1 + k_2 + \cdots + k_r = n$时上式才成立，否则为0。

多项分布用于有放回抽样场合。

2.多元超几何分布

袋中装有i 号球N_i 只，i=1, 2, …, r, $N_1 + N_1+...+ N_r=N$, 从中随机摸出n只，若以 $\xi_1, \xi_2, \cdots, \xi_r$ 分别记 i=1, 2, …, r 号球的出现数，则

$$P\{\xi_1 = n_1, \xi_2 = n_2, \cdots \xi_r = n_r\} = \frac{\binom{N_1}{n_1}\binom{N_2}{n_2}\cdots\binom{N_r}{n_r}}{\binom{N}{n}}.$$

这里整数 $n_i \ge 0$，且仅当$n_1 + n_2 + \cdots + n_r = n$时上式才成立，否则为0。

多元超几何分布用于不放回场合。

◆ 常见的多元连续型分布

在连续型场合，存在着非负函数 $p(x_1, x_2, \cdots x_n)$，使

$$F(x_1, x_2, \cdots x_n) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} p(y_1, y_2, \cdots y_n) dy_1 \cdots dy_n$$

这里的 $p(x_1, x_2, \cdots x_n)$ 称为联合密度函数，满足：

(1) $p(x_1, x_2, \cdots x_n) \ge 0$;

(2) $\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p(x_1, x_2, \cdots x_n) dx_1 \cdots dx_n = 1.$

联合密度函数的充要条件：(1)+(2)

均匀分布和n元正态分布是比较常见的多元连续型分布.

1. 均匀分布

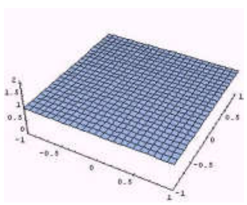
若G为R^n 中有限区域，其测度S(G)>0，则由密度函数

$$p(x_1, x_2, \cdots x_n) = \begin{cases} \frac{1}{S(G)}, & (x_1, x_2, \cdots x_n) \in G \\ 0, & (x_1, x_2, \cdots x_n) \notin G \end{cases}$$

给出的分布称为 G 上的均匀分布。

二维均匀分布所反映的背景:

向平面上有界区域G上任投一质点，若质点落在G内任一小区域B的概率与小区域的面积成正比，而与B的位置无关，则质点的坐标(X, Y)在G上服从均匀分布。



2. 多元正态分布

［多元正态分布］ 若 $\boldsymbol{\Sigma} = (\sigma_{ij})$ 是 n 阶正定对称矩阵，以 $\boldsymbol{\Sigma}^{-1} = (\gamma_{ij})$ 表示 $\boldsymbol{\Sigma}$ 的逆阵；det $\boldsymbol{\Sigma}$ 表示 $\boldsymbol{\Sigma}$ 的行列式的值. $\boldsymbol{\mu} = (\mu_1, \cdots, \mu_n)$ 是任意实值行向量，则由密度函数

$$p(x_1, \cdots, x_n) = \frac{1}{(2\pi)^{\frac{n}{2}} (\det \boldsymbol{\Sigma})^{\frac{1}{2}}} \exp\left\{-\frac{1}{2} \sum_{j,k=1}^{n} r_{jk}(x_j - \mu_j)(x_k - \mu_k)\right\} \tag{3.2.12}$$

定义的分布称为 n **元正态分布**，简记为 $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

这个密度函数也可以写成如下向量形式：

$$p(\boldsymbol{x}) = \frac{1}{(2\pi)^{\frac{n}{2}} (\det \boldsymbol{\Sigma})^{\frac{1}{2}}} \exp\left\{-\frac{1}{2} (\boldsymbol{x} - \boldsymbol{\mu}) \boldsymbol{\Sigma}^{-1} (\boldsymbol{x} - \boldsymbol{\mu})^{\mathrm{T}}\right\} \tag{3.2.13}$$

这里$(\boldsymbol{x} - \boldsymbol{\mu})^{\mathrm{T}}$ 表示行向量$(\boldsymbol{x} - \boldsymbol{\mu})$的转置.

◆ 二元正态分布

定义 若二维随机向量(X,Y)具有概率密度

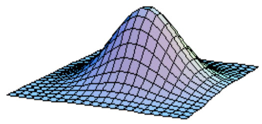
$$p(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\{-\frac{1}{2(1-\rho^2)}[(\frac{x-\mu_1}{\sigma_1})^2-2\rho(\frac{x-\mu_1}{\sigma_1})(\frac{y-\mu_2}{\sigma_2})+(\frac{y-\mu_2}{\sigma_2})^2]\}$$

则称(X,Y)服从参数为$\mu_1,\mu_2,\sigma_1,\sigma_2,\rho$ 的二元正态分布。

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

二维正态分布的图形特点：

二维正态分布的联合密度函数$p(x,y)$ 的图形是一个钟形曲面，它与平行于坐标平面OXY的水平平面相交的截口为椭圆，而与平行 于另外两个坐标平面的竖直平面相交 的截口为正态曲线。



二、边际分布

为方便起见，讨论将对二维场合进行.

1. 离散型场合

设ξ取值$x_1,x_2,\cdots$，η取值$y_1,y_2,\cdots$，记

$$p(x_i,y_j)=P\{\xi=x_i,\eta=y_j\},\quad i,j=1,2,\cdots$$

$\{p(x_i,y_j)\}$称为(ξ,η)的联合分布律. 显然

$$(1)\,p(x_i,y_j)\ge 0,\quad (2)\sum_{i,j}p(x_i,y_j)=1.$$

ξ 的分布律：

$$p_1(x_i)=P\{\xi=x_i\}=\sum_j p(x_i,y_j).$$

η 的分布律：

$$p_2(y_j)=P\{\eta=y_j\}=\sum_i p(x_i,y_j).$$

$\{p_1(x_i),i=1,2,\cdots\},\{p_2(y_j),j=1,2,\cdots\}$ 称为边际分布或边缘分布。

通常用以下表格表示(ξ,η)的联合概率分布和边际分布，这种表称为列联表.

$\xi \backslash \eta$	y_1	y_2	$\cdots$	y_j	$\cdots$	$p_1(\cdot)$
x_1	$p(x_1,y_1)$	$p(x_1,y_2)$	$\cdots$	$p(x_1,y_j)$	$\cdots$	$p_1(x_1)$
x_2	$p(x_2,y_1)$	$p(x_2,y_2)$	$\cdots$	$p(x_2,y_j)$	$\cdots$	$p_1(x_2)$
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
x_i	$p(x_i,y_1)$	$p(x_i,y_2)$	$\cdots$	$p(x_i,y_j)$	$\cdots$	$p_1(x_i)$
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
$p_2(\cdot)$	$p_2(y_1)$	$p_2(y_2)$	$\cdots$	$p_2(y_j)$	$\cdots$	1

［例 1］ 袋中装有 2 只白球及 3 只黑球. 现进行有放回的摸球，定义下列随机变量

$$\xi=\begin{cases}1, & \text{第 1 次摸出白球}\\ 0, & \text{第 1 次摸出黑球}\end{cases}\qquad \eta=\begin{cases}1, & \text{第 2 次摸出白球}\\ 0, & \text{第 2 次摸出黑球}\end{cases}$$

则(ξ,η)的联合概率分布与边际分布由表 3.2.1 给出.

［例 2］ 前例中若采用不放回摸球，则(ξ,η)的联合概率分布及边际分布由表 3.2.2 给出.

表 3.2.1 有放回摸球的概率分布

$\eta \backslash \xi$	0	1	$p_2(y_j)$
0	$\frac{3}{5}\cdot\frac{3}{5}$	$\frac{2}{5}\cdot\frac{3}{5}$	$\frac{3}{5}$
1	$\frac{3}{5}\cdot\frac{2}{5}$	$\frac{2}{5}\cdot\frac{2}{5}$	$\frac{2}{5}$
$p_1(x_i)$	$\frac{3}{5}$	$\frac{2}{5}$	

表 3.2.2 不放回摸球的概率分布

$\eta \backslash \xi$	0	1	$p_2(y_j)$
0	$\frac{3}{5}\cdot\frac{2}{4}$	$\frac{2}{5}\cdot\frac{3}{4}$	$\frac{3}{5}$
1	$\frac{3}{5}\cdot\frac{2}{4}$	$\frac{2}{5}\cdot\frac{1}{4}$	$\frac{2}{5}$
$p_1(x_i)$	$\frac{3}{5}$	$\frac{2}{5}$	

由此可见：联合分布不能由边际分布唯一确定，也就是说二维随机向量的性质不能由它两个分量的个别性质来确定，这时还必须考虑它们之间的联系，这也说明了研究多维随机向量的必要性。

例3.2.1 从 1,2,3,4 中随机取一个数ξ，再从1,…,ξ中随机地取一个数η，计算ξ、η的联合分布律。

解. 首先确定ξ、η的取值范围。

ξ可能取 1,2,3,4； η可能的取值仍然是 1,2,3,4， 并且具有关系$\eta \leq \xi$。

关键是这里必须使用条件概率:

$$P(\xi=i) = 1/4,$$

$$P(\eta=j|\xi=i) = 1/i, \quad 当\ j \leqslant i$$

根据乘法公式， ξ、η的联合分布律为

$$\begin{aligned} &P(\xi=i, \eta=j) \\ =\ &P(\xi=i)\times P(\eta=j|\xi=i) \\ =\ &1/4i, \quad 1\leqslant j \leqslant i \leqslant 4\ 。\end{aligned}$$

$\xi \backslash \eta$	1	2	3	4
1	1/4	0	0	0
2	1/8	1/8	0	0
3	1/12	1/12	1/12	0
4	1/16	1/16	1/16	1/16

一般地, 若(ξ,η)是二维随机向量(离散型或连续型)，其分布函数为$F(x,y)$，我们能得出ξ或η的分布函数

$$F_1(x) = P\{\xi < x\} = P\{\xi < x, \eta < +\infty\} = F(x,+\infty)$$

$$F_2(y) = P\{\eta < y\} = F(+\infty, y).$$

$F_1(x)$ 及 $F_2(y)$ 称为$F(x,y)$的边际分布函数。

例如 已知(X,Y)的分布函数为

$$F(x,y) = \begin{cases} 1-e^{-x}-xe^{-y} & 0\le x\le y \\ 1-e^{-y}-ye^{-y} & 0\le y\le x \\ 0 & 其它 \end{cases}$$

求 $F_X(x)$与$F_Y(y)$。

解：

$$F_X(x) = F(x,+\infty) = \begin{cases} 1-e^{-x} & x\ge 0 \\ 0 & x<0 \end{cases}$$

$$F_Y(y) = F(+\infty,y) = \begin{cases} 1-e^{-y}-ye^{-y} & y\ge 0 \\ 0 & y<0 \end{cases}$$

2. 连续型场合

连续型随机向量(ξ,η)的联合分布函数$F(x,y)$, 联合密度函数$p(x,y)$, 则

$$F(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} p(u,v)\,du\,dv,$$

$$p(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y}, \quad a.e.$$

在连续点(x,y)，(ξ,η)落在小矩形 $(x, x+\Delta x)\times(y, y+\Delta y)$ 上的概率近似为：$p(x,y)\Delta x \Delta y$.

进一步, (ξ,η)落在平面上任意区域D的概率为：

$$P\{(\xi,\eta)\in D\} = \iint_D p(x,y)\,dx\,dy$$

例3.2.2 ξ、η 具有联合密度函数

$$p(x,y)=\begin{cases}2e^{-(2x+y)}, & 当\ x、y>0\\ 0, & 其它\end{cases}$$

(1) 求联合分布函数 $F(x,y)$;

(2) 计算 $P\{0\le\xi<1, 0\le\eta<1\}$

(3) 计算 $P\{\xi\ge\eta\}$

(1) 对任意的 $x>0$、$y>0$,

$$F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}p(u,v)\,dudv$$

$$=\int_0^x\int_0^y 2e^{-(2u+v)}\,dudv=\int_0^y e^{-v}\left[\int_0^x 2e^{-2u}\,du\right]dv$$

$$=(1-e^{-2x})\int_0^y e^{-v}\,dv=(1-e^{-2x})(1-e^{-y})$$

最后得到联合分布函数,

$$F(x,y)=\begin{cases}(1-e^{-2x})(1-e^{-y}), & 当\ x、y>0\\ 0, & 其它\end{cases}$$

(2) $P\{0\le\xi<1, 0\le\eta<1\}$

$=F(1,1)-F(0,1)-F(1,0)+F(0,0)$

$=(1-e^{-2})(1-e^{-1})$

或利用

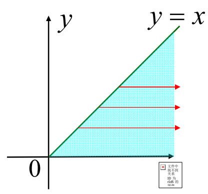
$$P\{0\le\xi<1, 0\le\eta<1\}=\int_0^1\int_0^1 p(x,y)\,dxdy.$$

(3) $P(\xi\geqslant\eta)$

$$=\iint_{x>y>0}2e^{-(2x+y)}\,dxdy$$

$$=\int_0^{+\infty}e^{-y}\left[\int_y^{+\infty}2e^{-2x}\,dx\right]dy$$

$$=\int_0^{+\infty}e^{-3y}\,dy=\frac{1}{3}$$



若F(x,y)是连续型分布函数,有密度函数p(x,y),则

$$F_1(x)=F(x,\infty)=\int_{-\infty}^{x}\left[\int_{-\infty}^{\infty}p(u,v)dv\right]du$$

则 $F_1(x)$ 也是连续型分布函数,其密度函数为

$$p_1(x)=\int_{-\infty}^{\infty}p(x,y)dy.$$

同理, $F_2(x)=F(\infty,x)$ 也是连续型分布函数,其密度函数为

$$p_2(y)=\int_{-\infty}^{\infty}p(x,y)dx$$

$p_1(x)$ 及 $p_2(y)$ 称为 $p(x,y)$ 的边际(分布)密度函数.

◆ 二元正态分布的边际分布:

二元正态分布的边际分布仍为正态分布。

$$(\xi,\eta)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,r) \Rightarrow \begin{cases}\xi\sim N(\mu_1,\sigma_1^2)\\ \eta\sim N(\mu_2,\sigma_2^2)\end{cases}$$

$$p(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}}\exp\{-\frac{1}{2(1-r^2)}[(\frac{x-\mu_1}{\sigma_1})^2-2r(\frac{x-\mu_1}{\sigma_1})(\frac{y-\mu_2}{\sigma_2})+(\frac{y-\mu_2}{\sigma_2})^2]\}$$

解. 作变换 $u=\frac{x-\mu_1}{\sigma_1}$, $v=\frac{y-\mu_2}{\sigma_2}$ ，则

ξ的边缘密度函数为

$$p_1(x)=\frac{1}{2\pi\sigma_1\sqrt{1-r^2}}\int_{-\infty}^{\infty}\exp[-\frac{u^2-2ruv+v^2}{2(1-r^2)}]dv$$

$$=\frac{1}{2\pi\sigma_1\sqrt{1-r^2}}\int_{-\infty}^{\infty}\exp\{-\frac{1}{2}[(\frac{v-ru}{\sqrt{1-r^2}})^2+u^2]\}dv$$

$$=\frac{\exp(-\frac{1}{2}u^2)}{\sqrt{2\pi}\sigma_1}\times\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sqrt{1-r^2}}\exp\{-\frac{(v-ru)^2}{2(1-r^2)}\}dv$$

因此ξ的边缘密度函数为

$$p_1(x)=\frac{\exp(-\frac{1}{2}u^2)}{\sqrt{2\pi}\sigma_1}=\frac{1}{\sqrt{2\pi}\sigma_1}e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}$$

即 $\xi\sim N(\mu_1,\sigma_1^2)$；同理 $\eta\sim N(\mu_2,\sigma_2^2)$。

注意: $\left\{\begin{array}{l}\xi\sim N(\mu_1,\sigma_1^2)\\ \eta\sim N(\mu_2,\sigma_2^2)\end{array}\right. \not\Rightarrow (\xi,\eta)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,r)$

例如: $p(x,y)=\frac{1}{2\pi}e^{-\frac{x^2+y^2}{2}}(1+xy)$, $(-\infty<x,y<+\infty)$

问题：均匀分布的边际分布是否还是均匀分布？

例 设二维随机向量(X,Y)服从区域D上的均匀分布，

其中区域 $D=\{(x,y)|x^2+y^2\le 1\}$

试求二维随机向量(X,Y)关于X,Y的边际概率密度。

解：由题意有

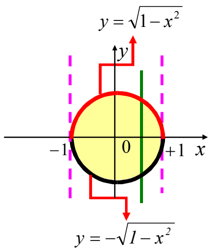
$$p(x,y)=\begin{cases}\frac{1}{\pi} & x^2+y^2\le 1\\ 0 & 其它\end{cases}$$

关于X 的边际概率密度为

$$p_X(x)=\int_{-\infty}^{+\infty}p(x,y)dy$$

当 $|x|>1$ 时，由于有 $p(x,y)=0$

所以 $p_X(x)=0$



当 $|x|\le 1$时，$p_X(x)=\int_{-\infty}^{+\infty}p(x,y)dy$

$$=\int_{-\infty}^{-\sqrt{1-x^2}}p(x,y)dy+\int_{-\sqrt{1-x^2}}^{+\sqrt{1-x^2}}p(x,y)dy+\int_{+\sqrt{1-x^2}}^{+\infty}p(x,y)dy$$

$$=\int_{-\sqrt{1-x^2}}^{+\sqrt{1-x^2}}\frac{1}{\pi}dy=\frac{2}{\pi}\sqrt{1-x^2}$$

所以，二维随机向量 (X,Y) 关于X的边际概率密度为

$$p_X(x)=\begin{cases}\frac{2}{\pi}\sqrt{1-x^2} & |x|\le 1\\ 0 & |x|>1\end{cases}$$

同理可得，随机向量(X,Y)关于Y的边际概率密度。

可见，圆面上均匀分布的边际分布不是一维均匀分布。

例3.2.3 在例3.2.2中已知联合密度函数，

$p(x,y)=2e^{-(2x+y)}$，当 $x、y>0$；

已经计算出联合分布函数:

$F(x,y)=(1-e^{-2x})(1-e^{-y})$，当 $x、y>0$

分别对x,y取极限得到边缘分布函数

$F_1(x)=1-e^{-2x}$，当 $x>0$，

$F_2(y)=1-e^{-y}$，当 $y>0$；

分别对边缘分布函数求导或者分别对联合密度函数积分都将得到

$p_1(x)=2e^{-2x}$，当 $x>0$；

$p_2(y)=e^{-y}$， 当 $y>0$。

三、条件分布

对于多个随机事件可以讨论它们的条件概率，同样地，对于多个随机变量也可以讨论它们的条件分布。仍对二维的场合进行讨论。

1. 离散型场合:

若已知 $\xi=x_i(p_1(x_i)>0)$，则事件 $\{\eta=y_j\}$的条件概率为

$$\boxed{P\{\eta=y_j\,|\,\xi=x_i\}=\frac{P\{\xi=x_i,\eta=y_j\}}{P\{\xi=x_i\}}=\frac{p(x_i,y_j)}{p_1(x_i)}}$$

此式定义了随机变量η关于ξ 的条件分布。

- 例如：前面例中的随机取数问题，

$\xi \backslash \eta$	1	2	3	4	$p_{i\bullet}$
1	1/4	0	0	0	1/4
2	1/8	1/8	0	0	1/4
3	1/12	1/12	1/12	0	1/4
4	1/16	1/16	1/16	1/16	1/4
$p_{\bullet j}$	25/48	13/48	7/48	3/48	

对每个固定的i，η关于$(\xi=i)$的条件分布显然是$p_{j|i}=1/i$，这里$1\leqslant j\leqslant i\leqslant 4$

(1) ξ关于$(\eta=1)$的条件分布律

1	2	3	4
12/25	6/25	4/25	3/25

(2) ξ关于$(\eta=2)$的条件分布律

2	3	4
6/13	4/13	3/13

(3) ξ关于$(\eta=3)$的条件分布律

3	4
4/7	3/7

(4) ξ关于$(\eta=4)$的条件分布？

2. 连续型场合：

对于一般随机向量(ξ, η)，我们也想定义条件分布函数，但是由于会出现 $P(\xi=x)=0$ ，因此我们定义

$$F(y\,|\,x)=P\{\eta<y\,|\,\xi=x\}=\lim_{\Delta x\to 0}P\{\eta<y\,|\,x\le\xi<x+\Delta x\}$$

$$=\lim_{\Delta x\to 0}\frac{P\{x\le\xi<x+\Delta x,\eta<y\}}{P\{x\le\xi<x+\Delta x\}}=\lim_{\Delta x\to 0}\frac{F(x+\Delta x,y)-F(x,y)}{F(x+\Delta x,\infty)-F(x,\infty)}.$$

特别对于有连续密度函数的场合，这定义导出

$$F(y\,|\,x)=P\{\eta<y\,|\,\xi=x\}=\lim_{\Delta x\to 0}\frac{\int_x^{x+\Delta x}\left[\int_{-\infty}^{y}p(u,v)dv\right]du}{\int_x^{x+\Delta x}\left[\int_{-\infty}^{\infty}p(u,v)dv\right]du}$$

当 $p_1(x)\neq 0$, 条件分布函数为

$$F(y\,|\,x)=P\{\eta<y\,|\,\xi=x\}=\frac{\int_{-\infty}^{y}p(x,v)dv}{p_1(x)}=\int_{-\infty}^{y}\frac{p(x,v)}{p_1(x)}dv.$$

因此在给定$\xi=x$的条件下，η的条件密度函数为

$$\boxed{p(y\,|\,x)=\frac{p(x,y)}{p_1(x)}.}$$

类似地，在给定$\eta=y$的条件下，ξ 的分布密度函数为

$$p(x\,|\,y)=\frac{p(x,y)}{p_2(y)}.$$

这里当然也要求：$p_2(y)\neq 0$.

例 二元正态分布$(\xi,\eta)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$的条件分布仍然是正态分布，即

$$\eta\,|\,\xi=x\ \sim N\left(\mu_2+\rho\frac{\sigma_2}{\sigma_1}(x-\mu_1),\ \sigma_2^2(1-\rho^2)\right)$$

$$\xi\,|\,\eta=y\ \sim N\left(\mu_1+\rho\frac{\sigma_1}{\sigma_2}(y-\mu_2),\ \sigma_1^2(1-\rho^2)\right)$$

证：$p(x,y)=\dfrac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\{-\dfrac{1}{2(1-\rho^2)}[(\dfrac{x-\mu_1}{\sigma_1})^2$

$-2\rho(\dfrac{x-\mu_1}{\sigma_1})(\dfrac{y-\mu_2}{\sigma_2})+(\dfrac{y-\mu_2}{\sigma_2})^2]\}$,

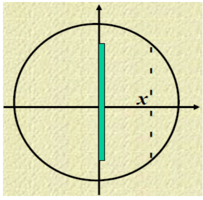
而$p_1(x)=\dfrac{1}{\sqrt{2\pi}\sigma_1}\exp\left[-\dfrac{1}{2}\left(\dfrac{(x-\mu_1)}{\sigma_1}\right)^2\right]$,

所以，$p(y\,|\,x)=\dfrac{p(x,y)}{p_1(x)}$

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

例: 若(ξ,η)服从单位圆上的均匀分布, 在ξ=x的条件下η 的条件分布?

$p(x,y)=1/\pi,\quad x^2+y^2<1$



已经知道ξ、η的边缘分布都不是均匀分布。

$p_1(x)=\frac{2}{\pi}\sqrt{1-x^2},\quad -1<x<1$

当$-1<x<1$时,

$p(y\mid\xi=x)=\frac{1}{2\sqrt{1-x^2}},\ -\sqrt{1-x^2}<y<\sqrt{1-x^2}.$

故:η|ξ=x ~$U(-\sqrt{1-x^2},\sqrt{1-x^2})$

四、随机变量之间的独立性

定义 设$\xi_1,\cdots\xi_n$ 是概率空间$(\Omega,\mathcal{F},P)$上的n个随机变量, 如果他们的联合分布函数等于各自边缘分布函数之积, 即

$$F(x_1,\cdots x_n)=F_1(x_1)\cdots F_n(x_n),\quad \forall x_i, i=1,\cdots,n.$$

则称 $\xi_1,\cdots,\xi_n$ 相互独立。

一族无限多个随机变量称作相互独立的，如果其中任意有限个相互独立。

定理 如果随机变量 $\xi_1,\cdots,\xi_n$ 相互独立，则其中任何一部分随机变量仍独立。

于是，整体独立的多个随机变量是两两独立的，但其逆命题不真。

下面，以 n=2为例讨论独立性定义的等价形式。

定理1 随机变量 ξ与η 相互独立，当且仅当

$$P\{\xi\in B_1,\eta\in B_2\}=P\{\xi\in B_1\}P\{\eta\in B_2\},\quad \forall B_1,B_2\in\mathcal{B}.$$

定理2 如果(ξ,η)为离散型随机向量，则ξ与η独立

$$\Leftrightarrow\ p(x_i,y_j)=p_1(x_i)p_2(y_j),\forall\, i,j=1,2,\cdots$$

定理3 如果(ξ,η) 为连续型随机向量，则ξ与 η独立

$$\Leftrightarrow\ p(x,y)=p_1(x)p_2(y),\quad a.e.\ x,y\in\mathbf{R}.$$

命题 二维正态随机变量相互独立 ⇔ **ρ= 0.**

推论 若随机变量 ξ 与η 独立, 则条件分布化为无条件分布:

$$P\{\eta<y\mid\xi=x\}=P\{\eta<y\}.$$

随机向量之间的独立性:

定义 对于n维随机向量ξ和m维随机向量η，如果

$$P\{\xi\in A,\boldsymbol{\eta}\in B\}=P\{\xi\in A\}P\{\boldsymbol{\eta}\in B\}$$

其中 A, B分别是任意一个n 维和m 维的Borel点集, 则称ξ与 η 相互独立。

显然，若ξ与η独立，则ξ的子向量与η的子向量独立。

§ 3.3 随机变量的函数及其分布

一、博雷尔函数与随机变量的函数

二、单个随机变量的函数的分布

三、随机向量的函数的分布

四、随机向量的变换

五、随机变量的函数的独立性

一、博雷尔函数与随机变量的函数

定义 设$y=g(x)$是 $\mathbf{R}^1$ 到 $\mathbf{R}^1$上的一个映照，若对于一切 $\mathbf{R}^1$中的博雷尔点集 B_1 均有

$$\{x: g(x)\in B_1\}\in\mathcal{B}_1$$

则称$g(x)$是一元博雷尔(可测)函数。

博雷尔函数是很广泛的一类函数，我们所碰到的大部分函数都是博雷尔函数，例如，所有的连续函数和单调函数都是博雷尔函数。

命题 若ξ是概率空间$(\Omega,\mathcal{F},P)$上的随机变量，而$g(x)$是一元博雷尔函数，则$g(\xi)$是$(\Omega,\mathcal{F},P)$上的随机变量.

证明：对一切一元博雷尔点集$B_1\in\mathcal{B}_1$，有

$$\{\omega:g(\xi(\omega))\in B_1\}=\{\omega:\xi(\omega)\in A_1\}\in\mathcal{F}$$

这里，$A_1=\{x:g(x)\in B_1\}$，

由博雷尔函数的定义知它是一维博雷尔点集，再根据随机变量的定义知上式成立。

定义 设$y=g(x_1,\cdots,x_n)$是$\mathbf{R}^n$到$\mathbf{R}^1$上的一个映照，若对于一切$\mathbf{R}^1$中的博雷尔点集B_1均有

$$\{(x_1,\cdots,x_n):g(x_1,\cdots,x_n)\in B_1\}\in\mathcal{B}_n$$

则称$g(x_1,\cdots,x_n)$是***n*** **元博雷尔(可测)函数.**

命题 若$(\xi_1,\cdots\xi_n)$是概率空间$(\Omega,\mathcal{F},P)$上的随机向量，而$g(x_1,\cdots,x_n)$是 ***n*** 元博雷尔函数，则$g(\xi_1,\cdots,\xi_n)$是$(\Omega,\mathcal{F},P)$上的随机变量。

更一般地，***n***维随机向量$(\xi_1,\cdots\xi_n)$的***m***个函数

$$g_1(\xi_1,\cdots,\xi_n),\cdots,g_m(\xi_1,\cdots,\xi_n)$$

这里$g_1,\cdots,g_m$都是***n***元博雷尔函数，则得到一个***m***维随机向量$(g_1(\xi_1,\cdots,\xi_n),\cdots,g_m(\xi_1,\cdots,\xi_n))$.

◆ 离散型随机变量的函数的分布

一般地来说，对于离散性随机变量，求出它的函数的分布并不困难，例如，若ξ的分布列为

$$\begin{pmatrix} x_1 & x_2 & \cdots & x_n & \cdots \\ p_1 & p_2 & \cdots & p_n & \cdots \end{pmatrix}$$

则 g(ξ) 的分布列可由下法得到，列出

$$\begin{pmatrix} g(x_1) & g(x_2) & \cdots & g(x_n) & \cdots \\ p_1 & p_2 & \cdots & p_n & \cdots \end{pmatrix}$$

当然这里可能有某些$g(x_i)$相等，把它们作适当并项即可.

◆ 表上作业法：更一般场合的方法

我们也可用一种更直接的方式——表上作业法来求上述概率分布. 这种表上作业法一般分为三步:

第一步:将(X,Y)的概率分布表达为下列形式.

$X \backslash Y$	-1	0	2
0	0.1	0.2	0
1	0.3	0.05	0.1
2	0.15	0	0.1

第二步,将每一对(x_i,y_j)对应的函数值$g(x_i,y_j)$算出标在对应小方框的右下角($X+Y$见下列左表,XY见下列右表).

$X \backslash Y$	-1	0	2
0	0.1 -1	0.2 0	0 2
1	0.3 0	0.05 1	0.1 3
2	0.15 1	0 2	0.1 4

$X \backslash Y$	-1	0	2
0	0.1 0	0.2 0	0 0
1	0.3 -1	0.05 0	0.1 2
2	0.15 -2	0 0	0.1 4

第三步:将第二部表格中$g(x_i,y_j,)$的值相同的项合并，即得g(X,Y)的分布律.

◆ 离散卷积(convolution)公式:

---独立随机变量和的分布的公式.

若ξ与η是相互独立的随机变量，它们都取非负整数值，其概率分布为$\{a_k\}$及$\{b_k\}$，下面是$\zeta=\xi+\eta$的分布:

因为 $\{\zeta=r\}=\{\xi=0,\eta=r\}+\{\xi=1,\eta=r-1\}$
$+\cdots+\{\xi=r,\eta=0\}$.

利用独立性假定得

$$c_r=P\{\zeta=r\}=a_0b_r+a_1b_{r-1}+\cdots+a_rb_0\,,\quad r=0,1,2,\cdots$$

$$\boxed{P(\xi+\eta=r)=\sum_{k=0}^{r}P(\xi=k)P(\eta=r-k)=\sum_{k=0}^{r}a_kb_{r-k}}$$

二项分布的加法定理:

若 $\xi \sim B(m,p)$, $\eta \sim B(n,p)$, ξ,η 相互独立, 则

$$\xi+\eta \sim B(m+n,p).$$

注. 从二项分布的背景很容易理解这一结论:

已知$\xi \sim B(m,p)$, 则ξ表示在m次独立重复试验中事件A出现的次数, $p=P(A)$; 已知 $\eta \sim B(n,p)$, 则 η表示在 n 次独立重复试验中事件 A 出现的次数. 由于ξ与η相互独立, 则 $\xi+\eta$表示在 $m+n$ 次独立重复试验中事件 A 出现的次数, 即 $\xi+\eta \sim B(m+n,p)$.

Poisson分布的加法定理:

若$\xi \sim P(\lambda_1)$,$\eta \sim P(\lambda_2)$,且相互独立, 则 $\xi+\eta \sim P(\lambda_1+\lambda_2)$.

二、单个随机变量的函数的分布

这里的一般问题是:已知随机变量 ξ 的分布函数 $F(x)$或密度函数 $p(x)$,要求 $\eta=g(\xi)$的分布函数 $G(y)$或密度函数 $q(y)$.

由(3.1.29)可知

$$G(y)=P\{\eta<y\}=P\{g(\xi)<y\}$$
$$=\int_{g(x)<y} p(x)\,dx \qquad (3.3.7)$$

上述积分计算的难易既与被积函数即 ξ 的密度函数 $p(x)$ 的表达式有关,更与积分区域 $\{x:g(x)<y\}$ 的形状相关,差别很大,因此这类问题通常采用个案处理的方式,但在方法上大体可分为直接法与变换法两类.

◆ 直接法:

直接法通过把 $\{g(\xi)<y\}$ 直接化为关于 ξ 的等价事件而求得 η 的分布函数或密度函数.

[例 1] 若随机变量 ξ 有密度函数 $p(x)$,而 $\eta=a\xi+b$,这里 $a\neq 0$. 求 η 的密度函数 $q(y)$.

[解] 分别记 ξ 与 η 的分布函数为 $F(x)$ 及 $G(y)$,显然有

$$G(y)=P\{\eta<y\}=P\{a\xi+b<y\}$$
$$=\begin{cases} P\left\{\xi<\dfrac{y-b}{a}\right\}=F\left(\dfrac{y-b}{a}\right), & 若\ a>0 \\ P\left\{\xi>\dfrac{y-b}{a}\right\}=1-F\left(\dfrac{y-b}{a}\right), & 若\ a<0 \end{cases}$$

因此

$$q(y)=\begin{cases} \dfrac{1}{a}p\left(\dfrac{y-b}{a}\right), & 若\ a>0 \\ -\dfrac{1}{a}p\left(\dfrac{y-b}{a}\right), & 若\ a<0 \end{cases}$$

或统一起来,写成

$$q(y)=\frac{1}{|a|}p\left(\frac{y-b}{a}\right) \qquad (3.3.8)$$

把例 1 的结果应用到正态分布的场合. 若 ξ 服从 $N(\mu,\sigma^2)$,则 $\zeta=\dfrac{\xi-\mu}{\sigma}$的密度函数

$$q(z)=\sigma\cdot\frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{\left[\sigma\left(z+\frac{\mu}{\sigma}\right)-\mu\right]^2}{2\sigma^2}\right\}=\frac{1}{\sqrt{2\pi}}e^{-z^2/2}$$

即 ζ 服从 $N(0,1)$,这是我们早已提到过的.

[例 2] 若 $\xi\sim N(\mu,\sigma^2)$,求 $\eta=e^{\xi}$的密度函数.

[解] 当 $y>0$ 时,

$$P\{\eta<y\}=P\{e^{\xi}<y\}=P\{\xi<\ln y\}$$
$$=\int_{-\infty}^{\ln y}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx$$

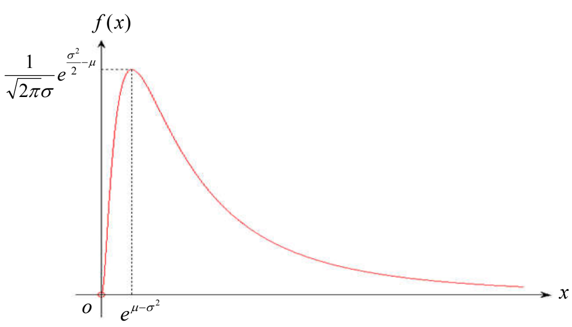
所以,η 的密度函数为

$$q(y)=\frac{1}{\sqrt{2\pi}\sigma y}e^{-\frac{(\ln y-\mu)^2}{2\sigma^2}}, \qquad y>0. \qquad (3.3.9)$$

η 的对数即 $\ln\eta=\xi$ 服从正态分布,故称 η 所服从的分布为**对数正态分布**.

当代金融学用对数正态分布取代正态分布作为资产价格分布建立起了十分漂亮而合理的理论. 另外,销售量,元件寿命等也已普遍使用对数正态分布作为模型. 这个分布的重要性正在提高.

Figure: 对数正态分布的密度函数



［例 3］ 若 $\zeta\sim N(0,1)$，求 $\eta=\zeta^2$ 的密度函数.

［解］ 当 $y\leqslant 0$ 时，$G(y)=P\{\eta<y\}=0$，显然，此时 $q(y)=0$.

当 $y>0$，

$$G(y)=P\{\eta<y\}=P\{\zeta^2<y\}=P\{-\sqrt{y}<\zeta<\sqrt{y}\}$$

$$=\int_{-\sqrt{y}}^{\sqrt{y}}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx=2\int_{0}^{\sqrt{y}}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx$$

因此 $\eta=\zeta^2$ 的密度函数为

$$q(y)=\frac{1}{\sqrt{2\pi}}y^{-\frac{1}{2}}e^{-\frac{y}{2}},\quad y>0 \qquad (3.3.10)$$

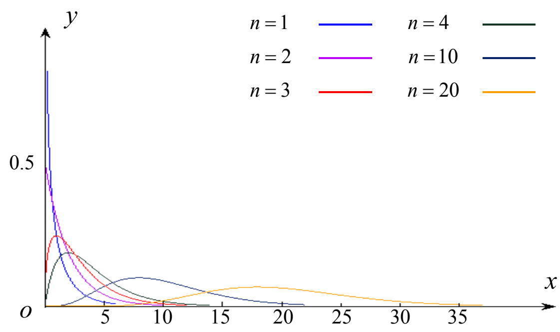
顺便指出，(3.3.10)是下列分布当 $n=1$ 时的特例.

［χ^2分布］ 具有密度函数

$$p(x)=\frac{1}{2^{n/2}\Gamma\left(\frac{n}{2}\right)}x^{\frac{n}{2}-1}e^{-\frac{x}{2}},\quad x>0 \qquad (3.3.11)$$

的分布称为**具有自由度 n 的χ^2分布**.

Figure χ^2-分布



例4 (分段单调情形) 设随机变量X的密度为

$$p_X(x)=\begin{cases}\dfrac{2x}{\pi^2} & 0<x<\pi\\ 0 & \text{其它}\end{cases}$$

求$Y=\sin X$的概率密度.

解：随机变量$Y=\sin X$的分布函数为

$$F_Y(y)=P\{Y<y\}=P\{\sin X<y\}.$$

当 $y\le 0$时， $F_Y(y)=0$.

当 $1\le y$时， $F_Y(y)=1$.

当$0<y<1$时， 由于$0<x<\pi$， 则有

$$F_Y(y)=P\{0<X<\arcsin y\}+P\{\pi-\arcsin y<X<\pi\}$$

$$=\int_0^{\arcsin y}\frac{2x}{\pi^2}dx+\int_{\pi-\arcsin y}^{\pi}\frac{2x}{\pi^2}dx$$

$$=\left(\frac{\arcsin y}{\pi}\right)^2+1-\left(\frac{\pi-\arcsin y}{\pi}\right)^2$$

$$\text{所以有}:F_Y(y)=\begin{cases}0 & y\le 0\\ \left(\dfrac{\arcsin y}{\pi}\right)^2+1-\left(\dfrac{\pi-\arcsin y}{\pi}\right)^2 & 0<y<1\\ 1 & y\ge 1\end{cases}$$

$$\text{故有随机变量}Y\text{的概率密度为 } p_Y(y)=\begin{cases}\dfrac{2}{\pi\sqrt{1-y^2}}, & 0<y<1\\ 0 & ,\text{其他}\end{cases}$$

法二：

当$0<y<1$时， 由于$0<x<\pi$， 则有

$$F_Y(y)=P\{0<X<\arcsin y\}+P\{\pi-\arcsin y<X<\pi\}$$

$$=F_X(\arcsin y)+1-F_X(\pi-\arcsin y)$$

所以有

$$p_Y(y)=p_X(\arcsin y)\frac{d(\arcsin y)}{dy}-p_X(\pi-\arcsin y)\frac{d(\pi-\arcsin y)}{dy}$$

$$=\frac{2\arcsin y}{\pi^2\sqrt{1-y^2}}+\frac{2(\pi-\arcsin y)}{\pi^2\sqrt{1-y^2}}=\frac{2}{\pi\sqrt{1-y^2}}$$

$$\text{故有随机变量}Y\text{的概率密度为 } p_Y(y)=\begin{cases}\dfrac{2}{\pi\sqrt{1-y^2}}, & 0<y<1\\ 0 & ,\text{其他}\end{cases}$$

◆ 变换法：

一般地，若ξ为连续型随机变量，其密度函数为$p(x)$，而$\eta=g(\xi)$, 我们对下面两种情形进行讨论.

（1）若 $g(x)$ 严格单调，且反函数 $g^{-1}(y)$ 有连续导函数，则 $\eta=g(\xi)$是具有密度函数

$$p(g^{-1}(y))\cdot\left|[g^{-1}(y)]'\right|$$

的连续型随机变量。

注：这里和后面均约定，对使反函数无意义的点y，密度函数定义为0.

证明: 对于任一实数a，$E_a=\{x:g(x)<a\}\xrightarrow{y=g(x)}\{y:y<a\}$

$$P\{\eta<a\}=\int_{-\infty}^{a}p_\eta(y)dy=P\{g(\xi)<a\}=P(\xi\in E_a)$$

$$=\int_{E_a}p_\xi(x)dx\overset{x=g^{-1}(y)}{=}\int_{-\infty}^{a}p_\xi[g^{-1}(y)]\cdot|[g^{-1}(y)]'|dy.$$

这里利用了“变量代换”，由上式知结论成立。

例3.3.2 (随机变量的线性变换)

设连续随机变量ξ有密度函数$p_1(x)$，定义随机变量$\eta=a+b\xi$，则η的密度函数为

$$p_2(y)=\frac{1}{|b|}p_1\left(\frac{y-a}{b}\right)$$

特别地，

(1) $\xi\sim U(c,d)$, $b>0$, 则$a+b\xi\sim U(a+bc,a+bd)$;

(2) $\xi\sim G(\lambda,\alpha)$, 则 $b\xi\sim G(\lambda/b,\alpha)$;

(3) $\xi\sim N(\mu,\sigma^2)$, 则$a+b\xi\sim N(a+b\mu,b^2\sigma^2)$;

例3.3.3 (Cauchy分布)

设$\xi\sim U(-\pi/2,\pi/2)$, 定义$\eta=\mathrm{tg}\,\xi$，则η的密度函数为

$$p(y)=\frac{1}{\pi}\frac{1}{1+y^2},\quad y\in\mathrm{R}^1$$

（2）若 $g(x)$在不相重叠的区间 $I_1,I_2,\cdots$ 上逐段严格单调，其反函数分别为$h_1(y),h_2(y),\cdots$，而且它们均有连续导数，那么 $\eta=g(\xi)$ 是连续型随机变量，其密度函数为

$$p[h_1(y)]\cdot|h_1'(y)|+p[h_2(y)]\cdot|h_2'(y)|+\cdots$$

证明: 给定实数a，$E_i(a)=\{x:g(x)<a,x\in I_i\}$.

显然诸$E_i(a)$不相交，且 $\cup E_i(a)=\{x:g(x)<a\}$

$$P\{\eta<a\}=\int_{-\infty}^{a}p_\eta(y)dy=P\{g(\xi)<a\}=P\{\xi\in\sum_i E_i(a)\}$$

$$=\sum_i\int_{E_i(a)}p(x)dx=\sum_i\int_{-\infty}^{a}p[h_i(y)]\cdot|h_i'(y)|dy$$

$$=\int_{-\infty}^{a}\sum_i p[h_i(y)]\cdot|h_i'(y)|dy$$

因此结论成立。

例3.3.4 (随机变量平方的分布)

设ξ有密度函数$p_1(x)$，定义$\eta=\xi^2$，则η的密度函数为

$$p_2(y)=\frac{1}{2\sqrt{y}}[p_1(\sqrt{y})+p_1(-\sqrt{y})],\ y>0$$

例如：若ξ服从N(0,1) ,计算ξ^2 的密度函数。

解：$$q(y)=\varphi(-\sqrt{y})\left|-\frac{1}{2\sqrt{y}}\right|+\varphi(\sqrt{y})\left|\frac{1}{2\sqrt{y}}\right|$$

$$=\frac{1}{\sqrt{2\pi}}y^{-\frac{1}{2}}e^{-\frac{y}{2}},\ y>0.$$

这恰为自由度为n=1的 χ^2分布, 或 Γ(0.5,0.5).

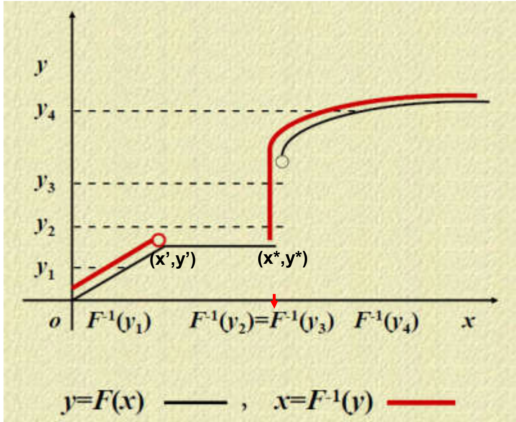
◆ 均匀分布的应用

1. 分布函数的单调逆函数

每个**R**上的严格单调的连续函数$F(x)$都有唯一的反函数$F^{-1}(x)$，反函数与$F(x)$有相同的单调性和连续性. 分布函数只有非降性和左连续性，需要将反函数的概念加以推广。

定义：称 $F^{-1}(y)=\inf\{x: F(x)>y\},\quad y\in(0,1)$

为分布函数$F(x)$的单调逆。



分布函数的单调逆函数的性质：

（1）当F(x)是严格上升的连续函数时，它的单调逆就是普通的反函数。

（2）单调不降。

(3) $F(F^{-1}(y))\le y,\ \forall y\in(0,1).$

当F(x)在点 $F^{-1}(y)$连续时等号成立，反之不真。

(4) $F^{-1}(F(x))\ge x,\ \forall x\in(-\infty,\infty).$ 严格不等号有时成立.

推论 对任意给定的 $x\in(-\infty,\infty),\ y\in(0,1)$，有

$1^0\ F^{-1}(y)\ge x \Leftrightarrow y\ge F(x),$

$2^0\ F^{-1}(y)<x \Leftrightarrow y<F(x).$

（5）$F^{-1}(y)$在任意点$y\in(0,1)$处均右连续。

2. 均匀分布的特殊地位

命题1 设随机变量ξ的分布函数为F(x), 且F(x) 为连续函数, 则θ=F(ξ)服从[0,1]上的均匀分布。

证明：对任意的 $x\in[0,1]$，有

$$P\{\theta<x\}=P\{F(\xi)<x\}=P\{\xi<F^{-1}(x)\}=F(F^{-1}(x))=x.$$

命题2 设F(x) 为分布函数, θ 服从[0,1]上的均匀分布, 则随机变量 $\xi=F^{-1}(\theta)$以F(x)为分布函数。

证明：对任意的 $x\in(-\infty,\infty)$，有

$$P\{\xi<x\}=P\{F^{-1}(\theta)<x\}=P\{\theta<F(x)\}=F(x).$$

这样，只要我们能产生[0,1]中均匀分布的随机变量的样本（观察值），那么我们也就能通过(3.3.17)产生分布函数为 $F(x)$ 的随机变量的样本，这结论在蒙特卡罗方法中具有基本的重要性. 通常的做法是利用数学或物理的方法产生[0,1]中均匀分布随机变量的样本（称为**均匀分布随机数**），再利用变换(3.3.17)得到任意分布 $F(x)$ 的随机数.

随机变量的存在性定理:

定理：若$F(x)$是左连续的单调不减函数，且 $F(-\infty)=0$，$F(+\infty)=1$，则存在一个概率空间$(\Omega,\mathcal{F},P)$及其上的随机变量$\xi(\omega)$，使ξ的分布函数正好为$F(x)$。

三、随机向量函数的分布

若$\eta=g(\xi_1,\cdots,\xi_n)$，而$(\xi_1,\cdots,\xi_n)$的密度函数为$p(x_1,\cdots,x_n)$，则

$$G(y)=P\{\eta<y\}=\int\cdots\int_{g(x_1,\cdots,x_n)<y} p(x_1,\cdots,x_n)dx_1\cdots dx_n,$$

更进一步，$\eta=g(\xi_1,\cdots,\xi_n)$的密度函数为

$$q(y)=G'(y),\quad a.e.\ \ y\in R.$$

◆ 和的分布

定理. 设(X,Y)的密度函数为$p(x,y)$，则$Z=X+Y$的密度函数为

$$p_Z(z)=\int_{-\infty}^{+\infty}p(x,z-x)dx=\int_{-\infty}^{+\infty}p(z-y,y)dy.$$

特别地，当 X,Y 相互独立时，

$$p_Z(z)=\int_{-\infty}^{+\infty}p_X(x)p_Y(z-x)dx$$
$$=\int_{-\infty}^{+\infty}p_X(z-y)p_Y(y)dy.$$

习惯上，函数$f(x),g(y)$的**卷积**定义为

$$f*g(z)=\int_{-\infty}^{+\infty}f(x)g(z-x)dx.$$

所以，当X,Y相互独立时，有

$$p_{X+Y}(z)=p_X*p_Y(z)=p_Y*p_X(z).$$

◆ 商、差、积的分布

定理. 设 (X,Y) 的密度函数为 $p(x,y)$，则 $Z=\frac{X}{Y}$，$Z=X-Y,Z=XY$ 的密度函数分别为

$$p_{\frac{X}{Y}}(z)=\int_{-\infty}^{+\infty}|y|p(zy,y)dy.$$

$$p_{X-Y}(z)=\int_{-\infty}^{+\infty}p(x,x-z)dx=\int_{-\infty}^{+\infty}p(y+z,y)dy;$$

$$p_{XY}(z)=\int_{-\infty}^{+\infty}\frac{1}{|x|}p(x,\frac{z}{x})dx=\int_{-\infty}^{+\infty}\frac{1}{|y|}p(\frac{z}{y},y)dy.$$

◆ 最大值，最小值的分布

定理. 设X, Y 相互独立，分布函数分别为$F_X(x)$，$F_Y(y)$，则 M=max{X,Y}，N=min{X,Y}的分布函数为

$$F_{max}(z)=F_X(z)F_Y(z)$$
$$F_{min}(z)=1-(1-F_X(z))(1-F_Y(z))$$

- 推广到n个相互独立的随机变量的情况

设$X_1,X_2,\cdots,X_n$是n个相互独立的随机变量，它们的分布函数分别为：$F_{X_i}(x_i)\quad i=1,2,\cdots n,$

则：$M=\max\limits_{1\le i\le n}X_i$及 $N=\min\limits_{1\le i\le n}X_i$的分布函数$F_{max}(z)$和$F_{min}(z)$为：

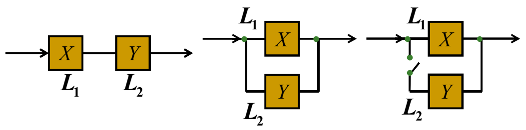
$$\boxed{\begin{aligned}&F_{max}(z)=F_{X_1}(z)F_{X_2}(z)\cdots F_{X_n}(z)\\&F_{min}(z)=1-[1-F_{X_1}(z)][1-F_{X_2}(z)]\cdots[1-F_{X_n}(z)]\end{aligned}}$$

特别，当$X_1,X_2,\cdots,X_n$相互独立且具有相同分布函数$F(x)$时，

$$\boxed{\begin{aligned}&F_{max}(z)=(F(z))^n\\&F_{min}(z)=1-[1-F(z)]^n\end{aligned}}$$

例 设系统 L 由两个相互独立的子系统 L_1,L_2 连接而成,连接的方式分别为 (1) 串联, (2) 并联, (3)备用 (当系统L_1损坏时, 系统 L_2开始工作) , 如下图所示。设 L_1,L_2 的寿命分别为X,Y，已知它们的概率密度分别为 $p_X(x)$，$p_Y(y)$.

试分别就以上三种连接方式写出L的寿命Z的概率密度.



$Z=\min(X,Y)$　　$Z=\max(X,Y)$　　$Z=X+Y$

四、连续型随机向量的变换

若$(\xi_1,\cdots,\xi_n)$的密度函数为$p(x_1,\cdots,x_n)$，求：

$\eta_1=g_1(\xi_1,\cdots,\xi_n),\cdots,\eta_m=g_m(\xi_1,\cdots,\xi_n)$的联合分布.

$$G(y_1,\cdots,y_m)=P\{\eta_1<y_1,\cdots,\eta_m<y_m\}$$
$$=\int\cdots\int_{\substack{g_1(x_1,\cdots x_n)<y_1\\ \cdots\\ g_m(x_1,\cdots x_n)<y_m}}p(x_1,\cdots,x_n)dx_1\cdots dx_n$$

显然，这是最一般的场合，当 m=1时便是随机向量的情形，当 m=n=1时得到单个随机变量的函数的情形。

下面考虑另一个重要的特殊情形，即当$(\xi_1,\cdots,\xi_n)$与$(\eta_1,\cdots,\eta_m)$有一一对应变换关系时，显然此时有m=n.

若对 $y_i = g_i(x_1,\cdots,x_n)$ 存在唯一的反函数 $x_i = x_i(y_1,\cdots,y_n)$ $(i=1,\cdots,n)$，且 $(\eta_1,\cdots,\eta_n)$ 的密度函数为 $q(y_1,\cdots,y_n)$，那么

$$G(y_1,\cdots,y_n)=\int\cdots\int_{\substack{u_1<y_1\\ \cdots\\ u_n<y_n}} q(u_1,\cdots,u_n)du_1\cdots du_n$$

比较前面两式可知：

$$q(y_1,\cdots,y_n)=\begin{cases} p(x_1,\cdots,x_n)\,|J|, & (y_1,\cdots,y_n)\in R(g_1,\cdots,g_n) \\ 0, & \text{其它} \end{cases}$$

其中 J 为坐标变换的雅可比行列式：

$$J=\begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \cdots & \frac{\partial x_n}{\partial y_1} \\ \cdots & \cdots & \cdots \\ \frac{\partial x_1}{\partial y_n} & \cdots & \frac{\partial x_n}{\partial y_n} \end{vmatrix}.$$

这里假定上述偏导数存在且连续。

例6 若 (ξ_1,ξ_2) 的密度函数为 $p(x_1,x_2)$，而 $\begin{cases}\eta_1 = a\xi_1 + b\xi_2 \\ \eta_2 = c\xi_1 + d\xi_2\end{cases}$

这里 $\Delta=\begin{vmatrix} a & b \\ c & d \end{vmatrix}\neq 0$，试求 (η_1,η_2) 的密度函数.

解：本例中

$$\begin{cases} y_1 = g_1(x_1,x_2)=ax_1+bx_2, \\ y_2 = g_2(x_1,x_2)=cx_1+dx_2 \end{cases}$$

因此，

$$x_1=\frac{d}{\Delta}y_1-\frac{b}{\Delta}y_2,\quad x_2=-\frac{c}{\Delta}y_1+\frac{a}{\Delta}y_2.$$

而

$$J=\begin{vmatrix} \frac{d}{\Delta} & -\frac{b}{\Delta} \\ -\frac{c}{\Delta} & \frac{a}{\Delta} \end{vmatrix}=\frac{ad-bc}{\Delta^2}=\frac{1}{ad-bc}.$$

最后，根据公式得

$$q(y_1,y_2)=\frac{p(\frac{d}{\Delta}y_1-\frac{b}{\Delta}y_2,-\frac{c}{\Delta}y_1+\frac{a}{\Delta}y_2)}{|ad-bc|}.$$

［例 7］ 若 ξ 与 η 相互独立，分别服从自由度为 m 和 n 的 χ^2 分布，试求 $\alpha=\xi+\eta$ 与 $\beta=\frac{\xi}{\eta}\cdot\frac{n}{m}$ 的密度函数 $q(u,v)$.

［解］ 由题设据(3.3.11)知 (ξ,η) 的联合密度函数为

$$p(x,y)=\frac{1}{2^{\frac{m+n}{2}}\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)}x^{\frac{m}{2}-1}y^{\frac{n}{2}-1}e^{-\frac{x+y}{2}}$$

当 $x>0,y>0$.

对 $u>0,v>0$ 作变换 $u=x+y,v=\frac{x}{y}\cdot\frac{n}{m}$，其逆变换为

$$x=\frac{muv}{n+mv},\quad y=\frac{nu}{n+mv}.$$

由于

$$J^{-1}=\begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix}=\begin{vmatrix} 1 & 1 \\ \frac{n}{ym} & -\frac{xn}{y^2m} \end{vmatrix}=-\frac{n(x+y)}{my^2}$$

$$=-\frac{n}{m}\cdot\frac{\left(1+\frac{m}{n}v\right)^2}{u}$$

因此

$$|J|=\frac{m}{n}\cdot\frac{u}{\left(1+\frac{m}{n}v\right)^2}$$

于是 (α,β) 的联合分布密度函数为

$$q(u,v)=\frac{1}{2^{\frac{m+n}{2}}\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)}e^{-\frac{u}{2}}\left(\frac{m}{n}\right)^{\frac{m}{2}-1}u^{\frac{m+n}{2}-2}$$

$$\times\frac{v^{\frac{m}{2}-1}}{\left(1+\frac{m}{n}v\right)^{\frac{m+n}{2}-2}}\cdot\frac{m}{n}\cdot\frac{u}{\left(1+\frac{m}{n}v\right)^2}$$

$$=\frac{1}{2^{\frac{m+n}{2}}\Gamma\left(\frac{m+n}{2}\right)}u^{\frac{m+n}{2}-1}e^{-\frac{u}{2}}$$

$$\times\frac{\Gamma\left(\frac{m+n}{2}\right)\cdot\left(\frac{m}{n}\right)^{\frac{m}{2}}v^{\frac{m}{2}-1}}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)\cdot\left(1+\frac{m}{n}v\right)^{\frac{m+n}{2}}}\qquad(3.3.37)$$

因此，α 与 β 独立，而且 α 服从自由度为 $m+n$ 的 χ^2 分布.

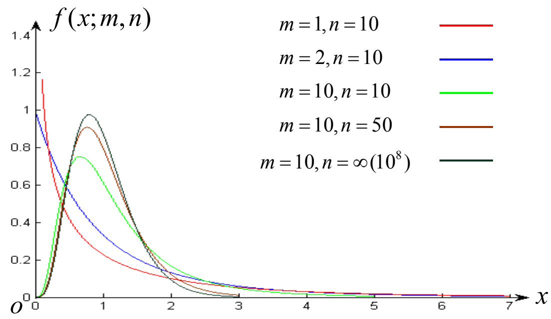
随机变量 $\beta=\frac{\xi}{\eta}\cdot\frac{n}{m}$ 的密度函数为

$$f(x;m,n)=\begin{cases}\dfrac{\Gamma\left(\frac{m+n}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)}\dfrac{\left(\frac{m}{n}\right)^{\frac{m}{2}}x^{\frac{m}{2}-1}}{\left(1+\frac{m}{n}x\right)^{\frac{m+n}{2}}}, & x>0\\ 0, & x\leq 0\end{cases}\qquad (3.3.39)$$

这个分布称为 **F 分布**，是数理统计中重要分布之一.

注：这个例子说明，即使由相同的随机向量构成的不同函数也是可能是独立的.

Figure: F 分布



■ 增补变量法：

在上一小节讨论随机向量的函数的分布律中，主要只讲直接法，下面我们通过一个例子来说明在这种场合如何通过增补变量使用变换法.

［例 8］ 设 ξ,η 为两个独立随机变量，ξ 服从 $N(0,1)$，η 服从自由度为 n 的 χ^2 分布(3.3.11)，令 $T=\xi\Big/\sqrt{\frac{\eta}{n}}$，试求 T 的密度函数.

［解］ 为求得 T 的密度函数，引进增补变量 $S=\eta$，先求 (S,T) 的联合密度函数.

ξ,η 相互独立，故 (ξ,η) 的联合密度函数为

$$p(x,y)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\cdot\frac{1}{2^{n/2}\Gamma\left(\frac{n}{2}\right)}y^{\frac{n}{2}-1}e^{-\frac{y}{2}},\quad -\infty<x<\infty,\ y>0$$

变换 $s=y,\ t=\frac{x}{\sqrt{y/n}}$ 的逆变换为 $x=t\left(\frac{s}{n}\right)^{1/2},\ y=s$，其雅可比行列式

$$J=\begin{vmatrix}\frac{\partial x}{\partial s} & \frac{\partial x}{\partial t}\\ \frac{\partial y}{\partial s} & \frac{\partial y}{\partial t}\end{vmatrix}=\begin{vmatrix}\frac{t}{2n}\left(\frac{s}{n}\right)^{-\frac{1}{2}} & \left(\frac{s}{n}\right)^{\frac{1}{2}}\\ 1 & 0\end{vmatrix}=-\left(\frac{s}{n}\right)^{\frac{1}{2}}$$

$$|J|=\left(\frac{s}{n}\right)^{\frac{1}{2}}$$

故 (S,T) 的联合密度函数为

$$q(s,t)=p\left(t\left(\frac{s}{n}\right)^{\frac{1}{2}},s\right)|J|$$

$$=\frac{1}{\sqrt{2\pi}}e^{-\frac{st^2}{2n}}\cdot\frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}s^{\frac{n}{2}-1}e^{-\frac{s}{2}}\cdot\left(\frac{s}{n}\right)^{\frac{1}{2}}$$

因而 T 的密度函数为

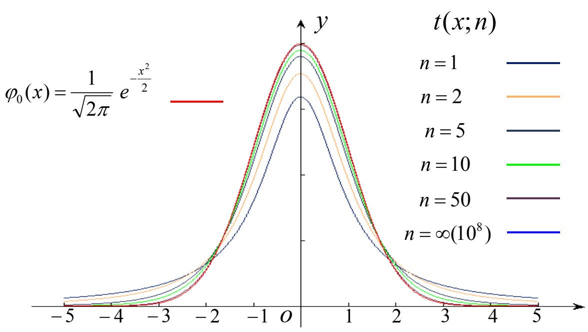
$$p_T(t)=\int_0^\infty q(s,t)\,ds=\frac{1}{2^{\frac{n+1}{2}}\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)}\int_0^\infty e^{-\left(1+\frac{t^2}{n}\right)\frac{s}{2}}s^{\frac{n+1}{2}-1}ds$$

$$=\frac{\left(1+\frac{t^2}{n}\right)^{-\frac{n+1}{2}}}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)}\int_0^\infty e^{-u}u^{\frac{n+1}{2}-1}du$$

$$=\frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{t^2}{n}\right)^{-\frac{n+1}{2}}$$

称为**自由度为 n 的 t 分布**，它是数理统计中另一重要分布.

Figure: t 分布



五、随机变量的函数的独立性

定理 若 $\xi_1,\cdots,\xi_n$ 是相互独立的随机变量，则它们各自的函数 $f_1(\xi_1),\cdots,f_n(\xi_n)$ 也是相互独立的，这里 $f_i(i=1,\cdots,n)$ 是任意的一元博雷尔函数。

证明：对任意的一元博雷尔点集 $A_1,\cdots,A_n$ 有

$$\begin{aligned}&P\{f_1(\xi_1)\in A_1,\cdots,f_n(\xi_n)\in A_n\}\\&=P\{\xi_1\in f_1^{-1}(A_1),\cdots,\xi_n\in f_n^{-1}(A_n)\}\\&=P\{\xi_1\in f_1^{-1}(A_1)\}\cdots P\{\xi_n\in f_n^{-1}(A_n)\}\\&=P\{f_1(\xi_1\}\in A_1\}\cdots P\{f_n(\xi_n)\in A_n\}.\end{aligned}$$

证毕.

定理的结论是明显的，它可以推广到随机向量的场合。

前面的例子说明，由相同的随机向量构成的不同函数也可能是独立的，下面再来看几个例子。

例9 若ξ与η是相互独立的随机变量，均服从 N(0,1)，试证化为极坐标后，$\rho=\sqrt{\xi^2+\eta^2}$ 及 $\varphi=\operatorname{arctg}\left(\frac{\eta}{\xi}\right)$（取值于 $[0,2\pi]$），是相互独立的。

解：因为 $\begin{cases} r=\sqrt{x^2+y^2}\\ \operatorname{tg}\theta=\dfrac{y}{x}\end{cases}$ 故 $\begin{cases} x=r\cos\theta\\ y=r\sin\theta\end{cases}$ $(0\le\theta\le2\pi)$

因此，

$$J=\begin{vmatrix}\frac{\partial x}{\partial r} & \frac{\partial x}{\partial\theta}\\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial\theta}\end{vmatrix}=\begin{vmatrix}\cos\theta & -r\sin\theta\\ \sin\theta & r\cos\theta\end{vmatrix}=r.$$

因为 (ξ,η) 的密度函数为

$$p(x,y)=\frac{1}{2\pi}e^{-\frac{x^2+y^2}{2}},$$

故 (ρ,φ) 的密度函数为

$$q(r,\theta)=\frac{1}{2\pi}e^{-\frac{x^2+y^2}{2}}\cdot r=\frac{1}{2\pi}\cdot re^{-\frac{r^2}{2}},\ r\ge0,0\le\theta\le2\pi.$$

因而，ρ 与 φ 的边际密度函数分别为

$$q_1(r)=\int_0^{2\pi}q(r,\theta)d\theta=re^{-\frac{r^2}{2}},\ r\ge0.$$ 称为瑞利分布

$$q_2(\theta)=\int_0^{\infty}q(r,\theta)dr=\frac{1}{2\pi},\ 0\le\theta\le2\pi.$$ 均匀分布

并且 $q(r,\theta)=q_1(r)q_2(\theta)$. 于是 ρ 与 φ 独立。

例10 设 U_1 与 U_2 相互独立，且均服从 $[0,1]$ 上的均匀分布，试证：$V_1=\sqrt{-2\ln U_1}\cos2\pi U_2$ 与 $V_2=\sqrt{-2\ln U_1}\sin2\pi U_2$ 相互独立且均服从标准正态分布。

证明：因为 $\begin{cases} y_1=\sqrt{-2\ln x_1}\cos2\pi x_2\\ y_2=\sqrt{-2\ln x_1}\sin2\pi x_2\end{cases}$ 则 $\begin{cases} y_1^2+y_2^2=-2\ln x_1\\ y_2=y_1\operatorname{tg}2\pi x_2\end{cases}$

因此 $\begin{cases} x_1=e^{-\frac{y_1^2+y_2^2}{2}}\\ x_2=\dfrac{1}{2\pi}\operatorname{arctg}\dfrac{y_2}{y_1}\end{cases}$ 故

$$J=\begin{vmatrix}\frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2}\\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2}\end{vmatrix}=\frac{1}{2\pi}e^{-\frac{y_1^2+y_2^2}{2}}.$$

因为 (U_1,U_2) 的密度函数为

$$p(x_1,x_2)=p_1(x_1)p_2(x_2)=1,\quad 0\le x_1,x_2\le1.$$

故 (V_1,V_2) 的密度函数为

$$q(y_1,y_2)=\frac{1}{2\pi}e^{-\frac{y_1^2+y_2^2}{2}}\cdot1=\frac{1}{\sqrt{2\pi}}e^{-\frac{y_1^2}{2}}\cdot\frac{1}{\sqrt{2\pi}}e^{-\frac{y_2^2}{2}}.$$

因而，V_1 与 V_2 的边际密度函数分别为

$$q_1(y_1)=\int_{-\infty}^{\infty}q(y_1,y_2)dy_2=\frac{1}{\sqrt{2\pi}}e^{-\frac{y_1^2}{2}}.$$

$$q_2(y_2)=\int_{-\infty}^{\infty}q(y_1,y_2)dy_1=\frac{1}{\sqrt{2\pi}}e^{-\frac{y_2^2}{2}}.$$

并且 $q(y_1,y_2)=q_1(y_1)q_2(y_2)$. 因而结论成立。

注：该结果常用来产生相互独立的N(0,1)随机数。