第三章　数字特征与特征函数

Tianxiao Pang

Zhejiang University

September 12, 2021

内容

内容

内容

内容

分布函数可以全面描述一个随机变量. 但在实际工作中, 完全掌握随机变量的分布函数是比较困难的.

分布函数可以全面描述一个随机变量. 但在实际工作中, 完全掌握随机变量的分布函数是比较困难的.

同时, 人们发现某些随机变量服从某类分布, 它们的一些参数可由某些数字特征确定. 因此, 对这些随机变量, 数字特征有更重要的意义.

主要的数字特征有描述随机变量取值的平均水平的数学期望(或叫均值)和描述随机变量取值相对于均值的离散程度的方差. 另外, 还有描述两个随机变量间线性关系密切程度的相关系数等.

主要的数字特征有描述随机变量取值的平均水平的数学期望(或叫均值)和描述随机变量取值相对于均值的离散程度的方差. 另外, 还有描述两个随机变量间线性关系密切程度的相关系数等.

在本章中, 我们将重点讨论数学期望、方差、相关系数这些数字特征. 同时, 我们还将讨论描述随机变量的一个强有力的工具——特征函数.

数学期望

3.1.1 离散型随机变量的数学期望

例1 为评价甲的射击技术, 随机观察甲的10次射击, 统计各次击中环数x_k和频数v_k如下表. 求他单次射击击中的平均环数(其中$N=\sum_k v_k=10$).

击中环数x_k	8	9	10
频数v_k	2	5	3
频率$f_k=v_k/N$	0.2	0.5	0.3

数学期望

3.1.1 离散型随机变量的数学期望

例1 为评价甲的射击技术，随机观察甲的10次射击，统计各次击中环数x_k和频数v_k如下表. 求他单次射击击中的平均环数(其中$N=\sum_k v_k=10$).

击中环数x_k	8	9	10
频数v_k	2	5	3
频率$f_k=v_k/N$	0.2	0.5	0.3

解: 这问题其实是求平均数$\bar{x}$. 易知

$$\begin{aligned}\bar{x}=&(8\times 2+9\times 5+10\times 3)/10\\ =&8\times 2/10+9\times 5/10+10\times 3/10\\ =&9.1.\end{aligned}$$

一般地，我们有

$$\bar{x} = (\sum_k x_k v_k)/N = \sum_k x_k \cdot v_k/N = \sum_k x_k f_k.$$

一般地, 我们有

$$\bar{x} = (\sum_k x_k v_k)/N = \sum_k x_k \cdot v_k/N = \sum_k x_k f_k.$$

若要全面考察甲的射击水平, 仅凭10次的观察是不够的. 因为频率$f_k = v_k/N$的大小与该次射击(试验)的结果有关.

一般地，我们有

$$\bar{x} = (\sum_k x_k v_k)/N = \sum_k x_k \cdot v_k/N = \sum_k x_k f_k.$$

若要全面考察甲的射击水平，仅凭10次的观察是不够的．因为频率$f_k = v_k/N$的大小与该次射击(试验)的结果有关．

但当观察次数N不断增加，由于频率稳定于概率p_k，则$\sum_k x_k f_k$稳定于$\sum_k x_k p_k$．后者是一个不依赖于具体试验的确定值，应该能更客观地表示甲的射击水平．因此，我们提出了下列的定义:

定义（离散型随机变量的数学期望）

设离散型随机变量ξ的分布列为

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

如果级数$\sum_k x_k p_k$绝对收敛，即$\sum_k |x_k| p_k < \infty$，就称$\sum_k x_k p_k$为$\xi$的数学期望(mathematical expectation)或均值(mean)．记作

$$E\xi = E(\xi) = \sum_k x_k p_k. \tag{3.1.1}$$

定义（离散型随机变量的数学期望）

设离散型随机变量ξ的分布列为

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

如果级数$\sum_k x_k p_k$绝对收敛，即$\sum_k |x_k| p_k < \infty$，就称$\sum_k x_k p_k$为$\xi$的数学期望(mathematical expectation)或均值(mean)．记作

$$E\xi = E(\xi) = \sum_k x_k p_k. \tag{3.1.1}$$

注：在定义中加了$\sum_k |x_k| p_k < \infty$的条件，是为了保证(3.1.1)的值不受求和次序的影响，从而确保了Eξ是一个确定的值．若条件$\sum_k |x_k| p_k < \infty$不被满足，则称$\xi$的数学期望不存在．

常见的离散型随机变量的数学期望:

常见的离散型随机变量的数学期望:

例2 求退化分布$\xi \equiv a$(即$\mathsf{P}(\xi = a) = 1$)的数学期望.

常见的离散型随机变量的数学期望:

例2 求退化分布$\xi \equiv a$(即$\mathsf{P}(\xi = a) = 1$)的数学期望.

解: 绝对收敛性满足. 按定义,

$$\mathsf{E}\xi = a \cdot \mathsf{P}(\xi = a) = a \cdot 1 = a.$$

即常数的数学期望就是它本身.

例3 $\xi \sim B(n,p)$. 求Eξ.

例3 $\xi \sim B(n,p)$. 求Eξ.

解: 绝对收敛性满足. 按定义,

$$
\begin{aligned}
\mathsf{E}\xi = \sum_{k=0}^{n} k\mathsf{P}(\xi = k) &= \sum_{k=0}^{n} kC_n^k p^k q^{n-k} \\
=& np\sum_{k=1}^{n} \frac{(n-1)!}{(k-1)![(n-1)-(k-1)]!} p^{k-1} q^{(n-1)-(k-1)} \\
=& np\sum_{r=0}^{n-1} \frac{(n-1)!}{r![(n-1)-r]!} p^r q^{(n-1)-r} \\
=& np(p+q)^{n-1} \\
=& np.
\end{aligned}
$$

例3 $\xi \sim B(n,p)$. 求Eξ.

解: 绝对收敛性满足. 按定义,

$$\begin{aligned}
\mathsf{E}\xi &= \sum_{k=0}^{n} k\mathsf{P}(\xi = k) = \sum_{k=0}^{n} kC_n^k p^k q^{n-k} \\
&= np\sum_{k=1}^{n} \frac{(n-1)!}{(k-1)![(n-1)-(k-1)]!} p^{k-1} q^{(n-1)-(k-1)} \\
&= np\sum_{r=0}^{n-1} \frac{(n-1)!}{r![(n-1)-r]!} p^r q^{(n-1)-r} \\
&= np(p+q)^{n-1} \\
&= np.
\end{aligned}$$

特别地, 当$n=1$时, 可知两点分布$0-1(p)$的数学期望为p.

例4 $\xi \sim P(\lambda)$. 求Eξ.

例4 $\xi \sim P(\lambda)$. 求Eξ.

解: 绝对收敛性满足. 按定义,

$$
\begin{aligned}
\mathsf{E}\xi = \sum_{k=0}^{\infty} k\mathsf{P}(\xi = k) &= \sum_{k=0}^{\infty} k\frac{\lambda^k e^{-\lambda}}{k!} \\
&= \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} \\
&= \lambda e^{-\lambda} \cdot e^{\lambda} \\
&= \lambda.
\end{aligned}
$$

例5 ξ服从几何分布, 分布列为

$$\mathsf{P}(\xi = k) = pq^{k-1}, \ \ k = 1, 2, \cdots, \ 0 < p < 1, \ q = 1 - p.$$

求$\mathsf{E}\xi$.

例5 ξ服从几何分布, 分布列为

$$\mathrm{P}(\xi=k)=pq^{k-1},\ \ k=1,2,\cdots,\ 0<p<1,\ q=1-p.$$

求$\mathrm{E}\xi$.

解: 绝对收敛性满足. 按定义,

$$\begin{aligned}
\mathrm{E}\xi &= \sum_{k=1}^{\infty} k\mathrm{P}(\xi=k) = p\sum_{k=1}^{\infty} kq^{k-1} \\
&= p\sum_{k=1}^{\infty}(x^k)'|_{x=q} = p(\sum_{k=1}^{\infty} x^k)'|_{x=q} \\
&= p(\frac{x}{1-x})'|_{x=q} = p\frac{1}{(1-x)^2}|_{x=q} \\
&= p\cdot\frac{1}{p^2} = \frac{1}{p}.
\end{aligned}$$

例6 ξ的分布列为

$$\mathrm{P}(\xi = (-1)^k 2^k / k) = 1/2^k, \quad k = 1, 2, \cdots .$$

求Eξ.

例6 ξ的分布列为

$$\mathsf{P}(\xi=(-1)^k 2^k/k)=1/2^k,\quad k=1,2,\cdots.$$

求Eξ.

解: 验证绝对收敛性:

$$\sum_{k=1}^{\infty}|x_k|p_k=\sum_{k=1}^{\infty}\frac{1}{k}=\infty.$$

所以, Eξ不存在.

例6 ξ的分布列为

$$\mathsf{P}(\xi=(-1)^k 2^k/k)=1/2^k,\quad k=1,2,\cdots.$$

求Eξ.

解: 验证绝对收敛性:

$$\sum_{k=1}^{\infty}|x_k|p_k=\sum_{k=1}^{\infty}\frac{1}{k}=\infty.$$

所以, Eξ不存在.

注: 若不验证绝对收敛性, 则会得出错误的结论:

$$\mathsf{E}\xi=\sum_{k=1}^{\infty}x_k p_k=\sum_{k=1}^{\infty}(-1)^k\frac{1}{k}<\infty.$$

3.1.2 连续型随机变量的数学期望

3.1.2 连续型随机变量的数学期望

设ξ是连续型随机变量, 密度函数为$p(x)$. 借鉴离散型随机变量的数学期望定义, 来定义连续型随机变量的数学期望.

3.1.2 连续型随机变量的数学期望

设ξ是连续型随机变量, 密度函数为$p(x)$. 借鉴离散型随机变量的数学期望定义, 来定义连续型随机变量的数学期望.

设ξ在区间$[a,b]$上取值(若ξ在区间$(-\infty,\infty)$上取值, 则令$a\to-\infty$, $b\to\infty$). 将$[a,b]$作分割:

$$a = x_0 < x_1 < \cdots < x_n = b,$$

3.1.2 连续型随机变量的数学期望

设ξ是连续型随机变量, 密度函数为$p(x)$. 借鉴离散型随机变量的数学期望定义, 来定义连续型随机变量的数学期望.

设ξ在区间$[a, b]$上取值(若ξ在区间$(-\infty, \infty)$上取值, 则令$a \to -\infty$, $b \to \infty$). 将$[a, b]$作分割:

$$a = x_0 < x_1 < \cdots < x_n = b,$$

ξ落在各小段的概率为

$$\mathsf{P}(x_k < \xi \leq x_{k+1}) = \int_{x_k}^{x_{k+1}} p(x)\mathrm{d}x \approx p(x_k)\Delta x_k, \ \ k = 0, \cdots, n-1.$$

定义一个离散型随机变量ξ_n:

$$\xi_n = x_k, \text{ 若} x_k < \xi \leq x_{k+1}.$$

定义一个离散型随机变量ξ_n:

$$\xi_n = x_k,\ \text{若} x_k < \xi \le x_{k+1}.$$

则

$$\mathsf{E}\xi_n = \sum_{k=0}^{n-1} x_k \mathsf{P}(x_k < \xi \le x_{k+1}) \approx \sum_{k=0}^{n-1} x_k p(x_k)\Delta x_k.$$

令$n \to \infty$, $\max_k\{\Delta x_k\} \to 0$, 则

$$|\xi_n - \xi| \le \max_k \Delta x_k \to 0$$

以及

$$\sum_{k=0}^{n-1} x_k p(x_k)\Delta x_k \to \int_a^b xp(x)\mathrm{d}x.$$

所以, 自然地, 我们把$\int_a^b xp(x)\mathrm{d}x$视为$\xi$的数学期望.

定义 (连续型随机变量的数学期望)

设ξ是连续型随机变量, 密度函数为$p(x)$. 则当$\int_{-\infty}^{\infty}|x|p(x)\mathrm{d}x<\infty$, 称

$$E\xi=\int_{-\infty}^{\infty}xp(x)\mathrm{d}x$$

为ξ的数学期望. 如果$\int_{-\infty}^{\infty}|x|p(x)\mathrm{d}x=\infty$, 则称$\xi$的数学期望不存在.

常见的连续型随机变量的数学期望:

常见的连续型随机变量的数学期望:

例7 $\xi \sim U[a,b]$, 求Eξ.

常见的连续型随机变量的数学期望:

例7 $\xi \sim U[a,b]$, 求Eξ.

解: 容易验证积分的绝对收敛性. 按定义:

$$\mathsf{E}\xi = \int_{-\infty}^{\infty} xp(x)\mathrm{d}x = \int_a^b x \cdot \frac{1}{b-a}\mathrm{d}x = \frac{a+b}{2}.$$

常见的连续型随机变量的数学期望:

例7 $\xi \sim U[a,b]$, 求Eξ.

解: 容易验证积分的绝对收敛性. 按定义:

$$\mathrm{E}\xi = \int_{-\infty}^{\infty} xp(x)\mathrm{d}x = \int_a^b x \cdot \frac{1}{b-a}\mathrm{d}x = \frac{a+b}{2}.$$

例8 $\xi \sim$Exp(λ), 求Eξ.

常见的连续型随机变量的数学期望:

例7 $\xi \sim U[a,b]$, 求$\mathrm{E}\xi$.

解: 容易验证积分的绝对收敛性. 按定义:

$$\mathrm{E}\xi = \int_{-\infty}^{\infty} xp(x)\mathrm{d}x = \int_a^b x \cdot \frac{1}{b-a}\mathrm{d}x = \frac{a+b}{2}.$$

例8 $\xi \sim$Exp(λ), 求$\mathrm{E}\xi$.

解: 容易验证积分的绝对收敛性. 按定义:

$$\mathrm{E}\xi = \int_{-\infty}^{\infty} xp(x)\mathrm{d}x = \int_0^{\infty} x \cdot \lambda e^{-\lambda x}\mathrm{d}x = \frac{1}{\lambda}.$$

例9 $\xi \sim N(a, \sigma^2)$, 求Eξ.

例9 $\xi \sim N(a, \sigma^2)$, 求Eξ.

解: 容易验证积分的绝对收敛性. 按定义:

$$\begin{aligned}
\mathsf{E}\xi &= \int_{-\infty}^{\infty} xp(x)\mathrm{d}x \\
&= \int_{-\infty}^{\infty} x \cdot \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-a)^2}{2\sigma^2}\right\}\mathrm{d}x \\
&\overset{z=\frac{x-a}{\sigma}}{=} \frac{a}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left\{-\frac{z^2}{2}\right\}\mathrm{d}z + \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z\exp\left\{-\frac{z^2}{2}\right\}\mathrm{d}z \\
&= a + 0 \\
&= a.
\end{aligned}$$

例10 ξ服从标准Cauchy分布，求Eξ.

例10 ξ服从标准Cauchy分布，求Eξ.

解：先来验证积分的绝对收敛性. 由于

$$\int_{-\infty}^{\infty}|x|p(x)\mathrm{d}x=2\int_{0}^{\infty}\frac{x}{\pi(1+x)^2}\mathrm{d}x=\infty.$$

所以Eξ不存在.

3.1.3 数学期望的一般定义

3.1.3 数学期望的一般定义

我们已经对离散型随机变量以及连续型随机变量分别定义了数学期望, 现在自然希望找到一种能适合一切随机变量的数学期望定义, 并把上述两种情况作为特例.

3.1.3 数学期望的一般定义

我们已经对离散型随机变量以及连续型随机变量分别定义了数学期望, 现在自然希望找到一种能适合一切随机变量的数学期望定义, 并把上述两种情况作为特例.

为了这一点, 我们需要斯梯尔吉斯(Stieltjes)积分.

若随机变量ξ的分布函数为$F(x)$, 类似于连续型随机变量的场合, 作很密的分割$x_0 < x_1 < x_2 < \cdots < x_n$, 则$\xi$落在$(x_k, x_{k+1}]$中的概率等于$F(x_{k+1}) - F(x_k) =: \Delta F(x_k)$. 定义离散型随机变量

$$\xi_n = x_k,\ \ 若 x_k < \xi \le x_{k+1}.$$

则

$$\mathsf{E}\xi_n = \sum_{k=0}^{n-1} x_k \mathsf{P}(x_k < \xi \le x_{k+1}) = \sum_{k=0}^{n-1} x_k \Delta F(x_k),$$

令$n \to \infty$, $\max_k\{\Delta x_k\} \to 0$, 则

$$|\xi_n - \xi| \le \max_k \Delta x_k \to 0,\ \sum_{k=0}^{n-1} x_k \Delta F(x_k) \to \int_{-\infty}^{\infty} x \mathrm{d}F(x).$$

这启发我们可把$\int_{-\infty}^{\infty} x\mathrm{d}F(x)$作为一般随机变量数学期望的定义.

定义（一般随机变量的数学期望）

若随机变量ξ的分布函数为$F(x)$, 则$\int_{-\infty}^{\infty}|x|\mathrm{d}F(x)<\infty$, 则称

$$E\xi=\int_{-\infty}^{\infty}x\mathrm{d}F(x)$$

为ξ的数学期望. 若$\int_{-\infty}^{\infty}|x|\mathrm{d}F(x)=\infty$, 则称$\xi$的数学期望不存在.

定义 (一般随机变量的数学期望)

若随机变量ξ的分布函数为$F(x)$, 则$\int_{-\infty}^{\infty}|x|\mathrm{d}F(x)<\infty$, 则称

$$E\xi=\int_{-\infty}^{\infty}x\mathrm{d}F(x)$$

为ξ的数学期望. 若$\int_{-\infty}^{\infty}|x|\mathrm{d}F(x)=\infty$, 则称$\xi$的数学期望不存在.

注: 对于斯梯尔吉斯积分, 有$F(x)=\int_{-\infty}^{x}\mathrm{d}F(t)$. 因此对任何随机变量$\xi$, $\mathrm{P}(a<\xi\leq b)=\int_a^b\mathrm{d}F(x)$, 概率$\mathrm{P}(\xi\in B)$可写成

$$\mathrm{P}(\xi\in B)=\int_{x\in B}\mathrm{d}F(x).$$

3.1.4 随机变量函数的数学期望

定理

设ξ是随机变量，$f(x)$是一元Borel函数，记$\eta = f(\xi)$，ξ和η的分布函数分别为$F_\xi(x)$和$F_\eta(y)$. 若η的数学期望存在，则

$$E\eta = Ef(\xi) = \int_{-\infty}^{\infty} y \mathrm{d}F_\eta(y) = \int_{-\infty}^{\infty} f(x)\mathrm{d}F_\xi(x). \qquad (3.1.2)$$

3.1.4 随机变量函数的数学期望

定理

设ξ是随机变量，$f(x)$是一元Borel函数，记$\eta = f(\xi)$，ξ和η的分布函数分别为$F_\xi(x)$和$F_\eta(y)$. 若η的数学期望存在，则

$$E\eta = Ef(\xi) = \int_{-\infty}^{\infty} y \mathrm{d}F_\eta(y) = \int_{-\infty}^{\infty} f(x)\mathrm{d}F_\xi(x). \qquad (3.1.2)$$

注: 我们无需先计算η的分布函数$F_\eta(x)$再求其数学期望，而可以直接从ξ的分布函数$F_\xi(x)$出发利用上述定理来计算. 在离散情形，(3.1.2)化为

$$\mathsf{E}f(\xi) = \sum_{i=1}^{\infty} f(x_i)p_i.$$

在连续情形，(3.1.2)化为

$$\mathsf{E}f(\xi) = \int_{-\infty}^{\infty} f(x)p(x)\mathrm{d}x,$$

我们只对特殊情形来验证上述定理:

设ξ有密度函数$p(x)$, $f(x)$严格单调增加, 它的反函数$\phi(y)$有连续导函数,

我们只对特殊情形来验证上述定理:

设ξ有密度函数$p(x)$, $f(x)$严格单调增加, 它的反函数$\phi(y)$有连续导函数, 那么可求得$\eta = f(\xi)$的密度函数为

$$g(y) = p(\phi(y))\phi^{'}(y),$$

我们只对特殊情形来验证上述定理:

设ξ有密度函数$p(x)$, $f(x)$严格单调增加, 它的反函数$\phi(y)$有连续导函数, 那么可求得$\eta = f(\xi)$的密度函数为

$$g(y) = p(\phi(y))\phi^{'}(y),$$

而$F_{\eta}^{'}(y) = g(y)$. 所以,

$$\mathsf{E}f(\xi) = \mathsf{E}\eta = \int_{-\infty}^{\infty} y \mathrm{d}F_{\eta}(y) = \int_{-\infty}^{\infty} yp(\phi(y))\phi^{'}(y)\mathrm{d}y.$$

我们只对特殊情形来验证上述定理:

设ξ有密度函数$p(x)$, $f(x)$严格单调增加, 它的反函数$\phi(y)$有连续导函数, 那么可求得$\eta = f(\xi)$的密度函数为

$$g(y) = p(\phi(y))\phi'(y),$$

而$F'_\eta(y) = g(y)$. 所以,

$$\mathsf{E}f(\xi) = \mathsf{E}\eta = \int_{-\infty}^{\infty} y\mathrm{d}F_\eta(y) = \int_{-\infty}^{\infty} yp(\phi(y))\phi'(y)\mathrm{d}y.$$

另一方面, 若令$f(x) = y$, 则$x = \phi(y)$, 上式右边等于

$$\int_{-\infty}^{\infty} f(x)p(x)\mathrm{d}x = \int_{-\infty}^{\infty} f(x)\mathrm{d}F_\xi(x).$$

所以, (3.1.2)式成立.

推广:

推广:

定理

设$(\xi_1,\cdots,\xi_n)$的分布函数为$F(x_1,\cdots,x_n)$, 而$g(x_1,\cdots,x_n)$是n元Borel函数. 若$g(\xi_1,\cdots,\xi_n)$的数学期望存在, 则

$$Eg(\xi_1,\cdots,\xi_n)=\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}g(x_1,\cdots,x_n)\mathrm{d}F(x_1,\cdots,x_n).$$

特别地, 有

$$E\xi_i=\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}x_i\mathrm{d}F(x_1,\cdots,x_n)=\int_{-\infty}^{\infty}x_i\mathrm{d}F_i(x_i),$$

其中$F_i(x)$是ξ_i的边际分布函数. 对于二元分布函数$F(x,y)$, 有

$$E(\xi\eta)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xy\mathrm{d}F(x,y),\ E\xi^2=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}x^2\mathrm{d}F(x,y).$$

例11 设某报童每日的潜在卖报数ζ服从参数为λ的Poisson分布. 如果每卖出一份报可得报酬a, 卖不掉而退回则每份赔偿b. 若某日该报童买进n份报, 试求其期望所得.

例11 设某报童每日的潜在卖报数ζ服从参数为λ的Poisson分布. 如果每卖出一份报可得报酬a, 卖不掉而退回则每份赔偿b. 若某日该报童买进n份报, 试求其期望所得.

解: 记其真正的卖报数为ξ, 则ξ与ζ的关系为

$$\xi = \begin{cases} \zeta, & \zeta < n, \\ n, & \zeta \geq n. \end{cases}$$

例11 设某报童每日的潜在卖报数ζ服从参数为λ的Poisson分布. 如果每卖出一份报可得报酬a, 卖不掉而退回则每份赔偿b. 若某日该报童买进n份报, 试求其期望所得.

解: 记其真正的卖报数为ξ, 则ξ与ζ的关系为

$$\xi = \begin{cases} \zeta, & \zeta < n, \\ n, & \zeta \geq n. \end{cases}$$

所以ξ的分布列为

$$\mathsf{P}(\xi = k) = \begin{cases} \frac{\lambda^k e^{-\lambda}}{k!}, & k = 0, \cdots, n-1, \\ \sum_{i=n}^{\infty} \frac{\lambda^i e^{-\lambda}}{i!}, & k = n. \end{cases}$$

记报童所得为η, 则η与ξ的关系为

$$\eta = g(\xi) = a\xi - b(n-\xi).$$

记报童所得为η, 则η与ξ的关系为

$$\eta = g(\xi) = a\xi - b(n-\xi).$$

因此期望所得为

$$\mathsf{E}g(\xi) = \sum_{k=0}^{n-1} \frac{\lambda^k e^{-\lambda}}{k!}[ka - (n-k)b] + na \sum_{k=n}^{\infty} \frac{\lambda^k e^{-\lambda}}{k!}.$$

例12 设$\xi \sim U[a,b]$, 求$\eta = \xi^2$的数学期望.

例12 设$\xi \sim U[a,b]$, 求$\eta=\xi^2$的数学期望.

解:

$$\begin{aligned}
\mathsf{E}\xi^2 &= \int_{-\infty}^{\infty} x^2 p(x)\mathrm{d}x \\
&= \int_a^b x^2 \frac{1}{b-a}\mathrm{d}x \\
&= \frac{a^2+ab+b^2}{3}.
\end{aligned}$$

3.1.5 数学期望的基本性质

3.1.5 数学期望的基本性质

性质1. 若$a \leq \xi \leq b$, 则$\mathsf{E}\xi$存在且$a \leq \mathsf{E}\xi \leq b$. 特别地, 若$\xi = c$, 则$\mathsf{E}\xi = \mathsf{E}c = c$.

3.1.5 数学期望的基本性质

性质1. 若$a \le \xi \le b$, 则$\mathsf{E}\xi$存在且$a \le \mathsf{E}\xi \le b$. 特别地, 若$\xi = c$, 则$\mathsf{E}\xi = \mathsf{E}c = c$.

性质2. 若$\mathsf{E}\xi_1, \cdots, \mathsf{E}\xi_n$都存在, 则对任意常数$c_1, \cdots, c_n$及$b$, $\mathsf{E}\big(\sum_{i=1}^n c_i\xi_i + b\big)$存在, 且

$$\mathsf{E}\Big(\sum_{i=1}^n c_i\xi_i + b\Big) = \sum_{i=1}^n c_i\mathsf{E}\xi_i + b.$$

特别地, 我们有

$$\mathsf{E}\Big(\sum_{i=1}^n \xi_i\Big) = \sum_{i=1}^n \mathsf{E}\xi_i, \quad \mathsf{E}(c\xi) = c\mathsf{E}\xi.$$

性质3. 若$\xi_1,\cdots,\xi_n$相互独立, 各$\mathsf{E}\xi_i$ $(i=1,\cdots,n)$存在, 则

$$\mathsf{E}(\xi_1\cdots\xi_n)=\mathsf{E}\xi_1\cdots\mathsf{E}\xi_n.$$

性质3. 若$\xi_1,\cdots,\xi_n$相互独立, 各$\mathsf{E}\xi_i\ (i=1,\cdots,n)$存在, 则

$$\mathsf{E}(\xi_1\cdots\xi_n)=\mathsf{E}\xi_1\cdots\mathsf{E}\xi_n.$$

证明: 因为

$$\begin{aligned}\mathsf{E}|\xi_1\cdots\xi_n| &= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}|x_1\cdots x_n|\mathrm{d}F(x_1,\cdots,x_n)\\ &= \int_{-\infty}^{\infty}|x_1|\mathrm{d}F_1(x_1)\cdots\int_{-\infty}^{\infty}|x_n|\mathrm{d}F_n(x_n)<\infty.\end{aligned}$$

所以$\mathsf{E}(\xi_1\cdots\xi_n)$存在. 因此,

$$\begin{aligned}\mathsf{E}(\xi_1\cdots\xi_n) &= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}x_1\cdots x_n\mathrm{d}F(x_1,\cdots,x_n)\\ &= \int_{-\infty}^{\infty}x_1\mathrm{d}F_1(x_1)\cdots\int_{-\infty}^{\infty}x_n\mathrm{d}F_n(x_n)\\ &= \mathsf{E}\xi_1\cdots\mathsf{E}\xi_n.\end{aligned}$$

定理（Markov不等式）

对于随机变量ξ, 对任意的$\varepsilon > 0$都有

$$P(|\xi| \geq \varepsilon) \leq \frac{E|\xi|}{\varepsilon}.$$

定理（Markov不等式）

对于随机变量ξ, 对任意的$\varepsilon > 0$都有

$$P(|\xi| \geq \varepsilon) \leq \frac{E|\xi|}{\varepsilon}.$$

证明: 令

$$\eta = \begin{cases} 1, & 若|\xi| \geq \varepsilon, \\ 0, & 若|\xi| < \varepsilon. \end{cases}$$

定理（Markov不等式）

对于随机变量ξ, 对任意的$\varepsilon > 0$都有

$$P(|\xi| \geq \varepsilon) \leq \frac{E|\xi|}{\varepsilon}.$$

证明：令

$$\eta = \begin{cases} 1, & 若|\xi| \geq \varepsilon, \\ 0, & 若|\xi| < \varepsilon. \end{cases}$$

则$\eta - |\xi|/\varepsilon \leq 0.$

定理（Markov不等式）

对于随机变量ξ, 对任意的$\varepsilon > 0$都有

$$P(|\xi| \geq \varepsilon) \leq \frac{E|\xi|}{\varepsilon}.$$

证明: 令

$$\eta = \begin{cases} 1, & \text{若} |\xi| \geq \varepsilon, \\ 0, & \text{若} |\xi| < \varepsilon. \end{cases}$$

则$\eta - |\xi|/\varepsilon \leq 0$. 由性质1和性质2可得

$$\mathrm{P}(|\xi| \geq \varepsilon) = \mathrm{E}\eta \leq \mathrm{E}\left(\frac{|\xi|}{\varepsilon}\right) = \frac{\mathrm{E}|\xi|}{\varepsilon}.$$

定理（Cauchy-Schwarz不等式）

对任意的随机变量ξ与η都有

$$|E(\xi\eta)| \le \sqrt{E\xi^2 \cdot E\eta^2}. \tag{3.1.3}$$

等号成立当且仅当

$$P(\eta = t_0\xi) = 1, \tag{3.1.4}$$

这里t_0是某一个常数.

定理（Cauchy-Schwarz不等式）

对任意的随机变量ξ与η都有

$$|E(\xi\eta)| \leq \sqrt{E\xi^2 \cdot E\eta^2}. \tag{3.1.3}$$

等号成立当且仅当

$$P(\eta = t_0\xi) = 1, \tag{3.1.4}$$

这里t_0是某一个常数.

证明: 对任意实数t, 定义

$$f(t) = \mathrm{E}(t\xi - \eta)^2 = t^2\mathrm{E}\xi^2 - 2t\mathrm{E}(\xi\eta) + \mathrm{E}\eta^2.$$

定理（Cauchy-Schwarz不等式）

对任意的随机变量ξ与η都有

$$|E(\xi\eta)| \le \sqrt{E\xi^2 \cdot E\eta^2}. \tag{3.1.3}$$

等号成立当且仅当

$$P(\eta = t_0\xi) = 1, \tag{3.1.4}$$

这里t_0是某一个常数.

证明: 对任意实数t, 定义

$$f(t) = \mathsf{E}(t\xi - \eta)^2 = t^2\mathsf{E}\xi^2 - 2t\mathsf{E}(\xi\eta) + \mathsf{E}\eta^2.$$

显然对一切t, 都有$f(t) \ge 0$. 因此二次方程$f(t) = 0$或者没有实根或者有一个重根. 所以

$$[\mathsf{E}(\xi\eta)]^2 - \mathsf{E}\xi^2 \cdot \mathsf{E}\eta^2 \le 0,$$

(3.1.3)得证.

此外, 方程$f(t)=0$有一个重根t_0的充要条件是

$$[\mathsf{E}(\xi\eta)]^2-\mathsf{E}\xi^2\cdot\mathsf{E}\eta^2=0,$$

这时$\mathsf{E}(t_0\xi-\eta)^2=0$. 因此, 由Markov不等式,

$$\begin{aligned}
&\mathsf{P}(t_0\xi-\eta=0)=1-\mathsf{P}(|t_0\xi-\eta|>0)\\
=&1-\lim_{n\to\infty}\mathsf{P}(|t_0\xi-\eta|^2\geq 1/n^2)\\
\geq&1-\lim_{n\to\infty}n^2\mathsf{E}(t_0\xi-\eta)^2\\
=&1,
\end{aligned}$$

(3.1.4)得证.

定理（Jensen不等式）

设X为非负随机变量，$g(x)$是定义在$[0,\infty)$上的Borel可测的凸函数，如果EX存在，则有

$$g(EX) \le Eg(X). \tag{3.1.5}$$

定理（Jensen不等式）

设X为非负随机变量，$g(x)$是定义在$[0,\infty)$上的Borel可测的凸函数，如果EX存在，则有

$$g(EX) \le Eg(X). \tag{3.1.5}$$

*证明：若X退化到0，则(3.1.5)两边都是$g(0)$，不等式自然成立. 若X不退化到0，则必有$0 < \mathrm{E}X < \infty$，此时由凸函数的性质知(何琛等编，数学分析，1985，pp.183)：存在实数c，使得

$$g(x) \ge c(x - \mathrm{E}X) + g(\mathrm{E}X), \quad x \in [0,\infty),$$

于是有

$$g(X) \ge c(X - \mathrm{E}X) + g(\mathrm{E}X),$$

上式两端同取数学期望，即得(3.1.5).

例13 设$\xi \sim B(n,p)$, 求Eξ.

例13 设$\xi \sim B(n,p)$, 求Eξ.

解: 设计一个Bernoulli试验, 记

$$\xi_i = \begin{cases} 1, & 第i次试验A发生, \\ 0, & 第i次试验A不发生. \end{cases}$$

此时, $\xi_i \sim 0-1(p)$分布, $p = \mathrm{P}(A)$, $\mathrm{E}\xi_i = p$且$\xi = \sum_{i=1}^n \xi_i$. 所以,

$$\mathrm{E}\xi = \sum_{i=1}^n \mathrm{E}\xi_i = np.$$

例14 设ξ服从超几何分布, 分布列为

$$\mathsf{P}(\xi=m)=C_M^m C_{N-M}^{n-m}/C_N^n,\ \ m=0,1,\cdots,\min(n,M),$$

求$\mathsf{E}\xi$.

例14 设ξ服从超几何分布, 分布列为

$$\mathsf{P}(\xi = m) = C_M^m C_{N-M}^{n-m} / C_N^n,\ \ m = 0, 1, \cdots, \min(n, M),$$

求$\mathsf{E}\xi$.

解: 设计一个不放回抽样. 令ξ_i为第i次抽取时的废品数, 则$\xi = \sum_{i=1}^n \xi_i$. 同时可知

$$\mathsf{P}(\xi_i = 1) = M/N,\ \ i = 1, 2, \cdots, n.$$

因此

$$\mathsf{E}\xi = \sum_{i=1}^n \mathsf{E}\xi_i = nM/N.$$

例15 设正随机变量$\xi_1,\cdots,\xi_n$独立同分布, 有密度函数$p(x)$. 试证对任意的$1\leq k\leq n$,

$$\mathsf{E}\left(\frac{\xi_1+\cdots+\xi_k}{\xi_1+\cdots+\xi_n}\right)=\frac{k}{n}.$$

例15 设正随机变量$\xi_1,\cdots,\xi_n$独立同分布, 有密度函数$p(x)$. 试证对任意的$1\le k\le n$,

$$\mathsf{E}\left(\frac{\xi_1+\cdots+\xi_k}{\xi_1+\cdots+\xi_n}\right)=\frac{k}{n}.$$

证明: 因为$\xi_1,\cdots,\xi_n$为正随机变量, 所以$0<\frac{\xi_i}{\xi_1+\cdots+\xi_n}<1$. 因此$\frac{\xi_i}{\xi_1+\cdots+\xi_n}$的数学期望存在且

$$\begin{aligned}
&\mathsf{E}\left(\frac{\xi_i}{\xi_1+\cdots+\xi_n}\right)\\
=&\int_0^\infty\cdots\int_0^\infty\frac{x_i}{x_1+\cdots+x_n}p(x_1)\cdots p(x_n)\mathrm{d}x_1\cdots\mathrm{d}x_n\\
=&\int_0^\infty\cdots\int_0^\infty\frac{y_1}{y_1+\cdots+y_n}p(y_1)\cdots p(y_n)\mathrm{d}y_1\cdots\mathrm{d}y_n\\
=&\mathsf{E}\left(\frac{\xi_1}{\xi_1+\cdots+\xi_n}\right).
\end{aligned}$$

又因为

$$\mathsf{E}\left(\frac{\xi_1}{\xi_1+\cdots+\xi_n}\right)+\cdots+\mathsf{E}\left(\frac{\xi_n}{\xi_1+\cdots+\xi_n}\right)=\mathsf{E}\left(\frac{\xi_1+\cdots+\xi_n}{\xi_1+\cdots+\xi_n}\right)=1,$$

所以

$$\mathsf{E}\left(\frac{\xi_1}{\xi_1+\cdots+\xi_n}\right)=\cdots=\mathsf{E}\left(\frac{\xi_n}{\xi_1+\cdots+\xi_n}\right)=\frac{1}{n}.$$

由此可得

$$\mathsf{E}\left(\frac{\xi_1+\cdots+\xi_k}{\xi_1+\cdots+\xi_n}\right)=\mathsf{E}\left(\frac{\xi_1}{\xi_1+\cdots+\xi_n}\right)+\cdots+\mathsf{E}\left(\frac{\xi_k}{\xi_1+\cdots+\xi_n}\right)=\frac{k}{n}.$$

例16 在一个人数很多的单位中普查某疾病, N个人验血, 可用两种方法: (1)每个人化验一次, 共需化验N次. (2)k个人的血混合化验, 如果结果是不含该病菌, 说明k个人都无该病, 这样k个人化验一次即可; 如果结果是含该病菌, 则该组每个人再分别化验一次, k个人共化验$k+1$次. 问哪一种方法可减少化验次数?

例16 在一个人数很多的单位中普查某疾病，N个人验血，可用两种方法：(1)每个人化验一次，共需化验N次．(2)k个人的血混合化验，如果结果是不含该病菌，说明k个人都无该病，这样k个人化验一次即可；如果结果是含该病菌，则该组每个人再分别化验一次，k个人共化验$k+1$次．问哪一种方法可减少化验次数？

解：记用第二种方法化验时，k人中每人所需化验次数为ξ．

例16 在一个人数很多的单位中普查某疾病, N个人验血, 可用两种方法: (1)每个人化验一次, 共需化验N次. (2)k个人的血混合化验, 如果结果是不含该病菌, 说明k个人都无该病, 这样k个人化验一次即可; 如果结果是含该病菌, 则该组每个人再分别化验一次, k个人共化验$k+1$次. 问哪一种方法可减少化验次数?

解: 记用第二种方法化验时, k人中每人所需化验次数为ξ. 第一种情况是k个人的血混合化验一次, 每个人化验$\xi = 1/k$次, 概率为

$$\mathsf{P}(\xi = 1/k) = (1-p)^k.$$

第二种情况是k个人化验$k+1$次, 每人化验$\xi = 1 + 1/k$次, 概率为

$$\mathsf{P}(\xi = 1 + 1/k) = 1 - (1-p)^k.$$

所以,

$$\begin{aligned} \mathsf{E}\xi =&(1/k)\cdot(1-p)^k+(1+1/k)\cdot(1-(1-p)^k) \\ =&1-(1-p)^k+1/k. \end{aligned}$$

当$1/k-(1-p)^k<0$时, $\mathsf{E}\xi<1$, 平均起来能减少化验次数.

*例17 设$A_1, A_2, \cdots, A_n$为来自同一概率空间中的任意n个事件，使得$\mathsf{P}\big(\bigcup_{k=1}^n A_k\big) > 0$. 证明：利用数学期望的Cauchy-Schwarz不等式(即：$|\mathsf{E}(\xi\eta)| \le \sqrt{\mathsf{E}\xi^2 \cdot \mathsf{E}\eta^2}$)证明下列的概率不等式

$$\mathsf{P}\Big(\bigcup_{k=1}^n A_k\Big) \ge \frac{\Big(\sum_{i=1}^n \mathsf{P}(A_i)\Big)^2}{\sum\limits_{i=1}^n \mathsf{P}(A_i) + 2\sum\limits_{1\le i<j\le n} \mathsf{P}(A_iA_j)}.$$

*例17 设$A_1, A_2, \cdots, A_n$为来自同一概率空间中的任意n个事件，使得$\mathsf{P}\left(\bigcup_{k=1}^n A_k\right) > 0$. 证明：利用数学期望的Cauchy-Schwarz不等式(即：$|\mathsf{E}(\xi\eta)| \le \sqrt{\mathsf{E}\xi^2 \cdot \mathsf{E}\eta^2}$)证明下列的概率不等式

$$\mathsf{P}\left(\bigcup_{k=1}^n A_k\right) \ge \frac{\left(\sum_{i=1}^n \mathsf{P}(A_i)\right)^2}{\sum_{i=1}^n \mathsf{P}(A_i) + 2\sum_{1\le i<j\le n} \mathsf{P}(A_iA_j)}.$$

证明：令

$$X_i = \begin{cases} 1, & \text{若}A_i\text{发生}, \\ 0, & \text{若}A_i\text{不发生}, \end{cases} \qquad i = 1, 2, \cdots, n.$$

由示性函数的性质和Cauchy-Schwarz不等式知

$$
\begin{aligned}
\Big(\sum_{i=1}^{n}\mathsf{P}(A_i)\Big)^2 =&\Big(\mathsf{E}\sum_{i=1}^{n}X_i\Big)^2 = \Big[\mathsf{E}\Big(I\Big(\sum_{i=1}^{n}X_i>0\Big)\sum_{i=1}^{n}X_i\Big)\Big]^2\\
\le&\mathsf{P}\Big(\sum_{i=1}^{n}X_i>0\Big)\cdot\mathsf{E}\Big(\sum_{i=1}^{n}X_i\Big)^2\\
=&\mathsf{P}\Big(\bigcup_{k=1}^{n}A_k\Big)\cdot\mathsf{E}\Big(\sum_{i=1}^{n}X_i\Big)^2,
\end{aligned}
$$

而

$$
\begin{aligned}
\mathsf{E}\Big(\sum_{i=1}^{n}X_i\Big)^2 =&\Big(\sum_{i=1}^{n}\mathsf{E}X_i^2\Big)+2\sum_{1\le i<j\le n}\mathsf{E}(X_iX_j)\\
=&\sum_{i=1}^{n}\mathsf{P}(A_i)+2\sum_{1\le i<j\le n}\mathsf{P}(A_iA_j).
\end{aligned}
$$

将后一式子代入前一式子，整理后即得所证.

3.1.6 条件数学期望(conditional expectation)

3.1.6 条件数学期望(conditional expectation)

在第二章第四节中曾引入条件分布列和条件概率密度函数的概念, 它们具有通常的分布列和概率密度函数的一切性质. 因此也可以关于它们求数学期望, 称为条件数学期望.

3.1.6 条件数学期望(conditional expectation)

在第二章第四节中曾引入条件分布列和条件概率密度函数的概念, 它们具有通常的分布列和概率密度函数的一切性质. 因此也可以关于它们求数学期望, 称为条件数学期望.

- 设在$\xi = x$的条件下, η有条件概率密度函数$p_{\eta|\xi}(y|x)$. 如果$\int_{-\infty}^{\infty} |y| p_{\eta|\xi}(y|x)\mathrm{d}y < \infty$. 则它的条件数学期望为

$$\mathrm{E}(\eta|\xi = x) = \int_{-\infty}^{\infty} y p_{\eta|\xi}(y|x)\mathrm{d}y. \tag{3.1.6}$$

3.1.6 条件数学期望(conditional expectation)

在第二章第四节中曾引入条件分布列和条件概率密度函数的概念, 它们具有通常的分布列和概率密度函数的一切性质. 因此也可以关于它们求数学期望, 称为条件数学期望.

- 设在$\xi = x$的条件下, η有条件概率密度函数$p_{\eta|\xi}(y|x)$. 如果$\int_{-\infty}^{\infty} |y| p_{\eta|\xi}(y|x) \mathrm{d}y < \infty$. 则它的条件数学期望为

$$\mathsf{E}(\eta|\xi = x) = \int_{-\infty}^{\infty} y p_{\eta|\xi}(y|x) \mathrm{d}y. \tag{3.1.6}$$

- 设在$\xi = x$的条件下, η有条件分布列$\mathsf{P}(\eta = y_j|\xi = x_i)$, $j = 1, 2, \cdots$. 如果$\sum_{j=1}^{\infty} |y_j| \mathsf{P}(\eta = y_j|\xi = x_i) < \infty$. 则它的条件数学期望为

$$\mathsf{E}(\eta|\xi = x_i) = \sum_{j=1}^{\infty} y_j \mathsf{P}(\eta = y_j|\xi = x_i). \tag{3.1.7}$$

例18 若$(\xi,\eta)\sim N(a,b,\sigma_1^2,\sigma_2^2,r)$. 求$\mathsf{E}(\eta|\xi=x)$.

例18 若$(\xi,\eta)\sim N(a,b,\sigma_1^2,\sigma_2^2,r)$. 求$\mathsf{E}(\eta|\xi=x)$.

解: 在第二章中, 我们曾得到:

$$p_{\eta|\xi}(y|x)=\frac{1}{\sqrt{2\pi}\sigma_2\sqrt{1-r^2}}\exp\Big\{-\frac{[y-b-\frac{r\sigma_2}{\sigma_1}(x-a)]^2}{2\sigma_2^2(1-r^2)}\Big\},\ y\in\mathbb{R}.$$

例18 若$(\xi, \eta) \sim N(a, b, \sigma_1^2, \sigma_2^2, r)$. 求$\mathsf{E}(\eta|\xi = x)$.

解: 在第二章中, 我们曾得到:

$$p_{\eta|\xi}(y|x) = \frac{1}{\sqrt{2\pi}\sigma_2\sqrt{1-r^2}} \exp\left\{ -\frac{[y - b - \frac{r\sigma_2}{\sigma_1}(x-a)]^2}{2\sigma_2^2(1-r^2)} \right\}, \ y \in \mathbb{R}.$$

代入(3.1.6), 得

$$\mathsf{E}(\eta|\xi = x) = b + r\frac{\sigma_2}{\sigma_1}(x-a).$$

若以$\mathsf{E}(\eta|\xi)$记ξ的如下函数：当$\xi = x$时它取值$g(x) = \mathsf{E}(\eta|\xi = x)$，这样定义的$g(\xi) = \mathsf{E}(\eta|\xi)$是一个随机变量．对$g(\xi)$求数学期望，我们得到如下的结论：

$$\mathsf{E}[\mathsf{E}(\eta|\xi)] = \mathsf{E}\eta. \tag{3.1.8}$$

若以$\mathsf{E}(\eta|\xi)$记ξ的如下函数：当$\xi=x$时它取值$g(x)=\mathsf{E}(\eta|\xi=x)$，这样定义的$g(\xi)=\mathsf{E}(\eta|\xi)$是一个随机变量．对$g(\xi)$求数学期望，我们得到如下的结论：

$$\mathsf{E}[\mathsf{E}(\eta|\xi)]=\mathsf{E}\eta. \tag{3.1.8}$$

当ξ是离散型随机变量时，记$p_i=\mathsf{P}(\xi=x_i)$，则(3.1.8)成为

$$\mathsf{E}\eta=\mathsf{E}g(\xi)=\sum_{i=1}^{\infty}g(x_i)\mathsf{P}(\xi=x_i)=\sum_{i=1}^{\infty}p_i\mathsf{E}(\eta|\xi=x_i). \tag{3.1.9}$$

我们称上式为全数学期望公式(total expectation formula).

我们对连续型随机变量给出(3.1.8)证明：设(ξ,η)有联合密度函数$p(x,y)$，此时ξ有边际密度函数$p_\xi(x)$，则

$$p_{\eta|\xi}(y|x) = p(x,y)/p_\xi(x) \quad (p_\xi(x) \neq 0),$$

$$g(x) = \mathsf{E}(\eta|\xi = x) = \int_{-\infty}^{\infty} yp(x,y)/p_\xi(x)\mathrm{d}y. \quad (3.1.10)$$

我们对连续型随机变量给出(3.1.8)证明：设(ξ,η)有联合密度函数$p(x,y)$，此时ξ有边际密度函数$p_\xi(x)$，则

$$p_{\eta|\xi}(y|x) = p(x,y)/p_\xi(x) \quad (p_\xi(x)\neq 0),$$
$$g(x) = \mathrm{E}(\eta|\xi = x) = \int_{-\infty}^{\infty} yp(x,y)/p_\xi(x)\mathrm{d}y. \quad (3.1.10)$$

由函数的数学期望公式(3.1.2)得：

$$\begin{aligned}\mathrm{E}[\mathrm{E}(\eta|\xi)] =& \mathrm{E}g(\xi) = \int_{-\infty}^{\infty} g(x)p_\xi(x)\mathrm{d}x\\ =& \int_{-\infty}^{\infty}\Big(\int_{-\infty}^{\infty} yp(x,y)/p_\xi(x)\mathrm{d}y\Big)p_\xi(x)\mathrm{d}x\\ =& \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} yp(x,y)\mathrm{d}y\mathrm{d}x\\ =& \mathrm{E}\eta.\end{aligned}$$

由(3.1.10)知, 若ξ与η独立, 则对任意的x,

$$\begin{aligned}\mathsf{E}(\eta|\xi=x) &= \int_{-\infty}^{\infty} yp(x,y)/p_\xi(x)\mathrm{d}y \\ &= \int_{-\infty}^{\infty} yp_\eta(y)\mathrm{d}y \\ &= \mathsf{E}\eta.\end{aligned}$$

所以在独立情形下:

$$\mathsf{E}(\eta|\xi) = \mathsf{E}\eta.$$

例19 设$\{\xi_k, k \geq 1\}$是同分布随机变量序列，服从二项分布$B(n,p)$，ν服从Poisson分布$P(\lambda)$，且与$\{\xi_k, k \geq 1\}$独立. 求$\mathrm{E}(\sum_{k=1}^{\nu} \xi_k)$.

例19 设$\{\xi_k, k \geq 1\}$是同分布随机变量序列, 服从二项分布$B(n,p)$, ν服从Poisson分布$P(\lambda)$, 且与$\{\xi_k, k \geq 1\}$独立. 求$\mathrm{E}(\sum_{k=1}^{\nu} \xi_k)$.

解: 记$\eta = \sum_{k=1}^{\nu} \xi_k$, 则

$\mathrm{E}(\eta|\nu = r) =$

例19 设$\{\xi_k, k \geq 1\}$是同分布随机变量序列，服从二项分布$B(n,p)$，ν服从Poisson分布$P(\lambda)$，且与$\{\xi_k, k \geq 1\}$独立. 求$\mathsf{E}(\sum_{k=1}^{\nu}\xi_k)$.

解：记$\eta = \sum_{k=1}^{\nu}\xi_k$，则

$$\mathsf{E}(\eta|\nu = r) = \mathsf{E}\left(\sum_{k=1}^{r}\xi_k|\nu = r\right) = \mathsf{E}\left(\sum_{k=1}^{r}\xi_k\right) = \sum_{k=1}^{r}\mathsf{E}\xi_k = rnp.$$

例19 设$\{\xi_k, k\geq 1\}$是同分布随机变量序列，服从二项分布$B(n,p)$，ν服从Poisson分布$P(\lambda)$，且与$\{\xi_k, k\geq 1\}$独立. 求$\mathsf{E}(\sum_{k=1}^{\nu}\xi_k)$.

解：记$\eta=\sum_{k=1}^{\nu}\xi_k$，则

$$\mathsf{E}(\eta|\nu=r)=\mathsf{E}\left(\sum_{k=1}^{r}\xi_k|\nu=r\right)=\mathsf{E}\left(\sum_{k=1}^{r}\xi_k\right)=\sum_{k=1}^{r}\mathsf{E}\xi_k=rnp.$$

所以由全数学期望公式(3.1.9)，

$$\begin{aligned}\mathsf{E}\eta=&\sum_{r=0}^{\infty}\mathsf{E}(\eta|\nu=r)\mathsf{P}(\nu=r)\\=&np\sum_{r=0}^{\infty}r\mathsf{P}(\nu=r)\\=&np\cdot\mathsf{E}\nu\\=&np\lambda.\end{aligned}$$

方差、协方差与相关系数

3.2.1 方差

方差、协方差与相关系数

3.2.1 方差

例1 比较甲乙两人的射击技术，已知两人每次击中环数ξ,η的概率分布为

$$\xi:\begin{pmatrix} 7 & 8 & 9 \\ 0.1 & 0.8 & 0.1 \end{pmatrix},\qquad \eta:\begin{pmatrix} 6 & 7 & 8 & 9 & 10 \\ 0.1 & 0.2 & 0.4 & 0.2 & 0.1 \end{pmatrix}.$$

问哪一个射击技术较好?

方差、协方差与相关系数

3.2.1 方差

例1 比较甲乙两人的射击技术，已知两人每次击中环数ξ,η的概率分布为

$$\xi:\begin{pmatrix} 7 & 8 & 9 \\ 0.1 & 0.8 & 0.1 \end{pmatrix},\qquad \eta:\begin{pmatrix} 6 & 7 & 8 & 9 & 10 \\ 0.1 & 0.2 & 0.4 & 0.2 & 0.1 \end{pmatrix}.$$

问哪一个射击技术较好?

解: $\mathsf{E}\xi=\mathsf{E}\eta=8$, 从均值来看无法分辨孰优孰劣.

方差、协方差与相关系数

3.2.1 方差

例1 比较甲乙两人的射击技术，已知两人每次击中环数ξ,η的概率分布为

$$\xi:\begin{pmatrix} 7 & 8 & 9 \\ 0.1 & 0.8 & 0.1 \end{pmatrix}, \qquad \eta:\begin{pmatrix} 6 & 7 & 8 & 9 & 10 \\ 0.1 & 0.2 & 0.4 & 0.2 & 0.1 \end{pmatrix}.$$

问哪一个射击技术较好?

解: $\mathsf{E}\xi=\mathsf{E}\eta=8$, 从均值来看无法分辨孰优孰劣. 但直观上看, 甲基本稳定在8环左右, 乙有时会击中10环, 有时6环, 较不稳定. 所以直观上应认为甲的射击技术较好.

方差、协方差与相关系数

3.2.1 方差

例1 比较甲乙两人的射击技术，已知两人每次击中环数ξ,η的概率分布为

$$\xi:\begin{pmatrix} 7 & 8 & 9 \\ 0.1 & 0.8 & 0.1 \end{pmatrix},\qquad \eta:\begin{pmatrix} 6 & 7 & 8 & 9 & 10 \\ 0.1 & 0.2 & 0.4 & 0.2 & 0.1 \end{pmatrix}.$$

问哪一个射击技术较好?

解: $\mathsf{E}\xi=\mathsf{E}\eta=8$, 从均值来看无法分辨孰优孰劣. 但直观上看, 甲基本稳定在8环左右, 乙有时会击中10环, 有时6环, 较不稳定. 所以直观上应认为甲的射击技术较好.

这个例子说明, 对于一随机变量, 除考虑它的平均取值外, 还要考虑它取值的稳定性(即离散程度).

称$\xi - E\xi$为随机变量ξ对于均值$E\xi$的离差(deviation), 它是一随机变量.

称$\xi - \mathsf{E}\xi$为随机变量ξ对于均值$\mathsf{E}\xi$的离差(deviation)，它是一随机变量. 我们可先尝试用$\mathsf{E}(\xi - \mathsf{E}\xi)$来度量$\xi$取值的离散程度. 但由于对任意的数学期望存在的$\xi$, $\mathsf{E}(\xi - \mathsf{E}\xi) = 0$. 因此用$\mathsf{E}(\xi - \mathsf{E}\xi)$来度量$\xi$取值的离散程度是无效的.

称$\xi - \mathsf{E}\xi$为随机变量ξ对于均值$\mathsf{E}\xi$的离差(deviation), 它是一随机变量. 我们可先尝试用$\mathsf{E}(\xi - \mathsf{E}\xi)$来度量$\xi$取值的离散程度. 但由于对任意的数学期望存在的$\xi$, $\mathsf{E}(\xi - \mathsf{E}\xi) = 0$. 因此用$\mathsf{E}(\xi - \mathsf{E}\xi)$来度量$\xi$取值的离散程度是无效的. 我们考虑用$\mathsf{E}(\xi - \mathsf{E}\xi)^2$ ($\mathsf{E}|\xi - \mathsf{E}\xi|$在数学处理上不大方便)来描述$\xi$取值的离散程度, 这就是方差.

称$\xi - \mathsf{E}\xi$为随机变量ξ对于均值$\mathsf{E}\xi$的离差(deviation)，它是一随机变量．我们可先尝试用$\mathsf{E}(\xi - \mathsf{E}\xi)$来度量$\xi$取值的离散程度．但由于对任意的数学期望存在的$\xi$，$\mathsf{E}(\xi - \mathsf{E}\xi) = 0$．因此用$\mathsf{E}(\xi - \mathsf{E}\xi)$来度量$\xi$取值的离散程度是无效的．我们考虑用$\mathsf{E}(\xi - \mathsf{E}\xi)^2$ ($\mathsf{E}|\xi - \mathsf{E}\xi|$在数学处理上不大方便)来描述$\xi$取值的离散程度，这就是方差．

定义（方差）

若$E(\xi - E\xi)^2 < \infty$，就称它是随机变量ξ的方差(*variance*)，记作$Var\xi$(或$D\xi$)，即

$$Var\xi = E(\xi - E\xi)^2.$$

称$\xi-\mathrm{E}\xi$为随机变量ξ对于均值$\mathrm{E}\xi$的离差(deviation), 它是一随机变量. 我们可先尝试用$\mathrm{E}(\xi-\mathrm{E}\xi)$来度量$\xi$取值的离散程度. 但由于对任意的数学期望存在的$\xi$, $\mathrm{E}(\xi-\mathrm{E}\xi)=0$. 因此用$\mathrm{E}(\xi-\mathrm{E}\xi)$来度量$\xi$取值的离散程度是无效的. 我们考虑用$\mathrm{E}(\xi-\mathrm{E}\xi)^2$ ($\mathrm{E}|\xi-\mathrm{E}\xi|$在数学处理上不大方便)来描述$\xi$取值的离散程度, 这就是方差.

定义 (方差)

若$E(\xi-E\xi)^2<\infty$, 就称它是随机变量ξ的方差(*variance*), 记作$Var\xi$(或$D\xi$), 即
$$Var\xi=E(\xi-E\xi)^2.$$

注: $\mathrm{Var}\xi$与ξ的量纲不一致, 为了统一量纲, 有时采用$\sqrt{\mathrm{Var}\xi}$, 称为ξ的标准差(standard deviation).

方差的计算公式:

方差的计算公式:

$$
\begin{aligned}
\mathsf{Var}\xi &= \int_{-\infty}^{\infty} (x - \mathsf{E}\xi)^2 \mathsf{d}F_\xi(x) \\
&= \begin{cases} \sum_i (x_i - \mathsf{E}\xi)^2 \mathsf{P}(\xi = x_i) & (\text{离散型}), \\ \int_{-\infty}^{\infty} (x - \mathsf{E}\xi)^2 p_\xi(x) \mathsf{d}x & (\text{连续型}). \end{cases}
\end{aligned}
$$

方差的计算公式:

$$
\begin{aligned}
\mathsf{Var}\xi &= \int_{-\infty}^{\infty} (x-\mathsf{E}\xi)^2 \mathrm{d}F_\xi(x) \\
&= \begin{cases} \sum_i (x_i - \mathsf{E}\xi)^2 \mathsf{P}(\xi = x_i) & \text{(离散型)}, \\ \int_{-\infty}^{\infty} (x-\mathsf{E}\xi)^2 p_\xi(x)\mathrm{d}x & \text{(连续型)}. \end{cases}
\end{aligned}
$$

另外,

$$
\mathsf{Var}\xi = \mathsf{E}(\xi-\mathsf{E}\xi)^2 = \mathsf{E}\left[\xi^2 - 2\xi\mathsf{E}\xi + (\mathsf{E}\xi)^2\right] = \mathsf{E}\xi^2 - (\mathsf{E}\xi)^2.
$$

例1(续) 计算ξ与η的方差.

$$\xi:\begin{pmatrix} 7 & 8 & 9 \\ 0.1 & 0.8 & 0.1 \end{pmatrix}, \qquad \eta:\begin{pmatrix} 6 & 7 & 8 & 9 & 10 \\ 0.1 & 0.2 & 0.4 & 0.2 & 0.1 \end{pmatrix}.$$

例1(续) 计算ξ与η的方差.

$$\xi:\begin{pmatrix} 7 & 8 & 9 \\ 0.1 & 0.8 & 0.1 \end{pmatrix}, \qquad \eta:\begin{pmatrix} 6 & 7 & 8 & 9 & 10 \\ 0.1 & 0.2 & 0.4 & 0.2 & 0.1 \end{pmatrix}.$$

解:

$$\begin{aligned}\mathsf{E}\xi^2 =& \sum_i x_i^2 \mathsf{P}(\xi = x_i) \\ =& 7^2 \times 0.1 + 8^2 \times 0.8 + 9^2 \times 0.1 = 64.2,\end{aligned}$$

所以

$$\mathsf{Var}\xi = \mathsf{E}\xi^2 - (\mathsf{E}\xi)^2 = 64.2 - 8^2 = 0.2.$$

同理可得

$$\mathsf{Var}\eta = \mathsf{E}\eta^2 - (\mathsf{E}\eta)^2 = 65.2 - 8^2 = 1.2.$$

例1(续) 计算ξ与η的方差.

$$\xi : \begin{pmatrix} 7 & 8 & 9 \\ 0.1 & 0.8 & 0.1 \end{pmatrix}, \qquad \eta : \begin{pmatrix} 6 & 7 & 8 & 9 & 10 \\ 0.1 & 0.2 & 0.4 & 0.2 & 0.1 \end{pmatrix}.$$

解:

$$\begin{aligned} \mathsf{E}\xi^2 =& \sum_i x_i^2 \mathsf{P}(\xi = x_i) \\ =& 7^2 \times 0.1 + 8^2 \times 0.8 + 9^2 \times 0.1 = 64.2, \end{aligned}$$

所以

$$\mathsf{Var}\xi = \mathsf{E}\xi^2 - (\mathsf{E}\xi)^2 = 64.2 - 8^2 = 0.2.$$

同理可得

$$\mathsf{Var}\eta = \mathsf{E}\eta^2 - (\mathsf{E}\eta)^2 = 65.2 - 8^2 = 1.2.$$

所以η取值较ξ分散. 这说明甲的射击技术较好.

一些重要的随机变量的方差:

一些重要的随机变量的方差:

例2 设$\xi \sim 0-1(p)$分布, 求Varξ.

一些重要的随机变量的方差:

例2 设$\xi \sim 0-1(p)$分布, 求$\mathsf{Var}\xi$.

解: 已知$\mathsf{E}\xi = p$. 又

$$\mathsf{E}\xi^2 = 1^2 \times p + 0^2 \times q = p.$$

一些重要的随机变量的方差:

例2 设$\xi \sim 0-1(p)$分布, 求$\mathrm{Var}\xi$.

解: 已知$\mathrm{E}\xi = p$. 又

$$\mathrm{E}\xi^2 = 1^2 \times p + 0^2 \times q = p.$$

所以

$$\mathrm{Var}\xi = \mathrm{E}\xi^2 - (\mathrm{E}\xi)^2 = p - p^2 = pq.$$

例3 设$\xi \sim P(\lambda)$分布, 求Varξ.

例3 设$\xi \sim P(\lambda)$分布, 求Varξ.

解: 已知E$\xi = \lambda$. 又

$$
\begin{aligned}
\mathsf{E}\xi^2 &= \sum_{k=0}^{\infty} k^2 \frac{\lambda^k}{k!} e^{-\lambda} = \sum_{k=1}^{\infty} k \frac{\lambda^k}{(k-1)!} e^{-\lambda} \\
&= \sum_{k=1}^{\infty} (k-1) \frac{\lambda^k}{(k-1)!} e^{-\lambda} + \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} e^{-\lambda} \\
&= \lambda \sum_{j=0}^{\infty} j \frac{\lambda^j}{j!} e^{-\lambda} + \lambda \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} e^{-\lambda} \\
&= \lambda^2 + \lambda.
\end{aligned}
$$

例3 设$\xi \sim P(\lambda)$分布, 求Varξ.

解: 已知E$\xi = \lambda$. 又

$$\begin{aligned}
\mathrm{E}\xi^2 =& \sum_{k=0}^{\infty} k^2 \frac{\lambda^k}{k!} e^{-\lambda} = \sum_{k=1}^{\infty} k \frac{\lambda^k}{(k-1)!} e^{-\lambda} \\
=& \sum_{k=1}^{\infty} (k-1) \frac{\lambda^k}{(k-1)!} e^{-\lambda} + \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} e^{-\lambda} \\
=& \lambda \sum_{j=0}^{\infty} j \frac{\lambda^j}{j!} e^{-\lambda} + \lambda \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} e^{-\lambda} \\
=& \lambda^2 + \lambda.
\end{aligned}$$

所以

$$\mathrm{Var}\xi = \mathrm{E}\xi^2 - (\mathrm{E}\xi)^2 = \lambda^2 + \lambda - \lambda^2 = \lambda.$$

例4 ξ服从几何分布, 分布列为

$$\mathsf{P}(\xi = k) = pq^{k-1}, \quad k = 1, 2, \cdots, \; 0 < p < 1, \; q = 1 - p.$$

求Varξ.

例4 ξ服从几何分布，分布列为

$$\mathsf{P}(\xi = k) = pq^{k-1}, \quad k = 1, 2, \cdots, \ 0 < p < 1, \ q = 1 - p.$$

求$\mathsf{Var}\xi$.

解：已知道$\mathsf{E}\xi = 1/p$，而

$$\begin{aligned}
\mathsf{E}\xi^2 = \sum_{k=1}^{\infty} k^2 \mathsf{P}(\xi = k) &= p \sum_{k=1}^{\infty} k^2 q^{k-1} \\
&= p\Big(\sum_{k=1}^{\infty} kx^k\Big)'\Big|_{x=q} \\
&= p\Big(\frac{x}{(1-x)^2}\Big)'\Big|_{x=q} = p\frac{1+x}{(1-x)^3}\Big|_{x=q} \\
&= p \cdot \frac{1+q}{p^3} = \frac{2-p}{p^2}.
\end{aligned}$$

所以

$$\mathsf{Var}\xi = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2}.$$

例5 设$\xi \sim U[a,b]$分布，求$\mathrm{Var}\xi$.

例5 设$\xi \sim U[a,b]$分布, 求Varξ.

解: 已知E$\xi = \frac{1}{2}(a+b)$. 又

$$\mathrm{E}\xi^2 = \int_a^b x^2 \frac{1}{b-a} \mathrm{d}x = \frac{1}{3}(a^2+ab+b^2).$$

例5 设$\xi\sim U[a,b]$分布，求$\mathrm{Var}\xi$.

解：已知$\mathrm{E}\xi=\frac{1}{2}(a+b)$. 又

$$\mathrm{E}\xi^2=\int_a^b x^2\frac{1}{b-a}\mathrm{d}x=\frac{1}{3}(a^2+ab+b^2).$$

所以

$$\mathrm{Var}\xi=\mathrm{E}\xi^2-(\mathrm{E}\xi)^2=\frac{1}{3}(a^2+ab+b^2)-\frac{1}{4}(a+b)^2=\frac{1}{12}(b-a)^2.$$

例6 设$\xi \sim \mathrm{Exp}(\lambda)$分布, 求$\mathrm{Var}\xi$.

例6 设$\xi \sim \mathsf{Exp}(\lambda)$分布, 求$\mathsf{Var}\xi$.

解: 因为

$$\begin{aligned}\mathsf{E}\xi^2 &= \int_0^\infty x^2 \cdot \lambda e^{-\lambda x} \mathrm{d}x \\ &= -x^2 e^{-\lambda x}\big|_0^\infty + 2\int_0^\infty x e^{-\lambda x} \mathrm{d}x \\ &= \frac{2}{\lambda}\mathsf{E}\xi = \frac{2}{\lambda^2}.\end{aligned}$$

例6 设$\xi \sim \mathrm{Exp}(\lambda)$分布，求$\mathrm{Var}\xi$.

解：因为

$$
\begin{aligned}
\mathrm{E}\xi^2 &= \int_0^\infty x^2 \cdot \lambda e^{-\lambda x} \mathrm{d}x \\
&= -x^2 e^{-\lambda x}|_0^\infty + 2\int_0^\infty x e^{-\lambda x} \mathrm{d}x \\
&= \frac{2}{\lambda}\mathrm{E}\xi = \frac{2}{\lambda^2}.
\end{aligned}
$$

所以

$$
\mathrm{Var}\xi = \mathrm{E}\xi^2 - (\mathrm{E}\xi)^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}.
$$

例7 设$\xi \sim N(a, \sigma^2)$分布，求Varξ.

例7 设$\xi \sim N(a,\sigma^2)$分布，求Varξ.

解：由于E$\xi = a$，所以

$$\begin{aligned}
\mathrm{Var}\xi &= \mathrm{E}(\xi-a)^2 = \int_{-\infty}^{\infty}(x-a)^2\frac{1}{\sqrt{2\pi}\sigma}\exp\Big\{-\frac{(x-a)^2}{2\sigma^2}\Big\}\mathrm{d}x \\
&\overset{z=\frac{x-a}{\sigma}}{=} \frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{\infty} z^2 e^{-z^2/2}\mathrm{d}z \\
&= \frac{\sigma^2}{\sqrt{2\pi}}\Big(-ze^{-z^2/2}\big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} e^{-z^2/2}\mathrm{d}z\Big) \\
&= \frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-z^2/2}\mathrm{d}z \\
&= \sigma^2.
\end{aligned}$$

定理（Chebyshev不等式）

设ξ为随机变量，则对任意给定的$\varepsilon > 0$，恒有

$$P(|\xi - E\xi| \geq \varepsilon) \leq \frac{Var\xi}{\varepsilon^2}.$$

定理（Chebyshev不等式）

设ξ为随机变量，则对任意给定的$\varepsilon > 0$，恒有

$$P(|\xi - E\xi| \geq \varepsilon) \leq \frac{Var\xi}{\varepsilon^2}.$$

证明：设ξ的分布函数为$F(x)$，则

$$\begin{aligned}
\mathrm{P}(|\xi - \mathrm{E}\xi| \geq \varepsilon) &= \int_{|x-\mathrm{E}\xi|\geq\varepsilon} \mathrm{d}F(x) \\
&\leq \int_{|x-\mathrm{E}\xi|\geq\varepsilon} \frac{(x-\mathrm{E}\xi)^2}{\varepsilon^2} \mathrm{d}F(x) \\
&\leq \frac{1}{\varepsilon^2}\int_{-\infty}^{\infty} (x-\mathrm{E}\xi)^2 \mathrm{d}F(x) \\
&= \frac{\mathrm{Var}\xi}{\varepsilon^2}.
\end{aligned}$$

定理（Chebyshev不等式）

设ξ为随机变量，则对任意给定的$\varepsilon > 0$，恒有

$$P(|\xi - E\xi| \geq \varepsilon) \leq \frac{Var\xi}{\varepsilon^2}.$$

证明：设ξ的分布函数为$F(x)$，则

$$\begin{aligned}
\mathrm{P}(|\xi - \mathrm{E}\xi| \geq \varepsilon) &= \int_{|x-\mathrm{E}\xi|\geq\varepsilon} \mathrm{d}F(x) \\
&\leq \int_{|x-\mathrm{E}\xi|\geq\varepsilon} \frac{(x-\mathrm{E}\xi)^2}{\varepsilon^2} \mathrm{d}F(x) \\
&\leq \frac{1}{\varepsilon^2} \int_{-\infty}^{\infty} (x-\mathrm{E}\xi)^2 \mathrm{d}F(x) \\
&= \frac{\mathrm{Var}\xi}{\varepsilon^2}.
\end{aligned}$$

注：事实上，可根据Markov不等式直接证得Chebyshev不等式.

Chebyshev不等式只用数学期望和方差就可对随机变量落在某些区间的概率进行估计. 利用该不等式, 可知

$$\mathsf{P}(|\xi - \mathsf{E}\xi| \le 3\sqrt{\mathsf{Var}\xi}) \ge$$

Chebyshev不等式只用数学期望和方差就可对随机变量落在某些区间的概率进行估计. 利用该不等式, 可知

$$\mathsf{P}(|\xi - \mathsf{E}\xi| \leq 3\sqrt{\mathsf{Var}\xi}) \geq 1 - \mathsf{Var}\xi/(3\sqrt{\mathsf{Var}\xi})^2 = 8/9 \approx 0.89.$$

这里, 我们也可以看出Chebyshev不等式不是一个非常精细的不等式, 因为在正态分布情形, 我们已知道

$$\mathsf{P}(|\xi - \mathsf{E}\xi| \leq 3\sqrt{\mathsf{Var}\xi}) \approx$$

Chebyshev不等式只用数学期望和方差就可对随机变量落在某些区间的概率进行估计. 利用该不等式, 可知

$$\mathsf{P}(|\xi - \mathsf{E}\xi| \le 3\sqrt{\mathsf{Var}\xi}) \ge 1 - \mathsf{Var}\xi/(3\sqrt{\mathsf{Var}\xi})^2 = 8/9 \approx 0.89.$$

这里, 我们也可以看出Chebyshev不等式不是一个非常精细的不等式, 因为在正态分布情形, 我们已知道

$$\mathsf{P}(|\xi - \mathsf{E}\xi| \le 3\sqrt{\mathsf{Var}\xi}) \approx 0.9973.$$

但在很多场合, 这个不够精细的不等式却已够用、管用.

方差的性质:

方差的性质:

性质1: $\mathrm{Var}\xi = 0$的充要条件是$\mathrm{P}(\xi = c) = 1$, 其中c是某常数.

方差的性质:

性质1: $\mathrm{Var}\xi = 0$的充要条件是$\mathrm{P}(\xi = c) = 1$, 其中c是某常数.

证明: 先证充分性. 若$\mathrm{P}(\xi = c) = 1$, 则$\mathrm{E}\xi = c$, $\mathrm{P}(\xi - \mathrm{E}\xi = 0) = 1$. 因此

$$\mathrm{Var}\xi = \mathrm{E}(\xi - \mathrm{E}\xi)^2 = 0^2 \times \mathrm{P}(\xi - \mathrm{E}\xi = 0) = 0.$$

方差的性质:

性质1: $\mathsf{Var}\xi = 0$的充要条件是$\mathsf{P}(\xi = c) = 1$, 其中c是某常数.

证明: 先证充分性. 若$\mathsf{P}(\xi = c) = 1$, 则$\mathsf{E}\xi = c$, $\mathsf{P}(\xi - \mathsf{E}\xi = 0) = 1$. 因此

$$\mathsf{Var}\xi = \mathsf{E}(\xi - \mathsf{E}\xi)^2 = 0^2 \times \mathsf{P}(\xi - \mathsf{E}\xi = 0) = 0.$$

再证必要性. 若$\mathsf{Var}\xi = 0$, 则$\mathsf{E}\xi$存在, 记为$\mathsf{E}\xi = c$. 由Chebyshev不等式, 对任意的$\varepsilon > 0$, 均有

$$\mathsf{P}(|\xi - \mathsf{E}\xi| \geq \varepsilon) = 0.$$

从而

$$\begin{aligned}\mathsf{P}(\xi = c) =& 1 - \mathsf{P}(|\xi - c| > 0) \\ =& 1 - \lim_{n\to\infty} \mathsf{P}(|\xi - c| \geq 1/n) \\ =& 1.\end{aligned}$$

性质2: 设b,c都是常数, 则$\mathrm{Var}(c\xi+b)=c^2\mathrm{Var}\xi$.

性质2: 设b, c都是常数, 则$\mathrm{Var}(c\xi+b)=c^2\mathrm{Var}\xi$.

证明:

$$\begin{aligned}\mathrm{Var}(c\xi+b)=&\mathrm{E}[(c\xi+b)-\mathrm{E}(c\xi+b)]^2\\=&\mathrm{E}[c\xi+b-c\mathrm{E}\xi-b]^2\\=&c^2\mathrm{E}(\xi-\mathrm{E}\xi)^2\\=&c^2\mathrm{Var}\xi.\end{aligned}$$

性质3: 若$c \neq \mathsf{E}\xi$, 则$\mathsf{Var}(\xi) < \mathsf{E}(\xi - c)^2$.

性质3: 若$c \neq \mathrm{E}\xi$, 则$\mathrm{Var}(\xi) < \mathrm{E}(\xi - c)^2$.

证明: 因为

$$\mathrm{E}(\xi - c)^2 = \mathrm{E}(\xi - \mathrm{E}\xi + \mathrm{E}\xi - c)^2 = \mathrm{Var}\xi + \mathrm{E}(\mathrm{E}\xi - c)^2 \geq \mathrm{Var}\xi,$$

“$=$”成立当且仅当$c = \mathrm{E}\xi$, 得证. 性质3说明随机变量ξ对$\mathrm{E}\xi$的离散度最小.

性质4:

$$\mathrm{Var}(\sum_{i=1}^{n}\xi_i)=\sum_{i=1}^{n}\mathrm{Var}(\xi_i)+2\sum_{1\le i<j\le n}\mathrm{E}[(\xi_i-\mathrm{E}\xi_i)(\xi_j-\mathrm{E}\xi_j)].$$

特别地, 若$\xi_1,\cdots,\xi_n$两两独立, 则

$$\mathrm{Var}(\sum_{i=1}^{n}\xi_i)=\sum_{i=1}^{n}\mathrm{Var}(\xi_i).$$

证明:

$$
\begin{aligned}
\mathsf{Var}\Big(\sum_{i=1}^{n}\xi_i\Big) =& \mathsf{E}\Big[\sum_{i=1}^{n}\xi_i - \mathsf{E}\Big(\sum_{i=1}^{n}\xi_i\Big)\Big]^2 = \mathsf{E}\Big[\sum_{i=1}^{n}(\xi_i - \mathsf{E}\xi_i)\Big]^2 \\
=& \mathsf{E}\Big[\sum_{i=1}^{n}(\xi_i - \mathsf{E}\xi_i)^2 + 2\sum_{1\le i<j\le n}(\xi_i - \mathsf{E}\xi_i)(\xi_j - \mathsf{E}\xi_j)\Big] \\
=& \sum_{i=1}^{n}\mathsf{Var}(\xi_i) + 2\sum_{1\le i<j\le n}\mathsf{E}[(\xi_i - \mathsf{E}\xi_i)(\xi_j - \mathsf{E}\xi_j)].
\end{aligned}
$$

当$\xi_1, \cdots, \xi_n$两两独立时,

$$
\mathsf{E}[(\xi_i - \mathsf{E}\xi_i)(\xi_j - \mathsf{E}\xi_j)] = \mathsf{E}(\xi_i\xi_j) - \mathsf{E}\xi_i \cdot \mathsf{E}\xi_j = 0.
$$

所以

$$
\mathsf{Var}\Big(\sum_{i=1}^{n}\xi_i\Big) = \sum_{i=1}^{n}\mathsf{Var}(\xi_i).
$$

例8 设$\xi \sim B(n,p)$, 求Varξ.

例8 设$\xi \sim B(n,p)$, 求Varξ.

解: 在Bernoulli概型中, 构造独立同分布的ξ_i:

$$\xi_i = \begin{cases} 1, & \text{第}i\text{次试验}A\text{发生}, \\ 0, & \text{第}i\text{次试验}A\text{不发生}, \end{cases} \qquad i = 1, \cdots, n.$$

则$\xi = \sum_{i=1}^n \xi_i$. 同时我们已知道Var$\xi_i = pq$. 所以

$$\mathrm{Var}\xi = \mathrm{Var}(\sum_{i=1}^n \xi_i) = \sum_{i=1}^n \mathrm{Var}\xi_i = npq.$$

例9 设随机变量$\xi_1, \cdots, \xi_n$相互独立同分布, $\mathrm{E}\xi_i = a$, $\mathrm{Var}\xi_i = \sigma^2$, $i = 1, \cdots, n$. 记$\bar{\xi} = \frac{1}{n}\sum_{i=1}^{n} \xi_i$. 求$\mathrm{E}\bar{\xi}$, $\mathrm{Var}\bar{\xi}$.

例9 设随机变量$\xi_1, \cdots, \xi_n$相互独立同分布, $\mathsf{E}\xi_i = a$, $\mathsf{Var}\xi_i = \sigma^2$, $i = 1, \cdots, n$. 记$\bar{\xi} = \frac{1}{n}\sum_{i=1}^{n} \xi_i$. 求$\mathsf{E}\bar{\xi}$, $\mathsf{Var}\bar{\xi}$.

解: 利用数学期望和方差的性质可得

$$\mathsf{E}\bar{\xi} = \frac{1}{n}\sum_{i=1}^{n} \mathsf{E}\xi_i = a.$$

$$\mathsf{Var}\bar{\xi} = \frac{1}{n^2}\sum_{i=1}^{n} \mathsf{Var}\xi_i = \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n}.$$

例9 设随机变量$\xi_1,\cdots,\xi_n$相互独立同分布, $\mathsf{E}\xi_i=a$, $\mathsf{Var}\xi_i=\sigma^2$, $i=1,\cdots,n$. 记$\bar{\xi}=\frac{1}{n}\sum_{i=1}^n\xi_i$. 求$\mathsf{E}\bar{\xi}$, $\mathsf{Var}\bar{\xi}$.

解: 利用数学期望和方差的性质可得

$$\mathsf{E}\bar{\xi}=\frac{1}{n}\sum_{i=1}^n\mathsf{E}\xi_i=a.$$

$$\mathsf{Var}\bar{\xi}=\frac{1}{n^2}\sum_{i=1}^n\mathsf{Var}\xi_i=\frac{1}{n^2}n\sigma^2=\frac{\sigma^2}{n}.$$

这说明在独立同分布情形, $\bar{\xi}$的数学期望与各ξ_i的数学期望相同, 而方差只有ξ_i的$1/n$倍. 这一事实在数理统计中有重要意义.

例10 设随机变量ξ的数学期望与方差都存在, $\mathrm{Var}\xi > 0$. 令

$$\xi^* = \frac{\xi - \mathrm{E}\xi}{\sqrt{\mathrm{Var}\xi}},$$

称它为随机变量ξ的标准化. 求$\mathrm{E}\xi^*$, $\mathrm{Var}\xi^*$.

例10 设随机变量ξ的数学期望与方差都存在, $\mathrm{Var}\xi > 0$. 令

$$\xi^* = \frac{\xi - \mathrm{E}\xi}{\sqrt{\mathrm{Var}\xi}},$$

称它为随机变量ξ的标准化. 求$\mathrm{E}\xi^*$, $\mathrm{Var}\xi^*$.

解: 根据数学期望与方差的性质,

$$\mathrm{E}\xi^* = \frac{\mathrm{E}(\xi - \mathrm{E}\xi)}{\sqrt{\mathrm{Var}\xi}} = 0,$$

$$\mathrm{Var}\xi^* = \frac{\mathrm{Var}(\xi - \mathrm{E}\xi)}{\mathrm{Var}\xi} = \frac{\mathrm{Var}\xi}{\mathrm{Var}\xi} = 1.$$

3.2.2 协方差

3.2.2 协方差

对于随机向量, 我们除了关心它的每个分量的情况外, 还希望知道各个分量之间的联系, 这光靠数学期望和方差是办不到的. 因此, 我们引入下面的一个概念—协方差.

定义（协方差）

记ξ_i和ξ_j的联合分布函数为$F_{ij}(x,y)$. 若

$$E|(\xi_i - E\xi_i)(\xi_j - E\xi_j)| < \infty,$$

就称

$$E(\xi_i - E\xi_i)(\xi_j - E\xi_j) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x - E\xi_i)(y - E\xi_j)\mathrm{d}F_{ij}(x,y)$$

为ξ_i和ξ_j的协方差(*covariance*), 记作$Cov(\xi_i, \xi_j)$.

定义（协方差）

记ξ_i和ξ_j的联合分布函数为$F_{ij}(x,y)$. 若

$$E|(\xi_i - E\xi_i)(\xi_j - E\xi_j)| < \infty,$$

就称

$$E(\xi_i - E\xi_i)(\xi_j - E\xi_j) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x - E\xi_i)(y - E\xi_j)\mathrm{d}F_{ij}(x,y)$$

为ξ_i和ξ_j的协方差(*covariance*)，记作$Cov(\xi_i,\xi_j)$.

注：显然，$\mathrm{Cov}(\xi_i,\xi_i) = \mathrm{Var}\xi_i$，且方差的性质4可写为

$$\mathrm{Var}(\sum_{i=1}^{n}\xi_i) = \sum_{i=1}^{n}\mathrm{Var}(\xi_i) + 2\sum_{1\le i<j\le n}\mathrm{Cov}(\xi_i,\xi_j).$$

协方差的性质(通过简单的计算可推得前3个性质):

协方差的性质(通过简单的计算可推得前3个性质):

性质1: $\mathrm{Cov}(\xi,\eta)=\mathrm{Cov}(\eta,\xi)=\mathrm{E}(\xi\eta)-\mathrm{E}\xi\cdot\mathrm{E}\eta.$

协方差的性质(通过简单的计算可推得前3个性质):

性质1: $\mathsf{Cov}(\xi,\eta)=\mathsf{Cov}(\eta,\xi)=\mathsf{E}(\xi\eta)-\mathsf{E}\xi\cdot\mathsf{E}\eta$.

性质2: 设a,b是常数, 则$\mathsf{Cov}(a\xi,b\eta)=ab\mathsf{Cov}(\xi,\eta)$.

协方差的性质(通过简单的计算可推得前3个性质):

性质1: $\mathrm{Cov}(\xi,\eta)=\mathrm{Cov}(\eta,\xi)=\mathrm{E}(\xi\eta)-\mathrm{E}\xi\cdot\mathrm{E}\eta$.

性质2: 设a,b是常数, 则$\mathrm{Cov}(a\xi,b\eta)=ab\mathrm{Cov}(\xi,\eta)$.

性质3: $\mathrm{Cov}(\sum_{i=1}^n\xi_i,\eta)=\sum_{i=1}^n\mathrm{Cov}(\xi_i,\eta)$.

协方差的性质(通过简单的计算可推得前3个性质):

性质1: $\mathrm{Cov}(\xi,\eta)=\mathrm{Cov}(\eta,\xi)=\mathrm{E}(\xi\eta)-\mathrm{E}\xi\cdot\mathrm{E}\eta$.

性质2: 设a,b是常数, 则$\mathrm{Cov}(a\xi,b\eta)=ab\mathrm{Cov}(\xi,\eta)$.

性质3: $\mathrm{Cov}(\sum_{i=1}^n\xi_i,\eta)=\sum_{i=1}^n\mathrm{Cov}(\xi_i,\eta)$.

性质3可推广到:

$$\mathrm{Cov}(\sum_{i=1}^n\xi_i,\sum_{j=1}^m\eta_j)=\sum_{i=1}^n\sum_{j=1}^m\mathrm{Cov}(\xi_i,\eta_j).$$

对于n维随机向量$\boldsymbol{\xi} = (\xi_1, \cdots, \xi_n)'$, 定义它的协方差矩阵为

$$B = \mathsf{Cov}(\boldsymbol{\xi}) = \mathsf{E}[(\boldsymbol{\xi} - \mathsf{E}\boldsymbol{\xi})(\boldsymbol{\xi} - \mathsf{E}\boldsymbol{\xi})'] = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{pmatrix},$$

其中$b_{ij} = \mathsf{Cov}(\xi_i, \xi_j)$. 显然$B$是一个对称矩阵, 且对任何实数$t_j$, $j = 1, \cdots, n$, 二次型

$$\sum_{j,k=1}^{n} t_j t_k b_{jk} \quad =$$

对于n维随机向量$\boldsymbol{\xi}=(\xi_1,\cdots,\xi_n)'$, 定义它的协方差矩阵为

$$B=\mathsf{Cov}(\boldsymbol{\xi})=\mathsf{E}[(\boldsymbol{\xi}-\mathsf{E}\boldsymbol{\xi})(\boldsymbol{\xi}-\mathsf{E}\boldsymbol{\xi})']=\begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{pmatrix},$$

其中$b_{ij}=\mathsf{Cov}(\xi_i,\xi_j)$. 显然$B$是一个对称矩阵, 且对任何实数$t_j$, $j=1,\cdots,n$, 二次型

$$\begin{aligned}\sum_{j,k=1}^{n} t_j t_k b_{jk} &= \sum_{j,k=1}^{n} t_j t_k \mathsf{E}(\xi_j-\mathsf{E}\xi_j)(\xi_k-\mathsf{E}\xi_k) \\ &= \mathsf{E}\left[\sum_{j=1}^{n} t_j(\xi_j-\mathsf{E}\xi_j)\right]^2 \geq 0.\end{aligned}$$

所以, 随机向量$\boldsymbol{\xi}$的协方差矩阵B是非负定的.

性质4: 设

$$\boldsymbol{\xi} = (\xi_1, \cdots, \xi_n)', \quad C = \begin{pmatrix} c_{11} & \cdots & c_{1n} \\ \vdots & \ddots & \vdots \\ c_{m1} & \cdots & c_{mn} \end{pmatrix},$$

则$C\boldsymbol{\xi}$(C为常数矩阵)的协方差阵为CBC', 其中B是$\boldsymbol{\xi}$的协方差阵.

性质4: 设

$$\boldsymbol{\xi} = (\xi_1, \cdots, \xi_n)', \quad C = \begin{pmatrix} c_{11} & \cdots & c_{1n} \\ \vdots & \ddots & \vdots \\ c_{m1} & \cdots & c_{mn} \end{pmatrix},$$

则$C\boldsymbol{\xi}$(C为常数矩阵)的协方差阵为CBC', 其中B是$\boldsymbol{\xi}$的协方差阵.

证明: 事实上,

$$\begin{aligned} &\mathsf{E}[C(\boldsymbol{\xi} - \mathsf{E}\boldsymbol{\xi})(C(\boldsymbol{\xi} - \mathsf{E}\boldsymbol{\xi}))'] \\ =&\mathsf{E}[C(\boldsymbol{\xi} - \mathsf{E}\boldsymbol{\xi})(\boldsymbol{\xi} - \mathsf{E}\boldsymbol{\xi})'C'] \\ =&C\mathsf{E}[(\boldsymbol{\xi} - \mathsf{E}\boldsymbol{\xi})(\boldsymbol{\xi} - \mathsf{E}\boldsymbol{\xi})']C' \\ =&CBC'. \end{aligned}$$

3.2.3 相关系数

3.2.3 相关系数

协方差虽然在某种意义上表示了两个随机变量间的关系，但 $\mathrm{Cov}(\xi, \eta)$的取值大小与ξ, η的量纲有关. 为避免这一点，我们用ξ, η的标准化随机变量来讨论.

3.2.3 相关系数

协方差虽然在某种意义上表示了两个随机变量间的关系，但$\mathrm{Cov}(\xi,\eta)$的取值大小与ξ,η的量纲有关. 为避免这一点，我们用ξ,η的标准化随机变量来讨论.

定义（相关系数）

称

$$r_{\xi\eta} = Cov(\xi^*,\eta^*) = E(\xi^*\eta^*) = \frac{E(\xi - E\xi)(\eta - E\eta)}{\sqrt{Var\xi \cdot Var\eta}}$$

为ξ,η的相关系数(*correlation coefficient*).

相关系数的性质:

相关系数的性质:

性质1: 对相关系数$r_{\xi\eta}$, 恒有$|r_{\xi\eta}| \le 1$. $r_{\xi\eta} = 1$当且仅当

$$\mathsf{P}\left(\frac{\xi - \mathsf{E}\xi}{\sqrt{\mathsf{Var}\xi}} = \frac{\eta - \mathsf{E}\eta}{\sqrt{\mathsf{Var}\eta}}\right) = 1;$$

$r_{\xi\eta} = -1$当且仅当

$$\mathsf{P}\left(\frac{\xi - \mathsf{E}\xi}{\sqrt{\mathsf{Var}\xi}} = -\frac{\eta - \mathsf{E}\eta}{\sqrt{\mathsf{Var}\eta}}\right) = 1.$$

证明：由Cauchy-Schwarz不等式，

$$|r_{\xi\eta}| = |\mathsf{E}(\xi^*\eta^*)| \leq \sqrt{\mathsf{E}\xi^{*2} \cdot \mathsf{E}\eta^{*2}} = \sqrt{\mathsf{Var}\xi^* \cdot \mathsf{Var}\eta^*} = 1.$$

证明：由Cauchy-Schwarz不等式，

$$|r_{\xi\eta}| = |\mathsf{E}(\xi^*\eta^*)| \le \sqrt{\mathsf{E}\xi^{*2} \cdot \mathsf{E}\eta^{*2}} = \sqrt{\mathsf{Var}\xi^* \cdot \mathsf{Var}\eta^*} = 1.$$

为证其它结论，首先可知$|r_{\xi\eta}| = 1$等价于

$$|\mathsf{E}(\xi^*\eta^*)| = \sqrt{\mathsf{E}\xi^{*2} \cdot \mathsf{E}\eta^{*2}}.$$

证明：由Cauchy-Schwarz不等式，

$$|r_{\xi\eta}| = |\mathsf{E}(\xi^*\eta^*)| \le \sqrt{\mathsf{E}\xi^{*2}\cdot\mathsf{E}\eta^{*2}} = \sqrt{\mathsf{Var}\xi^*\cdot\mathsf{Var}\eta^*} = 1.$$

为证其它结论，首先可知$|r_{\xi\eta}| = 1$等价于

$$|\mathsf{E}(\xi^*\eta^*)| = \sqrt{\mathsf{E}\xi^{*2}\cdot\mathsf{E}\eta^{*2}}.$$

然后回忆Cauchy-Schwarz不等式的证明过程知：$|\mathsf{E}(\xi^*\eta^*)| = \sqrt{\mathsf{E}\xi^{*2}\cdot\mathsf{E}\eta^{*2}}$等价于$\mathsf{P}(\eta^* = t_0\xi^*) = 1$且$t_0$为

$$f(t) = t^2\mathsf{E}\xi^{*2} - 2t\mathsf{E}(\xi^*\eta^*) + \mathsf{E}\eta^{*2}$$

的重根，而

$$t_0 = \frac{2\mathsf{E}(\xi^*\eta^*)}{2\mathsf{E}\xi^{*2}} = \mathsf{E}(\xi^*\eta^*) = r_{\xi\eta}.$$

证明：由Cauchy-Schwarz不等式，

$$|r_{\xi\eta}| = |\mathsf{E}(\xi^*\eta^*)| \le \sqrt{\mathsf{E}\xi^{*2}\cdot\mathsf{E}\eta^{*2}} = \sqrt{\mathsf{Var}\xi^*\cdot\mathsf{Var}\eta^*} = 1.$$

为证其它结论，首先可知$|r_{\xi\eta}| = 1$等价于

$$|\mathsf{E}(\xi^*\eta^*)| = \sqrt{\mathsf{E}\xi^{*2}\cdot\mathsf{E}\eta^{*2}}.$$

然后回忆Cauchy-Schwarz不等式的证明过程知：$|\mathsf{E}(\xi^*\eta^*)| = \sqrt{\mathsf{E}\xi^{*2}\cdot\mathsf{E}\eta^{*2}}$等价于$\mathsf{P}(\eta^* = t_0\xi^*) = 1$且$t_0$为

$$f(t) = t^2\mathsf{E}\xi^{*2} - 2t\mathsf{E}(\xi^*\eta^*) + \mathsf{E}\eta^{*2}$$

的重根，而

$$t_0 = \frac{2\mathsf{E}(\xi^*\eta^*)}{2\mathsf{E}\xi^{*2}} = \mathsf{E}(\xi^*\eta^*) = r_{\xi\eta}.$$

所以$r_{\xi\eta} = 1$当且仅当$\mathsf{P}(\eta^* = \xi^*) = 1$; $r_{\xi\eta} = -1$当且仅当$\mathsf{P}(\eta^* = -\xi^*) = 1$.

性质1表明相关系数$r_{\xi\eta}=\pm 1$时, ξ与η以概率1存在着线性关系. 另一极端情形是$r_{\xi\eta}=0$, 此时我们称ξ与η不相关(uncorrelated).

性质1表明相关系数$r_{\xi\eta} = \pm 1$时, ξ与η以概率1存在着线性关系. 另一极端情形是$r_{\xi\eta} = 0$, 此时我们称ξ与η不相关(uncorrelated).

性质2: 对随机变量ξ与η, 下列事实等价:

(1) $\mathrm{Cov}(\xi, \eta) = 0$; (2) ξ与η不相关;
(3) $\mathrm{E}(\xi\eta) = \mathrm{E}\xi \cdot \mathrm{E}\eta$; (4) $\mathrm{Var}(\xi + \eta) = \mathrm{Var}\xi + \mathrm{Var}\eta$.

证明: 略.

性质3: 若ξ与η独立, 则ξ与η不相关.

性质3: 若ξ与η独立, 则ξ与η不相关.

证明: 若ξ与η独立, 则$\mathsf{E}(\xi\eta)=\mathsf{E}\xi\cdot\mathsf{E}\eta$. 所以由性质2知$\xi$与$\eta$不相关.

性质3: 若ξ与η独立, 则ξ与η不相关.

证明: 若ξ与η独立, 则$\mathsf{E}(\xi\eta)=\mathsf{E}\xi\cdot\mathsf{E}\eta$. 所以由性质2知$\xi$与$\eta$不相关.

注: 性质3的逆命题不真.

例11 设随机变量$\theta \sim U[0, 2\pi]$, $\xi = \cos\theta$, $\eta = \sin\theta$.

例11 设随机变量$\theta \sim U[0,2\pi]$, $\xi = \cos\theta$, $\eta = \sin\theta$.

分析: 通过简单计算得到:

$$\mathsf{E}\xi = \mathsf{E}\cos\theta = \int_0^{2\pi} \cos\varphi \cdot \frac{1}{2\pi}\mathsf{d}\varphi = 0,$$

$$\mathsf{E}\eta = \mathsf{E}\sin\theta = \int_0^{2\pi} \sin\varphi \cdot \frac{1}{2\pi}\mathsf{d}\varphi = 0,$$

$$\mathsf{E}(\xi\eta) = \mathsf{E}(\sin\theta\cdot\cos\theta) = \int_0^{2\pi} \sin\varphi\cdot\cos\varphi \cdot \frac{1}{2\pi}\mathsf{d}\varphi = 0.$$

例11 设随机变量$\theta \sim U[0,2\pi]$, $\xi = \cos\theta$, $\eta = \sin\theta$.

分析: 通过简单计算得到:

$$\mathsf{E}\xi = \mathsf{E}\cos\theta = \int_0^{2\pi} \cos\varphi \cdot \frac{1}{2\pi}\mathsf{d}\varphi = 0,$$
$$\mathsf{E}\eta = \mathsf{E}\sin\theta = \int_0^{2\pi} \sin\varphi \cdot \frac{1}{2\pi}\mathsf{d}\varphi = 0,$$
$$\mathsf{E}(\xi\eta) = \mathsf{E}(\sin\theta \cdot \cos\theta) = \int_0^{2\pi} \sin\varphi \cdot \cos\varphi \cdot \frac{1}{2\pi}\mathsf{d}\varphi = 0.$$

所以, $\mathsf{Cov}(\xi,\eta) = 0$, 即$\xi$与$\eta$不相关.

例11 设随机变量$\theta \sim U[0, 2\pi]$, $\xi = \cos\theta$, $\eta = \sin\theta$.

分析: 通过简单计算得到:

$$\begin{aligned}
\mathsf{E}\xi &= \mathsf{E}\cos\theta = \int_0^{2\pi} \cos\varphi \cdot \frac{1}{2\pi}\mathsf{d}\varphi = 0, \\
\mathsf{E}\eta &= \mathsf{E}\sin\theta = \int_0^{2\pi} \sin\varphi \cdot \frac{1}{2\pi}\mathsf{d}\varphi = 0, \\
\mathsf{E}(\xi\eta) &= \mathsf{E}(\sin\theta \cdot \cos\theta) = \int_0^{2\pi} \sin\varphi \cdot \cos\varphi \cdot \frac{1}{2\pi}\mathsf{d}\varphi = 0.
\end{aligned}$$

所以, $\mathsf{Cov}(\xi, \eta) = 0$, 即$\xi$与$\eta$不相关.

下面我们来证明ξ与η不相互独立. 若它们相互独立, 则对任意的$B_1, B_2 \in \mathscr{B}$, 有

$$\mathsf{P}(\xi \in B_1, \eta \in B_2) = \mathsf{P}(\xi \in B_1)\mathsf{P}(\eta \in B_2).$$

现在取$B_1 = B_2 = (0, 1/2)$. 则有

$$\{\xi \in B_1\} = \{0 < \cos\theta < \frac{1}{2}\} = \{\frac{\pi}{3} < \theta < \frac{\pi}{2}\} \cup \{\frac{3\pi}{2} < \theta < \frac{5\pi}{3}\},$$
$$\{\eta \in B_2\} = \{0 < \sin\theta < \frac{1}{2}\} = \{0 < \theta < \frac{\pi}{6}\} \cup \{\frac{5\pi}{6} < \theta < \pi\}.$$

这表明

$$\{\xi \in B_1\} \cap \{\eta \in B_2\} = \phi.$$

因此,

$$\mathsf{P}(\xi \in B_1) = \mathsf{P}(\eta \in B_2) = \frac{1}{6}, \quad \mathsf{P}(\xi \in B_1, \eta \in B_2) = 0.$$

所以ξ与η不相互独立.

现在取$B_1 = B_2 = (0, 1/2)$. 则有

$$\{\xi \in B_1\} = \{0 < \cos\theta < \frac{1}{2}\} = \{\frac{\pi}{3} < \theta < \frac{\pi}{2}\} \cup \{\frac{3\pi}{2} < \theta < \frac{5\pi}{3}\},$$
$$\{\eta \in B_2\} = \{0 < \sin\theta < \frac{1}{2}\} = \{0 < \theta < \frac{\pi}{6}\} \cup \{\frac{5\pi}{6} < \theta < \pi\}.$$

这表明

$$\{\xi \in B_1\} \cap \{\eta \in B_2\} = \phi.$$

因此,

$$\mathsf{P}(\xi \in B_1) = \mathsf{P}(\eta \in B_2) = \frac{1}{6}, \ \ \mathsf{P}(\xi \in B_1, \eta \in B_2) = 0.$$

所以ξ与η不相互独立. 事实上, ξ与η存在着非线性关系:

$$\xi^2 + \eta^2 = \cos^2\theta + \sin^2\theta = 1.$$

综上讨论可知: 相关系数$r_{\xi\eta}$是ξ与η的线性相关关系的一种刻画. 当$|r_{\xi\eta}|$达到1时, ξ与η之间以概率1存在线性关系. 当$|r_{\xi\eta}|$ 达到最小值0时, ξ与η不相关, 即它们之间不存在线性关系, 但是它们之间仍然可能存在其它相依关系, 所以并不一定相互独立.

综上讨论可知: 相关系数$r_{\xi\eta}$是ξ与η的线性相关关系的一种刻画. 当$|r_{\xi\eta}|$达到1时, ξ与η之间以概率1存在线性关系. 当$|r_{\xi\eta}|$ 达到最小值0时, ξ与η不相关, 即它们之间不存在线性关系, 但是它们之间仍然可能存在其它相依关系, 所以并不一定相互独立.

但是对于二维正态分布却是例外:

性质4: 对于二元正态分布, 两个分量不相关与相互独立是等价的.

性质4: 对于二元正态分布, 两个分量不相关与相互独立是等价的.

证明: 设$(\xi,\eta)\sim N(a,b,\sigma_1^2,\sigma_2^2,r)$. 在第二章, 我们已知道$\xi$与$\eta$相互独立等价于$r=0$. 而$\xi$与$\eta$的不相关等价于$r_{\xi\eta}=0$. 若能证明$r_{\xi\eta}=r$, 就能得到性质4了.

性质4: 对于二元正态分布, 两个分量不相关与相互独立是等价的.

证明: 设$(\xi,\eta)\sim N(a,b,\sigma_1^2,\sigma_2^2,r)$. 在第二章, 我们已知道$\xi$与$\eta$相互独立等价于$r=0$. 而$\xi$与$\eta$的不相关等价于$r_{\xi\eta}=0$. 若能证明$r_{\xi\eta}=r$, 就能得到性质4了.

$$
\begin{aligned}
\mathrm{Cov}(\xi,\eta) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-a)(y-b)p(x,y)\mathrm{d}x\mathrm{d}y \\
=&\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-a)(y-b) \\
&\cdot\exp\left\{-\frac{1}{2(1-r^2)}\left(\frac{x-a}{\sigma_1}-r\frac{y-b}{\sigma_2}\right)^2-\frac{(y-b)^2}{2\sigma_2^2}\right\}\mathrm{d}x\mathrm{d}y
\end{aligned}
$$

作变量替换:

$$z=\frac{x-a}{\sigma_1}-r\frac{y-b}{\sigma_2},\quad t=\frac{y-b}{\sigma_2},$$

则

$$\frac{x-a}{\sigma_1}=z+rt,\quad y-b=\sigma_2 t,\quad J=\frac{\partial(x,y)}{\partial(z,t)}=\sigma_1\sigma_2.$$

于是,

$$\begin{aligned}\mathrm{Cov}(\xi,\eta)=&\frac{\sigma_1\sigma_2}{2\pi\sqrt{1-r^2}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(zt+rt^2)\exp\Big\{-\frac{z^2}{2(1-r^2)}\Big\}\\&\cdot\exp\Big\{-\frac{t^2}{2}\Big\}\mathrm{d}z\mathrm{d}t\end{aligned}$$

$$
\begin{aligned}
=&\sigma_1\sigma_2\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} t\exp\Big\{-\frac{t^2}{2}\Big\}\mathrm{d}t \\
&\qquad\cdot\frac{1}{\sqrt{2\pi}\sqrt{1-r^2}}\int_{-\infty}^{\infty} z\exp\Big\{-\frac{z^2}{2(1-r^2)}\Big\}\mathrm{d}z \\
&+\frac{r\sigma_1\sigma_2}{\sqrt{2\pi}}\int_{-\infty}^{\infty} t^2\exp\Big\{-\frac{t^2}{2}\Big\}\mathrm{d}t \\
&\qquad\cdot\frac{1}{\sqrt{2\pi}\sqrt{1-r^2}}\int_{-\infty}^{\infty} \exp\Big\{-\frac{z^2}{2(1-r^2)}\Big\}\mathrm{d}z \\
=&r\sigma_1\sigma_2.
\end{aligned}
$$

$$
\begin{aligned}
=&\sigma_1\sigma_2\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} t\exp\Big\{-\frac{t^2}{2}\Big\}\mathrm{d}t \\
&\qquad\cdot\frac{1}{\sqrt{2\pi}\sqrt{1-r^2}}\int_{-\infty}^{\infty} z\exp\Big\{-\frac{z^2}{2(1-r^2)}\Big\}\mathrm{d}z \\
&+\frac{r\sigma_1\sigma_2}{\sqrt{2\pi}}\int_{-\infty}^{\infty} t^2\exp\Big\{-\frac{t^2}{2}\Big\}\mathrm{d}t \\
&\qquad\cdot\frac{1}{\sqrt{2\pi}\sqrt{1-r^2}}\int_{-\infty}^{\infty}\exp\Big\{-\frac{z^2}{2(1-r^2)}\Big\}\mathrm{d}z \\
=&r\sigma_1\sigma_2.
\end{aligned}
$$

所以,

$$
r_{\xi\eta}=\frac{\mathrm{Cov}(\xi,\eta)}{\sqrt{\mathrm{Var}\xi\cdot\mathrm{Var}\eta}}=\frac{r\sigma_1\sigma_2}{\sqrt{\sigma_1^2\sigma_2^2}}=r.
$$

同时我们可以得到向量$(\xi,\eta)'$的协方差矩阵为

$$\mathsf{Cov}\begin{pmatrix}\xi\\ \eta\end{pmatrix}=\begin{pmatrix}\sigma_1^2 & r\sigma_1\sigma_2\\ r\sigma_1\sigma_2 & \sigma_2^2\end{pmatrix}.$$

3.2.4 矩(moment)

3.2.4 矩(moment)

数学期望, 方差, 协方差是随机变量最常用的数字特征, 它们都是某种矩. 矩是最广泛使用的一种数字特征, 在概率论和数理统计中占有重要地位. 最常用的矩有两种: 原点矩和中心矩.

3.2.4 矩(moment)

数学期望, 方差, 协方差是随机变量最常用的数字特征, 它们都是某种矩. 矩是最广泛使用的一种数字特征, 在概率论和数理统计中占有重要地位. 最常用的矩有两种: 原点矩和中心矩.

定义 (原点矩)

对正整数k, 称

$$m_k = E\xi^k$$

为ξ的k阶(原点)矩(*origin moment*).

注: 数学期望是一阶(原点)矩. 若高阶矩存在, 则低阶矩也存在.

注: 数学期望是一阶(原点)矩. 若高阶矩存在, 则低阶矩也存在.

证明: 假设$\mathsf{E}\xi^n$存在(这意味着$\mathsf{E}|\xi|^n < \infty$). 若$k \le n$, 则

$$
\begin{aligned}
\mathsf{E}|\xi|^k &= \int_{-\infty}^{\infty} |x|^k \mathrm{d}F(x) \\
&= \int_{|x|\le 1} |x|^k \mathrm{d}F(x) + \int_{|x|>1} |x|^k \mathrm{d}F(x) \\
&\le \mathsf{P}(|\xi| \le 1) + \int_{|x|>1} |x|^n \mathrm{d}F(x) \\
&\le 1 + \mathsf{E}|\xi|^n < \infty.
\end{aligned}
$$

所以$\mathsf{E}\xi^k (k \le n)$存在. (或直接利用Jensen不等式证明)

定义（中心矩）

对正整数k, 称
$$c_k = E(\xi - E\xi)^k$$
为ξ的k阶中心矩(*center moment*).

定义（中心矩)

对正整数k, 称
$$c_k = E(\xi - E\xi)^k$$
为ξ的k阶中心矩(*center moment*).

注: 方差是二阶中心矩.

定义（中心矩）

对正整数k, 称
$$c_k = E(\xi - E\xi)^k$$
为ξ的k阶中心矩(*center moment*).

注: 方差是二阶中心矩.

此外, 三阶和四阶中心矩也是常用的. 我们称$c_3/c_2^{3/2}$为偏态系数(skewness coefficient), 当它大于0时为正偏态, 小于0时为负偏态. 称$c_4/c_2^2 - 3$为峰态系数(kurtosis coefficient), 当它大于0时表明该分布密度比正态分布更为尖峭.

例12 设$\xi \sim N(0,\sigma^2)$. 此时, $\mathsf{E}\xi = 0$且

$$m_n = c_n = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} x^n \exp\Big\{-\frac{x^2}{2\sigma^2}\Big\}\mathrm{d}x.$$

例12 设$\xi \sim N(0,\sigma^2)$. 此时, $\mathsf{E}\xi = 0$且

$$m_n = c_n = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} x^n \exp\Big\{-\frac{x^2}{2\sigma^2}\Big\}\mathrm{d}x.$$

显然, n为奇数时, $m_n = 0$;

例12 设$\xi \sim N(0,\sigma^2)$. 此时, E$\xi = 0$且

$$m_n = c_n = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} x^n \exp\Big\{-\frac{x^2}{2\sigma^2}\Big\}\mathrm{d}x.$$

显然, n为奇数时, $m_n = 0$; n为偶数时,

$$m_n =$$

例12 设$\xi \sim N(0,\sigma^2)$. 此时, $\mathsf{E}\xi = 0$且

$$m_n = c_n = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} x^n \exp\Big\{-\frac{x^2}{2\sigma^2}\Big\}\mathrm{d}x.$$

显然, n为奇数时, $m_n = 0$; n为偶数时,

$$\begin{aligned} m_n &= \sqrt{\frac{2}{\pi}}\int_0^{\infty}\frac{x^n}{\sigma}\exp\Big\{-\frac{x^2}{2\sigma^2}\Big\}\mathrm{d}x \\ &= \sqrt{\frac{2}{\pi}}\sigma^n 2^{\frac{n-1}{2}}\int_0^{\infty} z^{\frac{n-1}{2}}e^{-z}\mathrm{d}z \\ &= \sqrt{\frac{2}{\pi}}\sigma^n 2^{\frac{n-1}{2}}\Gamma\Big(\frac{n+1}{2}\Big) \\ &= 1\times 3\times\cdots\times(n-1)\sigma^n. \end{aligned}$$

例12 设$\xi \sim N(0,\sigma^2)$. 此时，E$\xi = 0$且

$$m_n = c_n = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} x^n \exp\Big\{-\frac{x^2}{2\sigma^2}\Big\}\mathrm{d}x.$$

显然，n为奇数时，$m_n = 0$； n为偶数时，

$$\begin{aligned} m_n &= \sqrt{\frac{2}{\pi}}\int_0^{\infty}\frac{x^n}{\sigma}\exp\Big\{-\frac{x^2}{2\sigma^2}\Big\}\mathrm{d}x \\ &= \sqrt{\frac{2}{\pi}}\sigma^n 2^{\frac{n-1}{2}}\int_0^{\infty} z^{\frac{n-1}{2}}e^{-z}\mathrm{d}z \\ &= \sqrt{\frac{2}{\pi}}\sigma^n 2^{\frac{n-1}{2}}\Gamma\Big(\frac{n+1}{2}\Big) \\ &= 1\times 3\times\cdots\times(n-1)\sigma^n. \end{aligned}$$

特别地，$m_4 = c_4 = 3\sigma^4$. 因此$N(0,\sigma^2)$分布的偏态系数与峰态系数都为0.

称$M_\alpha = \mathsf{E}|\xi|^\alpha$为$\xi$的$\alpha$阶绝对矩. 那么当$\xi \sim N(0, \sigma^2)$时, 有

$$\mathsf{E}|\xi|^n = \begin{cases} \sqrt{\frac{2}{\pi}}\sigma^n 2^{\frac{n-1}{2}}\Gamma(\frac{n+1}{2}) = \sqrt{\frac{2}{\pi}}2^k k!\sigma^{2k+1}, & n = 2k+1, \\ 1 \times 3 \times \cdots \times (n-1)\sigma^n, & n = 2k. \end{cases}$$

例13 若$\xi \sim \mathsf{Exp}(\lambda)$, 那么对任意的$k \geq 1$,

例13 若$\xi \sim \mathsf{Exp}(\lambda)$, 那么对任意的$k \geq 1$,

$$\mathsf{E}|\xi|^k = \mathsf{E}\xi^k \quad =$$

例13 若$\xi \sim \mathsf{Exp}(\lambda)$, 那么对任意的$k \geq 1$,

$$\begin{aligned} \mathsf{E}|\xi|^k = \mathsf{E}\xi^k &= \int_0^\infty x^k \lambda e^{-\lambda x} \mathrm{d}x = \frac{k}{\lambda}\mathsf{E}\xi^{k-1} \\ &= \cdots = \frac{k!}{\lambda^k}\mathsf{E}\xi^0 \\ &= \frac{k!}{\lambda^k}. \end{aligned}$$

例13 若$\xi \sim \mathrm{Exp}(\lambda)$, 那么对任意的$k \geq 1$,

$$
\begin{aligned}
\mathsf{E}|\xi|^k = \mathsf{E}\xi^k &= \int_0^\infty x^k \lambda e^{-\lambda x} \mathrm{d}x = \frac{k}{\lambda} \mathsf{E}\xi^{k-1} \\
&= \cdots = \frac{k!}{\lambda^k} \mathsf{E}\xi^0 \\
&= \frac{k!}{\lambda^k}.
\end{aligned}
$$

所以指数分布的任意阶矩存在.

特征函数

数字特征只反映了概率分布的某些侧面, 一般说来, 数字特征不能完全确定随机变量的分布. 本节我们将介绍特征函数这个工具, 它既能完全决定分布函数而又具有良好的分析性质.

3.3.1 定义

为了定义特征函数, 我们拓广一下随机变量的概念, 引进复随机变量.

3.3.1 定义

为了定义特征函数，我们拓广一下随机变量的概念，引进复随机变量.

定义

设ξ, η为实值随机变量，称$\zeta = \xi + \mathrm{i}\eta$为复随机变量，这里$\mathrm{i}^2 = -1$. 同时定义$E\zeta = E\xi + \mathrm{i}E\eta$为$\zeta$的数学期望.

3.3.1 定义

为了定义特征函数, 我们拓广一下随机变量的概念, 引进复随机变量.

定义

设ξ,η为实值随机变量, 称$\zeta=\xi+\mathrm{i}\eta$为复随机变量, 这里$\mathrm{i}^2=-1$. 同时定义$E\zeta=E\xi+\mathrm{i}E\eta$为$\zeta$的数学期望.

注: 对复随机变量的研究本质上是对二维随机变量的研究. 举一例: 如果二维随机变量(ξ_1,η_1)与(ξ_2,η_2)是相互独立的, 则我们称复随机变量$\zeta_1=\xi_1+\mathrm{i}\eta_1$与$\zeta_2=\xi_2+\mathrm{i}\eta_2$是相互独立的; 若$\zeta_1,\cdots,\zeta_n$是相互独立的, 则可以证明

$$\mathrm{E}(\zeta_1\cdots\zeta_n)=\mathrm{E}\zeta_1\cdots\mathrm{E}\zeta_n.$$

定义（特征函数）

设ξ为实随机变量，称

$$f(t) = Ee^{\mathrm{i}t\xi} = \int_{-\infty}^{\infty} e^{\mathrm{i}tx} \mathrm{d}F(x)$$

为ξ的特征函数(characteristic function)，这里t是任意实数.

定义 (特征函数)

设ξ为实随机变量, 称

$$f(t) = Ee^{\mathrm{i}t\xi} = \int_{-\infty}^{\infty} e^{\mathrm{i}tx} \mathrm{d}F(x)$$

为ξ的特征函数(characteristic function), 这里t是任意实数.

注: (1)对于$e^{\mathrm{i}tx}$, 我们有时候用欧拉公式来处理:

$$e^{\mathrm{i}tx} = \cos(tx) + \mathrm{i}\sin(tx).$$

(2)由于$\mathrm{E}|e^{\mathrm{i}t\xi}| = 1$, 所以对一切实数$t$特征函数都是有意义的. 也就是说, 任何随机变量都有相应的特征函数.

(3)特征函数只与分布函数有关, 因此也称为某一分布函数的特征函数.

离散型随机变量的特征函数:

离散型随机变量的特征函数: 若ξ的分布列为

$$\mathsf{P}(\xi = x_n) = p_n, n = 1, 2, \cdots .$$

则

$$f(t) = \mathsf{E}e^{\mathrm{i}t\xi} =$$

离散型随机变量的特征函数: 若ξ的分布列为

$$\mathsf{P}(\xi = x_n) = p_n, n = 1, 2, \cdots .$$

则

$$f(t) = \mathsf{E}e^{\mathrm{i}t\xi} = \sum_{n=1}^{\infty} p_n e^{\mathrm{i}tx_n}.$$

离散型随机变量的特征函数: 若ξ的分布列为

$$\mathsf{P}(\xi = x_n) = p_n, n = 1, 2, \cdots .$$

则

$$f(t) = \mathsf{E}e^{\mathrm{i}t\xi} = \sum_{n=1}^{\infty} p_n e^{\mathrm{i}tx_n}.$$

连续型随机变量的特征函数:

离散型随机变量的特征函数: 若ξ的分布列为

$$\mathsf{P}(\xi = x_n) = p_n, n = 1, 2, \cdots .$$

则

$$f(t) = \mathsf{E}e^{\mathrm{i}t\xi} = \sum_{n=1}^{\infty} p_n e^{\mathrm{i}tx_n}.$$

连续型随机变量的特征函数: 若ξ有密度函数$p(x)$, 则

$$f(t) = \mathsf{E}e^{\mathrm{i}t\xi} =$$

离散型随机变量的特征函数: 若ξ的分布列为

$$\mathrm{P}(\xi = x_n) = p_n, n = 1, 2, \cdots .$$

则

$$f(t) = \mathrm{E}e^{\mathrm{i}t\xi} = \sum_{n=1}^{\infty} p_n e^{\mathrm{i}tx_n}.$$

连续型随机变量的特征函数: 若ξ有密度函数$p(x)$, 则

$$f(t) = \mathrm{E}e^{\mathrm{i}t\xi} = \int_{-\infty}^{\infty} e^{\mathrm{i}tx} p(x) \mathrm{d}x,$$

这正好是$p(x)$的Fourier变换.

一些重要分布的特征函数:

一些重要分布的特征函数:

例1 退化分布$P(\xi = c) = 1$的特征函数为

一些重要分布的特征函数:

例1 退化分布$\mathrm{P}(\xi = c) = 1$的特征函数为$f(t) = e^{\mathrm{i}tc}$.

一些重要分布的特征函数:

例1 退化分布$P(\xi = c) = 1$的特征函数为$f(t) = e^{itc}$.

例2 二项分布$B(n, p)$的特征函数为:

一些重要分布的特征函数:

例1 退化分布$\mathrm{P}(\xi=c)=1$的特征函数为$f(t)=e^{\mathrm{i}tc}$.

例2 二项分布$B(n,p)$的特征函数为:

$$f(t)=\sum_{k=0}^{n}C_n^k p^k q^{n-k} e^{\mathrm{i}tk}=\sum_{k=0}^{n}C_n^k (pe^{\mathrm{i}t})^k q^{n-k}=(pe^{\mathrm{i}t}+q)^n.$$

一些重要分布的特征函数:

例1 退化分布$\mathrm{P}(\xi = c) = 1$的特征函数为$f(t) = e^{\mathrm{i}tc}$.

例2 二项分布$B(n,p)$的特征函数为:

$$f(t) = \sum_{k=0}^{n} C_n^k p^k q^{n-k} e^{\mathrm{i}tk} = \sum_{k=0}^{n} C_n^k (pe^{\mathrm{i}t})^k q^{n-k} = (pe^{\mathrm{i}t} + q)^n.$$

例3 $0-1(p)$分布的特征函数为

一些重要分布的特征函数:

例1 退化分布$\mathsf{P}(\xi=c)=1$的特征函数为$f(t)=e^{\mathrm{i}tc}$.

例2 二项分布$B(n,p)$的特征函数为:

$$f(t)=\sum_{k=0}^{n}C_n^k p^k q^{n-k} e^{\mathrm{i}tk}=\sum_{k=0}^{n}C_n^k (pe^{\mathrm{i}t})^k q^{n-k}=(pe^{\mathrm{i}t}+q)^n.$$

例3 $0-1(p)$分布的特征函数为$f(t)=pe^{\mathrm{i}t}+q$.

一些重要分布的特征函数:

例1 退化分布$\mathrm{P}(\xi=c)=1$的特征函数为$f(t)=e^{\mathrm{i}tc}$.

例2 二项分布$B(n,p)$的特征函数为:

$$f(t)=\sum_{k=0}^{n}C_n^k p^k q^{n-k}e^{\mathrm{i}tk}=\sum_{k=0}^{n}C_n^k(pe^{\mathrm{i}t})^k q^{n-k}=(pe^{\mathrm{i}t}+q)^n.$$

例3 $0-1(p)$分布的特征函数为$f(t)=pe^{\mathrm{i}t}+q$.

例4 泊松分布$P(\lambda)$的特征函数为:

一些重要分布的特征函数:

例1 退化分布$P(\xi = c) = 1$的特征函数为$f(t) = e^{\mathrm{i}tc}$.

例2 二项分布$B(n, p)$的特征函数为:

$$f(t) = \sum_{k=0}^{n} C_n^k p^k q^{n-k} e^{\mathrm{i}tk} = \sum_{k=0}^{n} C_n^k (pe^{\mathrm{i}t})^k q^{n-k} = (pe^{\mathrm{i}t} + q)^n.$$

例3 $0-1(p)$分布的特征函数为$f(t) = pe^{\mathrm{i}t} + q$.

例4 泊松分布$P(\lambda)$的特征函数为:

$$f(t) = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} e^{\mathrm{i}tk} = \sum_{k=0}^{\infty} \frac{(\lambda e^{\mathrm{i}t})^k}{k!} e^{-\lambda} = e^{\lambda(e^{\mathrm{i}t}-1)}.$$

例5 指数分布Exp(λ)的特征函数为$f(t) = (1 - \frac{\mathrm{i}t}{\lambda})^{-1}$.

例5 指数分布Exp(λ)的特征函数为$f(t) = (1-\frac{\mathrm{i}t}{\lambda})^{-1}$.

证:

$$
\begin{aligned}
f(t) =&\lambda\int_0^\infty e^{-\lambda x}\cos(tx)\mathrm{d}x + \lambda\mathrm{i}\int_0^\infty e^{-\lambda x}\sin(tx)\mathrm{d}x \\
:=&\lambda(J_1(t)+\mathrm{i}J_2(t)).
\end{aligned}
$$

利用分部积分得到

$$J_1(t) = \frac{1}{\lambda} - \frac{t}{\lambda}J_2(t),\ \ J_2(t) = \frac{t}{\lambda}J_1(t).$$

解此方程组得出

$$J_1(t) = \frac{\lambda}{\lambda^2+t^2},\ \ J_2(t) = \frac{t}{\lambda^2+t^2}.$$

所以指数分布Exp(λ)的特征函数为

$$f(t) = \lambda(J_1(t)+\mathrm{i}J_2(t)) = \frac{\lambda(\lambda+\mathrm{i}t)}{\lambda^2+t^2} = \frac{\lambda}{\lambda-\mathrm{i}t} = (1-\frac{\mathrm{i}t}{\lambda})^{-1}.$$

*例6 如果ξ服从标准Cauchy分布，则$f(t) = e^{-|t|}$.

*例6 如果ξ服从标准Cauchy分布，则$f(t)=e^{-|t|}$.

证:

$$\begin{aligned}
f(t) =& \frac{1}{\pi}\int_{-\infty}^{\infty} e^{itx}\frac{1}{1+x^2}\mathrm{d}x \\
=& \frac{1}{\pi}\int_{-\infty}^{\infty} \cos(tx)\frac{1}{1+x^2}\mathrm{d}x \\
=& \frac{2}{\pi}\int_{0}^{\infty} \frac{\cos(tx)}{1+x^2}\mathrm{d}x \\
=& \frac{2}{\pi}\int_{0}^{\infty} \frac{\cos(|t|x)}{1+x^2}\mathrm{d}x.
\end{aligned}$$

问题归结为计算如下积分:

$$J=\int_0^{\infty}\frac{\cos(\beta x)}{1+x^2}\mathrm{d}x,\quad \beta=|t|.$$

通过一些计算, 不难得到

$$\int_0^\infty e^{-xy}\sin y\,\mathrm{d}y = \frac{1}{1+x^2}, \quad x>0;$$
$$\int_0^\infty e^{-xy}\cos(\beta x)\,\mathrm{d}x = \frac{y}{\beta^2+y^2}.$$

于是,

$$\begin{aligned}
J &= \int_0^\infty \frac{\cos(\beta x)}{1+x^2}\mathrm{d}x = \int_0^\infty \cos(\beta x)\left(\int_0^\infty e^{-xy}\sin y\,\mathrm{d}y\right)\mathrm{d}x \\
&= \int_0^\infty \sin y\left(\int_0^\infty e^{-xy}\cos(\beta x)\mathrm{d}x\right)\mathrm{d}y \\
&\overset{y=\beta x(\beta>0)}{=} \int_0^\infty \frac{x\sin(\beta x)}{1+x^2}\mathrm{d}x.
\end{aligned}$$

所以有

$$\frac{dJ}{d\beta} = -J.$$

解此微分方程得$J = Ce^{-\beta}$. 当$\beta = 0$时, $J = C = \pi/2$. 因此

$$J = \frac{\pi}{2}e^{-\beta}.$$

回到标准Cauchy分布的特征函数, 有

$$f(t) = \frac{2}{\pi} \cdot \frac{\pi}{2}e^{-\beta} = e^{-|t|}.$$

所以有

$$\frac{dJ}{d\beta} = -J.$$

解此微分方程得$J = Ce^{-\beta}$. 当$\beta = 0$时, $J = C = \pi/2$. 因此

$$J = \frac{\pi}{2}e^{-\beta}.$$

回到标准Cauchy分布的特征函数, 有

$$f(t) = \frac{2}{\pi} \cdot \frac{\pi}{2}e^{-\beta} = e^{-|t|}.$$

其它重要分布的特征函数: 待续··.

3.3.2 性质

性质1: $|f(t)| \leq f(0) = 1,\ f(-t) = \overline{f(t)}$.

3.3.2 性质

性质1: $|f(t)| \le f(0) = 1,\ f(-t) = \overline{f(t)}$.

证: 第一个结论:

$$f(0) = \int_{-\infty}^{\infty} e^{\mathrm{i}0x} \mathrm{d}F(x) = 1,$$

而

$$|f(t)| = \left| \int_{-\infty}^{\infty} e^{\mathrm{i}tx} \mathrm{d}F(x) \right| \le \int_{-\infty}^{\infty} |e^{\mathrm{i}tx}| \mathrm{d}F(x) = 1.$$

3.3.2 性质

性质1: $|f(t)| \le f(0) = 1,\ f(-t) = \overline{f(t)}$.

证: 第一个结论:

$$f(0) = \int_{-\infty}^{\infty} e^{\mathrm{i}0x}\mathrm{d}F(x) = 1,$$

而

$$|f(t)| = \left|\int_{-\infty}^{\infty} e^{\mathrm{i}tx}\mathrm{d}F(x)\right| \le \int_{-\infty}^{\infty} |e^{\mathrm{i}tx}|\mathrm{d}F(x) = 1.$$

第二个结论:

$$f(-t) = \int_{-\infty}^{\infty} e^{-\mathrm{i}tx}\mathrm{d}F(x) = \overline{\int_{-\infty}^{\infty} e^{\mathrm{i}tx}\mathrm{d}F(x)} = \overline{f(t)}.$$

性质2: $f(t)$在$(-\infty,\infty)$上一致连续.

性质2: $f(t)$在$(-\infty,\infty)$上一致连续.

证: 对任意的$t\in(-\infty,\infty)$及$\varepsilon>0$,

$$
\begin{aligned}
|f(t+h)-f(t)| &= \left|\int_{-\infty}^{\infty}(e^{\mathrm{i}(t+h)x}-e^{\mathrm{i}tx})\mathrm{d}F(x)\right| \\
&\le \left(\int_{|x|\ge A}+\int_{|x|<A}\right)|e^{\mathrm{i}hx}-1|\mathrm{d}F(x) \\
&:= I_1+I_2.
\end{aligned}
$$

注意上式右边已与t无关.

性质2: $f(t)$在$(-\infty,\infty)$上一致连续.

证: 对任意的$t\in(-\infty,\infty)$及$\varepsilon>0$,

$$\begin{aligned}|f(t+h)-f(t)| &= \left|\int_{-\infty}^{\infty}(e^{\mathrm{i}(t+h)x}-e^{\mathrm{i}tx})\mathrm{d}F(x)\right| \\ &\le \left(\int_{|x|\ge A}+\int_{|x|<A}\right)|e^{\mathrm{i}hx}-1|\mathrm{d}F(x) \\ &:= I_1+I_2.\end{aligned}$$

注意上式右边已与t无关. 由$\int_{-\infty}^{\infty}\mathrm{d}F(x)=1$容易知道, 存在充分大的$A>0$使得

$$\int_{|x|\ge A}\mathrm{d}F(x)<\frac{\varepsilon}{4}.$$

因此

$$I_1\le\int_{|x|\ge A}(|e^{\mathrm{i}hx}|+1)\mathrm{d}F(x)=2\int_{|x|\ge A}\mathrm{d}F(x)<\frac{\varepsilon}{2}.$$

又因为

$$|e^{\mathrm{i}hx}-1| = |e^{\mathrm{i}hx/2}|\cdot|e^{\mathrm{i}hx/2}-e^{-\mathrm{i}hx/2}| = 2|\sin(hx/2)| \le |hx|,$$

所以对已取定的A, 取$\delta = \varepsilon/(2A)$. 当$|x| < A$及$0 < h < \delta$时, 有

$$|e^{\mathrm{i}hx}-1| \le |hx| \le A\delta = \frac{\varepsilon}{2}.$$

因此

$$I_2 \le \frac{\varepsilon}{2}\int_{-A}^{A} \mathrm{d}F(x) \le \frac{\varepsilon}{2}.$$

又因为

$$|e^{\mathrm{i}hx}-1|=|e^{\mathrm{i}hx/2}|\cdot|e^{\mathrm{i}hx/2}-e^{-\mathrm{i}hx/2}|=2|\sin(hx/2)|\leq|hx|,$$

所以对已取定的A，取$\delta=\varepsilon/(2A)$. 当$|x|<A$及$0<h<\delta$时，有

$$|e^{\mathrm{i}hx}-1|\leq|hx|\leq A\delta=\frac{\varepsilon}{2}.$$

因此

$$I_2\leq\frac{\varepsilon}{2}\int_{-A}^{A}\mathrm{d}F(x)\leq\frac{\varepsilon}{2}.$$

综上所述，当$0<h<\delta$时$|f(t+h)-f(t)|<\varepsilon$，且$\delta$的选取与$t$无关. 所以$f(t)$在$(-\infty,\infty)$上一致连续.

性质3: $f(t)$是非负定的: 对任意的正整数n及任意实数$t_1, \cdots, t_n$, 复数$\lambda_1, \cdots, \lambda_n$, 有

$$\sum_{k=1}^{n}\sum_{j=1}^{n} f(t_k - t_j)\lambda_k\overline{\lambda_j} \geq 0.$$

性质3: $f(t)$是非负定的: 对任意的正整数n及任意实数$t_1,\cdots,t_n$, 复数$\lambda_1,\cdots,\lambda_n$, 有

$$\sum_{k=1}^{n}\sum_{j=1}^{n} f(t_k-t_j)\lambda_k\overline{\lambda_j}\geq 0.$$

证:

$$\begin{aligned}\sum_{k=1}^{n}\sum_{j=1}^{n} f(t_k-t_j)\lambda_k\overline{\lambda_j} &=\sum_{k=1}^{n}\sum_{j=1}^{n}\mathsf{E}e^{\mathrm{i}(t_k-t_j)\xi}\lambda_k\overline{\lambda_j}\\ &=\mathsf{E}\left(\sum_{k=1}^{n}e^{\mathrm{i}t_k\xi}\lambda_k\right)\overline{\left(\sum_{j=1}^{n}e^{\mathrm{i}t_j\xi}\lambda_j\right)}\\ &=\mathsf{E}\left|\sum_{k=1}^{n}e^{\mathrm{i}t_k\xi}\lambda_k\right|^2\geq 0.\end{aligned}$$

波赫纳尔–辛钦(Bochner–Khinchine)定理(证明略):

定理

函数$f(t)$为特征函数的充要条件是$f(t)$非负定、连续且$f(0)=1$.

波赫纳尔–辛钦(Bochner–Khinchine)定理(证明略):

定理

函数$f(t)$为特征函数的充要条件是$f(t)$非负定、连续且$f(0)=1$.

这个定理在理论上给出了一个判定特征函数的方法, 但在实际应用中却是不方便使用的.

性质4: 若$\xi_1, \xi_2, \cdots, \xi_n$相互独立, $\eta = \xi_1 + \xi_2 + \cdots + \xi_n$, ξ_k的特征函数为$f_k(t)$, 则

$$f_\eta(t) = f_1(t) f_2(t) \cdots f_n(t).$$

性质4: 若$\xi_1, \xi_2, \cdots, \xi_n$相互独立, $\eta = \xi_1 + \xi_2 + \cdots + \xi_n$, ξ_k的特征函数为$f_k(t)$, 则

$$f_\eta(t) = f_1(t) f_2(t) \cdots f_n(t).$$

证: 因为$\xi_1, \cdots, \xi_n$相互独立, 所以$e^{\mathrm{i}t\xi_1}, \cdots, e^{\mathrm{i}t\xi_n}$也相互独立. 因此,

$$\mathsf{E}e^{\mathrm{i}t\eta} = \mathsf{E}(e^{\mathrm{i}t\xi_1} \cdots e^{\mathrm{i}t\xi_n}) = \mathsf{E}e^{\mathrm{i}t\xi_1} \cdots \mathsf{E}e^{\mathrm{i}t\xi_n}.$$

性质5: 若$\mathsf{E}\xi^n$存在, 则$f(t)$是n次可微的, 且当$k \le n$时,

$$f^{(k)}(0) = \mathrm{i}^k \mathsf{E}\xi^k.$$

性质5: 若$\mathsf{E}\xi^n$存在, 则$f(t)$是n次可微的, 且当$k \le n$时,

$$f^{(k)}(0) = \mathrm{i}^k \mathsf{E}\xi^k.$$

证: 由于

$$\int_{-\infty}^{\infty} \left| \frac{\mathrm{d}^k}{\mathrm{d}t^k} e^{\mathrm{i}tx} \right| \mathrm{d}F(x) = \int_{-\infty}^{\infty} \left| \mathrm{i}^k x^k e^{\mathrm{i}tx} \right| \mathrm{d}F(x)$$

$$= \int_{-\infty}^{\infty} \left| x^k \right| \mathrm{d}F(x) = \mathsf{E}|\xi^k| < \infty,$$

因此$\int_{-\infty}^{\infty} \frac{\mathrm{d}^k}{\mathrm{d}t^k} e^{\mathrm{i}tx} \mathrm{d}F(x)$对$t$一致收敛, 故$f^{(k)}(t)$存在, 且

$$\begin{aligned} f^{(k)}(t) &= \frac{\mathrm{d}^k}{\mathrm{d}t^k} \int_{-\infty}^{\infty} e^{\mathrm{i}tx} \mathrm{d}F(x) \\ &= \int_{-\infty}^{\infty} \frac{\mathrm{d}^k}{\mathrm{d}t^k} e^{\mathrm{i}tx} \mathrm{d}F(x) \\ &= \mathrm{i}^k \int_{-\infty}^{\infty} x^k e^{\mathrm{i}tx} \mathrm{d}F(x). \end{aligned}$$

所以有

$$f^{(k)}(0) = \mathrm{i}^k \int_{-\infty}^{\infty} x^k \mathrm{d}F(x) = \mathrm{i}^k \mathsf{E}\xi^k.$$

所以有

$$f^{(k)}(0) = \mathrm{i}^k \int_{-\infty}^{\infty} x^k \mathrm{d}F(x) = \mathrm{i}^k \mathsf{E}\xi^k.$$

注: 由性质5可知，当$\mathsf{E}\xi^2$存在时，有

$$\mathsf{E}\xi = -\mathrm{i}f'(0),\ \mathsf{E}\xi^2 = -f''(0),\ \mathsf{Var}\xi = -f''(0) + [f'(0)]^2.$$

性质6: 设$\eta = a\xi + b$, a, b为任意常数, 则

$$f_\eta(t) = e^{\mathrm{i}bt} f_\xi(at).$$

性质6: 设$\eta = a\xi + b$, a, b为任意常数, 则

$$f_\eta(t) = e^{\mathrm{i}bt} f_\xi(at).$$

证: $\mathsf{E}e^{\mathrm{i}t\eta} = \mathsf{E}e^{\mathrm{i}t(a\xi+b)} = e^{\mathrm{i}bt}\mathsf{E}e^{\mathrm{i}at\xi} = e^{\mathrm{i}bt} f_\xi(at).$

例7 均匀分布$U[a,b]$的特征函数为$f(t)=\frac{e^{\mathrm{i}tb}-e^{\mathrm{i}ta}}{(b-a)\mathrm{i}t}$. 特别地, $U[-b,b]$的特征函数为$f(t)=\frac{\sin(bt)}{bt}$.

例7 均匀分布$U[a,b]$的特征函数为$f(t)=\frac{e^{itb}-e^{ita}}{(b-a)it}$. 特别地, $U[-b,b]$的特征函数为$f(t)=\frac{\sin(bt)}{bt}$.

证: 假设随机变量$\xi\sim U[a,b]$. 记

$$\eta=\frac{2}{b-a}\left(\xi-\frac{a+b}{2}\right),$$

则易知$\eta\sim$

例7 均匀分布$U[a,b]$的特征函数为$f(t)=\frac{e^{\mathrm{i}tb}-e^{\mathrm{i}ta}}{(b-a)\mathrm{i}t}$. 特别地, $U[-b,b]$的特征函数为$f(t)=\frac{\sin(bt)}{bt}$.

证: 假设随机变量$\xi\sim U[a,b]$. 记

$$\eta=\frac{2}{b-a}\left(\xi-\frac{a+b}{2}\right),$$

则易知$\eta\sim U[-1,1]$. 容易求得

$$f_\eta(t)=\frac{1}{2}\int_{-1}^{1}e^{\mathrm{i}tx}\mathrm{d}x=\frac{1}{2}\int_{-1}^{1}\cos(tx)\mathrm{d}x=\frac{\sin t}{t}.$$

由于

$$\xi=\frac{a+b}{2}+\frac{b-a}{2}\eta,$$

所以由性质6得

$$f_\xi(t)=e^{\mathrm{i}\frac{a+b}{2}t}f_\eta\left(\frac{b-a}{2}t\right)=\frac{e^{\mathrm{i}tb}-e^{\mathrm{i}ta}}{(b-a)\mathrm{i}t}.$$

例8 正态分布$N(a,\sigma^2)$的特征函数为$f(t)=e^{\mathrm{i}at-\sigma^2t^2/2}$.

例8 正态分布$N(a,\sigma^2)$的特征函数为$f(t)=e^{\mathrm{i}at-\sigma^2t^2/2}$.

证: 先求$N(0,1)$的特征函数. 设$\xi\sim N(0,1)$, 则

$$f_\xi(t)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{\mathrm{i}tx}e^{-x^2/2}\mathrm{d}x=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\cos(tx)e^{-x^2/2}\mathrm{d}x.$$

因此

$$\begin{aligned}
f_\xi'(t)=&-\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}x\sin(tx)e^{-x^2/2}\mathrm{d}x\\
=&\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\sin(tx)\mathrm{d}e^{-x^2/2}\\
=&-\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}t\cos(tx)e^{-x^2/2}\mathrm{d}x\\
=&-tf_\xi(t).
\end{aligned}$$

求解以上的微分方程得

$$f_\xi(t) = Ce^{-t^2/2}.$$

又因为$f_\xi(0) = 1$, 所以$C = 1$,

$$f_\xi(t) = e^{-t^2/2}.$$

设$\eta = a + \sigma\xi$, 则$\eta \sim N(a, \sigma^2)$. 由性质6可得

$$f_\eta(t) = e^{\mathrm{i}at} f_\xi(\sigma t) = e^{\mathrm{i}at - \sigma^2 t^2/2}.$$

例9 若$\xi \sim N(a, \sigma^2)$, 用特征函数方法求Eξ, Varξ.

例9 若$\xi \sim N(a, \sigma^2)$, 用特征函数方法求Eξ, Varξ.

解: ξ的特征函数为$f(t) = e^{\mathrm{i}at - \sigma^2 t^2/2}$.

例9 若$\xi \sim N(a,\sigma^2)$, 用特征函数方法求Eξ, Varξ.

解: ξ的特征函数为$f(t) = e^{\mathrm{i}at-\sigma^2t^2/2}$. 因此

$$f'(t) = (\mathrm{i}a - \sigma^2 t)e^{\mathrm{i}at-\sigma^2t^2/2}, \quad f'(0) = \mathrm{i}a,$$
$$f''(t) = [-\sigma^2 + (\mathrm{i}a - \sigma^2 t)^2]e^{\mathrm{i}at-\sigma^2t^2/2}, \quad f''(0) = -a^2 - \sigma^2.$$

例9 若$\xi \sim N(a, \sigma^2)$, 用特征函数方法求Eξ, Varξ.

解: ξ的特征函数为$f(t) = e^{\mathrm{i}at-\sigma^2 t^2/2}$. 因此

$$f'(t) = (\mathrm{i}a - \sigma^2 t)e^{\mathrm{i}at-\sigma^2 t^2/2}, \quad f'(0) = \mathrm{i}a,$$
$$f''(t) = [-\sigma^2 + (\mathrm{i}a - \sigma^2 t)^2]e^{\mathrm{i}at-\sigma^2 t^2/2}, \quad f''(0) = -a^2 - \sigma^2.$$

故E$\xi = a$, E$\xi^2 = a^2 + \sigma^2$, Var$\xi = \sigma^2$.

例10 问: 下列函数是否是某随机变量的特征函数?
(1) $f(t) = \sin t$;
(2) $f(t) = \ln(e + |t|)$;
(3) $f(t) = 0$ 当$t < 0$, $f(t) = 1$ 当$t \geq 0$.

例10 问: 下列函数是否是某随机变量的特征函数?
(1) $f(t) = \sin t$;
(2) $f(t) = \ln(e + |t|)$;
(3) $f(t) = 0$ 当$t < 0$, $f(t) = 1$ 当$t \geq 0$.

解: (1) $f(0) = 0 \neq 1$;

例10 问: 下列函数是否是某随机变量的特征函数?

(1) $f(t) = \sin t$;

(2) $f(t) = \ln(e + |t|)$;

(3) $f(t) = 0$ 当$t < 0$, $f(t) = 1$ 当$t \geq 0$.

解: (1) $f(0) = 0 \neq 1$; (2) $|t| > 0$时, $|f(t)| > \ln e = 1$;

例10 问: 下列函数是否是某随机变量的特征函数?
(1) $f(t) = \sin t$;
(2) $f(t) = \ln(e + |t|)$;
(3) $f(t) = 0$ 当$t < 0$, $f(t) = 1$ 当$t \geq 0$.

解: (1) $f(0) = 0 \neq 1$; (2) $|t| > 0$时, $|f(t)| > \ln e = 1$; (3) $f(t)$在$t = 0$处不连续.

例10 问: 下列函数是否是某随机变量的特征函数?
(1) $f(t) = \sin t$;
(2) $f(t) = \ln(e + |t|)$;
(3) $f(t) = 0$ 当$t < 0$, $f(t) = 1$ 当$t \geq 0$.

解: (1) $f(0) = 0 \neq 1$; (2) $|t| > 0$时, $|f(t)| > \ln e = 1$; (3) $f(t)$在$t = 0$处不连续. 因此它们都不是特征函数.

3.3.3 逆转公式与唯一性定理

3.3.3 逆转公式与唯一性定理

我们已经知道随机变量的分布函数可以唯一确定它的特征函数; 反之, 从特征函数是否可以唯一确定相应的分布函数?

接下来要介绍的逆转公式和唯一性定理对这个问题作了肯定的回答.

我们先来介绍一个分析结果:

引理 (狄利克雷积分)

有

$$\lim_{x\to\infty}\int_0^x \frac{\sin au}{u}\mathrm{d}u = \frac{\pi}{2}\,\mathrm{sgn}\{a\},$$

其中$\mathrm{sgn}\{a\}$是a的符号函数.

我们先来介绍一个分析结果:

引理 (狄利克雷积分)

有

$$\lim_{x\to\infty}\int_0^x \frac{\sin au}{u}\mathrm{d}u = \frac{\pi}{2}\,\mathrm{sgn}\{a\},$$

其中$\mathrm{sgn}\{a\}$是a的符号函数.

*证: 记

$$I(a,x) = \int_0^x \frac{\sin au}{u}\mathrm{d}u,$$

通过作变量替换$y = au$, 并注意被积函数是偶函数, 便知

$$\lim_{x\to\infty} I(a,x) = \lim_{x\to\infty} \mathrm{sgn}\{a\} I(1,x),$$

所以只需证明

$$\lim_{x\to\infty} I(1,x) = \lim_{x\to\infty}\int_0^x \frac{\sin u}{u}\mathrm{d}u = \frac{\pi}{2}.$$

将

$$\frac{1}{u} = \int_0^\infty e^{-uv} \mathrm{d}v$$

代入$I(1,x)$, 并交换积分顺序得

$$\begin{aligned} I(1,x) &= \int_0^x \sin u \left(\int_0^\infty e^{-uv} \mathrm{d}v \right) \mathrm{d}u \\ &= \int_0^\infty \left(\int_0^x e^{-uv} \sin u \mathrm{d}u \right) \mathrm{d}v \\ &= \int_0^\infty \left(\frac{1}{1+v^2} - \frac{v \sin x + \cos x}{1+v^2} e^{-vx} \right) \mathrm{d}v \\ &= \frac{\pi}{2} - \int_0^\infty \frac{s \sin x + x \cos x}{x^2+s^2} e^{-s} \mathrm{d}s. \end{aligned}$$

后一个积分中的被积函数被e^{-s}所控制, 再由Lebesgue控制收敛定理可知当$x \to \infty$时, 该积分趋于0. 证毕.

引理

设$x_1 < x_2$,

$$g(T, x, x_1, x_2) = \frac{1}{\pi}\int_0^T \left[\frac{\sin t(x-x_1)}{t} - \frac{\sin t(x-x_2)}{t}\right] \mathrm{d}t \tag{3.3.1}$$

则

$$\begin{aligned} \lim_{T\to\infty} g(T, x, x_1, x_2) &\triangleq h(x, x_1, x_2) \\ &= \begin{cases} 0, & x < x_1 \text{ 或 } x > x_2, \\ \frac{1}{2}, & x = x_1 \text{ 或 } x = x_2, \\ 1, & x_1 < x < x_2. \end{cases} \end{aligned} \tag{3.3.2}$$

证: 由狄利克雷积分知

$$D(a) = \frac{1}{\pi} \lim_{T\to\infty} \int_0^T \frac{\sin at}{t} \mathrm{d}t = \frac{1}{2} \operatorname{sgn}\{a\} = \begin{cases} 1/2, & a > 0, \\ 0, & a = 0, \\ -1/2, & a < 0. \end{cases}$$

而

$$\begin{aligned} \lim_{T\to\infty} g(T, x, x_1, x_2) =& D(x - x_1) - D(x - x_2) \\ =& \frac{1}{2}[\operatorname{sgn}\{x - x_1\} - \operatorname{sgn}\{x - x_2\}], \end{aligned}$$

易知引理成立.

定理（逆转公式）

设分布函数$F(x)$的特征函数为$f(t)$，又x_1, x_2是$F(x)$的两个连续点，则

$$F(x_2) - F(x_1) = \lim_{T\to\infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-\mathrm{i}tx_1} - e^{-\mathrm{i}tx_2}}{\mathrm{i}t} f(t)\mathrm{d}t. \quad (3.3.3)$$

证: 不妨设$x_1 < x_2$. 记

$$
\begin{aligned}
I_T =& \frac{1}{2\pi}\int_{-T}^{T} \frac{e^{-\mathrm{i}tx_1} - e^{-\mathrm{i}tx_2}}{\mathrm{i}t} f(t)\mathrm{d}t \\
=& \frac{1}{2\pi}\int_{-T}^{T}\int_{-\infty}^{\infty} \frac{e^{-\mathrm{i}tx_1} - e^{-\mathrm{i}tx_2}}{\mathrm{i}t} e^{\mathrm{i}tx}\mathrm{d}F(x)dt.
\end{aligned}
$$

为证被积函数的有界性, 用到不等式

$$
|e^{\mathrm{i}\alpha} - 1| \le |\alpha|.
$$

事实上, 对$\alpha \ge 0$

$$
|e^{\mathrm{i}\alpha} - 1| = \left|\int_0^{\alpha} e^{\mathrm{i}x} dx\right| \le \int_0^{\alpha} |e^{\mathrm{i}x}|\mathrm{d}x = \alpha,
$$

对$\alpha < 0$, 取共轭即知不等式也成立. 因此,

$$
\left|\frac{e^{-\mathrm{i}tx_1} - e^{-\mathrm{i}tx_2}}{\mathrm{i}t} e^{\mathrm{i}tx}\right| \le x_2 - x_1.
$$

交换上述二次积分顺序得到

$$
\begin{aligned}
I_T =& \frac{1}{2\pi}\int_{-\infty}^{\infty}\Big[\int_{-T}^{T}\frac{e^{-\mathrm{i}tx_1}-e^{-\mathrm{i}tx_2}}{\mathrm{i}t}e^{\mathrm{i}tx}\mathrm{d}t\Big]\mathrm{d}F(x)\\
=& \frac{1}{2\pi}\int_{-\infty}^{\infty}\Big[\int_{0}^{T}\frac{e^{\mathrm{i}t(x-x_1)}-e^{-\mathrm{i}t(x-x_1)}}{\mathrm{i}t}\\
&\qquad -\frac{e^{\mathrm{i}t(x-x_2)}-e^{-\mathrm{i}t(x-x_2)}}{\mathrm{i}t}\mathrm{d}t\Big]\mathrm{d}F(x)\\
=& \frac{1}{\pi}\int_{-\infty}^{\infty}\Big[\int_{0}^{T}\left(\frac{\sin t(x-x_1)}{t}-\frac{\sin t(x-x_2)}{t}\right)\mathrm{d}t\Big]\mathrm{d}F(x)\\
=& \int_{-\infty}^{\infty}g(T,x,x_1,x_2)\mathrm{d}F(x)\\
=& \mathrm{E}g(T,\xi,x_1,x_2).
\end{aligned}
$$

此处$g(T,x,x_1,x_2)$按(3.3.1)定义. 由(3.3.2)可以知道, 对任意的$x\in\mathbb{R}$, $|g(T,x,x_1,\,x_2)|$有界. 因此由Lebesgue控制收敛定理并利用引理的结果可得(同时注意到x_1和x_2为$F(x)$的连续点):

$$
\begin{aligned}
\lim_{T\to\infty} I_T = & \lim_{T\to\infty} \mathsf{E} g(T,\xi,x_1,x_2) \\
= & \mathsf{E} \lim_{T\to\infty} g(T,\xi,x_1,x_2) \\
= & \mathsf{E} h(\xi,x_1,x_2) \\
= & 0\times \mathsf{P}(\xi<x_1 \text{ or } \xi>x_2) + \frac{1}{2}\mathsf{P}(\xi=x_1 \text{ or } \xi=x_2) \\
& + 1\times \mathsf{P}(x_1<\xi<x_2) \\
= & \mathsf{P}(x_1<\xi<x_2) \\
= & F(x_2-0)-F(x_1)=F(x_2)-F(x_1).
\end{aligned}
$$

定理（唯一性定理）

分布函数可由特征函数唯一确定.

定理（唯一性定理）

分布函数可由特征函数唯一确定.

证：在(3.3.3)中令$y = x_1$沿$F(x)$的连续点趋向$-\infty$，令$x = x_2$，则

定理（唯一性定理）

分布函数可由特征函数唯一确定.

证: 在(3.3.3)中令$y = x_1$沿$F(x)$的连续点趋向$-\infty$, 令$x = x_2$, 则

$$F(x) = \lim_{y\to-\infty}\lim_{T\to\infty}\frac{1}{2\pi}\int_{-T}^{T}\frac{e^{-\mathrm{i}ty}-e^{-\mathrm{i}tx}}{\mathrm{i}t}f(t)\mathrm{d}t.$$

于是在$F(x)$的连续点, 由$f(t)$决定了$F(x)$. 至于在$F(x)$的间断点上, 由于$F(x)$的右连续性, 只需沿连续点取右极限, 就唯一确定了在间断点处的$F(x)$值.

定理 (唯一性定理)

分布函数可由特征函数唯一确定.

证: 在(3.3.3)中令$y=x_1$沿$F(x)$的连续点趋向$-\infty$, 令$x=x_2$, 则

$$F(x)=\lim_{y\to-\infty}\lim_{T\to\infty}\frac{1}{2\pi}\int_{-T}^{T}\frac{e^{-\mathrm{i}ty}-e^{-\mathrm{i}tx}}{\mathrm{i}t}f(t)\mathrm{d}t.$$

于是在$F(x)$的连续点, 由$f(t)$决定了$F(x)$. 至于在$F(x)$的间断点上, 由于$F(x)$的右连续性, 只需沿连续点取右极限, 就唯一确定了在间断点处的$F(x)$值.

注: (1) 结合此定理和特征函数的定义可知: 分布函数和特征函数相互唯一确定; (2) 用唯一性定理和逆转公式计算分布函数是困难的, 它的意义主要是理论上的.

定理（逆Fourier变换）

设$f(t)$是特征函数，且$\int_{-\infty}^{\infty}|f(t)|\mathrm{d}t<\infty$($f(t)$绝对可积)，则分布函数$F(x)$的导数存在且连续，此时

$$F'(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-\mathrm{i}tx}f(t)\mathrm{d}t. \tag{3.3.4}$$

定理（逆Fourier变换）

设$f(t)$是特征函数，且$\int_{-\infty}^{\infty}|f(t)|\mathrm{d}t<\infty$($f(t)$绝对可积)，则分布函数$F(x)$的导数存在且连续，此时

$$F'(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-\mathrm{i}tx}f(t)\mathrm{d}t. \tag{3.3.4}$$

注：$F'(x)$连续，则$F(x)=\int_{-\infty}^{x}F'(t)\mathrm{d}t$，从而$F'(x)=p(x)$是$\xi$的密度函数，$\xi$为连续型随机变量．这个定理说明：$f(t)$绝对可积时，对应的随机变量必为连续型，其密度函数由(3.3.4)决定．(3.3.4)与连续型随机变量的特征函数公式$f(t)=\int_{-\infty}^{\infty}e^{\mathrm{i}tx}p(x)\mathrm{d}x$恰为一对Fourier变换．

证: 我们先来证明$F(x)$在$\mathbb{R}$上是连续的. 任取一点$x \in \mathbb{R}$, 选择$\delta > 0$使得$x \pm \delta$都是分布函数的连续点, 那么根据逆转公式可得

$$F(x+\delta) - F(x-\delta) = \frac{1}{2\pi} \lim_{T\to\infty} \int_{-T}^{T} \frac{e^{-\mathrm{i}t(x-\delta)} - e^{-\mathrm{i}t(x+\delta)}}{\mathrm{i}t} f(t)\mathrm{d}t.$$

证: 我们先来证明$F(x)$在$\mathbb{R}$上是连续的. 任取一点$x \in \mathbb{R}$, 选择$\delta > 0$使得$x \pm \delta$都是分布函数的连续点, 那么根据逆转公式可得

$$F(x+\delta) - F(x-\delta) = \frac{1}{2\pi} \lim_{T\to\infty} \int_{-T}^{T} \frac{e^{-\mathrm{i}t(x-\delta)} - e^{-\mathrm{i}t(x+\delta)}}{\mathrm{i}t} f(t)\mathrm{d}t.$$

利用不等式$|e^{\mathrm{i}\alpha} - 1| \le |\alpha|$, 可得

$$\left| \frac{e^{-\mathrm{i}t(x-\delta)} - e^{-\mathrm{i}t(x+\delta)}}{\mathrm{i}t} \right| \le 2\delta.$$

根据假设$\int_{-\infty}^{\infty} |f(t)|dt < \infty$, 可得

$$F(x+\delta) - F(x-\delta) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{-\mathrm{i}t(x-\delta)} - e^{-\mathrm{i}t(x+\delta)}}{\mathrm{i}t} f(t)\mathrm{d}t.$$

上式右边关于变量δ是连续的. 令$\delta \to 0$, 同时使得$x \pm \delta$都是$F(x)$的连续点, 那么由控制收敛定理得

$$\lim_{\delta \to 0} \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{-\mathrm{i}t(x-\delta)} - e^{-\mathrm{i}t(x+\delta)}}{\mathrm{i}t} f(t)\mathrm{d}t = 0.$$

所以

$$F(x) = F(x+0) = F(x-0).$$

所以$F(x)$在$\mathbb{R}$上是连续的.

接下来, 由逆转公式, 对任意的$x, x+\Delta x \in \mathbb{R}$, 有

$$\frac{F(x+\Delta x)-F(x)}{\Delta x}=\lim_{T\to\infty}\frac{1}{2\pi}\int_{-T}^{T}\frac{e^{-\mathrm{i}tx}-e^{-\mathrm{i}t(x+\Delta x)}}{\mathrm{i}t\Delta x}f(t)\mathrm{d}t.$$

接下来, 由逆转公式, 对任意的$x, x+\Delta x \in \mathbb{R}$, 有

$$\frac{F(x+\Delta x)-F(x)}{\Delta x}=\lim_{T\to\infty}\frac{1}{2\pi}\int_{-T}^{T}\frac{e^{-\mathrm{i}tx}-e^{-\mathrm{i}t(x+\Delta x)}}{\mathrm{i}t\Delta x}f(t)\mathrm{d}t.$$

再次利用不等式$|\mathrm{e}^{\mathrm{i}\alpha}-1|\le|\alpha|$以及$f(t)$的绝对可积性, 可得

$$\frac{F(x+\Delta x)-F(x)}{\Delta x}=\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{e^{-\mathrm{i}tx}-e^{-\mathrm{i}t(x+\Delta x)}}{\mathrm{i}t\Delta x}f(t)\mathrm{d}t.$$

利用控制收敛定理得到

$$
\begin{aligned}
F^{'}(x) &= \lim_{\Delta x \to 0} \frac{F(x+\Delta x) - F(x)}{\Delta x} \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} \lim_{\Delta x \to 0} \frac{e^{-\mathrm{i}tx} - e^{-\mathrm{i}t(x+\Delta x)}}{\mathrm{i}t\Delta x} f(t)\mathrm{d}t \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-\mathrm{i}tx} f(t)\mathrm{d}t.
\end{aligned}
$$

因此$p(x) = F^{'}(x)$存在且有界.

利用控制收敛定理得到

$$
\begin{aligned}
F'(x) &= \lim_{\Delta x\to 0}\frac{F(x+\Delta x)-F(x)}{\Delta x}\\
&=\frac{1}{2\pi}\int_{-\infty}^{\infty}\lim_{\Delta x\to 0}\frac{e^{-\mathrm{i}tx}-e^{-\mathrm{i}t(x+\Delta x)}}{\mathrm{i}t\Delta x}f(t)\mathrm{d}t\\
&=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-\mathrm{i}tx}f(t)\mathrm{d}t.
\end{aligned}
$$

因此$p(x)=F'(x)$存在且有界. 再次利用控制收敛定理可得

$$
\begin{aligned}
\lim_{h\to 0}[F'(x+h)-F'(x)] &= \lim_{h\to 0}\frac{1}{2\pi}\int_{-\infty}^{\infty}(e^{-\mathrm{i}t(x+h)}-e^{-\mathrm{i}tx})f(t)\mathrm{d}t\\
&=\frac{1}{2\pi}\int_{-\infty}^{\infty}\lim_{h\to 0}(e^{-\mathrm{i}t(x+h)}-e^{-\mathrm{i}tx})f(t)\mathrm{d}t\\
&=0.
\end{aligned}
$$

所以$F'(x)$是连续的.

对于离散型随机变量，我们有类似的结果:

对于离散型随机变量，我们有类似的结果：假设ξ是取非负整数值的随机变量，分布列为

$$\mathsf{P}(\xi = k) = p_k, \quad k = 0, 1, 2, \cdots,$$

那么其特征函数为

$$f(t) = \sum_{k=0}^{\infty} p_k e^{\mathrm{i}tk}.$$

由于

$$\int_0^{2\pi} e^{\mathrm{i}nt}\mathrm{d}t = \begin{cases} 2\pi, & n = 0, \\ 0, & n \neq 0. \end{cases}$$

那么，我们有

$$p_k = \frac{1}{2\pi}\int_0^{2\pi} e^{-\mathrm{i}tk} f(t)\mathrm{d}t.$$

例11 求证$f(t)=\cos t$是某随机变量的特征函数, 并求出它的概率分布.

例11 求证$f(t)=\cos t$是某随机变量的特征函数, 并求出它的概率分布.

解:
$$f(t)=\cos t=\frac{1}{2}(e^{\mathrm{i}t}+e^{-\mathrm{i}t})=\frac{1}{2}e^{\mathrm{i}t}+\frac{1}{2}e^{-\mathrm{i}t},$$
这是分布列为

ξ	1	-1
P	$1/2$	$1/2$

的随机变量的特征函数.

例11 求证$f(t) = \cos t$是某随机变量的特征函数, 并求出它的概率分布.

解:

$$f(t) = \cos t = \frac{1}{2}(e^{\mathrm{i}t} + e^{-\mathrm{i}t}) = \frac{1}{2}e^{\mathrm{i}t} + \frac{1}{2}e^{-\mathrm{i}t},$$

这是分布列为

ξ	1	-1
P	$1/2$	$1/2$

的随机变量的特征函数.

注: 一般地, 若$f(t)$能写成$f(t) = \sum_n a_n e^{\mathrm{i}tx_n}$的形式, 其中$a_n > 0$, $\sum_n a_n = 1$, 则$f(t)$是特征函数, 相应的随机变量的分布列为

$$\mathrm{P}(\xi = x_n) = a_n, \ n = 1, 2, \cdots.$$

例12 若$f(t)$是某随机变量的特征函数, 求证$\overline{f(t)}$, $|f(t)|^2$以及$f^n(t)$ $(n \in \mathbb{N})$ 都是特征函数.

例12 若$f(t)$是某随机变量的特征函数, 求证$\overline{f(t)}$, $|f(t)|^2$以及$f^n(t)$ $(n \in \mathbb{N})$ 都是特征函数.

解: 假设$f(t)$是随机变量ξ的特征函数. 易知$\overline{f(t)} = f(-t)$是$-\xi$的特征函数;

例12 若$f(t)$是某随机变量的特征函数, 求证$\overline{f(t)}$, $|f(t)|^2$以及$f^n(t)$ $(n \in \mathbb{N})$ 都是特征函数.

解: 假设$f(t)$是随机变量ξ的特征函数. 易知$\overline{f(t)} = f(-t)$是$-\xi$的特征函数; 设$\xi_1, \cdots, \xi_n$相互独立且与ξ具有相同的分布. 则$\eta = -\xi_2$与ξ_1独立, 且特征函数为$\overline{f(t)}$. 易知$\xi_1 + \eta = \xi_1 - \xi_2$的特征函数为

$$f(t) \cdot \overline{f(t)} = |f(t)|^2;$$

例12 若$f(t)$是某随机变量的特征函数, 求证$\overline{f(t)}$, $|f(t)|^2$以及$f^n(t)$ $(n \in \mathbb{N})$ 都是特征函数.

解: 假设$f(t)$是随机变量ξ的特征函数. 易知$\overline{f(t)} = f(-t)$是$-\xi$的特征函数; 设$\xi_1, \cdots, \xi_n$相互独立且与ξ具有相同的分布. 则$\eta = -\xi_2$与ξ_1独立, 且特征函数为$\overline{f(t)}$. 易知$\xi_1 + \eta = \xi_1 - \xi_2$的特征函数为

$$f(t) \cdot \overline{f(t)} = |f(t)|^2;$$

$f^n(t)$显然是$\zeta = \xi_1 + \cdots + \xi_n$的特征函数.

*例13 证明函数$f(t)=|\cos t|$不是特征函数.

*例13 证明函数$f(t)=|\cos t|$不是特征函数.

证: 因为$\cos t$是特征函数, 所以$\cos^2 t$是也是特征函数, 事实上它是下列随机变量X的特征函数:

X	-2	0	2
P	$1/4$	$1/2$	$1/4$

假设$f(t)=|\cos t|$是随机变量ξ的特征函数. 令η与ξ相互独立且具有相同的分布, 则$\xi+\eta$的特征函数为$f^2(t)=\cos^2 t$. 而我们已经知道特征函数为$\cos^2 t$的随机变量为X, 它只能取三个值. 所以ξ只能取两个值. 另一方面, 因为ξ的特征函数是实的, 所以ξ的分布是对称的, 即存在非零常数a使得ξ的分布列为

$$\mathsf{P}(\xi=a)=\mathsf{P}(\xi=-a)=\frac{1}{2}.$$

从这个分布列可求得$f(t)=\cos(at)$, 这与$f(t)=|\cos t|$矛盾. 所以$f(t)=|\cos t|$不是特征函数.

3.3.4 分布函数的可加性

3.3.4 分布函数的可加性

许多重要的分布函数具有一个有趣的性质—-可加性. 这个性质用特征函数来研究最为方便. 下面通过几个例子来说明.

例14 若$\xi_j, j = 1, \cdots, k$ 各自服从二项分布$B(n_j, p)$, 且相互独立, 则

$$\sum_{j=1}^{k} \xi_j \sim B(\sum_{j=1}^{k} n_j, p).$$

例14 若$\xi_j, j=1,\cdots,k$ 各自服从二项分布$B(n_j,p)$，且相互独立，则

$$\sum_{j=1}^{k}\xi_j \sim B(\sum_{j=1}^{k} n_j, p).$$

解: ξ_j的特征函数为

$$f_j(t)=(pe^{\mathrm{i}t}+q)^{n_j},\ (p+q=1).$$

由独立性，可知$\sum_{j=1}^{k}\xi_j$的特征函数为

$$\prod_{j=1}^{k} f_j(t)=(pe^{\mathrm{i}t}+q)^{\sum_{j=1}^{k} n_j}.$$

根据唯一性定理可知$\sum_{j=1}^{k}\xi_j \sim B(\sum_{j=1}^{k} n_j, p)$.

例15 若$\xi_j, j=1,\cdots,k$ 各自服从Poisson分布$P(\lambda_j)$, 且相互独立, 则

$$\sum_{j=1}^{k}\xi_j \sim P(\sum_{j=1}^{k}\lambda_j).$$

例15 若$\xi_j, j=1,\cdots,k$ 各自服从Poisson分布$P(\lambda_j)$, 且相互独立, 则

$$\sum_{j=1}^{k}\xi_j \sim P(\sum_{j=1}^{k}\lambda_j).$$

解: ξ_j的特征函数为

$$f_j(t) = \exp\{\lambda_j(e^{\mathrm{i}t}-1)\}.$$

由独立性, 可知$\sum_{j=1}^{k}\xi_j$的特征函数为

$$\prod_{j=1}^{k} f_j(t) = \exp\left\{\sum_{j=1}^{k}\lambda_j(e^{\mathrm{i}t}-1)\right\}.$$

根据唯一性定理可知$\sum_{j=1}^{k}\xi_j \sim P(\sum_{j=1}^{k}\lambda_j)$.

例16 若$\xi_j, j=1,\cdots,k$ 各自服从正态分布分布$N(a_j,\sigma_j^2)$, 且相互独立, 则

$$\sum_{j=1}^k \xi_j \sim N(\sum_{j=1}^k a_j, \sum_{j=1}^k \sigma_j^2).$$

例16 若$\xi_j, j=1,\cdots,k$ 各自服从正态分布分布$N(a_j,\sigma_j^2)$, 且相互独立, 则

$$\sum_{j=1}^{k}\xi_j \sim N(\sum_{j=1}^{k}a_j, \sum_{j=1}^{k}\sigma_j^2).$$

解: ξ_j的特征函数为

$$f_j(t)=\exp\{\mathrm{i}a_jt-\sigma_j^2t^2/2\}.$$

由独立性, 可知$\sum_{j=1}^{k}\xi_j$的特征函数为

$$\prod_{j=1}^{k}f_j(t)=\exp\left\{\mathrm{i}(\sum_{j=1}^{k}a_j)t-(\sum_{j=1}^{k}\sigma_j^2)t^2/2\right\}.$$

根据唯一性定理可知$\sum_{j=1}^{k}\xi_j \sim N(\sum_{j=1}^{k}a_j, \sum_{j=1}^{k}\sigma_j^2)$.

3.3.5 多元特征函数

3.3.5 多元特征函数

定义

设随机向量$\boldsymbol{\xi} = (\xi_1, \cdots, \xi_n)'$的分布函数为$F(x_1, \cdots, x_n)$, 称

$$f(t_1, \cdots, t_n) = Ee^{\mathrm{i}(t_1\xi_1+\cdots+t_n\xi_n)}$$
$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{\mathrm{i}(t_1x_1+\cdots+t_nx_n)}\mathrm{d}F(x_1, \cdots, x_n) \qquad (3.3.5)$$

为$\boldsymbol{\xi}$的特征函数.

3.3.5 多元特征函数

定义

设随机向量$\boldsymbol{\xi}=(\xi_1,\cdots,\xi_n)'$的分布函数为$F(x_1,\cdots,x_n)$, 称

$$\begin{aligned}f(t_1,\cdots,t_n)&=Ee^{\mathrm{i}(t_1\xi_1+\cdots+t_n\xi_n)}\\&=\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}e^{\mathrm{i}(t_1x_1+\cdots+t_nx_n)}\mathrm{d}F(x_1,\cdots,x_n)\end{aligned}\tag{3.3.5}$$

为$\boldsymbol{\xi}$的特征函数.

向量形式: 记$\boldsymbol{t}=(t_1,\cdots,t_n)'$, $\boldsymbol{x}=(x_1,\cdots,x_n)'$, 则(3.3.5)可改写为

$$f(\boldsymbol{t})=\mathrm{E}e^{\mathrm{i}\boldsymbol{t}'\boldsymbol{\xi}}=\int_{\mathbb{R}^n}e^{\mathrm{i}\boldsymbol{t}'\boldsymbol{x}}\mathrm{d}F(\boldsymbol{x}).$$

可以类似于一元的场合，可建立多元特征函数的理论，例如一致连续性、唯一性等等．这里我们只介绍某些新的性质，证明从略．

可以类似于一元的场合，可建立多元特征函数的理论，例如一致连续性、唯一性等等. 这里我们只介绍某些新的性质，证明从略.

性质1$'$：$\eta = a_1\xi_1 + \cdots + a_n\xi_n$的特征函数为

$$
\begin{aligned}
f_\eta(t) =& \mathsf{E}e^{\mathsf{i}t\eta} = \mathsf{E}\exp\left\{\mathsf{i}t\sum_{k=1}^{n} a_k\xi_k\right\} \\
=& \mathsf{E}\exp\left\{\mathsf{i}\sum_{k=1}^{n}(a_k t)\xi_k\right\} \\
=& f(a_1t, \cdots, a_nt).
\end{aligned}
$$

可以类似于一元的场合，可建立多元特征函数的理论，例如一致连续性、唯一性等等. 这里我们只介绍某些新的性质，证明从略.

性质1$'$：$\eta = a_1\xi_1 + \cdots + a_n\xi_n$的特征函数为

$$\begin{aligned} f_\eta(t) =& \mathsf{E}e^{\mathrm{i}t\eta} = \mathsf{E}\exp\left\{\mathrm{i}t\sum_{k=1}^{n} a_k\xi_k\right\} \\ =& \mathsf{E}\exp\left\{\mathrm{i}\sum_{k=1}^{n}(a_k t)\xi_k\right\} \\ =& f(a_1 t, \cdots, a_n t). \end{aligned}$$

性质2$'$：若$(\xi_1, \cdots, \xi_n)'$的特征函数为$f(t_1, \cdots, t_n)$，则k维子向量$(\xi_{l_1}, \cdots, \xi_{l_k})'$的特征函数为

$$f(0, \cdots, 0, t_{l_1}, 0, \cdots, 0, t_{l_2}, 0, \cdots, 0, t_{l_k}, 0, \cdots, 0).$$

性质$3'$: 设随机向量$\boldsymbol{\xi}=(\xi_1,\cdots,\xi_n)'$的特征函数为$f_{\boldsymbol{\xi}}$, 假设$\boldsymbol{L}$是$m\times n$矩阵,

$$\boldsymbol{a}=(a_1,\cdots,a_m)',\ \ \boldsymbol{\eta}=\boldsymbol{L}\boldsymbol{\xi}+\boldsymbol{a}.$$

那么

$$f_{\boldsymbol{\eta}}(t_1,\cdots,t_m)=f_{\boldsymbol{\xi}}(\boldsymbol{L}'\boldsymbol{t})e^{\mathrm{i}\boldsymbol{t}'\boldsymbol{a}}.$$

性质$3'$：设随机向量$\boldsymbol{\xi} = (\xi_1, \cdots, \xi_n)'$的特征函数为$f_{\boldsymbol{\xi}}$, 假设$\boldsymbol{L}$是$m \times n$矩阵,

$$\boldsymbol{a} = (a_1, \cdots, a_m)', \ \ \boldsymbol{\eta} = \boldsymbol{L}\boldsymbol{\xi} + \boldsymbol{a}.$$

那么

$$f_{\boldsymbol{\eta}}(t_1, \cdots, t_m) = f_{\boldsymbol{\xi}}(\boldsymbol{L}'\boldsymbol{t})e^{\mathrm{i}\boldsymbol{t}'\boldsymbol{a}}.$$

性质$4'$：设ξ_j的特征函数为$f_j(t)$, $j = 1, \cdots, n$. 则$\xi_1, \cdots, \xi_n$独立的充要条件为$(\xi_1, \cdots, \xi_n)'$的特征函数为

$$f(t_1, \cdots, t_n) = f_1(t_1) \cdots f_n(t_n).$$

性质5$'$：随机向量$(\xi_1,\cdots,\xi_k)'$与$(\xi_{k+1},\cdots,\xi_n)'$独立的充要条件是它们的特征函数的乘积恰为$(\xi_1,\cdots,\xi_n)'$的特征函数.

性质5$'$：随机向量$(\xi_1,\cdots,\xi_k)'$与$(\xi_{k+1},\cdots,\xi_n)'$独立的充要条件是它们的特征函数的乘积恰为$(\xi_1,\cdots,\xi_n)'$的特征函数.

性质6$'$：设随机向量$(\xi_1,\cdots,\xi_n)'$的特征函数为$f(t_1,\cdots,t_n)$. 如果

$$\mathsf{E}\prod_{l=1}^{m}\xi_{i_l}\ (1\le m\le n)$$

存在，那么

$$\left.\frac{\partial f(t_1,\cdots,t_n)}{\partial t_{i_1}\cdots\partial t_{i_m}}\right|_{t_1,\cdots,t_n=0}=\mathrm{i}^m\mathsf{E}\prod_{l=1}^{m}\xi_{i_l}.$$

*多元正态分布

多元分布以多元正态分布最为重要, 在多元分析中多元正态分布更为其立论之本. 这一节我们将借助多元特征函数详细讨论多元正态分布的定义和性质.

*多元正态分布

多元分布以多元正态分布最为重要, 在多元分析中多元正态分布更为其立论之本. 这一节我们将借助多元特征函数详细讨论多元正态分布的定义和性质.

3.4.1 密度函数和特征函数

*多元正态分布

多元分布以多元正态分布最为重要, 在多元分析中多元正态分布更为其立论之本. 这一节我们将借助多元特征函数详细讨论多元正态分布的定义和性质.

3.4.1 密度函数和特征函数

在第二章中, 我们已给出n元正态分布$\boldsymbol{\xi} \sim N(\boldsymbol{a}, \boldsymbol{\Sigma})$的密度函数:

$$p(\boldsymbol{x}) = \frac{1}{(2\pi)^{n/2}|\boldsymbol{\Sigma}|^{1/2}} \exp\left\{ -\frac{1}{2}(\boldsymbol{x}-\boldsymbol{a})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{a})\right\}, \ \boldsymbol{x} \in \mathbb{R}^n, \tag{3.4.1}$$

其中$\boldsymbol{\Sigma} = (\sigma_{ij})$为$n$阶正定对称矩阵,

$$\boldsymbol{x} = (x_1, \cdots, x_n)', \ \ \boldsymbol{a} = (a_1, \cdots, a_n)'.$$

$\boldsymbol{\xi} \sim N(\boldsymbol{a}, \boldsymbol{\Sigma})$的特征函数:

$\boldsymbol{\xi} \sim N(\boldsymbol{a}, \boldsymbol{\Sigma})$的特征函数:

记$\boldsymbol{\Sigma} = \boldsymbol{B}\boldsymbol{B}'$ $(\boldsymbol{B} = \boldsymbol{\Sigma}^{1/2})$. 令$\boldsymbol{\eta} = \boldsymbol{B}^{-1}(\boldsymbol{\xi} - \boldsymbol{a})$. 则$\boldsymbol{\eta}$的密度函数为

$$\begin{aligned}
&p_\eta(\boldsymbol{y}) = p(\boldsymbol{x})|\boldsymbol{B}|, \quad (\boldsymbol{x} = \boldsymbol{B}\boldsymbol{y} + \boldsymbol{a}) \\
=&\frac{1}{(2\pi)^{n/2}} \exp\left\{ -\frac{1}{2}(\boldsymbol{x} - \boldsymbol{a})'\boldsymbol{B}'^{-1}\boldsymbol{B}^{-1}(\boldsymbol{x} - \boldsymbol{a})\right\} \\
=&\frac{1}{(2\pi)^{n/2}} \exp\left\{ -\frac{1}{2}\boldsymbol{y}'\boldsymbol{y}\right\} \\
=&\prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp\left\{ -\frac{y_i^2}{2}\right\}.
\end{aligned}$$

即, $\eta_1, \cdots, \eta_n$相互独立且都服从$N(0,1)$. 因此$\boldsymbol{\eta}$的特征函数为

$$f_{\boldsymbol{\eta}}(\boldsymbol{t}) = \prod_{i=1}^{n} \exp\{-t_i^2/2\} = \exp\left\{ -\frac{1}{2}\boldsymbol{t}'\boldsymbol{t}\right\}.$$

由于$\boldsymbol{\xi} = \boldsymbol{B}\boldsymbol{\eta} + \boldsymbol{a}$, 所以$\boldsymbol{\xi}$的特征函数为

$$\begin{aligned} f(\boldsymbol{t}) &= \mathsf{E}e^{\mathrm{i}\boldsymbol{t}'\boldsymbol{\xi}} = e^{\mathrm{i}\boldsymbol{t}'\boldsymbol{a}}\mathsf{E}e^{\mathrm{i}\boldsymbol{t}'\boldsymbol{B}\boldsymbol{\eta}} = e^{\mathrm{i}\boldsymbol{t}'\boldsymbol{a}}\mathsf{E}e^{\mathrm{i}(\boldsymbol{B}'\boldsymbol{t})'\boldsymbol{\eta}} \\ &= e^{\mathrm{i}\boldsymbol{t}'\boldsymbol{a}}\exp\left\{-\frac{1}{2}(\boldsymbol{B}'\boldsymbol{t})'(\boldsymbol{B}'\boldsymbol{t})\right\} \\ &= \exp\left\{\mathrm{i}\boldsymbol{a}'\boldsymbol{t} - \frac{1}{2}\boldsymbol{t}'\boldsymbol{\Sigma}\boldsymbol{t}\right\}, \end{aligned} \tag{3.4.2}$$

或记为

$$f(t_1, \cdots, t_n) = \exp\left\{\mathrm{i}\sum_{k=1}^{n} a_k t_k - \frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n}\sigma_{jk}t_j t_k\right\}. \tag{3.4.3}$$

上述只对$\boldsymbol{\Sigma}$是正定对称矩阵场合定义了多元正态分布. 当$\boldsymbol{\Sigma}$是非负定时, (3.4.1)可能没有意义, 但(3.4.2)仍有意义. 事实上, 它是下面随机变量的特征函数. 令$\boldsymbol{\Sigma}=\boldsymbol{B}\boldsymbol{B}'$, $\boldsymbol{\xi}=(\xi_1,\cdots,\xi_n)'$服从$n$元正态$N(\boldsymbol{0},\boldsymbol{I}_n)$, 则$\boldsymbol{\eta}=\boldsymbol{B}\boldsymbol{\xi}+\boldsymbol{a}$的特征函数为$\exp\left\{\mathrm{i}\boldsymbol{a}'\boldsymbol{t}-\frac{1}{2}\boldsymbol{t}'\boldsymbol{\Sigma}\boldsymbol{t}\right\}$. 当$\boldsymbol{\Sigma}$的秩为$r(r<n)$时, 称它对应的分布为奇异正态分布(singular normal distribution)或退化正态分布. 它实际上只是r维子空间上的一个分布.

所以特征函数(3.4.2)比密度函数(3.4.1)适用范围更广. 因此有时就通过特征函数(3.4.2)来定义多元正态分布.

3.4.2 性质

3.4.2 性质

下面的讨论中，我们总假定随机向量$\boldsymbol{\xi} = (\xi_1, \cdots, \xi_n)'$服从$n$元正态分布$N(\boldsymbol{a}, \boldsymbol{\Sigma})$.

性质1: $\boldsymbol{\xi}$的任一子向量$(\xi_{l_1},\cdots,\xi_{l_k})$也服从正态分布, 分布为$N(\tilde{\boldsymbol{a}},\tilde{\boldsymbol{\Sigma}})$, 其中$\tilde{\boldsymbol{a}}=(a_{l_1},\cdots,a_{l_k})'$, $\tilde{\boldsymbol{\Sigma}}$为取$\boldsymbol{\Sigma}$的第$l_1,\cdots,l_k$ 行与第$l_1,\cdots,l_k$列交叉点所得的k阶矩阵.

性质1: $\boldsymbol{\xi}$的任一子向量$(\xi_{l_1},\cdots,\xi_{l_k})$也服从正态分布，分布为$N(\tilde{\boldsymbol{a}},\tilde{\boldsymbol{\Sigma}})$，其中$\tilde{\boldsymbol{a}}=(a_{l_1},\cdots,a_{l_k})'$，$\tilde{\boldsymbol{\Sigma}}$为取$\boldsymbol{\Sigma}$的第$l_1,\cdots,l_k$行与第$l_1,\cdots,l_k$列交叉点所得的$k$阶矩阵.

证: $(\xi_{l_1},\cdots,\xi_{l_k})$的特征函数为

$$f(0,\cdots,0,t_{l_1},0,\cdots,0,t_{l_k},0,\cdots,0).$$

在(3.4.3)中除上述各t_{l_j}外，令其余的$t_i=0$，并记$\tilde{\boldsymbol{t}}=(t_{l_1},\cdots,t_{l_k})'$，则

$$\begin{aligned}\tilde{f}(\tilde{t}) &= \exp\left\{\mathrm{i}\sum_{j=1}^{k}a_{l_j}t_{l_j}-\sum_{j=1}^{k}\sum_{s=1}^{k}\sigma_{l_jl_s}t_{l_j}t_{l_2}/2\right\}\\ &= \exp\left\{\mathrm{i}\tilde{\boldsymbol{a}}'\tilde{\boldsymbol{t}}-\tilde{\boldsymbol{t}}'\tilde{\boldsymbol{\Sigma}}\tilde{\boldsymbol{t}}\right\}.\end{aligned}$$

这正是$N(\tilde{\boldsymbol{a}},\tilde{\boldsymbol{\Sigma}})$的特征函数.

性质1: $\boldsymbol{\xi}$的任一子向量$(\xi_{l_1},\cdots,\xi_{l_k})$也服从正态分布，分布为$N(\tilde{\boldsymbol{a}},\tilde{\boldsymbol{\Sigma}})$，其中$\tilde{\boldsymbol{a}}=(a_{l_1},\cdots,a_{l_k})'$，$\tilde{\boldsymbol{\Sigma}}$为取$\boldsymbol{\Sigma}$的第$l_1,\cdots,l_k$ 行与第$l_1,\cdots,l_k$列交叉点所得的k阶矩阵.

证: $(\xi_{l_1},\cdots,\xi_{l_k})$的特征函数为

$$f(0,\cdots,0,t_{l_1},0,\cdots,0,t_{l_k},0,\cdots,0).$$

在(3.4.3)中除上述各t_{l_j}外，令其余的$t_i=0$，并记$\tilde{\boldsymbol{t}}=(t_{l_1},\cdots,t_{l_k})'$，则

$$\begin{aligned}\tilde{f}(\tilde{t})&=\exp\left\{\mathrm{i}\sum_{j=1}^{k}a_{l_j}t_{l_j}-\sum_{j=1}^{k}\sum_{s=1}^{k}\sigma_{l_jl_s}t_{l_j}t_{l_2}/2\right\}\\&=\exp\left\{\mathrm{i}\tilde{\boldsymbol{a}}'\tilde{\boldsymbol{t}}-\tilde{\boldsymbol{t}}'\tilde{\boldsymbol{\Sigma}}\tilde{\boldsymbol{t}}\right\}.\end{aligned}$$

这正是$N(\tilde{\boldsymbol{a}},\tilde{\boldsymbol{\Sigma}})$的特征函数.

注: 性质1说明多元正态分布的边际分布还是正态分布. 特别地，当$k=1$时，$\xi_j\sim N(a_j,\sigma_{jj})$.

性质2: 若$\boldsymbol{\xi} \sim N(\boldsymbol{a}, \boldsymbol{\Sigma})$, 则$\boldsymbol{\xi}$的数学期望向量为$\boldsymbol{a}$, 协方差矩阵为$\boldsymbol{\Sigma}$.

性质2: 若$\boldsymbol{\xi} \sim N(\boldsymbol{a}, \boldsymbol{\Sigma})$, 则$\boldsymbol{\xi}$的数学期望向量为$\boldsymbol{a}$, 协方差矩阵为$\boldsymbol{\Sigma}$.

证: 由性质1可知$\mathsf{E}\xi_j = a_j$, $\mathsf{Var}\xi_j = \sigma_{jj}$, 所以我们只需证明$\boldsymbol{\Sigma}$的非对角线元素为相互协方差. 由Cauchy-Schwarz不等式可知$\mathsf{E}|\xi_j\xi_k| \le \sqrt{\mathsf{E}\xi_j^2\mathsf{E}\xi_k^2}$存在. 由(3.4.3)得

$$\frac{\partial f(t_1,\cdots,t_n)}{\partial t_j} = \Big[\mathrm{i}a_j - \frac{1}{2}\Big(\sum_{l=1}^n \sigma_{lj}t_l + \sum_{s=1}^n \sigma_{js}t_s\Big)\Big] \cdot \exp\Big\{\mathrm{i}\sum_{k=1}^n a_k t_k - \frac{1}{2}\sum_{l=1}^n\sum_{s=1}^n \sigma_{ls}t_l t_s\Big\},$$

$$\begin{aligned}\frac{\partial^2 f(t_1,\cdots,t_n)}{\partial t_j \partial t_k} = &\Big\{-\frac{1}{2}(\sigma_{jk}+\sigma_{kj}) + \Big[\mathrm{i}a_j - \frac{1}{2}\Big(\sum_{l=1}^n \sigma_{lj}t_l + \sum_{s=1}^n \sigma_{js}t_s\Big)\Big] \\ &\cdot\Big[\mathrm{i}a_k - \frac{1}{2}\Big(\sum_{l=1}^n \sigma_{lk}t_l + \sum_{s=1}^n \sigma_{ks}t_s\Big)\Big]\Big\}\exp\Big\{\mathrm{i}\sum_{k=1}^n a_k t_k - \frac{1}{2}\sum_{l=1}^n\sum_{s=1}^n \sigma_{ls}t_l t_s\Big\}.\end{aligned}$$

那么由多元特征函数的性质可得

$$\mathsf{E}(\xi_j\xi_k) = \frac{1}{\mathrm{i}^2}\frac{\partial^2 f(t_1,\cdots,t_n)}{\partial t_j \partial t_k}\Big|_{t_1=\cdots=t_n=0} = \sigma_{jk} + a_j a_k.$$

故

$$\mathsf{E}[(\xi_j - a_j)(\xi_k - a_k)] = \mathsf{E}(\xi_j\xi_k) - a_j a_k = \sigma_{jk}.$$

那么由多元特征函数的性质可得

$$\mathsf{E}(\xi_j\xi_k) = \frac{1}{\mathrm{i}^2}\frac{\partial^2 f(t_1,\cdots,t_n)}{\partial t_j \partial t_k}\Big|_{t_1=\cdots=t_n=0} = \sigma_{jk} + a_j a_k.$$

故

$$\mathsf{E}[(\xi_j - a_j)(\xi_k - a_k)] = \mathsf{E}(\xi_j\xi_k) - a_j a_k = \sigma_{jk}.$$

注: n元正态分布由它的前面二阶矩完全确定.

正如在二元场合一样, 对n元正态分布而言, 独立性与不相关性有密切联系.

正如在二元场合一样, 对n元正态分布而言, 独立性与不相关性有密切联系.

性质3: 服从n元正态分布的随机变量$\xi_1, \cdots, \xi_n$相互独立的充要条件是它们两两不相关.

正如在二元场合一样, 对n元正态分布而言, 独立性与不相关性有密切联系.

性质3: 服从n元正态分布的随机变量$\xi_1, \cdots, \xi_n$相互独立的充要条件是它们两两不相关.

证:

$$
\begin{aligned}
\xi_1, \cdots, \xi_n \text{相互独立} \quad & \Longleftrightarrow \quad f(t_1, \cdots, t_n) = f_1(t_1) \cdots f_n(t_n) \\
& \Longleftrightarrow \quad \boldsymbol{t}' \boldsymbol{\Sigma} \boldsymbol{t} = \sum_{j=1}^{n} \sigma_{jj} t_j^2 \\
& \Longleftrightarrow \quad \sigma_{ij} = 0, \ \ i, j = 1, \cdots, n, \ i \neq j \\
& \Longleftrightarrow \quad \xi_i \text{与} \xi_j \text{不相关}, \ \ i, j = 1, \cdots, n, \ i \neq j.
\end{aligned}
$$

性质4: 若$\boldsymbol{\xi} = (\boldsymbol{\xi}_1', \boldsymbol{\xi}_2')'$, 这里$\boldsymbol{\xi}_1$与$\boldsymbol{\xi}_2$是$\boldsymbol{\xi}$的子向量. 记

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix}, \tag{3.4.4}$$

其中$\boldsymbol{\Sigma}_{11}$及$\boldsymbol{\Sigma}_{22}$分别是$\boldsymbol{\xi}_1$及$\boldsymbol{\xi}_2$的协方差矩阵, $\boldsymbol{\Sigma}_{12}$则是由$\boldsymbol{\xi}_1$与$\boldsymbol{\xi}_2$的相互协方差矩阵, 则$\boldsymbol{\xi}_1$与$\boldsymbol{\xi}_2$独立的充要条件是$\boldsymbol{\Sigma}_{12} = \mathbf{0}$.

性质4: 若$\boldsymbol{\xi} = (\boldsymbol{\xi}_1', \boldsymbol{\xi}_2')'$, 这里$\boldsymbol{\xi}_1$与$\boldsymbol{\xi}_2$是$\boldsymbol{\xi}$的子向量. 记

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix}, \tag{3.4.4}$$

其中$\boldsymbol{\Sigma}_{11}$及$\boldsymbol{\Sigma}_{22}$分别是$\boldsymbol{\xi}_1$及$\boldsymbol{\xi}_2$的协方差矩阵, $\boldsymbol{\Sigma}_{12}$则是由$\boldsymbol{\xi}_1$与$\boldsymbol{\xi}_2$的相互协方差矩阵, 则$\boldsymbol{\xi}_1$与$\boldsymbol{\xi}_2$独立的充要条件是$\boldsymbol{\Sigma}_{12} = \mathbf{0}$.

证: 若$\boldsymbol{\xi}_1$与$\boldsymbol{\xi}_2$独立, 则$\boldsymbol{\xi}_1$的任一分量与$\boldsymbol{\xi}_2$的任一分量独立, 因此其协方差为0, 从而由它们构成的矩阵$\boldsymbol{\Sigma}_{12} = \mathbf{0}$, 这就证明了必要性. 下证充分性: 由$\sigma_{ij} = \sigma_{ji}$, 因此$\boldsymbol{\Sigma}_{21} = \boldsymbol{\Sigma}_{12}{}' = \mathbf{0}$, 若$\boldsymbol{t} = (\boldsymbol{t}_1', \boldsymbol{t}_2')'$, 这里$\boldsymbol{t}_1$与$\boldsymbol{\xi}_1$有相同的维数, $\boldsymbol{t}_2$与$\boldsymbol{\xi}_2$有相同的维数, 则

$$\boldsymbol{t}'\boldsymbol{\Sigma}\boldsymbol{t} = \boldsymbol{t}_1{}'\boldsymbol{\Sigma}_{11}\boldsymbol{t}_1 + 2\boldsymbol{t}_1{}'\boldsymbol{\Sigma}_{12}\boldsymbol{t}_2 + \boldsymbol{t}_2{}'\boldsymbol{\Sigma}_{22}\boldsymbol{t}_2 = \boldsymbol{t}_1{}'\boldsymbol{\Sigma}_{11}\boldsymbol{t}_1 + \boldsymbol{t}_2{}'\boldsymbol{\Sigma}_{22}\boldsymbol{t}_2.$$

因此若记$\boldsymbol{a}=(\boldsymbol{a}_1^{'},\boldsymbol{a}_2^{'})^{'}$, 其中$\boldsymbol{a}_1,\boldsymbol{a}_2$分别为$\boldsymbol{\xi}_1$以$\boldsymbol{\xi}_2$的数学期望, 则

$$
\begin{aligned}
f_{\boldsymbol{\xi}}(\boldsymbol{t}) &= \exp\left\{\mathrm{i}\boldsymbol{a}^{'}\boldsymbol{t}-\frac{1}{2}\boldsymbol{t}^{'}\boldsymbol{\Sigma}\boldsymbol{t}\right\} \\
&= \exp\left\{\mathrm{i}\boldsymbol{a}_1^{'}\boldsymbol{t}_1+\mathrm{i}\boldsymbol{a}_2^{'}\boldsymbol{t}_2-\frac{1}{2}\boldsymbol{t}_1^{'}\boldsymbol{\Sigma}_{11}\boldsymbol{t}_1-\frac{1}{2}\boldsymbol{t}_2^{'}\boldsymbol{\Sigma}_{22}\boldsymbol{t}_2\right\} \\
&= \exp\left\{\mathrm{i}\boldsymbol{a}_1^{'}\boldsymbol{t}_1-\frac{1}{2}\boldsymbol{t}_1^{'}\boldsymbol{\Sigma}_{11}\boldsymbol{t}_1\right\}\cdot\exp\left\{\mathrm{i}\boldsymbol{a}_2^{'}\boldsymbol{t}_2-\frac{1}{2}\boldsymbol{t}_2^{'}\boldsymbol{\Sigma}_{22}\boldsymbol{t}_2\right\} \\
&= f_{\boldsymbol{\xi}_1}(\boldsymbol{t}_1)f_{\boldsymbol{\xi}_2}(\boldsymbol{t}_2).
\end{aligned}
$$

所以$\boldsymbol{\xi}_1$与$\boldsymbol{\xi}_2$独立.

例1 $\boldsymbol{\xi} = (\xi_1, \xi_2, \xi_3)' \sim N(\boldsymbol{a}, \boldsymbol{\Sigma})$, 其中$\boldsymbol{\Sigma} = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$. 求证$\xi_1$与$\xi_2$相互不独立, 而$(\xi_1, \xi_2)'$与$\xi_3$独立.

例1 $\boldsymbol{\xi} = (\xi_1, \xi_2, \xi_3)' \sim N(\boldsymbol{a}, \boldsymbol{\Sigma})$, 其中$\boldsymbol{\Sigma} = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$. 求证$\xi_1$与$\xi_2$相互不独立, 而$(\xi_1, \xi_2)'$与$\xi_3$独立.

证: 因为$\text{Cov}(\xi_1, \xi_2) = \sigma_{12} = 1 \neq 0$, 所以$\xi_1$与$\xi_2$相互不独立. 又

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix}, \quad \boldsymbol{\Sigma}_{12} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

故$(\xi_1, \xi_2)'$与ξ_3独立.

服从正态分布的随机变量在线性变换下具有许多特殊的性质，这些性质有很大的理论和实用价值，下面只讨论这类性质中最基本的一些.

一般，若$\boldsymbol{\xi}=(\xi_1,\cdots,\xi_n)'$是$n$维随机向量，其数学期望为$\boldsymbol{a}$，协方差矩阵为$\boldsymbol{\Sigma}$.

考虑$\boldsymbol{\xi}$的分量的线性组合$\zeta=\sum_{j=1}^n l_j\xi_j=\boldsymbol{l}'\boldsymbol{\xi}$，显然

$$\mathsf{E}\zeta=\sum_{j=1}^n l_j a_j=\boldsymbol{l}'\boldsymbol{a},$$

$$\mathsf{Var}\zeta=\sum_{j=1}^n\sum_{k=1}^n l_j l_k\sigma_{jk}=\boldsymbol{l}'\boldsymbol{\Sigma}\boldsymbol{l}.$$

同样的，若$\boldsymbol{C}=(c_{ij})$是$m\times n$矩阵，则对m维随机向量$\boldsymbol{\eta}=\boldsymbol{C}\boldsymbol{\xi}$有

$$\mathsf{E}\boldsymbol{\eta}=\boldsymbol{C}\boldsymbol{a},\ \ \mathsf{Cov}(\boldsymbol{\eta})=\boldsymbol{C}\boldsymbol{\Sigma}\boldsymbol{C}'.$$

性质5: $\boldsymbol{\xi}$服从多元正态分布的充要条件是它的各分量的任意非退化的线性组合仍服从正态分布. 如记$\boldsymbol{l} = (l_1, \cdots, l_n)'$为不等于$\mathbf{0}$的任意$n$维实向量, 则

$$\begin{aligned}
\boldsymbol{\xi} \sim N(\boldsymbol{a}, \boldsymbol{\Sigma}) \quad & \Longleftrightarrow \quad \zeta = \boldsymbol{l}'\boldsymbol{\xi} \sim N(\boldsymbol{l}'\boldsymbol{a}, \boldsymbol{l}'\boldsymbol{\Sigma}\boldsymbol{l}) \\
& \Longleftrightarrow \quad \zeta = \sum_{j=1}^{n} l_j \xi_j \sim N(\sum_{j=1}^{n} l_j a_j, \sum_{j=1}^{n}\sum_{k=1}^{n} l_j l_k \sigma_{jk}).
\end{aligned}$$

证：必要性：设$\boldsymbol{\xi} \sim N(\boldsymbol{a}, \boldsymbol{\Sigma})$，对任意实数$u$，记$u\boldsymbol{l} = \boldsymbol{t}$，则$\zeta$的特征函数为

$$\begin{aligned} f_\zeta(u) =& \mathsf{E} e^{\mathrm{i}u\zeta} = \mathsf{E} e^{\mathrm{i}(u\boldsymbol{l})'\boldsymbol{\xi}} = f_{\boldsymbol{\xi}}(u\boldsymbol{l}) \\ &= \exp\left\{\mathrm{i}\boldsymbol{a}'u\boldsymbol{l} - (u\boldsymbol{l})'\boldsymbol{\Sigma}(u\boldsymbol{l})/2\right\} \\ &= \exp\left\{\mathrm{i}(\boldsymbol{a}'\boldsymbol{l})u - (\boldsymbol{l}'\boldsymbol{\Sigma}\boldsymbol{l})u^2/2\right\}, \end{aligned} \tag{3.4.5}$$

所以$\zeta \sim N(\boldsymbol{l}'\boldsymbol{a}, \boldsymbol{l}'\boldsymbol{\Sigma}\boldsymbol{l})$.

充分性：设$\zeta \sim N(\boldsymbol{l}'\boldsymbol{a}, \boldsymbol{l}'\boldsymbol{\Sigma}\boldsymbol{l})$，则(3.4.5)对一切实数$u$都成立. 取$u = 1$，则有

$$f_{\boldsymbol{\xi}}(\boldsymbol{l}) = \exp\left\{\mathrm{i}\boldsymbol{a}'\boldsymbol{l} - (\boldsymbol{l}'\boldsymbol{\Sigma}\boldsymbol{l})/2\right\}.$$

因此$\boldsymbol{\xi} \sim N(\boldsymbol{a}, \boldsymbol{\Sigma})$.

性质5中的必要条件可以推广到m维线性变换.

性质5中的必要条件可以推广到m维线性变换.

性质6: 设$\boldsymbol{\xi} = (\xi_1, \cdots, \xi_n)' \sim N(\boldsymbol{a}, \boldsymbol{\Sigma})$, $\boldsymbol{C} = (c_{ij})$为$m \times n$矩阵, 则$\boldsymbol{\eta} = \boldsymbol{C}\boldsymbol{\xi}$服从$m$元正态分布$N(\boldsymbol{C}\boldsymbol{a}, \boldsymbol{C}\boldsymbol{\Sigma}\boldsymbol{C}')$.

性质5中的必要条件可以推广到m维线性变换.

性质6: 设$\boldsymbol{\xi} = (\xi_1, \cdots, \xi_n)' \sim N(\boldsymbol{a}, \boldsymbol{\Sigma})$, $\boldsymbol{C} = (c_{ij})$为$m \times n$矩阵, 则$\boldsymbol{\eta} = \boldsymbol{C}\boldsymbol{\xi}$服从$m$元正态分布$N(\boldsymbol{C}\boldsymbol{a}, \boldsymbol{C}\boldsymbol{\Sigma}\boldsymbol{C}')$.

证:

$$\begin{aligned} f_{\boldsymbol{\eta}}(\boldsymbol{t}) &= \mathsf{E}e^{\mathrm{i}\boldsymbol{t}'\boldsymbol{\eta}} = \mathsf{E}e^{\mathrm{i}(\boldsymbol{t}'\boldsymbol{C})\boldsymbol{\xi}} = \mathsf{E}e^{\mathrm{i}(\boldsymbol{C}'\boldsymbol{t})'\boldsymbol{\xi}} = f_{\boldsymbol{\xi}}(\boldsymbol{C}'\boldsymbol{t}) \\ &= \exp\left\{\mathrm{i}\boldsymbol{a}'\boldsymbol{C}'\boldsymbol{t} - (\boldsymbol{C}'\boldsymbol{t})'\boldsymbol{\Sigma}(\boldsymbol{C}'\boldsymbol{t})/2\right\} \\ &= \exp\left\{\mathrm{i}(\boldsymbol{C}\boldsymbol{a})'\boldsymbol{t} - \boldsymbol{t}'(\boldsymbol{C}\boldsymbol{\Sigma}\boldsymbol{C}')\boldsymbol{t}/2\right\}, \end{aligned}$$

得证.

性质5中的必要条件可以推广到m维线性变换.

性质6: 设$\boldsymbol{\xi} = (\xi_1, \cdots, \xi_n)' \sim N(\boldsymbol{a}, \boldsymbol{\Sigma})$, $\boldsymbol{C} = (c_{ij})$为$m \times n$矩阵, 则$\boldsymbol{\eta} = \boldsymbol{C}\boldsymbol{\xi}$服从$m$元正态分布$N(\boldsymbol{C}\boldsymbol{a}, \boldsymbol{C}\boldsymbol{\Sigma}\boldsymbol{C}')$.

证:

$$\begin{aligned} f_{\boldsymbol{\eta}}(\boldsymbol{t}) &= \mathrm{E}e^{\mathrm{i}\boldsymbol{t}'\boldsymbol{\eta}} = \mathrm{E}e^{\mathrm{i}(\boldsymbol{t}'\boldsymbol{C})\boldsymbol{\xi}} = \mathrm{E}e^{\mathrm{i}(\boldsymbol{C}'\boldsymbol{t})'\boldsymbol{\xi}} = f_{\boldsymbol{\xi}}(\boldsymbol{C}'\boldsymbol{t}) \\ &= \exp\left\{\mathrm{i}\boldsymbol{a}'\boldsymbol{C}'\boldsymbol{t} - (\boldsymbol{C}'\boldsymbol{t})'\boldsymbol{\Sigma}(\boldsymbol{C}'\boldsymbol{t})/2\right\} \\ &= \exp\left\{\mathrm{i}(\boldsymbol{C}\boldsymbol{a})'\boldsymbol{t} - \boldsymbol{t}'(\boldsymbol{C}\boldsymbol{\Sigma}\boldsymbol{C}')\boldsymbol{t}/2\right\}, \end{aligned}$$

得证.

注: 通常称性质6为正态变量的线性变换不变性.

推论 (I)

若$\boldsymbol{\xi}$服从n元正态分布$N(\boldsymbol{a}, \boldsymbol{\Sigma})$, 则存在一个正交变换$\boldsymbol{U}$, 使得$\boldsymbol{\eta} = \boldsymbol{U}\boldsymbol{\xi}$是一个具有独立正态分量的随机向量, 它的数学期望是$\boldsymbol{U}\boldsymbol{a}$, 而它的方差分量是$\boldsymbol{\Sigma}$的特征值.

推论 (I)

若$\boldsymbol{\xi}$服从n元正态分布$N(\boldsymbol{a},\boldsymbol{\Sigma})$，则存在一个正交变换$\boldsymbol{U}$，使得$\boldsymbol{\eta}=\boldsymbol{U}\boldsymbol{\xi}$是一个具有独立正态分量的随机向量，它的数学期望是$\boldsymbol{U}\boldsymbol{a}$，而它的方差分量是$\boldsymbol{\Sigma}$的特征值.

证：从矩阵论知道，对实对称矩阵$\boldsymbol{\Sigma}$，存在正交矩阵$\boldsymbol{U}$，使得$\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{U}'=\boldsymbol{D}$，其中

$$\boldsymbol{D}=\begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{pmatrix},$$

这里$d_1,\cdots,d_n$是$\boldsymbol{\Sigma}$的特征值. 若$\boldsymbol{\Sigma}$的秩为r，则有r个特征值不为零. 此处的$\boldsymbol{U}$是以特征向量为列构成的正交阵. 把这里的$\boldsymbol{U}$作为性质6中的变换矩阵即可得此推论.

此推论说明, 对于多元正态向量, 可以进行正交变换, 使其既保持正态性不变又可让各分量独立, 这种方法在数理统计中十分有用.

此推论说明，对于多元正态向量，可以进行正交变换，使其既保持正态性不变又可让各分量独立，这种方法在数理统计中十分有用.

推论 (II)

在正交变换下，多元正态变量保持其独立性、同方差性不变.

此推论说明, 对于多元正态向量, 可以进行正交变换, 使其既保持正态性不变又可让各分量独立, 这种方法在数理统计中十分有用.

推论 (II)

在正交变换下, 多元正态变量保持其独立性、同方差性不变.

证: 设$\boldsymbol{\xi} = (\xi_1, \cdots, \xi_n)'$, 诸$\xi_i$相互独立具有相同的方差$\sigma^2$, 则它们的协方差矩阵$\mathrm{Cov}(\boldsymbol{\xi}) = \sigma^2 \boldsymbol{I}$, 其中$\boldsymbol{I}$是$n$阶单位阵. 若$\boldsymbol{U}$是一个正交阵, $\boldsymbol{\eta} = \boldsymbol{U}\boldsymbol{\xi}$, 则由性质6可知$\boldsymbol{\eta}$服从正态分布, 其协方差矩阵为

$$\boldsymbol{U}\sigma^2\boldsymbol{I}\boldsymbol{U}' = \sigma^2\boldsymbol{I}.$$

因此$\boldsymbol{\eta} = (\eta_1, \cdots, \eta_n)'$的诸分量仍是相互独立且具有相同方差.

推论（III）

若$\boldsymbol{\xi} \sim N(\boldsymbol{a}, \boldsymbol{\Sigma})$, 其中$\boldsymbol{\Sigma}$是$n$阶正定矩阵, 则

$$(\boldsymbol{\xi} - \boldsymbol{a})' \boldsymbol{\Sigma}^{-1} (\boldsymbol{\xi} - \boldsymbol{a}) \sim \chi_n^2.$$

推论 (III)

若$\boldsymbol{\xi} \sim N(\boldsymbol{a}, \boldsymbol{\Sigma})$, 其中$\boldsymbol{\Sigma}$是$n$阶正定矩阵, 则

$$(\boldsymbol{\xi}-\boldsymbol{a})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\xi}-\boldsymbol{a}) \sim \chi^2_n.$$

证: 利用分解式$\boldsymbol{\Sigma} = \boldsymbol{L}\boldsymbol{L}'$($\boldsymbol{L}$是非奇异阵),

$$\begin{aligned}(\boldsymbol{\xi}-\boldsymbol{a})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\xi}-\boldsymbol{a}) =&(\boldsymbol{\xi}-\boldsymbol{a})'(\boldsymbol{L}\boldsymbol{L}')^{-1}(\boldsymbol{\xi}-\boldsymbol{a})\\ =&[\boldsymbol{L}^{-1}(\boldsymbol{\xi}-\boldsymbol{a})]'[\boldsymbol{L}^{-1}(\boldsymbol{\xi}-\boldsymbol{a})]\\ :=&\boldsymbol{\zeta}'\boldsymbol{\zeta}.\end{aligned}$$

由性质6可知$\boldsymbol{\zeta}$是均值为$\boldsymbol{0}$的n维正态变量, 其协方差矩阵为

$$\boldsymbol{L}^{-1}\boldsymbol{\Sigma}(\boldsymbol{L}^{-1})' = \boldsymbol{L}^{-1}\boldsymbol{L}\boldsymbol{L}'(\boldsymbol{L}^{-1})' = \boldsymbol{I}.$$

从而$\boldsymbol{\zeta} = (\zeta_1, \cdots, \zeta_n)'$的各分量是相互独立的标准正态分布, 因此

$$\boldsymbol{\zeta}'\boldsymbol{\zeta} = \zeta_1^2 + \cdots + \zeta_n^2 \sim \chi^2_n.$$

例2 设$\boldsymbol{\xi} = (\xi_1, \xi_2)' \sim N(a_1, a_2, \sigma^2, \sigma^2, r)$, 求证$\eta_1 = \xi_1 + \xi_2$与$\eta_2 = \xi_1 - \xi_2$相互独立. 并问$\eta_1, \eta_2$各自服从怎样的分布?

例2 设$\boldsymbol{\xi}=(\xi_1,\xi_2)'\sim N(a_1,a_2,\sigma^2,\sigma^2,r)$, 求证$\eta_1=\xi_1+\xi_2$与$\eta_2=\xi_1-\xi_2$相互独立. 并问$\eta_1,\eta_2$各自服从怎样的分布?

解: 我们利用随机向量的线性变换来求解.

$$\begin{pmatrix}\eta_1\\ \eta_2\end{pmatrix}=\begin{pmatrix}\xi_1+\xi_2\\ \xi_1-\xi_2\end{pmatrix}=\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}\begin{pmatrix}\xi_1\\ \xi_2\end{pmatrix}:=\boldsymbol{C}\boldsymbol{\xi},$$

$$\boldsymbol{\Sigma}=\sigma^2\begin{pmatrix}1 & r\\ r & 1\end{pmatrix},\quad \boldsymbol{C}\boldsymbol{\Sigma}\boldsymbol{C}'=2\sigma^2\begin{pmatrix}1+r & 0\\ 0 & 1-r\end{pmatrix}.$$

所以$\mathrm{Cov}(\eta_1,\eta_2)=0$, η_1与η_2相互独立. 又因为

$$\boldsymbol{C}\boldsymbol{a}=\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}\begin{pmatrix}a_1\\ a_2\end{pmatrix}=\begin{pmatrix}a_1+a_2\\ a_1-a_2\end{pmatrix},$$

所以

$$\eta_1\sim N(a_1+a_2,2\sigma^2(1+r)),\quad \eta_2\sim N(a_1-a_2,2\sigma^2(1-r)).$$

假设$\boldsymbol{\xi} = (\boldsymbol{\xi}_1', \boldsymbol{\xi}_2')'$服从$n$元正态分布$N(\boldsymbol{a}, \boldsymbol{\Sigma})$, $\mathsf{E}\boldsymbol{\xi}_1 = \boldsymbol{a}_1$, $\mathsf{E}\boldsymbol{\xi}_2 = \boldsymbol{a}_2$, 协方差矩阵$\boldsymbol{\Sigma}$由(3.4.4)表出.

下面求在给定$\boldsymbol{\xi}_1 = \boldsymbol{x}_1$的条件下$\boldsymbol{\xi}_2$的条件密度函数. 假定$|\boldsymbol{\Sigma}| \neq 0$, 即只讨论非奇异的场合, 这时$|\boldsymbol{\Sigma}_{11}| \neq 0$.

首先, 我们来寻找一个线性变换

$$\boldsymbol{\eta}_1 = \boldsymbol{\xi}_1, \quad \boldsymbol{\eta}_2 = \boldsymbol{T}\boldsymbol{\xi}_1 + \boldsymbol{\xi}_2.$$

这个线性变换使$\boldsymbol{\eta}_1 = \boldsymbol{\xi}_1$, 此外我们还要求它使得$\boldsymbol{\eta}_1$与$\boldsymbol{\eta}_2$独立. 由性质4知道为使$\boldsymbol{\eta}_1$与$\boldsymbol{\eta}_2$独立, 只须它们的相互协方差矩阵为零矩阵. 但是

$$\begin{aligned}
&\mathsf{E}(\boldsymbol{\eta}_1 - \mathsf{E}\boldsymbol{\eta}_1)(\boldsymbol{\eta}_2 - \mathsf{E}\boldsymbol{\eta}_2)' \\
=&\mathsf{E}(\boldsymbol{\xi}_1 - \mathsf{E}\boldsymbol{\xi}_1)(\boldsymbol{T}\boldsymbol{\xi}_1 + \boldsymbol{\xi}_2 - \boldsymbol{T}\mathsf{E}\boldsymbol{\xi}_1 - \mathsf{E}\boldsymbol{\xi}_2)' \\
=&\boldsymbol{\Sigma}_{11}\boldsymbol{T}' + \boldsymbol{\Sigma}_{12}.
\end{aligned}$$

因此应取

$$\boldsymbol{T} = -\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}.$$

这时,

$$\boldsymbol{\eta}_2 = -\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\xi}_1 + \boldsymbol{\xi}_2,$$

所求的线性变换为

$$\begin{pmatrix} \boldsymbol{\eta}_1 \\ \boldsymbol{\eta}_2 \end{pmatrix} = \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ -\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1} & \boldsymbol{I} \end{pmatrix} \begin{pmatrix} \boldsymbol{\xi}_1 \\ \boldsymbol{\xi}_2 \end{pmatrix}. \tag{3.4.6}$$

为求$\boldsymbol{\eta}_1$及$\boldsymbol{\eta}_2$的概率密度函数, 先计算

$$\mathsf{E}\boldsymbol{\eta}_1 = \boldsymbol{a}_1,\ \mathsf{Var}\boldsymbol{\eta}_1 = \boldsymbol{\Sigma}_{11},\ \mathsf{E}\boldsymbol{\eta}_2 = \boldsymbol{a}_2 - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{a}_1,$$

$$\begin{aligned}
&\mathsf{Cov}(\boldsymbol{\eta}_2) = \mathsf{E}(\boldsymbol{\eta}_2 - \mathsf{E}\boldsymbol{\eta}_2)(\boldsymbol{\eta}_2 - \mathsf{E}\boldsymbol{\eta}_2)' \\
=&\mathsf{E}[(\boldsymbol{\xi}_2 - \boldsymbol{a}_2) - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}(\boldsymbol{\xi}_1 - \boldsymbol{a}_1)][(\boldsymbol{\xi}_2 - \boldsymbol{a}_2) - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}(\boldsymbol{\xi}_1 - \boldsymbol{a}_1)]' \\
=&\boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12} + \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{11}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12} \\
=&\boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}.
\end{aligned}$$

又变换(3.4.6)的雅可比行列式为1, 所以有

$$p_{\boldsymbol{\xi}}(\boldsymbol{x}_1,\boldsymbol{x}_2)=p_{\boldsymbol{\eta}}(\boldsymbol{y}_1,\boldsymbol{y}_2)=p_{\boldsymbol{\eta}_1}(\boldsymbol{y}_1)p_{\boldsymbol{\eta}_2}(\boldsymbol{y}_2),$$

这里$\boldsymbol{y}_1=\boldsymbol{x}_1$, $\boldsymbol{y}_2=-\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{x}_1+\boldsymbol{x}_2$, 显然$p_{\boldsymbol{\xi}_1}(\boldsymbol{x}_1)=p_{\boldsymbol{\eta}_1}(\boldsymbol{y}_1)$, 因此在给定$\boldsymbol{\xi}_1=\boldsymbol{x}_1$的条件下, $\boldsymbol{\xi}_2$的条件密度函数为

$$\begin{aligned}p_{\boldsymbol{\xi}_2|\boldsymbol{\xi}_1}(\boldsymbol{x}_2|\boldsymbol{\xi}_1=\boldsymbol{x}_1)=&\frac{p_{\boldsymbol{\xi}}(\boldsymbol{x}_1,\boldsymbol{x}_2)}{p_{\boldsymbol{\xi}_1}(\boldsymbol{x}_1)}\\=&\frac{p_{\boldsymbol{\eta}_1}(\boldsymbol{y}_1)p_{\boldsymbol{\eta}_2}(\boldsymbol{y}_2)}{p_{\boldsymbol{\eta}_1}(\boldsymbol{y}_1)}=p_{\boldsymbol{\eta}_2}(\boldsymbol{y}_2)\\=&p_{\boldsymbol{\eta}_2}(\boldsymbol{x}_2-\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{x}_1).\end{aligned}$$

因为

$$\boldsymbol{\eta}_2\sim N(\boldsymbol{a}_2-\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{a}_1,\boldsymbol{\Sigma}_{22}-\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}),$$

所以在$\boldsymbol{\xi}_1=\boldsymbol{x}_1$的条件下$\boldsymbol{\xi}_2$的条件分布为

$$N(\boldsymbol{a}_2+\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}(\boldsymbol{x}_1-\boldsymbol{a}_1),\boldsymbol{\Sigma}_{22}-\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}).$$

综上所述, 我们有下面的结论:

综上所述, 我们有下面的结论:

性质7: 若$\boldsymbol{\xi} = (\boldsymbol{\xi}_1', \boldsymbol{\xi}_2')'$服从$n$元正态分布$N(\boldsymbol{a}, \boldsymbol{\Sigma})$, $\mathsf{E}\boldsymbol{\xi}_1 = \boldsymbol{a}_1$, $\mathsf{E}\boldsymbol{\xi}_2 = \boldsymbol{a}_2$, 协方差矩阵$\boldsymbol{\Sigma}$由(3.4.4)表出. 则在给定$\boldsymbol{\xi}_1 = \boldsymbol{x}_1$的条件下$\boldsymbol{\xi}_2$的条件分布还是正态分布, 其条件数学期望为

$$\boldsymbol{a}_{2\cdot 1} = \mathsf{E}(\boldsymbol{\xi}_2|\boldsymbol{\xi}_1 = \boldsymbol{x}_1) = \boldsymbol{a}_2 + \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}(\boldsymbol{x}_1 - \boldsymbol{a}_1), \tag{3.4.7}$$

其条件方差为

$$\boldsymbol{\Sigma}_{22\cdot 1} = \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}. \tag{3.4.8}$$

这里$\mathsf{E}(\boldsymbol{\xi}_2|\boldsymbol{\xi}_1)$称为$\boldsymbol{\xi}_2$关于$\boldsymbol{\xi}_1$的回归. 注意它是$\boldsymbol{\xi}_1$的线性函数, 条件方差$\boldsymbol{\Sigma}_{22\cdot 1}$与$\boldsymbol{\xi}_1$无关. 此外, $\boldsymbol{\xi}_1$与$\boldsymbol{\xi}_2 - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\xi}_1$独立.

例3 二元场合，若$\boldsymbol{\xi} = (\xi_1, \xi_2)' \sim N(a_1, a_2, \sigma_1^2, \sigma_2^2, r)$，则

$$\boldsymbol{a} = \begin{pmatrix} \boldsymbol{a}_1 \\ \boldsymbol{a}_2 \end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & r\sigma_1\sigma_2 \\ r\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}.$$

在给定$\xi_1 = x_1$的条件下ξ_2的条件分布还是正态分布，其条件期望由(3.4.7)推知为

$$\mathsf{E}(\xi_2 | \xi_1 = x_1) = a_2 + r\frac{\sigma_2}{\sigma_1}(x_1 - a_1),$$

条件方差由(3.4.8)推知为

$$\sigma_2^2 - \frac{(r\sigma_1\sigma_2)^2}{\sigma_1^2} = \sigma_2^2(1 - r^2).$$