$F'(x)$,则 $F'(x)$就是 X 的分布密度."

这个事实的证明,作为习题(习题七第 16 题)留给读者.

例 4.6 已知 $X \sim N(\mu,\sigma^2)$, $Y=a+bX$(a,b 为常数,$b\neq 0$).求 Y 的概率分布.

解 设 $b>0$,易知 $F_Y(y)=P\{Y\leqslant y\}=P\{a+bX\leqslant y\}$ $=P\left\{X\leqslant\frac{y-a}{b}\right\}=F_X\left(\frac{y-a}{b}\right)$.

类似地,若 $b<0$,则

$$F_Y(y)=P\left\{X\geqslant\frac{y-a}{b}\right\}=1-F_X\left(\frac{y-a}{b}\right)$$

由于 $F_Y(y)$有连续的导函数 $F'_Y(y)$.故得到 Y 的密度函数:

$$p_Y(y)=F'_Y(y)=\begin{cases}\frac{1}{b}p_X\left(\frac{y-a}{b}\right) & \text{当}\, b>0\\ -\frac{1}{b}p_X\left(\frac{y-a}{b}\right) & \text{当}\, b<0\end{cases}$$

这里 $p_X(x)$是 X 的密度函数.总之,

$$p_Y(y)=\frac{1}{|b|}p_X\left(\frac{y-a}{b}\right)=\frac{1}{\sqrt{2\pi}|b|\sigma}\mathrm{e}^{-\frac{(y-a-b\mu)^2}{2b^2\sigma^2}}$$

由此知 $Y\sim N(a+b\mu,b^2\sigma^2)$.本例表明,服从正态分布的随机变量经线性变换后,仍服从正态分布.

特别地,有以下重要事实(正态随机变量的标准化):

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

例 4.7 对圆片直径进行测量,其值在[5,6]上均匀分布,求圆片面积的概率分布.

解 设圆片的直径的测量值为 X,面积为 Y,则有

$$Y=\frac{\pi X^2}{4}$$

按已知条件,X 的分布函数为

$$F_X(x)=\begin{cases}0 & \text{当} \quad x<5 \\ x-5 & \text{当} \quad 5\leqslant x<6 \\ 1 & \text{当} \quad x\geqslant 6\end{cases}$$

于是，当 $y>0$ 时 $F_Y(y)=P\left\{\frac{\pi X^2}{4}\leqslant y\right\}=P\left\{-\sqrt{\frac{4y}{\pi}}\leqslant X\leqslant\sqrt{\frac{4y}{\pi}}\right\}=P\left\{X\leqslant\sqrt{\frac{4y}{\pi}}\right\}=F_X\left(\sqrt{\frac{4y}{\pi}}\right)$. 当 $y\leqslant 0$ 时显然有 $F_Y(y)=0$. 总之可得：

$$F_Y(y)=\begin{cases}0 & \text{当} \quad y<\frac{25}{4}\pi \\ \sqrt{\frac{4y}{\pi}}-5 & \text{当} \quad \frac{25}{4}\pi\leqslant y<9\pi \\ 1 & \text{当} \quad y\geqslant 9\pi\end{cases}$$

可求得密度函数如下：

$$p_Y(y)=\begin{cases}\frac{1}{\sqrt{\pi y}} & \text{当} \quad \frac{25}{4}\pi\leqslant y\leqslant 9\pi \\ 0 & \text{其他}\end{cases}$$

本例的证明过程中，还用到了以下事实：

“如果随机变量 X 的分布函数 $F(x)$ 满足以下条件：

① $F(x)$ 连续；

② 存在 $x_1<x_2<\cdots<x_n\ (n\geqslant 1)$，在区间 $(-\infty,x_1),(x_1,x_2),\cdots,(x_{n-1},x_n),(x_n,\infty)$ 上 $F'(x)$ 存在且连续.

令 $$f(x)=\begin{cases}F'(x) & \text{当} \quad F'(x)\text{存在时} \\ 0 & \text{当} \quad F'(x)\text{不存在时}\end{cases}$$

则 $f(x)$ 是 X 的分布密度.”(见习题七第 17 题.)

例 4.8 设 X 的密度函数为 $p_X(x)$，又函数 $f(x)$ 的导数 $f'(x)$ 连续且处处大于零. 求 $Y=f(X)$ 的分布密度 $p_Y(y)$.

解 因为 $f(x)$ 是严格增加的连续函数，可设其值域为 $(A,B)\ (-\infty\leqslant A<B\leqslant\infty)$，反函数为 $g(y)$（它在 (A,B) 上有定

义)，当然 $g'(y)$存在. 任给定 $y\in(A,B)$，易知$\{f(X)\leqslant y\}=\{X\leqslant g(y)\}$，故 $F_Y(y)=P\{Y\leqslant y\}=P\{f(X)\leqslant y\}=P\{X\leqslant g(y)\}=F_X[g(y)]$.

当 $y\leqslant A$ 时，$\{f(X)\leqslant y\}$ 是不可能事件，故 $F_Y(y)=P\{f(X)\leqslant y\}=0$.

当 $y\geqslant B$ 时，$\{f(X)\leqslant y\}$ 是必然事件，故 $F_Y(y)=P\{f(X)\leqslant y\}=1$.

总之，

$$F_Y(y)=\begin{cases}0 & \text{当}\quad y\leqslant A\\ F_X[g(y)] & \text{当}\quad y\in(A,B)\\ 1 & \text{当}\quad y\geqslant B\end{cases}$$

利用关系式(4.2)，不难知道，Y 的密度函数

$$p_Y(y)=\begin{cases}p_X[g(y)]g'(y) & \text{当}\quad y\in(A,B)\\ 0 & \text{当}\quad y\overline{\in}(A,B)\end{cases}$$

如果本例中的条件“$f'(x)$连续且处处大于零”改为“$f'(x)$连续且处处小于零”，则上面的推导过程需要稍作改变，此时求得的密度函数是

$$p_Y(y)=\begin{cases}p_X[g(y)][-g'(y)] & \text{当}\quad y\in(A,B)\\ 0 & \text{当}\quad y\overline{\in}(A,B)\end{cases}$$

需要注意的是，本例中的条件“$f'(x)$处处大于零”(或“$f'(x)$处处小于零”)是相当苛刻的，许多常见的函数 $f(x)$就不能满足. 此时只好直接去求所要的分布函数，而不要去考虑什么反函数了(参看例 4.7).

作为本节末尾，我们还要给出两个重要例子，由于严格的证明涉及较复杂的数学推理，初学者不必看所述的证明过程. 这两个例子实际是两个定理，在理论上和应用上都相当重要，例 4.10 还是“随机模拟”的理论基础.

* **例 4.9** 设随机变量 X 的分布函数是$F(x)$且 $F(x)$是连续

函数，则随机变量 $Y=F(X)$ 服从 $(0,1)$ 上均匀分布.

实际上，Y 取值属于 $(0,1)$. 往下只需证明对一切 $y\in(0,1)$，$P(Y\leqslant y)=y$. 不难知道：对给定的 $y\in(0,1)$，利用连续函数的性质必有 x_0 满足 $F(x_0)=y$ 且 $\{x:F(x)\leqslant y\}=(-\infty,x_0]$①. 于是 $P(Y\leqslant y)=P(F(X)\leqslant y)=P(X\leqslant x_0)=F(x_0)=y$.

*例 4.10** 设函数 $F(x)$ 具有下列性质：

(1) $0\leqslant F(x)\leqslant 1$ （一切实数 x）

(2) $F(x)$ 是 x 的不减函数

(3) $\lim\limits_{x\to-\infty}F(x)=0$, $\lim\limits_{x\to+\infty}F(x)=1$

(4) $F(x)$ 是右连续函数

$$g(y)=\min\{x:F(x)\geqslant y\}^{②}\quad(0<y<1)$$

若 U 是服从 $(0,1)$ 上均匀分布的随机变量，则随机变量 $X=g(U)$ 的分布函数恰好是上述的 $F(x)$.

实际上，从 $g(y)$ 的定义知，对任何固定的 $y\in(0,1)$，$F(x)\geqslant y$ 的充要条件是 $x\geqslant g(y)$. 于是 $P(X\leqslant x)=P(g(U)\leqslant x)=P(U\leqslant F(x))=F(x)$. 这表明 X 的分布函数是 $F(x)$.

习　题　七

1. 设随机变量 X 的概率密度为

$$p(x)=\begin{cases}Cx & 0\leqslant x\leqslant 1\\ 0 & \text{其他}\end{cases}$$

求(1) 常数 C；(2) X 落在区间 $(0.3,0.7)$ 内的概率.

2. 随机变量 X 的概率密度为

$$p(x)=\begin{cases}\dfrac{C}{\sqrt{1-x^2}} & |x|<1\\ 0 & \text{其他}\end{cases}$$

① 任给定 $y\in(0,1)$. 由 $F(x)$ 是连续函数，集合 $\{x:F(x)=y\}$ 中必有最小的数，取 x_0 为这个最小数即可.

② 利用 $F(x)$ 的右连续性，可以证明实数集合 $\{x:F(x)\geqslant y\}$（y 固定，$0<y<1$）里有最小数（记作 $g(y)$）. $g(y)$ 是 $F(x)$ 的（广义）反函数.

求(1) 常数 C;(2) X 落在区间$\left(-\frac{1}{2},\frac{1}{2}\right)$内的概率.

3. 随机变量 X 的概率密度为

$$p(x)=Ce^{-|x|},\ -\infty<x<+\infty$$

求(1) 常数 C;(2) X 落在区间(0,1)内的概率.

4. 设 $X\sim N(1,0.6^2)$,求 $P(X>0)$和 $P(0.2<X<1.8)$.

5. 设 $X\sim N(\mu,\sigma^2)$,对

$$P(\mu-k\sigma<X<\mu+k\sigma)=0.95,0.90,0.99$$

分别找出相应的 k 值(查表).又对于 k 的什么值有$P(X>\mu-k\sigma)=0.95$?

6. 乘以什么常数 C 将使Ce^{-x^2+x}变成概率密度函数?

7. 设 X 的密度为

$$p_X(x)=\begin{cases}\dfrac{2}{\pi(x^2+1)} & x>0\\ 0 & x\leqslant 0\end{cases}$$

求 $Y=\ln X$ 的密度.(当 $x\leqslant 0$ 时,规定 $\ln x=0$.)

8. 设 X 服从自由度为k 的χ^2分布:

$$p_X(x)=\begin{cases}\dfrac{1}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)}x^{\frac{k}{2}-1}e^{-\frac{x}{2}} & x>0\\ 0 & x\leqslant 0\end{cases}$$

求 $Y=\sqrt{X/k}$的密度.

9. 由统计物理学知道分子运动的速率 X 服从麦克斯韦(Maxwell)分布,即密度为

$$p_X(x)=\begin{cases}\dfrac{4x^2}{\alpha^3\sqrt{\pi}}e^{-\frac{x^2}{\alpha^2}} & x>0\\ 0 & x\leqslant 0\end{cases}$$

其中参数 $\alpha>0$.求分子的动能

$$Y=\frac{1}{2}mX^2$$

的密度.

10. 设 $\ln X\sim N(1,2^2)$,求 $P\left(\frac{1}{2}<X<2\right)$.($\ln 2=0.693$)

11. 对球的直径作测量,设其值均匀地分布在$[a,b]$内.求体积的密度函数.

12. 点随机地落在中心在原点，半径为 R 的圆周上，并且对弧长是均匀地分布的．求落点的横坐标的概率密度．

13. 设随机变量 X 的分布函数为

$$F(x)=\begin{cases}1-e^{-x} & x>0\\ 0 & x\leqslant 0\end{cases}$$

(1) 求 $P\{X\leqslant 2\}, P\{X>3\}$；

(2) 求 X 的密度函数 $p(x)$．

14. 设随机变量 X 的分布密度为

(1) $$p(x)=\begin{cases}\dfrac{2}{\pi}\sqrt{1-x^2} & -1<x<1\\ 0 & \text{其他}\end{cases}$$

(2) $$p(x)=\begin{cases}x & 0\leqslant x<1\\ 2-x & 1\leqslant x\leqslant 2\\ 0 & \text{其他}\end{cases}$$

求 X 的分布函数 $F(x)$，并作出(2)中 $p(x)$ 与 $F(x)$ 的图形．

15. 某产品的质量指标 $X\sim N(160,\sigma^2)$，若要求 $P\{120<X<200\}\geqslant 0.80$，问：允许 σ 最多为多少？

16. 如果 X 的分布函数 $F(x)$ 具有连续的导函数 $F'(x)$，试证：$F'(x)$ 是 X 的分布密度．

17*. 设 X 的分布函数 $F(x)$ 满足下列条件：

① $F(x)$ 连续；

② 存在 $x_1<x_2<\cdots<x_n\ (n\geqslant 1)$，在区间 $(-\infty,x_1),(x_1,x_2),\cdots,(x_{n-1},x_n),(x_n,+\infty)$ 上 $F'(x)$ 存在且连续．

令

$$f(x)=\begin{cases}F'(x) & \text{当 } F'(x) \text{ 存在时}\\ 0 & \text{当 } F'(x) \text{ 不存在时}\end{cases}$$

试证：$f(x)$ 是 X 的分布密度．

18*. 设随机变量 X 的分布函数 $F(x)$ 不是连续函数，试证明：随机变量 $Y=F(X)$ 一定不服从 $(0,1)$ 上的均匀分布．

第三章 随机变量的数字特征

知道了随机变量 X 的概率分布或概率密度(下面统称为概率分布)以后,X 的全部概率特性就都知道了. 但是在实际问题中概率分布较难确定,而它的某些数字特征却比较容易估算出来;并且不少问题中只要知道它的某些数字特征也就够了,而不必细致地了解它的详细的概率特性. 因此在对随机变量的研究中,某些数字特征的确定就很重要. 在这些数字特征中,期望和方差是最常用到的.

§1 离散型随机变量的期望

1. 期望的概念

设随机变量 X 的概率分布是:

X	x_1	x_2	$\cdots$	x_k	$\cdots$
p	p_1	p_2	$\cdots$	p_k	$\cdots$

我们希望能找到这样一个数值(仅仅是"一个数值"),它体现了 X 取值的"平均"大小,就类似通常一堆数字的平均数那样.

对于一堆数,比如:$-1.1,1.9,0.2,0.5,0.5$ 这 5 个数,它们的平均数是$\frac{1}{5}(-1.1+1.9+0.2+0.5+0.5)=0.4$. 可是,对于随机变量 X 而言,X 的可能值 $x_1,x_2,\cdots,x_k,\cdots$ 的和再除以总个数那种方式的"平均数",并不真正起到平均的作用(当可能值有无穷多个时,也无法确定那样的"平均数"). 例如,X 的分布如下:

X	100	200
p	0.01	0.99

(1.1)

作为可能值的平均数，是$\frac{1}{2}(100+200)=150$.但是(从直觉看来)这150并不真正体现 X 的取值的平均，这是对 X 的可能值"100"，"200"一视同仁的结果；然而，实际上，从分布看出，X 取"200"的机会比 X 取"100"的机会多得多.总之，要真正体现 X 的取值的平均，不能只由它取的什么值来决定，还要考虑到它取那些值的相应的概率.

那么，体现 X 的取值的"平均"那样一个数字特征究竟怎样来确定呢？

我们再来分析一下上一章那个放射性物质放出的粒子数的著名的例子.

我们问：对于这 2 608 次观察而言，放射性物质平均每次放出几个粒子？由例 2.2 的记录可知，它是：

$$\frac{1}{2\ 608}[0\times 57+1\times 203+2\times 383+3\times 525+\cdots$$
$$+10\times 16]^{①}=10\ 086/2\ 608=3.87$$

即平均每次放出 3.87 个粒子.我们可将上面的算式

$$\frac{1}{2\ 608}[0\times 57+1\times 203+2\times 383+3\times 525+\cdots+10\times 16]$$

改写为

$$0\times\frac{57}{2\ 608}+1\times\frac{203}{2\ 608}+2\times\frac{383}{2\ 608}+3\times\frac{525}{2\ 608}+\cdots$$
$$+10\times\frac{16}{2\ 608}$$
$$=0\times p_0^*+1\times p_1^*+2\times p_2^*+3\times p_3^*+\cdots+10\times p_{10}^*$$

---

① 这最后一项 10×16 并不准确，但影响很小.

这是一个和式；而每项又都是两个数的乘积，其中一个数是放射粒子数，而另一个数是相应的频率. 注意，这对放射粒子数 0,1,2,…,10 而言，形式上是加权平均. 由此，我们有

定义 1.1 设离散型随机变量 X 的概率分布是：

X	x_1	x_2	$\cdots$	x_k	$\cdots$
p	p_1	p_2	$\cdots$	p_k	$\cdots$

（即 $P\{X=x_k\}=p_k, k=1,2,\cdots$），则称和数

$$\sum_k x_k p_k（即\ x_1p_1+x_2p_2+\cdots+x_kp_k+\cdots）^{①} \qquad (1.2)$$

为随机变量 X 的**期望**（或数学期望），记作 $E(X)$.

对于(1.1)式定义的随机变量 X，它的期望是

$$E(X)=100\times0.01+200\times0.99=199$$

它与 200 非常靠近，而远不是 150.

显然，$E(X)$ 是一个实数. 当 X 的概率分布为已知时，$E(X)$ 可由(1.2)式算得. 它形式上是 X 的可能值的加权平均，实质上它体现了随机变量 X 取值的真正的“平均”. 为此，我们也称它为 X 的**均值**. 有时也称为分布的均值.

2. 几个常用分布的期望

(1) 二点分布

设 X 服从二点分布：

X	1	0
q	p	p

按(1.2)式，此时有

$$E(X)=1\cdot p+0\cdot q=p$$

(2) 二项分布

① 当 X 可能取的值有无穷多个时，定义要求该级数绝对收敛.

设 X 服从二项分布 $B(n,p)$，即

$$P\{X=k\}=C_n^k p^k q^{n-k} \quad (k=0,1,\cdots,n)$$

按(1.2)式，此时有

$$E(X)=\sum_{k=0}^{n} k\cdot P\{X=k\}=\sum_{k=1}^{n} kC_n^k p^k q^{n-k}$$

$$=\sum_{k=1}^{n}\frac{kn!}{k!\ (n-k)!}p^k q^{n-k}$$

$$=\sum_{k=1}^{n}\frac{np\cdot(n-1)!}{(k-1)!\ [(n-1)-(k-1)]!}\cdot p^{k-1}q^{(n-1)-(k-1)}$$

$$\xlongequal{令\ k'=k-1} np\sum_{k'=0}^{n-1}\frac{(n-1)!}{k'!\ [(n-1)-k']!}p^{k'}q^{(n-1)-k'}$$

$$=np(p+q)^{n-1}=np$$

(3) 泊松分布

设 X 服从泊松分布，即

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

按(1.2)式，此时有

$$E(X)=\sum_{k=0}^{\infty} k\cdot\frac{\lambda^k}{k!}e^{-\lambda}=e^{-\lambda}\sum_{k=1}^{\infty}\frac{\lambda^{k-1}}{(k-1)!}\cdot\lambda$$

$$=\lambda e^{-\lambda}\cdot e^{\lambda}=\lambda$$

(4) 超几何分布

设 X 服从参数为 N,M,n 的超几何分布[①]，即

$$P\{X=m\}=\frac{C_M^m C_{N-M}^{n-m}}{C_N^n} \quad (m=0,1,2,\cdots,l)$$

这里 $l=\min(M,n)$. 根据(1.2)式，并利用等式

$$\sum_{m=0}^{l} P\{X=m\}=1,$$

① 为简单计，我们假设 $n\leqslant N-M$.

不难推得

$$E(X)=\frac{nM}{N}$$

习　题　八

1. 已知随机变量 X 的概率分布为

$$P\{X=k\}=\frac{1}{10},\ k=2,4,\cdots,18,20$$

求 $E(X)$.

2. 两台生产同一种零件的车床，一天生产中次品数的概率分布分别是

甲	0	1	2	3	(次品数)
p	0.4	0.3	0.2	0.1	

乙	0	1	2	3	(次品数)
p	0.3	0.5	0.2	0	

如果两台机床的产量相同，问哪台机床好？

3. 某射手每次射击打中目标的概率都是 0.8，现连续向一目标射击，直到第一次击中为止. 求“射击次数”X 的期望.

4. 推导超几何分布的期望计算公式.

5. 盒中有五个球，其中有三白二黑，从中随机抽取两个球，求“抽得的白球数”X 的期望.

6. 射击比赛，每人射四次(每次一发)，约定全部不中得 0 分，只中一弹得 15 分，中二弹得 30 分，中三弹得 55 分，中四弹得 100 分. 甲每次射击命中率为 $\frac{3}{5}$，问他期望能得多少分？

7*. 某射手每次射中目标的概率是 p，现携有 10 发子弹准备对一目标连续射击(每次打一发)，一旦射中或子弹打完了就立刻转移到别的地方. 问：他在转移前平均射击几次？

§2　连续型随机变量的期望

我们的目的仍然是想找一个能反映随机变量取值的“平均”的一个数字特征，但是就连续型随机变量而言，显然如(1.2)那样的

和式已无意义.然而对于熟悉微积分的读者,不难理解下述定义.

定义 2.1 设连续型随机变量 X 的密度为 $p(x)$,称

$$\int_{-\infty}^{+\infty} xp(x)\mathrm{d}x \tag{2.1}$$

为 X 的**期望**(或均值)[①],记作 $E(X)$.

这里,先对定义作些解释.设 X 的密度为 $p(x)$.如图 3.1,在数轴上取分点(为方便起见,取等分点,而记相邻两点的距离为 λ):

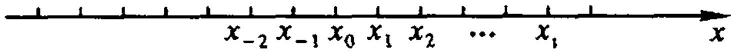

图 3.1

由上章(3.1)式知,X 落在区间 $(x_i, x_{i+1}]$ 中的概率为:

$$P\{x_i < X \leqslant x_{i+1}\} = \int_{x_i}^{x_{i+1}} p(x)\mathrm{d}x$$

$$(i = 0, \pm 1, \pm 2, \cdots)$$

现将 X 离散化,定义一个新的离散型随机变量:

$$X^* \triangleq x_i \text{ 当 } X \text{ 取值属于}(x_i, x_{i+1}]\text{时}$$

$$(i = 0, \pm 1, \pm 2, \cdots)$$

我们看到,X^* 的取值情况与 X 是相近的.而且 λ 越小,它们越接近.按(1.2)有

$$E(X^*) = \sum_{i=-\infty}^{+\infty} x_i P\{X^* = x_i\}$$

注意到 X^* 的定义,我们有

$$P\{X^* = x_i\} = P\{x_i < X \leqslant x_{i+1}\}$$

于是

---

① 本定义要求 $\int_{-\infty}^{+\infty} |x| p(x)\mathrm{d}x$ 收敛.

$$E(X^*) = \sum_i x_i P\{x_i < X \leqslant x_{i+1}\}$$ [1]

$$= \sum_i x_i \int_{x_i}^{x_{i+1}} p(x)\mathrm{d}x = \sum_i \int_{x_i}^{x_{i+1}} x_i p(x)\mathrm{d}x$$

由微积分有关知识可知,当 $\lambda \to 0$ 时,

$$E(X^*) \to \int_{-\infty}^{+\infty} xp(x)\mathrm{d}x$$

这表明数值 $\int_{-\infty}^{+\infty} xp(x)\mathrm{d}x$,确实反映了随机变量 X 取值的"平均".

下面求出几个常用分布的均值.

(1) 均匀分布

设 X 有密度 $p(x)$:

$$p(x) = \begin{cases} \dfrac{1}{b-a} & a \leqslant x \leqslant b \\ 0 & \text{其他} \end{cases} \quad (a < b)$$

按(2.1),

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

$$= \frac{1}{b-a} \cdot \frac{x^2}{2}\Big|_a^b = \frac{1}{2} \cdot \frac{b^2 - a^2}{b-a}$$

$$= \frac{1}{2}(b+a)$$

它恰是区间$[a,b]$的中点.这跟 $E(X)$的概率意义相符.

(2) 指数分布

设 X 有密度 $p(x)$:

---

① 类似地,对于任何随机变量 X(不必是连续型的),令

$$m(\lambda) = \sum_{i=-\infty}^{+\infty} x_i P(x_i < X \leqslant x_{i+1})\text{(假定级数绝对收敛)}$$

若 $\lim\limits_{\lambda \to 0} m(\lambda) = m$ 存在,则称 m 是 X 的期望.

$$p(x)=\begin{cases}\lambda e^{-\lambda x} & x\geqslant 0\\ 0 & \text{其他}\end{cases}\quad(\lambda>0)$$

于是

$$E(X)=\int_{-\infty}^{+\infty}xp(x)\mathrm{d}x=\lambda\int_{0}^{+\infty}xe^{-\lambda x}\mathrm{d}x$$

$$\xlongequal{\text{令 } t=\lambda x}\frac{1}{\lambda}\int_{0}^{+\infty}te^{-t}\mathrm{d}t$$

$$=\frac{1}{\lambda}\left[(-te^{-t})\Big|_{0}^{+\infty}+\int_{0}^{+\infty}e^{-t}\mathrm{d}t\right]=\frac{1}{\lambda}$$

(3) 正态分布

设 $X\sim N(\mu,\sigma^2)$,则

$$E(X)=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{+\infty}xe^{-\frac{1}{2\sigma^2}(x-\mu)^2}\mathrm{d}x$$

$$\xlongequal{\text{令 } x-\mu=t}\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{+\infty}(t+\mu)e^{-\frac{1}{2\sigma^2}t^2}\mathrm{d}t$$

$$=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{+\infty}te^{-\frac{t^2}{2\sigma^2}}\mathrm{d}t+\mu\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{+\infty}e^{-\frac{t^2}{2\sigma^2}}\mathrm{d}t$$

不难看出上式右边第一项等于零;第二项按上章(3.5)式,等于 μ.所以,

$$E(X)=\mu$$

这表明,正态分布的参数 μ 恰是该分布的均值.

顺便提一下,对于正态分布 $N(\mu,\sigma^2)$的参数 μ,在上一章中已指出,它的几何含义是,密度函数以 $x=\mu$ 为对称轴;这里又给出了它的概率含义,即 μ 是该分布的均值.其实从均值的定义不难看出,这是具有一般性的:即,若某密度函数以 $x=c$ 为对称轴(如果均值存在),则其均值必是 c.

(4) Γ 分布

设 $X\sim\Gamma(\alpha,\beta)$,即其密度函数为:

$$p(x)=\begin{cases}\dfrac{\beta^{\alpha}}{\Gamma(\alpha)}x^{\alpha-1}\mathrm{e}^{-\beta x} & x>0\\ 0 & x\leqslant 0\end{cases}\quad (\alpha>0,\beta>0)$$

于是

$$\begin{aligned}E(X)&=\frac{\beta^{\alpha}}{\Gamma(\alpha)}\int_{0}^{+\infty}x^{\alpha}\mathrm{e}^{-\beta x}\mathrm{d}x\\ &\xlongequal{令\ t=\beta x}\frac{1}{\Gamma(\alpha)\cdot\beta}\int_{0}^{+\infty}t^{\alpha}\mathrm{e}^{-t}\mathrm{d}t\\ &=\frac{\Gamma(\alpha+1)}{\Gamma(\alpha)\cdot\beta}=\frac{\alpha}{\beta}\end{aligned}$$

§3 期望的简单性质及随机变量函数的期望公式

1. 期望的简单性质

我们在上两节介绍了期望的定义及几个常用分布的期望. 期望是分布的最重要和最基本的数字特征. 这里,对它的性质作一些初步的讨论. 下面先列出这些性质:

(1) $E(c)=c$;

(2) $E(kX)=kE(X)$;

(3) $E(X+b)=E(X)+b$; (3.1)

(4) $E(kX+b)=kE(X)+b$;

其中 k,b,c 都是常数. 我们分别给出证明.

对于(1)式,因为常量 c 作为随机变量而言,是个离散型随机变量,它只有一个可能值 c,概率为 1. 按(1.2)式,

$$E(c)=c\cdot 1=c.$$

对于(2)式. 当 $k=0$ 时,显然成立. 至于 $k\neq 0$ 的情形,先设 X 是离散型随机变量.

设 X 的概率分布是:

X	x_1	x_2	$\cdots$	x_m	$\cdots$
p	p_1	p_2	$\cdots$	p_m	$\cdots$

则随机变量 kX 的概率分布显然是:

kX	kx_1	kx_2	$\cdots$	kx_m	$\cdots$
p	p_1	p_2	$\cdots$	p_m	$\cdots$

于是按(1.2)式有

$$\begin{aligned} E(kX) &= kx_1p_1 + kx_2p_2 + \cdots + kx_mp_m + \cdots \\ &= k(x_1p_1 + x_2p_2 + \cdots + x_mp_m + \cdots) \\ &= kE(X) \end{aligned}$$

现在设 X 是连续型的随机变量,其密度函数为 $p(x)$,此时容易看出随机变量 kX 的密度函数为 $\frac{1}{|k|}p\left(\frac{x}{k}\right)$. 于是

$$\begin{aligned} E(kX) &= \int_{-\infty}^{+\infty} x\cdot\frac{1}{|k|}p\left(\frac{x}{k}\right)\mathrm{d}x \xlongequal{\text{令 } x=kt} \int_{-\infty}^{+\infty} ktp(t)\mathrm{d}t \\ &= k\int_{-\infty}^{+\infty} tp(t)\mathrm{d}t = kE(X) \end{aligned}$$

这表明,对于连续型的随机变量 X,(2)式也成立.

对于(3)式,这里只给出 X 是连续型时的证明.(X 是离散型时的证明,留给读者.)

设 X 的密度函数是 $p(x)$,$X+b$ 作为随机变量 X 的函数,不难看出它的密度函数 $q(y)$为:

$$q(y) = p(y-b)$$

于是按(2.1)式有

$$\begin{aligned} E(X+b) &= \int_{-\infty}^{+\infty} yq(y)\mathrm{d}y = \int_{-\infty}^{+\infty} yp(y-b)\mathrm{d}y \\ &\xlongequal{\text{令 } y-b=x} \int_{-\infty}^{+\infty} (x+b)p(x)\mathrm{d}x \\ &= \int_{-\infty}^{+\infty} xp(x)\mathrm{d}x + b = E(X)+b \end{aligned}$$

至于(4)式,由(2)、(3)两式可直接推出.

我们看到,(4)式实际上包括了(2)、(3)式.期望的更深刻的性质将在下一章谈到.

2. 随机变量函数的期望公式

在这里,我们给出两个重要公式.

设 X 的密度是 $p(x)$, Y 是 X 的函数: $Y=f(X)$,如果下式右边绝对收敛,则

$$E[f(X)]=\int_{-\infty}^{+\infty}f(x)p(x)\mathrm{d}x \qquad (3.2)$$

当 X 是离散型时,也有类似的公式.设 X 的概率分布是 $P\{X=x_i\}=p_i(i=1,2,\cdots)$,则 $Y=f(X)$的均值 $E[f(X)]$可按下式直接计算:

$$E[f(X)]=\sum_i f(x_i)p_i \qquad (3.3)$$

对于公式(3.2)作一点说明.

如果按定义计算 $E[f(X)]$,先需找出 $Y=f(X)$的密度 $p_Y(y)$,然后计算 $\int_{-\infty}^{+\infty}yp_Y(y)\mathrm{d}y$.然而,找出 $p_Y(y)$有时是很麻烦的.但(3.2)式告诉我们,不必去找 $p_Y(y)$,而直接利用已知的 $p(x)$及 $f(x)$,相乘后求积分即得.

我们先举些例子说明公式(3.2)如何应用以及(3.2)带来的好处.然后在后面用小字排印部分给出(3.2)的严格证明.

例 3.1 已知 $X\sim N(0,1)$,求 $E(X^2)$.

解法 1 用公式(3.2).(注意这里的 $f(x)$是 x^2.)

$$\begin{aligned}E(X^2)&=\int_{-\infty}^{+\infty}x^2\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}\mathrm{d}x\\&=-\int_{-\infty}^{+\infty}x\mathrm{d}\left(\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}\right)\\&=-\left[x\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}\right]_{-\infty}^{+\infty}+\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}\mathrm{d}x\end{aligned}$$

因 $x\mathrm{e}^{-\frac{x^2}{2}}\to0(x\to\infty)$,故第一项为零.至于第二项,因被积函数是

标准正态分布的密度函数,故为1.所以

$$E(X^2)=1$$

解法2 用分布函数法先求 $Y=X^2$ 的分布密度.显然,

$$F_Y(y)=P\{Y\leqslant y\}=P\{X^2\leqslant y\}$$

(1) 若 $y<0$,则$\{X^2\leqslant y\}$是不可能事件,故 $F_Y(y)=0$.

(2) 若 $y\geqslant 0$,则 $P\{X^2\leqslant y\}=P\{-\sqrt{y}\leqslant X\leqslant\sqrt{y}\}$

$$=\int_{-\sqrt{y}}^{\sqrt{y}}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt=2\int_0^{\sqrt{y}}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt=\int_0^{y}\frac{1}{\sqrt{2\pi}}x^{-\frac{1}{2}}e^{-\frac{x}{2}}dx$$

不难看出 Y 的分布密度是

$$p_Y(y)=\begin{cases}\dfrac{1}{\sqrt{2\pi}}y^{-\frac{1}{2}}e^{-\frac{y}{2}} & y>0\\ 0 & y\leqslant 0\end{cases}$$

于是

$$E(Y)=\int_{-\infty}^{+\infty}yp_Y(y)dy=\int_0^{+\infty}y\cdot\frac{1}{\sqrt{2\pi}}y^{-\frac{1}{2}}e^{-\frac{y}{2}}dy$$

$$\xlongequal{令\ y=t^2}\int_0^{+\infty}\frac{2t^2}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt$$

$$=-\int_0^{+\infty}\frac{1}{\sqrt{2\pi}}2td(e^{-\frac{t^2}{2}})=2\int_0^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt$$

$$=\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt=1$$

比较这两种解法,我们看到,解法1比解法2简便.解法1由于利用公式(3.2)而不必找出 X^2 的密度.

例3.2 已知 X 遵从$[0,2\pi]$上的均匀分布,求 $E(\sin X)$.

解 按公式(3.2),

$$E(\sin X)=\int_{-\infty}^{+\infty}\sin x\cdot p_X(x)dx$$

但已知

$$p_X(x)=\begin{cases}\dfrac{1}{2\pi} & 0\leqslant x\leqslant 2\pi\\ 0 & 其他\end{cases}$$

所以

$$E(\sin X)=\frac{1}{2\pi}\int_0^{2\pi}\sin x\mathrm{d}x=0$$

当然也可先求出 $Y=\sin X$ 的密度 $p_Y(y)$. 再由 $E(Y)=\int_{-\infty}^{+\infty}yp_Y(y)\mathrm{d}y$ 求出 $E(\sin X)=0$. 请读者自己完成，并跟上法比较.

我们还指出，(3.1)式的(4)式，即 $E(kX+b)=kE(X)+b$，也可用公式(3.2)直接得出.

事实上，记 $f(X)=kX+b$，则由公式(3.2)知

$$\begin{aligned}E(kX+b)&=\int_{-\infty}^{+\infty}(kx+b)p_X(x)\mathrm{d}x\\&=k\int_{-\infty}^{+\infty}xp_X(x)\mathrm{d}x+b\int_{-\infty}^{+\infty}p_X(x)\mathrm{d}x\\&=kE(X)+b\end{aligned}$$

这又一次得到(3.1)式的(4).

从以上的说明中，我们看到公式(3.2)确实提供了一个计算随机变量函数的期望的有力工具，望读者学会正确运用它.

均值公式的证明

我们给出公式(3.2)的数学证明. 由于证明中要用到较多的积分知识，只关心应用的读者不必阅读这个证明. 公式(3.3)的证明则比较容易，请读者自己完成.

设随机变量 X 的概率密度为 $p(x)$，$Y=f(X)$. 分两种情况. 首先考虑 Y 是离散型情形，其可能值是 $y_1,y_2,\cdots$(有限个或无穷个). 令 $A_i=\{x:f(x)=y_i\}$，则 $E(Y)=\sum_i y_iP(Y=y_i)=\sum_i y_iP(X$ 取值属于 $A_i)$ $=\sum_i y_i\int_{A_i}p(x)\mathrm{d}x=\sum_i\int_{A_i}f(x)p(x)\mathrm{d}x=\int_{-\infty}^{+\infty}f(x)p(x)\mathrm{d}x$ 这表明公式(3.2)成立. 其次考虑 Y 是连续型情形. 其概率密度为 $g(y)$. 这种情形推理较为复杂，先证下列引理：

引理 设随机变量 Y 有概率密度 $g(y)$ 且 $\int_{-\infty}^{+\infty}yg(y)\mathrm{d}y$ 绝对收敛，则

$$E(Y)=\int_0^{+\infty} P(Y>u)\mathrm{d}u-\int_0^{+\infty} P(Y<-u)\mathrm{d}u \tag{3.4}$$

实际上 $P(Y>u)=\int_u^{+\infty} g(y)\mathrm{d}y, P(Y<-u)=\int_{-\infty}^{-u} g(y)\mathrm{d}y$. 利用积分交换次序的性质,知

$$\int_0^{+\infty} P(Y>u)\mathrm{d}u=\int_0^{+\infty}\left[\int_u^{+\infty} g(y)\mathrm{d}y\right]\mathrm{d}u=\iint_{0<u<y} g(y)\mathrm{d}y\mathrm{d}u$$

$$=\int_0^{+\infty}\left[\int_0^y \mathrm{d}u\right]g(y)\mathrm{d}y=\int_0^{+\infty} yg(y)\mathrm{d}y$$

$$\int_0^{+\infty} P(Y<-u)\mathrm{d}u=\int_0^{+\infty}\left[\int_{-\infty}^{-u} g(y)\mathrm{d}y\right]\mathrm{d}u=\iint_{0<u<-y} g(y)\mathrm{d}y\mathrm{d}u$$

$$=\int_{-\infty}^0\left[\int_0^{-y}\mathrm{d}u\right]g(y)\mathrm{d}y=-\int_{-\infty}^0 yg(y)\mathrm{d}y$$

由此可见(3.4)中等号右端等于 $\int_{-\infty}^{+\infty} yg(y)\mathrm{d}y=E(Y)$. 故(3.4)成立.

既然 $Y=f(X)$,从(3.4)知

$$E(Y)=E[f(X)]=\int_0^{+\infty} P(f(X)>y)\mathrm{d}y-\int_0^{+\infty} P(f(X)<-y)\mathrm{d}y. \tag{3.5}$$

利用积分交换次序的性质,知

$$\int_0^{+\infty} P(f(X)>y)\mathrm{d}y=\int_0^{+\infty}\left[\int_{x:f(x)>y} p(x)\mathrm{d}x\right]\mathrm{d}y$$

$$=\int_{x:f(x)>0}\left[\int_0^{f(x)}\mathrm{d}y\right]p(x)\mathrm{d}x=\int_{\{x:f(x)>0\}} f(x)p(x)\mathrm{d}x$$

类似地, $\int_0^{+\infty} P(f(X)<-y)\mathrm{d}y=\int_0^{+\infty}\left[\int_{x:f(x)<-y} p(x)\mathrm{d}x\right]\mathrm{d}y=$

$$\int_{x:f(x)<0}\left[\int_0^{-f(x)}\mathrm{d}y\right]p(x)\mathrm{d}x=-\int_{x:f(x)<0} f(x)p(x)\mathrm{d}x$$

再利用(3.5)知, $EY=\int_{-\infty}^{+\infty} f(x)p(x)\mathrm{d}(x)$, 即(3.2)成立.

在证明过程中,我们使用了函数在集合上的积分,其含义是: $\int_E h(x)\mathrm{d}x \triangleq \int_{-\infty}^{+\infty} I_E(x)h(x)\mathrm{d}x$, 这里 $I_E(x)$是 E 的示性函数.

我们还指出,当 $f(X)$不是离散型也不是连续型的随机变量时,公式(3.2)仍成立(这时期望的定义见第 82 页上的注). 由于要用到更多的积分知

识，我们就不证了.

习 题 九

1. 设随机变量 X 的概率密度为

$$p(x)=\begin{cases}2x & 0\leqslant x\leqslant 1\\ 0 & \text{其他}\end{cases}$$

求 $E(X)$.

2. 设随机变量 X 的密度为

$$p(x)=\begin{cases}\dfrac{1}{\pi\sqrt{1-x^2}} & |x|<1\\ 0 & \text{其他}\end{cases}$$

求 $E(X)$.

3. 设随机变量 X 的概率密度为

$$p(x)=\frac{1}{2}e^{-|x|},\ -\infty<x<+\infty$$

求 $E(X)$.

4. 设 $X\sim N(0,\sigma^2)$，求 $E(X^n)$.

5. 对球的直径作近似测量，设其值均匀地分布在区间$[a,b]$内，求球体积的均值.

6. 点随机地落在中心在原点、半径为 R 的圆周上，并对弧长是均匀分布的.求落点横坐标的均值.

7. 设 X 的密度函数 $p(x)$满足：

$$p(c+x)=p(c-x)\quad(x>0)$$

其中 c 为一常数，又 $\int_{-\infty}^{+\infty}|x|p(x)\mathrm{d}x$ 收敛.求证：

$$E(X)=c$$

§4 方差及其简单性质

1. 方差的概念

随机变量的均值体现了随机变量取值平均的大小，它是随机变量的重要数字特征.但只知道均值是不够的，还应该知道随机变量

的取值如何在均值周围变化.正如一批统计数字,只知道它们的平均数也是不够的,还要知道它们的分散程度.

举一个通俗的例子,设有甲、乙两个女声小合唱队,都由五名队员组成,她们的身高为

甲队:1米60厘米,1米62厘米,1米59厘米,1米60厘米,1米59厘米.

乙队:1米80厘米,1米60厘米,1米50厘米,1米50厘米,1米60厘米.

不难算出甲、乙两队的平均身高都是1米60厘米.但乙队身高波动大,甲队身高波动小.单从身高来看,甲队比较整齐,演出的效果会好些.实际工作中,数据的波动程度是反映客观现象的一种指标.例如,产品的某种特性(如硬度)波动大,说明生产不够稳定.又如生物的某种特性(如血压,血球)波动大,表示该生物处于病态.所以对于一批数据,除了研究它的平均值以外,还应该研究它的波动程度.

对于给定的一批数据 $x_1, x_2, \cdots, x_n$,通常用数量

$$\frac{1}{n}[(x_1-\bar{x})^2+(x_2-\bar{x})^2+\cdots+(x_n-\bar{x})^2] \qquad (4.1)$$

或

$$\frac{1}{n-1}[(x_1-\bar{x})^2+(x_2-\bar{x})^2+\cdots+(x_n-\bar{x})^2]^{①} \qquad (4.1')$$

来刻画这批数据的分散程度,其中 $\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$.

我们看到:

这个数量((4.1)或(4.1′))综合考虑了这 n 个数值与它们的平均值 $\bar{x}$ 的偏离程度,因而它很好地刻画了这批数据的分散程度.(请读者分别对于甲、乙两个小合唱队的身高,按(4.1)式算出两个数量,并跟直观结果进行比较.)

---

① 当 n 较大时,(4.1)与(4.1′)所表示的数量差别不大

同样，对于随机变量的取值情况，我们也希望通过一个数字来反映该随机变量取值的分散程度.

定义 4.1 设离散型随机变量的概率分布是

$$P(X=x_k)=p_k, k=1,2,\cdots$$

则称和数①

$$\sum_k [x_k-E(X)]^2 p_k \tag{4.2}$$

为 X 的**方差**. 记作 $D(X)$. [显然，$D(X)\geqslant 0$.]

定义 4.2 设连续型随机变量 X 的密度是 $p(x)$，则称②

$$\int_{-\infty}^{+\infty} [x-E(X)]^2 p(x)\mathrm{d}x \tag{4.2'}$$

为 X 的方差，记作 $D(X)$. [同样有 $D(X)\geqslant 0$.]

由方差的定义，按上节随机变量函数的期望公式(3.2)与(3.3)，可知

$$D(X)=E[X-E(X)]^2 \tag{4.3}$$

这表示 X 的方差，就是 $Y=[X-E(X)]^2$ 的均值(注意 $E(X)$ 是一个确定的实数).

其实，也可以直接用(4.3)式作为方差的定义. 这样做可将离散型和连续型随机变量的方差的定义统一起来. 另外，由(4.3)式更能看出，方差是刻画分散性的指标. 我们也常常把随机变量的方差称为它的分布的方差.

为了便于计算方差，我们来推导下列计算公式：

$$D(X)=E(X^2)-[E(X)]^2 \tag{4.4}$$

证 (只就 X 是连续型随机变量的情形给出证明. 当 X 是离散型的情形，由读者自行证明.)

---

① 当 X 的可能值不是有限个时，要求级数 $\sum_k [x_k-E(X)]^2 p_k$ 收敛；如级数发散，则称 X 的方差不存在.

② 若积分发散则称 X 的方差不存在.

$$
\begin{aligned}
D(X) &= \int_{-\infty}^{+\infty} [x - E(X)]^2 p(x)\mathrm{d}x \\
&= \int_{-\infty}^{+\infty} [x^2 - 2xE(X) + E^2(X)] p(x)\mathrm{d}x \\
&= \int_{-\infty}^{+\infty} x^2 p(x)\mathrm{d}x - 2E(X)\int_{-\infty}^{+\infty} xp(x)\mathrm{d}x + \\
&\quad E^2(X)\int_{-\infty}^{+\infty} p(x)\mathrm{d}x \\
&= E(X^2) - 2E(X)\cdot E(X) + E^2(X)\cdot 1 \\
&= E(X)^2 - [E(X)]^2
\end{aligned}
$$

以上推导中出现的 $E^2(X)$,就是$[E(X)]^2$的另一种记法.上面倒数第二个等号用到了(3.2)式及 $E(X)$的定义.

2. 常用分布的方差

(1) 二点分布

$$E(X^2) = 1^2\cdot p + 0^2\cdot q = p$$

由 §1 知,$E(X) = p$,于是按(4.4)有

$$D(X) = E(X^2) - [E(X)]^2 = p - p^2 = pq$$

(2) 二项分布

由 §1 知,$E(X) = np$.

$$
\begin{aligned}
E(X^2) &= \sum_{k=0}^{n} k^2 C_n^k p^k q^{n-k} \\
&= \sum_{k=1}^{n} [k(k-1) + k]\frac{n!}{k!\ (n-k)!}p^k q^{n-k} \\
&= \sum_{k=1}^{n} [(k-1) + 1]\frac{n!}{(k-1)!\ (n-k)!}p^k q^{n-k} \\
&= \sum_{k=2}^{n} (k-1)\frac{n(n-1)(n-2)!}{(k-1)!\ (n-k)!}p^2\cdot p^{k-2} q^{(n-2)-(k-2)} \\
&\quad + \sum_{k=1}^{n} \frac{n!}{(k-1)!\ (n-k)!}p^k q^{n-k} \\
&\xlongequal{令\ k' = k-2} n(n-1)p^2\sum_{k'=0}^{n-2}\frac{(n-2)!}{k'!\ (n-2-k')!}\cdot
\end{aligned}
$$

$$p^{k'}q^{(n-2)-k'}+E(X)$$
$$=n(n-1)p^2+np$$

于是

$$D(X)=E(X^2)-[E(X)]^2$$
$$=n(n-1)p^2+np-n^2p^2=npq$$

(3) 泊松分布

由§1知,$E(X)=\lambda$.

$$E(X^2)=\sum_{k=0}^{\infty}k^2\cdot\frac{\lambda^k}{k!}e^{-\lambda}$$
$$=\sum_{k=1}^{\infty}(k-1+1)\frac{\lambda^k}{(k-1)!}e^{-\lambda}$$
$$=\sum_{k=2}^{\infty}\frac{\lambda^{k-2}\cdot\lambda^2}{(k-2)!}e^{-\lambda}+\sum_{k=1}^{\infty}\frac{\lambda^k}{(k-1)!}e^{-\lambda}$$
$$=\lambda^2+\lambda$$

于是

$$D(X)=(\lambda^2+\lambda)-\lambda^2=\lambda$$

(4) 均匀分布

$$p(x)=\begin{cases}\dfrac{1}{b-a} & a\leqslant x\leqslant b\\ 0 & 其他\end{cases}$$

由§1知,$E(X)=\frac{1}{2}(b+a)$.

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

于是

$$D(X)=\frac{1}{3}(b^2+ab+a^2)-\frac{1}{4}(b^2+2ab+a^2)$$
$$=\frac{1}{12}(b-a)^2$$

(5) 指数分布

$$p(x)=\lambda e^{-\lambda x},(x\geqslant 0)\quad (\lambda>0)^{①}$$

由§1知，$E(X)=\frac{1}{\lambda}$.

$$E(X^2)=\lambda\int_0^{+\infty}x^2e^{-\lambda x}dx=\frac{1}{\lambda^2}\int_0^{+\infty}t^2e^{-t}dt=\frac{2}{\lambda^2}$$

于是

$$D(X)=\frac{2}{\lambda^2}-\frac{1}{\lambda^2}=\frac{1}{\lambda^2}$$

(6) 正态分布

对于 $X\sim N(\mu,\sigma^2)$，由§1知，$E(X)=\mu$. 于是

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

$$\xlongequal{\text{令 } t=\frac{x-\mu}{\sigma}}\frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}t^2e^{-\frac{t^2}{2}}dt=\frac{-\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}t\,d(e^{-\frac{t^2}{2}})$$

$$=\frac{-\sigma^2}{\sqrt{2\pi}}\left[\left(te^{-\frac{t^2}{2}}\right)_{-\infty}^{+\infty}-\int_{-\infty}^{+\infty}e^{-\frac{t^2}{2}}dt\right]$$

$$=\frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{-\frac{t^2}{2}}dt=\sigma^2$$

(7) Γ分布

设 $X\sim\Gamma(\alpha,\beta)$，即其密度为

$$p(x)=\frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x}\quad (x>0)$$

由§1知，$E(X)=\frac{\alpha}{\beta}$.

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

$$\xlongequal{\text{令 } t=\beta x}\frac{1}{\Gamma(\alpha)\beta^2}\int_0^{+\infty}t^{\alpha+1}e^{-t}dt$$

① 当 $x<0$ 时 $p(x)=0$. 以后，为方便计，密度函数等于零的部分常不写出来.

$$=\frac{\Gamma(\alpha+2)}{\Gamma(\alpha)\beta^2}=\frac{(\alpha+1)\alpha}{\beta^2}$$

于是

$$D(X)=E(X^2)-E^2(X)=\frac{\alpha}{\beta^2}$$

3. 方差的简单性质

由均值的性质，当 k,b,c 为常数时，不难推证：

(1) $D(c)=0$;

(2) $D(kX)=k^2D(X)$;

(3) $D(X+b)=D(X)$; (4.5)

(4) $D(kX+b)=k^2D(X)$.

这里只给出(2)的证明，其余的证明留给读者.

证 由(4.3)及(3.1)的(2)，我们有

$$\begin{aligned}D(kX)&=E[kX-E(kX)]^2=E[kX-kE(X)]^2\\&=k^2E[X-E(X)]^2=k^2D(X)\end{aligned}$$

§5 其 他

1. 切比雪夫不等式

这里顺便介绍一个重要的不等式——切比雪夫不等式.

定理 5.1 设随机变量 X 存在均值 $E(X)$ 与方差 $D(X)$，则有

$$P\{|X-E(X)|\geqslant\varepsilon\}\leqslant\frac{D(X)}{\varepsilon^2}\quad(\varepsilon>0)\tag{5.1}$$

证 （这里仅对连续型给出证明.离散型的情形，由读者给出.）

$$\begin{aligned}D(X)&=\int_{-\infty}^{+\infty}[x-E(X)]^2p(x)\mathrm{d}x\\&\geqslant\int_{-\infty}^{E(X)-\varepsilon}[x-E(X)]^2p(x)\mathrm{d}x+\end{aligned}$$

$$\int_{E(X)+\varepsilon}^{+\infty}[x-E(X)]^2 p(x)\mathrm{d}x$$
$$\geqslant \varepsilon^2\int_{-\infty}^{E(X)-\varepsilon} p(x)\mathrm{d}x+\varepsilon^2\int_{E(X)+\varepsilon}^{+\infty} p(x)\mathrm{d}x$$
$$=\varepsilon^2 P\{X\leqslant E(X)-\varepsilon\}+\varepsilon^2 P\{X\geqslant E(X)+\varepsilon\}$$
$$=\varepsilon^2 P\{|X-E(X)|\geqslant\varepsilon\}$$

由此即得(5.1).

在(5.1)中,取 $\varepsilon=k\sqrt{D(X)}$,则有

$$P\{|X-E(X)|\geqslant k\sqrt{D(X)}\}\leqslant\frac{1}{k^2} \tag{5.2}$$

特别地,取 $k=3$,则有

$$P\{|X-E(X)|\geqslant 3\sqrt{D(X)}\}\leqslant\frac{1}{9} \tag{5.3}$$

(请读者想一想当 $X\sim N(\mu,\sigma^2)$时,$P\{|X-E(X)|\geqslant 3\sigma\}=?$ 并与(5.3)式比较.)

由切比雪夫不等式(5.1)知,$D(X)$越小,则 X 取值越集中在 $E(X)$附近.我们由此进一步体会到特征数方差的概率含义——它刻画了随机变量的取值的分散程度.以后将会看到,切比雪夫不等式还是著名的大数定律的理论基础.

2. 原点矩与中心矩

这里只介绍一下名词.称

$$E(X^k)\quad(k=1,2,\cdots)$$

为 X 的 k 阶原点矩,记为 ν_k;称

$$E(X-E(X))^k\quad(k=1,2,\cdots)$$

为 X 的 k 阶中心矩,记为 μ_k.

显然 ν_1 就是均值 $E(X)$;μ_2 是方差 $D(X)$.高阶原点矩与高阶中心矩较少用到.

3. 分位数与中位数

给定 $p\in(0,1)$.称 x_p 是随机变量 X 的 p 分位数,若

$P(X<x_p)\leqslant p\leqslant P(X\leqslant x_p)$. 当 $p=\dfrac{1}{2}$时，p 分位数又叫中位数.
可以证明，p 分位数一定存在，但有时不惟一.

习 题 十

1. 对于习题八的习题 1,3 中的随机变量，分别求出它们的方差.

2. 对于习题九的习题 1,2,3,6 中的随机变量，分别求出它们的方差.

3. 设 X 服从参数为 $N,M,n(n\leqslant N-M)$ 的超几何分布，即

$$P\{X=m\}=\frac{C_M^m C_{N-M}^{n-m}}{C_N^n}\quad (m=0,1,2,\cdots,l)$$

这里 $l=\min(M,n)$. 试证明：

$$D(X)=\frac{nM(N-n)(N-M)}{N^2(N-1)}$$

4. 已知 $X\sim N(\mu,\sigma^2)$，$Y=e^X$，求 Y 的分布密度，并计算 $E(Y)$ 和 $D(Y)$.

5. 设轮船横向摇摆的随机振幅 X 的概率密度为

$$p(x)=Axe^{-x^2/2\sigma^2}\quad (x>0)$$

求(1) A;(2) 遇到大于其振幅均值的概率是多少？(3) X 的方差.

6. 设 X 的密度为

$$p(x)=\frac{x^m}{m!}e^{-x}\quad (x\geqslant 0)$$

试证

$$P\{0<X<2(m+1)\}\geqslant\frac{m}{m+1}$$

(提示：用切比雪夫不等式.)

7. 设 $X\sim$ 贝塔分布，即它的密度为

$$p(x)=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}\quad (0<x<1)$$

$$(\alpha>0,\beta>0)$$

求 $E(X),D(X)$.

8. 对某一目标进行射击，直到击中 r 次为止. 如果每次射击的命中率为 p，求需射击次数的均值与方差.

常用分布表

名称	概　率　分　布	均值	方差	参数的范围
二点分布	$P(X=x)=p^x q^{1-x}$ $(x=0,1)$	p	pq	$0<p<1$ $q=1-p$
二项分布	$P(X=x)=C_n^x p^x q^{n-x}$ $(x=0,1,\cdots,n)$	np	npq	$0<p<1$ $q=1-p$ n 自然数
泊松分布	$P(X=x)=\frac{\lambda^x}{x!}e^{-\lambda}$ $(x=0,1,2,\cdots)$	λ	λ	$\lambda>0$
超几何分布	$P(X=x)=\frac{C_{N-M}^{n-x}C_M^x}{C_N^n}$① $(x=0,1,\cdots,\min(M,n))$	$\frac{nM}{N}$	$\frac{n(N-n)(N-M)M}{N^2(N-1)}$	n,M,N 自然数 $n\leqslant N$ $M\leqslant N$
负二项分布	$P(X=x)=C_{r+x-1}^{r-1}p^r q^x$ $(x=0,1,2,\cdots)$	$\frac{rq}{p}$	$\frac{rq}{p^2}$	$0<p<1$ $q=1-p$ r 自然数
均匀分布	$p(x)=\frac{1}{b-a}(a\leqslant x\leqslant b)$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$b>a$
指数分布	$p(x)=\lambda e^{-\lambda x}(x>0)$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$\lambda>0$
正态分布	$p(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	μ	σ^2	μ 任意 $\sigma>0$
Γ分布	$p(x)=\frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x}$ $(x>0)$	$\frac{\alpha}{\beta}$	$\frac{\alpha}{\beta^2}$	$\alpha>0$ $\beta>0$

续表

名称	概　率　分　布	均值	方差	参数的范围
贝塔分布	$p(x)=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}\cdot(1-x)^{\beta-1}$ $(0<x<1)$	$\frac{\alpha}{\alpha+\beta}$	$\frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}$	$\alpha>0$ $\beta>0$
对数正态分布	$p(x)=\frac{1}{\sqrt{2\pi}\sigma x}e^{-\frac{(\lg x-\mu)^2}{2\sigma^2}}$ $(x>0)$	$e^{\mu+\frac{1}{2}\sigma^2}$	$e^{2\mu+\sigma^2}(e^{\sigma^2}-1)$	μ 任意 $\sigma>0$
韦布尔分布(Weibull)	$p(x)=\frac{mx^{m-1}}{\eta^m}e^{-\left(\frac{x}{\eta}\right)^m}$ $(x>0)$	$\eta\Gamma\left(1+\frac{1}{m}\right)$	$\eta^2\left[\Gamma\left(1+\frac{2}{m}\right)-\Gamma^2\left(1+\frac{1}{m}\right)\right]$	$m>0$ $\eta>0$

① 当 $i>k$ 时，规定 $C_k^i=0$.

第四章　随 机 向 量

在第二、三章中，我们讨论了随机变量的分布及其数字特征——均值和方差．但是，在很多随机现象中往往涉及多个随机变量．例如打靶时，弹着点就由两个随机变量——弹着点的横坐标 X 和纵坐标 Y——所构成．又如炼钢厂中炼出的每炉钢中钢的硬度、含碳量、含硫量都必须考察，就要用三个随机变量 X,Y,Z 来描述，这里 X 代表硬度，Y 代表含碳量，Z 代表含硫量．这类例子多得很，值得强调的是，这些随机变量之间一般说来又有某种联系，因而需要把这些随机变量作为一个整体(即向量)来研究．

定义 0.1　我们称 n 个随机变量 $X_1,X_2,\cdots,X_n$ 的整体 $\boldsymbol{\xi}=(X_1,X_2,\cdots,X_n)$ 为 n 维随机向量．

例如炮弹落点的位置 $\boldsymbol{\xi}=(X,Y)$ 就是一个二维随机向量，每炉钢的基本指标(X,Y,Z)，(硬度、含碳量、含硫量)就是一个三维随机向量．又若从一大批螺钉中，随机抽取五个，那么它们的直径 $X_1,X_2,\cdots,X_5$ 是五个随机变量，而 $\boldsymbol{\eta}=(X_1,X_2,\cdots,X_5)$ 就是一个五维随机向量．

"维数"的概念表示"共有几个分量"，从几何图形上看，二维随机向量可以看作是平面(二维空间！)上的"随机点"．三维随机向量可以看成是空间(三维空间！)中的"随机点"．上两章研究的随机变量是一维"随机向量"．

本章重点讨论二维随机向量．

§1 随机向量的(联合)分布与边缘分布

1. 二维离散型随机向量

和一维随机变量的情形类似.对二维随机向量,我们也只讨论离散型和连续型两大类.

定义 1.1 如果二维随机向量 $\xi=(X,Y)$可能取的值(向量!)只有有限个或者可列个(即可排成一个序列),则称 ξ 为离散型的①.

显然,如果 $\xi=(X,Y)$是离散型的,则 X,Y 都是一维离散型的随机变量,反过来也成立.

设 X 可能取的值是 $x_1,x_2,\cdots$(有限个或可列个),Y 的可能值是 $y_1,y_2,\cdots$(有限个或可列个),令 $E=\{(x_i,y_j):i=1,2,\cdots,j=1,2,\cdots\}$.显然 $\xi=(X,Y)$取的值都在 E 中.我们可以把 E 看作是 $\xi=(X,Y)$取值的范围,当然对某些 i,j,$\{\xi=(x_i,y_j)\}$可能是“不可能事件”.

和随机变量的情形一样,我们更关心 $\xi=(X,Y)$取值的概率:

$$P\{(X,Y)=(x_i,y_j)\}=p_{ij}\quad\begin{pmatrix}i=1,2,\cdots\\j=1,2,\cdots\end{pmatrix}\tag{1.1}$$

一般称(1.1)中的 p_{ij}(对一切 i,j)为 $\xi=(X,Y)$的概率分布,也称为(X,Y)的联合分布,有时用如下的概率分布表来表示:

① 数学上更确切的定义是:若存在有限个或可列个向量组成的集合 E,使“ξ 属于 E”为必然事件,则称 ξ 为离散型随机向量.

$$
\begin{array}{c|cccccc}
X \backslash Y & y_1 & y_2 & \cdots & y_j & \cdots \\
\hline
x_1 & p_{11} & p_{12} & \cdots & p_{1j} & \cdots \\
x_2 & p_{21} & p_{22} & \cdots & p_{2j} & \cdots \\
\vdots & \vdots & \vdots & & \vdots & \\
x_i & p_{i1} & p_{i2} & \cdots & p_{ij} & \cdots \\
\vdots & \vdots & \vdots & & \vdots &
\end{array}
\tag{1.1'}
$$

这些 p_{ij} 具有性质：

(1) $p_{ij} \geqslant 0 (i=1,2,\cdots, j=1,2,\cdots)$；

(2) $\sum_i \sum_j p_{ij} = 1.$ (1.2)

(1)式是显然的；至于(2)，利用概率的完全可加性（见第一章），就可推导出来．实际上，

$$
\begin{aligned}
\sum_i \sum_j p_{ij} &= \sum_i \sum_j P\{(X,Y)=(x_i,y_j)\} \\
&= P\left\{\sum_i \sum_j \{(X,Y)=(x_i,y_j)\}\right\} = P(U) \\
&= 1 (\text{这里 } U \text{ 表示必然事件.})
\end{aligned}
$$

例 1.1 设二维随机向量 (X,Y) 仅取 $(1,1)$，$(1.2,1)$，$(1.4,1.5)$，$(1,1.3)$，$(0.9,1.2)$ 五个点，且取它们的概率相同，则 (X,Y) 的联合分布为：

$$P\{(X,Y)=(1,1)\} = \frac{1}{5}$$

$$P\{(X,Y)=(1.2,1)\} = \frac{1}{5}$$

$$P\{(X,Y)=(1.4,1.5)\} = \frac{1}{5}$$

$$P\{(X,Y)=(1,1.3)\} = \frac{1}{5}$$

$$P\{(X,Y)=(0.9,1.2)\} = \frac{1}{5}$$

概率分布表为：

X \ Y	1	1.5	1.3	1.2
1	$\frac{1}{5}$	0	$\frac{1}{5}$	0
1.2	$\frac{1}{5}$	0	0	0
1.4	0	$\frac{1}{5}$	0	0
0.9	0	0	0	$\frac{1}{5}$

显然，本例中的 p_{ij} 满足(1.2)中的(1)，(2).

例 1.2 设(X,Y)的联合分布是：

$$P\{(X,Y)=(k_1,k_2)\}$$

$$=\frac{n!}{k_1!\ k_2!\ (n-k_1-k_2)!}p_1^{k_1}p_2^{k_2}(1-p_1-p_2)^{n-k_1-k_2}$$

$$k_1=0,1,\cdots,n,k_2=0,1,\cdots,n,k_1+k_2\leqslant n \tag{1.3}$$

其中 n 是给定的正整数；$0<p_1<1,0<p_2<1,p_1+p_2<1$.

它称为三项分布（参数 $n;p_1,p_2$）. 我们来验证它满足(1.2)式. 首先，显然有 $P\{(X,Y)=(k_1,k_2)\}\geqslant 0$，下面验证(1.2)中的(2)：

$$\sum_{k_1=0}^{n}\sum_{k_2=0}^{n-k_1}P\{(X,Y)=(k_1,k_2)\}$$

$$=\sum_{k_1=0}^{n}\sum_{k_2=0}^{n-k_1}\frac{n!}{k_1!\ k_2!\ (n-k_1-k_2)!}\quad p_1^{k_1}p_2^{k_2}(1-p_1-p_2)^{n-k_1-k_2}$$

$$=\sum_{k_1=0}^{n}\sum_{k_2=0}^{n-k_1}\frac{n!\ p_1^{k_1}}{k_1!\ (n-k_1)!}\frac{(n-k_1)!}{k_2!\ [(n-k_1)-k_2]!}$$
$$\cdot p_2^{k_2}(1-p_1-p_2)^{(n-k_1)-k_2}$$

$$=\sum_{k_1=0}^{n}\frac{n!}{k_1!\ (n-k_1)!}p_1^{k_1}\left\{\sum_{k_2=0}^{n-k_1}\frac{(n-k_1)!}{k_2!\ [(n-k_1)-k_2]!}\right.$$
$$\left.\cdot p_2^{k_2}(1-p_1-p_2)^{(n-k_1)-k_2}\right\}$$

$$= \sum_{k_1=0}^{n} \frac{n!}{k_1!\ (n-k_1)!} p_1^{k_1} [p_2 + (1 - p_1 - p_2)]^{n-k_1} = 1$$

(最后两个等号的依据是二项式定理.)

例 1.3 (三项分布的实例)一大批粉笔.其中 60% 是白的,25% 是黄的,15% 是红的.现从中随机地、顺序地取出 6 支,问这 6 支中恰有 3 支白、1 支黄、2 支红的概率.

解 用(白,白,白,黄,红,红)表示第一支是白的.第二支是白的,第三支是白的,第四支是黄的,第五支是红的,第六支是红的.由于是大批量,我们可以认为各次抽取是独立的,且抽取到黄、白、红的概率不变,有

$$\begin{aligned} &P\{(\text{白},\text{白},\text{白},\text{黄},\text{红},\text{红})\} \\ &= P(\text{白})P(\text{白})P(\text{白})P(\text{黄})P(\text{红})P(\text{红}) \\ &= (0.6)^3(0.25)(0.15)^2 \end{aligned}$$

于是

$$\begin{aligned} &P\{6\text{ 支中恰有 3 支白、1 支黄、2 支红}\} \\ &= m\cdot(0.6)^3(0.25)(0.15)^2 \end{aligned}$$

其中 m 是恰有 3 个白、1 个黄、2 个红的六维向量的个数.用关于组合的知识可知:

$$m = \frac{6!}{3!\ 1!\ 2!} = 60$$

因此,所求的概率为 $60\cdot(0.6)^3(0.25)(0.15)^2 = 0.072\,9$.

用随机向量的术语来说,若令

$$X = 6\text{ 支中白粉笔的数目}$$

$$Y = 6\text{ 支中黄粉笔的数目}$$

则事件"恰有 3 支白、1 支黄、2 支红"就是事件

$$\{X = 3, Y = 1\}$$

亦即事件

$$\{(X, Y) = (3, 1)\}$$

上面的结果表示为:

$$P\{(X, Y) = (3, 1)\} = \frac{6!}{3!\ 1!\ 2!}(0.6)^3(0.25)^1(0.15)^2$$

一般地,有(对于 $0 \leqslant k_1 \leqslant 6, 0 \leqslant k_2 \leqslant 6, k_1 + k_2 \leqslant 6$)

$$P\{6\text{ 支中恰有 } k_1 \text{ 支白、} k_2 \text{ 支黄、}(6 - k_1 - k_2)\text{ 支红}\}$$

$$=P\{(X,Y)=(k_1,k_2)\}$$

$$=\frac{6!}{k_1!\ k_2!\ (6-k_1-k_2)!}(0.6)^{k_1}(0.25)^{k_2}(0.15)^{6-k_1-k_2}$$

这就是例 1.2 中所说的，参数为 $n=6$；$p_1=0.6$，$p_2=0.25$ 的三项分布.

2. 边缘分布及其与联合分布的关系

对于二维随机向量 (X,Y)，分量 X 的概率分布称为 (X,Y) 的关于 X 的边缘分布；分量 Y 的概率分布称为 (X,Y) 的关于 Y 的边缘分布.

由于 (X,Y) 的联合分布全面地反映了 (X,Y) 的取值情况.因此，当我们已知 (X,Y) 的联合分布时，是可以求得关于 X 的或关于 Y 的边缘分布的.具体来说，

若已知

$$P\{(X,Y)=(x_i,y_j)\}=p_{ij}\quad\begin{pmatrix}i=1,2,\cdots\\ j=1,2,\cdots\end{pmatrix}$$

则随机变量 X 的概率分布(即关于 X 的边缘分布)：

$$P\{X=x_i\}=P\{X=x_i,U\}=P\left\{X=x_i,\ \sum_j\{Y=y_j\}\right\}$$

$$=P\left\{\sum_j\{X=x_i,Y=y_j\}\right\}=\sum_j P\{X=x_i,Y=y_j\}$$

$$=\sum_j P\{(X,Y)=(x_i,y_j)\}=\sum_j p_{ij}$$

(这里 U 是必然事件)

这样，我们得到了关于 X 的边缘分布：

$$P\{X=x_i\}=\sum_j p_{ij}\quad(i=1,2,\cdots)\tag{1.4}$$

类似可得关于 Y 的边缘分布：

$$P\{Y=y_j\}=\sum_i p_{ij}\quad(j=1,2,\cdots)\tag{1.4'}$$

例 1.4 继续讨论例 1.1.

(X,Y) 的概率分布表为：

$X \backslash Y$	1	1.5	1.3	1.2	
1	$\frac{1}{5}$	0	$\frac{1}{5}$	0	$\frac{2}{5}$
1.2	$\frac{1}{5}$	0	0	0	$\frac{1}{5}$
1.4	0	$\frac{1}{5}$	0	0	$\frac{1}{5}$
0.9	0	0	0	$\frac{1}{5}$	$\frac{1}{5}$
	$\frac{2}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	

求(X,Y)的边缘分布.

解 这里 X 的全体可能值 x_1,x_2,x_3,x_4 分别是 1,1.2,1.4,0.9;Y 的全体可能值 y_1,y_2,y_3,y_4 分别是 1,1.5,1.3,1.2.显然有:

$$\begin{pmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \\ p_{41} & p_{42} & p_{43} & p_{44} \end{pmatrix} = \begin{pmatrix} 1/5 & 0 & 1/5 & 0 \\ 1/5 & 0 & 0 & 0 \\ 0 & 1/5 & 0 & 0 \\ 0 & 0 & 0 & 1/5 \end{pmatrix}$$

于是按(1.4)式,

$$P\{X=x_1\} = \sum_{j=1}^{4} p_{1j} = p_{11}+p_{12}+p_{13}+p_{14}$$
$$= \frac{1}{5}+0+\frac{1}{5}+0=\frac{2}{5}$$

(即分布表第一行的数值的和)

类似地,

$$P\{X=x_2\} = \text{分布表第二行的数值的和} = \frac{1}{5}$$

$$P\{X=x_3\} = \text{分布表第三行的数值的和} = \frac{1}{5}$$

$$P\{X=x_4\}=\text{分布表第四行的数值的和}=\frac{1}{5}$$

而按(1.4′)式，

$$P\{Y=y_j\}=\sum_{i=1}^{4}p_{ij}=p_{1j}+p_{2j}+p_{3j}+p_{4j}$$
$$=\text{分布表第 } j \text{ 列的数值的和}(j=1,2,3,4)$$

其具体值见(X,Y)概率分布表第二个横线的下方.

例 1.5 (继续讨论上一段例 1.2)设(X,Y)的联合分布由(1.3)式给出,求关于 X 的边缘分布.

解 按(1.4)式,

$$P\{X=k_1\}=\sum_{k_2}P\{(X,Y)=(k_1,k_2)\}$$
$$(k_1=0,1,2,\cdots,n)$$

注意到(1.3)式中要求 $k_1+k_2\leqslant n$,即 $k_2\leqslant n-k_1$,所以

$$P\{X=k_1\}=\sum_{k_2=0}^{n-k_1}\frac{n!}{k_1!\;k_2!\;(n-k_1-k_2)!}p_1^{k_1}p_2^{k_2}(1-p_1-p_2)^{n-k_1-k_2}$$
$$=\frac{n!}{k_1!\;(n-k_1)!}p_1^{k_1}\sum_{k_2=0}^{n-k_1}\frac{(n-k_1)!}{k_2!\;(n-k_1-k_2)!}\cdot p_2^{k_2}(1-p_1-p_2)^{n-k_1-k_2}$$
$$(n,k_1\text{等跟求和指标 } k_2\text{无关})$$
$$=\frac{n!}{k_1!\;(n-k_1)!}p_1^{k_1}(p_2+1-p_1-p_2)^{n-k_1}$$
$$(\text{按二项式定理})$$
$$=\frac{n!}{k_1!\;(n-k_1)!}p_1^{k_1}(1-p_1)^{n-k_1}$$
$$(k_1=0,1,\cdots,n)$$

这表明 X 服从参数为 n,p_1 的二项分布.其实,由三项分布的实际背景,这个结果是很自然的.

3. 二维连续型随机向量的分布密度

和一维情形类似,我们给出以下定义:

定义 1.2 对于二维随机向量 $\xi=(X,Y)$,如果存在非负函

数 $p(x,y)(-\infty<x<+\infty,-\infty<y<+\infty)$，使对任意一个邻边分别平行于坐标轴的矩形区域 D："即由不等式$a<x<b,c<y<d$ 确定的区域"，有

$$P\{(X,Y)\in D\}=\iint_D p(x,y)\mathrm{d}x\mathrm{d}y \tag{1.5}$$

则称随机向量 $\xi=(X,Y)$为连续型的；并称 $p(x,y)$为 ξ 的分布密度，也称 $p(x,y)$为(X,Y)的联合分布密度(简称联合密度).

对于连续型的随机向量 $\xi=(X,Y)$，可以证明，对于平面上相当任意的集合 D，均有

$$P\{(X,Y)\in D\}=\iint_D p(x,y)\mathrm{d}x\mathrm{d}y \tag{1.6}$$

这里 $p(x,y)$是(X,Y)的联合分布密度.

公式(1.6)是本章最重要的公式之一，由于它的证明要用到较多的数学知识(包括对"相当任意的集合"下确切的定义)，我们就不证明了.希望读者注意理解这个公式的意义与用法.

下面对于联合密度 $p(x,y)$这个概念作几点说明：

(1) 这里的联合密度 $p(x,y)$与物理学中的质量面密度的概念相仿，请读者进行比较①.

(2) $p(x,y)$是一个全平面(即 $-\infty<x<+\infty,-\infty<y<+\infty$)上有定义的二元非负函数.以下，我们总假定 $p(x,y)$在 xy

---

① 由(1.5)式，当 $p(x,y)$在 $x=x_0,y=y_0$ 处连续时，利用重积分的性质容易推知

$$\lim_{\substack{\Delta x\to 0^+\\ \Delta y\to 0^+}}\frac{1}{\Delta x\Delta y}P\left\{x_0-\frac{\Delta x}{2}<X<x_0+\frac{\Delta x}{2},y_0-\frac{\Delta y}{2}<Y<y_0+\frac{\Delta y}{2}\right\}$$

$$=\lim_{\substack{\Delta x\to 0^+\\ \Delta y\to 0^+}}\frac{1}{\Delta x\Delta y}\int_{x_0-\frac{\Delta x}{2}}^{x_0+\frac{\Delta x}{2}}\int_{y_0-\frac{\Delta y}{2}}^{y_0+\frac{\Delta y}{2}}p(x,y)\mathrm{d}x\mathrm{d}y=p(x_0,y_0)$$

由此可见，若 $p(x,y),q(x,y)$都是(X,Y)的联合密度，而且两函数都是连续的，则 $p(x,y)\equiv q(x,y)$.

平面上是连续的，或者除个别几条线外是连续的.

(3) 由(1.6)式容易看出

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

另外，由(1.6)式又看出，二维随机向量(X,Y)落在平面上任一区域D的概率，就等于联合密度$p(x,y)$在D上的积分.这就把概率的计算转化为一个二重积分的计算.由此，顺便指出$\{(X,Y)\in D\}$的概率，数值上就等于以曲面$z=p(x,y)$为顶、以平面区域D为底的曲顶柱体的体积.这就给出了$p(x,y)$的几何意义.

例 1.6 设(X,Y)的联合密度为

$$p(x,y)=\begin{cases}Ce^{-(x+y)} & \text{当 } x\geqslant 0, y\geqslant 0\\ 0 & \text{其他}\end{cases}$$

(1) 求常数C,(2)求$P\{0<X<1,0<Y<1\}$.

解 (1) 因为$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} p(x,y)\mathrm{d}x\mathrm{d}y=1$,故

$$1=\int_0^{+\infty}\int_0^{+\infty} Ce^{-(x+y)}\mathrm{d}x\mathrm{d}y = C\cdot\int_0^{+\infty} e^{-x}\mathrm{d}x\int_0^{+\infty} e^{-y}\mathrm{d}y$$

于是$C=1$.

(2) 记$D=\{(x,y):0<x<1,0<y<1\}$

则由(1.5)式有

$$\begin{aligned}&P\{0<X<1,0<Y<1\}=P\{(X,Y)\in D\}\\&=\iint_D p(x,y)\mathrm{d}x\mathrm{d}y=\iint_D e^{-x-y}\mathrm{d}x\mathrm{d}y\\&=\int_0^1 e^{-x}\mathrm{d}x\int_0^1 e^{-y}\mathrm{d}y=\left(1-\frac{1}{e}\right)^2\end{aligned}$$

与离散型随机向量相仿，现在来介绍连续型随机向量的边缘分布.

定义 1.3 对于随机向量(X,Y)，作为其分量的随机变量X(或Y)的密度函数$p_X(x)$(或$p_Y(y)$)，称为(X,Y)的关于

X(或 Y)的边缘分布密度.

当(X,Y)的联合密度 $p(x,y)$已知时,由下面定理中的(1.7)式,容易求得 X,Y 的边缘密度$p_X(x),p_Y(y)$.

定理 1.1 若(X,Y)的联合密度是 $p(x,y)$,则

$$\begin{aligned} p_1(x) &= \int_{-\infty}^{+\infty} p(x,y)\mathrm{d}y \\ p_2(y) &= \int_{-\infty}^{+\infty} p(x,y)\mathrm{d}x \end{aligned} \tag{1.7}$$

分别是 X,Y 的分布密度.

证 由于$\{-\infty<Y<+\infty\}$是必然事件,故

$$P\{a<X<b\}=P\{a<X<b,-\infty<Y<+\infty\}$$

令 $D=\{(x,y):a<x<b,-\infty<y<+\infty\}$(带形区域),利用(1.6)式得

$$\begin{aligned} P\{a<X<b\} &= \iint_D p(x,y)\mathrm{d}x\mathrm{d}y \\ &= \int_a^b \left[\int_{-\infty}^{+\infty} p(x,y)\mathrm{d}y\right]\mathrm{d}x \end{aligned}$$

再根据随机变量的分布密度的定义(见第二章(3.1)式),不难看出 $p_1(x)=\int_{-\infty}^{+\infty} p(x,y)\mathrm{d}y$ 就是X 的分布密度.同理,

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

就是 Y 的分布密度.证完.

定义 1.4 设 G 是平面上面积为$a(0<a<+\infty)$的区域,称(X,Y)服从 G 上的均匀分布,若 $P\{(X,Y)\in G\}=1$,而且(X,Y)取值属于 G 之任何部分A(A 是G 的子区域)的概率与 A 之面积成正比.

此时,容易推知(X,Y)的联合密度可取为:

$$p(x,y)=\begin{cases} \dfrac{1}{a} & \text{当}(x,y)\in G \\ 0 & \text{其他} \end{cases}$$

例 1.7 设(X,Y)服从如图 4.1 区域 G(抛物线 $y=x^2$和直线 $y=x$ 所夹的区域)上的均匀分布,求联合密度与边缘密度.

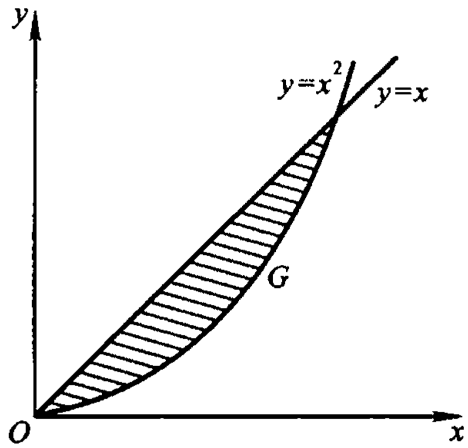

图 4.1

解 由于 G 的面积为

$$\int_0^1 (x-x^2)\mathrm{d}x=\frac{1}{6}$$

故联合密度

$$p(x,y)=\begin{cases}6 & (x,y)\in G\\ 0 & \text{其他}\end{cases}$$

再由(1.7)可得(X,Y)的边缘密度

$$p_X(x)=\int_{-\infty}^{+\infty} p(x,y)\mathrm{d}y=\int_{x^2}^{x} 6\mathrm{d}y=6(x-x^2)$$

$$(0\leqslant x\leqslant 1)$$

$$p_Y(y)=\int_{-\infty}^{+\infty} p(x,y)\mathrm{d}x=\int_{y}^{\sqrt{y}} 6\mathrm{d}x=6(\sqrt{y}-y)$$

$$(0\leqslant y\leqslant 1)$$

当然,$x\overline{\in}[0,1]$ 时 $p_X(x)=0$,$y\overline{\in}[0,1]$时 $p_Y(y)=0$.

4. 随机变量的独立性

随机变量的独立性是概率统计中的一个重要概念.

我们在研究随机现象时,经常碰到这样的一些随机变量:其中

一些的取值对其余随机变量没有什么影响.例如,两个人分别向一目标射击,各自命中的环数 X,Y 就属于这种情形.为了描述这类情况,引进下列定义.

定义 1.5 设 X,Y 是两个随机变量,如果对任意 $a<b,c<d$,事件 $\{a<X<b\}$ 与 $\{c<Y<d\}$ 相互独立,则称 X 与 Y 是相互独立的."相互独立"简称"独立".

定理 1.2 设 X,Y 分别有分布密度 $p_X(x),p_Y(y)$,则 X 与 Y 相互独立的充要条件是:二元函数

$$p_X(x)p_Y(y) \tag{1.8}$$

是随机向量 (X,Y) 的联合密度.

证 先证充分性.设 $p_X(x)p_Y(y)$ 是 (X,Y) 的联合密度,则

$$\begin{aligned}
&P\{\{a<X<b\}\cdot\{c<Y<d\}\}\\
&=P\{a<X<b,c<Y<d\}\\
&=\iint\limits_{\substack{a<x<b\\ c<y<d}} p_X(x)p_Y(y)\mathrm{d}x\mathrm{d}y\\
&=\int_a^b p_X(x)\mathrm{d}x\cdot\int_c^d p_Y(y)\mathrm{d}y\\
&=P\{a<X<b\}\cdot P\{c<Y<d\}
\end{aligned}$$

可见 X 与 Y 是相互独立的.

再证必要性.设 X 与 Y 相互独立,令 $D=\{(x,y):a<x<b,c<y<d\}$,则

$$\begin{aligned}
&P\{(X,Y)\in D\}=P\{a<X<b,c<Y<d\}\\
&=P\{a<X<b\}\cdot P\{c<Y<d\}\\
&=\int_a^b p_X(x)\mathrm{d}x\cdot\int_c^d p_Y(y)\mathrm{d}y=\iint\limits_D[p_X(x)p_Y(y)]\mathrm{d}x\mathrm{d}y
\end{aligned}$$

根据(1.5)知函数 $p_X(x)p_Y(y)$ 是 (X,Y) 的联合密度.定理全部证完.

对于离散型随机变量有下列结论:

定理 1.3 设 X 可能取的值是 $x_1,x_2,x_3,\cdots$(有限个或可列

个)，Y 可能取的值是 $y_1, y_2, y_3, \cdots$(有限个或可列个)，则 X 与 Y 相互独立的充分必要条件是：对一切 i,j 成立

$$P\{X=x_i, Y=y_j\}=P\{X=x_i\}P\{Y=y_j\} \qquad (1.9)$$

证 只证充分性．任给定 $a<b, c<d$．令

$$A=\{x_i: a<x_i<b\}, B=\{y_j: c<y_j<d\}$$

从(1.9)知

$$\begin{aligned}
&\sum_{x_i\in A}\sum_{y_j\in B}P\{X=x_i, Y=y_j\}\\
&=\sum_{x_i\in A}\sum_{y_j\in B}P\{X=x_i\}P\{Y=y_j\}\\
&=\left(\sum_{x_i\in A}P\{X=x_i\}\right)\left(\sum_{y_j\in B}P\{Y=y_j\}\right)\\
&=P\{a<X<b\}\cdot P\{c<Y<d\}
\end{aligned}$$

但

$$\begin{aligned}
&\sum_{x_i\in A}\sum_{y_j\in B}P\{X=x_i, Y=y_j\}\\
&=P\{X\text{ 取值属于}(a,b), Y\text{ 取值属于}(c,d)\}\\
&=P\{a<X<b, c<Y<d\}
\end{aligned}$$

这就证明了 X 与 Y 相互独立．

必要性的证明稍微复杂一些，从略．请读者自行完成(提示：利用第一章(4.23)式)．

定理 1.2 是概率统计中的一条重要的定理．我们在前面已讨论过联合密度与边缘密度的关系：联合密度决定了边缘密度．反之，边缘密度能否决定联合密度呢？一般来讲是不能的．然而，这个定理告诉我们，当 X, Y 独立时，两个边缘密度 $p_X(x)$ 和 $p_Y(y)$ 的乘积就是它们的联合密度．就是说，当 X, Y 独立时，边缘密度也能确定联合密度．下面看一个例子．

例 1.8 设 $X_1\sim N(\mu_1,\sigma_1^2)$，$X_2\sim N(\mu_2,\sigma_2^2)$，且 X_1 与 X_2 相互独立，求 (X_1, X_2) 的联合密度．

解 按已知条件，X_1, X_2 的分布密度分别是：

$$p_{X_1}(x_1)=\frac{1}{\sqrt{2\pi}\sigma_1}e^{-\frac{1}{2\sigma_1^2}(x_1-\mu_1)^2}$$

$$p_{X_2}(x_2)=\frac{1}{\sqrt{2\pi}\sigma_2}e^{-\frac{1}{2\sigma_2^2}(x_2-\mu_2)^2}$$

据定理 1.2 知函数 $p_{X_1}(x_1)\cdot p_{X_2}(x_2)$，即

$$\frac{1}{2\pi\sigma_1\sigma_2}e^{-\frac{1}{2}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2+\left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right]}$$

为(X_1,X_2)的联合密度.

5．二维正态分布

最常见最重要的二维随机向量是二维正态随机向量.

定义 1.6 称 $\xi=(X,Y)$服从二维正态分布，如果它的密度函数是这样的：

$$p(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}e^{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2-\frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\left(\frac{y-\mu_2}{\sigma_2}\right)^2\right]} \tag{1.10}$$

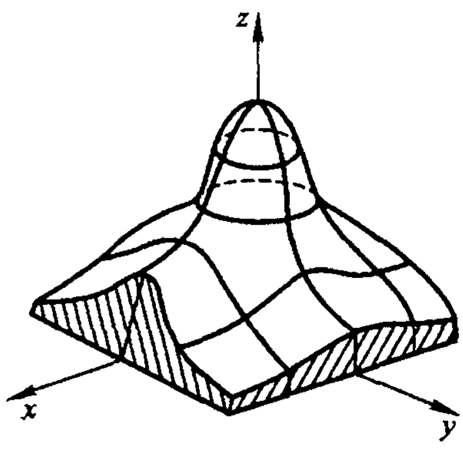

图 4.2

其中 $\mu_1,\mu_2,\sigma_1>0,\sigma_2>0,|\rho|<1$是 5 个参数. $p(x,y)$称为二维正态密度.服从二维正态分布的随机向量叫做二维正态随机向量.

(1) 二维正态分布的边缘密度

按(1.7)，

$$
\begin{aligned}
p_X(x) &= \int_{-\infty}^{+\infty} p(x,y)\mathrm{d}y \\
&= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \mathrm{e}^{-\frac{1}{2(1-\rho^2)}\left(\frac{x-\mu_1}{\sigma_1}\right)^2} \\
&\quad \cdot \int_{-\infty}^{+\infty} \mathrm{e}^{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{y-\mu_2}{\sigma_2}\right)^2 - 2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}\right]}\mathrm{d}y \\
&\xlongequal{\text{令 } t=\frac{y-\mu_2}{\sigma_2}} \frac{1}{2\pi\sigma_1\sqrt{1-\rho^2}} \mathrm{e}^{-\frac{1}{2(1-\rho^2)}\left(\frac{x-\mu_1}{\sigma_1}\right)^2} \\
&\quad \cdot \int_{-\infty}^{+\infty} \mathrm{e}^{-\frac{1}{2(1-\rho^2)}\left[t^2 - 2\rho\frac{x-\mu_1}{\sigma_1}t\right]}\mathrm{d}t \\
&= \frac{1}{2\pi\sigma_1\sqrt{1-\rho^2}} \mathrm{e}^{-\frac{1}{2(1-\rho^2)}\left(\frac{x-\mu_1}{\sigma_1}\right)^2} \\
&\quad \cdot \int_{-\infty}^{+\infty} \mathrm{e}^{-\frac{1}{2(1-\rho^2)}\left[\left(t-\rho\frac{x-\mu_1}{\sigma_1}\right)^2 - \rho^2\left(\frac{x-\mu_1}{\sigma_1}\right)^2\right]}\mathrm{d}t \\
&= \frac{1}{\sqrt{2\pi}\sigma_1} \mathrm{e}^{-\frac{1}{2(1-\rho^2)}\left(\frac{x-\mu_1}{\sigma_1}\right)^2} + \mathrm{e}^{\frac{\rho^2}{2(1-\rho^2)}\left(\frac{x-\mu_1}{\sigma_1}\right)^2} \\
&\quad \cdot \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}\sqrt{1-\rho^2}} \mathrm{e}^{-\frac{1}{2(1-\rho^2)}\left(t-\rho\frac{x-\mu_1}{\sigma_1}\right)^2}\mathrm{d}t
\end{aligned}
$$

注意到积分号内的被积函数是均值为$\dfrac{\rho(x-\mu_1)}{\sigma_1}$,方差为 $1-\rho^2$ 的正态分布的密度函数,所以积分值等于 1. 于是,

$$
p_X(x) = \frac{1}{\sqrt{2\pi}\sigma_1} \mathrm{e}^{-\frac{(1-\rho^2)}{2(1-\rho^2)}\left(\frac{x-\mu_1}{\sigma_1}\right)^2} = \frac{1}{\sqrt{2\pi}\sigma_1} \mathrm{e}^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}
$$

同样可求得

$$
p_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_2} \mathrm{e}^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}}
$$

上面的结果表明,二维正态的边缘密度是(一维)正态密度;而

且还顺便告诉我们：二维正态 5 个参数中的 μ_1,μ_2 和 σ_1^2,σ_2^2 分别是两个边缘密度的均值和方差. 除此之外，还可说明由(1.10)式所给出的 $p(x,y)$ 确是一个二维联合密度函数. 这是因为

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} p(x,y)\mathrm{d}x\mathrm{d}y = \int_{-\infty}^{+\infty}\left[\int_{-\infty}^{+\infty} p(x,y)\mathrm{d}y\right]\mathrm{d}x$$

$$= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma_1}\mathrm{e}^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}\mathrm{d}x = 1$$

(2) 分量独立的充要条件

我们指出下列重要事实：若 (X,Y) 服从二维正态分布(参数为 $\mu_1,\mu_2,\sigma_1,\sigma_2,\rho$)，则 X 与 Y 相互独立的充分必要条件是 $\rho=0$.

先证充分性. 设 (X,Y) 的密度是

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

其中 $\mu_1,\mu_2,\sigma_1>0,\sigma_2>0,|\rho|<1$ 是 5 个参数.

此时 X,Y 的密度分别是

$$p_X(x)=\frac{1}{\sqrt{2\pi}\sigma_1}\mathrm{e}^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}$$

$$p_Y(y)=\frac{1}{\sqrt{2\pi}\sigma_2}\mathrm{e}^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}}$$

当 $\rho=0$ 时，显然有 $p(x,y)=p_X(x)p_Y(y)$. 由定理 1.2 知 X 与 Y 是相互独立的.

下面证明必要性. 设 X,Y 相互独立，由定理 1.2 知

$$p_X(x)p_Y(y)=\frac{1}{2\pi\sigma_1\sigma_2}\mathrm{e}^{-\frac{1}{2}\left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2+\left(\frac{y-\mu_2}{\sigma_2}\right)^2\right]} \tag{1.11}$$

为 (X,Y) 的联合密度.

于是(1.10)与(1.11)都是 (X,Y) 的联合密度，由于这两个密度函数都是连续的，它们应该处处相等. 特别应该有 $p(\mu_1,\mu_2)=$

$p_X(\mu_1)p_Y(\mu_2)$. 从(1.10)知 $p(\mu_1,\mu_2)=\dfrac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}$，从(1.11)知 $p_X(\mu_1)p_Y(\mu_2)=\dfrac{1}{2\pi\sigma_1\sigma_2}$，于是 $\sqrt{1-\rho^2}=1$，从而 $\rho=0$，必要性得证.

作为本节的结束，我们介绍一下二维随机向量的分布函数的概念.

定义 1.7 设 $\boldsymbol{\xi}=(X,Y)$ 是二维随机向量，称函数

$$F(x,y)=P\{X\leqslant x,Y\leqslant y\}$$

为它的分布函数.

若 $\boldsymbol{\xi}=(X,Y)$ 的分布函数有二阶连续偏微商，则 $\dfrac{\partial^2 F(x,y)}{\partial x\partial y}$ 就是 $\boldsymbol{\xi}$ 的分布密度.

习题十一

1. 离散型随机向量(X,Y)有如下的概率分布：

X \ Y	0	1	2	3	4	5	6
0	0.202	0.174	0.113	0.062	0.049	0.023	0.004
1	0	0.099	0.064	0.040	0.031	0.020	0.006
2	0	0	0.031	0.025	0.018	0.013	0.008
3	0	0	0	0.001	0.002	0.004	0.011

求边缘分布．又：随机变量 X,Y 是否独立？

2. 随机向量(X,Y)在矩形区域 $D=\{(x,y):a<x<b,c<y<d\}$上服从均匀分布，求联合密度与边缘密度.又问随机变量 X,Y 是否独立？

3. 随机向量(X,Y)的联合密度

$$p(x,y)=\begin{cases}c(R-\sqrt{x^2+y^2}) & \text{当 } x^2+y^2<R^2\\ 0 & \text{当 } x^2+y^2\geqslant R^2\end{cases}$$

求：(1) 系数 c；(2) 随机向量落在圆 $x^2+y^2\leqslant r^2(r<R)$内的概率.

4. 设(X,Y)的联合密度是

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

求:(1) 系数 c;(2) (X,Y)落在以(0,0),(0,1),(1,0),(1,1)为顶点的正方形内的概率;(3) 问 X,Y 是否独立?

5. 设(X,Y)的联合密度是

$$p(x,y)=\begin{cases}A\sin(x+y) & \text{当 } 0<x<\frac{\pi}{2}, 0<y<\frac{\pi}{2}\\ 0 & \text{其他}\end{cases}$$

求:(1) 系数 A;(2) 边缘密度.

6. 一机器制造直径为 X 的圆轴,另一机器制造内径为 Y 的轴衬,设(X,Y)的联合密度为

$$p(x,y)=\begin{cases}2\,500 & \text{当 } 0.49<x<0.51, 0.51<y<0.53\\ 0 & \text{其他}\end{cases}$$

若轴衬的内径与轴的直径之差大于 0.004 且小于 0.036,则两者可以相适衬.求任一轴与任一轴衬相适衬的概率.

7. 设(X,Y)服从

$$D=\left\{(x,y):\frac{(x+y)^2}{2a^2}+\frac{(x-y)^2}{2b^2}\leqslant 1\right\}$$

上的均匀分布,求(X,Y)的联合密度 $p(x,y)$.

8. 对于下列三组参数,写出二维正态随机向量的联合密度与边缘密度.

	μ_1	μ_2	σ_1	σ_2	ρ
(1)	3	0	1	1	$\frac{1}{2}$
(2)	1	1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
(3)	1	2	1	$\frac{1}{2}$	0

§2 两个随机变量的函数的分布

在第二章中,我们讨论了一个随机变量的函数的分布问题,即:已知 X 的分布,求 X 的函数 $Y=f(X)$的分布问题.在那里,对于 X 是连续型的情形,我们介绍了一个基本方法,该方法的要点是:通过 Y 与 X 的关系,先找出 Y 的分布函数,再利用分布函数与密度函数

的关系，最终找到 Y 的密度函数.

本节要讨论的是两个随机变量的函数的分布问题. 具体来说：已知(X, Y)的联合密度，求

$$Z = f(X, Y)$$

的密度函数(如果密度函数存在的话).

1. 和的分布

已知(X, Y)的联合密度是 $p(x, y)$，求 $Z = X + Y$ 的密度.

还用"分布函数法"，我们先来找出 $P\{Z \leqslant z\}$. 显然有

$$P\{Z \leqslant z\} = P\{X + Y \leqslant z\}$$

注意到事件$\{X + Y \leqslant z\}$形式上可以换成我们所熟悉的表达形式：

$$\{X + Y \leqslant z\} = \{(X, Y) \in D\}$$

这里的 $D = \{(x, y): x + y \leqslant z\}$，它是 xy 平面上的一个区域(见图 4.3). 由于(X, Y)的联合密度是 $p(x, y)$，所以有

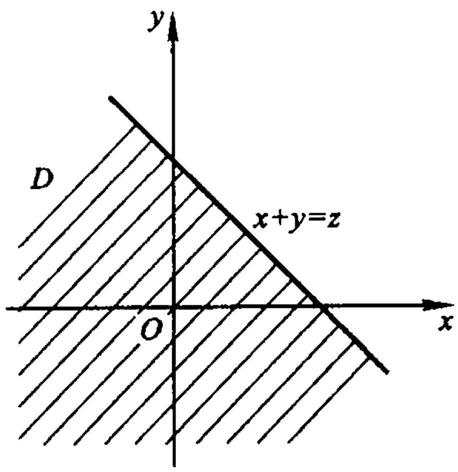

图 4.3

$$\begin{aligned} P\{Z \leqslant z\} &= P\{X + Y \leqslant z\} \\ &= P\{(X, Y) \in D\} \\ &= \iint\limits_{D} p(x, y) \mathrm{d}x \mathrm{d}y \\ &= \iint\limits_{x+y \leqslant z} p(x, y) \mathrm{d}x \mathrm{d}y \end{aligned} \tag{2.1}$$

至此，对于每个确定的 z，$P\{Z \leqslant z\}$的值已归结为计算一个二重积分，而该积分的被积函数与积分区域都是已知的. 利用二重积分与累次积分的关系，我们有

$$\iint\limits_{x+y \leqslant z} p(x, y) \mathrm{d}x \mathrm{d}y = \int_{-\infty}^{+\infty} \mathrm{d}x \int_{-\infty}^{z-x} p(x, y) \mathrm{d}y$$

$$\xlongequal{\text{令 } u=y+x} \int_{-\infty}^{+\infty} \mathrm{d}x \int_{-\infty}^{z} p(x, u-x)\mathrm{d}u$$

$$= \int_{-\infty}^{z} \mathrm{d}u \int_{-\infty}^{+\infty} p(x, u-x)\mathrm{d}x$$

因此,得

$$F_Z(z) = P\{Z \leqslant z\} = \int_{-\infty}^{z} \left[\int_{-\infty}^{+\infty} p(x, u-x)\mathrm{d}x\right]\mathrm{d}u$$

于是,Z 的分布密度 $p_Z(z)$ 求得如下:

$$p_Z(z) = \int_{-\infty}^{+\infty} p(x, z-x)\mathrm{d}x \tag{2.2}$$

例 2.1 设 X 与 Y 相互独立,服从相同的分布 $N(\mu, \sigma^2)$,求 $X+Y$ 的分布密度.

解 由题设知(X, Y)的联合密度 $p(x, y)$是

$$\frac{1}{2\pi\sigma^2}\mathrm{e}^{-\frac{1}{2\sigma^2}[(x-\mu)^2+(y-\mu)^2]}$$

于是按(2.2)式,$X+Y$ 的密度 $p(z)$:

$$p(z) = \int_{-\infty}^{+\infty} \frac{1}{2\pi\sigma^2}\mathrm{e}^{-\frac{1}{2\sigma^2}[(x-\mu)^2+(z-x-\mu)^2]}\mathrm{d}x$$

$$\xlongequal{\text{令 } t=x-\mu} \int_{-\infty}^{+\infty} \frac{1}{2\pi\sigma^2}\mathrm{e}^{-\frac{1}{2\sigma^2}[t^2+(z-2\mu-t)^2]}\mathrm{d}t$$

$$= \int_{-\infty}^{+\infty} \frac{1}{2\pi\sigma^2}\mathrm{e}^{-\frac{1}{2\sigma^2}[2t^2-2(z-2\mu)t+(z-2\mu)^2]}\mathrm{d}t$$

$$= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}\left(\frac{\sigma}{\sqrt{2}}\right)}\mathrm{e}^{-\frac{1}{2\sigma^2}2\left(t-\frac{z-2\mu}{2}\right)^2}$$

$$\cdot \frac{1}{\sqrt{2\pi}(\sqrt{2}\sigma)}\mathrm{e}^{-\frac{1}{2\sigma^2}\frac{(z-2\mu)^2}{2}}\mathrm{d}t$$

$$= \frac{1}{\sqrt{2\pi}(\sqrt{2}\sigma)}\mathrm{e}^{-\frac{(z-2\mu)^2}{2(\sqrt{2}\sigma)^2}}$$

这表明 $X+Y \sim N(2\mu, 2\sigma^2)$,即 $X+Y$ 也服从正态分布,其均值与方差都是原来的两倍.

2. 两个例子

前面推导了二维随机向量(X,Y)之和 $Z=X+Y$ 的分布，得到了公式(2.2)，然后利用公式(2.2)很容易推导出正态独立同分布的随机变量之和的分布. 这些推导过程是有一般性的. 为了说明这种一般性，我们先来解剖一下在公式(2.2)的推导过程中的两个特点：

(1) 为了求和的密度，先求和的分布，即

$$P\{X+Y\leqslant z\}$$

(2) 在求 $P\{X+Y\leqslant z\}$ 的过程中，用到下列等式：

$$P\{X+Y\leqslant z\}=\iint\limits_{x+y\leqslant z}p(x,y)\mathrm{d}x\mathrm{d}y$$

即(2.1)式，其中 $p(x,y)$为(X,Y)的联合密度.

事实上，这(1)，(2)两点很容易推广到以下(1′)，(2′)两点：

(1′) 为求随机变量函数 $f(X,Y)$的密度，先求它的分布，即

$$P\{f(X,Y)\leqslant z\}$$

(2′) 在求 $P\{f(X,Y)\leqslant z\}$的过程中，用到下列等式

$$P\{f(X,Y)\leqslant z\}=\iint\limits_{f(x,y)\leqslant z}p(x,y)\mathrm{d}x\mathrm{d}y \qquad (2.3)$$

其中 $p(x,y)$是(X,Y)的联合密度.

下面利用(1′)，(2′)两点，计算两个重要的例子.

例2.2 设 X,Y 相互独立服从相同分布 $N(0,1)$，求 $\sqrt{X^2+Y^2}$ 的密度.

解 记 $Z=\sqrt{X^2+Y^2}$ 的分布函数为 $F_Z(z)$，则

$$F_Z(z)=P\{Z\leqslant z\}=P\{\sqrt{X^2+Y^2}\leqslant z\}$$

$$=\iint\limits_{\sqrt{x^2+y^2}\leqslant z}\frac{1}{2\pi}\mathrm{e}^{-\frac{1}{2}(x^2+y^2)}\mathrm{d}x\mathrm{d}y$$

作极坐标变换：$x=r\cos\theta, y=r\sin\theta(r\geqslant 0, 0\leqslant\theta<2\pi)$，易知 $z>0$ 时

$$F_Z(z) = \int_0^{2\pi} d\theta \int_0^z \frac{1}{2\pi} e^{-\frac{1}{2}r^2} r dr = \int_0^z r e^{-\frac{1}{2}r^2} dr$$

当 $z \leqslant 0$ 时 $F_Z(z) = 0$. 于是$\sqrt{X^2 + Y^2}$的密度 $p(z)$是

$$p(z) = \begin{cases} z e^{-\frac{z^2}{2}} & z > 0 \\ 0 & z \leqslant 0 \end{cases}$$

这就是所谓瑞利(**Rayleigh**)分布

例 2.3 设 X, Y 独立同分布，共同的密度函数为 $p(\cdot)$，分布函数为 $F(\cdot)$.（对于函数 $h(x)$，如果我们只注意函数关系，而对自变量采用什么记号并不关心时，习惯上就写作 $h(\cdot)$.）求 $Z = \max(X, Y)$的密度函数.（为简单计，我们设 $p(\cdot)$ 是连续函数.）

解 $F_Z(z) = P\{Z \leqslant z\}$

$= P\{\max(X, Y) \leqslant z\}$

注意到事件等式：

$$\{\max(X, Y) \leqslant z\} = \{X \leqslant z\} \cdot \{Y \leqslant z\}$$

（左边的事件发生意味着 X, Y 中的较大者$\leqslant z$，因而有 $X \leqslant z$ 及 $Y \leqslant z$，即$\{X \leqslant z\}$发生以及$\{Y \leqslant z\}$发生，亦即$\{X \leqslant z\} \cdot \{Y \leqslant z\}$发生；这表明左边$\subset$右边. 类似可证右边$\subset$左边. 于是得上面的等式.）

再由 X, Y 的独立性得：

$$\begin{aligned} P\{\max(X, Y) \leqslant z\} &= P\{X \leqslant z, Y \leqslant z\} \\ &= P\{X \leqslant z\} \cdot P\{Y \leqslant z\} \\ &= F(z) \cdot F(z) = F^2(z) \end{aligned}$$

于是，$\max(X, Y)$的密度 $p_Z(z)$为

$$p_Z(z) = [F^2(z)]' = 2F(z) \cdot F'(z) = 2F(z) \cdot p(z)$$

随机变量函数的联合密度

已知(X, Y)的联合密度，而

$$\begin{cases} U = f(X, Y) \\ V = g(X, Y) \end{cases}$$

如何求出(U,V)的联合密度？我们可以证明下列重要定理，由于严格的数学论述较为抽象，初学者可以略去不读.

定理 2.1 设(X,Y)有联合密度 $p(x,y)$，且区域 A(可以是全平面)满足：$P\{(X,Y)\in A\}=1$.

又函数 $f(x,y)$，$g(x,y)$满足：

① 对任何实数 u,v，方程组

$$\begin{cases} f(x,y)=u \\ g(x,y)=v \end{cases} \tag{2.4}$$

在 A 中至多有一个解 $x=x(u,v)$，$y=y(u,v)$.

② f,g 在A 中有连续偏导数.

③ 雅可比行列式$\dfrac{\partial(f,g)}{\partial(x,y)}$在 A 中处处不等于 0.

设 $U=f(X,Y)$，$V=g(X,Y)$.

$G=\{(u,v)$：方程组(2.4)在 A 中有解$\}$

$$q(u,v)=\begin{cases} p[x(u,v),y(u,v)]\left|\dfrac{\partial(x,y)}{\partial(u,v)}\right| & \text{当}(u,v)\in G \\ 0 & \text{当}(u,v)\bar{\in} G \end{cases}$$

这里$\left|\dfrac{\partial(x,y)}{\partial(u,v)}\right|$是函数 $x(u,v)$，$y(u,v)$的雅可比行列式的绝对值. 则 $q(u,v)$是(U,V)的联合密度.

证 给定 $a<b$，$c<d$. 设 $D=\{(u,v):a<u<b,c<v<d\}$，$D^*=\{(x,y):(f(x,y),g(x,y))\in D\}$，易知$(f(x,y),g(x,y))$是$D^*\cap A$到$D\cap G$上之一一映射，其逆映射是$(x(u,v),y(u,v))$. 据重积分的变数替换公式知

$$\iint\limits_{D^*\cap A} p(x,y)\mathrm{d}x\mathrm{d}y=\iint\limits_{D\cap G} p[x(u,v),y(u,v)]\left|\frac{\partial(x,y)}{\partial(u,v)}\right|\mathrm{d}u\mathrm{d}v$$

于是

$$P\{(U,V)\in D\}=P\{(f(X,Y),g(X,Y))\in D\}$$

$$=P\{(X,Y)\in D^*\}=P\{(X,Y)\in D^*\cap A\}=\iint\limits_{D^*\cap A} p(x,y)\mathrm{d}x\mathrm{d}y$$

$$=\iint\limits_{D\cap G} p[x(u,v),y(u,v)]\left|\frac{\partial(x,y)}{\partial(u,v)}\right|\mathrm{d}u\mathrm{d}v=\iint\limits_{D} q(u,v)\mathrm{d}u\mathrm{d}v$$

这就证明了 $q(u,v)$是(U,V)的联合密度.

例 2.4 设 X,Y 相互独立，都服从$[0,1]$上的均匀分布，

$$U=\sqrt{-2\ln X}\cos 2\pi Y$$

$$V=\sqrt{-2\ln X}\sin 2\pi Y$$

求(U,V)的联合密度.

解 在定理2.1中令$f(x,y)=\sqrt{-2\ln x}\cos 2\pi y, g(x,y)=\sqrt{-2\ln x}\cdot\sin 2\pi y$, $A=\left\{(x,y):0<x<1,0<y<1,\text{但 } y\neq\frac{1}{4},\frac{2}{4},\frac{3}{4}\right\}$,易知那里的$G=\{(u,v):u\neq 0,v\neq 0\}$, $x=x(u,v)=e^{-\frac{1}{2}(u^2+v^2)}$, $y=y(u,v)$的表达式较复杂.当$u>0,v>0$时$y(u,v)=\frac{1}{2\pi}\arctan\frac{v}{u}$;当$u>0,v<0$时,$y(u,v)=1+\frac{1}{2\pi}\arctan\frac{v}{u}$;当$u<0$时,$y(u,v)=\frac{1}{2}+\frac{1}{2\pi}\arctan\frac{v}{u}$.不难推知

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

但(X,Y)的联合密度为

$$p(x,y)=\begin{cases}1 & 0\leqslant x\leqslant 1,0\leqslant y\leqslant 1\\ 0 & \text{其他}\end{cases}$$

故(U,V)的联合密度函数为

$$q(u,v)=\begin{cases}\frac{1}{2\pi}e^{-\frac{1}{2}(u^2+v^2)} & \text{当 } u\neq 0, v\neq 0 \text{ 时}\\ 0 & \text{当 } u=0 \text{ 或 } v=0 \text{ 时}\end{cases}$$

当然,函数$\varphi(u,v)\equiv\frac{1}{2\pi}e^{-\frac{1}{2}(u^2+v^2)}$也是$(U,V)$的联合密度函数.由于$\varphi(u,v)=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}u^2}\cdot\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}v^2}$,不难看出,$U,V$是相互独立的而且都服从$N(0,1)$.

例2.5 设X,Y相互独立,都服从$N(0,1)$,

$$\begin{aligned}X&=R\cos\Theta\\ Y&=R\sin\Theta\end{aligned}\quad\begin{pmatrix}R\geqslant 0\\ 0\leqslant\Theta<2\pi\end{pmatrix}$$

求(R,Θ)的联合密度与边缘密度.

解 利用定理2.1.(取$A=\{(x,y):x\neq 0,y\neq 0\}$并改变密度函数$q(\cdot,\cdot)$在个别点上的值)可以求得$(R,\Theta)$的联合密度为:

$$\varphi(r,\theta)=\begin{cases}\frac{1}{2\pi}re^{-\frac{1}{2}r^2} & r>0,0<\theta<2\pi\\ 0 & \text{其他}\end{cases}$$

如令

$$f(r)=\begin{cases} re^{-\frac{1}{2}r^2} & r>0 \\ 0 & \text{其他} \end{cases}$$

$$g(\theta)=\begin{cases} \dfrac{1}{2\pi} & 0<\theta<2\pi \\ 0 & \text{其他} \end{cases}$$

则有 $\varphi(r,\theta)=f(r)g(\theta)$. 由此不难看出:$R$ 的密度是 $f(r)$, Θ 的密度是 $g(\theta)$, 且 R 与 Θ 相互独立.

***例 2.6** 设 R 与 Θ 相互独立,都服从(0,1)上的均匀分布, r_1 和 r_2 是固定的数($0\leqslant r_1<r_2$),

$$X=r_2\sqrt{\frac{r_1^2}{r_2^2}+\left(1-\frac{r_1^2}{r_2^2}\right)R}\cos(2\pi\Theta)$$

$$Y=r_2\sqrt{\frac{r_1^2}{r_2^2}+\left(1-\frac{r_1^2}{r_2^2}\right)R}\sin(2\pi\Theta)$$

则 (X,Y) 服从环 $\{(x,y):r_1^2\leqslant x^2+y^2\leqslant r_2^2\}$ 上的均匀分布.

实际上,令

$$f(r,\theta)=r_2\sqrt{\frac{r_1^2}{r_2^2}+\left(1-\frac{r_1^2}{r_2^2}\right)r}\cos 2\pi\theta$$

$$g(r,\theta)=r_2\sqrt{\frac{r_1^2}{r_2^2}+\left(1-\frac{r_1^2}{r_2^2}\right)r}\sin 2\pi\theta$$

则
$$X=f(R,\Theta),\ Y=g(R,\Theta)$$

记 $x=f(r,\theta)$, $y=g(r,\theta)$, 则

$$r=\frac{x^2+y^2-r_1^2}{r_2^2-r_1^2}\xlongequal{\text{记}}r(x,y)$$

$$\theta=\begin{cases} \dfrac{1}{2\pi}\arctan\dfrac{y}{x}, & x>0,y>0 \\ \dfrac{1}{2\pi}\left(\arctan\dfrac{y}{x}+\pi\right), & x<0 \\ \dfrac{1}{2\pi}\left(\arctan\dfrac{y}{x}+2\pi\right), & x>0,y<0 \end{cases}$$

$$\xlongequal{\text{记}}\theta(x,y)$$

易知雅可比行列式 $J=\dfrac{\partial(r,\theta)}{\partial(x,y)}=\dfrac{1}{\pi(r_2^2-r_1^2)}$, (R,Θ) 的联合密度 $p(r,\theta)=$

$I_{\{0,1\}}(r)I_{\{0,1\}}(\theta)$,这里

$$I_F(x)=\begin{cases}1 & x\in E\\ 0 & x\overline{\in} E\end{cases}$$

设 D 是环 $\{(x,y):r_1^2\leqslant x^2+y^2\leqslant r_2^2\}$ 中任一区域,则

$$\begin{aligned}P((X,Y)\in D)&=\iint\limits_{\{(f(r,\theta),g(r,\theta))\in D\}} p(r,\theta)\mathrm{d}r\mathrm{d}\theta\\ &=\iint\limits_{D} I_{[0,1]}(r(x,y))I_{\{0,1\}}(\theta(x,y))|J|\mathrm{d}x\mathrm{d}y\\ &=\iint\limits_{D}|J|\mathrm{d}x\mathrm{d}y=\frac{m(D)}{\pi(r_2^2-r_1^2)}\end{aligned}$$

这里 $m(D)$是 D 的面积.可见(X,Y)服从环上的均匀分布.

习 题 十 二

1. 设 X,Y 相互独立,其密度分别为

$$p_X(x)=\begin{cases}1 & 0\leqslant x\leqslant 1\\ 0 & \text{其他}\end{cases}$$

$$p_Y(y)=\begin{cases}\mathrm{e}^{-y} & y>0\\ 0 & y\leqslant 0\end{cases}$$

求 $X+Y$ 的密度

2. 设 X,Y 相互独立,分别服从自由度为 k_1,k_2的 χ^2分布,即

$$p_X(x)=\begin{cases}\dfrac{1}{2^{\frac{k_1}{2}}\Gamma\left(\dfrac{k_1}{2}\right)}x^{\frac{k_1}{2}-1}\mathrm{e}^{-\frac{x}{2}} & x>0\\ 0 & x\leqslant 0\end{cases}$$

$$p_Y(y)=\begin{cases}\dfrac{1}{2^{\frac{k_2}{2}}\Gamma\left(\dfrac{k_2}{2}\right)}y^{\frac{k_2}{2}-1}\mathrm{e}^{-\frac{y}{2}} & y>0\\ 0 & y\leqslant 0\end{cases}$$

证明 $X+Y$ 也服从χ^2分布,自由度为 k_1+k_2.

3. 设 $X_1\sim N(\mu_1,\sigma_1^2)$,$X_2\sim N(\mu_2,\sigma_2^2)$,且 X_1,X_2相互独立,求证:

(1) $X_1+X_2\sim N(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2)$;

(2) $\dfrac{X_1+X_2}{2}\sim N\left(\dfrac{\mu_1+\mu_2}{2},\dfrac{\sigma_1^2+\sigma_2^2}{4}\right)$.

4. 设 X,Y 独立同分布,密度为 $p(\cdot)$,分布函数为 $F(\cdot)$.求 $\min\{X,Y\}$ 的密度.

5. 设系统 L 由两个相互独立的子系统 L_1,L_2联接而成,联接的方式分别为(1)串联,(2)并联,(3)备用(当系统 L_1损坏时,系统 L_2开始工作),如图 4.4 所示.已知 L_1,L_2的寿命分别为 X 和 Y,概率密度分别为

$$p_X(x)=\begin{cases}\alpha e^{-\alpha x} & x>0\\ 0 & x\leqslant 0\end{cases}$$

$$p_Y(y)=\begin{cases}\beta e^{-\beta y} & y>0\\ 0 & y\leqslant 0\end{cases}$$

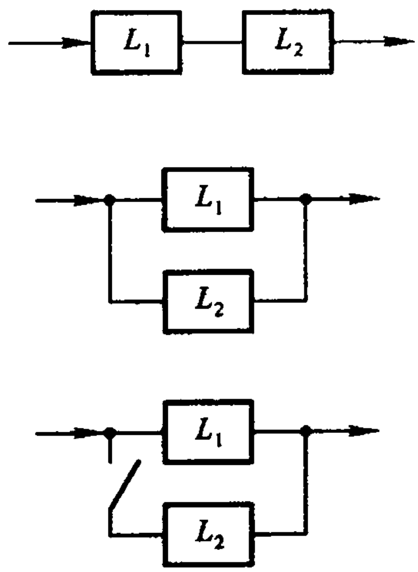

图 4.4

其中 $\alpha>0,\beta>0$ 且 $\alpha\neq\beta$,试分别就这三种联接方式写出系统 L 的寿命 Z 的概率密度.

6. 设某种商品一周的需要量是一个随机变量,其密度为

$$p(x)=\begin{cases}xe^{-x} & x>0\\ 0 & x\leqslant 0\end{cases}$$

如果各周的需要量是互相独立的.试求(1) 两周,(2) 三周的需要量的概率密度.

§3 随机向量的数字特征

本节讨论二维随机向量的数字特征,并对均值与方差的性质

作进一步的讨论.作为基本工具,先介绍均值公式.

1. 两个随机变量的函数的均值公式

对于两个随机变量的函数 $Z=f(X,Y)$ 的均值,也有均值公式:

设 (X,Y) 的联合密度为 $p(x,y)$,$Z=f(X,Y)$,则有①

$$E(Z)=E[f(X,Y)]=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f(x,y)p(x,y)\mathrm{d}x\mathrm{d}y \tag{3.1}$$

由于(3.1)式的证明中要用到较多的数学知识,我们就不证明了.

例 3.1 设 X,Y 独立同分布,共同分布是 $N(0,1)$,求

$$E[\sqrt{X^2+Y^2}]$$

解法 1 用公式(3.1).

$$\begin{aligned}E[\sqrt{X^2+Y^2}]&=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\sqrt{x^2+y^2}\frac{1}{2\pi}\cdot\mathrm{e}^{-\frac{1}{2}(x^2+y^2)}\mathrm{d}x\mathrm{d}y\\&=\int_0^{2\pi}\mathrm{d}\theta\int_0^{+\infty}r\cdot\frac{1}{2\pi}\mathrm{e}^{-\frac{1}{2}r^2}r\mathrm{d}r\text{(作极坐标变换)}\\&=\int_0^{+\infty}r^2\mathrm{e}^{-\frac{1}{2}r^2}\mathrm{d}r=\frac{\sqrt{2\pi}}{2}\end{aligned}$$

解法 2 由§2,此时 $\sqrt{X^2+Y^2}$ 服从瑞利分布.即 $\sqrt{X^2+Y^2}$ 的密度是

$$p(z)=\begin{cases}z\mathrm{e}^{-\frac{1}{2}z^2} & z>0\\0 & z\leqslant 0\end{cases}$$

所以

$$E(\sqrt{X^2+Y^2})=\int_{-\infty}^{+\infty}zp(z)\mathrm{d}z=\int_0^{+\infty}z^2\mathrm{e}^{-\frac{1}{2}z^2}\mathrm{d}z=\frac{\sqrt{2\pi}}{2}$$

---

① 严格地说,还要求该二重积分绝对收敛.(X,Y) 为离散型时,也有与(3.1)相应的均值公式.

比较这两种解法可知，利用了公式(3.1)的解法 1 要方便些，因为它不需要知道$\sqrt{X^2+Y^2}$的分布，而只需直接计算一个二重积分即可.

另外，对于二维随机向量(X,Y)，计算两个分量 X，Y 之和以及它们本身的均值和方差，都可以利用(3.1)式(将在下面 2 中介绍)，由此可见(3.1)式的重要性.

2. 均值与方差的性质

(1) 设(X,Y)的联合密度为 $p(x,y)$，X，Y 的边缘密度分别为 $p_X(x)$，$p_Y(y)$，则由上一章的讨论，知道有

$$E(X)=\int_{-\infty}^{+\infty} xp_X(x)\mathrm{d}x,\ E(Y)=\int_{-\infty}^{+\infty} yp_Y(y)\mathrm{d}y$$

$$D(X)=\int_{-\infty}^{+\infty}[x-E(X)]^2 p_X(x)\mathrm{d}x$$

$$D(Y)=\int_{-\infty}^{+\infty}[y-E(Y)]^2 p_Y(y)\mathrm{d}y$$

这里，我们再列出另一套由联合密度 $p(x,y)$给出的计算公式①.

$$\begin{aligned}
E(X)&=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xp(x,y)\mathrm{d}x\mathrm{d}y\\
E(Y)&=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} yp(x,y)\mathrm{d}x\mathrm{d}y\\
D(X)&=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}[x-E(X)]^2 p(x,y)\mathrm{d}x\mathrm{d}y\\
D(Y)&=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}[y-E(Y)]^2 p(x,y)\mathrm{d}x\mathrm{d}y
\end{aligned} \tag{3.2}$$

由公式(3.1)立即可推得(3.2)各式成立，请读者自行完成.

(2) 我们有

$$E(X+Y)=E(X)+E(Y) \tag{3.3}$$

① 当 X，Y 是离散型时，也有与(3.2)各式相应的结果.

$$D(X+Y)=D(X)+D(Y)$$
$$+2E\{[X-E(X)][(Y-E(Y)]\} \tag{3.4}$$

当 X,Y 独立时,有

$$E(X\cdot Y)=E(X)\cdot E(Y) \tag{3.5}$$
$$D(X+Y)=D(X)+D(Y) \tag{3.6}$$

证 先证(3.3)式.由(3.1)式[①],

$$\begin{aligned}E(X+Y)&=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x+y)p(x,y)\mathrm{d}x\mathrm{d}y\\&=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xp(x,y)\mathrm{d}x\mathrm{d}y+\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}yp(x,y)\mathrm{d}x\mathrm{d}y\\&=E(X)+E(Y)\end{aligned}$$

再证(3.4)式.

$$\begin{aligned}D(X+Y)&=E[(X+Y)-E(X+Y)]^2\\&=E\{[X-E(X)]+[Y-E(Y)]\}^2\text{(由(3.3)式)}\\&=E\{[X-E(X)]^2+[Y-E(Y)]^2+\\&\quad 2[X-E(X)][Y-E(Y)]\}\\&=E[X-E(X)]^2+E[Y-E(Y)]^2+\\&\quad 2E\{[X-E(X)][Y-E(Y)]\}\\&=D(X)+D(Y)+2E\{[X-E(X)][Y-E(Y)]\}\end{aligned}$$

(3.4)式得证.下面来证(3.5)式.设 X,Y 的密度分别是 $p_X(x),p_Y(y)$.从 X,Y 独立知,(X,Y)的联合密度为 $p_X(x)\cdot p_Y(y)$,由(3.1)式知

$$\begin{aligned}E(X\cdot Y)&=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xyp_X(x)p_Y(y)\mathrm{d}x\mathrm{d}y\\&=\left[\int_{-\infty}^{+\infty}xp_X(x)\mathrm{d}x\right]\cdot\left[\int_{-\infty}^{+\infty}yp_Y(y)\mathrm{d}y\right]\end{aligned}$$

---

① 我们是从均值公式(3.1)出发推导(3.3),即假定了(X,Y)是连续型的二维随机向量.这个假定显然太苛刻了.用较多的数学知识可以证明,对任意两个随机变量,只要它们的期望存在,则(3.3)仍成立.

$$=E(X)\cdot E(Y)$$

最后来证(3.6).由于

$$E\{[X-E(X)][Y-E(Y)]\}$$
$$=E\{XY-XE(Y)-YE(X)+E(X)\cdot E(Y)\}$$
$$=E(XY)-E(X)\cdot E(Y)-E(Y)\cdot E(X)+E(X)\cdot E(Y) \qquad (\text{由 } 3.3)$$
$$=E(XY)-E(X)\cdot E(Y)=0$$

(最后一个等号用到了 X 与 Y 的独立性及(3.5))于是,从(3.4)立即得到(3.6).

定义 3.1 称向量$(E(X),E(Y))$为随机向量(X,Y)的均值,称数值 $E\{[X-E(X)][Y-E(Y)]\}$为 X,Y 的协方差.

3. 协方差(斜方差)

协方差 $E\{[X-E(X)][Y-E(Y)]\}$是二维随机向量(X,Y)的一个重要的数字特征,它刻画了 X,Y 取值间的相互联系.通常采用的记号为

$$\operatorname{cov}(X,Y)\triangleq E\{[X-E(X)][Y-E(Y)]\}$$

或

$$\sigma_{XY}\triangleq E\{[X-E(X)][Y-E(Y)]\}$$

由公式(3.1)知:

$$\sigma_{XY}=\operatorname{cov}(X,Y)$$
$$=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}[x-E(X)][y-E(Y)]p(x,y)\mathrm{d}x\mathrm{d}y \qquad (3.7)$$

其中 $p(x,y)$为(X,Y)的联合密度.

与记号 σ_{XY}对应,对于 $D(X),D(Y)$也分别可采用记号 σ_{XX},σ_{YY}.

根据前面的讨论,当 X 与 Y 相互独立时,协方差 $\sigma_{XY}=0$(如果协方差存在的话).值得注意的是,由 $\sigma_{XY}=0$ 并不能保证 X 与 Y 独立.

例 3.2 设(X,Y)的联合密度是

$$p(x,y)=\begin{cases}\dfrac{1}{\pi} & x^2+y^2\leqslant 1\\ 0 & \text{其他}\end{cases}$$

求 σ_{XX}，σ_{YY}，σ_{XY}.

解 先求 $E(X)$，$E(Y)$.由公式(3.2)，有：

$$E(X)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xp(x,y)\mathrm{d}x\mathrm{d}y$$

$$=\iint\limits_{x^2+y^2\leqslant 1}x\cdot\frac{1}{\pi}\mathrm{d}x\mathrm{d}y$$

同样有

$$E(Y)=\iint\limits_{x^2+y^2\leqslant 1}y\cdot\frac{1}{\pi}\mathrm{d}x\mathrm{d}y$$

由于以上两个积分里被积函数为奇函数，且积分区域有对称性，则有 $E(X)=E(Y)=0$.

对于 σ_{XX}，σ_{YY}，σ_{XY}，有

$$\sigma_{XX}=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x-E(X))^2p(x,y)\mathrm{d}x\mathrm{d}y$$

$$=\frac{1}{\pi}\iint\limits_{x^2+y^2\leqslant 1}x^2\mathrm{d}x\mathrm{d}y$$

$$=\frac{1}{\pi}\int_0^{2\pi}\mathrm{d}\theta\int_0^1 r^2\cos^2\theta\cdot r\mathrm{d}r \quad \text{（作极坐标变换）}$$

$$=\frac{1}{\pi}\left(\int_0^{2\pi}\cos^2\theta\mathrm{d}\theta\right)\cdot\left(\int_0^1 r^3\mathrm{d}r\right)=\frac{1}{4}$$

自然也有 $\sigma_{YY}=\dfrac{1}{4}$.而

$$\sigma_{XY}=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}[x-E(X)][y-E(Y)]p(x,y)\mathrm{d}x\mathrm{d}y$$

$$=\frac{1}{\pi}\iint\limits_{x^2+y^2\leqslant 1}xy\mathrm{d}x\mathrm{d}y$$

这个二重积分值，按积分区域 $x^2+y^2\leqslant 1$ 与被积函数 xy 的对称

性，是等于 0 的. 因此有 $\sigma_{XY}=0$.

可是，本例中 X 与 Y 并不相互独立（请读者自行验证）.

例 3.3 设 (X,Y) 服从二维正态分布，密度函数为

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

求 σ_{XY}.（前面已求得：$E(X)=\mu_1$，$E(Y)=\mu_2$；$\sigma_{XX}=\sigma_1^2$，$\sigma_{YY}=\sigma_2^2$.）

解

$$\begin{aligned}\sigma_{XY}&=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}[x-E(X)][y-E(Y)]p(x,y)\mathrm{d}x\mathrm{d}y\\&=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x-\mu_1)(y-\mu_2)\\&\quad\cdot e^{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2-\frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\left(\frac{y-\mu_2}{\sigma_2}\right)^2\right]}\mathrm{d}x\mathrm{d}y\\&\xlongequal{\text{令 } u=\frac{x-\mu_1}{\sigma_1},\,v=\frac{y-\mu_2}{\sigma_2}}\frac{\sigma_1\sigma_2}{2\pi\sqrt{1-\rho^2}}\\&\quad\cdot\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}uve^{-\frac{1}{2(1-\rho^2)}(u^2-2\rho uv+v^2)}\mathrm{d}u\mathrm{d}v=\frac{\sigma_1\sigma_2}{2\pi\sqrt{1-\rho^2}}\\&\quad\cdot\int_{-\infty}^{+\infty}\left[\int_{-\infty}^{+\infty}uve^{-\frac{1}{2(1-\rho^2)}[(u-\rho v)^2+(1-\rho^2)v^2]}\mathrm{d}u\right]\mathrm{d}v\\&=\frac{\sigma_1\sigma_2}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\left[ve^{-\frac{1}{2}v^2}\frac{1}{\sqrt{2\pi}\sqrt{1-\rho^2}}\right.\\&\quad\left.\cdot\int_{-\infty}^{+\infty}ue^{-\frac{1}{2(1-\rho^2)}(u-\rho v)^2}\mathrm{d}u\right]\mathrm{d}v\end{aligned}$$

注意到上式方括号内的因子

$$\frac{1}{\sqrt{2\pi}\sqrt{1-\rho^2}}\int_{-\infty}^{+\infty}ue^{-\frac{1}{2(1-\rho^2)}(u-\rho v)^2}\mathrm{d}u$$

是 $N(\rho v,(\sqrt{1-\rho^2})^2)$的均值,所以它等于 ρv. 于是得

$$\sigma_{XY}=\frac{\sigma_1\sigma_2}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\rho v^2 e^{-\frac{1}{2}v^2}\mathrm{d}v=\rho\sigma_1\sigma_2$$

这个结果给出了二维正态密度第五个参数 ρ 的概率意义:

$$\rho=\frac{\sigma_{XY}}{\sqrt{\sigma_{XX}}\sqrt{\sigma_{YY}}}\tag{3.8}$$

对于二维正态来说,由本章§1中第5段的讨论知道,$\rho=0$ 是 X,Y 独立的充分必要条件. 由上面(3.8)式容易看出 $\rho=0$ 和 $\sigma_{XY}=0$ 是等价的. 所以说,对于二维正态分布,$\sigma_{XY}=0$ 也是 X,Y 独立的充分条件. 而对于一般的二维分布,充分性不成立.

4. 相关系数

(1) **定义 3.2** 称

$$\frac{\sigma_{XY}}{\sqrt{\sigma_{XX}}\sqrt{\sigma_{YY}}}$$

为 X,Y 的**相关系数**(自然要求分母不为0),记作 ρ_{XY}. 在不会引起混淆的情况下,简记作 ρ. 即

$$\rho_{XY}=\frac{\sigma_{XY}}{\sqrt{\sigma_{XX}}\sqrt{\sigma_{YY}}}\tag{3.9}$$

从定义看到,ρ 跟协方差 σ_{XY} 只差一个常数倍数. 另外,对于二维正态分布,正好有:

$$\rho_{XY}=\rho$$

即二维正态分布的第五个参数 ρ 就是相关系数.

(2) 相关系数 ρ 满足:

$$|\rho|\leqslant 1\tag{3.10}$$

理由如下:对于任意实数 λ,有

$$\begin{aligned}D(Y-\lambda X)&=E[Y-\lambda X-E(Y-\lambda X)]^2\\&=E\{[Y-E(Y)]-\lambda[X-E(X)]\}^2\\&=E[Y-E(Y)]^2+\lambda^2E[X-E(X)]^2\end{aligned}$$

$$-2\lambda E\{[Y-E(Y)][X-E(X)]\}$$

$$=\lambda^2\sigma_{XX}-2\lambda\sigma_{XY}+\sigma_{YY} \tag{3.11}$$

在上式中令 $\lambda=b\triangleq\dfrac{\sigma_{XY}}{\sigma_{XX}}$,则有

$$D(Y-bX)=\sigma_{YY}\left(1-\frac{\sigma_{XY}^2}{\sigma_{XX}\sigma_{YY}}\right)=\sigma_{YY}(1-\rho^2)$$

由于方差是非负的,故 $\sigma_{YY}(1-\rho^2)\geqslant 0$,所以 $|\rho|\leqslant 1$. 这就证明了(3.10). 我们还看出,$|\rho|=1$ 的充要条件是 $D(Y-bX)=0$.

经过数学研究知道,$D(\xi)=0$ 的充要条件是存在常数 a 使 $P\{\xi=a\}=1$. 所以 $|\rho|=1$ 的充要条件是存在常数 a 使 $P\{Y-bX=a\}=1$,即 $P\{Y=a+bX\}=1$.

相关系数 ρ 的实际意义是:它刻画了 X,Y 间线性关系的近似程度. 一般说来,$|\rho|$ 越接近 1,X 与 Y 越近似地有线性关系. 要注意的是,ρ 只刻画 X 与 Y 间线性关系的近似程度,当 X,Y 之间有很密切的曲线关系时,$|\rho|$ 的数值也可能很小,例如,X 服从 $N(0,1)$,$Y=X^2$,此时 Y 与 X 有很密切的曲线关系,但是 $\rho_{XY}=0$.

线性预测与相关系数

这里介绍一下所谓**线性预测**问题. 以此强调 ρ 只是刻画 X,Y 间线性关系的程度的一个数量特征.

给定两个随机变量 $X,Y(\sigma_{XX}\neq 0,\sigma_{YY}\neq 0)$,如果希望用 X 的线性函数 $a+bX$ 来近似代替 Y(或说来预测 Y),问:a,b 应如何选择,使 $a+bX$ 最接近 Y;接近的程度又如何?

要解决这个问题,首先要对所谓"接近"给出明确而合理的标准. 注意,这里的"接近",是对两个随机变量 $a+bX$ 与 Y 而言的. 一个自然而合理的标准是用量

$$Q\triangleq E[Y-(a+bX)]^2$$

作为 Y 与 $a+bX$ 之间的"距离";它越小就表示 Y 与 $a+bX$ 越接近.

a,b 如何选择呢? 问题已很清楚,就是选这样的 a,b,它们使 Q 达到极小. 下面我们用配方的办法来具体找出这样的 a,b(也可用微积分的办法).

$$
\begin{aligned}
Q &= E[Y-(a+bX)]^2 \\
&= E\{[(Y-E(Y))-b(X-E(X))]+ \\
&\quad [E(Y)-bE(X)-a]\}^2
\end{aligned}
$$

记第一个方括号内的随机变量为 Z,即

$$Z \triangleq Y-E(Y)-b(X-E(X))$$

容易验证有 $E(Z)=0$.

再注意到第二个方括号内的量是常量.于是有

$$
\begin{aligned}
Q &= E(Z^2)+[E(Y)-bE(X)-a]^2 \\
&= \sigma_{YY}-2b\sigma_{XY}+b^2\sigma_{XX}+[E(Y)-bE(X)-a]^2 \\
&= \sigma_{XX}\left(b-\frac{\sigma_{XY}}{\sigma_{XX}}\right)^2+\left(\sigma_{YY}-\frac{\sigma_{XY}^2}{\sigma_{XX}}\right)+ \\
&\quad [E(Y)-bE(X)-a]^2 \quad (\text{配方})
\end{aligned}
$$

我们看到,Q 由三部分构成:

(i) $\sigma_{YY}-\dfrac{\sigma_{XY}^2}{\sigma_{XX}}$;

(ii) $\sigma_{XX}\left(b-\dfrac{\sigma_{XY}}{\sigma_{XX}}\right)^2$;

(iii) $[E(Y)-bE(X)-a]^2$.

a,b 取什么值能使“(i)”+“(ii)”+“(iii)”最小呢?“(i)”是由 X,Y 本身的量决定的.不论 a,b 如何选择对它都无影响;“(ii)”是非负的.显然,选 $b=b^*=\dfrac{\sigma_{XY}}{\sigma_{XX}}$ 时,它最小,其值为 0;再由“(iii)”,决定选 $a=a^*=E(Y)-b^*E(X)$.

综合上述,取

$$b^*=\frac{\sigma_{XY}}{\sigma_{XX}}$$

$$a^*=E(Y)-b^*E(X)$$

可使 Q 达到最小,最小值是

$$Q_{\min}=\sigma_{YY}-\frac{\sigma_{XY}^2}{\sigma_{XX}}=\sigma_{YY}(1-\rho^2) \tag{3.12}$$

回顾 $Q_{\min}$ 的意义,我们知道它是 X 的线性函数 $a+bX$ 与 Y 之间的最小“距离”.它越小就表示用 X 的线性函数来近似 Y 的程度越高,也即 X,Y 之间的线性关系越密切.而(3.12)式表明,$|\rho|$ 越接近 1(即越大),$Q_{\min}$ 就越小.这说明 $|\rho|$ 越大,X,Y 间的线性关系越密切.ρ 刻画了 X,Y 间线性关系的密

切程度.

习 题 十 三

1. 设(X,Y)的联合密度

$$p(x,y)=\begin{cases}4xye^{-(x^2+y^2)} & \text{当 } x>0,y>0\\ 0 & \text{其他}\end{cases}$$

求 $Z=\sqrt{X^2+Y^2}$ 的均值(用两种方法).

2. 证明:如果 X 与 Y 独立,则

$$D(XY)=D(X)D(Y)+[E(X)]^2D(Y)+[E(Y)]^2D(X)$$

3. 设(X,Y)的联合密度

$$p(x,y)=\frac{A}{(x^2+y^2+1)^2}\text{(}A\text{ 是常数)}$$

试求出 A 的数值,并问 σ_{XX} 与 σ_{YY} 是否存在?

4. 设(X,Y)服从区域 $D=\{(x,y):0<x<1,0<y<x\}$上的均匀分布,求相关系数 ρ.

5. 设 X_1,X_2独立,概率密度分别为

$$p_1(x)=\begin{cases}2x & 0\leqslant x\leqslant 1\\ 0 & \text{其他}\end{cases}$$

$$p_2(x)=\begin{cases}e^{-(x-5)} & x>5\\ 0 & \text{其他}\end{cases}$$

求 $E(X_1\cdot X_2)$.

6. 已知 $D(X)=25,D(Y)=36,\rho=0.4$.求 $D(X+Y)$及$D(X-Y)$.

7. 设(X,Y)服从二维正态分布,$E(X)=E(Y)=0,D(X)=a^2,D(Y)=b^2,\rho=0$,求$(X,Y)$落在区域 $D=\left\{(x,y):\frac{x^2}{a^2}+\frac{y^2}{b^2}\leqslant k^2\right\}$中的概率.$(k>0)$

8. 直接验证:若 $Y=a+bX$,则

$$\rho=\begin{cases}1 & \text{当 } b>0\\ -1 & \text{当 } b<0\end{cases}$$

9. 设 $X\sim N(0,1)$,而 $Y=X^n$(n 是正整数).求 ρ_{XY}.

§4 关于 n 维随机向量

对于一般的 n 维随机向量,可仿照二维的情形进行讨论.这

里，我们只是把基本的概念与结果列举出来，读者不难看出这些是二维随机向量情形的推广．

1．联合密度与边缘密度

定义 4.1 对于 n 维随机向量 $\boldsymbol{\xi}=(X_1,X_2,\cdots,X_n)$，如果存在非负函数 $p(x_1,x_2,\cdots,x_n)$，使对于任意 n 维长方体 $D=\{(x_1,x_2,\cdots,x_n):a_1<x_1<b_1,a_2<x_2<b_2,\cdots,a_n<x_n<b_n\}$均成立：

$$P\{\boldsymbol{\xi}\in D\}=\iint\limits_D\cdots\int p(x_1,x_2,\cdots,x_n)\mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_n \quad (4.1)$$

则称 $\boldsymbol{\xi}=(X_1,X_2,\cdots,X_n)$是连续型的，并称 $p(x_1,x_2,\cdots,x_n)$为 $\boldsymbol{\xi}$ 的分布密度，也称 $p(x_1,x_2,\cdots,x_n)$为$(X_1,X_2,\cdots,X_n)$的联合分布密度(简称联合密度)．

利用较多的数学知识，从(4.1)出发可以证明，对于 n 维空间中相当任意的集合D，仍成立

$$P\{\boldsymbol{\xi}\in D\}=\iint\limits_D\cdots\int p(x_1,x_2,\cdots,x_n)\mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_n \quad (4.2)$$

(4.2)是连续型的 n 维随机向量的基本公式．

我们还称$(X_1,X_2,\cdots,X_n)$的一部分分量所构成的向量——例如(X_1,X_2)——的分布密度为边缘密度．特别地，每个分量 X_i 的分布密度 $p_i(x_i)$当然也是$(X_1,X_2,\cdots,X_n)$的边缘密度，称它们为单个密度．

设 $X_1,X_2,\cdots,X_n$的联合密度为$p(x_1,x_2,\cdots,x_n)$，则边缘密度就都可由 $p(x_1,x_2,\cdots,x_n)$求得．例如，X_1的边缘密度：

$$p_1(x_1)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}p(x_1,x_2,\cdots,x_n)\mathrm{d}x_2\mathrm{d}x_3\cdots\mathrm{d}x_n \quad (4.3)$$

又如(X_1,X_2)的边缘密度：

$$p_{12}(x_1,x_2)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}p(x_1,x_2,\cdots,x_n)\mathrm{d}x_3\mathrm{d}x_4\cdots\mathrm{d}x_n$$

在众多的 n 维随机向量中，最重要的是所谓 n 维正态随机向量．

定义 4.2 称随机向量 $\xi=(X_1,X_2,\cdots,X_n)$服从 n 维正态分布(也称 ξ 是 n 维正态随机向量),如果它的分布密度是这样的:

$$p(x_1,x_2,\cdots,x_n)=\frac{1}{(2\pi)^{\frac{n}{2}}|\boldsymbol{\Sigma}|^{\frac{1}{2}}}\exp\left\{-\frac{1}{2}(x-\mu)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(x-\mu)\right\} \tag{4.4}$$

其中 $\mu=(\mu_1,\mu_2,\cdots,\mu_n)^{\mathrm{T}}$是固定的向量,$\boldsymbol{\Sigma}=(\sigma_{ij})_{n\times n}$是固定的 n 阶正定矩阵,$\boldsymbol{\Sigma}^{-1}$是 $\boldsymbol{\Sigma}$ 的逆矩阵,$|\boldsymbol{\Sigma}|$是 $\boldsymbol{\Sigma}$ 的行列式,$x=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$,T 表示矩阵的转置运算.

当随机向量 ξ 有密度(4.4)时,常记作 $\xi\sim N(\boldsymbol{\mu},\boldsymbol{\Sigma})$.表达式(4.4)比较复杂,初学者不必记住它.

2. 独立性

定义 4.3 设 $X_1,X_2,\cdots,X_n$是 n 个随机变量,如果对任意 $a_i<b_i(i=1,2,\cdots,n)$,事件$\{a_1<X_1<b_1\}$,$\{a_2<X_2<b_2\}$,$\cdots$,$\{a_n<X_n<b_n\}$相互独立,则称 $X_1,X_2,\cdots,X_n$是**相互独立的**.

定理 4.1 设 $X_1,X_2,\cdots,X_n$的分布密度分别是 $p_1(x_1)$,$p_2(x_2)$,$\cdots$,$p_n(x_n)$,则 $X_1,X_2,\cdots,X_n$相互独立的充要条件是:n 元函数 $p_1(x_1)p_2(x_2)\cdots p_n(x_n)$是$(X_1,X_2,\cdots,X_n)$的联合密度.

证明很容易,和§1 定理 1.2 的证法相仿,从略.

3. n 个随机变量的函数的分布

这里只列出一个类似于(2.3)的公式,而不进行具体讨论.

$$P\{f(X_1,X_2,\cdots,X_n)\leqslant z\}$$
$$=\underset{f(x_1,x_2,\cdots,x_n)\leqslant z}{\iint\cdots\int}p(x_1,x_2,\cdots,x_n)\mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_n \tag{4.5}$$

公式(4.5)把找 $f(X_1,X_2,\cdots,X_n)$的分布函数的问题,归结为一个 n 重积分的计算问题.

4. 数字特征

(1) 均值公式

$$E[f(X_1,X_2,\cdots,X_n)]=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}f(x_1,x_2,\cdots,x_n).$$

$$p(x_1,x_2,\cdots,x_n)\mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_n \tag{4.6}$$

其中 $p(x_1,x_2,\cdots,x_n)$是$(X_1,X_2,\cdots,X_n)$的联合密度(要求右端的积分绝对收敛).

以上的讨论形式上限于连续型;其实对于离散型也有相应的结果.通常称向量$(E(X_1),E(X_2),\cdots,E(X_n))$为随机向量$(X_1,X_2,\cdots,X_n)$的均值.

(2) 均值与方差的性质

$$E(X_1+X_2+\cdots+X_n)=E(X_1)+E(X_2)+\cdots+E(X_n) \tag{4.7}$$

当 $X_1,X_2,\cdots,X_n$独立时有:

$$E(X_1X_2\cdots X_n)=E(X_1)E(X_2)\cdots E(X_n) \tag{4.8}$$

$$D(X_1+X_2+\cdots+X_n)=D(X_1)+D(X_2)+\cdots+D(X_n)$$

(3) 协方差与协差阵

记 $$\sigma_{ij}=E[(X_i-E(X_i))(X_j-E(X_j))]$$
$$(i=1,2,\cdots,n,j=1,2,\cdots,n) \tag{4.9}$$

很明显,对于 $i\neq j$,σ_{ij}是第i个分量X_i与第j个分量X_j的协方差;而σ_{ii}是第i个分量X_i的方差.

称矩阵

$$\begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2n} \\ \vdots & \vdots & & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_{nn} \end{pmatrix} \tag{4.10}$$

为$(X_1,X_2,\cdots,X_n)$的协差阵,记为 $\boldsymbol{\Sigma}$($\boldsymbol{\Sigma}$ 显然是对称矩阵).

可以验证矩阵 $\boldsymbol{\Sigma}$ 是非负定的[①].

---

① 称 n 阶实对称矩阵 $\boldsymbol{A}=(a_{ij})$是非负定的,若对任何实数 $t_1,t_2,\cdots,t_n$ 均有 $\sum\limits_{i,j}a_{ij}t_it_j\geqslant 0$.

(4) 相关系数与相关阵

记

$$\rho_{ij} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}}\sqrt{\sigma_{jj}}}(i=1,2,\cdots,n, j=1,2,\cdots,n)$$

显然,对于 $i \neq j$, ρ_{ij} 是 X_i, X_j 的相关系数;而

$$\rho_{ii} = 1 \quad (i=1,2,\cdots,n)$$

称矩阵

$$\begin{pmatrix} \rho_{11} & \rho_{12} & \cdots & \rho_{1n} \\ \rho_{21} & \rho_{22} & \cdots & \rho_{2n} \\ \vdots & \vdots & & \vdots \\ \rho_{n1} & \rho_{n2} & \cdots & \rho_{nn} \end{pmatrix} \tag{4.11}$$

为 $(X_1, \cdots, X_n)$ 的相关阵,记为 $\boldsymbol{R}$($\boldsymbol{R}$ 也是对称阵).

若记

$$\boldsymbol{C} \triangleq \begin{pmatrix} \frac{1}{\sqrt{\sigma_{11}}} & & & 0 \\ & \frac{1}{\sqrt{\sigma_{22}}} & & \\ & & \ddots & \\ 0 & & & \frac{1}{\sqrt{\sigma_{nn}}} \end{pmatrix}$$

则有

$$\boldsymbol{R} = \boldsymbol{C\Sigma C}$$

作为本节的结束,我们还要介绍一个名词:n 维分布函数.

定义 4.4 设 $\boldsymbol{\xi} = (X_1, X_2, \cdots, X_n)$ 是 n 维随机向量,称 n 元函数 $F(x_1, x_2, \cdots, x_n) = P\{X_1 \leqslant x_1, X_2 \leqslant x_2, \cdots, X_n \leqslant x_n\}$ 为 $\boldsymbol{\xi}$ 的分布函数(也称为 $X_1, X_2, \cdots, X_n$ 的联合分布函数).

如果 $\boldsymbol{\xi}$ 有分布密度 $p(x_1, x_2, \cdots, x_n)$,则有下列关系式:

$$F(x_1, x_2, \cdots, x_n) = \int_{-\infty}^{x_1}\int_{-\infty}^{x_2}\cdots\int_{-\infty}^{x_n} p(u_1, u_2, \cdots, u_n)\mathrm{d}u_1\mathrm{d}u_2\cdots\mathrm{d}u_n$$

习题十四

1. 已知(X,Y,Z)的联合密度为

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

分别求出 X,Y,Z 的单个密度. 又：X,Y,Z 相互独立吗？

2. 设 $X_1,X_2,\cdots,X_n$独立同分布，都服从 $N(\mu,\sigma^2)$，求$(X_1,X_2,\cdots,X_n)$的联合密度.

3. 设 X,Y,Z 独立同分布，服从 $N(0,1)$，求$\sqrt{X^2+Y^2+Z^2}$的概率密度.

4. 设 $X_1,X_2,\cdots,X_n$独立同分布，分布的均值是 μ，方差是 σ^2. 而 $Y=\frac{1}{n}(X_1+X_2+\cdots+X_n)$，求 $E(Y),D(Y)$.

5. 设 $X_1,X_2,\cdots,X_n$ 独立，都服从参数为 m,η 的韦布尔分布$\left(p(x)=\frac{m}{\eta^m}x^{m-1}e^{-\left(\frac{x}{\eta}\right)^m},m>0,\eta>0,x>0\right)$，试证明：$\min(X_1,X_2,\cdots,X_n)$仍然服从韦布尔分布.

6. 将 n 只球放入 M 只盒子中去，设每只球落入各个盒子是等可能性的. 求有球的盒子数 X 的均值.

(提示：引进随机变量：

$$X_i=\begin{cases}1 & 第\ i\ 只盒子中有球\\ 0 & 第\ i\ 只盒子中无球\end{cases}$$

显然有 $X=\sum_{i=1}^{M}X_i$.)

7. 求事件在 n 次独立试验中发生次数的均值与方差，如果该事件在第 i 次试验中发生的概率等于 $p_i(i=1,2,\cdots,n)$.

8. 对于随机变量 X,Y,Z，已知

$$E(X)=E(Y)=1,E(Z)=-1$$

$$D(X)=D(Y)=D(Z)=1$$

$$\rho_{XY}=0,\rho_{XZ}=\frac{1}{2},\rho_{YZ}=-\frac{1}{2}$$

求 $E(X+Y+Z),D(X+Y+Z)$.

9．设 $X_1, X_2, \cdots, X_n$ 相互独立，且 $X_i \sim N(\mu, \sigma^2)(i=1,2,\cdots,n)$．试证明①：

$$\frac{X_1+X_2+\cdots+X_n}{n} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

（提示：应用习题十二中第 3 题的结论．）

*§5　条件分布与条件期望

设 X 与 Y 是两个随机变量．给定实数 y，如果 $P(Y=y)>0$，则称 x 的函数 $P(X \leqslant x \mid Y=y)$ 为 $Y=y$ 的条件下 X 的条件分布函数，记作 $F_{X|Y}(x|y)$．显然，根据条件概率的定义有

$$F_{X|Y}(x|y)=P(X \leqslant x, Y=y)/P(Y=y) \tag{5.1}$$

如果 $P(Y=y)=0$（例如 Y 为连续型随机变量），怎样定义 X 的条件分布函数呢？这就不能从条件概率的初等定义（见第一章§5）出发了．我们采用下列很自然的处理方法．

定义 5.1　设对任何 $\varepsilon>0, P(y-\varepsilon<Y \leqslant y+\varepsilon)>0$．若极限

$$\lim_{\varepsilon \to 0} P(X \leqslant x \mid y-\varepsilon<Y \leqslant y+\varepsilon)$$

存在，则称此极限为 $Y=y$ 的条件下 X 的条件分布函数，记作 $P(X \leqslant x \mid Y=y)$ 或 $F_{X|Y}(x|y)$，即

$$F_{X|Y}(x|y)=\lim_{\varepsilon \to 0} P(X \leqslant x \mid y-\varepsilon<Y \leqslant y+\varepsilon) \tag{5.2}$$

不难看出，当 $P(Y=y)>0$ 时(5.2)与(5.1)有相同的结果．

应注意的是，条件分布涉及联合分布，由后者所确定．分两种情形进行讨论．

（一）离散型情形

设 (X, Y) 是二维离散型随机向量，其概率分布为

$P(X=x_i, Y=y_j)=p_{ij}(i=1,2,\cdots; j=1,2,\cdots)$，这里 $P(Y=y_j)>0(j \geqslant 1)$．则在 $Y=y_j$ 的条件下 X 的条件分布是

① 此结论在数理统计中常常用到．

$$P(X=x_i \mid Y=y_j)=\frac{P(X=x_i, Y=y_j)}{P(Y=y_j)}$$

$$=\frac{p_{ij}}{\sum_{k\geqslant 1} p_{kj}} \quad (i=1,2,\cdots) \tag{5.3}$$

例 5.1 一射手进行射击,单发击中目标的概率为 $p(0<p<1)$,射击进行到击中目标两次为止.设以 X 表示第一次击中目标所需的射击次数,以 Y 表示总共进行的射击次数.试求(X,Y)的联合分布和条件分布?

解 显然 $P(X=m, Y=n)=p^2q^{n-2}(m=1,2,\cdots,n-1; n=2,3,\cdots)$,这里 $q=1-p$.

对其他的 m,n 显然 $P(X=m, Y=n)=0$.

于是 $P(X=m)=\sum_{n=m+1}^{\infty} P(X=m, Y=n)$

$$=\sum_{n=m+1}^{\infty} p^2q^{n-2}=pq^{m-1} \quad (m=1,2,\cdots)$$

$$P(Y=n)=\sum_{m=1}^{n-1} P(X=m, Y=n)=\sum_{m=1}^{n-1} p^2q^{n-2}$$

$$=(n-1)p^2q^{n-2} \quad (n=2,3,\cdots)$$

所以条件分布是:

$$P(X=m \mid Y=n)=\begin{cases}\dfrac{1}{n-1} & (n\geqslant 2, m=1,\cdots,n-1)\\ 0 & \text{其他情形}\end{cases}$$

$$P(Y=n \mid X=m)=\begin{cases}pq^{n-m-1} & (n\geqslant m+1)\\ 0 & (n\leqslant m)\end{cases}$$

(二) 连续型情形

设(X,Y)有联合分布函数 $F(x,y)$,联合密度 $p(x,y)$.我们在对 $p(x,y)$作出若干假定(这些假定在实际应用中碰到的大多数情形下是满足的)后,可找出条件分布的表达式.

实际上,Y 的分布密度 $p_Y(y)=\int_{-\infty}^{+\infty} p(x,y)\mathrm{d}x$,$Y$ 的分布函

数 $F_Y(y)=\int_{-\infty}^{y} p_Y(u)\mathrm{d}u$.若 $p_Y(u)$在 $u=y$ 处连续,则$\frac{\mathrm{d}F_Y(y)}{\mathrm{d}y}=p_Y(y)$.从而$\lim\limits_{\varepsilon\to 0}\frac{F_Y(y+\varepsilon)-F_Y(y-\varepsilon)}{2\varepsilon}=p_Y(y)$.

若$\int_{-\infty}^{x} p(u,v)\mathrm{d}u$ 在 $v=y$处连续,则$\frac{\partial F(x,y)}{\partial y}=\int_{-\infty}^{x} p(u,y)\mathrm{d}u$,从而$\lim\limits_{\varepsilon\to 0}\frac{F(x,y+\varepsilon)-F(x,y-\varepsilon)}{2\varepsilon}=\int_{-\infty}^{x} p(u,y)\mathrm{d}u$.于是 $F_{X|Y}(x\mid y)$

$$
\begin{aligned}
&=\lim_{\varepsilon\to 0}P(X\leqslant x\mid y-\varepsilon<Y\leqslant y+\varepsilon)\\
&=\lim_{\varepsilon\to 0}\frac{P(X\leqslant x,y-\varepsilon<Y\leqslant y+\varepsilon)}{P(y-\varepsilon<Y\leqslant y+\varepsilon)}\\
&=\lim_{\varepsilon\to 0}\frac{F(x,y+\varepsilon)-F(x,y-\varepsilon)}{F_Y(y+\varepsilon)-F_Y(y-\varepsilon)}=\frac{\int_{-\infty}^{x} p(u,y)\mathrm{d}u}{p_Y(y)}\\
&=\int_{-\infty}^{x}\frac{p(u,y)}{p_Y(y)}\mathrm{d}u \quad (\text{当 } p_Y(y)>0 \text{ 时})
\end{aligned}
$$

这就求出了 $Y=y$ 的条件下 X 的条件分布函数.自然称$\frac{p(x,y)}{p_Y(y)}$为 $Y=y$ 的条件下 X 的条件分布密度,记作 $p_{X|Y}(x|y)$,即

$$
p_{X|Y}(x|y)=\frac{p(x,y)}{p_Y(y)} \tag{5.4}
$$

这与离散型情形下的条件分布的形式(5.3)很相似.(应注意,(5.4)式的前提是假定 $p_Y(y)>0$).

例 5.2 设(X,Y)服从二维正态分布,其联合密度为

$p(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2-\frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\left(\frac{y-\mu_2}{\sigma_2}\right)^2\right]\right\}$ (μ_1,μ_2是实数,$\sigma_1>0,\sigma_2>0,|\rho|<1$).

易知 Y 的分布密度 $p_Y(y)=\frac{1}{\sqrt{2\pi}\sigma_2}\exp\left\{-\frac{(y-\mu_2)^2}{2\sigma_2^2}\right\}$代入(5.4)

得

$$p_{X|Y}(x|y)=\frac{1}{\sqrt{2\pi(1-\rho^2)}\sigma_1}\exp\left\{-\frac{(x-m)^2}{2(1-\rho^2)\sigma_1^2}\right\},$$

其中 $m=\mu_1+\rho\frac{\sigma_1}{\sigma_2}(y-\mu_2)$，它是 y 的线性函数.

定义 5.2 （条件期望）设 X,Y 是两个随机变量，有联合密度 $p(x,y)$，设 $Y=y$ 的条件下 X 有条件分布密度 $p_{X|Y}(x|y)$. 则 $\int_{-\infty}^{+\infty}xp_{X|Y}(x|y)\mathrm{d}x$（当积分绝对收敛时）叫做 $Y=y$ 的条件下 X 的**条件期望**，记作 $E(X|Y=y)$.（当(X,Y)是离散型随机向量时，可类似地下定义，从略）.

从(5.4)知

$$E(X|Y=y)=\frac{1}{p_Y(y)}\int_{-\infty}^{+\infty}xp(x,y)\mathrm{d}x \tag{5.5}$$

这里 $p_Y(y)=\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}x$ 是 Y 的分布密度. 条件期望 $E(X|Y=y)$的含义是：在 $Y=y$ 的条件下，X 取值的平均大小. 可以证明：

$$E(X)=\int_{\{y:p_Y(y)>0\}}E(X|Y=y)p_Y(y)\mathrm{d}y \tag{5.6}$$

我们首先指出，若 $p_Y(y)=0$，则

$$\int_{-\infty}^{+\infty}xp(x,y)\mathrm{d}x=0$$

实际上，对任何 $A>0$，$\left|\int_{-A}^{A}xp(x,y)\mathrm{d}x\right|\leqslant A\int_{-A}^{A}p(x,y)\mathrm{d}x\leqslant A\int_{-\infty}^{+\infty}p(x,y)\mathrm{d}x=Ap_Y(y)=0$. 于是 $\int_{-\infty}^{+\infty}xp(x,y)\mathrm{d}x=\lim\limits_{A\to+\infty}\int_{-A}^{A}xp(x,y)\mathrm{d}x=0$. 从而(5.6)式等号的右端 $=\int_{\{y:p_Y(y)>0\}}\left[\int_{-\infty}^{+\infty}xp(x,y)\mathrm{d}x\right]\mathrm{d}y=\int_{-\infty}^{+\infty}\left[\int_{-\infty}^{+\infty}xp(x,y)\mathrm{d}x\right]\mathrm{d}y=$

$\int_{-\infty}^{+\infty} x\left[\int_{-\infty}^{+\infty} p(x,y)\mathrm{d}y\right]\mathrm{d}x = \int_{-\infty}^{+\infty} x p_X(x)\mathrm{d}x = E(X)$（$p_X(x)$是 X 的分布密度）. 这就证明了(5.6)成立.

有时把(5.6)写成下列形式

$$E(X) = \int_{-\infty}^{+\infty} E(X \mid Y=y) p_Y(y)\mathrm{d}y \tag{5.7}$$

这时要注意，当 $p_Y(y)=0$ 时，规定 $E(X \mid Y=y)$为 0.

公式(5.6)的意义在于：为求 $E(X)$，有时 $E(X)$不便直接求，而条件期望 $E(X \mid Y=y)$从含义出发反而易于求出，此时利用(5.6)就可算出 $E(X)$来.(5.6)是第一章中全概公式的一种推广.

类似地，可求出 $X=x$ 的条件下 Y 的条件分布密度 $p_{Y|X}(y|x)$，即

$$p_{Y|X}(y|x) = \frac{p(x,y)}{p_X(x)}, \tag{5.8}$$

其中

$$p_X(x) = \int_{-\infty}^{+\infty} p(x,y)\mathrm{d}y \tag{5.9}$$

是 X 的分布密度. 在 $X=x$ 的条件下 Y 的条件期望是

$$\begin{aligned} E(Y \mid X=x) &= \int_{-\infty}^{+\infty} y p_{Y|X}(y|x)\mathrm{d}y \\ &= \frac{1}{p_X(x)}\int_{-\infty}^{+\infty} y p(x,y)\mathrm{d}y \end{aligned} \tag{5.10}$$

与公式(5.6)类似，有公式

$$E(Y) = \int_{\{x:\, p_X(x)>0\}} E(Y \mid X=x) p_X(x)\mathrm{d}x \tag{5.11}$$

例 5.3 设 $U_1, U_2, \cdots$是相互独立同分布的随机变量列，共同分布是(0,1)上均匀分布，

$$N = \min\left\{n : \sum_{i=1}^{n} U_i > 1\right\}$$

试求出 $E(N)$.

解 对任何 $x\in[0,1]$,令

$$N(x)=\min\left\{n:\sum_{i=1}^{n}U_i>x\right\}$$
$$m(x)=E[N(x)]$$

据公式(5.6)知

$$m(x)=\int_0^1 E[N(x)\mid U_1=y]\mathrm{d}y$$

由于

$$E[N(x)\mid U_1=y]=\begin{cases}1 & y>x\\ 1+m(x-y) & y\leqslant x\end{cases}$$

于是

$$m(x)=\int_x^1 \mathrm{d}y+\int_0^x[1+m(x-y)]\mathrm{d}y$$
$$=1+\int_0^x m(u)\mathrm{d}u$$

$m'(x)=m(x)$.由此知 $m(x)=k\mathrm{e}^x$.因为 $m(0)=1$.故 $k=1$.所以 $m(x)=\mathrm{e}^x$.从而 $E(N)=m(1)=\mathrm{e}$①.

例 5.4[8] 设一家供电公司新建不久,生产不够稳定,每月可以供应某工厂的电力服从[10,30]上均匀分布(单位:万度),而该工厂每月实际生产所需要的电力服从[10,20]上的均匀分布.如果工厂能从这家供电公司得到足够的电力,则每 1 万度电可创造 30 万元的利润.若工厂从供电公司得不到足够的电力,则不足部分由工厂通过其他途径自行解决,此时每 1 万度电只能产生 10 万元的利润.问:该工厂每月的平均利润是多少?

解 设工厂每月实际生产所需要的电力为 X(万度),供电公司每月供应该厂的电力为 Y(万度),工厂每月的利润为 R(万

① 对于任何两个随机变量 X 与 Y(不必限制(X,Y)是离散型的或连续型的),均可定义 $Y=y$ 的条件下 X 的条件分布和条件期望,由于要用到较深的数学知识(测度论),我们就不去对这种一般情形下定义了.