§ 3.1 二维随机向量及其分布函数

一.随机向量的定义

- **定义**： 设X_1，X_2，…，X_n是n个随机变量，由它们构成的一个n维向量(X_1，X_2，…，X_n)称为**n维随机向量**，亦称**n维随机变量**，X_1，X_2，…，X_n叫做随机向量的分量。
- 例如，打靶时弹着点的位置(X，Y)就是一个二维随机向量。
- 本章主要讨论二维随机向量及其概率分布，并把它们推广到n维随机向量。

二.分布函数的定义

- **定义**：设（X，Y）是二维随机向量，对于任意的实数x，y，称二元函数

$$F(x,y)=P(X\leq x,Y\leq y)$$

为二维随机向量（X，Y）的**分布函数**，或称为（X，Y）的**联合分布函数。**

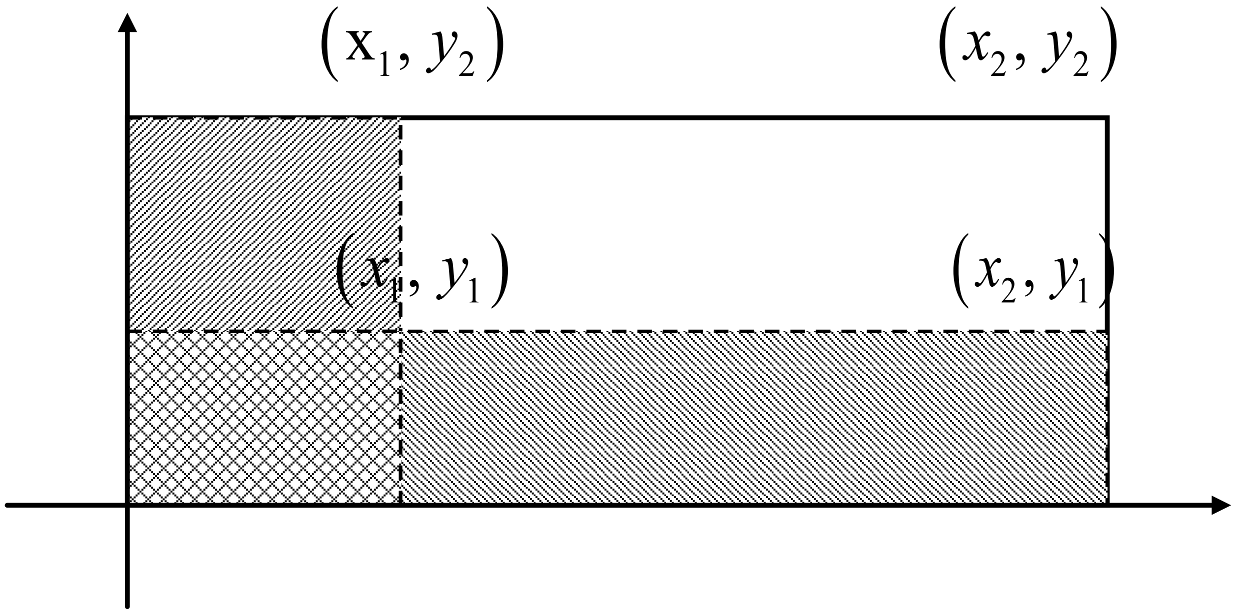
随机向量的几何意义

- 如果把二维随机向量（X，Y）看作平面上具有随机坐标（X，Y）的点，那么，分布函数F（x，y）在（x，y）处的函数值就是随机点（X，Y）落在以点（x，y）为顶点而位于该点左下方的矩形区域内的概率。

- 随机点（X，Y）落在矩形区域$\{x_1<X\le x_2，y_1<Y\le y_2\}$内的概率可用分布函数表示为

$$P(x_1<X\le x_2，\ y_1<Y\le y_2)$$
$$=F(x_2，y_2)-F(x_1，y_2)-F(x_2，y_1)+F(x_1，y_1)$$

如图所示：



分布函数的性质：

- 1. F(x,y)对x或y都是单调不减的。对任意固定的y，当$x_2>x_1$时，$F(x_2,y)\geqslant F(x_1,y)$；对任意固定的x，当$y_2>y_1$时，$F(x,y_2)\geqslant F(x,y_1)$。
- 2.0≤F(x,y)≤1。对任意固定的y, $F(-\infty, y)=\lim\limits_{x\to-\infty}F(x, y)=0$

  对任意固定的x, $F(x,-\infty)=\lim\limits_{y\to-\infty}F(x, y)=0$

  且$F(-\infty,-\infty)=\lim\limits_{\substack{x\to-\infty\\ y\to-\infty}}F(x, y)=0$　$F(+\infty,+\infty)=\lim\limits_{\substack{x\to+\infty\\ y\to+\infty}}F(x, y)=1$
- 3.F(x,y)关于x,y右连续，即

  $F(x_0+0, y)=\lim\limits_{x\to x_0+}F(x, y)=F(x_0, y)$　$F(x, y_0+0)=\lim\limits_{y\to y_0+}F(x, y)=F(x, y_0)$
- 4.对任意(x_1,y_1)和$(x_2,y_2)(x_1<x_2,y_1<y_2)$,有

$$F(x_2,y_2)-F(x_1,y_2)-F(x_2,y_1)+F(x_1,y_1)\geq 0$$

反例如下：

- **例3.1.1** 定义二元函数

$$F(x,\ y)=\begin{cases}1, & 当x+y\geq 0\\ 0, & 当x+y<0\end{cases}$$

显然，此二元函数满足性质（1）-（3）但不满足性质（4）。

因为取$x_1=-1$, $x_2=1$, $y_1=-1$ $y_2=1$，则有

$$F(1,\ 1)-F(1,\ -1)-F(-1,\ 1)+F(-1,\ -1)<0$$

所以，F(x，y)不是一个二元分布函数。

§3.2. 二维离散型随机向量

- **定义：** 如果二维随机向量（X，Y）的所有可能取值是至多可列的，则称（X，Y）为**二维离散型随机向量**。

- 显然，若（X，Y）为二维离散型随机向量，则随机变量X，Y均为一维离散型随机变量；反之也成立。

定义： 设二维离散型随机向量（X，Y）的所有可能取值的集合为$\{(x_i,\ y_j)\ i,j=1,2,\ldots\}$，则称

$$p_{ij}=P(X=x_i,\ Y=y_j)\ i,j=1,2,\ldots$$

为（X，Y）的**联合分布列（或分布列）**。二维离散型随机向量（X，Y）的联合分布列可用如下的概率分布表来表示：

X \ Y	y_1	y_2	…	y_j	…
x_1	p_{11}	p_{12}	…	p_{1j}	…
x_2	p_{21}	p_{22}	…	p_{2j}	…
…	…	…	…	…	…
x_i	p_{i1}	p_{i2}	…	p_{ij}	…
…	…	…	…	…	…

- p_{ij}具有性质：

$$（1）p_{ij} \geq 0; \quad (2)\sum_i \sum_j p_{ij} = 1$$

- 1.三项分布：
- 在n重独立试验中，若每次试验有三种结果，A_1，A_2，A_3，$P(A_i)=p_i$，$p_1+p_2+p_3=1$。设A_1发生X次，A_2发生Y次，则(X, Y)服从三项分布$M(n; p_1,p_2,p_3)$；

$$P\left(X = k_1,\ Y = k_2\right) = \frac{n!}{k_1!k_2!(n-k_1-k_2)!} p_1^{k_1} p_2^{k_2} (1 - p_1 - p_2)^{n-k_1-k_2}$$

$$k_1 + k_2 \leq n, k_1 \geq 0, k_2 \geq 0.$$

- 2.多元超几何分布。
- 总共N个元素分为A，B，C三类，每类分别有N_1，N_2，N_3个，从中不放回地取出n个，设取到A类X个，B类Y个，则(X, Y)服从二维超几何分布。

$$P\left(X=n_1,\ Y=n_2\right)=\frac{C_{N_1}^{n_1}C_{N_2}^{n_2}C_{N_3}^{n_3}}{C_N^n},$$

$$N=N_1+N_2+N_3, n=n_1+n_2+n_3.$$

- **例3.2.1** 甲乙两盒分别存放形状相同的十个球和九个球，其中甲盒4个红球3个白球3个黑球，乙盒4个红球2个白球3个黑球。现从甲盒中任取一球放入乙盒，充分混合后再从乙盒中取出一球放回甲盒，设X，Y分别表示经如此交换后，甲盒中的红球数与白球数，试求(X,Y)的联合分布列。
- 解： $P(X=3,Y=3)=P($取出一个红球，拿进一个黑球$)=0.4*0.3$
- $P(X=4,Y=3)=P($取出一个红球，拿进一个红球；取出一个白球，拿进一个白球；取出一个黑球，拿进一个黑球；$)=0.4*0.5+ 0.3*0.3+ 0.3*0.4$

$X \backslash Y$	2	3	4	$p_{i.}$
3	0	$\frac{4\times3}{10\times10}$	$\frac{4\times2}{10\times10}$	$\frac{2}{10}$
4	$\frac{3\times3}{10\times10}$	$\frac{4\times5+3\times3+3\times4}{10\times10}$	$\frac{3\times2}{10\times10}$	$\frac{56}{100}$
5	$\frac{3\times4}{10\times10}$	$\frac{3\times4}{10\times10}$	0	$\frac{24}{100}$
$p_{.j}$	$\frac{21}{100}$	$\frac{65}{100}$	$\frac{14}{100}$	1

§ 3.3. 二维连续型随机向量

- **定义**: 设（X，Y）是二维随机向量，F（x，y）为其联合分布函数，若存在非负函数p（x，y），使得对任意的实数x，y，都有

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

则称（X，Y）是**二维连续性随机向量**，p（x，y）称为（X，Y）的**密度函数**（或**联合密度函数**）。

密度函数的性质

- 1. $p(x,y) \geq 0;\quad \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} p(x,y)dxdy = 1$
- 2. $P((X,Y)\in D) = \iint_D p(x,y)dxdy$
- 3.若p(x,y)在点(x,y)处连续，则

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

- 4.p(x,y)的大小反映了二维随机变量(X,Y)在点(x,y)附近的概率大小。

- 设D为平面上的有界区域，面积为S_D，若二维随机变量(X,Y)的密度函数为：

$$p(x,y)=\begin{cases}\dfrac{1}{S_D}, & (x,y)\in D\\ 0, & \text{其它}\end{cases}$$

则称(X,Y)服从D上的**均匀分布**。

- 若二维随机向量(X,Y)的密度函数为

$$p(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\times$$

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

其中μ_1，μ_2，$\sigma_1>0$，$\sigma_2>0$，$|\rho|<1$为参数，则称(X,Y)服从**二元正态分布**$N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$，简记为$(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$。

- 一般的，n维正态分布的密度函数如下：

二次型

$$p(x_1,\cdots,x_n) = \frac{1}{(2\pi)^{\frac{n}{2}}(\det\Sigma)^{\frac{1}{2}}}\exp\left\{-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right\},\quad (x_1,\cdots,x_n)\in\mathbb{R}^n$$

其中$\mu=(\mu_1,\cdots,\mu_n)$，$\Sigma=\left(\sigma_{ij}\right)$是正定阵。记为$N_n(\mu,\Sigma)$

- 特别，在二维正态分布中

$$\Sigma=\begin{pmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2\\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{pmatrix},\Sigma^{-1}=\frac{1}{(1-\rho^2)\sigma_1^2\sigma_2^2}\begin{pmatrix}\sigma_2^2 & -\rho\sigma_1\sigma_2\\ -\rho\sigma_1\sigma_2 & \sigma_1^2\end{pmatrix}。$$

- 代入即得到上式。

- **例3.3.1** 设二维随机向量（X，Y）的密度函数为

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

求 （1）常数a；

（2）（X，Y）的分布函数F（x，y）；

（3）（X，Y）落在区域G={(x, y)|x+y<1}内的概率。

- 解：1）a=2；　　3）　1/5

$\int_0^1\int_0^{1-x} 2x^2+2xy^2\,dy\,dx$

$\int_0^1 2x^2(1-x)+\frac{2}{3}x(1-x)^3\,dx$

$$= \frac{2}{3}x^3 - \frac{1}{2}x^4 \Big|_0^1 + \frac{2}{3}\int_0^1 (1-t)\, t^3\, dt$$

$$= \frac{2}{3} - \frac{1}{2} + \left(\frac{1}{4} - \frac{1}{5}\right) \times \frac{2}{3}$$

$$= \frac{1}{6} + \frac{1}{30} = \frac{1}{5}.$$

- 2）

$$F(x,y)=\begin{cases} 0 & x<0\text{或}y<0 \\ \frac{2}{3}x^3y+\frac{1}{3}x^2y^3 & 0\le x<1,0\le y<1 \\ \frac{2}{3}x^3+\frac{1}{3}x^2 & 0\le x<1, y\ge 1 \\ \frac{2}{3}y+\frac{1}{3}y^3 & x\ge 1, 0\le y<1 \\ 1 & x\ge 1, y\ge 1 \end{cases}$$

§ 3.4. 边缘分布

- 对于二维随机向量（X，Y），随机变量X和Y都各自有自己的概率分布。随机向量的分布函数完全决定了X和Y的概率特征，这样，我们可以通过联合分布函数得到各个分量的分布函数，这样得到的分布函数称为**边缘分布函数**。

- **问题**：边缘分布能否唯一决定联合分布？

一.一般情形

- **定义**：设（X, Y）是二维随机变量，F（x，y）是它的联合分布函数，令
  F_X（x）= F（x, +∞）， F_Y（y）=F（+∞, y）
  则称F_X（x）为（X，Y）**关于X的边缘分布函数;**
  称F_Y（y）为（X, Y）**关于Y的边缘分布函数。**

$$F_X(x)=F(x,+\infty)=\lim_{y\to+\infty}F(x,y)=P(X\le x,Y<+\infty)$$

$$F_Y(y)=F(+\infty,y)=\lim_{x\to+\infty}F(x,y)=P(X<+\infty,Y\le y)$$

二. 离散型随机向量

- 设二维随机向量（X, Y）的联合分布列为

$p_{ij}=P（X=x_i，Y=y_j）i , j=1,2,\ldots$ 则

$$P(X=x_i)=P(X=x_i,Y<+\infty)=\sum_{i=1}^{+\infty}P(X=x_i,Y=y_j)=\sum_{i=1}^{+\infty}p_{ij}$$

$$P(Y=y_j)=P(X<+\infty,Y=y_j)=\sum_{i=1}^{+\infty}P(X=x_i,Y=y_j)=\sum_{i=1}^{+\infty}p_{ij}$$

记 $$p_{i.}=\sum_{j=1}^{+\infty}p_{ij} \qquad p_{.j}=\sum_{i=1}^{+\infty}p_{ij}$$

则

$$P(X=x_i)=p_{i.} \quad i=1,2\cdots \qquad P(Y=y_j)=p_{.j} \quad j=1,2\cdots$$

称为X,Y的**边缘分布列**。

- **例3.4.1：** 证明三项分布的边缘分布是二项式分布。
- **证：** 当k=0,1,…,n时

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

$$=\sum_{j=0}^{n-k}\frac{n!}{k!\,j!(n-k-j)!}p_1^k p_2^j(1-p_1-p_2)^{n-k-j}$$

$$=\frac{n!}{k!(n-k)!}p_1^k\sum_{j=0}^{n-k}\frac{(n-k)!}{j!((n-k)-j)!}p_2^j(1-p_1-p_2)^{(n-k)-j}$$

$$=\frac{n!}{k!(n-k)!}p_1^k(1-p_1)^{n-k}=C_n^k p_1^k(1-p_1)^{n-k}$$

- **例3.4.2**：设随机变量X,Y同分布，

-1	0	1
0.25	0.5	0.25

且P(XY=0)=1，求(X,Y)的联合分布律。

- **解**：

X \ Y	-1	0	1	$p_{i\cdot}$
-1	p_{11}	p_{12}	p_{13}	0.25
0	p_{21}	p_{22}	p_{23}	0.5
1	p_{31}	p_{32}	p_{33}	0.25
$p_{\cdot j}$	0.25	0.5	0.25	1

- 由P(AB)=0得$p_{11}=p_{13}=p_{31}=p_{33}=0$;
- 利用边缘分布计算得：
- $p_{21}=p_{12}=p_{23}=p_{32}=0.25$；$p_{22}=0$。

X \ Y	**-1**	**0**	**1**
-1	0	0.25	0
0	0.25	0	0.25
1	0	0.25	0

三. 连续型随机向量

- 设（X, Y）的联合分布函数为F(x，y)，联合密度函数为p(x，y)，X的边缘分布函数为$F_X(x)$，Y的边缘分布函数为$F_Y(y)$，则

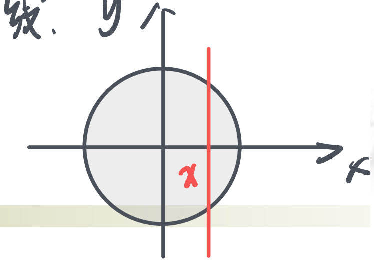
$$F_X(x)=F(x,+\infty)=\int_{-\infty}^{x}\left[\int_{-\infty}^{+\infty}p(u,v)dv\right]du$$

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

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

为的X,Y**边缘概率密度函数**。

① 画图 ② 画线.



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

- **例3.4.3：** 设（X, Y）在单位圆G={(x, y)|$x^2+y^2\leqslant 1$}上服从均匀分布，试求X和Y的边缘密度函数。

$$p_X(x)=F_X'(x)=\int_{-\infty}^{+\infty}p(x,y)\,dy=\begin{cases}\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}}\frac{1}{\pi}\,dy, & -1\le x\le 1\\ 0, & 其它\end{cases}$$

$$=\frac{2\sqrt{1-x^2}}{\pi},\quad -1\le x\le 1$$

- **例3.4.4：** 若二维随机向量（X, Y）服从二元正态分布N($\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho$)，试求X和Y的边缘密度函数。

$$p_Y(y)=\int_{-\infty}^{+\infty}p(x,y)\,dy=\begin{cases}\int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}}\frac{1}{\pi}\,dx, & -1\le y\le 1.\\ 0, & 其它\end{cases}$$

$$=\frac{2\sqrt{1-y^2}}{\pi},\quad -1\le y\le 1$$

- **解:**

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

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

$$=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\underline{(u-\rho v)^2}+(1-\rho^2)v^2\right]\right\}$$

面前面

$$p_Y(y)=\int_{-\infty}^{+\infty}p(x,y)dx=\int_{-\infty}^{+\infty}\frac{1}{2\pi\sigma_2\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2(1-\rho^2)}\left[(u-\rho v)^2+(1-\rho^2)v^2\right]\right\}du$$

$$=\frac{1}{\sqrt{2\pi}\sigma_2}e^{-\frac{v^2}{2}}\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2(1-\rho^2)}\left[(u-\rho v)^2\right]\right\}du=\frac{1}{\sqrt{2\pi}\sigma_2}e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}}$$

- 于是X~ N(μ_1,σ_1^2)，Y~N(μ_2,σ_2^2)。

- **例3.4.5**：若二维随机向量(X,Y)服从平面区域D上的均匀分布，其中D为$y=x^2$与$y=x$围成的有界区域，试求X和Y的边缘密度函数。
- **解：**

$$p_X(x)=\int_{-\infty}^{+\infty}p(x,y)dy=\begin{cases}\int_{x^2}^{x}6dy=6x-6x^2 & 0\le x\le 1\\ 0 & \text{其它}\end{cases}$$

$$p_Y(y)=\int_{-\infty}^{+\infty}p(x,y)dx=\begin{cases}\int_{y}^{\sqrt{y}}6dx=6\sqrt{y}-6y & 0\le y\le 1\\ 0 & \text{其它}\end{cases}$$

§ 3.5 条件分布

- 设X的分布函数为F(x)，则在A(假设P(A)>0)发生的条件下X的条件分布函数为：

$$F(x \mid A) = P(X \le x \mid A)$$

- **例3.5.1：** 设随机变量X的密度函数为

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

求X>1/2条件下X的条件分布函数。

一.离散型随机变量的条件分布

- 设二维随机向量（X, Y）的联合分布列为

$$p_{ij}=P（X=x_i,\ Y=y_j）i\ ,\ j=1,2,\ldots$$

记

$$p_{i.}=\sum_{j=1}^{+\infty}p_{ij}\qquad p_{.j}=\sum_{i=1}^{+\infty}p_{ij}$$

则

$$P\{X=x_i\mid Y=y_j\}=\frac{P\{X=x_i,Y=y_j\}}{P\{Y=y_j\}}=\frac{p_{ij}}{P_{\cdot j}}$$

$$P\{Y=y_j\mid X=x_i\}=\frac{P\{X=x_i,Y=y_j\}}{P\{X=x_i\}}=\frac{P_{ij}}{P_{i\cdot}}$$

分别称为在$Y=y_j$条件下X的**条件分布**及在$X=x_i$条件下Y的**条件分布**.

- **例3.5.2**: 设二维随机向量（X，Y）服从三项分布 $M(n; p_1,p_2,p_3)$，求在给定 $Y= k_2$ 条件下X的条件分布。
- **解:**

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

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

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

$$=\frac{n!}{k_2!\left(n-k_2\right)!} p_2^{k_2}(1-p_2)^{n-k_2}$$

- 在给定Y = k_2的条件下X的条件分布

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

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

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

- 可见此时X的条件分布为为B(n- k_2, $p_1/(1-p_2)$)

二. 连续型随机变量的条件密度

- **定义**：设（X, Y）的联合密度函数为p(x，y)，若对于任意x，极限

$$\lim_{\varepsilon \to 0} P\{X \le x \mid y < Y \le y+\varepsilon\}$$

存在，则称此极限为在条件Y=y下X的**条件分布函数**，记为 $F_{X|Y=y}(x)$ 。

若存在 $p_{X|Y=y}(x) \ge 0$ ，使得

$$F_{X|Y=y}(x) = \int_{-\infty}^{x} p_{X|Y=y}(x)dx$$

则称$p_{X|Y=y}(x)$为在条件Y = y下X的条件密度函数。

- **定理：** 设(X, Y)是二维连续型随机向量，p(x, y)为它的联合密度函数，$p_X(x)$和$p_Y(y)$分别为X和Y的密度函数。若p(x,y)在点(x,y)处连续， $p_Y(y)$在y处连续，且$p_Y(y)>0$，则

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

类似地可以证明

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

- 证明：

$$F_{Y|X=x}(y)=\lim_{\varepsilon\to 0}\frac{P(x<X\le x+\varepsilon,Y\le y)}{P(x<X\le x+\varepsilon)}=\lim_{\varepsilon\to 0}\frac{F(x+\varepsilon,y)-F(x,y)}{F_X(x+\varepsilon)-F_X(x)}$$

$$=\frac{\lim\limits_{\varepsilon\to 0}\dfrac{F(x+\varepsilon,y)-F(x,y)}{\varepsilon}}{\lim\limits_{\varepsilon\to 0}\dfrac{F_X(x+\varepsilon)-F_X(x)}{\varepsilon}}=\frac{\lim\limits_{\varepsilon\to 0}\dfrac{\int_x^{x+\varepsilon}(\int_{-\infty}^{y}p(u,v)dv)du}{\varepsilon}}{F_X\,'(x)}$$

$$=\frac{\int_{-\infty}^{y}p(x,v)dv}{p_X(x)}=\int_{-\infty}^{y}\frac{p(x,v)}{p_X(x)}dv\Rightarrow p_{Y|X=x}(y)=\frac{p(x,y)}{p_X(x)}$$

- **例3.5.3**: 设二维随机向量（X，Y）的密度函数为

$$p(x,y)=\begin{cases}\dfrac{15}{2}x(2-x-y), & 0\le x\le 1, 0\le y\le 1\\ 0, & \text{其它}\end{cases}$$

求X在Y=y条件下的条件密度。

- 解：当$0<y<1$时，

$$p_{X|Y=y}(x)=\frac{p(x,y)}{p_Y(y)}=\begin{cases}\dfrac{6x(2-x-y)}{4-3y} & 0<x<1\\ 0 & \text{其他}\end{cases}$$

- **例3.5.4：** 多元正态分布的条件分布。
- 解：

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

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

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

- 于是Y|X=x~ $N(\mu_2+\rho(x-\mu_1)\dfrac{\sigma_2}{\sigma_1},\sigma_2^2(1-\rho^2))$

- **例3.5.5**：设随机变量X在(0,1)上随机取值，在X=x的条件下，Y在区间(0，x)上随机取值，1）求Y的密度函数，2）求概率$P(X^2+Y^2<1|X=x),0<x<1$
- **解：**

$$1.\ p(x,y)=p_X(x)p_{Y|X=x}(y)=\begin{cases}1/x & 0<y<x<1\\ 0 & \text{其他}\end{cases},$$

$$p_Y(y)=\int_{-\infty}^{+\infty}p(x,y)dx=\begin{cases}-\ln y & 0<y<1\\ 0 & \text{其他}\end{cases}$$

$$2.\ P(X^2+Y^2<1|X=x)=P(Y^2<1-X^2|X=x)$$

$$=P(0<Y<\sqrt{1-X^2}|X=x)$$

$$=\int_0^{\sqrt{1-x^2}}p_{Y|X=x}(y)dy=\int_0^{\min(x,\sqrt{1-x^2})}\frac{1}{x}dy$$

$$=\frac{1}{x}\min(x,\sqrt{1-x^2})=\min(1,\sqrt{1/x^2-1})$$

§3.6 随机变量的独立性

- 对于两个随机变量X和Y，其中任意一个随机变量取值对另一个随机变量没有什么影响，这时就称这两个随机变量相互独立。

- 当然，这只是对随机变量独立性的一种直观描述。严格地讲，若对任意事件A，B，事件{ $X\in A$ }与事件{ $Y\in B$ }相互独立，则称X和Y**相互独立。**

随机变量相互独立的定义

- **定义:** 设二维随机向量(X, Y)的联合分布函数为F (x, y)，X和Y的分布函数分别为F_X (x)和F_Y (y)，如果对任意实数x, y，有

$$F(x, y) = F_X(x) F_Y(y)$$

成立，则称**X与Y相互独立**，简称为**X与Y独立**。

- **注1:**

$$\forall x, y, F(x,y) = F_X(x)F_Y(y) \Leftrightarrow \forall A, B, P(X \in A, Y \in B) = P(X \in A)P(Y \in B)$$

- **注2**：若X与Y相互独立，则f(X)与g(Y)也相互独立。

离散型和连续型情况：

- （1）如果(X, Y)是二维离散型随机向量，其联合分布列 为$\{ p_{ij} \}$ i , j=1,2,…，那么X与Y独立的充要条件是

$$P(X=x_i，Y=y_j)= P(X=x_i) P(Y=y_j) \quad i , j=1,2,\ldots$$

即

$$p_{ij} = p_{i.}\, p_{.j} \qquad i , j=1,2,\ldots$$

- （2）如果(X, Y)是二维连续型随机向量，p(x, y)为它的联合密度函数，$p_X (x)$和$p_Y (y)$分别为X和Y的密度函数，那么X与Y独立的充要条件是

$$p(x, y)= p_X (x)\, p_Y (y)$$

- **例3.6.1**　若二维随机向量（X，Y）在单位圆上服从均匀分布，问X与Y是否相互独立?

- **例3.6.2**　若二维随机向量（X，Y）服从二元正态分布$N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$，则X与Y相互独立的充要条件是$\rho=0$。

- **例3.6.3**　若二维随机向量（X，Y）服从

$$p(x,y)=\begin{cases} 6x, & 0<x<y<1 \\ 0, & \text{其它}. \end{cases}$$

试问X,Y是否独立?

- 解:

$$p_X(x)=\int_{-\infty}^{+\infty}p(x,y)dy=\begin{cases}\int_x^1 6xdy=6x-6x^2 & 0\le x\le 1\\ 0 & \text{其它}\end{cases}$$

$$p_Y(y)=\int_{-\infty}^{+\infty}p(x,y)dx=\begin{cases}\int_0^y 6xdx=3y^2 & 0\le y\le 1\\ 0 & \text{其它}\end{cases}$$

- 所以X和Y不独立。

- **例3.6.4：** 设二维连续型随机向量(X, Y)的联合分布函数为
  $F(x,y)=A(B+\arctan(x/2))(C+\arctan(y/3))$,试求:
  (1) 系数A, B及C;
  (2) (X,Y)的联合密度函数;
  (3) X,Y的边缘分布函数及边缘密度函数;
  (4) X和Y是否独立? 说明理由。

- 解：1. $F(x,-\infty)=A(B+\arctan(x/2))(C-\pi/2)=0$

$$F(-\infty,y)=A(B-\pi/2)(C+\arctan(y/3))=0$$

$$F(+\infty,+\infty)=A(B+\pi/2)(C+\pi/2)=1$$

$$\therefore A=1/\pi^2, B=C=\pi/2$$

- 2. $p(x,y)=\dfrac{\partial^2 F(x,y)}{\partial x\partial y}=\dfrac{6}{\pi^2(4+x^2)(9+y^2)}$

- 3. $F_X(x)=F(x,+\infty)=\dfrac{1}{\pi}(\dfrac{\pi}{2}+\arctan\dfrac{x}{2}),\therefore p_X(x)=\dfrac{2}{\pi(4+x^2)}$

$$F_Y(y)=F(+\infty,y)=\frac{1}{\pi}(\frac{\pi}{2}+\arctan\frac{y}{3}),\therefore p_Y(y)=\frac{3}{\pi(9+y^2)}$$

§ 3.7 随机向量函数的分布

已知二维随机向量(X, Y)的联合分布，如何求(X, Y)的函数g (X , Y)的分布？

设二维连续型随机向量(X, Y)的联合密度函数为p(x , y)，要求Z= g (X , Y)的密度函数，这类问题的**基本方法**是先求Z的分布函数，然后再通过求导得到相应的密度函数。

根据定义，Z= g (X , Y)的分布函数为

$$F(z)=P(g(X,Y)\le z)=\iint\limits_{g(x,y)\le z} p(x,y)dxdy$$

再通过对z求导即可得Z= g (X , Y)的密度函数。

- **例3.7.1**：设随机变量X,Y相互独立，服从参数为λ_1 ，λ_2的泊松分布，求Z=X+Y的分布。
- 解：

$$P(Z=n)=P(X+Y=n)=P\left(\bigcup_{k=0}^{n}(X=k,\ Y=n-k)\right)$$

$$=\sum_{k=0}^{n}P(X=k)P(Y=n-k)=\sum_{k=0}^{n}\frac{\lambda_1^k}{k!}e^{-\lambda_1}\frac{\lambda_2^{n-k}}{(n-k)!}e^{-\lambda_2}$$

$$=\frac{e^{-(\lambda_1+\lambda_2)}}{n!}\sum_{k=0}^{n}\frac{n!}{k!(n-k)!}\lambda_1^k\lambda_2^{n-k}=\frac{e^{-(\lambda_1+\lambda_2)}}{n!}(\lambda_1+\lambda_2)^n \quad .$$

- **注**：二项分布也具有分布可加性：若X~B(n_1, p)，Y~B(n_2, p)且X，Y独立，则Z=X+Y~B(n_1+n_2, p)。

和与差的分布

- 设(X, Y)的密度函数为p(x , y)，求Z=X+Y的密度函数。

$$F_Z(z) = P(X+Y \le z) = \iint\limits_{x+y \le z} p(x,y)dxdy$$

$$= \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{z-x} p(x,y)dy\right)dx \underline{\underline{令u=y+x}} \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{z} p(x,u-x)du\right)dx$$

$$= \int_{-\infty}^{z} \left(\int_{-\infty}^{+\infty} p(x,u-x)dx\right)du$$

- 求导得:

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

- 特别的，若X，Y相互独立，$p_X(x)$，$p_Y(y)$分别为它们的密度函数。则上述公式可表示为：

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

- 类似的，差Z=X-Y的卷积公式为：

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

- **例3.7.2**：设设随机变量X,Y相互独立，服从N(0,1)，求X+Y的分布。
- **解：**

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

$$=\frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-x^2/2}\cdot e^{-(z-x)^2/2}dx=\frac{1}{2\pi}e^{-z^2/4}\int_{-\infty}^{+\infty} e^{-(x-z/2)^2}dx$$

$$=\frac{1}{2\sqrt{\pi}}e^{-z^2/4}\int_{-\infty}^{+\infty}\frac{1}{\sqrt{\pi}}e^{-(x-z/2)^2}dx=\frac{1}{2\sqrt{\pi}}e^{-z^2/4},$$

- 若X,Y相互独立, $X\sim N(\mu_1,\sigma_1^2),Y\sim N(\mu_2,\sigma_2^2)$

则$X+Y\sim N(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2)$

- **例3.7.3：** 设随机变量X与Y相互独立，且X～U[0,1]，Y～E(1)，求Z=2X+Y的概率密度。
- **解：** 2X～U[0,2]，运用卷积公式

$$Z = 2X + Y : p_Z(z) = \int_{-\infty}^{+\infty} p_{2X}(x) p_Y(z-x)dx = \int_0^2 \frac{1}{2} p_Y(z-x)dx$$

当$0 < z < 2$时，$p_Z(z) = \int_0^z \frac{1}{2} e^{-(z-x)}dx = \frac{1}{2}(1-e^{-z})$

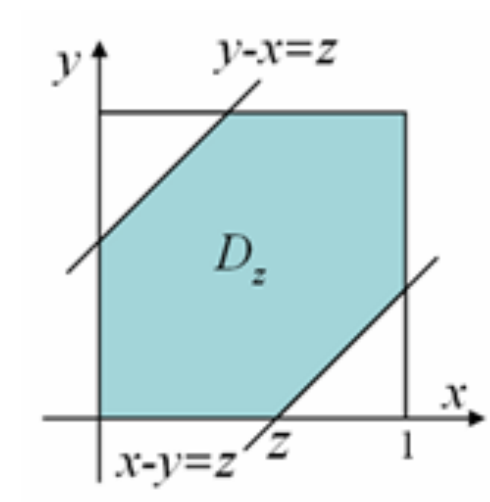
当$z > 2$时，$p_Z(z) = \int_0^2 \frac{1}{2} e^{-(z-x)}dx = \frac{1}{2}(e^2-1)e^{-z}$

- **例3.7.4**：设设随机变量X,Y相互独立，服从[0,1]上的均匀分布，求Z=|X-Y|的密度函数。
- **解**：

$$F_Z(z)=P(Z\le z)=P(|X-Y|\le z)=\iint\limits_{|x-y|\le z} p(x,y)dxdy$$

$$(\text{当}0<z<1\text{时})=\iint\limits_{|x-y|\le z} dxdy=S_{D_z}=1-(1-z)^2$$

$$p_z(z)=\begin{cases}2-z, & 0<z<1\\ 0, & \text{其它}.\end{cases}$$



最大值与最小值的分布

- 设X 与Y是两个相互独立的随机变量，它们的分布函数分别为$F_X(x)$和$F_Y(y)$，密度函数分别为$f_X(x)$和$f_Y(y)$。求M=max{X,Y}和N=min{X,Y}的分布函数。

$$F_M(z)=P(M\le z)=P(X\le z,\ Y\le z)=P(X\le z)P(Y\le z)$$

- M的分布函数为

$$F_M(z)=F_X(z)\,F_Y(z)$$

- N的分布函数为

$$F_N(z)=1-(1-F_X(z))(1-F_Y(z))$$

M和N的密度函数可通过对分布函数求导数得到。

- 以上结果可以推广到n个独立随机变量的情形：
- 设随机向量$(X_1, X_2, \ldots, X_n)$相互独立， X_i的分布函数为$F_i(x)$，$i=1,2,\ldots,n$。 求$M=\max\{X_1, X_2, \ldots, X_n\}$， $N=\min\{X_1, X_2, \ldots, X_n\}$的分布函数。
- M的分布函数为

$$F_M(z)= F_1(z)\, F_2(z)\ldots F_n(z)$$

- N的分布函数为

$$F_N(z)=1- [1- F_1(z)][(1- F_2(z)]\ldots [(1-F_n(z)]$$

- **例3.7.5**：设随机向量$(X_1, X_2, \ldots, X_n)$相互独立，服从[0,1]上的均匀分布，求$X_{(n)}=\max\{X_1, X_2, \ldots, X_n\}$，$X_{(1)}=\min\{X_1, X_2, \ldots, X_n\}$的密度函数。
- **解：**

$$p_{\max}(z)=\begin{cases} nz^{n-1}, & 0<z<1 \\ 0, & \text{其它}. \end{cases}$$

$$p_{\min}(z)=\begin{cases} n(1-z)^{n-1}, & 0<z<1 \\ 0, & \text{其它}. \end{cases}$$

- **例3.7.6**：设随机变量X,Y独立，均服从标准正态分布，求 $Z=\sqrt{X^2+Y^2}$ 的密度函数。
- **解**：

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

$$=\iint\limits_{\{(x,y)|x^2+y^2\le z^2\}}\frac{1}{2\pi}\exp\{-\frac{x^2+y^2}{2}\}dxdy$$

$$=\int_0^{2\pi}\int_0^{z}\frac{1}{2\pi}e^{-r^2/2}rdrd\theta=1-e^{-z^2/2}$$

$$p_Z(z)=F_Z{}'(z)=\begin{cases} ze^{-z^2/2} & z\ge 0 \\ 0 & z<0 \end{cases}$$

- **例3.7.7**：设随机变量X,Y独立，P(X=k)=1/3，k=1，2，3。Y服从指数分布E(1)，求Z=X+Y的密度函数。
- **解**：

$$F_z(z)=P(Z\le z)=P(X+Y\le z)$$

$$=\sum_{k=1}^{3}P(X=k)P(X+Y\le z\mid X=k)$$

$$=\frac{1}{3}\sum_{k=1}^{3}P(Y\le z-k)=\frac{1}{3}\sum_{k=1}^{3}F_Y(z-k)$$

$$p_Z(z) = F_Z'(z) = \frac{1}{3}[p_Y(z-1) + p_Y(z-2) + p_Y(z-3)]$$

$$= \begin{cases} 0 & z \le 1 \\ e^{1-z}/3 & 1 \le z < 2 \\ (e^{1-z} + e^{2-z})/3 & 2 \le z < 3 \\ (e^{1-z} + e^{2-z} + e^{3-z})/3 & z \ge 3 \end{cases}$$