概率论部分习题讲解

王子豪

全概率公式与贝叶斯公式

全概率公式： $P(A)=\sum_{i=1}^{n}P(B_i)P(A\mid B_i).$

利用B_i对A事件进行一个划分，将A划分成i个互不相容事件，以便于计算。

贝叶斯公式： $P(A|B)=\frac{P(B|A)P(A)}{\sum_{i=1}^{n}P(B|A_i)P(A_i)}$

在已知P(B|A)的情况下需要计算P(A|B)时会使用到，因此又叫做逆概率公式。

已知100件产品中有10件绝对可靠的正品，每次使用时不会发生故障，其余的产品每次使用时有0.1的可能性发生故障。今从这100件产品中任取一件，若使用了n次均未发生故障，问n为多大时才能有70%以上的把握认为所取产品为绝对可靠的产品。

A　　B

设 A="任取一件产品是绝对可靠的"

B="任取一件产品，使用n次未出故障"

由题意有

$$P(A|B) \geq 0.70$$

而

$$P(A|B) = \frac{P(A)P(B|A)}{P(A)P(B|A) + P(\bar{A})P(B|\bar{A})}$$

$$= \frac{0.1}{0.1 + 0.9 \times (0.9)^n}$$

$$\frac{P(AB)}{P(B)} = P(A|B)$$

即

$$\frac{1}{1 + 9 \times 0.9^n} \geq 0.7$$

$$0.9^n \leq \frac{1}{21}$$

$$n \geq -\frac{\ln 21}{\ln 0.9}$$

一维情形下随机游走问题

u_{a-1}

u_{a+1}

在 x 轴上有一个质点可以在整个数轴的整数点上游动，记 X_n 表示时刻 n 时质点的位置。该质点移动的规则是：每隔单位时间，分别以概率 p 及概率 $q=1-p$ $(0 < p < 1)$ 向正的及负的方向移动一个单位。假设质点在时刻 $t=0$ 时，位于 a，即 $X_0=a$ $(a>0)$，而在 0 和 $a+b$ $(b>0)$ 处各有一个吸收壁（即质点移动到 0 和 $a+b$ 时，将不能再移动）。求质点的初始位置为 a 而最终在 $a+b$ 被吸收的概率 u_a（$u_0=0$， $u_{a+b}=1$）.

分析：我们通过简单的分析可以产生以下几个认识，首先从a到a+b被吸收的情况有无限种是无法直接计算的。同时我们发现由于质点每次只能走一个单位，故其下一次游走的状态实际上只有两种，前一个点与后一个点。这样就把u_a与u_{a+1}和u_{a-1}用全概率公式联系了起来，得到了递推式：

$$u_n = pu_{n+1} + qu_{n-1}$$

0 a a+1 a+b

k 1-k

$a u_n - b u_{n-1} = H(a u_{n+1} - b u_n)$

$$(p+q)q_n = pq_{n+1} + qq_{n-1}, n = 1,2,\Lambda, a+b-1$$

$$\therefore p(q_{n+1} - q_n) = q(q_n - q_{n-1})$$

$$\therefore q_{n+1} - q_n = c_n = \frac{q}{p}(q_n - q_{n-1}) = (\frac{q}{p})^2(q_{n-1} - q_{n-2})$$

$$\Lambda \qquad = (\frac{q}{p})^n(q_1 - q_0) \hat{=} r^n c_0$$

$f_2(x)$ — $f_1(x)$

$f_{n+1}(x)$ — $f_n(x)$

f_{n-1} — f_1

当 $r=1$ 时，$p = q = \frac{1}{2}$，亦称为对称的随机游动的场合，此时 $c_n = c_{n-1}$，因此，

$$q_{n+1} - q_n = q_n - q_{n-1} = \Lambda = q_1 - q_0 \hat{=} d$$

$a+b$

则 $q_n = q_0 + nd$，而 $q_0 = 0, q_{a+b} = 1$，$\therefore q_n = \frac{n}{a+b}$，特别地，$q_a = \frac{a}{a+b}$

$a+b$

当 $r \neq 1$ 时，$p \neq q$ 的场合

这时 $c_n = rc_{n-1} = \Lambda = r^n c_0$，从而，

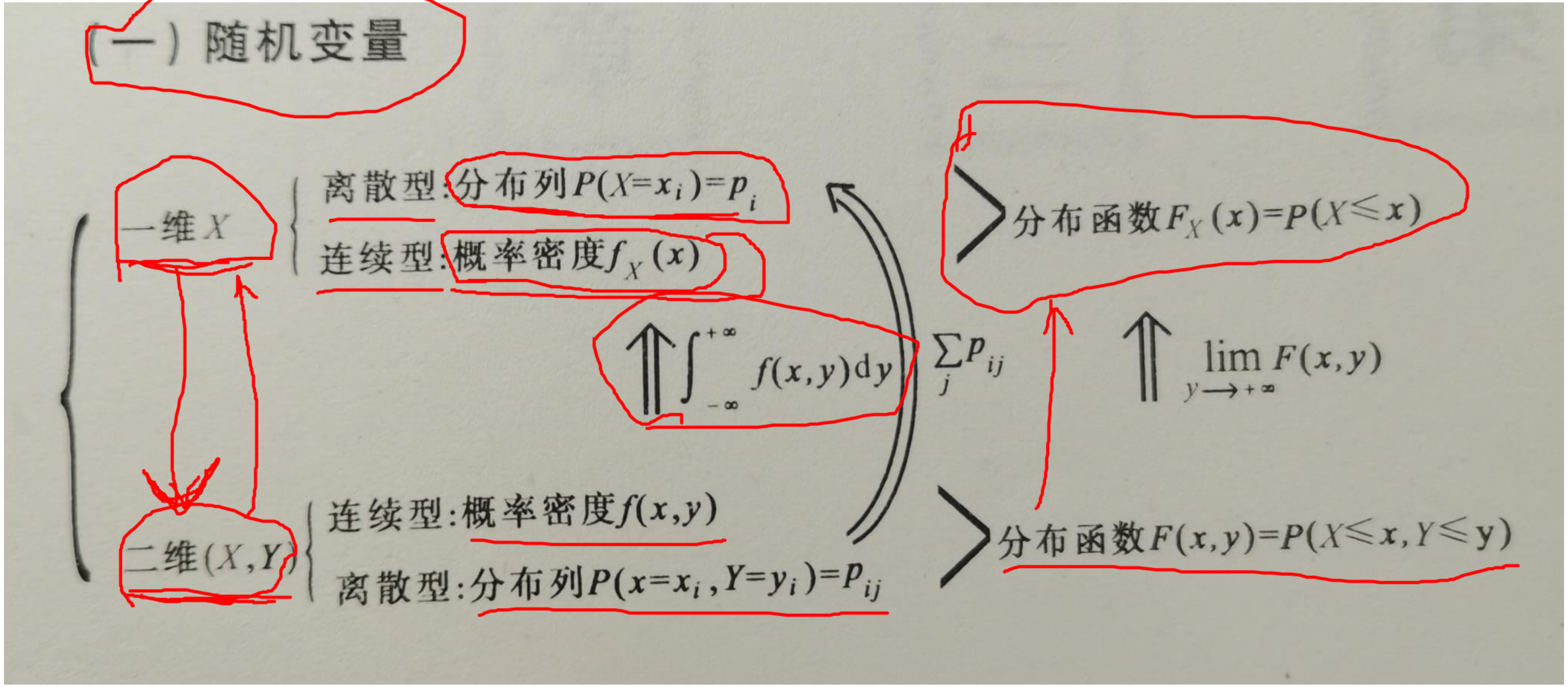
$$q_n - q_0 = \sum_{k=0}^{n-1}(q_{k+1} - q_k) = \sum_{k=0}^{n-1} c_k = \sum_{k=0}^{n-1} r^k c_0 = \frac{1-r^n}{1-r} c_0$$

由于

$$q_0 = 0, q_{a+b} = 1$$

所以 $\frac{1-r^{a+b}}{1-r} = 1, c_0 = \frac{1-r}{1-r^{a+b}}$，因此 $q_n = \frac{1-r^n}{1-r^{a+b}}$

随机变量及其分布


(一) 随机变量
一维 X
离散型：分布列 $P(X=x_i)=p_i$
连续型：概率密度 $f_X(x)$
分布函数 $F_X(x)=P(X\leq x)$
$\int_{-\infty}^{+\infty} f(x,y)\mathrm{d}y$
$\sum_j p_{ij}$
$\lim_{y\to+\infty} F(x,y)$
二维 (X,Y)
连续型：概率密度 $f(x,y)$
离散型：分布列 $P(x=x_i,Y=y_i)=p_{ij}$
分布函数 $F(x,y)=P(X\leq x,Y\leq y)$

随机变量函数的分布往往是考试的重要内容，其中一维随机变量与二维随机变量都有自己的公式法。但是二维随机变量函数的分布确定比一维随机变量更为麻烦些，我们这里主要讲解二维随机变量函数的分布题型。

(X,Y) 的概率密度为 $f(x,y)=\begin{cases}\lambda^2 e^{-\lambda x}, & 0<y<x \\ 0 & 其他\end{cases}$

$\lambda>0$，求 $Z=X+Y$ 的概率密度 $f_Z(z)$

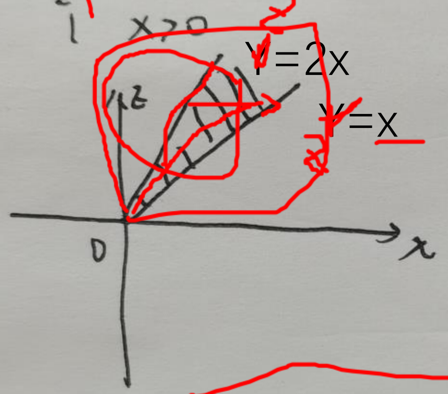
四、解：变量代换　公式法

$Y = Z - X$

$$\begin{cases} 0 < z - x < x \\ x > 0 \end{cases}$$

解得

$$\begin{cases} \frac{z}{2} < x < z \\ x > 0 \end{cases}$$



$$f_Z(z) = \int_{\frac{z}{2}}^{z} f(x, z-x)\,dx$$

$$= \int_{\frac{z}{2}}^{z} \lambda^2 e^{-\lambda x}\,dx$$

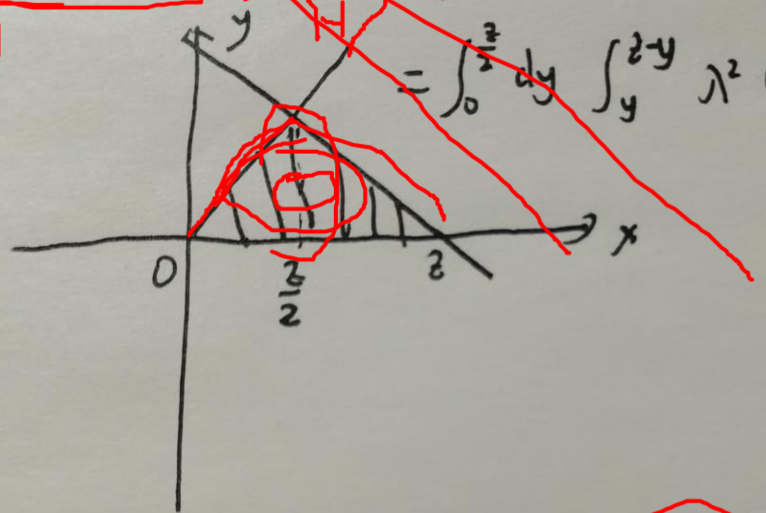
$$= -\lambda\left(e^{-\lambda z} - e^{-\frac{\lambda}{2}z}\right) \quad (z>0)$$

$$f_Z(z) = \begin{cases} -\lambda e^{-\lambda z} + \lambda e^{-\frac{\lambda z}{2}} & (z>0) \\ 0 & z \le 0 \end{cases}$$

分布函数法：

$$P(Z \le z) = P(X+Y \le z) = \iint_{x+y\le z} \lambda^2 e^{-\lambda x}\,d\sigma$$

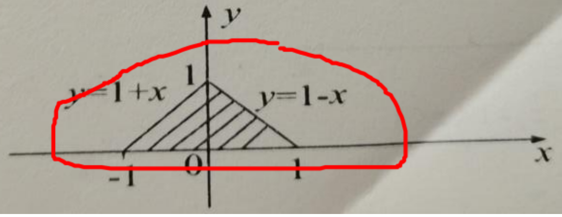
$$= \int_0^{\frac{z}{2}} dy \int_y^{z-y} \lambda^2 e^{-\lambda x}\,dx$$



$$原式 = -2e^{-\frac{\lambda z}{2}} + 1 + e^{-\lambda z} \quad z > 0$$

$$f(z) = \begin{cases} \lambda e^{-\frac{\lambda z}{2}} - \lambda e^{-\lambda z} & z > 0 \\ 0 & z \le 0 \end{cases}$$

【例 3.12】 设 (X,Y) 在如图 3.2 的区域 D 上服从均匀分布，求 $Z=X+Y$ 的概率密度.



$E(X)=\sum_{i=1}^{\infty} x_i p_i$

设 Z 的概率密度函数 $f_Z(z)=\int_{-\infty}^{+\infty} f(x,z-x)\,dx$

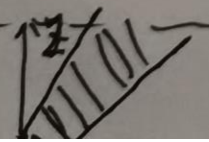
(X,Y) 在 D 上服从均匀分布，有

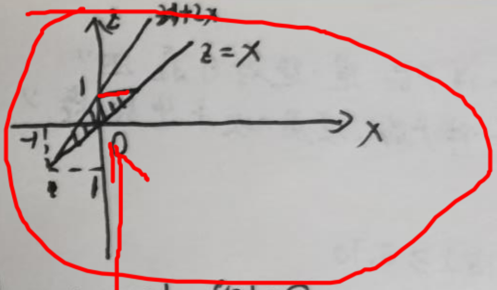
$$f(x,y)=\begin{cases}1 & 0\le y\le 1 \quad y-1\le x\le 1-y\\ 0 & \text{其他}\end{cases}$$

~~从而可得~~ 可得

$$f(x,z-x)=\begin{cases}1 & 0\le z-x\le 1 \quad z-x-1\le x\le 1-z+x\\ 0 & \text{其他}\end{cases}$$

$$\begin{cases}0\le z-x\le 1\\ z-x-1\le x\le 1-z+x\end{cases} \longrightarrow \begin{cases}\frac{z-1}{2}\le x\le z\\ z\le 1\end{cases}$$





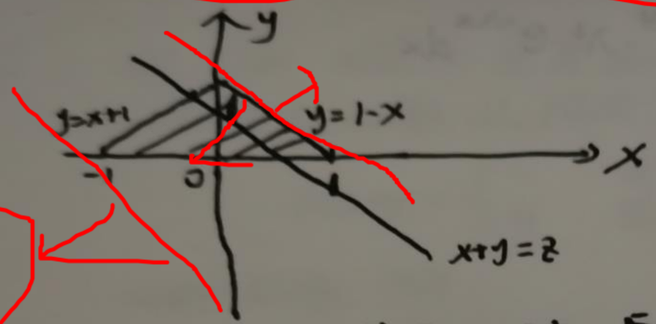
$|z|\ge 1$ 时 $f_Z(z)=0$

$|z|<1$ 时 $f_Z(z)=\int_{\frac{z-1}{2}}^{z} dx=\frac{z+1}{2}$

即 Z 的概率密度为

$$f_Z(z)=\begin{cases}\frac{z+1}{2} & |z|<1\\ 0 & |z|\ge 1\end{cases}$$

解2 $F_Z(z)=\iint\limits_{x+y\le z} f(x,y)\,dx\,dy$，记 $G=\{(x,y)\mid x+y\le z\}$



当 $z\le -1$ 时 ~~当 $z\ge 1$ 时~~ $F_Z(z)=0$

当 $z\ge 1$ 时 $F_Z(z)=1$

当 $-1<z<1$ 时

$$F_Z(z)=\int_0^{\frac{z+1}{2}} dy\int_{y-1}^{z-y} dx$$

$$=\int_0^{\frac{z+1}{2}}(z+1-2y)\,dy=\frac{(z+1)^2}{4}$$

$$F_Z(z)=\begin{cases}0 & z\le -1\\ \frac{(z+1)^2}{4} & -1<z<1\\ 1 & z\ge 1\end{cases}$$

$$f(z)=\begin{cases}0 & |z|\ge 1\\ \frac{z+1}{2} & |z|<1\end{cases}$$

(7) 设二维随机变量(X,Y)在矩形$G\{(x,y)\mid 0\leqslant x\leqslant 2, 0\leqslant y\leqslant 1\}$上服从均匀分布,试求边长为$X$和$Y$的矩形面积$S$的概率密度$f(s)$.

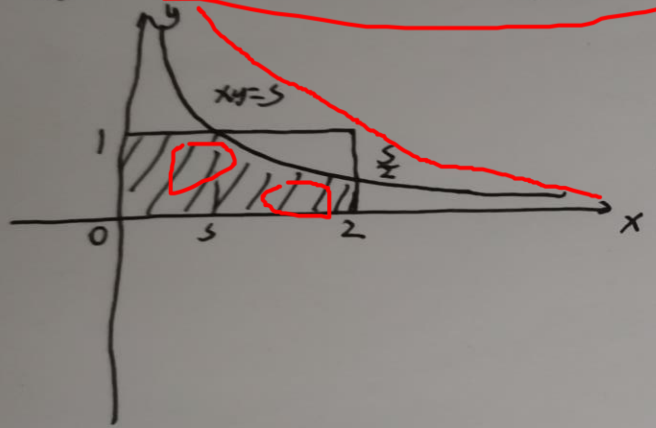
$\frac{X}{Y}$

解: $\varphi(x,y)$为(X,Y)的概率密度为

$$\varphi(x,y)=\begin{cases}\frac{1}{2} & (x,y)\in G\\ 0 & 其他\end{cases}$$

设S的分布函数$F(s)$

$$F(s)=P\{S\leq s\}=P(XY\leq s)=\iint_{xy\leq s}\varphi(x,y)\,dxdy$$



~~F(s)=f~~

当$s\geq 2$时 $F(s)=1$

当$s\leq 0$时 $F(s)=0$

$0<s<2$时

$$F(s)=\int_0^s dx\int_0^1\frac{1}{2}dy+\int_s^2 dx\int_0^{\frac{s}{x}}\frac{1}{2}dy$$

$$=\frac{s}{2}(1+\ln(2)-\ln(s))$$

$$f(s)=F'(s)=\begin{cases}\frac{1}{2}(\ln(2)-\ln(s)) & 0<s<2\\ 0 & 其他\end{cases}$$

解2 利用乘积的概率密度公式

$$f(s)=\int_{-\infty}^{+\infty}\varphi\left(\frac{s}{y},y\right)\frac{1}{|y|}dy$$

$$\begin{cases}0\leq\frac{s}{y}\leq 2\\ 0\leq y\leq 1\end{cases}\longrightarrow\begin{cases}\frac{s}{2}\leq y\leq 1\\ 2\geq s\geq 0\end{cases}$$

$$f(s)=\int_{\frac{s}{2}}^{1}\frac{1}{2}\cdot\frac{1}{y}dy$$

$$=\frac{1}{2}(\ln(2)-\ln(s))\quad(0<s<2)$$

$$f(s)=\begin{cases}\frac{1}{2}(\ln(2)-\ln(s)), & 0<s<2\\ 0, & 其他\end{cases}$$

X $Z=$ $Y=Z-X$

$aX+bY=Z$

$\frac{Z-bY}{a}$

$\frac{X}{Y}$

$|Y|$

书上的定理(公式法的来源)

定理：

定理 4.7.1 设(X,Y)的概率密度为$f(x,y)$，且区域A(可以是全平面)满足$P((X,Y)\in A)=1$，在变换

$$\begin{cases} u=g_1(x,y), \\ v=g_2(x,y) \end{cases} \tag{4.58}$$

中，当$(x,y)\in A$时，(u,v)的值域为G，且变换(4.58)满足

(1) 变换(4.58)是$A\to G$的一一对应，有逆变换$\begin{cases} x=h_1(u,v), \\ y=h_2(u,v). \end{cases}$

(2) 变换(4.58)的雅可比行列式$J(x,y)=\dfrac{\partial(u,v)}{\partial(x,y)}=\begin{vmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y} \\ \dfrac{\partial v}{\partial x} & \dfrac{\partial v}{\partial y} \end{vmatrix}$在$A$中处处不为0，则$U$和$V$的联合概率密度为

$$f_{(U,V)}(u,v)=\begin{cases} f(h_1(u,v),h_2(u,v))\left|\dfrac{\partial(x,y)}{\partial(u,v)}\right|, & (u,v)\in G, \\ 0, & \text{其他}. \end{cases} \tag{4.59}$$

其中$\left|\dfrac{\partial(x,y)}{\partial(u,v)}\right|$是函数$h_1(u,v),h_2(u,v)$的雅可比行列式的绝对值.

$x = x$

$y = \dfrac{z}{x}$

$\begin{vmatrix} 1 & 0 \\ -\dfrac{z}{x^2} & \dfrac{1}{x} \end{vmatrix}$

【例 3.16】 设(X,Y)的概率密度为

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

求：(1)$M=\max\{X,Y\}$的分布函数和概率密度；

(2)$N=\min\{X,Y\}$的分布函数和概率密度.

~~$X\leq n$~~ ~~$N\geq n$~~ $\underline{N\geq n}$

$X\geq n$ $Y\geq$

~~$N\geq n$~~ $N<n$ $\boxed{X\geq n\ \ Y\geq n}$

$X<n$ $Y<n$

$1-P(N\geq n)=\boxed{1-P(X\geq n,Y\geq n)}$

解：(1) 设 M 的分布函数为 $F_M(z)$，则

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

当 $z\le 0$ 时，$F_M(z)=0$；

当 $z>1$ 时，$F_M(z)=1$；

当 $0<z\le 1$ 时

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

$$=\iint\limits_{x\le z, y\le z} f(x,y)\,dx\,dy$$

$$=\int_0^z dx\int_0^z (x+y)\,dy=z^3$$

故 $M=\max\{X,Y\}$ 的分布函数为

$$F_M(z)=\begin{cases}0 & z\le 0\\ z^3 & 0<z\le 1\\ 1 & z>1\end{cases}$$

M 的概率密度为

$$f_M(z)=F'_M(z)=\begin{cases}3z^2 & 0<z\le 1\\ 0 & 其他\end{cases}$$

(2) 设 $N=\min\{X,Y\}$ 的分布函数为 $F_N(t)$ 则

$$F_N(t)=P(N\le t)=1-P(\min\{X,Y\}>t)$$

$$=1-P(X>t, Y>t)$$

当 $t\le 0$ 时，$F_N(t)=0$；

当 $t>1$ 时，$F_N(t)=1$；

当 $0<t\le 1$ 时，$F_N(t)=1-\int_t^1\int_t^1(x+y)\,dx\,dy=t+t^2-t^3$

所以 $N=\min\{X,Y\}$ 的分布函数为

$$F_N(t)=\begin{cases}0, & t\le 0\\ t^2+t-t^3 & 0<t\le 1\\ 1 & t>1\end{cases}$$

N 的概率密度为

$$f_N(t)=F'_N(t)=\begin{cases}1+2t-3t^2 & 0<t\le 1\\ 0 & 其他\end{cases}$$

总结

公式法：
1.将其中一个随机变量写成另外两个随机变量的函数形式，可直接套用相关公式也可以利用课本给出的变量代换公式。
2.对于二元随机变量函数产生的一维随机变量，由于原来的随机变量往往在特定区域U的概率密度不为0，故需要反解出变量代换后的积分域之后再求出边缘概率密度函数。对于二元随机变量函数产生的二维随机变量，仍然需要确定积分后的分段范围。
分布函数法：
1.首先求出分布函数，求出分布函数的过程中需要画图求出$Z\overset{<}{\cancel{>}}z$与原联合概率密度函数不为0的区域的交集区域，并写出分布函数在其他范围的取值。
2.在这个积分域进行积分求取分布函数。
3.对分布函数求导求出概率密度函数。

随机变量的数字特征

哈尔滨工业大学
HARBIN INSTITUTE OF TECHNOLOGY
邮编：150001

(收藏问题) 盒中有一套邮票，共N张。从中有放回地每次抽取一张，要收集k张不同的邮票，期望抽取多少次？

$$P(X_k = j) = p_k (1-p_k)^{j-1} \qquad j = 1, 2, \cdots$$

$$X = \sum_{i=1}^{N} X_i$$

$$E(X) = E\left(\sum_{i=1}^{N} X_i\right) = \sum_{i=1}^{N} E(X_i) = \sum_{i=1}^{n} \frac{1}{p_i}$$

$$\frac{1}{p_k} = \frac{N}{N-k+1} \qquad p_k = \frac{N-k+1}{N}$$

$$S_n = X_1 + X_2 + \cdots + X_n \quad \text{得}$$

$$= N\left(\sum_{j=1}^{N} \frac{1}{j} - \sum_{j=1}^{N-k} \frac{1}{j}\right)$$