第三章

第三节

多维随机变量函数的分布

$Z=g(X,Y)$ 的概率分布 — 离散型

例. 设随机变量(X,Y)的分布律为

Y \ X	0	1	2	3
0	0.1	0.05	0.01	0.12
1	0.04	0.06	0.07	0.08
2	0.13	0.08	0.11	0.15

试求: (1) $Z=2X-Y$的分布律;

(2) $M=\max\{X, Y\}$的分布律;

(3) $N=\min\{X, Y\}$的分布律.

一、离散型的情况

1. $Z=g(X,Y)$ 的分布律

(1) 确定Z的可能取值$z_{ij}=g(x_i, y_j)$ $i, j=1,2,\cdots$

(2) 确定$P\{Z=g(x_i,y_j)\}=P\{X=x_i,Y=y_j\}=p_{ij}$;

(3) 列出$Z=g(X,Y)$的分布律。

一、离散型的情况

2. 二项分布的可加性

若 $X\sim b(m, p)$, $Y\sim b(n, p)$, 且独立,

则 $Z=X+Y\sim b(m+n, p)$.

注意: 若 $X_i\sim b(1, p)$, 且独立, 则

$Z=X_1+X_2+\ldots\ldots+X_n\sim b(n, p)$.

一、离散型的情况

3. 泊松分布的可加性

若 $X\sim P(\lambda_1)$, $Y\sim P(\lambda_2)$, 且独立,

则 $Z=X+Y\sim P(\lambda_1+\lambda_2)$.

注意: $X-Y$ 不服从泊松分布.

二、连续型的情况

- 随机变量极值的分布
- 随机变量和的分布
- 变量变换法

二、连续型的情况

1. 随机变量极值的分布

设X,Y是两个相互独立的随机变量，

其概率密度为 $p_X(x), p_Y(y)$,

$$\left.\begin{matrix} M=\max\{X,Y\} \\ N=\min\{X,Y\} \end{matrix}\right\}\text{——统称为极值变量}$$

二、连续型的情况

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

$$\underline{\underline{ind}}\ P\{X\le z\}\cdot P\{Y\le z\}=F_X(z)F_Y(z)$$

$$F_N(z)=P\{N\le z\}=1-P\{N>z\}$$

$$=1-P\{X>z, Y>z\}$$

$$\underline{\underline{ind}}\ 1-P\{X>z\}\cdot P\{Y>z\}$$

$$=1-\left(1-F_X(z)\right)\left(1-F_Y(z)\right)$$

二、连续型的情况

推广：

$$X_1,\cdots,X_n\ \ ind. \Longrightarrow \begin{cases} F_M(z)=\prod_{i=1}^{n}F_{X_i}(z) \\ F_N(z)=1-\prod_{i=1}^{n}\left(1-F_{X_i}(z)\right) \end{cases}$$

$$X_1,\cdots,X_n\ \ i.i.d.\ F(x) \Longrightarrow \begin{cases} F_M(z)=\left[F(z)\right]^n \\ F_N(z)=1-\left[1-F(z)\right]^n \end{cases}$$

二、连续型的情况

课 堂 练 习.

有5个相互独立工作的电子装置,其寿命X_k

$(k=1,2,3,4,5)$ $i.i.d.\ Exp\left(\frac{1}{\theta}\right)$, $\theta>0$

(1)5个串联工作时,整机寿命N的$E(N)$；

(2)5个并联工作时,整机寿命M的$E(M)$.

解: $X_k\sim p(x)=\begin{cases}\frac{1}{\theta}e^{-\frac{x}{\theta}} & x>0\\ 0 & x\le 0\end{cases}$ $(\theta>0)$

$$F(x)=\begin{cases}1-e^{-\frac{x}{\theta}} & x>0\\ 0 & x\le 0\end{cases}\quad (k=1,2,3,4,5)$$

(1) $N=\min(X_1,X_2,\ldots,X_5)$

$$F_N(x)=1-\left[1-F(x)\right]^5=\begin{cases}1-e^{-\frac{5}{\theta}x} & x>0\\ 0 & x\le 0\end{cases}$$

$$\Rightarrow p_N(x)=\begin{cases}\frac{5}{\theta}e^{-\frac{5x}{\theta}} & x>0\\ 0 & \text{其它}\end{cases}\ \Rightarrow N\sim Z\left(\frac{5}{\theta}\right)$$

$$\Rightarrow E(N)=\frac{\theta}{5}$$

(2)$M=\max(X_1,X_2,\ldots,X_5)$

$$F_M(x)=\left[F(x)\right]^5=\begin{cases}\left(1-e^{-\frac{x}{\theta}}\right)^5 & x>0\\ 0 & x\le 0\end{cases}$$

$$\Rightarrow p_M(x)=\begin{cases}\frac{5}{\theta}\left[1-e^{-\frac{x}{\theta}}\right]^4 e^{-\frac{x}{\theta}} & x>0\\ 0 & x\le 0\end{cases}$$

$$\Rightarrow E(M)=\int_{-\infty}^{+\infty} x p_M(x)dx$$

$$=\int_0^{+\infty} x\cdot\frac{5}{\theta}\left(1-e^{-\frac{x}{\theta}}\right)^4 e^{-\frac{x}{\theta}}dx \quad \left\langle t=\frac{x}{\theta}\right\rangle$$

$$=5\theta\int_0^{+\infty} t\left(1-e^{-t}\right)^4 e^{-t}dt=\frac{137}{60}\theta$$

二、连续型的情况

2. 随机变量和的分布

若$(X,Y)\sim p(x,y)$, 则$Z=X+Y$的概率密度为

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

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

证明：$F_Z(z)=P\{Z\le z\}=P\{X+Y\le z\}$

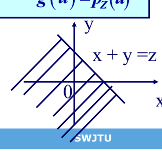
$$=\iint_{x+y\le z} p(x,y)dxdy=\int_{-\infty}^{+\infty}\left[\int_{-\infty}^{z-x} p(x,y)dy\right]dx$$

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

u 的一元函数 $g(u)=p_Z(u)$

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

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



当X与Y相互独立时，则

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

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

注：$p_X * p_Y=\int_{-\infty}^{+\infty} p_X(x)\,p_Y(z-x)\,dx$

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

称为卷积公式。

二、连续型的情况

例1. 设随机变量$(X，Y)$

$$p(x,y)=\begin{cases}\frac{1}{2}(x+y)e^{-(x+y)} & x>0,y>0\\ 0 & \text{其它}\end{cases}$$

试求：(1) $p_X(x),p_Y(y)$; (2) $Z=X+Y$的概率密度。

例2. 设随机变量X与Y相互独立，均服从均匀分布$U(-1,1)$,试求$Z=X+Y$的概率密度。

二、连续型的情况

例3. 设X和Y是两个相互独立的随机变量 它们都服从$N(0,1)$分布，试求Z=X+Y的概率分布。

一般地, 若X 与Y ind, 且$X\sim N(\mu_1,\sigma_1^2),Y\sim N(\mu_2,\sigma_2^2)$

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

推广. 有限个相互独立的正态随机变量的线性组合仍然正态分布。

课 堂 练 习

设随机变量X与Y相互独立，$X \sim U(-1,1)$，

$Y \sim Exp(2)$。试求$Z = X + Y$的概率密度。

课 堂 练 习

已知随机变量X与Y相互独立，其中

$P(X=-1)=0.4,\quad P(X=1)=0.6$

而随机变量$Y \sim N(\mu,\sigma^2)$。

试求$Z = X + Y$的分布。

二、连续型的情况

2. 随机变量和的分布

(1) 伽玛分布的可加性

若 $X \sim Ga(\alpha_1, \lambda)$，$Y \sim Ga(\alpha_2, \lambda)$，且相互独立，

则 $Z = X + Y \sim Ga(\alpha_1+\alpha_2, \lambda)$.

注意： $X - Y$ 不服从 $Ga(\alpha_1-\alpha_2, \lambda)$.

二、连续型的情况

2. 随机变量和的分布

(2) χ^2分布的可加性

若 $X \sim \chi^2(m)$，$Y \sim \chi^2(n)$，且独立，

则 $Z = X + Y \sim \chi^2(m+n)$.

注意： ① $X - Y$ 不服从 χ^2 分布.

② 若 $X_i \sim N(0,1)$，且独立，则

$$Z = \sum_{i=1}^{n} X_i^2 \sim \chi^2(n).$$

二、连续型的情况

2. 随机变量和的分布

(3) 两个结论

① 独立的0-1分布随机变量之和服从二项分布

② 独立同分布的指数分布随机变量之和服从伽玛分布.

$X_1, \cdots, X_n$ *i.i.d.* $Exp(\lambda) = Ga(1, \lambda)$

$\Longrightarrow Z = X_1 + \cdots + X_n \sim Ga(n, \lambda)$

二、连续型的情况

3. 变量变换法

已知 (X, Y) 的分布，(X, Y) 的函数

$$\begin{cases} U = g_1(X,Y) \\ V = g_2(X,Y) \end{cases}$$

求 (U, V) 的分布.

二、连续型的情况

若 $\begin{cases} u = g_1(x,y) \\ v = g_2(x,y) \end{cases}$ 有连续偏导、存在反函数

$\begin{cases} x = x(u,v) \\ y = y(u,v) \end{cases}$ 则 (U, V) 的联合密度为

$$p_{UV}(u,v) = p_{XY}(x(u,v), y(u,v))\,|J|$$

其中J为变换的雅可比行列式：$J = \dfrac{\partial(x,y)}{\partial(u,v)} = \left(\dfrac{\partial(u,v)}{\partial(x,y)}\right)^{-1}$

二、连续型的情况

4. 增补变量法

若要求 $U = g_1(X, Y)$ 的密度 $p_U(u)$，

可增补一个变量$V = g_2(X, Y)$，

先用变量变换法求出(U, V)的联合密度$p_{UV}(u, v)$,

然后再由联合密度$p_{UV}(u, v)$，去求出边际密度$p_U(u)$

用此方法可以求出卷积公式、积的公式、商的公式

二、连续型的情况

命题. （商的分布）

设(X,Y)的概率密度为$p(x,y)$，则$Z = \dfrac{X}{Y}$

的概率密度为

$$p(z) = \int_{-\infty}^{+\infty} p(yz, y)\,|y|\,dy \quad \text{—— 令} \begin{cases} Z = X/Y \\ V = Y \end{cases}$$

特别. 当X、Y独立时，

$$p(z) = \int_{-\infty}^{+\infty} p_X(yz)\,p_Y(y)\,|y|\,dy$$

用分布函数如何处理

设(X, Y)的概率密度为$p(x, y)$，则$Z = \dfrac{X}{Y}$的

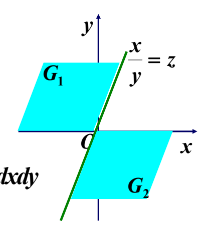
分布函数为

$$F_Z(z) = P\{Z \le z\} = P\left\{\frac{X}{Y} \le z\right\}$$

$$= \iint_{G_1} p(x,y)dxdy + \iint_{G_2} p(x,y)dxdy$$

$$= \int_0^{+\infty}\int_{-\infty}^{yz} p(x,y)dxdy + \int_{-\infty}^{0}\int_{yz}^{+\infty} p(x,y)dxdy$$

令$u = x/y$,



$$\iint_{G_1} p(x,y)dxdy = \int_0^{+\infty}\int_{-\infty}^{yz} p(x,y)dxdy$$

$$= \int_0^{+\infty}\int_{-\infty}^{z} yp(yu,y)dudy = \int_{-\infty}^{z}\int_0^{+\infty} yp(yu,y)dydu$$

同理可得

$$\iint_{G_2} p(x,y)dxdy = -\int_{-\infty}^{z}\int_{-\infty}^{0} yp(yu,y)dydu$$

故有 $F_Z(z) = P\{Z \le z\}$

$$= \iint_{G_1} p(x,y)dxdy + \iint_{G_2} p(x,y)dxdy$$

$$= \int_{-\infty}^{z}\left[\int_0^{+\infty} yp(yu,y)dy - \int_{-\infty}^{0} yp(yu,y)dy\right]du$$

由此可得分布密度为

$$p(z) = \int_0^{+\infty} yp(yz,y)dy - \int_{-\infty}^{0} yp(yz,y)dy$$

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