第三章 联合分布

§1 引言：联合累积分布函数

§2 (二维)离散随机变量

§3 (二维)连续随机变量

§4 独立随机变量

§5 条件分布

§6 联合分布随机变量函数

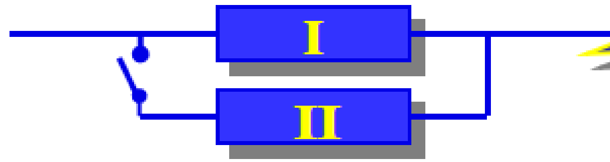
§7 极值和顺序统计量

(Extrema and Order Statistics)

实际背景

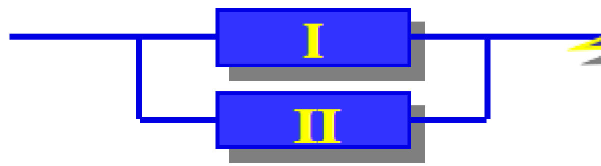
设有两个部件 I 、II, 其工作寿命分别为 X , Y

冷冗余系统： 部件 I 坏了, 换上备用部件 II 继续工作



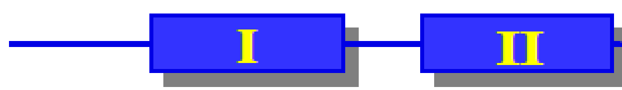
系统寿命 $X+Y$

热冗余系统： 部件I、II 并联同时工作, 仅当两个部件都损坏时，整个系统才失效

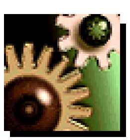


系统寿命 $\max\{X,Y\}$

串联系统： 部件I、II 串联同时工作, 只要有一个部件损坏，整个系统就失效



系统寿命 $\min\{X,Y\}$



问题 怎样确定上述各系统的寿命?

极值 max(X, Y), min(X, Y) 的分布

① 设 $X \sim F_X(x), Y \sim F_Y(y)$，且 X, Y 相互独立，则

$$\begin{aligned} F_{\max}(z) &= P\{\max(X,Y) \le z\} \\ &= P\{X \le z, Y \le z\} \\ &= P\{X \le z\} \cdot P\{Y \le z\} \\ &= F_X(z) \cdot F_Y(z) \end{aligned}$$

$$\begin{aligned} F_{\min}(z) &= P\{\min(X,Y) \le z\} \\ &= 1 - P\{\min(X,Y) > z\} \\ &= 1 - P\{X > z, Y > z\} \\ &= 1 - P\{X > z\} \cdot P\{Y > z\} \\ &= 1 - [1 - P\{X \le z\}] \cdot [1 - P\{Y \le z\}] \\ &= 1 - [1 - F_X(z)] \cdot [1 - F_Y(z)] \end{aligned}$$

极值 max(X, Y), min(X, Y) 的分布

① 设 $X \sim F_X(x), Y \sim F_Y(y)$ 且 X, Y 相互独立, 则

$$F_{\max}(z) = F_X(z) \cdot F_Y(z)$$

$$F_{\min}(z) = 1 - [1 - F_X(z)] \cdot [1 - F_Y(z)]$$

② 设 $X_i \sim F_{X_i}(x), i = 1, 2, \cdots, n,$ 且 $X_1, X_2, \cdots, X_n$ 相互独立,

则 $$F_{\max}(z) = P\{\max(X_1, X_2, \cdots, X_n) \le z\}$$
$$= F_{X_1}(z) F_{X_2}(z) \cdots F_{X_n}(z)$$

$$F_{\min}(z) = P\{\min(X_1, X_2, \cdots, X_n) \le z\}$$
$$= 1 - \prod_{i=1}^{n} [1 - F_{X_i}(z)]$$

③ 特别当 $X_1, X_2, \cdots, X_n$ 独立同分布于 $F(x)$ 时，有

$$F_{\max}(z) = F^n(z)$$

$$F_{\min}(z) = 1 - [1 - F(z)]^n$$

问题 question 设X,Y独立同分布，具有密度$f(x)$,怎样求 $\max(X,Y)$, $\min(X,Y)$ 的密度？

分析 $\because \quad F_{\max}(z) = F^2(z)$

$$\therefore \quad f_{\max}(z) = 2f(z)F(z) = 2f(z)\int_{-\infty}^{z} f(t)dt$$

$$\because \quad F_{\min}(z) = 1 - [1 - F(z)]^2$$

$$\therefore \quad f_{\min}(z) = 2f(z)[1 - F(z)] = 2f(z)[1 - \int_{-\infty}^{z} f(t)dt]$$

同理可推广，求出n个独立同分布的r.v.的极值密度为：

$$\boxed{f_{\max}(z) = nf(z)[F(z)]^{n-1}} \qquad \boxed{f_{\min}(z) = nf(z)[1 - F(z)]^{n-1}}$$

推广

设$X_1,\dots,X_n$是n个相互独立的随机变量，它们的分布函数分别为

$$F_{X_i}(x) \quad (i=0,1,\ \dots,n)$$

$M=\max(X_1,\dots,X_n)$的分布函数为:

$$F_M(z)=F_{X_1}(z)\cdots F_{X_n}(z)$$

N=min(X_1,…,X_n)的分布函数是

$$F_N(z) = 1 - [1 - F_{X_1}(z)] \cdots [1 - F_{X_n}(z)]$$

特别，当X_1,…,X_n相互独立且具有相同分布函数$F(x)$时，有

$$F_M(z)=[F(z)]^n$$

$$F_N(z)=1-[1-F(z)]^n$$

注意

若$X_1,\ldots,X_n$是连续型随机变量，在求得$M=\max(X_1,\ldots,X_n)$和$N=\min(X_1,\ldots,X_n)$的分布函数后，不难求得M和N的密度函数.

例 体育馆的大屏幕由信号处理机和显示屏构成，

它们的寿命分别为 X, Y，若它们的概率密度分别为

$$f_X(x) = \begin{cases} \alpha e^{-\alpha x}, x > 0 \\ 0 \ , \ x \le 0 \end{cases} \qquad f_Y(y) = \begin{cases} \beta e^{-\beta y}, y > 0 \\ 0 \ , \ y \le 0 \end{cases}$$

其中 $\alpha > 0, \beta > 0.$ 试求大屏幕系统的寿命 Z 的概率密度.

解 大屏幕系统寿命 $Z = \min(X, Y)$，由独立性有

$$F_{\min}(z) = 1 - [1 - F_X(z)][1 - F_Y(z)]$$

$$= \begin{cases} 1 - e^{-\alpha z} e^{-\beta z}, z > 0 \\ 0 \quad , \quad z \le 0 \end{cases}$$

$$\therefore \quad f_Z(z) = \begin{cases} (\alpha + \beta) e^{-(\alpha+\beta) z}, z > 0 \\ 0 \ , \qquad z \le 0 \end{cases}$$

可见指数分布的串联系统仍服从指数分布

其失效率是每个部件的失效率之和

例 设某种电子管的寿命（以天计）近似服从 $N(1195, 15^2)$，随机地选取3只，求
(1) 其中没有一只寿命超过1210天的概率;
(2) 其中没有一只寿命小于1210天的概率.

解：设X_1，X_2，X_3为3只电子管的寿命，
它们相互独立同分布，

$X_i \sim N(1195, 15^2)$ 分布函数为 $F(x)$

记 $M = \max\{X_1, X_2, X_3\}$,

$N = \min\{X_1, X_2, X_3\}$.

$$F_M(z)=[F(z)]^n$$

$$P(M \le 1210) = F_M(1210) = [F(1210)]^3$$

$$= \left[\Phi\left(\frac{1210-1195}{15}\right)\right]^3 = [\Phi(1)]^3$$

$$= 0.5955.$$

$$F(1210) = \Phi\left(\frac{1210-1195}{15}\right) = \Phi(1)$$

$$F_N(z)=1-[1-F(z)]^n$$

$$P(N\geq 1210)=1-P(N<1210)=1-F_N(1210)$$

$$F_N(1210)=1-[1-F(1210)]^3$$

$$\therefore P(N\geq 1210)=[1-\Phi(1)]^3=0.004$$

另解

独立性

$$p = P(X \le 1210) = F(1210) = \Phi(1) = 0.8413$$

没有一只寿命超过1210天

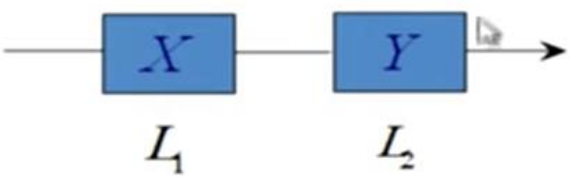
$$(1)\ p^3 = 0.5955$$

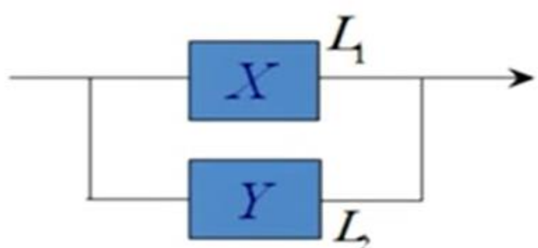
没有一只寿命小于1210天

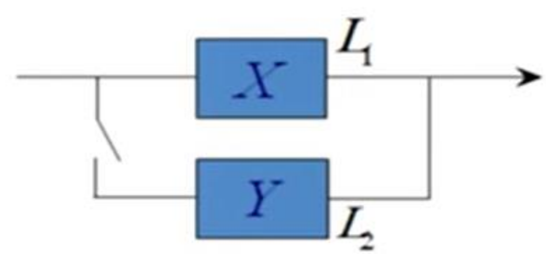
$$(2)\ (1-p)^3 = 0.004$$

设系统L由两个相互独立的子系统L_1、L_2联结而成，联结的方式分别为：(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，设L_1、L_2的寿命分别为X、Y，已知它们的概率密度分别如下所示，求系统寿命的概率密度.

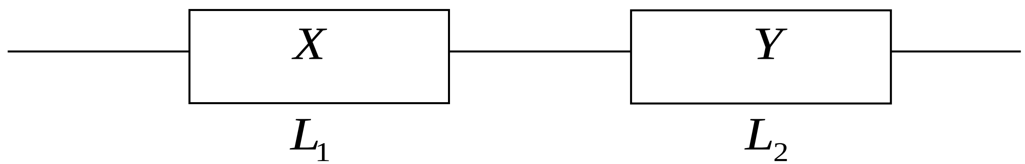
$$f_X(x)=\begin{cases}\alpha e^{-\alpha x}, & x>0,\\ 0, & x\le 0.\end{cases}\quad f_Y(y)=\begin{cases}\beta e^{-\beta y}, & y>0,\\ 0, & y\le 0.\end{cases}\quad \alpha>0,\beta>0,\alpha\neq\beta$$







解 (1) 串联的情况



因为有一个损坏时，系统 *L*就停止工作，
所以 *L*的寿命为 $Z=\min\{X, Y\}$,

X,Y都服从指数分布，分布函数为

$$F_X(x)=\begin{cases}1-e^{-\alpha x}, x>0;\\ 0, x\le 0,\end{cases} \quad , \quad F_Y(y)=\begin{cases}1-e^{-\beta y}, y>0;\\ 0, y\le 0,\end{cases}$$

故Z 的分布函数

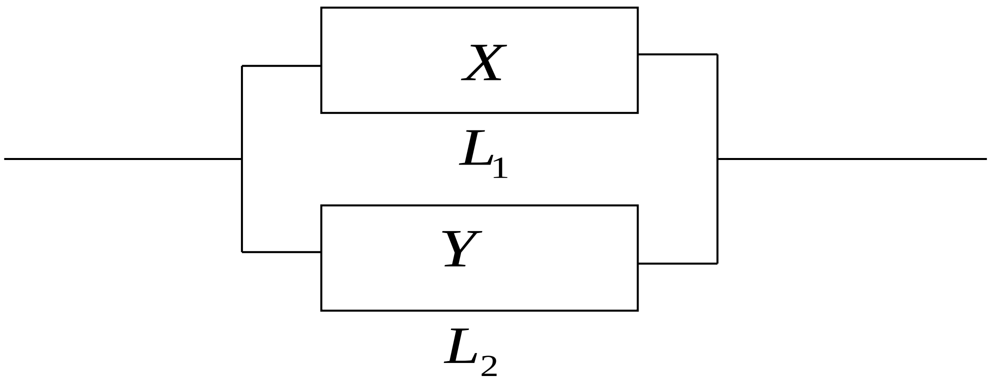
$$F_Z(z)=1-(1-F_X(z))(1-F_Y(z))$$

$$=\begin{cases}1-e^{-(\alpha+\beta)z}, & z>0,\\ 0, & z\le 0,\end{cases}$$

于是，得Z的密度函数

$$f_Z(z)=\begin{cases}(\alpha+\beta)e^{-(\alpha+\beta)z}, & z>0;\\ 0, & z\le 0.\end{cases}$$

(2) 并联的情况



因为当且仅当都损坏时，系统 L 才停止工作，所以L的寿命Z为

$$Z = \max\{X, Y\}$$

Z的分布函数

$$F_Z(z) = F_X(z)F_Y(z)$$

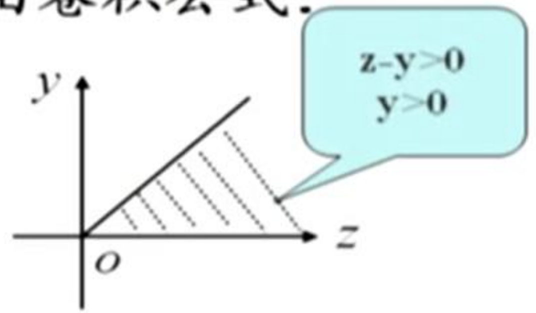
$$= \begin{cases} (1-e^{-\alpha z})(1-e^{-\beta z}), z > 0, \\ 0, z \le 0. \end{cases}$$

Z的密度函数

$$f_Z(z) = \begin{cases} \alpha e^{-\alpha z} + \beta e^{-\beta z} - (\alpha+\beta)e^{-(\alpha+\beta)z}, z > 0; \\ 0, \qquad z \le 0. \end{cases}$$

（3）备用的情况

由于这时当系统L_1损坏时，系统L_2才开始工作，因此整个系统L的寿命Z是L_1, L_2寿命之和，即$Z = X + Y$，由卷积公式：



当$z \le 0$时，$f_Z(z) = 0$

当$z > 0$时，$f_Z(z) = \int_{-\infty}^{+\infty} f_X(z-y) f_Y(y) dy$

$$= \int_0^z \alpha e^{-\alpha(z-y)} \beta e^{-\beta y} dy \quad = \alpha\beta e^{-\alpha z} \int_0^z e^{-(\beta-\alpha)y} dy$$

$$= \frac{\alpha\beta}{\beta-\alpha}[e^{-\alpha z} - e^{-\beta z}] \quad \text{即 } f_Z(z) = \begin{cases} \frac{\alpha\beta}{\beta-\alpha}[e^{-\alpha z} - e^{-\beta z}], & z > 0 \\ 0, & z \le 0 \end{cases}$$

设$X_i(i=1,\ldots,n)\sim f(x)$且相互独立,怎样求

$$Z=\max(X_1,\ldots,X_n)$$

的密度?

微分思路

剩余$n-1$个X_j 某个X_i

$[\quad]$

$z \qquad z+\mathrm{dz}$

对于充分小的区间 $:[z,z+\mathrm{dz}]$

$$P\{z\le Z\le z+\mathrm{dz}\}=n[F(z)]^{n-1}f(z)\mathrm{dz}$$

$$=f_{\max}(z)\mathrm{dz}$$

故

$$\boxed{f_{\max}(z)=nf(z)[F(z)]^{n-1}}$$

（二）顺序统计量 $X_{(k)}$ 的分布

设$X_i(i=1,\ldots,n)\sim f(x)$ 是独立同分布的连续型r.v.

将 $X_1,X_2,\ldots,X_n$ 由小到大排列为 $X_{(1)}\le X_{(2)}\le\ldots\le X_{(n)}$

顺序统计量 $(X_{(1)},X_{(2)},\ldots,X_{(n)})$

最小值 $X_{(1)}=\min\{X_1,X_2,\cdots,X_n\}$

最大值 $X_{(n)}=\max\{X_1,X_2,\cdots,X_n\}$

若 n 是奇数 $n=2m+1$，则 $X_{(m+1)}$ 称为中位数(median).

如何求$X_{(k)}$的密度？

微分思路

$k-1$个X_j $\quad [\quad X_{(k)} \quad] \quad$ $n-k$个X_j

$$x \qquad x+\mathrm{d}x$$

对于充分小的区间 : $[x, x+\mathrm{d}x]$

$$P\{x \le X_{(k)} \le x+\mathrm{d}x\}$$

$$=\frac{n!}{(k-1)!1!(n-k)!}F^{k-1}(x)[1-F(x)]^{n-k}f(x)dx$$

$$= f_k(x)\mathrm{d}x$$

故

$$\boxed{f_k(x)=\frac{n!}{(k-1)!(n-k)!}f(x)F^{k-1}(x)[1-F(x)]^{n-k}}$$

$$\boxed{f_k(x)=\frac{n!}{(k-1)!(n-k)!}f(x)F^{k-1}(x)[1-F(x)]^{n-k}}$$

若 $X_i(i=1,\dots,n)\sim \mathrm{U}[0,1]$ 且相互独立, 则 $X_{(k)}$ 的密度

$$\frac{n!}{(k-1)!(n-k)!}x^{k-1}(1-x)^{n-k},\quad 0\le x\le 1$$

$$\sim \mathrm{Beta}(k,n-k+1)$$

由于密度函数积分等于1，因此得到

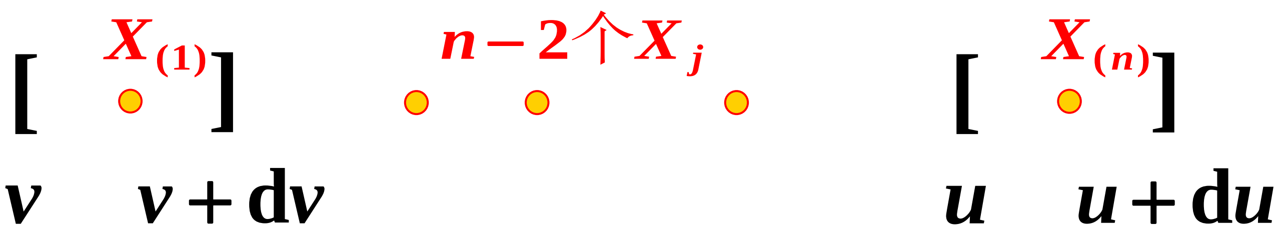
$$\boxed{\int_0^1 x^{k-1}(1-x)^{n-k}\mathrm{d}x=\frac{(k-1)!(n-k)!}{n!}.}$$

贝塔密度用来刻画[0,1] 区间上的r.v.:

$$f(u)=\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}u^{a-1}(1-u)^{b-1},\quad 0\le u\le 1$$

设 $V=X_{(1)}, U=X_{(n)}$，怎样求二者的联合密度？

微分思路



$$P\{v \le X_{(1)} \le v+\mathrm{d}v, u \le X_{(n)} \le u+\mathrm{d}u\}$$

$$= n(n-1)\cdot[F(u)-F(v)]^{n-2}\cdot f(v)\mathrm{d}v\cdot f(u)\mathrm{d}u$$

$$= f(u,v)\mathrm{d}u\mathrm{d}v$$

故

$$\boxed{f(u,v)=n(n-1)f(v)f(u)[F(u)-F(v)]^{n-2}, \quad u \ge v}$$

对于均匀分布，

$$f(u,v)=n(n-1)(u-v)^{n-2}, \quad 1 \ge u \ge v \ge 0$$

课后作业 P81：70、补充题

1. 从 1,2,3 中一次任取两个数，第一个数为 X，第二个为 Y，记 $Z=\max(X,Y)$，求(X,Y)和(X,Z)的联合频率函数和边缘频率函数。

2. 设 X 和 Y 是两个相互独立的随机变量，服从 $N(0,1)$，令 $Z=\min(X,Y)$，求 Z 的分布函数。

END