第三章：随机向量

3.1 二维随机向量及其分布

3.2 二维离散型随机变量

3.3 二维连续型随机向量

3.4 边缘分布

3.5 条件分布

3.6 随机变量的独立性

3.7 随机向量函数的分布

3.1 二维随机向量及其分布

二维随机向量及其分布

定义 1.1

对某个随机试验涉及的 n 个随机变量 $X_1, X_2, \cdots, X_n$, 记为 $(X_1, X_2, \cdots, X_n)$, 称为 n 维随机向量或 n 维随机变量.

定义 1.2

设 (X, Y) 是一个二维随机向量, 对任意实数 x, y, 称二元函数

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

为 (X, Y) 的分布函数.

分布函数 $F(x, y)$ 表示事件 $\{X \leq x\}$ 和 $\{Y \leq y\}$ 同时发生的概率!

二维随机向量及其分布: 性质

性质 1.1

任一二维联合分布函数 $F(x,y)$ 必具有如下四条性质:

(1) 单调性. $F(x,y)$ 分别对 x 或 y 是单调非减的, 即当 $x_1 < x_2$ 时, 有 $F(x_1,y) \le F(x_2,y)$; 当 $y_1 < y_2$ 时, 有 $F(x,y_1) \le F(x,y_2)$.

(2) 有界性. 对任意的 x 和 y, 有 $0 \le F(x,y) \le 1$, 且

$$F(-\infty,y) = \lim_{x\to-\infty} F(x,y) = 0, F(x,-\infty) = \lim_{y\to-\infty} F(x,y) = 0,$$

$$F(\infty,\infty) = \lim_{x,y\to\infty} F(x,y) = 1.$$

(3) 非负性. 对任意的 $a < b, c < d$ 有

$$P(a < X \le b, c < Y \le d) = F(b,d) - F(a,d) - F(b,c) + F(a,c).$$

证明: (1) 因为当 $x_1 < x_2$ 时, 有 $\{X \le x_1\} \subset \{X \le x_2\}$, 所以有

$$\{X \le x_1, Y \le y\} \subset \{X \le x_2, Y \le y\},$$

由此可得

$$F(x_1, y) = P\{X \le x_1, Y \le y\} \le P\{X \le x_2, Y \le y\} = F(x_2, y).$$

即 $F(x, y)$ 关于 x 是单调非减的. 同理 $F(x, y)$ 关于 y 是单调非减的.

(2) 由概率性质知, $0 \le F(x, y) \le 1$. 又因为对任意的正整数 n 有

$$\lim_{x \to -\infty} \{X \le x\} = \lim_{n \to \infty} \bigcap_{m=1}^{n} \{X \le -m\} = \varnothing,$$

联合分布函数: 性质证明

$$\lim_{x\to\infty}\{X\le x\} = \lim_{n\to\infty}\bigcup_{m=1}^{n}\{X\le m\} = \Omega.$$

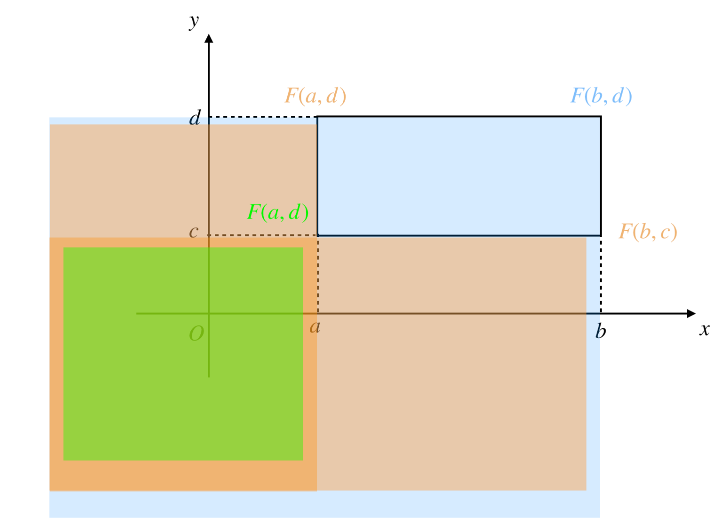
对 $\{Y\le y\}$ 也有类似的性质, 由概率的连续性可知:

$$F(-\infty, y) = F(x, -\infty) = 0, \quad F(\infty, \infty) = 1.$$

(3) 由下图分析易知, 将分布函数转换为落在某个区域的概率来考虑.

具有上述四条性质的二元函数 $F(x, y)$ 一定是某个二维随机变量的分布函数. 注意, 四条性质缺一不可, 因为 (1)-(3) 条无法推出第 (4) 条.

联合分布函数: 性质证明



3.2 二维离散型随机变量

二维离散型随机变量: 定义

定义 2.1

如果二维随机变量 (X, Y) 只取有限个或可列个数对 (x_i, y_i), 则称 (X, Y) 为二维离散随机变量, 称

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

为 (X, Y) 的联合分布列.

性质 2.1

联合分布列的基本性质:

(1) 非负性 $p_{ij} \geq 0$.

(2) 正则性 $\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} p_{ij} = 1$.

二维离散型随机变量: 例子

例 2.1

设有 *10* 件产品, 其中 *7* 件正品, *3* 件次品. 现从中任取两次, 每次取一件产品, 取后不放回, 令

$$X = \begin{cases} 1, & \text{第一次取到的产品是次品}, \\ 0, & \text{第一次取到的产品是正品}, \end{cases}$$

$$Y = \begin{cases} 1, & \text{第二次取到的产品是次品}, \\ 0, & \text{第二次取到的产品是正品}, \end{cases}$$

求二维随机向量 (X, Y) 的概率分布.

二维离散型随机变量: 例子

解: (X, Y) 所有可能的取值为: $(0,0), (0,1), (1,0), (1,1)$. 其中, $\{X=0, Y=0\}$ 表示第一次取到正品, 第二次也取到正品, 其概率为

$$P\{X=0, Y=0\} = \frac{7}{10} \times \frac{6}{9} = \frac{7}{15}.$$

类似地,

$$P\{X=0, Y=1\} = \frac{7}{10} \times \frac{3}{9} = \frac{7}{30},$$
$$P\{X=1, Y=0\} = \frac{3}{10} \times \frac{7}{9} = \frac{7}{30},$$
$$P\{X=1, Y=1\} = \frac{3}{10} \times \frac{2}{9} = \frac{1}{15}.$$

□

二维离散型随机变量: 例子

例 2.2

为了进行吸烟与肺癌关系的研究, 随机调查了 *23000* 个 *40* 岁以上的人, 其结果列在下表中.

肺癌 / 吸烟	患	不患	合计
吸	*3*	*4597*	*4600*
不吸	*1*	*18399*	*18400*
合计	*4*	*22996*	*2300*

$$X = \begin{cases} 1, & \text{被调查者不吸烟}, \\ 0, & \text{被调查者吸烟}, \end{cases} \quad Y = \begin{cases} 1, & \text{被调查者未患肺癌}, \\ 0, & \text{被调查者患肺癌}. \end{cases}$$

求二维随机向量 (X, Y) 的概率分布.

二维离散型随机变量: 例子

解: (X, Y) 所有可能的取值为: $(0,0)$, $(0,1)$, $(1,0)$, $(1,1)$. 其中, $\{X=0, Y=0\}$ 被调查者吸烟且患肺癌, 其概率为

$$P\{X=0, Y=0\} = \frac{3}{23000} = 0.00013.$$

类似地,

$$P\{X=0, Y=1\} = \frac{4597}{23000} = 0.19987,$$
$$P\{X=1, Y=0\} = \frac{1}{23000} = 0.00004,$$
$$P\{X=1, Y=1\} = \frac{18399}{23000} = 0.79996.$$

既吸烟又患肺癌的概率是 0.00013, 而不吸烟患肺癌的概率是 0.00004, 远小于 0.00013. □

3.3 二维连续型随机向量

二维连续型随机向量: 定义

定义 3.1

若存在二元非负函数 $f(x,y)$, 使得二维随机变量 (X,Y) 的分布函数 $F(x,y)$ 表示为

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

则称 (X,Y) 为二维连续随机变量, 称 $f(u,v)$ 为 (X,Y) 的联合密度函数, 简称概率密度.

二维连续型随机向量: 性质

性质 3.1

联合密度函数的基本性质:

(1) 非负性. $f(x,y) \geq 0$.

(2) 正则性.

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dydx = 1.$$

(3) 若 $F(x,y)$ 在点 (x,y) 处连续, 则有

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

(4) 若 G 为平面上的一个区域, 则事件 $\{(X,Y) \in G\}$ 的概率为

$$P((X,Y) \in G) = \iint_G f(x,y)dxdy.$$

二维连续型随机向量: 例子

例 3.1

设 (X,Y) 的概率密度函数为

$$f(x,y)=\frac{A}{\pi^2(16+x^2)(25+y^2)},$$

$-\infty<x<\infty$, $-\infty<y<\infty$. 其中 A 是常数. *(1)* 求常数 A; *(2)* 求 (X,Y) 的分布函数; *(3)* 计算 $P\{0<X<4,0<Y<5\}$.

解: (1) 由性质 (2) 可知

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{A}{\pi^2(16+x^2)(25+y^2)}dxdy=1.$$

二维连续型随机向量: 例子

上式等价于

$$\frac{A}{\pi^2}\left(\int_{-\infty}^{\infty}\frac{1}{16+x^2}dx\right)\left(\int_{-\infty}^{\infty}\frac{1}{25+y^2}dy\right)=1.$$

因为

$$\int_{-\infty}^{\infty}\frac{1}{16+x^2}dx=\frac{1}{4}\int_{-\infty}^{\infty}\frac{1}{1+(x/4)^2}dx=\frac{1}{4}\arctan x|_{-\infty}^{\infty}=\frac{\pi}{4}.$$

同理,

$$\int_{-\infty}^{\infty}\frac{1}{25+y^2}dy=\frac{\pi}{5}.$$

从而, $A=20$.

二维连续型随机向量: 例子

(2) 由分布函数定义可得

$$
\begin{aligned}
F(x,y) &= \int_{-\infty}^{x}\int_{-\infty}^{y} \frac{20}{\pi^2(16+x^2)(25+y^2)}dvdu \\
&= \frac{20}{\pi^2}\left(\int_{-\infty}^{x} \frac{1}{16+u^2}du\right)\left(\int_{-\infty}^{y} \frac{1}{25+v^2}dv\right) \\
&= \left(\frac{1}{\pi}\arctan\frac{x}{4}+\frac{1}{2}\right)\left(\frac{1}{\pi}\arctan\frac{y}{5}+\frac{1}{2}\right).
\end{aligned}
$$

(3) 将 (X,Y) 看作是平面上随机点的坐标, 则

$$
\begin{aligned}
&P\{0<X<4, 0<Y<5\} \\
&= \int_{-\infty}^{4}\int_{-\infty}^{5} \frac{20}{\pi^2(16+x^2)(25+y^2)}dvdu \\
&= \frac{1}{16}.
\end{aligned}
$$

二维连续型随机向量: 例子

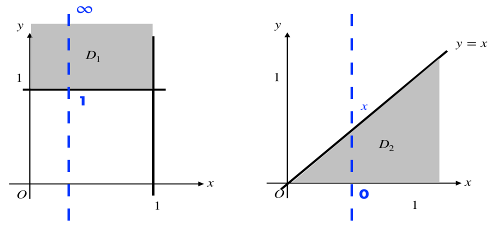
例 3.2

设 (X,Y) 的联合密度函数为

$$f(x,y)=\begin{cases}6e^{-2x-3y}, & x>0,y>0\\ 0, & \text{其他}.\end{cases}$$

试求: (1) $P(X<1,Y>1)$; (2) $P(X>Y)$.

解: 考虑两个题目中的积分区域



二维连续型随机向量: 例子

(1) 按照 D_1 区域积分

$$P(X<1,Y>1)=P\big((X,Y)\in D_1\big)=\int_0^1\int_1^\infty 6e^{-2x-3y}dydx$$
$$=6\int_0^1 e^{-2x}dx\int_1^\infty e^{-3y}dy$$
$$=(1-e^{-2})e^{-3}\approx 0.043.$$

(2) 按照 D_2 区域积分

$$P(X>Y)=P\big((X,Y)\in D_2\big)=\int_0^\infty\int_0^x 6e^{-2x-3y}dydx$$
$$=\left(-e^{-2x}+\frac{2}{5}e^{-5x}\right)\Big|_0^\infty=1-\frac{2}{5}=\frac{3}{5}.$$

$\square$

定义 3.2 (多维均匀分布)

设 D 为 $\mathbb{R}^n$ 中的一个有界区域, 其度量为 S_D (平面为面积, 空间为体积等), 如果多维随机变量 $(X_1, X_2, \cdots, X_n)$ 的联合密度函数为

$$p(x_1, x_2, \cdots, x_n) = \frac{1}{S_D}, \quad (x_1, x_2, \cdots, x_n) \in D,$$

则称 $(X_1, X_2, \cdots, X_n)$ 服从 D 上的多维均匀分布, 记为

$$(X_1, X_2, \cdots, X_n) \sim U(D).$$

例 3.3

设 D 为平面上以原点为圆心, 以 r 为半径的圆内区域, 如今向该圆内随机投点, 其坐标 (X, Y) 服从 D 上的二维均匀分布, 其密度函数为

$$f(x, y) = \begin{cases} \frac{1}{\pi r^2}, & x^2 + y^2 \leq r^2, \\ 0, & x^2 + y^2 > r^2. \end{cases}$$

试求概率 $P(|X| \leq r/2)$.

二维连续型随机向量: 二维均匀分布

解: 概率密度 $f(x,y)$ 的非零区域与事件 $\{|X| \le r/2\}$ 的交集部分如上图所示, 因此所求概率为

$$
\begin{aligned}
P(|X| \le \frac{r}{2}) &= \int_{-\frac{r}{2}}^{\frac{r}{2}} \int_{-\sqrt{r^2-x^2}}^{\sqrt{r^2-x^2}} \frac{1}{\pi r^2} dy dx = \frac{1}{\pi r^2} \int_{-\frac{r}{2}}^{\frac{r}{2}} 2\sqrt{r^2-x^2} dx \\
&= \frac{1}{\pi r^2} \left[x\sqrt{r^2-x^2} + r^2 \arcsin \frac{x}{r} \right] \Big|_{-\frac{r}{2}}^{\frac{r}{2}} \\
&= \frac{1}{\pi r^2} \left(r\sqrt{r^2 - \frac{r^2}{4}} + 2r^2 \arcsin \frac{1}{2} \right) \\
&= \frac{1}{\pi} \left(\frac{\sqrt{3}}{2} + \frac{\pi}{3} \right) = 0.609.
\end{aligned}
$$

$\square$

二维连续型随机向量: 二元正态分布

定义 3.3 (二元正态分布)

如果二维随机变量 (X,Y) 的联合密度函数为

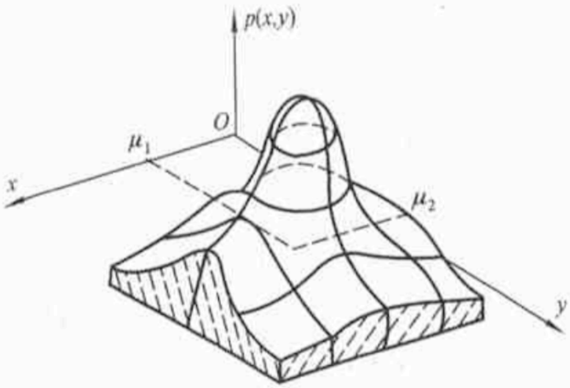
$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2}\right.\right.$$
$$\left.\left.-2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\right]\right\},\quad -\infty<x,y<\infty$$

则称 (X,Y) 服从二元正态分布, 记为 $(X,Y)\sim\mathcal{N}(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$. 其中五个参数的取值范围为

$$-\infty<\mu_1,\mu_2<\infty,\quad \sigma_1,\sigma_2>0,\quad -1\le\rho\le 1.$$

μ_1,μ_2 分别是 X 与 Y 的均值, σ_1^2,σ_2^2 分别是 X 与 Y 的方差, ρ 是 X 与 Y 的相关系数.

二维连续型随机向量: 二元正态分布



二维连续型随机向量: 二元正态分布例子

例 3.4

设二维随机变量 $(X,Y)\sim\mathcal{N}(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$, 求 (X,Y) 落在区域

$$D=\left\{(x,y):\frac{(x-\mu_1)^2}{\sigma_1^2}-2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\le\lambda\right\}$$

内的概率.

解: 所求概率为

$$p=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\iint_D\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2}\right.\right.$$

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

二维连续型随机向量: 二元正态分布例子

观察指数部分, 对 x 进行配方

$$\begin{aligned}&\frac{(x-\mu_1)^2}{\sigma_1^2}-2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\\=&\left(\frac{x-\mu_1}{\sigma_1}-\rho\frac{y-\mu_2}{\sigma_2}\right)^2+(1-\rho^2)\frac{(y-\mu_2)^2}{\sigma_2^2}\end{aligned}$$

可作变换

$$\begin{cases}u=\dfrac{x-\mu_1}{\sigma_1}-\rho\dfrac{y-\mu_2}{\sigma_2},\\ v=\dfrac{y-\mu_2}{\sigma_2}\sqrt{1-\rho^2}.\end{cases}$$

二维连续型随机向量: 二元正态分布例子

计算雅各比行列式的逆得

$$J^{-1} = \frac{\partial(u,v)}{\partial(x,y)} = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial v}{\partial x} \\ \frac{\partial u}{\partial y} & \frac{\partial v}{\partial y} \end{vmatrix} = \begin{vmatrix} \frac{1}{\sigma_1} & 0 \\ -\frac{\rho}{\sigma_2} & \frac{\sqrt{1-\rho^2}}{\sigma_2} \end{vmatrix} = \frac{\sqrt{1-\rho^2}}{\sigma_1\sigma_2}.$$

由此

$$p = \frac{1}{2\pi(1-\rho^2)} \iint_{u^2+v^2\le\lambda^2} \exp\left\{-\frac{u^2+v^2}{2(1-\rho^2)}\right\} du dv.$$

再做极坐标变换

$$\begin{cases} u = r\sin\alpha, \\ v = r\cos\alpha. \end{cases}$$

二维连续型随机向量: 二元正态分布例子

则

$$J = \begin{vmatrix} \frac{\partial u}{\partial r} & \frac{\partial v}{\partial r} \\ \frac{\partial u}{\partial \alpha} & \frac{\partial v}{\partial \alpha} \end{vmatrix} = -r(\sin^2\alpha + \cos^2\alpha) = -r.$$

最后得

$$\begin{aligned} p &= \frac{1}{2\pi(1-\rho^2)} \int_0^{2\pi} d\alpha \int_0^{\lambda} r \exp\left\{-\frac{r^2}{2(1-\rho^2)}\right\} dr \\ &= \int_0^{\lambda} \exp\left\{-\frac{r^2}{2(1-\rho^2)}\right\} d\left(\frac{r^2}{2(1-\rho^2)}\right) \\ &= -\exp\left\{-\frac{r^2}{2(1-\rho^2)}\right\}\Bigg|_0^{\lambda} = 1 - \exp\left\{-\frac{\lambda^2}{2(1-\rho^2)}\right\}. \end{aligned}$$

□

3.4 边缘分布

边缘分布函数

定义 4.1 (边缘分布)

如果在二维随机变量 (X,Y) 的联合分布函数 $F(x,y)$ 中令 $y\to\infty$, 由于 $\{Y<\infty\}$ 为必然事件, 故可得

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

这是由 (X,Y) 的联合分布函数 $F(x,y)$ 求得的 X 的分布函数, 被称为 X 的边际分布, 记为

$$F_X(x)=F(x,\infty).$$

类似地, 在 $F(x,y)$ 中令 $x\to\infty$, 可得 Y 的边际分布

$$F_Y(y)=F(\infty,y).$$

边缘分布函数: 例子

例 4.1

设二维随机变量 (X, Y) 的联合分布函数为二维指数分布,

$$F(x,y) = \begin{cases} 1 - e^{-x} - e^{-y} + e^{-x-y-\lambda xy}, & x > 0, y > 0. \\ 0, & \text{其他}. \end{cases}$$

求 X 和 Y 的边际分布函数 $F_X(x)$ 和 $F_Y(y)$.

解: 按定义

$$F_X(x) = F(x, \infty) = 1 - e^{-x}, x > 0,$$
$$F_Y(y) = F(\infty, y) = 1 - e^{-y}, y > 0,$$

在 $x \le 0$ 和 $y \le 0$ 的部分, $F_X(x)$ 和 $F_Y(y)$ 均为 0. □

二维离散型随机向量的边缘概率分布

定义 4.2

在二维离散随机变量 (X,Y) 的联合分布列为

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

对 j 求和所得的分布列

$$\sum_{j=1}^{\infty}P(X=x_i,Y=y_j)=P(X=x_i),\quad i=1,2,\cdots$$

被称为 X 的边缘概率分布. 类似地, 对 i 求和所得的分布列

$$\sum_{i=1}^{\infty}P(X=x_i,Y=y_j)=P(Y=y_j),\quad j=1,2,\cdots$$

被称为 Y 的边缘概率分布.

二维离散型随机向量的边缘概率分布: 例子

例 4.2

设二维随机变量 (X, Y) 有如下的联合分布列

	1	*2*	*3*
0	*0.09*	*0.21*	*0.24*
1	*0.07*	*0.12*	*0.27*

求 X 与 Y 的边际概率分布.

解: 按公式求得

	1	2	3	$P(X=i)$
0	0.09	0.21	0.24	0.54
1	0.07	0.12	0.27	0.46
$P(Y=j)$	0.16	0.33	0.51	1

定义 4.3

如果二维连续随机变量 (X, Y) 的联合密度函数为 $f(x, y)$, 其分布函数为

$$F_X(x) = F(x, \infty) = \int_{-\infty}^{x} \left(\int_{-\infty}^{\infty} f(u, v) dv \right) du = \int_{-\infty}^{x} f_X(u) du,$$

$$F_Y(y) = F(\infty, y) = \int_{-\infty}^{y} \left(\int_{-\infty}^{\infty} f(u, v) du \right) dv = \int_{-\infty}^{y} f_Y(v) dv,$$

其中 $f_X(x)$ 和 $f_Y(y)$ 分别为

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y) dy, \quad f_Y(y) = \int_{-\infty}^{\infty} f(x, y) dx,$$

其中 $f_X(x)$ 和 $f_Y(y)$ 分别称为 X 和 Y 的边缘概率密度.

二维连续型随机向量的边缘概率密度: 例子

例 4.3

设 (X,Y) 服从单位圆域 $D=\{(x,y): x^2+y^2=1\}$ 上的均匀分布, 求 X 和 Y 的边缘概率密度.

解: (X,Y) 的联合概率密度

$$f(x,y)=\frac{1}{\pi}, x^2+y^2\leq 1,$$

其他为 0. 先求 $f_X(x)$. 当 $x<-1$ 或 $x>1$ 时, $f(x,y)=0$, 从而 $f_X(x)=0$. 当 $-1\leq x\leq 1$ 时,

$$f_X(x)=\int_{-\infty}^{\infty} f(x,y)dy=\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}}\frac{1}{\pi}dy=\frac{2}{\pi}\sqrt{1-x^2}.$$

由于 X 和 Y 在问题中的对称性, $f_Y(y)=2\sqrt{1-y^2}/\pi$, $-1\leq y\leq 1$, 其他为 0. $\square$

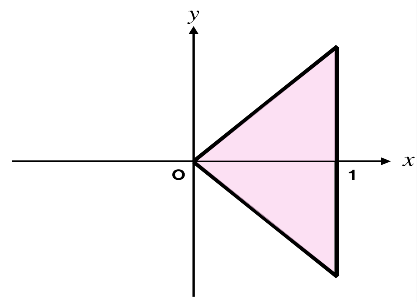
二维连续型随机向量的边缘概率密度: 例子

例 4.4

设二维随机变量 (X,Y) 的联合密度函数为

$$f(x,y)=\begin{cases}1, & 0<x<1, |y|<x\\ 0, & \text{其他}.\end{cases}$$

求边际密度函数 $f_X(x)$ 和 $f_Y(y)$.



二维连续型随机向量的边缘概率密度: 例子

解: 求 $f_X(x)$. 当 $x \le 0$ 或 $x \ge 1$ 时, 有 $f_X(x) = 0$. 当 $0 < x < 1$ 时,

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y) dy = \int_{-x}^{x} dy = 2x, 0 < x < 1.$$

再求 $f_Y(y)$. 当 $y \le -1$ 或 $y \ge 1$ 时, $f_Y(y) = 0$. 当 $-1 < y < 0$ 时,

$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y) dx = \int_{-y}^{1} dx = 1 + y;$$

当 $0 < y < 1$ 时,

$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y) dx = \int_{y}^{1} dx = 1 - y.$$

$\square$

二维连续型随机向量的边缘概率密度: 例子

例 4.5

证明二维正态分布的边际分布为一维正态分布.

证: 设二维随机变量 $(X, Y) \sim \mathcal{N}(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$. 将密度函数指数部分

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

改写为

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

二维连续型随机向量的边缘概率密度: 例子

作变换 (将 x 看成常量)

$$t = \rho\frac{x-\mu_1}{\sigma_1\sqrt{1-\rho^2}} - \frac{y-\mu_2}{\sigma_2\sqrt{1-\rho^2}},$$

则

$$\begin{aligned} f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \\ &\quad \times \exp\left\{-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right\}\sigma_2\sqrt{1-\rho^2}\int_{-\infty}^{\infty}\exp\left\{-\frac{t^2}{2}\right\}dt \\ &= \frac{1}{\sqrt{2\pi}\sigma_1}\exp\left\{-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right\}. \end{aligned}$$

即 $X \sim \mathcal{N}(\mu_1, \sigma_1^2)$. 同理, $Y \sim \mathcal{N}(\mu_2, \sigma_2^2)$. $\square$

二维连续型随机向量的边缘概率密度: 讨论

二维联合分布不仅含有每个分量的概率分布, 而且还含有两个变量 X 与 Y 间关系的信息.

- 二维正态分布的边际分布不含参数 ρ.
- 二维正态分布

$$\mathcal{N}(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, 0.8) \quad 与 \quad \mathcal{N}(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, -0.5)$$

  的边际分布是相同的!
- 具有相同边际分布的多维联合分布可以是不同的.

3.5 条件分布

离散型条件分布

定义 5.1

设二维离散随机变量 (X, Y) 的联合分布列为

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

对一切使 $P(Y = y_j) = \sum_{i=1}^{\infty} p_{ij} \equiv p_{\cdot j} > 0$ 的 y_j, 称

$$P(X = x_i | Y = y_j) = \frac{P(X = x_i, Y = y_j)}{P(Y = y_j)} = \frac{p_{ij}}{p_{\cdot j}}, i = 1, 2, \cdots$$

为给定 $Y = y_j$ 条件下 X 的条件分布列.

离散型条件分布: 例子

例 5.1

设 X 在 *1, 2, 3, 4* 四个整数中等可能地取值, Y 在 $1, 2, \cdots, X$ 中等可能地取值, 已求出 (X, Y) 的分布律以及 X 和 Y 的边沿分布, 如下表所示. 试求 $Y|X=3$ 条件下 Y 的条件分布以及 $X|Y=3$ 条件下 X 的条件分布.

Y / X	1	2	3	4	P_{i*}
1	1/4	0	0	0	1/4
2	1/8	1/8	0	0	1/4
3	1/12	1/12	1/12	0	1/4
4	1/16	1/16	1/16	1/16	1/4
P_{*j}	25/48	13/48	7/48	1/16	

离散型条件分布: 例子 (续)

解: 观察下表

Y / X	1	2	3	4	P_{i*}
1	1/4	0	0	0	1/4
2	1/8	1/8	0	0	1/4
3	1/12	1/12	1/12	0	1/4
4	1/16	1/16	1/16	1/16	1/4
P_{*j}	25/48	13/48	7/48	1/16	

$$P(Y=1|X=3)=\frac{p_{31}}{p_{3\cdot}}=\frac{1/12}{1/4}=\frac{1}{3};$$

$$P(X=4|Y=3)=\frac{p_{43}}{p_{\cdot 3}}=\frac{1/16}{7/48}=\frac{3}{7}.$$

连续型条件分布: 定义

定义 5.2

对一切使 $f_Y(y) > 0$ 的 y, 给定 $Y = y$ 条件下 X 的条件分布函数和条件密度函数分别为

$$F(x|y) = \int_{-\infty}^{x} \frac{f(u,y)}{f_Y(y)} du, \quad f(x|y) = \frac{f(x,y)}{f_Y(y)}.$$

对连续随机变量 $P(Y = y) = 0$, 无法直接除, 需转换为 $h \to 0$ 时,

$$P(X \le x | y \le Y \le y + h) = \lim_{h\to 0} P(X \le x | y \le Y \le y + h)$$

$$= \lim_{h\to 0} \frac{\int_{-\infty}^{x} \frac{1}{h} \int_{y}^{y+h} f(u,v) dv du}{\frac{1}{h} \int_{y}^{y+h} f_Y(v) dv} = \int_{-\infty}^{x} \frac{f(u,y)}{f_Y(y)} du$$ (积分中值定理).

连续型条件分布: 例子

例 5.2

设二维随机变量 (X, Y) 服从 $G = \{(x, y)|x^2 + y^2 \le 1\}$ 上的均匀分布, 试求给定 $Y = y$ 条件下 X 的条件密度 $f(x|y)$.

解: 先求 Y 的边缘密度

$$f_Y(y) = \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} \frac{1}{\pi} dx = \frac{2}{\pi}\sqrt{1-y^2}, -1 \le y \le 1.$$

当 $-1 \le y \le 1$ 时, 给定 Y 的条件下, $X|Y = y$ 的密度为

$$f(x|y) = \frac{f(x, y)}{f(y)} = \frac{(1/\pi)}{(2/\pi)\sqrt{1-y^2}} = \frac{1}{2\sqrt{1-y^2}},$$

其中 $-\sqrt{1-y^2} \le x \le \sqrt{1-y^2}$. □

连续型条件分布: 二元正态分布例子

例 5.3

设 (X, Y) 服从二维正态分布 $\mathcal{N}(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2, \rho)$, 试求条件概率密度函数 $f_{X|Y}(x|y)$.

解: 先考虑联合密度

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

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

注意技巧把 $f(x,y)$ 中关于 y 的部分看成常数, 将与 x 有关的部分配方,

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

连续型条件分布: 二元正态分布例子

二维正态分布的边缘分布都为正态分布, 即 $Y \sim \mathcal{N}(\mu_2, \sigma_2^2)$. 考虑

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

其中 $h(y)$ 为 y 的函数. 注意与 x 有关的部分, $X|Y \sim \mathcal{N}(\mu_3, \sigma_3^2)$, 其核心部分

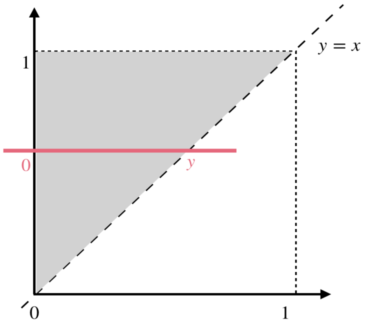
$$f(x|y) \propto \exp\left\{-\frac{1}{2\sigma_1^2(1-\rho^2)}\left[x - \left(\mu_1 + \rho\frac{\sigma_1}{\sigma_2}(y-\mu_2)\right)\right]^2\right\}.$$

观察其核心部分,

$$\mu_3 = \mu_1 + \rho\frac{\sigma_1}{\sigma_2}(y-\mu_2),\ \sigma_3^2 = \sigma_1^2(1-\rho^2).$$

例 5.4

设 X 在区间 $(0,1)$ 上随机地取值, 当观察到 $X = x$ $(0 < x < 1)$ 时, Y 在区间 $(x,1)$ 上随机地取值. 求 Y 的概率密度.



连续型条件分布: 例子

解: 由题意知

$$f_X(x) = 1, 0 < x < 1, \quad f_{Y|X}(x|y) = \frac{1}{1-x}, x < y < 1.$$

$$f(x,y) = f_X(x) f_{Y|X}(x|y) = \frac{1}{1-x}, 0 < x < 1, x < y < 1.$$

则

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{\infty} f(x,y)dx \\ &= \int_0^y \frac{1}{1-x}dx \\ &= -ln(1-y), 0 < y < 1. \end{aligned}$$

$\square$

连续型条件分布: 例子

例 5.5

设二维随机向量 (X,Y) 的概率密度函数为

$$f(x,y)=\begin{cases}\frac{21}{4}x^2y, & x^2\le y\le 1,\\ 0, & \text{其他}.\end{cases}$$

求条件概率密度和条件概率 $P\{Y>\frac{3}{4}|X=\frac{1}{2}\}$.

解: 先求得 X 和 Y 的边缘概率密度为

$$f_X(x)=\begin{cases}\frac{21}{8}x^2(1-x^4), & -1\le x\le 1,\\ 0, & \text{其他}.\end{cases}\quad f_Y(y)=\begin{cases}\frac{7}{2}y^{\frac{5}{2}}, & 0\le y\le 1,\\ 0, & \text{其他}.\end{cases}$$

连续型条件分布: 例子

从而求条件概率分布 $f_{X|Y}(x|y)$: 对 $0 < y \le 1$,

$$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)} = \begin{cases} \frac{3}{2}x^2y^{-\frac{3}{2}}, & -\sqrt{y} \le x \le \sqrt{y} \\ 0, & \text{其他}. \end{cases}$$

条件概率密度 $f_{Y|X}(y|x)$: 对 $-1 < x < 1$,

$$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)} = \begin{cases} \frac{2y}{1-x^4}, & x^2 \le y \le 1 \\ 0, & \text{其他}. \end{cases}$$

特别地, 对 $x = 1/2$, 有 $f_{Y|X}(y|\frac{1}{2}) = \frac{32}{15}y$, $\frac{1}{4} \le y \le 1$. 从而

$$P\{Y > \frac{3}{4}|X = \frac{1}{2}\} = \int_{\frac{3}{4}}^{\infty} f_{Y|X}(y|\frac{1}{2})dy = \int_{\frac{3}{4}}^{1} \frac{32}{15}ydy = \frac{7}{15}.$$

3.6 随机变量的独立性

定义 6.1

设二维随机变量 (X,Y) 的分布函数为 $F(x,y)$, X 和 Y 的边缘分布函数分别为 $F_X(x)$ 和 $F_Y(y)$. 若对任意的实数 x,y 有

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

则称随机变量 X 与 Y 相互独立.

性质 6.1

由分布函数的定义,

$$P\{X\le x,Y\le y\}=P\{X\le x\}\cdot P\{Y\le y\}.$$

因此, 随机变量 X 与 Y 相互独立是指对任意实数 x,y, 随机事件 $\{X\le x\}$ 与 $\{Y\le y\}$ 相互独立.

随机变量的独立性

性质 6.2

设 (X, Y) 是二维离散型随机向量, 其所有可能取的值为 (x_i, y_i), $i = 1, 2, \cdots, j = 1, 2, \cdots$, 则 X 与 Y 相互独立的条件可写为

$$P\{X = x_i, Y = y_i\} = P\{X = x_i\}P\{Y = y_i\},$$

其中 $i = 1, 2, \cdots, j = 1, 2, \cdots$.

性质 6.3

设 (X, Y) 是二维连续型随机向量, 则 X 与 Y 相互独立的条件可写为: 对任意的实数 x, y, 有

$$f(X, Y) = f_X(x)f_Y(y).$$

随机变量的独立性

对连续型情况加以说明.

- 若 X, Y 相互独立, 则 $\forall x, y \in \mathbb{R}$, 有 $F(x,y) = F_X(x)F_Y(y)$, 两边对 x, y 求导, 得

$$f(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y} = \frac{\partial F_X(x)}{\partial x}\frac{\partial F_Y(y)}{\partial y} = f_X(x)f_Y(y).$$

- 反之, 若对任意实数 $x, y, f(x,y) = f_X(x)f_Y(y)$, 则

$$\begin{aligned} F(x,y) &= \int_{-\infty}^{y}\int_{-\infty}^{x} f(u,v)dudv = \int_{-\infty}^{y}\int_{-\infty}^{x} f_X(u)f_Y(v)dudv \\ &= \left(\int_{-\infty}^{x} f_x(u)du\right)\left(\int_{-\infty}^{y} f_Y(v)dv\right) = F_X(x)F_Y(y). \end{aligned}$$

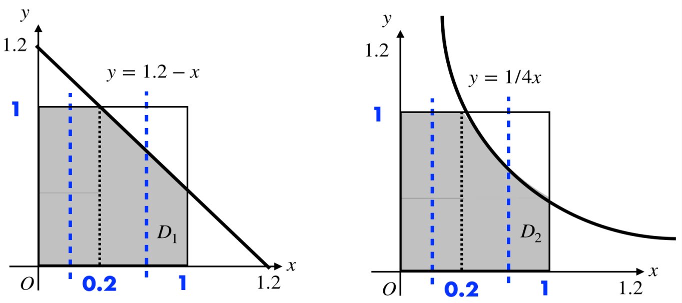
  表明 X 与 Y 相互独立.

例 6.1

从 (0, 1) 中任取两个数, 求下列事件的概率

(1) 两数之和小于 1.2;

(2) 两数之积小于 1/4.



随机变量间的独立性: 例子

解: 分别记这两个数为 X 和 Y, 则 X 和 Y 都服从 $(0,1)$ 上的均匀分布, 由于 X 与 Y 之间没有互相影响, 故 X 与 Y 之间相互独立, 其联合密度为

$$f(x,y)=f_X(x)f_Y(y)=1, 0<x<1, 0<y<1.$$

(1) 事件 $\{X+Y<1.2\}$ 的非零区域为 D_1, 其概率为

$$\begin{aligned} P(X+Y<1.2) &= \int\int_{D_1} 1dxdy \\ &= \int_0^{0.2}\int_0^1 dydx + \int_{0.2}^1\int_0^{1.2-x} dydx \\ &= 0.68. \end{aligned}$$

随机变量间的独立性: 例子

(2) 事件 $\{XY < 1/4\}$ 的非零区域为 D_2, 其概率为

$$
\begin{aligned}
P(XY < 1/4) &= \int\int_{D_2} 1dxdy \\
&= \int_0^{1/4}\int_0^1 dydx + \int_{1/4}^1\int_0^{1/4x} dydx \\
&= 0.5966.
\end{aligned}
$$

□

随机变量的独立性: 例子

例 6.2

若二维随机变量 (X, Y) 的联合密度函数为

$$f(x,y)=\begin{cases}8xy, & 0\le x\le y\le 1,\\ 0, & \text{其他}.\end{cases}$$

问 X 与 Y 是否相互独立?

解: 考虑 X 和 Y 支撑上密度.

$$f_X(x)=\int_x^1 8xydy=8x(\frac{1}{2}-\frac{x^2}{2})=4x(1-x^2), 0\le x\le 1,$$

$$f_Y(y)=\int_0^y 8xy=4y^3, 0\le y\le 1.$$

显然 $f_X(x)f_Y(y)\neq f(x,y)$, 故 X 与 Y 不相互独立. □

随机变量的独立性: 讨论

直观上看, 联合密度 $f(x,y)$ 似乎可分离变量. **但由于其非零区域相交织, X 的取值受 Y 的值影响 $(0 \le x \le y)$, Y 的取值也受 X 的取值的影响 $(x \le y \le 1)$**. 最后导致 $f(x,y)$ 不可分离, X 与 Y 相互不可能相互独立.

例 6.3

若二维随机变量 (X, Y) 的联合密度函数为

$$f(x,y) = \begin{cases} 12y^2, & 0 \le y \le x \le 1, \\ 0, & \text{其他}. \end{cases}$$

因为 $0 \le y \le x \le 1$ 不可分离, X 和 Y 一定不独立.

随机变量的独立性: 例子

例 6.4

设 $(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$, 则 X 与 Y 相互独立的充分必要条件是 $\rho = 0$.

证明: (1) 充分性. 设 $\rho = 0$, 则

$$f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2} e^{-\frac{1}{2}\left[\frac{(x-\mu_1)^2}{\sigma_1^2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right]} = f_X(x) f_Y(y),$$

从而 X 与 Y 相互独立. (2) 必要性. 设 X 与 Y 相互独立, 则对任意 $x, y, f(x, y) = f_X(x) f_Y(y)$. 特别取 $x = \mu_1, y = \mu_2$, 得到 $f(\mu_1, \mu_2) = f_X(\mu_1) f_Y(\mu_2)$, 即

$$\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} = \frac{1}{\sqrt{2\pi}\sigma_1} \frac{1}{\sqrt{2\pi}\sigma_2}.$$

于是 $\sqrt{1-\rho^2} = 1 \Rightarrow \rho = 0$. $\square$

3.7 随机向量函数的分布

例 7.1

设二维随机变量 (X, Y) 的联合分布列为

	-1	1	2
-1	5/20	2/20	6/20
2	3/20	3/20	1/20

试求:

(1) $Z_1 = X + Y$;

(2) $Z_2 = X - Y$;

(3) $Z_3 = \max\{X, Y\}$ 的分布列.

解: 将 (X, Y) 及其各个函数的取值对应于同一表中:

P	5/20	2/20	6/20	3/20	3/20	1/20
(X, Y)	(-1, 1)	(-1, 1)	(-1, 2)	(2, -1)	(2, 1)	(2, 2)
$Z_1 = X + Y$	-2	0	1	1	3	4
$Z_2 = X - Y$	0	-2	-3	3	1	0
$Z_3 = \max\{X, Y\}$	-1	1	2	2	2	2

离散情况: 例子

整理可得:

$Z_1 = X + Y$	-2	0	1	3	4
P	5/20	2/20	9/20	3/20	1/20

$Z_2 = X - Y$	-3	-2	0	1	3
P	6/20	2/20	6/20	3/20	3/20

$Z_3 = \max\{X, Y\}$	-1	1	2
P	5/20	2/20	13/20

$\square$

离散情况: 泊松分布的可加性

例 7.2

设随机变量 $X \sim \mathcal{P}(\lambda_1), Y \sim \mathcal{P}(\lambda_2)$, 且 X 与 Y 独立, 证明 $Z = X + Y \sim \mathcal{P}(\lambda_1 + \lambda_2)$.

解: $Z = X + Y$ 可取 $0, 1, 2, \cdots$ 所有非负整数. 事件 $\{Z = k\}$ 是如下诸互不相容事件

$$\{Z = k\} = \bigcup_{i=0}^{k} \{X = i, Y = k - i\}.$$

由于 X 和 Y 相互独立, 对任意非负整数 k, 有

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

离散情况: 泊松分布的可加性

这个等式被称为**离散场合下卷积公式**.

$$
\begin{aligned}
P(Z=k) &= \sum_{i=0}^{k}\left(\frac{\lambda_1^i}{i!}e^{-\lambda_1}\right)\left(\frac{\lambda_2^{k-i}}{(k-i)!}e^{-\lambda_2}\right) \\
&= \frac{(\lambda_1+\lambda_2)^k}{k!}e^{-(\lambda_1+\lambda_2)}\sum_{i=0}^{k}\frac{k!}{i!(k-i)!}\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^i\left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{k-i} \\
&= \frac{(\lambda_1+\lambda_2)^k}{k!}e^{-(\lambda_1+\lambda_2)}\left(\frac{\lambda_1}{\lambda_1+\lambda_2}+\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^k \\
&= \frac{(\lambda_1+\lambda_2)^k}{k!}e^{-(\lambda_1+\lambda_2)}, \quad k=0,1,\cdots
\end{aligned}
$$

即 $X+Y \sim P(\lambda_1+\lambda_2)$. $\square$

连续情况: 卷积公式

定理 7.1

设 X 与 Y 是两个相互独立的连续随机变量, 其密度函数分别为 $f_X(x)$ 和 $f_Y(y)$, 则其和 $Z = X + Y$ 的密度函数为

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(z-y) f_Y(y)dy = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)dx.$$

证: $Z = X + Y$ 的分布函数为

$$\begin{aligned}
F_Z(z) = P(X+Y \le z) &= \iint_{x+y\le z} f_X(x) f_Y(y)dxdy \\
&= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{z-y} f_X(x)dx \right) f_Y(y)dy \\
&= \int_{-\infty}^{z} \left(\int_{-\infty}^{\infty} f_X(t-y) f_Y(y)dy \right) dt.
\end{aligned}$$

连续情况: 卷积公式

由此得 X 的密度函数为

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(z-y) f_Y(y) dy.$$

在上式积分中令 $z - y = x$, 则可得

$$\int_{-\infty}^{\infty} f_X(x) f_Y(z-x) dx.$$

这就是**连续场合下的卷积公式.** $\square$

连续情况: 例子

例 7.3

设某种商品一周的需要量是一个随机变量, 其概率密度函数为

$$f(x) = \begin{cases} xe^{-x}, & x > 0, \\ 0, & x \leq 0. \end{cases}$$

如果各周的需要量相互独立, 求两周需要量的概率密度函数.

解: 令 X 和 Y 表示第一、二周的需要量, 则

$$f_X(x) = \begin{cases} xe^{-x}, & x > 0, \\ 0, & x \leq 0. \end{cases} \qquad f_Y(y) = \begin{cases} ye^{-y}, & y > 0, \\ 0, & y \leq 0. \end{cases}$$

两周需要量 $Z = X + Y$.

(1) 当 $z \le 0$ 时, 若 $x > 0$, 则 $z - x < 0, f_Y(z-x) = 0$.

(2) 若 $x \le 0$, 则 $f_X(x) = 0$, 从而 $f_Z(z) = 0$.

(3) 当 $z > 0$ 时, 若 $x \le 0$, 则 $f_X(x) = 0$. 若 $z - x \le 0$, 则 $f_Y(z-x) = 0$. 因此,

$$
\begin{aligned}
\int_{-\infty}^{\infty} f_X(x) f_Y(z-x)dx &= \int_0^z x e^{-x}(z-x)e^{-(z-x)}dx \\
&= e^{-z}\int_0^z x(z-x)dx = \frac{z^3}{6}e^{-z}.
\end{aligned}
$$

从而

$$
f_Z(z) = \begin{cases} \frac{z^3}{6}e^{-z}, & z > 0, \\ 0, & z \le 0. \end{cases}
$$

$\square$

连续情况: 正态分布的可加性

例 7.4

设随机变量 $X \sim \mathcal{N}(\mu_1, \sigma_1^2)$, $Y \sim \mathcal{N}(\mu_2, \sigma_2^2)$, 且 X 与 Y 独立, 证明 $Z = X + Y \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.

证: 因为 $Z = X + Y$ 仍在 $(-\infty, \infty)$ 上取值, 利用卷积公式

$$f_Z(z) = \frac{1}{2\pi\sigma_1\sigma_2}\int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}\left[\frac{(z-y-\mu_1)^2}{\sigma_1^2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right]\right) dy,$$

对 y 进行配方

$$-\frac{1}{2}\left[\frac{(z-y-\mu_1)^2}{\sigma_1^2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right] = A\left(y - \frac{B}{A}\right)^2 + \frac{(z-\mu_1-\mu_2)^2}{\sigma_1^2+\sigma_2^2},$$

其中

$$A = \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}, \quad B = \frac{z-\mu_1}{\sigma_1^2} + \frac{\mu_2}{\sigma_2^2}.$$

连续情况: 正态分布的可加性

代入原式可得

$$f_Z(z) = \frac{1}{2\pi\sigma_1\sigma_2}\exp\left(-\frac{1}{2}\frac{(z-\mu_1-\mu_2)^2}{\sigma_1^2+\sigma_2^2}\right)\cdot\int_{-\infty}^{\infty}\exp\left(-\frac{A}{2}\Big(y-\frac{B}{A}\Big)^2\right)dy.$$

利用正态分布的正则性,

$$\exp\left(-\frac{A}{2}\Big(y-\frac{B}{A}\Big)^2\right)dy = \frac{\sqrt{2\pi}}{A},$$

于是

$$f_Z(z) = \frac{1}{\sqrt{2\pi(\sigma_1^2+\sigma_2^2)}}\exp\left(-\frac{1}{2}\frac{(z-\mu_1-\mu_2)^2}{\sigma_1^2+\sigma_2^2}\right),$$

即 $Z \sim \mathcal{N}(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$. □

连续情况: 正态分布的可加性

定理 7.2

任意 n 个相互独立的正态变量的线性组合仍是正态变量, 即

$$a_1X_1 + a_2X_2 + \cdots + a_nX_n \sim \mathcal{N}(\mu_0, \sigma_0^2).$$

若记 $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2), i = 1, 2, \cdots, n$, 则参数 μ_0 与 σ_0^2 分别为

$$\mu_0 = \sum_{i=1}^{n} a_i\mu_i, \quad \sigma_0^2 = \sum_{i=1}^{n} a_i^2\sigma_i^2.$$

例 7.5

已知 $X \sim \mathcal{N}(-3, 1), Y \sim \mathcal{N}(2, 1)$, 且 X 与 Y 独立, 则

$$Z = X - 2Y + 7 \sim \mathcal{N}(-3 - 2 \times 1 + 7, 1 + 4 \times 1) = \mathcal{N}(0, 5).$$

最大值分布

例 7.6

设 $X_1, \cdots, X_n$ 是 n 个独立同分布连续随机变量, $X_i \sim F_i(x)$, 其密度函数为 $f(x)$, $i = 1, \cdots, n$. 求 $Y = \max\{X_1, \cdots, X_n\}$ 的分布.

解: Y 的分布函数为

$$\begin{aligned} F_Y(y) &= P(\max\{X_1, X_2, \cdots, X_n\} \le y) \\ &= P(X_1 \le y, X_2 \le y, \cdots, X_n \le y) \\ &= P(X_1 \le y) P(X_2 \le y) \cdots P(X_n \le y) \\ &= \prod_{i=1}^{n} F_i(y) = [F(y)]^n. \end{aligned}$$

密度函数为

$$f_Y(y) = F_Y'(y) = n[F(y)]^{n-1} f(y).$$

最小值分布

例 7.7

设 $X_1, \cdots, X_n$ 是 n 个独立同分布连续随机变量, $X_i \sim F_i(x)$, 其密度函数为 $f(x)$, $i = 1, \cdots, n$. 求 $Y = \min\{X_1, \cdots, X_n\}$ 的分布.

解: Y 的分布函数为

$$\begin{aligned} F_Y(y) &= P(\min\{X_1, X_2, \cdots, X_n\} \leq y) \\ &= 1 - P(X_1 \leq y, X_2 \leq y, \cdots, X_n > y) \\ &= 1 - P(X_1 > y)P(X_2 > y) \cdots P(X_n > y) \\ &= 1 - \prod_{i=1}^{n}[1 - F_i(y)] = 1 - [1 - F(y)]^n. \end{aligned}$$

密度函数为

$$f_Y(y) = F_Y'(y) = n[1 - F(y)]^{n-1} f(y).$$

最小值分布: 例子

例 7.8

某段道路原来有 5 个路灯, 道路改建后有 20 个路灯用于此道路的晚间照明. 改建后道路管理人员总认为灯泡更容易坏了, 请解释其中原因.

解: 设所有灯泡的使用寿命是独立同分布的随机变量, 其共同分布为指数分布 $Exp(2000^{-1})$.

- 道路改建后前 5 个灯泡中第一个灯泡烧坏的时间

$$T_1 = \min\{X_1, \cdots, X_5\}.$$

最小值分布: 例子

- 若每只灯泡每天用 10 小时, 则 30 天内需更换灯泡的概率

$$
\begin{aligned}
P(T_1 \le 300) &= 1 - \exp\{-5\lambda \cdot 300\} \\
&= 1 - \exp\Big\{-\frac{1500}{2000}\Big\} = 0.5276.
\end{aligned}
$$

- 道路改建后, 第一只烧坏时间 $T_2 = \min\{X_1, \cdots, X_{20}\}$, 则 30 天内需更换灯泡的概率

$$
\begin{aligned}
P(T_2 \le 300) &= 1 - \exp\{-20\lambda \cdot 300\} \\
&= 1 - \exp\Big\{-\frac{6000}{2000}\Big\} = 0.9502.
\end{aligned}
$$

□

例 7.9

设系统 L 由两个相互独立的子系统 L_1, L_2 连接而成, 连接的方式分别为: (1) 串联; (2) 并联; (3) 备用 (开关完全可靠, 子系统 L_2 在储备期内性能无变化, 当 L_1 损坏时, L_2 开始工作). 设 L_1, L_2 的寿命分别为 X 和 Y, 其概率密度分别为

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

其中 $\alpha > 0, \beta > 0$, 且 $\alpha \neq \beta$. 分别求上述三种连接方式写出 L 的寿命 Z 的概率密度函数.

解: 求得 X 和 Y 的分布函数

$$F_X(x) = \begin{cases} 1 - e^{-\alpha x}, & x > 0, \\ 0, & x \leq 0 \end{cases} \qquad F_Y(x) = \begin{cases} 1 - e^{-\beta y}, & y > 0, \\ 0, & y \leq 0 \end{cases}$$

(1) 串联时, $Z = \min\{X, Y\}$. 当 $z \leq 0$ 时, $F_{\min}(z) = 0$. 当 $z > 0$ 时, 其分布函数为

$$F_{\min}(z) = 1 - [1 - F_X(z)][1 - F_Y(z)] = 1 - e^{-(\alpha+\beta)z}.$$

其密度函数为

$$f_{\min}(z) = (\alpha + \beta)e^{-(\alpha+\beta)z}, \quad z > 0.$$

当 $z \leq 0$ 时,$f_{\min}(z) = 0$.

随机向量函数的分布: 综合练习

(2) 并联时, $Z = \min\{X, Y\}$, 其分布函数为

$$F_{\min}(z) = \begin{cases} (1 - e^{-\alpha z})(1 - e^{-\beta z}), & z > 0, \\ 0, & x \leq 0 \end{cases}$$

其密度函数为

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

随机向量函数的分布: 综合练习

(3) 备用时, $Z = X + Y$. 当 $z \leq 0$ 时, $f_Z(z) = 0$. 当 $z > 0$ 时,

$$f_Z(z) = \int_0^z \alpha e^{-\alpha x} \beta e^{-\beta(z-x)} dx = \frac{\alpha\beta}{\alpha - \beta}(e^{-\beta x} - e^{-\alpha x}).$$

从而,

$$f_{\min}(z) = \begin{cases} \frac{\alpha\beta}{\alpha-\beta}(e^{-\beta x} - e^{-\alpha x}), & z > 0, \\ 0, & z \leq 0 \end{cases}$$

$\square$