数据科学基础

Foundations of Data Science

8.1 多维概率分布

陈振宇
南京大学智能软件工程实验室
www.iselab.cn

二维连续概率分布律-课堂练习

例 8.1:

$$F(x,y)=1-e^{-0.01x}-e^{-0.01y}+e^{-0.01(x+y)}, x\geq 0, y\geq 0$$

其他 $F(x,y)=0$. 求:

(1) $P(X<120, Y<120)$

(2) $P(X>120, Y>120)$

(3) $P(Y\leq X)$

二维连续概率分布律-练习解答

（1） $P(X<120,Y<120)=\ F(120,120)=1-e^{-1.2}-e^{-1.2}+e^{-2.4}=(1-e^{-1.2})^2$

（2） $P(X>120,Y>120)$

$=\ 1-P(X\le 120)-P(Y\le 120)+P(X\le 120,Y\le 120)$

$=\ 1-F(120,+\infty)-F(+\infty,120)+F(120,120)$

$=\ 1-2(1-e^{-1.2})+1-2e^{-1.2}+e^{-2.4}\ =\ e^{-2.4}$

（3） $P(Y\leqslant \mathrm{X})=\iint_G f(x,y)dxdy$

$$=\int_0^\infty\int_y^\infty 0.01^2e^{-0.01(x+y)}dxdy$$

$$=\int_0^\infty(-0.01e^{-0.01x-0.01y}|_y^\infty)dy=\int_0^\infty 0.01e^{-0.02y}dy$$

$$=-\frac{1}{2}e^{-0.02y}|_y^\infty=\frac{1}{2}$$

二维连续概率分布律-课堂练习

设二维随机变量 (X,Y) 具有概率密度

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

试求

(1) 求分布函数 F;

(2) 求概率 $P(Y\leq X)$.

二维连续概率分布律-练习解答

解：由题意可知，$F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}f(x,y)dxdy$，并且当 $x>0,y>0$ 时 $f(x,y)$ 不为 0，则有

1.

$$F(x,y)=\int_{-\infty}^{y}\int_{-\infty}^{x}f(x,y)dxdy=\int_{-\infty}^{y}\int_{-\infty}^{x}2e^{-(2x+y)}dxdy$$
$$=-\int_{-\infty}^{y}e^{-(2x+y)}\Big|_{0}^{x}dy=-\int_{-\infty}^{y}e^{-(2x+y)}-e^{-y}dy$$
$$=-\left(-e^{-y}(e^{-2x}-1)\Big|_{0}^{y}\right)=\left(1-e^{-2x}\right)\left(1-e^{-y}\right)$$

$$F(x,y)=\begin{cases}(1-e^{-2x})(1-e^{-y}) & x>0,y>0\\ 0 & 其他\end{cases}$$

2. $Y\le X$ 为 45° 斜线的右下方区域, 记为 G，则对该区域积分，则有

$$P(Y\le X)=\iint_{G}2e^{-(2x+y)}dxdy=\int_{0}^{\infty}\int_{y}^{\infty}2e^{-(2x+y)}dxdy=\frac{1}{3}$$

条件概率-课堂练习

设二维随机变量（X，Y）的概率密度为：$f(x,y)=cx^2y,\ x^2<=y<1$

1　求常数c

2　求边缘密度函数

3　求当Y=1/2时，X的条件概率密度

4　求$P\{Y>=1/4|X=1/2\}$

条件概率-练习解答

（1） $$F(1,1)=\int_{-1}^{1}dx\int_{x^2}^{1}cx^2y\,dy$$

$$=\int_{-1}^{1}\frac{c}{2}x^2-\frac{c}{2}x^6\,dx$$

$$=\frac{4}{21}c=1$$

$$\therefore c=\frac{21}{4}$$

（2） $$f_y(y)=\int_{-\infty}^{+\infty}f(x,y)dx$$

$$=\int_{-\sqrt{y}}^{\sqrt{y}}\frac{21}{4}x^2y\,dx$$

$$=\frac{7}{4}yx^3\Big|_{-\sqrt{y}}^{\sqrt{y}}=\frac{7}{4}y^{\frac{5}{2}}+\frac{7}{4}y^{\frac{5}{2}}$$

$$=\frac{7}{2}y^{\frac{5}{2}}$$

$$f_x(x)=\int_{x^2}^{1}\frac{21}{4}x^2y\,dy=\frac{21}{8}x^2-\frac{21}{8}x^6$$

条件概率-练习解答

（3） $Y=\frac{1}{2}$

$$f_Y(\frac{1}{2})=\frac{7}{2}\times\left(\frac{1}{2}\right)^{\frac{5}{2}}=7\times\left(\frac{1}{2}\right)^{\frac{7}{2}}$$

$$f_{X|Y}\left(\frac{x}{y}\right)=\frac{f(x,y)}{f_Y(y)}=\frac{\frac{21}{4}x^2y}{7\times\left(\frac{1}{2}\right)^{\frac{7}{2}}}$$

$$=\frac{\frac{21}{4}x^2\cdot\frac{1}{2}}{7\times\left(\frac{1}{2}\right)^{\frac{7}{2}}}=3\times\left(\frac{1}{2}\right)^{-\frac{1}{2}}x^2$$

$$=3\sqrt{2}x^2$$

（4） $f_{y|x}(y\,|\,0.5)=\dfrac{f(0.5,y)}{f_x(0.5)}=\dfrac{64}{15}y$

$$P(Y\geq 0.25\,|\,X=0.5)=\int_{0.25}^{1}f_{y|x}(y\,|\,0.5)dy=1$$

独立性–课堂练习

二维随机向量(X,Y)的联合概率密度为$f(x,y)=e^{-y}, 0<x\leq y$;

(1) 求分别关于X 和Y 的边缘概率密度;

(2) 计算$P\{X+Y\leq 1\}$;

(3) 判断X与Y是否独立并说明理由.

独立性-练习解答

(1). $f_x(x)=\int_x^{+\infty} e^{-y}dy=e^{-x}$

$$f_y(y)=\int_0^{y} e^{-y}dx=ye^{-y}$$

(2). $P(X+Y\le 1)=\int_0^{0.5}dx\int_x^{1-x}e^{-y}dy=1+e^{-1}-2e^{-0.5}$

(3). $f(x,y)\ne f_x(x)\bullet f_y(y)$

所以不独立，x 与 y 存在相互限制

更多详情，请咨询课程助教。